

06.09.07

Deliverable DJ2.2.4: Findings of the Advanced Anomaly Detection Pilot



Deliverable DJ2.2.4

Contractual Date:	31/08/07
Actual Date:	06/09/07
Contract Number:	511082
Instrument type:	Integrated Infrastructure Initiative (I3)
Activity:	JRA2
Work Item:	2
Nature of Deliverable:	R (Report)
Dissemination Level	PU (Public)
Lead Partner	SurfNET
Document Code	GN2-07-218v2

Authors: Chelo Malagon (RED.es), Maurizio Molina (DANTE), Jacques Schuurman (SURFNET)

Abstract

Advanced anomaly detection is the process in which network traffic patterns are continuously and if possible near real-time monitored in order to detect deviations from the ordinary. These deviations usually form a strong first indication that traffic flows on a monitored network are not according to what was to be expected in a normal operational context. By developing sophisticated methods for anomaly detection and deploying these methods in pilot environments, an overall improvement of the security levels of a monitored network can be achieved

Table of Contents

0	Executive Summary	v
1	Introduction	1
2	Advanced Anomaly Detection Methods	3
2.1	BICHOS	3
2.1.1	Short description	3
2.1.2	Reported results	4
2.1.3	Current status and future	4
2.2	Holts-Winter	4
2.2.1	Short description	4
2.2.2	Reported results	4
2.2.3	Current status and future	5
2.3	DDoSVax	6
2.3.1	Short description	6
2.3.2	Reported results	7
2.3.3	Current status and future	7
3	Conclusions and Recommendations	8
3.1	Proposed additional methods	8
3.1.1	NfSen-overflow	9
3.1.2	NetReflex	9
3.1.3	DDoSVax host behaviour method	10
3.2	Other possible methods and/or tools	11
4	References	12
5	Acronyms	13

Table of Figures

Figure 1.1: The place of anomaly analysis (and detection) in the context of the complete Toolset 2

0 Executive Summary

A Toolset has been compiled within JRA2 that enables operational security staff of large IP networks to monitor traffic patterns on their networks and hence assess the underlying, possibly malicious behaviour of systems and/or users.

The paradigm being used is to observe traffic patterns as closely (in real-time) as possible to identify the “normal” behaviour of the network. It is then possible to search for and detect deviations (anomalies) from these normal patterns, which can then be isolated for further investigation as these deviations are strong indicators of malicious traffic.

This report discusses the various methods of detection piloted and tested by the JRA2 development partners, and puts forward recommendations for further testing and development based on the findings.

Project:	GN2
Deliverable Number:	DJ2.2.4
Date of Issue:	06/09/07
EC Contract No.:	511082
Document Code:	GN2-07-218v2

1 Introduction

Within the GÉANT2 project, the Joint Research Activity 2 (JRA2) focuses on the development of security services to help GÉANT users use the network in a secure manner. One of the deliverables is a Toolset to be deployed by security teams and/or network operators to monitor the traffic on their networks (connected to GÉANT2). The basic architecture of the toolset is sketched in figure 1.1. Of particular interest in the scope of this deliverable is the anomaly analysis function. It alerts the network operators of an unusual situation in two ways: either in semi-real time when such a situation is detected on the live network or when investigating a particular situation based on stored patterns.

