Quality Control in Function of Statistical Anomaly Detection in Intrusion Detection Systems

Petar Čisar

Telekom Srbija, Subotica, Serbia, petarc@telekom.yu

Sanja Maravić Čisar

Polytechnical Engineering College, Subotica, Serbia, sanjam@vts.su.ac.yu

Abstract: This paper deals with statistical anomaly detection as one of the essential parts of intrusion detection systems. Analyzing and introduction with computer network traffic behavior, threshold adjustment, optimizing its value in function of real traffic intensity and false positive alarm reduction is the key aim of statistical anomaly detection. There are several approaches to this area with different specific tools. One of them is mathematical statistics. The authors of this paper tried to work out a way of defining the most appropriate relatively fixed threshold using different statistical methods, applied to a real computer network. By comparing and discussing separate independent methods it is possible to converge to the optimal solution for thresholds.

Keywords: IDS, threshold, false alarms, mean, standard deviation

1 Introduction

An Intrusion Detection System (IDS) generally detects unwanted manipulations to systems. The manipulations may take form of attacks by skilled malicious hackers or using automated tools. An IDS is required to detect all types of malicious network traffic and computer usage that can not be detected by a conventional firewall. Even the best packet - filtering can miss quite a lot of intrusions.

The performance of a network IDS can be more effective if it includes not only signature matching but also traffic analysis. By using traffic analysis, anomalous traffic is identified as a potential intrusion. Traffic analysis does not deal with the payload of a message, but its other characteristics such as source, destination, routing, lenght of the message, time it was sent, the frequency of the communication etc. Traffic payload is not always available for analysis – the

traffic may be encrypted or it may simply be against policy to analyze packet payload.

An IDS may be categorized by its detection mechanism: anomaly - based, signature – based or hybrid (uses both of previous technologies).

Anomaly Detection Techniques:

- Protocol Anomaly Detection – refers to all expectations related to protocol format and protocol behaviour. Some examples: illegal field values and combinations, illegal command usage, running a protocol or service for a non – standard purpose or on a non – standard port.
- Application Payload Anomaly – Application anomaly must be supported by detailed analysis of application protocols. Application anomaly also requires understanding of the apllication semantics in order to be effective. Example: the presence of schellcode in unexpected fields.
- Statistical Anomaly – Statistical DDoS – To fully characterize the traffic behaviour in any network, various statistical measures are used to capture this behaviour. For example: there is a stable balance among different types of TCP packets in the absence of attacks. This balance can be learned and compared against short – term observations that will be affected by attack events. Additionally, the statistical algorithm must recognize the difference between the long – term (assumed normal) and the short – term observations to avoid generating false alarms on normal traffic variations.

2 Network Statistical Anomaly Detection (NSAD)

NSAD attempts to dynamically understand the network and statistically identify traffic that deviates from normal traffic usage and patterns.

NSAD systems can be broken down further into threshold, baseline and adaptive systems, with each looking for different triggers to identify anomalous behaviour.

Threshold NSAD Systems – These systems allow the administrator to configure thresholds to certain network usage parameters and report the passing of a configured threshold as a potential attack. For instance, threshold NSAD may allow the administrator to configure a threshold of 3000 request/minute to a Web server. Then, any time that the system measures more then 3000 requests in a minute it will be reported as an anomaly and a potential attack.

Baseline NSAD Systems – This systems detect and report statistical anomalies by establishing a baseline of some network usage pattern and then reporting deviations from that baseline as a potential intrusions. For example, baseline NSAD can look at total network traffic volume by hour and establish a range of

'normal' values for that parameter. For example, 'on Mondays between 9 am and 10 am the total traffic volume is expected to be between 80 and 120 megabytes'. Then, if the system detects more or less traffic in that hour, it is reported as an anomaly and a potential attack.

Adaptive NSAD Systems – Since usage patterns change over time, NSAD systems attempt to adapt to these changes continually. Adaptive systems accomplish this by using 'statistical usage profiling'. Basically, the system maintains two sets of usage data – a long – term usage profile and a short – term observed usage. To detect attacks, a modern NSAD system compares the short – term usage to the long – term profile and reports deviations that are considered 'statistically significant' as a potential attacks. The system further blends the short – term observed usage into the long – term usage profile to realize adaptation.

