A Cooperative Approach for a Global Intrusion Detection System for Internet Service Providers

Conference Paper · October 2017

DOI: 10.1109/CSNET.2017.8241992

Renato Souza Silva
Federal University of Rio de Janeiro

Evandro Macedo
Federal University of Rio de Janeiro

A Cooperative Approach for a Global Intrusion Detection System for Internet Service Providers

Renato S. Silva
Federal University of Rio de Janeiro
RAVEL/PESC/COPPE
Rio de Janeiro, RJ 21941-914
Email: renato@ravel.ufrj.br

Evandro L. C. Macedo
Federal University of Rio de Janeiro
RAVEL/PESC/COPPE
Rio de Janeiro, RJ 21941-914
Email: evandro@ravel.ufrj.br

Abstract—**Cyber-attacks have become more threatening as Internet evolves, particularly for Internet Service Providers (ISPs) that play a rule of carrying them to their subscribers. In order to protect themselves and their subscribers, ISPs invest in typical protection systems like IDS, IPS, or Firewalls, that are designed for perimeter-based operation. Even though these expensive systems are efficient to protect confined environments, they do not allow ISPs to anticipate cyber-attacks. At most, ISPs might only react to them as soon as possible to maintain network services for legitimate traffic. Based on what prior DIDS approaches have lacked, our approach relies on BGP protocol to interconnect distributed intrusion detection elements, each of which cooperating by sending information about a potential threatening flow that traverses its Autonomous System (AS). We present the architecture of our approach as well as the analytic model based on Dempster-Shafer's combination rule. The results show significant improvement in terms of reliability of the combined information, that enables better countermeasures decisions.**

1. Introduction

One of the main factors for cyber-attackss success is precisely the connectivity based design of the Internet [1]. As the connectivity of Internet evolves cyber-attacks also have become more threatening, particularly for Internet Service Providers (ISPs) that play a rule of carrying them to their subscribers. Regarding such compromising scenario, ISPs are indeed spending a lot of money in complex protection systems for defending themselves (and their clients) from cyber-attacks threatness [2]. Although such these systems can identify malicious flows and clean network traffic, they cannot foresee a new incoming attack. In order to prevent from incoming cyber-attacks, the ISP SIEM (Security Information and Event Management) often correlates information from several sources, like warning vendors, Cyber Emergency Response Teams (CERT) [3] and Internet forums. However needless to say this takes precious time.

Even having their systems full up-to-date, ISPs are not safe against zero-day attacks. This type of attack are often successful and bring on serious damages by taking advantage of greenfield protection systems [4]. For instance, a quite recent ransomware zero-day attack (May 12th, 2017), namelly WannaCry, have spread across the Internet, which infected more than $230,000$ computers around the world, impacting several systems. Such an example illustrates how an ISP can be naturally used as a cyber-attack vector towards its own clients.

Typical protection systems like IDS (Intrusion Detection Systems), IPS (Intrusion Prevention Systems) or Firewalls are designed for perimeter-based operation. That is, they are placed on the perimeter of network to avoid intrusive incomings flows. Even though these expensive systems are efficient to protect confined environments, they do not allow ISPs to anticipate blocking countermeasures. At most, relying on consolidated informations, ISPs might only react to them as soon as possible to protect network services. However, the closer to the protection perimeter, the slower the reaction time and more threatening the attack.

Just as the Internet connectivity-based design helps cyber-attacks reach their targets wherever they are, the number of networks traversed by a malicious flow might increase its detection likelihood. As a matter of fact, approximately 97% of Distributed Denial of Service (DDoS) attacks come from external Autonomous Systems (AS) [5]. Assuming that all these path networks have their own network detection system, which is able to identify an anomalous traffic in transit, the probability of detection increases as the number of detection systems increases along the path.

We propose a global intrusion detection system composed by autonomous Internet-distributed detection systems. In our approach distributed detection elements cooperate by sending information about a potential threatening flow that traverses its Autonomous System (AS). Distributed Intrusion Detection Systems (DIDS) use Border Gateway Protocol (BGP) updating capabilities in order to spread intrusion warning messages across Internet routing domain so as to notify the SIEM of the attack target.

When an anomalous in-transit traffic is detected, the AS-integrated IDS gathers all attributes of the anomalous flow in the extended BGP Network Layer Reachability Information (NLRI) field and advertises it towards the AS target of the intrusion. Then, the SIEM of the target AS can use such information set to manage related protection countermeasures. Since the well known BGP AS-path field shows the flow path, it also permits a partial knowledge of the intrusion source. Moreover, the combined information from correlated warning messages received at the destination system can be used to infer about alarm severity as well. Based on what prior DIDS approaches have lacked to become an Internet standard, we believe this work offers a new insight to leverage a cooperative global intrusion detection system based on distributed autonomous elements fully-associated.

The remaining of this paper is organized as follows. Section 2 analyzes the state-of-the-art in attacks detection systems. In Sections 3 to 5 we present some concepts for a better understanding of our underlying approach. Section 6 outlines key elements that compose the architecture of the proposal. In Section 7 we propose an analytic model to evaluate the system performance regarding the combined intrusion information. An overall analysis and conclusion are presented in Section 9.