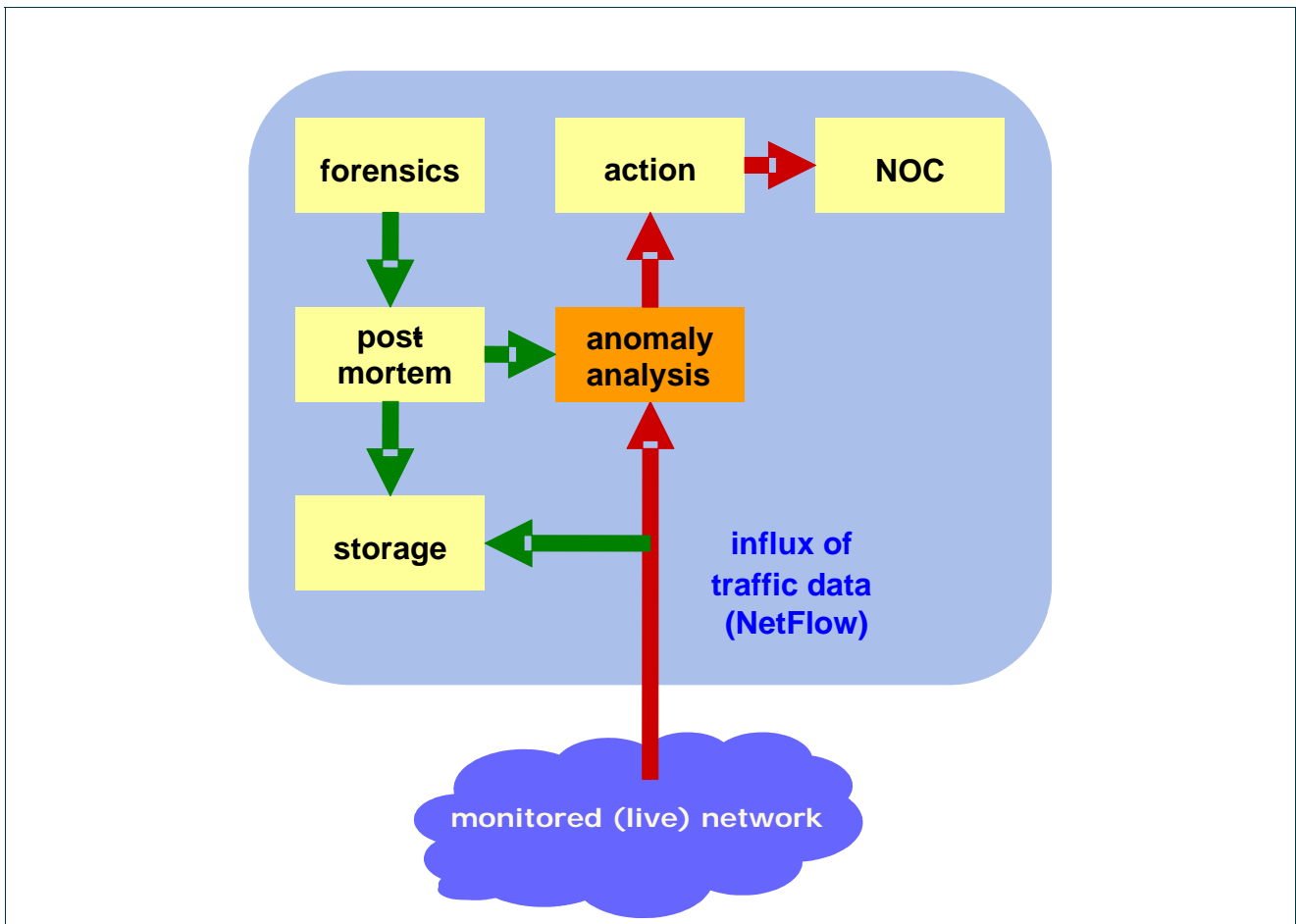


Figure 1.1: The place of anomaly analysis (and detection) in the context of the complete Toolset

The aim of the Toolset is to gain an aggregated view on the traffic patterns on a backbone scale network with a typical sustained load of tens of Gbps. Using earlier acquired knowledge of how these patterns should look like (how the traffic behaves on the monitored part of the network), it is possible to quickly search for deviations of those "normal" patterns. These deviations are called traffic anomalies.

During the third project year, several approaches of anomaly detection were tried, and this interim document aims to give an overview of these initiatives, along with a high-level report on the results. The various initiatives are discussed in Chapter 2. In Chapter 3, some conclusions and recommendations from the group are given as to how these developments could be taken further.

2 Advanced Anomaly Detection Methods

Various methods were tried during the course of the third project year. In Cambridge in January 2007, an overview was presented of these methods, and the stage in which they were tested by various project partners. These methods include:

- BICHOS [BICHOS]
- Holts-Winter
- DDoSVax [DDoSVax]

Each of these methods is discussed in more detail in the remainder of this chapter, along with a summarised report on their testing (if applicable). Also included is a global indication of the status at the time of writing of this report, and possible guidelines for future development.

