Available online at www.sciencedirect.com

SCIENCE DIRECT®

Engineering Applications of Artificial Intelligence 19 (2006) 429–438

Engineering Applications of
ARTIFICIAL
INTELLIGENCE

www.elsevier.com/locate/engappai

Alarm clustering for intrusion detection systems in computer networks

Roberto Perdisci*, Giorgio Giacinto, Fabio Roli

Department of Electrical and Electronic Engineering, University of Cagliari, Piazza D' Armi, 09123 Cagliari, Italy

Received 3 January 2006; accepted 3 January 2006
Available online 6 March 2006

Abstract

Until recently, network administrators manually arranged alarms produced by intrusion detection systems (IDS) to attain a high-level description of cyberattacks. As the number of alarms is increasingly growing, automatic tools for alarm clustering have been proposed to provide such a high-level description of the attack scenarios. In addition, it has been shown that effective threat analysis requires the *fusion* of different sources of information, such as different IDS. This paper proposes a new strategy to perform alarm clustering which produces unified descriptions of attacks from alarms produced by multiple IDS. In order to be effective, the proposed alarm clustering system takes into account two characteristics of IDS: (i) for a given attack, different sensors may produce a number of alarms reporting different attack descriptions; and (ii) a certain attack description may be produced by the IDS in response to different types of attack. Experimental results show that the high-level alarms produced by the alarm clustering module effectively summarize the attacks, drastically reducing the volume of alarms presented to the administrator. In addition, these high-level alarms can be used as the base to perform further higher-level threat analysis.

Keywords: Computer security; Intrusion detection; Alarm clustering

1. Introduction

The increasing number and value of services accessible through the Internet require an adequate protection against cyberattacks. Intrusion detection systems (IDS) are an essential component of a complete defense-in-depth architecture for computer network security. IDS collect and inspect audit data looking for evidence of intrusive behaviors. As soon as an intrusive event is detected, an alarm is raised giving the network administrator the opportunity to promptly react. The intrusion detection problem can be viewed as an instance of the generic signal-detection problem, whereby the attacks can be viewed as the signal to be detected, whereas normal system or network events can be viewed as the noise (Axelsson, 2000). Accordingly, IDS can be grouped into two broad categories, namely *misuse-based* and *anomaly-based* IDS. Misuse-based IDS base their decisions on signal characterization, whereas anomaly-based detectors base their decisions on noise characterization. In order to detect an attack, a misuse-based IDS must possess a description of the attack which can be matched to the attack manifestations (i.e., the signal). Such a description is often called *attack signature* and misuse-based IDS are also referred as *signature-based* IDS (Sy, 2005). Conversely, anomaly-based IDS rely on the assumption that attack manifestations are somehow distinguishable from the normal events (i.e., the noise). Therefore, for an anomaly-based IDS to detect an attack, it must possess a model of the normal events which can be compared to the attack manifestations. The attack can be detected if its manifestations deviate from the model of normal events. IDS can be further grouped into two categories with respect to the source of the audit data they analyze, namely *host-based* and *network-based* IDS. Host-based detectors collect audit data from operating system event logs, application logs, file system information, etc., whereas network-based detectors collect data from packets crossing a network segment.

At present, a number of commercial, open-source, and research IDS tools are available. Among them,

*Corresponding author. Tel.: +39 070 675 5776; fax: +39 070 675 5782.
E-mail addresses: roberto.perdisci@diee.unica.it (R. Perdisci), giacinto@diee.unica.it (G. Giacinto), roli@diee.unica.it (F. Roli).