2. Related Works

The authors in [6] assess how HIS identification process can be used to distinguish the normal network activities from abnormal network ones. The relative analogy proposed in [7] defines 3 goals to design an efficient network-based IDS: distributed, self-organizing and lightweight. The distributed requirement suggests that a zero-day attack experienced by an IDS may no longer be for another geographically distant IDS. Self-organizing implies that IDS agents can reconfigure themselves. Agents can be added or removed from the system without any noticeable availability loss. Being lightweight means easy to deploy and not requiring such a complex structure to operate. In other words, being a lightweight platform encourages in attracting new agents. According to [8], the larger the number of integrated IDS the better DIDS works.

The AAIFD system proposed in [9] introduces autonomous entities that operate in a distributed-hierarchical SNMP (Simple Network Management Protocol) [10] architecture for intrusion detection. The main drawbacks of the AAIFD system are the need of a third part component to provide detection system with redundancy and the delay to gather necessary information for detection.

CRIM [11] is part of MIRADOR [12] project as a cooperative alert module for correlating alarms from several IDS. The main objectives of CRIM are to reduce false alarms and increase the intrusion detection rate.

The authors of Indra [13] argue that the chances of an intrusion detection increases with the number of distributed elements, the heterogeneity of the elements and the level of currency of their security fixes. Indra relies on Peer-to-Peer (P2P) communication model to take proactive and reactive protection measures.

DOMINO architecture [14] is a dynamic infrastructure composed of heterogeneous Internet-spanning nodes organized as an overlay network. Its design mixes P2P and hierarchical components in order to gain in scalability and redundancy as well. Even though the remarkable concerning of DOMINO of inducing networks to join its infrastruture, its deployment still requires a wide development effort.

In a more recent work Igbe *et al.* [15] proposes a fully distributed NIDS approach where no central controller is needed. The detection system is based on an adaptive artificial immune mechanism whose classifying method uses an unsupervised machine learning to distinguish normal traffic (self) from the abnormal one (non-self). The authors advocate that zero-day attacks can be detected through interactions between distant IDS.

Bass [2] analyses employing multisensor data fusion systems for inferring about the identity and the location of an intruder, the intruder activity, the observed threats, the attack rates, and an assessment of the severity of a cyber-attack. Following the same approach the DIDS architecture proposed in [16] combines distributed monitoring and data reduction with centralized data analysis to address the shortcoming of current single host IDS. Trough an unique network-user information (NID) DIDS can infer detection even in some challenging environments since it can correlates data from multiple hosts and the network.

Currently there are several global data-base security services that collect intrusion data around the globe in order to help individuals and security personnel in anticipating cyber-attacks. The most of these services uses voluntary agents for gathering attack data. The technical report in [17] classifies some of these services and compare them according to their design objective, sources of data used, ability for performing anonymous uploads and availability of attacker notification and tracking tools. Although some of these services are free they do not offer specific analysis regarding threatness to an individual network.

As can be seen in the related works, the concept of distributing intrusion detection elements and make them cooperate with each other is not new. Actually there has been a huge amount of different approaches to fulfill the goal of providing Internet with its own auto-immunization system. In spite of having such a wide range of approaches, there still isn't an open global detection system able to act as a first line of protection against cyber-attacks. A natural question that comes up after studying the former works refers to what do they lack to become a *de facto* DIDS architecture running on the Internet. In our approach we suggest two key answers to this question: easy deploying and autonomy. Based on these two answers we propose a system that take advantage of stable Internet operation in order to fastly gather the largest possible number of federated members; and the heterogeneity of methods so as to extend the detection range.

3. Intrusion Detection Systems

An intrusion is a kind of cyber-attack where the attacker attempts to jump security mechanisms to gain access into a system so as to damage its integrity, availability, or confidentiality [18]. There are two similar systems that can identify and report intrusions in the network: Intrusion Detection

System (IDS) and Intrusion Prevention System (IPS). Both systems enhance the security level of networks, monitoring traffic, inspecting, and scanning packets for suspicious data. Whereas IDS is a passive element, IPS introduces online countermeasures component so that current and further attacks are blocked.

Regarding the source of audit data, there may be two different approaches, namelly, network-based and host-based detection. In the HIDS (Host-based Intrusion Detection System) approach suspicious activities such as application or data requests, network connection attempts, and read-/write attempts[1] are monitored directly on the operating system. Due to its tight integration with the host operating system, there may be some problems to update it when necessary. The NIDS (Network-based Detection System) approach relies on in-line network connection to inspect all the elements attached into the same network. Towards the detection methodology, systems may be classified as statistical anomaly detection and signature-based detection [19]. Whatever the approach or methodology the main drawback of any detection system is false-positives (FP) & false-negatives (FN) reports. A typical topology used for network IDS/IPS systems is depicted in Figure 1.

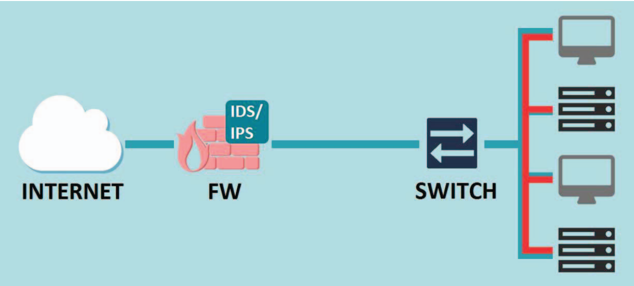


Figure 1. IDS/IPS network topology

Distributed Intrusion Detection Systems (DIDS) can be defined as multiple Intrusion Detection Systems (IDS) spread over a large network[2], all of which communicate with each other, or with a central server that facilitates advanced network monitoring, incident analysis, and instant attack data [17]. The DIDS concept came up as a strategy to bypass problems related to the monolithic architecture of traditional detection systems. DIDS architecture can be seen as a part of Artificial Immune System (AIS) where several network-distributed autonomous system cooperate with each other to detect an intrusion/anomaly.