2.1 BICHOS

2.1.1 Short description

BICHOS (Backbone Information Collector of Harmful Objects Specimens, see <http://www.redirie.es/cert/proyectos/bichos/>) was conceived by the project partner RedIRIS. Its objective is to obtain information about the malicious code that crosses the network at the Network/ISP level before arriving at the end user/customer network, then providing this information to security companies and sharing this information with the security community.

Some ports that are directly blocked at the backbone level (MS windows RPC services and others) are redirected to a collector machine, which after doing a reverse IP NAT acts as the attacked system and captures the binary, using a medium level interaction honey pot.

After collecting the binaries the files are distributed, using encrypted PGP messages to a repository. Information and statistics about the new binaries that are attacking the network are provided, as well as information about the origin of the attacks.

Project:	GN2
Deliverable Number:	DJ2.2.4
Date of Issue:	06/09/07
EC Contract No.:	511082
Document Code:	GN2-07-218v2

2.1.2 Reported results

RedIRIS reported on the successful installation of the BICHOS prototype and its piloting at the JRA2 meeting in Cambridge in January 2007. Despite BICHOS' well-recognised potential in detecting, assessing and quantifying malware, the results did not attract enough interest for further in-depth experimenting. A call for further testing and participation in Cambridge in January 2007 was issued, but yielded no further interest in participation.

2.1.3 Current status and future

Due to the low amount of feedback (and apprehension), JRA2 has decided not to pursue further development of BICHOS during the fourth project year of GÉANT2 JRA2.

2.2 Holts-Winter

2.2.1 Short description

Holts-Winter (HW) analysis uses a basic heuristic method (with NfSen as the underlying vehicle) to give a reliable prediction of expected change in traffic patterns. The method uses the assumption that the relative recent history of a specific (traffic) pattern is an adequate forecast of the patterns to be expected in the (near) future. Holts-Winter becomes more reliable if (a) the length of the time frame of historical data is large enough, and (b) if specific patterns are reappearing consistently. These patterns can be seen as a Fourier series, consisting of the same polynomial components over a longer period of time. The basic implementation was carried out by Gabor Kis (HUNGARNET), and experimentally deployed by four project partners: DANTE, NIIF/HUNGARNET, RedIRIS, SURFnet, and GRNET.

2.2.2 Reported results

GRNET reports that useful HW analysis starts only after collecting a minimal amount of data spanning 10 days. Even then, aberrant network traffic patterns appeared only in the ICMP parts of the traffic, possibly as a result of large port scans. GRNET decided not to include (yet) HW analysis in its production environment.

NIIF/HUNGARNET reports that they have chosen to use an un-sampled feed of NetFlow data from their backbone. Further processing (analysis and graphical representation) of data resulting from a hybrid flow of sampled and un-sampled NetFlow data has proven to be rather difficult.

The HW analysis has been in place since August 2006. The anomaly detection led NIIF/HUNGARNET to have an early warning system, eventually to help improve overall security of the backbone and its constituency. Two flaws were reported; alerts are not generated automatically (human intervention is always required) and alerts need further investigation by knowledgeable network security staff in order to prevent false positives. These include patterns observed during events such as the start of the academic year (notably with remote

registration procedures), network throughput tests conducted by connected sites to the backbone, or remote data backup operations.

DANTE, operating the GÉANT2 backbone in Europe, reports the deployment of flow analysis on 21 routers on all interfaces where GÉANT2 has a peering with an external AS. The sampling rate used is 1:1000, which leads to a sustained NetFlow stream of 1-2 Mbps, monitoring an amount of traffic that peaks at 25 Gbps. They started the deployment of HW analysis in the beginning of February 2007. It should be noted that DANTE does not currently use HW analysis (or any NetFlow analysis) to proactively trigger security alerts, but rather to investigate, in a reactive mode, alerts that were brought to the attention of their security response team (DANCERT).

RedIRIS reported the deployment of HW in a laboratory environment, using an indirect NetFlow feed from 34 router sources. In contrast to other partners, RedIRIS has tested HW in a hybrid environment of sample rates, varying from 1:1 to 1:800.(covering the range from full rate to the maximum recommended rate for Juniper routers) RedIRIS has deployed HW from mid October 2006 until mid February 2007. The best results reported by RedIRIS were anomalies resulting from a sudden decline (or complete death) of specific flows. In the context of a migration (to RedIRIS10), the unexpected cessation of specific flows, detected by HW analysis, quickly identified configuration errors, resulting in downtime for parts of the backbone.

