Collaborative Attack Detection in High-Speed Networks

Martin Rehák[1], Michal Pěchouček[2], Pavel Čeleda[3], Vojtěch Krmíček[3], Pavel Minařík[3], and David Medvigy[2]

[1] Center for Applied Cybernetics, Faculty of Electrical Engineering
[2] Department of Cybernetics, Faculty of Electrical Engineering,
Czech Technical University in Prague Technická 2, 166 27 Prague, Czech Republic
{mrehak,pechouc}@labe.felk.cvut.cz
[3] Institute of Computer Science, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
{celeda,vojtec}@ics.muni.cz

Abstract. We present a multi-agent system designed to detect malicious traffic in high-speed networks. In order to match the performance requirements related to the traffic volume, the network traffic data is acquired by hardware accelerated probes in NetFlow format and preprocessed before processing by the detection agent. The proposed detection algorithm is based on extension of trust modeling techniques with representation of uncertain identities, context representation and implicit assumption that significant traffic anomalies are a result of potentially malicious action. In order to model the traffic, each of the cooperating agents uses an existing anomaly detection method, that are then correlated using a reputation mechanism. The output of the detection layer is presented to operator by a dedicated analyst interface agent, which retrieves additional information to facilitate incident analysis. Our performance results illustrate the potential of the combination of high-speed hardware with cooperative detection algorithms and advanced analyst interface.

1 Introduction

The purpose of the presented work is to deliver an autonomous system able to detect malicious traffic on high-speed networks and to alert the operators efficiently. While the system reasoning is based on intelligent agents and multi-agent methods, the network traffic data acquisition and preprocessing in both dedicated adaptive hardware and specialized software is essential for project success, as the traditional agent techniques are not well suited for efficient low-level traffic processing.

In the work presented in this paper, we aggregate the network data to capture the information about network flows, unidirectional components of TCP connections (or UDP/ICMP equivalent) identified by shared source and destination addresses and ports, together with the protocol, and delimited by the time frame used for data acquisition (see Section 2.1). This information provides no hint about the content of the transmitted data, but by detecting the anomalies in the list of flows acquired over the monitoring period, we can detect the anomalies and possible attacks, albeit with limited effectiveness.

H.-D. Burkhard et al. (Eds.): CEEMAS 2007, LNAI 4696, pp. 73–82, 2007.

In order to detect an attack from the flow information on the backbone level, especially without any feedback from the affected hosts, we have to analyze the patterns in the traffic data, compare them with normal behavior and conclude whether the irregularity corresponds to a known attack profile or not. This approach to Network Intrusion Detection, typically based on the flow information captured by network flow monitor is currently an important field of research into *anomaly based intrusion detection*. Numerous existing systems, based on traffic volume analysis modeled by Principal Component Analysis methods [1], models of entropy of IP header fields for relevant subsets of traffic [2,3], or just count of the flows corresponding to the selected criteria [4] offer each a particular valid perspective on the network traffic. In our approach, we have decided not to develop a novel detection method, by rather to integrate each of the methods with an extended trust models of a specialized agent. This combination allows us to correlate the results of the used methods and to combine them to improve their effectiveness. Most anomaly detection methods today are not fit for commercial deployment due to the high ratio of false positives (i.e. legitimate traffic classified as malicious) or false negatives (malicious traffic classified as legitimate). While their current level of performance is a valid scientific achievement, the costs associated with supervision of such systems are prohibitive for most organizations. Therefore, our main research goal is to combine the efficient low-level methods for traffic observation, with multi-agent detection process to detect the attacks with comparatively lower error rate (see Table 1), and to provide the operator with efficient incident analysis layer presented in Section 2.3. Analysis layer supports operator's decisions about detected anomalies by providing additional information from related data sources. It is also responsible for visualization of the anomalies and the detection layer status.

2 Architecture

This section details the design of the system architecture. We present an overview of key techniques and technologies for each layer, and we also present the motivations behind the selection of individual techniques or methods.

The architecture consists of several layers with varying requirements on on-line processing characteristics, level of reasoning and responsiveness. While the low-level layers need to be optimized to handle network traffic at wire speed during the traffic acquisition and preprocessing, the higher layers will use the preprocessed data to infer the conclusions regarding the degree of anomaly and consecutively also the maliciousness of the particular flow or a group of flows. Therefore, while the computation in the higher layers must be still reasonably efficient, the preprocessing by the lower layers allows us to deploy more sophisticated algorithms.