Advantages:

- it can detect attacks that would be missed by other detection mechanisms, with applications confined to specific types of traffic that can be easily quantified

Weaknesses:

- attack reporting is hard to interpret or turn into an action
- traffic in large organizations is constantly changing, making it virtually impossible to establish a baseline
- attacks can be contained within the baseline and an organization would never know
- attackers can train the system to see attack traffic as normal
- false alarms generate a management burden for an organization

2.1 Malicious NSAD Training

As adaptive NSAD systems continually update their long – term usage profile to adapt to changing network usage patterns, the systems open themselves up to a serious and detrimental attack, usually referred to as the 'NSAD training attack'. An attacker that knows that there is an NSAD system monitoring the network can influence the monitored usage pattern slowly enough to not be detected and in such a way that the attacker will eventually get the adaptive baseline to a point where it recognizes an attack as normal traffic.

For example, imagine an NSAD system that monitors network volume. Assume that the current baseline is 80 to 120 megabytes per – hour and that the attack wants to flood the network with 500 megabytes per – hour. The attack can start by maintaining a constant network volume of 110 megabytes per – hour. This may bring the beseline to the range of 100 to 140 megabytes, at which time the attacker will increase the volume to 130 megabytes per – hour, and so on. The attacker can

repeat this process until the system's baseline is in the vicinity of 500 megabytes per – hour. At this point the attacker can launch his attack and the system will never spot it.

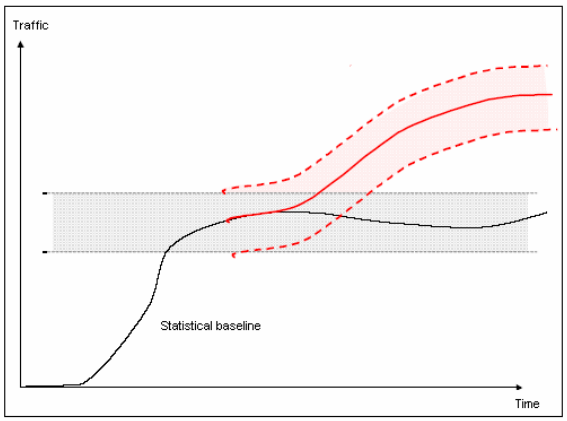


Figure 1
NSAD training

2.2 False Positives

False positives happen when an IDS falsely reports an attack. NSAD systems have the worst false positive rates among all intrusion detection mechanisms due to the way networks operate. Deviations from baseline usage patterns can happen both during an attack and as part of normal network operation.

For example, an NSAD system that monitors e-mail usage will establish a baseline in terms of the number of e-mail messages that a company receives in a given period, such as one day. The problem is that deviations from this baseline can happen at any time. Deviations are not just the result of a mail server being under a DoS (Denial of Service) attack. More likely, deviations could be the result of an e-mail – based marketing campaign, a holiday that prompts the exchange of greetings, a major news announcement that prompts a lot of requests to the sales department and many other scenarios.

The false positive rates of NSAD systems are so bad that industry experts consider systems that generate less than a 95% false positive rate (i.e. 19 false alarms for each real one) as outstanding [1].

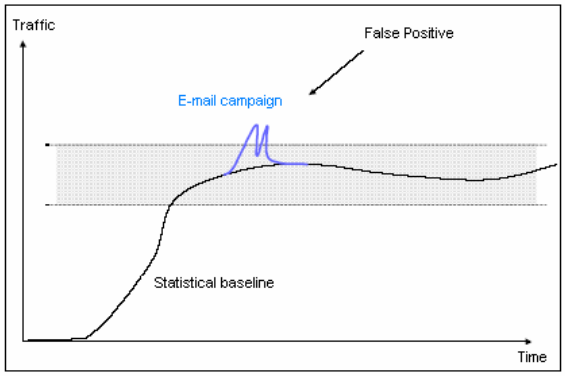


Figure 2
The false positive

3 Statistical Quality Control

Statistical quality control (SQC) is the term used to describe the set of statistical tools used by quality professionals. SQC can be divided into three broad categories:

- Descriptive statistics are used to describe quality characteristics and relationships. Included are statistics such as the mean, standard deviation, the range and a measure of the distribution of data.
- Statistical process control (SPC) involves inspecting a random sample of the output from a process and deciding whether the process is producing products with characteristics that fall within a predetermined range.
- Acceptance sampling is the process of randomly inspecting a sample of goods and deciding whether to accept the entire lot based on the results.

