Computer Communications 35 (2012) 772–783

Contents lists available at SciVerse ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Unsupervised Network Intrusion Detection Systems: Detecting the Unknown without Knowledge

Pedro Casas *, Johan Mazel, Philippe Owezarski

CNRS; LAAS; 7 avenue du colonel Roche, F-31077 Toulouse Cedex 4, France
Université de Toulouse; UPS, INSA, INP, ISAE; UT1, UTM, LAAS; F-31077 Toulouse Cedex 4, France

ARTICLE INFO

Article history:
Received 31 August 2011
Received in revised form 10 January 2012
Accepted 20 January 2012
Available online 28 January 2012

Keywords:
NIDS
Unsupervised Machine Learning
Sub-Space Clustering
Evidence Accumulation
Outliers detection

ABSTRACT

Traditional Network Intrusion Detection Systems (NIDSs) rely on either specialized signatures of previously seen attacks, or on expensive and difficult to produce labeled traffic datasets for user-profiling to hunt out network attacks. Despite being opposite in nature, both approaches share a common downside: they require the knowledge provided by an external agent, either in terms of signatures or as normal-operation profiles. In this paper we present UNIDS, an Unsupervised Network Intrusion Detection System capable of detecting unknown network attacks without using any kind of signatures, labeled traffic, or training. UNIDS uses a novel unsupervised outliers detection approach based on Sub-Space Clustering and Multiple Evidence Accumulation techniques to pin-point different kinds of network intrusions and attacks such as DoS/DDoS, probing attacks, propagation of worms, buffer overflows, illegal access to network resources, etc. We evaluate UNIDS in three different traffic datasets, including the well-known KDD99 dataset as well as real traffic traces from two operational networks. We particularly show the ability of UNIDS to detect unknown attacks, comparing its performance against traditional misuse-detection-based NIDSs. In addition, we also evidence the supremacy of our outliers detection approach with respect to different previously used unsupervised detection techniques. Finally, we show that the algorithms used by UNIDS are highly adapted for parallel computation, which permits to drastically reduce the overall analysis time of the system.

1. Introduction

The detection of network attacks is a paramount task for network operators in today's Internet. Botnets, Malwares, Distributed Denial of Service attacks (DDoS), buffer overflow attacks, network-scanning activities, and spreading worms or viruses are examples of the different threats that daily compromise the integrity and normal operation of the network. The principal challenge in automatically detecting network attacks is that these are a moving and ever-growing target [1,2]. Network Intrusion Detection Systems (NIDSs) are the war-horses of network security. Two different approaches are by far dominant in the research literature and commercial IDS security devices: signatures-based or **misuse detection** (*detect what I know*) and **anomaly detection** (*detect what it is different from what I know*).

Misuse detection is the de facto approach used in most IDSs. When an attack is discovered, generally after its occurrence during a diagnosis phase, the associated malicious pattern is coded as a *signature* by human experts, which is then used to detect a new occurrence of the same attack. To avoid costly and time-consuming human intervention, signatures can also be constructed by supervised machine-learning techniques, using instances of the discovered attack to build a detection model for it. Misuse detection systems are highly effective to detect those attacks which they are programmed to alert on. However, they cannot defend the network against new attacks, simply because they cannot recognize those attacks which do not match their lists of signatures. Indeed, networks protected by misused detection systems suffer from long periods of vulnerability between the diagnosis of a new attack and the construction of the new signature.

On the other hand, anomaly detection uses instances of normal-operation traffic to build normal-operation profiles, detecting anomalies as activities that deviate from this baseline. Such methods can detect new kinds of network attacks not seen before. Nevertheless, they require training to construct profiles, which is time-consuming and depends on the availability of anomaly-free traffic instances. In addition, it is not easy to maintain an accurate and up-to-date normal-operation profile, which induces high false-alarm rates.

* Corresponding author at: CNRS, LAAS, 7 avenue du colonel Roche, F-31077 Toulouse Cedex 4, France. Tel.: +33 (0)5 61 33 68 05; fax: +33 (0)5 61 33 64 11.
E-mail addresses: pcasas@laas.fr (P. Casas), jmazel@laas.fr (J. Mazel), owe@laas.fr (P. Owezarski).

0140-3664/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2012.01.016

Despite being opposite in nature, both misuse detection and anomaly detection share a common downside: they require the knowledge provided by an external agent to achieve their goal, either in terms of attack-signatures or as normal-operation profiles. As such, current network security looks more like a reactive countermeasure than a proactive prevention mechanism. Over the past years we have, however, witnessed an increased interest within the network security community in shifting away from reactive defense towards more proactive security systems [3]. Our thesis behind this work is that reactive, knowledge-based approaches are not sufficient to tackle the network security problem, and that a holistic solution should also include proactive, knowledge-independent analysis techniques.