System can be split into these layers, as shown in Figure 1:

2.1 Traffic Acquisition and Preprocessing Layer

The traffic acquisition and preprocessing layer is responsible for acquiring network traffic, preprocessing data and providing traffic characteristics to upper system layers. We use the flow characteristics, based on information from packet's headers.

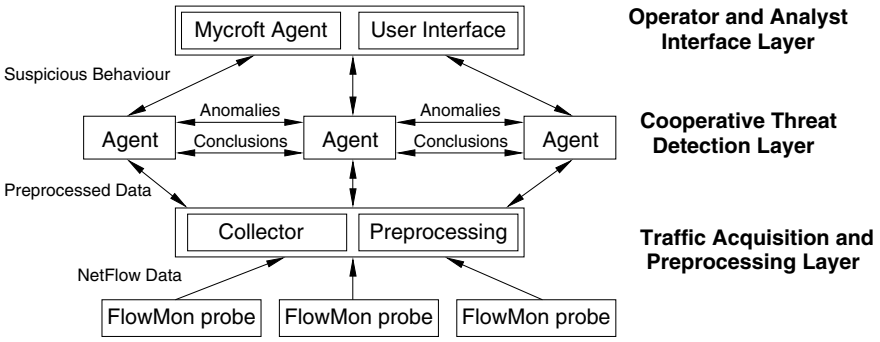

Fig. 1. System overview, with network probes, acquisition and preprocessing layer at the bottom, agent platform for anomalies detection in the middle and visualization on the top

In general, flows are a set of packets which share a common property. The simplest type of flow is a 5-tuple, with all its packets having the same source and destination IP addresses, port numbers and protocol. Flows are unidirectional and all their packets travel in the same direction. For the flow monitoring we use NetFlow protocol developed by Cisco Systems.

The amount of traffic in nowadays high-speed networks increases continuously and traffic characteristics change heavily in time (network throughput fluctuation due to time of day, server backups, DoS attacks, scanning attacks, etc.). Performance of network probes must be independent of such states and behave reliably in all possible cases. The quality of provided data significantly effects the upper layers and chances to detect traffic anomalies.

Therefore we use hardware accelerated NetFlow probes FlowMon [5] a passive network monitoring device based on the COMBO hardware [6], which provides high performance and accuracy. The FlowMon probe is preferred due to implemented features which contains packet/flow sampling, several sampling types, active/inactive timeouts, flow filtering, data anonymization, NetFlow protocol version 5 and 9 support. The FlowMon probe handles 1 Gb/s traffic at line rate in both directions and exports acquired NetFlow data to different collectors. Detailed evaluation of these crucial capabilities is described in Section 3.

The collector stores incoming packets with NetFlow data from FlowMon probes into database. The collector provides interface to graphical representation of network traffic, flow filtration, aggregation and statistics evaluation, using source and destination IP addresses, ports and protocol.

To process acquired IP flows by upper system layers the preprocessing must be performed on several levels and in different manners. Packets can be sampled (random, deterministic or adaptive sampling) on input and sampling information is added to NetFlow data. On the collector side the same flows are aggregated to reduce the amount of data without information loss and several statistic characteristics (average traffic values, entropy of flows) are computed.

After deploying probes in monitored network, the probes can be reprogrammed to acquire new traffic characteristics. The system is fully reconfigurable and the probes can

adapt their features and behavior. As we can see in Section 3, presented solution can acquire unsampled flow data from even very fast links. The proposed traffic acquisition and preprocessing layer is able to provide real-time traffic characteristics to detect anomalies by upper system layers.

2.2 Cooperative Threat Detection Layer

Cooperative threat detection is based on the principles of trust modeling [7] that are an integral part of agent research. However, there are three important features [8] that must be added to trust modeling to cover our domain-specific requirements:

• **Uncertain Identity Modeling:** Baseline trust models evaluate the behavior of individual agents, whose identity is guaranteed (to an extent) by the multi-agent platform or similar computational environment. In the network domain, we have to evaluate the trustfulness of network flows, and while they can be distinguished as unique identities, this distinction is unpractical from the intrusion detection perspective. We represent the connections in a metric space, and use the associated distance function to assess the similarity of the flow representations in this space. All detection agents use the same NetFlow quintuple to construct the identity representations, but may obtain different results regarding the similarity due to the use of different distance functions. For example, one agent can emphasize the similarity of srcIP address (and likely host), while the others may concentrate on ports that are more application specific. This variability makes the agent perspective on the system multi-faceted, and the attacks are less likely to avoid multiple agents.