0952-1976/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2006.01.003

signature-based network IDS are widely used in many organizations thanks to their ability to detect well-known patterns of intrusions while producing a low number of false alarms. Most signature-based network IDS analyze network traffic on a packet basis looking for packets that match the signature of known attacks. As soon as a signature is matched, an alarm is raised. Packet inspection provides a powerful source of fine-grain information related to suspicious activities in the protected network. Nevertheless, this fine-grain analysis causes IDS to produce a high number of alarms. The source of such a large number of alarms is motivated by the nature of some categories of attacks which send a large number of malicious packets. As signature-based IDS produce an alarm for each malicious packet, alarm flooding may occur. In addition, many attacks are performed as a sequence of steps. While each step can be easily detected by signature-based IDS, the valuable information for the network administrator relies on the aggregation of alarms related to the different steps, rather than on each single alarm. More importantly, the produced alarms are often imprecise and could report descriptions of the detected events that are either too specific or too generic. As each IDS implements different detection algorithms and signatures, the combination of complementary IDS is a promising technique that can be used to obtain a more precise and comprehensive view of suspicious network events (Bass, 2000). The use of multiple detection technologies can provide the following benefits: (i) for a given attack, different IDS may produce different outputs; (ii) for a given attack, only a limited number of IDS might be able to detect it; and (iii) the fusion of alarms raised by different IDS produces more comprehensive information about intrusion attempts than that attained using a single IDS technique. In order to gain an understanding of the intrusions against the protected network, a network administrator needs to correlate the alarms produced by different IDS to attain a high-level description of the threat. Obviously, it is infeasible for a network administrator to manually arrange the huge volume of alarms produced by multiple IDS.

Kruegel et al. (2005) presented a comprehensive view of the alarm correlation process. The proposed alarm correlation system consists of a sequence of components that transform IDS *elementary alarms* into high-level intrusion reports that are sent to the administrator. Alarm clustering is an essential part of the alarm correlation process. The aim of alarm clustering is to handle the elementary alarms produced by multiple IDS due to a certain attack, fusing them to produce an higher-level alarm message, called *meta-alarm*, that summarizes the characteristics of the detected attack and provides a reference to the related elementary alarm messages produced by the IDS. The obtained meta-alarms can be further processed by modules that perform attack scenario reconstruction and threat assessment.

This paper is a revised version of Giacinto et al. (2005), where we proposed a new *on-line* alarm clustering system. The proposed system consists of three main modules, namely an *alarm management interface* (AMI), an *alarm classifier* and an *alarm clustering* module. The AMI receives alarms from multiple IDS and translates them in a standard format. Then, the alarm classifier assigns a *class label* to the received alarms and sends them to the alarm clustering module, where the alarms are fused to obtain meta-alarms. In order to be effective, the proposed system takes into account two characteristics of IDS: (i) for a given attack, different sensors may produce a number of alarms reporting different attack descriptions; and (ii) a certain attack description may be produced by the IDS in response to different types of attack. Experimental results show that the high-level alarms produced by the proposed alarm clustering system effectively summarize the attacks, drastically reducing the volume of alarms presented to the administrator. Besides, the obtained information is suitable to be further processed by higher-level modules that perform scenario reconstruction and threat analysis.

The paper is organized as follows. A summary of the related works on alarm clustering and correlation is reported in Section 2. Section 3 presents the details of the proposed alarm clustering system. Experimental results attained by using commercial and open-source IDS are reported in Section 4. In particular, the structure of the meta-alarm is presented which can summarize a large number of elementary alarms. Conclusions are drawn in Section 5.

2. Related work

At present, only few products are available for alarm clustering and correlation, but their need is acknowledged by administrators which cannot cope with a very large number of alarms. For this reason, there is a huge effort to develop algorithms that effectively aggregate alarms. Without such tools, the logging function of IDS is likely to be turned off by administrators.

A heuristic/probabilistic approach to alarm correlation has been proposed in Valdes and Skinner (2001). Weighted distance functions are defined to aggregate alarms. An overall similarity index between alarms is obtained through a weighted sum of similarity indexes among features like the announced attack class, IP addresses, TCP/UDP source and destination ports, timestamps, etc. Expert systems have been also used to perform alarm correlation (Cuppens, 2001; Cuppens and Miege, 2002). Alarms are clustered according to suitable distance measures, and global alarms are produced. Distances among alarms are computed taking into account similarity between attack descriptions, source and target similarity, time similarity, etc.

A thorough threat analysis system has been proposed in the framework of the M-Correlator project (Porras et al., 2002). Alarms are filtered and clustered according to the knowledge of the network architecture and known vulnerabilities. For each alarm cluster, a relevance score