3.1 A Suggestion of a Method for Determining the Threshold Value

Maximal value of traffic is possible to determine, with accepted probability, by usage of descriptive statistics, method of confidence interval and 6σ concept. In order to achieve the highest possible accuracy in determining the maximal value of traffic, descriptive statistics will be applied on local maximums of the traffic curve at a definite time interval. Traffic samples are taken in moments when the traffic curve reaches the local maximums. Namely, if analysis is done with random samplings, the result would not include a sufficient number of

representative maximum points, so then the final result would not be precise enough. In that way, a set of local maximums is created on which further adequate methods will be apllied.

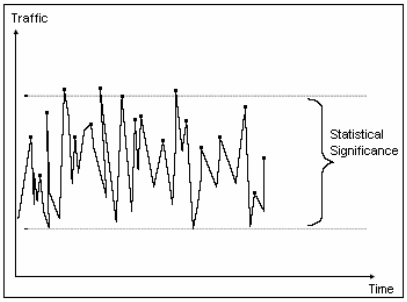


Figure 3
Local maximums

This approach to the problem is the more accurate the longer the time interval of observation is and the more local maximum points are included. In spite of that, it is necessary to pay attention to taking the largest values of all observed local maximums, because their influence on final result is the most significant. We can treat this sampling as a random sample. If the number of samples n is large enough (n>30), i.e. it is the case of a large sample, it can be considered that this sampling has normal distribution.

3.1.1 Method of Confidence Interval

Procedure of determining maximum value:

the mean $\bar{x}$ and standard deviation σ of all values are established from the set of n determined local maximums are:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \tag{2}$$

maximal value of all local maximums is obtained as confidence interval:

$$x_{\max} = \bar{x} + \sigma_{\bar{x}} . \text{ uz} \tag{3}$$

$$x_{\min} = \bar{x} - \sigma_{\bar{x}} . \text{ uz} \tag{4}$$

where $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ is standard error, while u_z is obtained from the table of normal distribution for accepted confidence level β (mostly 95% and 99%).

3.1.2 6σ Concept

This concept originated from the 'Motorola' company in 1987. It is implemented with goal of describing high level of quality, which a company tends to achieve.

This means that the upper specification limit (USL) and lower specification limit (LSL) are defined as:

$$UCL = x_{\max} = \bar{x} + 6\sigma \tag{5}$$

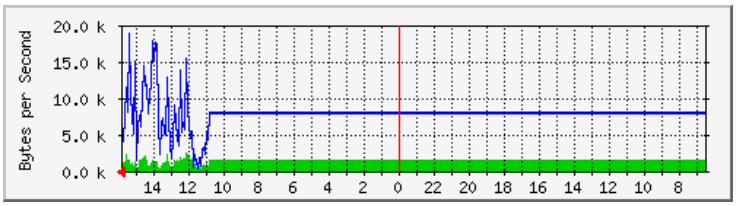
$$LCL = x_{\min} = \bar{x} - 6\sigma \tag{6}$$

In accordance with this concept, 99.999660% of all values are within the range (LCL, UCL) while 3,4 ppm is out of it.

3.2 Application of Suggested Methods

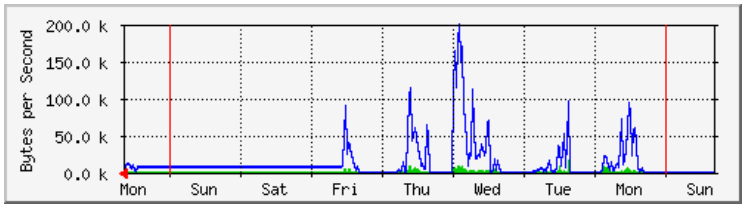
The methods above will be applied to authentical data obtained by traffic analizer in Politechnical Engineering College in Subotica, whose grafical review of traffic intensity for different period of observation is given by following diagrams. As the principle is identical, with aim of easier read-out, just 'outgoing traffic' will be observed and analyzed.

`Daily' Graph (5 Minute Average)



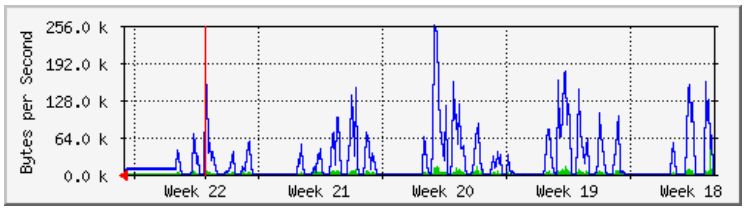
	Max	Average	Current
In	2576.0 B/s (0.2%)	1451.0 B/s (0.1%)	1056.0 B/s (0.1%)
Out	18.6 kB/s (1.5%)	7766.0 B/s (0.6%)	6708.0 B/s (0.5%)