• **Context Modeling:** Merely representing the flow identities in the metric space and evaluating their trustfulness gives unsatisfactory results, as it ignores the most important information from the NetFlow data – the information about the related flows in the current traffic sample. This information constitutes the context of the trusting decision [9], and together with the identity defines the Identity-Context metric space, where the detection agents assess the trustfulness of flow representations. Each of the agents uses its own particular context space, typically based on the existing anomaly detection methods. For instance, we can complement the information about the flow by the number of flows from the same srcIP and same dstPrt, or with an entropy of dstIP addresses computed over all flows with the same srcIP. The use of this information is twofold; each agent uses it to place the flow representations in the Identity-Context space of its trust model, and in the same time to provide the information about the degree of flow anomaly to other trusting agents.

• **Implicit Feedback Mechanism:** The principal input of classic trust models is a result of past cooperations with the partner: quality of service, degree of success, on-time delivery and other domain specific parameters. In our case, the system is deployed on backbone network and it is very difficult to obtain the feedback that can be associated with the current traffic on the network; cooperative IDS (like `dshield.org`) typically provide unsynchronized feedback, and not all the threats are Internet-wide. Obtaining the feedback from the connected operators or organizations is even more difficult: while the IETF has several working groups focusing on incident response interoperability, the bulk of the work is not suitable for real-time data processing, and concentrates on human-

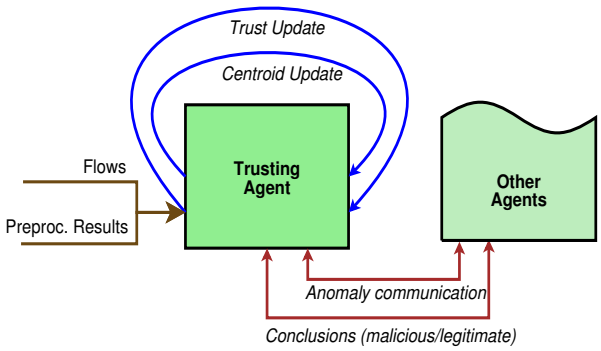

Fig. 2. Overview of detection (trusting) agent operations

to-human interaction. Therefore, we use the information regarding the flow anomaly *as assessed by the other agents* to replace the direct feedback, therefore connecting the anomaly detection between diverse agents.

While processing the information about the network flows, each trusting agent receives an identical copy of network flows list and associated pre-extracted statistics. Then, it uses its specific preprocessing to determine the anomaly of each flow (in most cases working only with already extracted statistics) and to communicate the list of anomalies to other agents. In its turn, the agent also receives the anomalies from the others, and starts the flow processing by its internal trust model. As we have implicitly suggested above, the trustfulness is not associated to individual flows, but rather to selected objects in the Identity-Context space. Individual flow is therefore represented by its identity (i.e. NetFlow quintuple) and the associated context is retrieved to determine its position in the Context subspace. Then, we retrieve the positions of nearby centroids from the current trust models and update their trustworthiness with an aggregated degree of flow anomaly as determined by the other agents. When there is no appropriate cluster in the vicinity of the observed flow, a new cluster is created. The details of the approach are presented in [8].

The decision whether a given flow is trusted or untrusted depends on the typical degree of anomaly in the observed network traffic. This parameter varies widely with network type – the number of anomalies is typically low on corporate or government WANs, but is significantly higher on public Internet backbone, or in the university networks. To avoid the problems with manual tuning of the system, we use a fuzzy-inference process integrated with the trust model to decide whether the given flow is malicious, by computing its inference with Low and High trust values. These values are determined similarly to the trustfulness of individual flow representations, but are represented as two fuzzy intervals [10].

2.3 Operator and Analyst Interface Layer

The Cooperative Threat Detection Layer is coordinated by a super-agent Mycroft. Mycroft is again a multi-agent system. This system is constructed for context based inference over information synthesized from various data sources. The name of this system refers

to the so called Mycroft problem well known from Doyle's Mycroft Holmes - a Sherlock Holmes' brother. Every detected suspicious behavior on the network is reported to the agent Mycroft by the detection layer. Mycroft opens the new case subsequently and retrieves relevant information from data sources to which it is connected. Then the network operator can explore and evaluate the reported case together with contextual information retrieved from connected data sources. One of the main Mycroft's abilities is the adaptability which consists in:

- integration of data sources relevant to observed network,
- adaptation to detection layer status,
- adaptation to current traffic situation on the network,
- personalization for given network operator.

