Proposals on Assessment Environments for Anomaly-Based Network Intrusion Detection Systems

M. Bermúdez-Edo, R. Salazar-Hernández, J. Díaz-Verdejo, and P. García-Teodoro

Dpt. of Signal Theory, Telematics and Communications, University of Granada
E.T.S. Ing. Informática y Telecomunicación, C/ Periodista Daniel Saucedo Aranda, s/n
18071 – Granada (Spain)
mbe@ugr.es,rsalaza@correo.ugr.es,jedv@ugr.es,pgteodor@ugr.es

Abstract. One of the key challenges that researchers should face when proposing a new intrusion detection approach (IDS) is that of demonstrating its general validity. This fact goes necessarily through the disposal of a real set of intrusion (as well as non-intrusion) related events, from which to compare and thus validate the performance of the novel proposed techniques. However, this a priori simple issue is far to be obvious because of the lack of a commonly accepted assessment methodology. In this line, the authors discuss a set of basic requirements that an intrusion-oriented framework should fulfill in order to deal with the normalization of the evaluation process in IDS environments. In its current preliminary state, the work is mainly focused to analyze, specify and manage traffic databases for developing and validating NIDS.

Keywords: Network security, Intrusion event, IDS, Assessment.

1 Introduction

As the social acceptance of communication systems has grown, especially by means of Internet, the dependence on ICT technologies has become more and more critical. Moreover, the increasing relationships and complexity of current ICT-related systems make global information and communication infrastructures highly vulnerable [1] [2]. A prominent aspect of such vulnerability concerns the dramatic increase of the number of security incidents in last years [3]. In this context, one of the most adopted tools to improve security in internetworking facilities is that of intrusion detection systems (IDS) [4].

IDS' can be classified according to several criteria [5]. One of them regards the type of analysis to be performed, from which an IDS can be either a signature-based intrusion detection system (S-IDS) or an anomaly-based one (A-IDS). In the first case attack patterns are specified, and the system will signal an intrusion event when a match between the monitored events and one of the patterns in the signature database is observed. In the anomaly-based IDS approach, the normal behavior of the target system is captured and modeled, and an alarm rose if the behavior of the monitored environment does not comply, within an accepted range, the expected one.

The main advantage of the S-IDS approach lies on the control of known attacks. On the contrary, the main disadvantage of this approach is the impossibility of

J. Lopez (Ed.): CRITIS 2006, LNCS 4347, pp. 210 – 221, 2006.

detecting unknown attacks, even if they are quite similar to a known one. On the other hand, the advantage of the A-IDS methodology is its hypothetical capacity to detect previously unobserved intrusive events; whilst the main disadvantage is related to the fact that an alarm will be triggered every time an "abnormal" event is accomplished, even if it is legitimate (false positives). Regarding this last point, it is important to indicate that false positives rate can be reduced through the so-called "specification-based" IDS approach, where a model is derived from formal protocol specifications.

One more accepted classification for IDS depends on the origin of the data to be analyzed: either the network or a host. The first case corresponds to a network-based IDS (NIDS), that is, the data to be analyzed is related to communication protocols and payloads. Instead, the approach is called host-based IDS (HIDS) when host events, such as processes, users, system calls, etc., are indeed the analyzed events.

Any combination of these techniques can be used to improve the effectiveness of the detection, like signature and anomaly detection based, or HIDS and NIDS [7], or anomaly based and specification based [8].

Because of the increasing impact of attacks in ICT, the intrusion detection technology is continuously evolving to improve the security and protection of systems and infrastructures. However, one of the main challenges that researchers face, when trying to implement and validate a new intrusion detection method, is to asses it, and to compare its performance to that of the currently available approaches. This is due to various reasons. First of all to data privacy, that forbids the real databases to be shared among researchers; on the other hand, the use of synthetic networking events has been widely criticized [9]. The second main obstacle concerns the lack of a methodology to use event databases for testing the new model, and benchmarking it with the existing ones [10].

At the present time, IDS methods are validated without any rigor from an objective technical point of view. Generally, every researcher has his/her own methodology for testing the work done, and hence, it is very difficult to actually decide on which technique performs better than another. In this line, and due to the relevance of the subject, the aim of this paper is to point out a number of issues to define equivalent frameworks to accomplish the assessment of IDS environments. Thus, every researcher could create his own real traffic database and use it with guaranty about the reliability of the obtained results. We will center the attention on the database management, preparing the real dataset for properly testing an IDS, and the steps followed in the process. The methodology to be described will be preliminary applied to a specific environment developed by the authors, which is mainly characterized by the use of a hybrid NIDS solution: signature-based detection, and anomaly-based detection (complemented with some specification-based aspects) are combined to take advantage of the two approaches.

The paper is organized as follows. In Section 2 already existing methodologies to assess IDS environments are described, and their main limitations discussed. Section 3 introduces a methodology for validating anomaly-based IDS, and how to manage the traffic databases for that. After that, Section 4 proposes and discusses a practical method to solve the implementation problems emphasized in the previous section. Section 5 describes the specific framework developed by the authors to test their own IDS approaches. Finally, main conclusions and future related work are given in Section 6.