`Weekly' Graph (30 Minute Average)



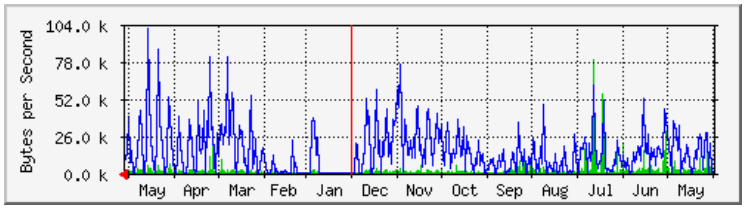
	Max	Average	Current
In	16.5 kB/s (1.3%)	1350.0 B/s (0.1%)	1167.0 B/s (0.1%)
Out	199.3 kB/s (15.9%)	12.7 kB/s (1.0%)	10.8 kB/s (0.9%)

Monthly' Graph (2 Hour Average)



	Max	Average	Current
In	62.1 kB/s (5.0%)	1913.0 B/s (0.2%)	1288.0 B/s (0.1%)
Out	254.3 kB/s (20.3%)	22.1 kB/s (1.8%)	7915.0 B/s (0.6%)

`Yearly' Graph (1 Day Average)



	Max	Average	Current
In	79.4 kB/s (6.4%)	2570.0 B/s (0.2%)	1497.0 B/s (0.1%)
Out	100.2 kB/s (8.0%)	15.6 kB/s (1.3%)	7839.0 B/s (0.6%)

GREEN ### Incoming Traffic in Bytes per Second

BLUE ### Outgoing Traffic in Bytes per Second

Data from this table is collected by direct readout from diagrams above.

Sample No.	Local max. x_i (kbytes/s)
1	90
2	115
3	60
4	65
5	205
6	120
7	70
8	100
9	75
10	95
11	60
12	66
13	150
14	62
15	60
16	70
17	100
18	130
19	140
20	72
21	64
22	254
23	125

Sample No.	Local max. x_i (kbytes/s)
24	170
25	126
26	105
27	90
28	70
29	160
30	180
31	150
32	64
33	115
34	115
35	60
36	150
37	155
38	100
39	85
40	82
41	82
42	78
43	60
44	60
45	60

On the basis of data from the table, the mean and standard deviation of samples of local maximums are:

$$n = 45 \tag{7}$$

$$\Sigma x_i = 4665 \tag{8}$$

$$\bar{x} = 103{,}6666667 \text{ kbytes/s} \tag{9}$$

$$\sigma = 44{,}66287863 \tag{10}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 6{,}6579489 \tag{11}$$

a) Method of confidence interval – For accepted level of confidence $\beta = 0{,}99999994 \rightarrow \Phi(z_\beta) = \beta / 2 = 0{,}49999997$, we obtain from table $u_z = 5$. Thus, thresholds of maximal and minimal traffic are:

x_{max} = 136,956 kbytes/s (12)

x_{min} = 70,377 kbytes/s (13)

b) 6σ concept – By using the calculated values $\bar{x}$ and σ, the values for maximal and minimal traffic are:

x_{max} = 371,644 kbytes/s (14)

x_{min} = 0 (because traffic can't be negative) (15)

3.3 Comparative Analysis and Discussion about Suggested Methods

The aspect of false alarms – The results obtained by these suggested methods considerably differ from each other. In order to decide which fit results best to the real situation shown on diagrams, it is necessary to compare the obtained results with the values from diagrams. It has to be stated that the diagrams represent the network status without attacks on it. Analyzing the result obtained by method of confidence interval, it can be concluded that the calculated maximal value is exceeded in 10 of 45 cases, which represents an initial error of 22.2%. If the threshold of an IDS is set on that value, then false alarms would be generated in the above-mentioned percentage within the observed interval. With identical analysis we come to conclusion that the maximal value of traffic received by the 6σ method is not overflowed, thus no false alarms would be generated – i.e. 0%. This means that the threshold set to 6σ-value gives less false alarms, but tolerates greater possible malicious traffic. But, lower threshold in the same time means that IDS will react to a real alarm earlier, so in that sense the method of confidence interval is more secure.

On the basis of statements made above, the authors of this paper are of the opinion that as optimal value, from the aspect of sensitivity to real alarm and the number of false alarms, the mean of maximal values obtained from described two methods could be accepted. In the actual case, that value would be (136,956 kBytes + 371,644 kBytes/s) / 2 = 254,3 kBytes/s, which excellently fits in real situation shown on diagrams.