All these adaptations are possible thanks to the following construction: Individuals are classified into categories to express their properties. The individual can be any object relevant to the suspicion detection or any elementary relation between such objects relevant to the suspicion detection. Individuals are classified into categories with a measure from the interval $< -1, 1 >$. Value -1 stands for "certainly not in given category", value 1 stands for "certainly is in given category" and value 0 stands for "cannot decide". This classification is called the elementary fact and is realized always in some context. The context is nothing else than a set of elementary facts. All the information retrieved from data sources is disassembled into elementary facts. This approach allows data utilization in a way that fits actual situation. Presented construction is a straight extension of the so called diamond of focus presented in [11] for the first time. Formal definition of systems referred to as Knowledge and Information Robots is provided in [12]. Multi-agent system Mycroft is one of the representatives of Knowledge and Information Robots class.

Different data sources can be used to support network operator's evaluations and decisions. Interoperability with data sources means that The multi-agent system Mycroft can be taught what kind of information given data source contains, how to combine this information with all other information that system already knows or is capable to acquire and how to access this information. Communication with data source is realized by the set of technical adapters. Learning methods are used throughout the Mycroft agent to provide the adaptation functionality. Teaching is realized using quite formalized natural language instead of manual customization and programming.

Adaptability on the detection layer status means that the multi-agent system Mycroft monitors current status of each agent present in the detection layer and is able to communicate with him. Main purpose of this communication is agent reconfiguration according to operator's demands, current situation in the monitored network or some other specified rules. These rules are taught in the same way as data sources using quite formalized natural language and can be modified.

Adaptability on the current network traffic means that the multi-agent system Mycroft can choose suitable data sources and provide additional knowledge to support operator's decisions. Using additional knowledge provided by these data sources Mycroft can perform basic evaluations of the traffic situation and present this evaluations to the network operator concurrently with the detection report. In some basic cases, using prelearned transmutation patterns, system can even evaluate this suspicion as false positive

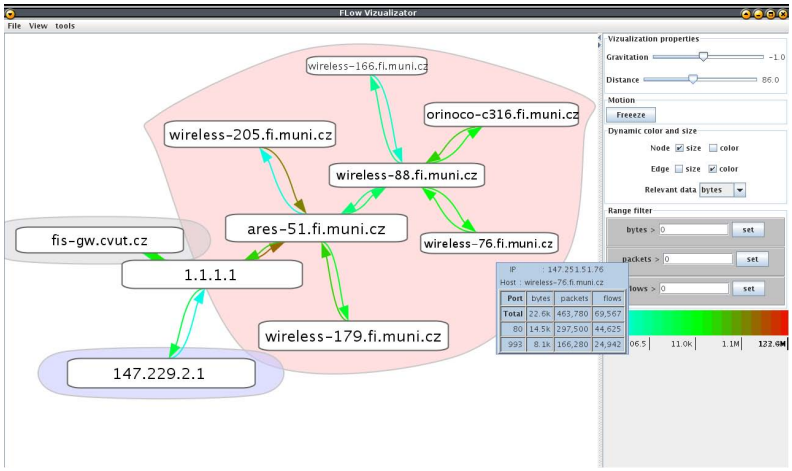

Fig. 3. Mycroft's flow data visualization example. Oval nodes in graph represent network devices identified by IP addresses. Arrows show a direction of the flows.

and drop it. Again this is possible by the means of the knowledge provided by suitable data sources.

Personalization for given operator means that the system is able to follow operator's procedures, habits and preferences. System learns during routine work. When operator is not satisfied he can ask the system for different presentation of given information.

3 System Evaluation and Performance

The performance is becoming a key concern of NIDS *(Network Intrusion Detection Systems)* in high-speed networks. The results of traffic acquisition and processing vary depending on the amount of acquired data. Numerous existing NIDS are based on commodity hardware with open-source software and very limited hardware acceleration.

The Figure 4 shows flow statistics for various IP packet's sizes transmitted on Ethernet at a rate of a gigabit per second. The tests were performed with the Spirent AX/4000 broadband test system [13]. The test system generated 5 flows with 500000 packets per flow for each packet size. The results of FlowMon probe correspond to the maximal theoretical throughput on gigabit Ethernet network. On the other hand the results of software NetFlow probe (nProbe [14]) are misrepresented for small IP packets. The nProbe was used on Linux OS with Intel PCI-X network interface card and default kernel configuration.