The main challenges of deploying a DIDS refers to the infrastructure to support its cooperative behavior and how distributed such infrastructure is.

4. Dissemination of Flow Specification Rules

The benefit of BGP Flow Spec is that it allows BGP speakers to use a new BGP NLRI defining flow filter information which can then be advertised to upsteam neighbors via BGP. The primary and immediate motivation of this protocol is to provide networks with intra and inter provider distribution of traffic filtering rules to filter DoS and DDoS attacks. However it can be used for a wide variety of applications in which filtering information must be dynamically distributed throughout a network.

RFC 5575 [20] relies on MP-BGP [21] interworking capabilities[3] so as to distribute traffic flow specifications through a new BGP Network Layer Reachability Information (NLRI). In order to link the encoded application with its Local Routing Information Base (LOC-RIB) at destination router, two new Address Family Identifiers (AFI) and Subsequent Address Family Identifiers (SAFI) pairs were specified, as follows.

- (AFI=1,SAFI=133) for Internet Protocol version 4 (IPv4)
- (AFI=1,SAFI=134) for Virtual Private Network version (VPNv4)

RFC 5575 also defines a minimum set of filtering actions related to a disseminated flow.

```
+--------+--------------------+--------------------------+
| type   | extended community | encoding                 |
+--------+--------------------+--------------------------+
| 0x8006 | traffic-rate       | 2-byte as#, 4-byte float |
| 0x8007 | traffic-action     | bitmask                  |
| 0x8008 | redirect           | 6-byte Route Target      |
| 0x8009 | traffic-marking    | DSCP value               |
+--------+--------------------+--------------------------+
```

Figure 2. Extended community values used to define particular actions.

The flow specification NRLI-type field comprises several options. A packet is considered to match flow specification when it matches the intersection of all the components present at the NRLI-type field. Table 1 shows these options.

TABLE 1. Flow features in NRLI-type field

Type	Description	Type	Description
1	Destination Prefix	7	ICMP type
2	Source Prefix	8	ICMP code
3	IP Protocol	9	TCP flags
4	Port	10	Packet length
5	Destination Port	11	DSCP
6	Source Port	12	Fragment

5. Dampster-Shafer's Evidence Theory

Dempster-Shafer's theory [22] is a formal framework to combine sources of evidence. It differs from probability theory in terms of explicit representation of uncertainty and combination of evidences. Dempster's rule of combination concerns combining $n \geq 2$ belief functions from distinct and independent sources of evidence.

Let $\Omega = \{\omega_1, \omega_2\}$ be the exhaustive and mutually exclusive frame of discernment that has only two elements

[1] Recent works also classify application-based as a host-based evolution

[2] The "large network" in the present discussion is presumed to be the global Internet, not merely a network or intranet of any single organization.

[3] MP-BGP protocol is affectionately called the swiss army knife by BGP specialists due to its wide range of application

(detected or not detected) for the fusion of n independent sources of evidence ($IDS_1, IDS_2, ..., IDS_n$) about a potential threat. The set of all hypothesis subsets of Ω is named as the power-set of Ω and is denoted by 2^Ω. Belief functions rely on mass functions $m : 2^\Omega \to [0, 1]$ and that conform to the following conditions.

$$m(\emptyset) = 0 \quad and \quad \sum_{A \subseteq \Omega} m(A) = 1$$

The belief of a set A, $Bel(A)$, is defined as the sum of all belief masses of subsets of the set of interest as

$$Bel(A) = \sum_{B|B\subseteq A} m(B) \; ; \; \forall A \subseteq \Omega \tag{1}$$

Plausibility of A comes from the the fact that the relation $Bel(A) + Bel(\widehat{A}) = 1$ does not hold. $Pl(A)$ relates to $Bel(A)$ through

$$Pl(A) = 1 - Bel(\widehat{A}) \tag{2}$$

Where $\widehat{A}$ is referring to its compliment “not A”. $Bel(\widehat{A})$ is often called the doubt in A. Clearly $Pl(A)$ represents the extent to which we fail to disbelieve A.

Let $m_i(A_s)$, $i = \{1, ..., n\}$; $s = \{1, ..., k\}$ be n independent sources of s kinds of evidence. The combined belief mass $m_C(A)$ is obtained as follows

$$m_C(A) = \frac{\sum_{\cap A_s = A} \prod_{1\leq i\leq n} m_i(A_s)}{1 - K} \tag{3}$$

K evaluates the amount of conflict among the mass sets.

$$K = \sum_{\cap A_s = \emptyset} \prod_{1\leq i\leq n} m_i(A_s) \tag{4}$$

6. Architecture

What is really being sought is empowering Internet with its own inborn self-defense and immunization system that detects threats wherever they rise to permit SIEM to anticipate countermeasures. In this regard, human immunization system has been inspiring the most of the similar proposes. However, one question we have been asking ourselves is precisely what these noteworthy DIDS approaches lack to become an Internet standard. The most likely answer to our question is **blood**, in an analogy, as a powerful carrier for whole human body and particularly to transport *B cells*[4] as ubiquitous defense agents.