On the other hand, RedIRIS also reported that anomalies associated to UDP traffic were detected more often by HW than those associated with other protocols. These UDP anomalies were real incidents in most of the cases.

2.2.3 Current status and future

NIIF/HUNGARNET asserts that there are (probably) not many false negatives (occurring traffic anomalies that remain undetected by HW analysis). DANTE agrees with this observation. This assertion should actually be confirmed by an expert assessment, or perhaps another method for anomaly detection run in parallel on the backbone. A suggested approach is to show all the anomalies for the same source in a single view and introduce sub-views for different protocols (TCP, UDP, ICMP, other).

Some short-term improvements are to be implemented in-house by NIIF/HUNGARNET:

- Automatic alerting (if possible, by using the generic alerting functionality offered by NFsen).
- Integration into the upcoming Toolset version (if possible, by relying on the plug-in API of NFsen).
- Comparison with the data-mining solution NetReflex, subject to its availability.

Improvements regarding the (CPU) load needed to complete an analysis stage are highly desirable. Especially as using NfSen's features for aggregation and sorting with respect to either large datasets or a longer timespan leads to disappointingly slow results.

Regarding false positives, some enhancements would help in reducing the alerts to a number more manageable by a security response team. Some of the alerts turn out to be "false" (that is due to non malicious

Project:	GN2
Deliverable Number:	DJ2.2.4
Date of Issue:	06/09/07
EC Contract No.:	511082
Document Code:	GN2-07-218v2

traffic) only after investigation. There is probably little that can be done to improve this. But for some other areas, enhancements are possible:

- When the anomaly has a "long lasting step" shape (i.e. an abrupt traffic increase, followed by a period of sustained traffic, followed by a return to the original level) HW raises two separate alerts, on the raising edge and on the decreasing edge. It would be preferable to have just one, i.e. the alert staying "on" for all the period of sustained traffic.
- Unfortunately, gaps in the NetFlow data exporting/collection resulting in short periods of zero traffic are quite frequent, at least in DANTE's configuration. HW always triggers alerts for this situation. These anomalies are actually the majority. It is suggested that a way is found to differentiate these anomalies from those triggered by sudden increase of traffic from a base level.
- Anomalies associated to UDP are occurring much more often than those associated with TCP or ICMP. It is asserted that most of these are port scans happening in the communication with IRC servers. These anomalies should definitely be signalled in a different way than other, more "alarming" anomalies.
- Some attempt should be made to correlate anomalies happening at the same time and for the same source for the UDP and TCP (and ICMP) view, or happening at the same time for the "packets" "flows" and "traffic" view. That is, looking at twelve subviews ((TCP, UDP, ICMP, any)*(packets, bytes, traffic)) is unpractical. There should be a way to look at fewer graphs grouping all the anomalies happening in more than one subviews, leaving in the single subviews anomalies happening only for that specific subview.

RedIRIS noted that fine-tuning HW analysis is not a trivial task. They find it a useful tool for detecting (specific) anomalies, but feel that more information about the algorithm underneath would be necessary. With respect to the performance, RedIRIS suggests that real-time investigations in the nfsen-hw web interface would be more useful and faster (rather than in the same NfSen instance, as is being done now). Recommendations about the different HW algorithm parameters and confidential interval (different from the ones set by default) would be useful, in order to find (sub)optimal parameter settings.

2.3 DDoSVax

2.3.1 Short description

DDoSVax is a set of methods conceived by the Faculty of Electrical Engineering of the Eidgenössische Technische Hochschule Zuerich in order to quickly detect (Distributed) Denial of Service traffic. DDoSVax is short for "In search of a vaccine (Vax) against Distributed Denial of Service (DDoS) attacks", which was the initial motivation for starting the project. In the context of JRA2, DDoS VaX is represented by the project partner SWITCH.

DDoSVaX research is still ongoing at ETH Zürich, but most of the work under consideration for our purposes was carried out in 2003-2005. It is important to note that it was based on a different threat model. While our anomaly detection tries to identify potentially subtle deviations from the regular, DDoSVax intends to develop capabilities to detect huge work outbreaks in an as early stage as possible. Common to both goals is that detection needs to take place in semi-realtime during a phase when obvious signals for the presence of

anomalies are still lacking. It was therefore the goal of this pilot to identify the most promising methods of DDoSVax for the purpose of identifying anomalies.