The aspect of attack - It should be noticed from diagrams that there are large intervals without traffic (during weekend when college doesn't work and by night as well). These empty intervals have great influence on average and give an unreal low value in longer periods. That's why the daily graph seems to be referential. We can see on it that the average is approximately equal to a half of the maximal value.

In [7] is given a classification according to intensity of an attack into:

intense attack – attack amplitude ~ 250% mean

small attack - attack amplitude ~ 10% mean

Generally,

$$\text{attack} = A_0 + \alpha_1 \cdot T_0 \cdot f_a(t) \tag{16}$$

$$T_0 = \alpha_2 \cdot T_{\max} \tag{17}$$

where fa(t) is a random function $0 \le fa(t) \le 1$, T_{max} is the calculated maximal value of traffic, T_0 is the mean of the traffic and A_0 is the mean of the attack for the observed interval. The elementary requirement for an IDS with fixed threshold for alarm generating is:

$$\text{traffic} + \text{attack} > \text{Tmax}, \tag{18}$$

which using (16) and (17) gives the following general condition:

$$\text{traffic} > T_{\max} - A_0 - \alpha_1 \cdot T_0 \cdot f_a(t) \tag{19}$$

Because of $T_{max} >> A_0$, (18) becomes:

$$\text{traffic} > (1 - \alpha_1 \cdot \alpha_2 \cdot f_a(t)) \cdot T_{\max} \tag{20}$$

For the network traffic in Politechnical Engineering College, the following approximative value can be accepted (based on adequate graphs): $\alpha_2 \approx 1/2$. For example, in the case of small attack (30% of the mean), $\alpha_1 = 1/3$. Then from (20) we get:

$$\text{traffic} > (1 - 1/6 \cdot f_a(t)) \cdot T_{\max} \tag{21}$$

3.4 The Aspect of 'Normal' Behaviour of Network Traffic

One of the possible ways for defining 'normal' behaviour of a network traffic is, in case of none attack, to create several levels (in kbytes/s). For example: 60, 100, 140, 180 and 220. After this, the next step is to follow the traffic curve and count the number of cutting points with these levels. This gives the referent number of cuttings related to some level: N60, N100, N140, N180 and N220. At the end of the observation period we get the final number of cutting points Nc. Thus, N60/Nc, N100/Nc, N140/Nc, N180/Nc and N220/Nc represent referential percentages for a concrete level. If during the traffic investigation some of the percentage becomes greater than the referential (for e.g. 1%), that moment indicates the appearance of attack. It is especially important for the highest levels.

This method provides successful detection of small attacks as well. In that case we should establish appropriate lower levels (e.g. N20 and N40) and define their 'normal' percentages. Also, due to easy setting the triggers, this method is suitable for appropriate network monitoring and alerting softwares.

Concluding Considerations

A great number of false positive alarms is one of the key problems in systems with statistical anomaly detection. Determining the thresholds is a vital problem. If the threshold is set in an adequate way, it is possible to significantly decrease the number of false alarms. At the same time, a threshold with too high value provides a space for latent intrusion into the system. Methods suggested in this paper offer a statistical tools for more precise definition of alarm events. As the traffic in a computer network is random process, alarm events calculated by these methods, with accepted level of confidence, are related to a given observation interval.

With the aim of getting more accurate result, having in mind the natural growth of traffic (because of incresing number of network users, more complex software which they use etc.), this procedures have to be repeated periodically. If applied to a greater number of obtained maximal values of thresholds, it will provide a sufficiently real and precise value, offering a satisfactory level of statistical anomaly detection and by that the protection of network itself.

References

[1] Sorensen Sarah: Competitive Overview of Statistical Anomaly Detection, White Paper, Juniper Networks, 2004

[2] Gong Fengmin: Deciphering Detection Techniques: Part II Anomaly – Based Intrusion Detection, White Paper, McAfee Security, 2003

[3] SANS Intrusion Detection FAQ: Can you explain traffic analysis and anomaly detection?, http://www.sans.org/resources/idfaq/anomaly_detection.php

[4] Intrusion - detection system – Wikipedia, http://en.wikipedia.org/wiki/Intrusion-detection_system

[5] Montgomery Douglas: Introduction to Statistical Quality Control, 5th Edition, John Wiley & Sons, 2005

[6] Statistical Quality Control, www.wiley.com/college/reid/0471347248/samplechapter/ch06.pdf

[7] Siris A. Vasilios: Denial of Service and Anomaly Detection, SCAMPI BoF, Zagreb, 2002