Intrusion information does need to be carried around whenever it is necessary. Actually, BGP is the only ubiquitous protocol on the Internet. BGP update messages can be sent from any AS and can be received by all other AS. Moreover, as mentioned in [20] modern IP routers contain both the capability to forward traffic according to IP prefixes as well as to classify, shape, rate limit, filter, or redirect packets based on administratively defined policies. One underlying conclusion is about IP/BGP routers as being B cells of the Internet. As a human being that is born with its own self-defense system, it is impossible to conceive a new AS without any IP/BGP router connected to the Internet.

Taking advantage from the global DIDS framework, the BGP protocol and the capability of the new routers with customized NetFPGA technology for integrated multi-service processing [23], we present a lightweight fully distributed intrusion detection architecture aiming to open new insights for leveraging a *de facto* standard intrusion detection system.

6.1. High Level Design

The high level design blends the distributed intrusion detection concept and BGP skills as the Internet information carrier and the modern IP routers processing capabilities. The main idea is to take advantage from the enhanced processing of modern routers for deploying a hardware-integrated intrusion detection system at the border layer of each AS. The distributed location of autonomous IDSs into each AS neighbor gives rise to an intrusion detection dataset from different elements, geographically spread across Internet. Given the full integration of these elements by Internet BGP routing domain, they can communicate with each other everywhere everytime they need by using BGP extended update messages. At the SIEM into the destination AS all warning messages received from BGP network are correlated for building an intrusion information base in order to support a countermeasure decision.

The scenario depicted in Figure 3 shows 6 integrated IDS that can detect an intrusion with a certain probability. During a typical Distributed Denial of Service (DDoS) attack, four botnets into ASs 1, 2 and 3 are remotely controlled to perform a simultaneous attack against a victim in AS 6. IDSs into ASs 3, 4 and 5 got successful in detecting the anomalous flow and advertised a BGP update message toward AS 6 so as to warn its SIEM about the anomalous flow. SIEM into AS 6 receives the 3 update messages and proceeds its protection script which includes correlating BGP messages according their flow features, combining correlated messages to get its reliability degree and analyzing it against the normal flow to define the best policy so as to preserve as much as possible the legitimate traffic. After mitigating the attack, SIEM carries on a further *post mortem* analysis over its IDSs rules to prevent from future attacks.

One strategical operation point of the proposed architecture concerns of enabling distributed IDS for warning only anomalous **in-transit**[5] flows. A distributed IDS that detects an anomalous flow incoming its own autonomous system cannot warn any other ASs. Such strategical point claims for the cooperative behavior of the system and prevents from possible legal issues related to ISP subscribers.

True-positive rate (TPR) parameter means the rate between the number of detected intrusions over the number of anomalous flows. It is related to the IDS capacity in

[4] B Cell is a type of lymphocyte normally involved in the production of antibodies for dynamically combating infection

[5] Flows that are traversing an AS

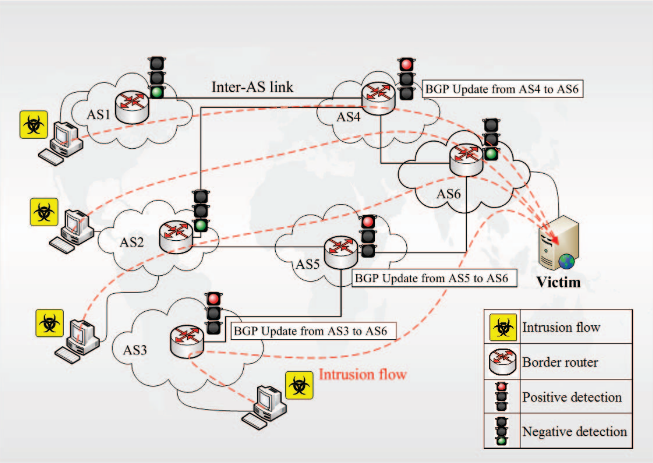


Figure 3. Four attackers performing an attack against a victim into $AS6$. IDSs into ASs 3, 4 and 5 detect the anomalous traffic and convey a warning message to $AS6$ by advertising an extended BGP update encoding the anomalous flow features.

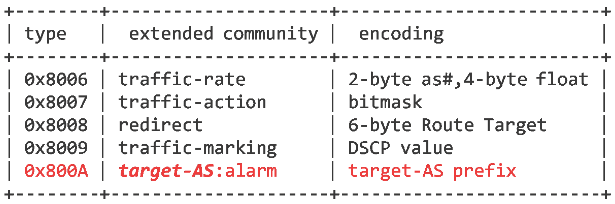
detecting an intrusion among the normal traffic. Since the number of received messages is the main way to infer about an event severity, true-positive rate play a rule of generate input data in such correlation system. Therefore having each IDS element a good true-positive rate assures efficiency to the cooperative approach while the false-negative[6] rate is reduced by combining multiple autonomous IDS reports.

Due to its inherent requirement to inspect all in-transit traffic, coupling the IDS elements at border gateway ensures both inspection access to incoming & outgoing AS traffic and the BGP connectivity. Nevertheless, questions related to router capacity in inspecting all AS incoming & outgoing traffic may be challenging to manage even considering the aforementioned capacity of modern routers. Moreover, even though the IDS elements are autonomous and independent it is necessary to encourage a reasonable true-positive rate in each companion detection element.