2.3.2 Reported results

The most promising candidate identified in our review was the method described in “Host Behaviour Based Early Detection of Worm Outbreaks in Internet Backbones”:

<http://www.tik.ee.ethz.ch/~ddosvax/publications/papers/WETICE-STCA-duebendorfer-host-behaviour.pdf>

The assumption is that traffic patterns of systems on the Internet often exhibit a typical behaviour, which can be classified. Systems changing their class indicate anomalies. The method fully relies on information available in netflow records, which makes it in principle compatible with our Toolset. The reference implementation by the authors of this method was analysed and an architectural plan for re-implementation in the framework of NfSen was established.

Another promising candidate from the DDoSVax framework was the method described in “Entropy Based Worm and Anomaly Detection in Fast IP Networks”:

http://www.tik.ee.ethz.ch/~ddosvax/publications/papers/wetice05_entropy.pdf

By measuring the entropy of information available in netflow -records, this method is capable of detecting the presence of certain types of anomalies, but has only limited analytical capabilities. We doubt the effectiveness of this method to detect other than wide-spread infections and recommend against further studies in the context of our pilot.

2.3.3 Current status and future

The DDoSVax trial identified the DDoSVax host behaviour method as the most promising method from the DDoSVax research to identify anomalies. It was confirmed at the JRA2 meeting in Prague in May 2007 that JRA2 should develop a plugin for NfSen based on the DDoSVax host behaviour method.

3 Conclusions and Recommendations

Based on the results so far, the Toolset appears to be in a promising stage of development. However, both extending the tests of already adopted methods for advanced anomaly detection (as described in Chapter 2), as well as other possible methods (to be described in more detail in this section) are desirable. The Toolset will only maximise its value, both for seasoned teams already familiar with the concepts of network flow monitoring and analysis, as well as for new teams looking for easily accessible ways to monitor their networks.

The overall response has potential for improvement. Further development of the Toolset and its quality (and therefore its usefulness) depends on a serious level of involvement from the project partners. Methods conceived by a single team or NREN should ideally be adopted and deployed (in an experimental, or preferably semi-operational way) by at least a couple of partner NRENS to get a reliable estimate as to the validity and usefulness of that proposed method.

3.1 Proposed additional methods

In light of the aforementioned observation, two other methods, not yet tested nor discussed in Chapter 2 have already been proposed within the JRA2 community:

- NfSen-overflow.
- NetReflex.
- DDoSVax host behaviour method.

Those methods will be discussed briefly in this section.

3.1.1 NfSen-overflow

3.1.1.1 Short description

NfSen-overflow is a plug-in developed by SURFnet (Werner Schram) that notifies the user about possibly malicious network traffic. It can be used to spot the occurrence of botnet control hosts (and the traffic generated to or from them) in the network being monitored.

3.1.1.2 Current status

The plugin has been added to the open source repository. It is currently being deployed by SURFcert to detect botnet controllers residing with any of the connected networks.

3.1.1.3 Goals of the test

The goal of the test is to assess its feasibility for detecting (active) botnet control hosts within the monitored network. Work has been carried out to automatically alert the responsible incident response team by means of an IODEF formatted XML message indicating the alleged infected host (by its IP address). Further testing by NRENs that run a base installation of NfSen of their own should give a more accurate indication of whether this plugin is a useful addition to the Toolset for the purpose already mentioned: to quickly take down (parts of) zombie networks.

3.1.2 NetReflex

3.1.2.1 Short description

NetReflex (manufactured by Guavus) is an anomaly detection tool presented to the JRA2 community on two occasions (Cambridge, January 2006, and Cambridge, January 2007) by GUAVUS (Anukool Lakhina). Some NRENs have already expressed interest in joining DANTE in beta testing it. This work fits well in JRA2 plans for Y4.

3.1.2.2 Current status

NetReflex is in a pre-beta stage. DANTE recently acquired access to an instance of the software running on a server hosted by Tompson Paris (one of the Guavus' sponsors), but fed with DANTE's data (there is an NDA in place between DANTE and Tompson Paris). So far, DANTE has roughly verified that some of the traffic matrixes produced by NetReflex are actually correct (comparing it with DANTE's NfSen installation). The verification of the anomaly detection capability of NetReflex has still to be done. Since a major release is expected by end of August, this testing will not start before then.