Armed with these ideas in mind, we present an Unsupervised Network Intrusion Detection System (UNIDS) capable of detecting network attacks without relying on signatures, training, or labeled traffic instances of any kind. Based on the observation that network attacks, and particularly the most difficult ones to detect, are contained in a small fraction of traffic instances with respect to normal-operation traffic [6] (we show that this hypothesis can always be verified by using traffic aggregation), their unsupervised detection consists in identifying *outliers*, i.e. instances that are remarkably different from the majority. UNIDS relies on robust clustering techniques to blindly extract the traffic instances that compose an attack. This unsupervised security system runs in three consecutive steps, analyzing packets captured in contiguous time slots of fixed length. Fig. 1 depicts a modular, high-level description of this system:

The **first step** consists in detecting an anomalous time slot in which the clustering analysis will be performed. For doing so, captured packets are first aggregated into *multi-resolution* traffic flows. Different time-series are then built on top of these flows, and any generic change-detection algorithm based on time-series analysis is finally used to flag an anomalous change. In this paper we shall use a standard change-detection algorithm [4] on three very simple and traditionally used volume metrics, consisting of # bytes, # packets, and # flows per time slot. The choice of volume metrics is based on [11], but change-detection can be performed on any other traffic metric sensitive to anomalies. The algorithm basically flags an anomaly when the derivative of any of these metrics exceeds a detection threshold, dynamically computed from the variance of previous anomaly-free measurements. The reader should bear in mind that this change-detection step is not a critical part of UNIDS, but that it is merely used to limit the frequency of usage of the clustering step, which is certainly more expensive in terms of computational resources.

The **second step** takes as input all the flows in the time slot flagged as anomalous. At this step, outlying flows are identified using a robust multi-clustering algorithm, based on a combination of Sub-Space Clustering (SSC) [18], Density-based Clustering [23], and Evidence Accumulation Clustering (EAC) [22] techniques. The knowledge provided by this clustering algorithm is used to rank the degree of *abnormality* of all the identified outlying flows, building an *outliers ranking*.

In the **third step**, the top-ranked outlying flows are flagged as anomalies, using a simple thresholding detection approach.

As we show through out the paper, the main contribution of UNIDS relies on its ability to detect unknown attacks in a completely unsupervised fashion, avoiding the need for signatures, training, or labeled traffic flows. This paper represents a continuation of our previous work on unsupervised anomaly detection [5]. In particular, we show that UNIDS can be used to detect unknown network attacks of very different nature, outperforming traditional misuse-detection-based systems; as such, we provide more evidence and solid results on the quality and relevance of our proposals for unsupervised anomaly detection. In addition, we show that the computational time involved in the unsupervised traffic analysis can be drastically reduced w.r.t the system presented in [5], by simply taking advantage of the parallel structure of the multi-clustering algorithm used in the core of UNIDS.

The remainder of the paper is organized as follows. Section 2 presents a brief state of the art in the network intrusion and anomaly detection fields, describing our main contributions. Section 3 describes the multi-resolution traffic aggregation and change-detection procedures used in the first step of the UNIDS system to identify an anomalous time-slot. Section 4 describes the core of UNIDS, presenting an in depth description of the different clustering techniques used to construct the outliers ranking. Section 5 presents the validation of UNIDS in real traffic traces obtained from two networking datasets: the public MAWI traffic repository of the WIDE project [25], and the METROSEC project dataset [27]. In this section we also compare the performance of UNIDS against previous proposals for unsupervised detection of attacks available in the literature. Section 6 evaluates the ability of UNIDS to detect unknown attacks in the well-known KDD99 network attacks dataset, comparing its performance with that obtained by an extensively investigated misuse NIDS based on decision trees. Implementation related issues of UNIDS, including evaluation of computational

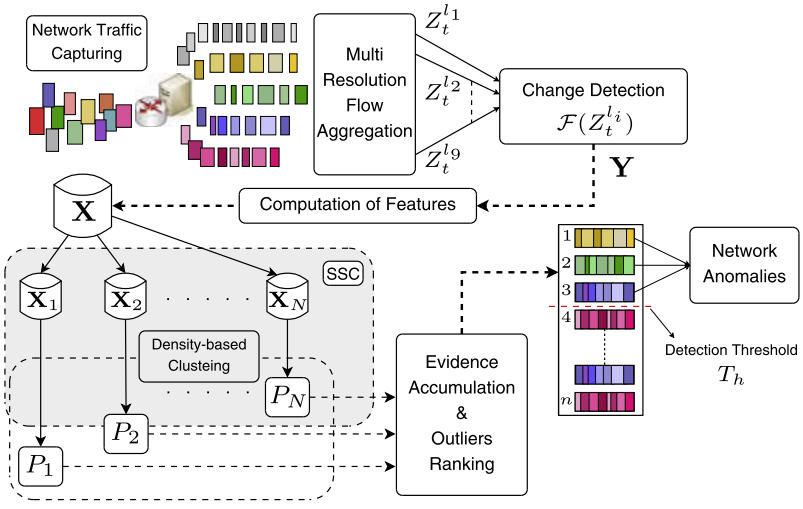

Fig. 1. High-level description of the Unsupervised Network Intrusion Detection System (UNIDS).