6.2. Low Level Design

RFC 5575 defines a minimum set of actions to be deployed from a specific disseminated flow. Instead of detecting a specific attack type and triggering an immediate action, our approach proposes let SIEM of the target AS analyze a number of BGP-disseminated warnings that include the features of the threatening flows. Such analysis can be accomplished by using a big-data [24] application configured according to the AS security assumptions. Instead of using one of the already defined RFC actions, a new extended community is proposed according to Figure 4 below.

Target-AS:alarm is a transitive extended community consisting of 6 bytes of which only the 4 least significant bytes are defined to the target IP prefix. The IDS application will interacts with BGP-aware element to fill this field with the intrusion target-AS number as well as the remaining fields according to RFC 5575. The same community configured at the target-AS assures to run the related action at the AS destination. Each BGP-aware element at the transit AS originates by it self the flow-spec BGP update. Other lower level configuration related to BGP attributes (i.e. next-hop and originator) to deploy flow-spec communication among different ASs depend on their neighboring policies.



```
+--------+---------------------+-------------------------+
| type   |  extended community |  encoding               |
+--------+---------------------+-------------------------+
| 0x8006 | traffic-rate        | 2-byte as#,4-byte float |
| 0x8007 | traffic-action      | bitmask                 |
| 0x8008 | redirect            | 6-byte Route Target     |
| 0x8009 | traffic-marking     | DSCP value              |
| 0x800A | target-AS:alarm     | target-AS prefix        |
+--------+---------------------+-------------------------+
```

Figure 4. Set of extended community values that can be used to define particular actions, including the new proposed action (red).

The diversity of information learned from autonomous IDS elements is the core of any distributed intrusion detection system. Thereby, the more heterogeneous the elements, the better the chance of detecting an intrusion, including a zero-day attack. Having heterogeneous intrusion messages means having various geographically-distributed IDSs which are able to learn the behavior of their own networks. However, even when the current behavior is captured perfectly, it is very likely that the sample gets outdated soon due to the continuous change of network software, devices and usage. The so-called self-learning systems can dynamically adjust to changes on network traffic behavior [25].

In the previous section it was considered the hypothesis of fitting the IDS as a card at the border gateway chassis. Such proposition solve both the problem in approaching the IDS to the incoming & outgoing traffic and in being a low cost solution. A relevant factor to be considered in the integrated structure concerns to the forwarding performance of the border gateway in inspecting this amount of traffic. Such a performance issue can be addressed by combining a sampling-based traffic monitoring method with the dense processing capacity provided by Netfpga technology [23]. Netfpga technology has recently raised as an option to build compact structures able to perform middleware tasks i. e. Firewall, Carrie Grade Network Address Translation (CGNAT), Deep Packet Inspection (DPI) and an IDS as well [26], [27].

7. Analytic Model

As mentioned in the Section 6 our proposal resides on analyzing a number of true-positive (TP) intrusion messages received at a SIEM from each independent DIDS in the attack flow path.

True-positive rate (TPR) of a DIDS refers to the fraction of intrusion that is correctly detected over the total intrusion number.

[6] False-negative parameter (FN) means that the system fails in detecting an intrusion.

$$TPR = \frac{TP}{TP + FN}$$

Despite being proportional to the number of intrusion messages to correlate, higher true-positive rate can also be undesirably achieved by lowering the true-positive threshold. On the other hand the precision (Pr) of a certain DIDS refers to its ability to identify positive results by gauging the fraction of test data detected as intrusion that are actually intrusion. Precision is indeed considered a balanced metric because it includes a false-positive (FP) parameter to compensate for a possible low true-positive threshold.

$$Pr = \frac{TP}{TP + FP}$$

It is obvious that there is a trade-off between true-positive (TPR) and precision (Pr) rate. The recall-precision characterization of an IDS, named F-score (Fs), is given by

$$Fs = \frac{2 \times TPR \times Pr}{TPR + Pr} \tag{5}$$

Like precision and true-positive rate, Fs also underlies on the probabilistic framework hence it can be used to evaluate the probability that a certain IDS is reliable.

Figure 5 models a distributed detection system composed by a total number of t weightage-equivalent and independent elements. It shows an intrusion being detected by a number $N \leq t$ of IDSs. Each IDS in the flow path advertises its own BGP update message (U_i) related to the detected intrusion. Once arriving at the destination SIEM, such $n \leq N$ messages are correlated and combined (fused) to support a consensus countermeasure decision.

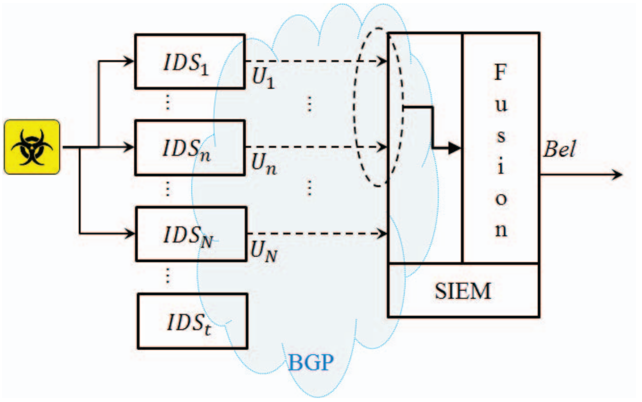


Figure 5. Parallel decision fusion network.