Existing flow monitoring systems are mostly based on exporting flow data from routers. The routers are dedicated for routing the data in networks and enabling the flow export has often negative impacts on overall router performance especially during attacks. The exported flows are sampled or limited to maximal number of exported

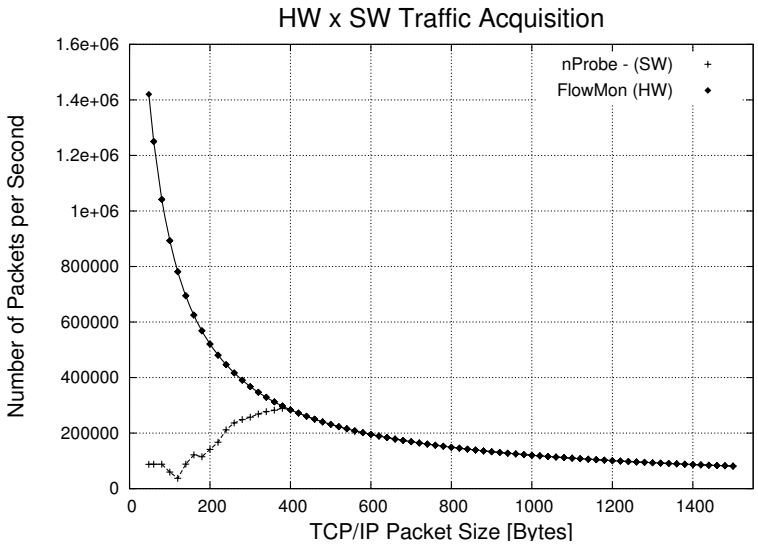

Fig. 4. Flow statistics acquired by SW and HW accelerated probe

flows per second e.g. 5000 flows/s. In comparison with FlowMon probe, user defined extensions can't be added to the routers and the adaptability is very limited.

In our work we are focusing on the impact of packet sampling on anomaly detection. The articles [15,16] study whether existing sampling techniques distort traffic features that are critical for effective anomaly detection. They show that packet sampling methods (random packet sampling, random flow sampling, smart sampling, and sample-and hold sampling) introduce fundamental bias that degrades the performance of the anomaly detection algorithms. To avoid such a misbehavior the FlowMon probe provides non-sampled data, without packet loss at a line rate.

The Table 1 shows the performance of multi-agent system. The observed network traffic is processed by several layers to handle high amount of data in nowadays networks. Each layer has specific physical and performance limits e.g. number of incidents which can be handled by human operator. To overcome such limitations, the system is fully scalable and all layers can be distributed. The multi-agent system adapts the detection behavior to reduce the number of false positives and negatives so the final number of incidents fits the limits of human operator.

4 Conclusions and Future Work

Our work presents a design of multi-agent system for network intrusion detection that is optimized for deployment on backbone networks. Our system addresses two main limitations of existing intrusion detection systems – efficiency and effectiveness. Deployment on high-speed links implies the need to process the important quantity of

Table 1. Multi-agent system performance overview. The system process backbone link traffic with average load of 800 Mb/s.

Layer	*Processed Data*	
	Input	*Output*
Operator	Security Incidents - incidents at a certain priority levels	Incident Handling - resolving up to 10 high priority incidents per hour
Operator and Analyst Interface Layer	Network Anomalies - detected threads with additional network information	Detected Incidents - priority-based incidents up to 100 incidents/minute
Cooperative Threat Detection Layer	Network Traffic Statistics - aggregated flow statistics up to 100000 flows/minute	Detected Threats - network traffic anomalies up to 10000 threats/minute
Traffic Acquisition and Preprocessing Layer	Network Traffic - packets 125000 packets/s	Flow Statistics - flows 3800 flows/s

data in near real-time, in order to prevent the spread of novel threats. Therefore, the individual agents do not acquire the data from the network directly, but receive the data already preprocessed, with the level of detail that is appropriate for anomaly-based intrusion detection. Each detection agent in the system is based on existing anomaly detection technique, which defines its perception of network flow identities in its trust model. Its private view of the data is complemented by the opinions of other agents regarding the anomaly of flows in the current traffic, therefore collaboratively improving the effectiveness of anomaly detection process. When the agents reach a conclusion regarding the untrustfulness of a particular subset of flows, they submit this observation to user-interface agent that automatically retrieves context information (DNS records, history, etc.) to allow rapid analysis by human supervisors.