3.1.2.3 Goals of the test

The main goal of the test is to verify the automatic anomaly detection and classification of NetReflex, which is one of its promising features. Basically it has to be seen if the promising research work done by A. Lakhina in the past years, and published at SIGCOMM [SIGCOMM], can be successfully turned into something useful by NRENs' CERTs.

3.1.2.4 Test draft plan

A staged approach is proposed: a first stage should involve a very limited number of partners with the goal to identify the potential benefit to the NREN's community of running such a tool. The main reason for restricting the first testing to a few partners is that NetReflex still requires some manual configuration for the data feed, and assistance by Guavus engineers may be needed. Guavus cannot commit to doing that with too many partners. The first stage should be completed by Jan 2008.

If the potential of the tool is assessed positively, we can enter stage 2 and go for a more in depth analysis and look into longer-term collaboration issues, and also cover financial and licensing matters (NetReflex will be a commercial tool). It is expected, however, that the whole NREN community will get a preferential commercial treatment in return for the testing. A Guavus draft commitment for that is expected to come before starting the first stage.

3.1.3 DDoSVax host behaviour method

3.1.3.1 Short description

A description of this method was given in 2.3.1 above.

3.1.3.2 Current status

The DDoSVax trial identified the host behaviour method as the most promising method from the DDoSVax research to identify anomalies. An architecture plan for its implementation in the framework of the Toolset exists. It was confirmed at the JRA2 meeting in Prague in May 2007 that JRA2 should develop a plugin for NfSen based this method

3.1.3.3 Goals of the test

The goal of this test is to assess the capability of this method to identify anomalies in an semi-operational environment with live data. Attention needs be given to the accuracy of its results (rate of false positives and negatives), the effort needed to interpret its results and its scalability on networks with large numbers of flows.

3.2 Other possible methods and/or tools

It should be noted that JRA2 is still open and eager to include any other method or tool or solution that aims to use NetFlow data as a basis for advanced anomaly detection. The objective of the group's work is, after all, to deliver a viable Toolset by the end of the fourth project year, with a maximum of usefulness and feasibility to achieve one of the basic objectives of this JRA: to drastically enhance the security of the GÉANT2 network and its connected NREN backbones. This can be accomplished by incorporating as many (proven) methods for reliable anomaly detection as possible and working closely with other groups in the same area to incorporate or investigate new developments and techniques.

Another observation within the group was that the deployment of essentially experimental methods for anomaly detection is crucial for their potential success rate. Using advanced technologies to detect aberrant network patterns, and therefore detecting potentially malicious traffic in an early stage, will be a great contribution to the improvement of the overall security of the GÉANT2 network. The synergy of the JRA2 community lies partly in the fact that this community is a mesh of interconnected networks, fundamentally using the same sorts of technologies, but within various contexts and under the governance of different policies, rules, appreciations, etc. Sharing a common technological base makes it relatively easy to discuss the potential usefulness of new service ideas. But the rather different operational environments make it rather difficult to build widely acceptable prototype services. The factor "deployability" needs more attention in future tests to suit a sufficiently large number of partners for sufficiently large-scale operational tests of new service prototypes.

4 References

- [BICHOS]** <https://forja.rediris.es/projects/bichos/>
- [DDoSVax]** <http://www.tik.ee.ethz.ch/~ddosvax/>
- [GUAVUS]** <http://www.guavus.com/>
- [NetReflex]** <http://www.guavus.com/technology.htm>
- [Nfsen-overflow]** <http://sourceforge.net/projects/nfsen-overflow/>
- [SIGCOMM]** Mining Anomalies Using Traffic Feature Distributions. A. Lakhina, M. Crovella and C. Diot. ACM SIGCOMM, Philadelphia, PA, August 2005, <http://www.sigcomm.org/>

5 Acronyms

CERT	Computer Emergency Response Team
HW	Holts-Winter forecast model
ICMP	Internet Control Message Protocol
IODEF	Incident Object Definition and Exchange Format
TCP	Transmission Control Protocol
UDP	User Datagram Protocol