Dempster-Shafer's theory has been used to model uncertainty in expert systems, specially in the domains of diagnostics for decisions. Among its further applications is of combining information from distributed sensors for improving the performance of intrusion detection systems [28]–[30]. In our research we have used Dempster-Shafer's framework to evaluate how the combined information from a number of correlated warning messages received at the supposed destination of a cyber-attack can improve a protection decision making. The warning message we use as input is in effect a high level information about a detected threatening flow traversing an AS.

Section 5 indicates the two main concerns of Dempster-Shafer theory: obtaining degrees of belief for an evidence from its intrinsic probability and Dempster's rule of combinations of such degrees of belief, considering independent sources of evidence. One important question that supports using Dempster-Shafer's framework is the subjective uncertainty of the combined information. The intrusion statement from a reliable source of evidence must be true but it is not necessarily false if such source of evidence is unreliable.

The Dempster-Shaffer's frame of discernment for the scenario depicted in Figure 5 is simply $\Omega = \{intrusion, nointrusion\}$. Lets assume Fs_i in Equation 5 as the belief mass $m_i(A)$ of each correlated message U_i and $(1 - Fs_i)$ representing its disbelief mass $m_i(\Omega)$. From the belief functions Bel_i on frame Ω associated to basic probability functions $m_i(A)$ we can form the combined belief function Bel on Ω with basic probability function $m_C(A)$ through Equation 3.

Table 2 helps to understand the application of Equation 3 for obtaining the combined belief masses $m_C(A)$ from three sources of evidence $m_1(A)$, $m_2(A)$ and $m_3(A)$ according to the notation proposed in the previous paragraph.

TABLE 2. EVALUATION OF THE COMBINED BELIEF MASS FROM EQUATION 3 FOR THREE INDEPENDENT SOURCES OF EVIDENCE $m_1(A)$, $m_2(A)$ AND $m_3(A)$

Fusion 1	$m_1(A) = Fs_1$	$m_1(\Omega) = 1 - Fs_1$
$m_2(A) = Fs_2$	$Fs_1 Fs_2$	$Fs_2(1 - Fs_1)$
$m_2(\Omega) = 1 - Fs_2$	$Fs_1(1 - Fs_2)$	$(1 - Fs_2)(1 - Fs_1)$
Fusion 2	$m_C(A) = \alpha$	$m_C(\Omega) = \beta$
$m_3(A) = Fs_3$	αFs_3	βFs_3
$m_3(\Omega) = 1 - Fs_3$	$\alpha(1 - Fs_3)$	$\beta(1 - Fs_3)$

The notations used in Table 2 are defined bellow.

$$m_C(A) = Fs_1 Fs_2 + Fs_2(1 - Fs_1) + Fs_1(1 - Fs_2) = \alpha$$
$$m_C(\Omega) = (1 - Fs_2)(1 - Fs_1) = \beta$$
$$m_{C_1}(A) = \alpha Fs_3 + \beta Fs_3 + \alpha(1 - Fs_3)$$
$$m_{C_1}(\Omega) = \beta(1 - Fs_3)$$

Extending the evaluation procedure showed in Table 2 for n sources of evidence and considering that there isn't conflicting propositions ($K = 0$), it is possible to calculate the combined degree of belief Bel through Equation 6.

$$Bel = 1 - [1 - m_1(A)] \times [1 - m_2(I)] \times ... \times [1 - m_n(I)] \tag{6}$$

The belief degree in Equation 6 also means the probability that at least one message received at SIEM is reliable. It measures the degree of reliability of the final information of an intrusion being detected from a combination of n correlated messages.

Figures 6 and 7 shows the combined degree of belief behavior considering the mean belief mass $m(A) = (Fs_1 + Fs_2 + ... + Fs_t)/t$ and the number of correlated messages n. In this case Equation 6 can be rewrited as follows.

$$Bel = 1 - [1 - m(A)]^n \tag{7}$$

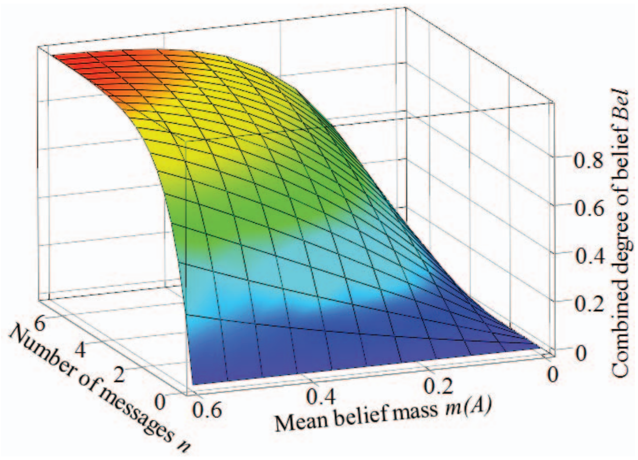


Figure 6. Degree of belief of a final intrusion information Bel combining n correlated messages and the mean value of the belief mass $0 \leq m(A) \leq 0.6$ of each IDS from Equation 7.

As can be seen in Figure 6, the highest value of the combined degree of belief ($Bel < 1$) is obtained with correlated messages from 6 different IDSs with mean belief mass $m(A) = 0.6$.