At the time of this writing, the first preliminary version of the complete system is being integrated and the whole system is still under active development. Therefore, we don't present any definitive experimental results regarding its effectiveness of the complete system, but only the performance evaluations of critical components at the current integration stage. The data used for system testing are acquired on Masaryk University network, connected to the Czech national educational network (CESNET). In our future work, we will provide detailed experimental evaluation of system deployment, and also analyze its performance in countering a wide selection of currently observed malicious traffic.

Acknowledgment. This material is based upon work supported by the European Research Office of the US Army under Contract No. N62558-07-C-0001. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the European Research Office of the US Army. Also supported by Czech Ministry of Education grants 1M0567 and 6840770038.

References

1. Lakhina, A., Crovella, M., Diot, C.: Characterization of Network-Wide Anomalies in Traffic Flows. In: ACM SIGCOMM conference on Internet measurement IMC '04, pp. 201–206. ACM Press, New York (2004)
2. Xu, K., Zhang, Z.L., Bhattacharrya, S.: Reducing Unwanted Traffic in a Backbone Network. In: USENIX Workshop on Steps to Reduce Unwanted Traffic in the Internet (SRUTI), Boston, MA (2005)
3. Lakhina, A., Crovella, M., Diot, C.: Mining Anomalies using Traffic Feature Distributions. In: ACM SIGCOMM, Philadelphia, PA, August 2005, pp. 217–228. ACM Press, New York (2005)
4. Ertoz, L., Eilertson, E., Lazarevic, A., Tan, P.N., Kumar, V., Srivastava, J., Dokas, P.: MINDS - Minnesota Intrusion Detection System. In: Next Generation Data Mining, MIT Press, Cambridge (2004)
5. Čeleda, P., Kováčik, M., Koníř, T., Krmíček, V., Špringl, P., Žádník, M.: FlowMon Probe. Technical Report 31/2006, CESNET, z. s. p. o. (2006), `http://www.cesnet.cz/doc/techzpravy/2006/flowmon-probe/`
6. CESNET, z. s. p. o.: Family of COMBO Cards (2007), `http://www.liberouter.org/hardware.php`
7. Sabater, J., Sierra, C.: Review on computational trust and reputation models. Artif. Intell. Rev. 24, 33–60 (2005)
8. Rehak, M., Pechoucek, M.: Trust modeling with context representation and generalized identities. In: Cooperative Information Agents XI. LNAI/LNCS, vol. 4676, Springer, Heidelberg (2007)
9. Rehak, M., Gregor, M., Pechoucek, M., Bradshaw, J.M.: Representing context for multiagent trust modeling. In: IAT'06. IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2006 Main Conference Proceedings), Los Alamitos, CA, USA, pp. 737–746. IEEE Computer Society, Los Alamitos (2006)
10. Rehák, M., Foltýn, L., Pěchouček, M., Benda, P.: Trust Model for Open Ubiquitous Agent Systems. In: Intelligent Agent Technology. 2005 IEEE/WIC/ACM International Conference, vol. PR2416, IEEE, Los Alamitos (2005)
11. Staníček, Z.: Universal Modeling and IS Construction. PhD thesis, Masaryk University, Brno (2003)
12. Procházka, F.: Universal Information Robots a way to the effective utilisation of cyberspace. PhD thesis, Masaryk University, Brno (2006)
13. Spirent, C.: Spirent AX/4000 Broadband Test System (2007), `http://www.spirentcom.com/`
14. Deri, L.: nProbe - An Extensible NetFlow v5/v9/IPFIX GPL Probe for IPv4/v6 (2007), `http://www.ntop.org/nProbe.html`
15. Mai, J., Chuah, C.N., Sridharan, A., Ye, T., Zang, H.: Is sampled data sufficient for anomaly detection? In: IMC '06. Proceedings of the 6th ACM SIGCOMM on Internet measurement, pp. 165–176. ACM Press, New York (2006)
16. Brauckhoff, D., Tellenbach, B., Wagner, A., May, M., Lakhina, A.: Impact of packet sampling on anomaly detection metrics. In: IMC '06. Proceedings of the 6th ACM SIGCOMM on Internet measurement, pp. 159–164. ACM Press, New York (2006)