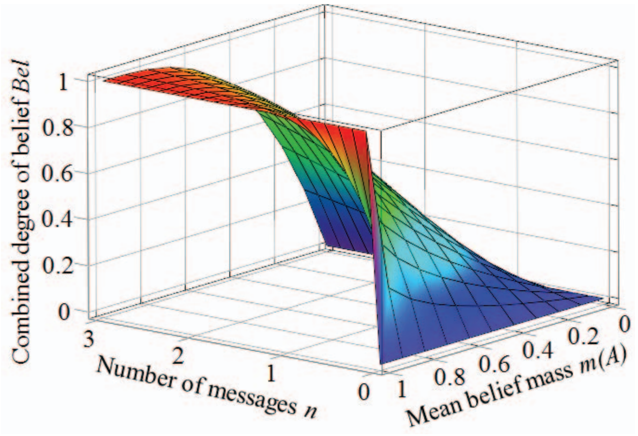


Figure 7. Degree of belief of final intrusion information Bel combining $0 \leq n \leq 3$ correlated messages and the mean value of the belief mass $m(A)$ of each IDS from Equation 7.

Figure 7 depicts the variation of combined degree of belief ($Bel < 1$) with higher values of mean belief mass $m(A)$. Even for a fewer warning messages, combined degree of belief increases faster. It aims to show how important is the detection technology used on each federated IDS.

8. Simulation Results

In order to test the RFC 5575 capabilities and to analyze the format of the extended BGP update messages received at the destination, we have developed a simulation model whose topology is depicted in Figure 8.

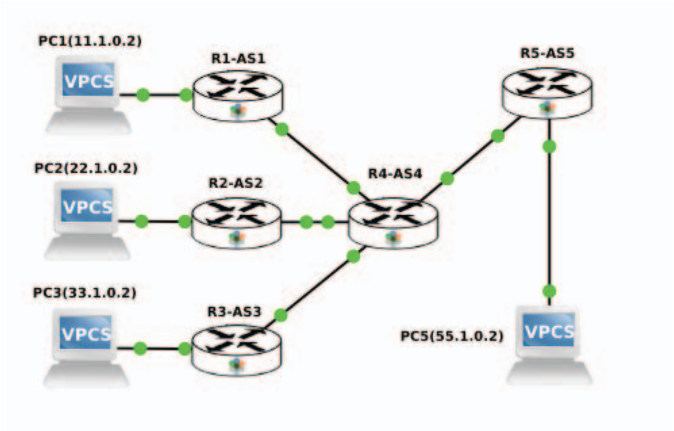


Figure 8. Network topology comprising 4 border routers, with each router in its respective Autonomous System (AS), simulating a coordinated ICMP attack against $PC5$ from $PC1$, $PC2$, and $PC3$.

The test scenario is modeled using the GNS3 [31] simulation tool installed on Ubuntu 16.04. We use the Junos 12.1 image to run virtual machines for each router. After configuring flow advertisements on router $R1$, $R2$ and $R3$, we capture the BGP update messages at the WAN interface of border router 5. The Wireshark logs below summarizes the NLRI field of three BGP update messages that can be collected at the $AS5$. It can be seen that the 3 messages can be correlated with each other according to their destination address (55.1.0.2/32) and protocol. This viabilizes our proposal of relying on BGP to interconnect distributed intrusion detection elements.

```
Network layer reachability information(19 bytes)
 FLOW_SPEC_NLRI (19 bytes)
   NRLI length: 18
      Filter: Destination prefix filter (55.1.0.2/32)
      Filter: Source prefix filter (11.1.0.2/32)
      Filter: Source prefix filter (22.1.0.2/32)
      Filter: Source prefix filter (33.1.0.2/32)
      Filter: IP protocol filter (=1)
      Filter: ICMP type filter (=8)
```

9. Conclusion

Anticipating protection measures at the destination is accomplished by correlating BGP updates as warning messages received from different IDSs along the intrusion path. The number of correlated messages at the destination can be also used to infer the severity of an incoming threat, since the belief level of the combined information tends to be higher.

As mentioned in Section 1, a relevant factor for the success of distributed intrusion detection system is its number of federated IDSs and how they are placed across networks. Another important factor that influences zero-day attacks detection is the heterogeneity of the IDS methodologies and their capacity of learning the behavior of their own network traffic. We argue that the more heterogeneous is the detection method the wider the detection range and the greater the likelihood of detecting zero-day attacks. The use of BGP capabilities for distributing cooperative warning messages of in-transit intrusions makes the system lightweight and plug-and-play which motivates federating new members. On the other hand, the independence and autonomy of each federated IDS to perform its own detection procedure as well as the cooperative behavior of the system as a whole, assures heterogeneity for extend the detection range.

As future works we intend to continue with the development of the simulation model in order to consolidate our approach and validate key performance indicators. Enhancing the detection system by including attack types identification is part of our plans as well.

Acknowledgments

We thank our colleague Celso Alves Vieira from Oi who provided practical BGP insights and expertise that greatly assisted the research.

References

[1] B. M. Leiner, V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. Wolff, “A brief history of the Internet,” *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 5, pp. 22–31, 2009.

[2] T. Bass, “Intrusion Detection Systems and Multisensor Data Fusion,” *Commun. ACM*, vol. 43, no. 4, pp. 99–105, Apr. 2000.

[3] (2017) Cert Coordination Center. Http://www.cert.org/.

[4] L. Bilge and T. Dumitras, “Before We Knew It: An Empirical Study of Zero-day Attacks in the Real World,” in *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, ser. CCS ’12. New York, NY, USA: ACM, 2012, pp. 833–844.

[5] (2016) Kaspersky DDOS intelligence report for Q3 2016. Https://securelist.com/kaspersky-ddos-intelligence-report-for-q3-2016/76464/.

[6] J. Kim and P. Bentley, “The artificial immune model for network intrusion detection,” in *7th European congress on intelligent techniques and soft computing (EUFIT’99)*, vol. 158. Citeseer, 1999.

[7] ——, “The Human Immune System and Network Intrusion Detection.” Citeseer, 2001, pp. 1244–1252.

[8] V. Yegneswaran, P. Barford, and J. Ullrich, “Internet Intrusions: Global Characteristics and Prevalence,” *SIGMETRICS Perform. Eval. Rev.*, vol. 31, no. 1, pp. 138–147, Jun. 2003.

[9] J. S. Balasubramaniyan, J. O. Garcia-Fernandez, D. Isacoff, E. Spafford, and D. Zamboni, “An architecture for intrusion detection using autonomous agents,” in *Computer security applications conference, 1998. Proceedings. 14th annual*. IEEE, 1998, pp. 13–24.

[10] J. D. Case, M. Fedor, M. L. Schoffstall, and J. Davin, “Simple Network Management Protocol (SNMP),” pp. 1–36, May 1990.

[11] F. Cuppens and A. Miege, “Alert correlation in a cooperative intrusion detection framework,” in *Proceedings 2002 IEEE Symposium on Security and Privacy*, 2002, pp. 202–215.

[12] C. Kruegel, F. Valeur, G. Vigna, and R. Kemmerer, “Stateful intrusion detection for high-speed networks,” in *Proceedings of the 2002 IEEE Symposium on Security and Privacy*, ser. SP ’02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 285–.

[13] R. Janakiraman, M. Waldvogel, and Q. Zhang, “Indra: A peer-to-peer approach to network intrusion detection and prevention,” in *12th IEEE International Workshops on Enabling Technologies (WETICE 2003), Infrastructure for Collaborative Enterprises, 9-11 June 2003, Linz, Austria*, 2003, pp. 226–231.

[14] V. Yegneswaran, P. Barford, and S. Jha, “Global Intrusion Detection in the DOMINO Overlay System,” in *In Proceedings of Network and Distributed System Security Symposium (NDSS)*, 2004.

[15] O. Igbe, I. Darwish, and T. Saadawi, “Distributed Network Intrusion Detection Systems: An Artificial Immune System Approach,” in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, June 2016, pp. 101–106.

[16] S. R. Snapp, J. Brentano, G. V. Dias, T. L. Goan, L. T. Heberlein, C.-L. Ho, K. N. Levitt, B. Mukherjee, S. E. Smaha, T. Grance *et al.*, “DIDS (distributed intrusion detection system)-motivation, architecture, and an early prototype,” in *Proceedings of the 14th national computer security conference*, vol. 1. Washington, DC, 1991, pp. 167–176.

[17] R. Robbins, “Distributed Intrusion Detection Systems: An Introduction and Review,” InfoSec Reading Room - SANS Institute, Tech. Rep., Jan 2002.

[18] S. Shiva, C. Simmons, C. Ellis, D. Dasgupta, S. Roy, and Wu, “AVOIDIT: A cyber attack taxonomy.” University of Memphis, Tech. Rep., August, 2009.

[19] S. Axelsson, “Intrusion detection systems: A survey and taxonomy,” Technical report, Tech. Rep., 2000.

[20] P. R. Marques, J. Mauch, N. Sheth, B. Greene, R. Raszuk, and D. R. McPherson, “Dissemination of Flow Specification Rules,” pp. 1–22, Aug 2009.

[21] T. Bates, R. Chandra, D. Katz, and Y. Rekhter, “Multiprotocol Extensions for BGP-4,” pp. 1–12, January 2007.

[22] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.

[23] A. Monemi, R. Zarei, and M. N. Marsono, “Online NetFPGA decision tree statistical traffic classifier ,” *Computer Communications*, vol. 36, no. 12, pp. 1329 – 1340, 2013.

[24] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.

[25] D. Hock and M. Kappes, “A Self-Learning Network Anomaly Detection System using Majority Voting,” in *Tenth International Network Conference, INC 2014, Plymouth, UK, July 8-10, 2014. Proceedings*, 2014, pp. 59–69.

[26] A. Goodney, S. Narayan, V. Bhandwalkar, and Y. H. Cho, “Pattern based packet filtering using NetFPGA in DETER infrastructure,” in *1st Asia NetFPGA developers workshop. Daejeon, Korea*, 2010.

[27] Nokia, “Nokia 7750 Service Router and 7450 Ethernet Service Switch Integrated Service Adapters,” https://resources.ext.nokia.com/asset/157673, 2015, original Datasheet.

[28] Y. Wang, H. Yang, X. Wang, and R. Zhang, “Distributed intrusion detection system based on data fusion method,” in *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on*, vol. 5. IEEE, 2004, pp. 4331–4334.

[29] V. Shah, A. K. Aggarwal, and N. Chaubey, “Performance improvement of intrusion detection with fusion of multiple sensors,” *Complex & Intelligent Systems*, vol. 3, no. 1, pp. 33–39, 2017.

[30] V. M. Shah and A. K. Agarwal, “Reliable Alert Fusion of Multiple Intrusion Detection Systems,” *I. J. Network Security*, vol. 19, no. 2, pp. 182–192, 2017.

[31] J. C. Neumann, *The Book of GNS3*, 1st ed. San Francisco, CA, USA: No Starch Press, 2014.