Optimization Methods of EWMA Statistics

Petar Čisar

Telekom Srbija, Subotica, petar.cisar@gmail.com

Sanja Maravić Čisar

Subotica Tech, Subotica, sanjam@vts.su.ac.rs

Abstract: Many intrusions which attempt to compromise the security of computer and network systems manifest themselves in changes in the intensity of events. Because of the ability of exponentially weighted moving average (EWMA) statistics to monitor the rate of occurrences of events based on their intensity, this technique is appropriate for implementation in control limits based algorithms. The research has shown that the usual application of this algorithm on computer network traffic, as applied in industrial processes, does not provide acceptable results. The paper also gives a review of possible optimization methods.

Keywords: intrusion detection; EWMA; control limits; optimization; autocorrelation; ARL

1 Introduction

The exponentially weighted moving average is a statistic for monitoring the process that averages the data in a way that gives less and less weight to data as they are further removed in time. For the EWMA control technique, the decision regarding the state of control of the process depends on the EWMA statistic, which is an exponentially weighted average of all prior data, including the most recent measurements.

By the choice of weighting factor λ, the EWMA control procedure can be made sensitive to a small or gradual drift in the process.

The statistic that is calculated is the following:

$$EWMA_t = \lambda\, Y_t + (1-\lambda)\, EWMA_{t-1} \quad \text{for } t = 1, 2, ..., n \tag{1}$$

where

- $EWMA_0$ is the mean of historical data

- Y_t is the observation at time t
- n is the number of observations to be monitored including $EWMA_0$
- $0 < \lambda \leq 1$ is a constant that determines the depth of memory.

This equation was established by Roberts as described in [4].

The parameter λ determines the rate at which "older" data enter into the calculation of the EWMA statistic. A value of λ = 1 implies that only the most recent measurement influences the EWMA. Thus, a large value of λ = 1 gives more weight to recent data and less weight to older data; a small value of λ gives more weight to older data. The value of λ is usually set between 0.2 and 0.3 [2], although this choice is somewhat arbitrary. Lucas and Saccucci [3] have shown that although the smoothing factor λ used in an EWMA chart is usually recommended to be in the interval between 0.05 to 0.25, in practice the optimally designed smoothing factor depends not only on the given size of the mean shift δ, but also on a given in-control Average Run Length (ARL). ARL represents the average number of determined process points before the first point indices the appearance of out-of-control state (exceeding one of the control limits).

The estimated variance of the EWMA statistic is approximately:

$$\sigma^2_{EWMA} = (\lambda / (2 - \lambda))\ \sigma^2 \quad (2)$$

where σ is the standard deviation calculated from the historical data.

The center line for the control chart is the target value or $EWMA_0$. The upper and lower control limits are:

$$UCL = EWMA_0 + k\sigma_{EWMA} \quad (3)$$

$$LCL = EWMA_0 - k\sigma_{EWMA} \quad (4)$$

where the factor k is either set equal to 3 (the 3-sigma control limits) or chosen using the Lucas and Saccucci tables (ARL = 370).

In addition to the aforementioned authors, the publications [6] - [13] and [19] have also dealt with the topic of EWMA statistics and statistical anomaly detection in computer networks.

Control charts are specialized time series plots which assist in determining whether a process is in statistical control. Some of the most widely used forms of control charts are X-R charts and Individuals charts. These are frequently referred to as "Shewhart" charts after the control charting pioneer, Walter Shewhart, who introduced such techniques. These charts are sensitive to detecting relatively large shifts in the process (i.e. of the order of 1.5σ or above). In computer network practice, shifts can be caused by intrusion or attack, for example. Two types of charts are usually used to detect smaller shifts (less than 1.5σ), namely cumulative sum (or CUSUM) charts and EWMA charts. A CUSUM chart plots the cumulative sums of the deviations of each sample value from a target value. An

alternative technique to detect small shifts is to use the EWMA methodology. This type of chart has some very attractive properties, in particular:

1 Unlike X-R and Individuals charts, all of the data collected over time may be used to determine the control status of a process.

2 Like the CUSUM, the EWMA utilizes all previous observations, but the weight attached to data exponentially decreases as the observations become older and older.

3 The EWMA is often superior to the CUSUM charting technique due to the fact that it detects larger shifts better.

4 EWMA schemes may be applied for monitoring standard deviations in addition to the process mean.

5 EWMA schemes can be used to forecast values of a process mean.

6 The EWMA methodology is not sensitive to normality assumptions.

In real situations, the exact value of the shift size is often unknown and can only be reasonably assumed to vary within a certain range. Such a range of shifts deteriorates the performance of existing control charts. One of the algorithms for determining the maximal shift in normal traffic is described in [16].

The paper describes the process of the application of the EWMA algorithm for one major user, given as an example. It can be shown that the obtained results are valid for the other analyzed users as well. This research uses samples of authentic network traffic (i.e. traffic intensity in a unit of time). Traffic analysis is realized in the form of statistical calculations on samples which derive from the traffic curve. From the appropriate pattern of Internet traffic, 35 samples of local maximums are taken in order to ensure that the statistical analysis is performed on a large sample (number of samples $n > 30$), thus supporting and leading to general conclusions.

The aim of this research is to determine those allowed EWMA values of traffic, so that when they are exceeded, it will be considered as the appearance of a statistical anomaly suspected to attack. In this sense, the choice of only local maximums for analysis can be accepted as logical, because the critical point of maximum value of aggregate traffic is in this way also included.

The proposed method of calculating the overall optimal value Λ is applied to traffic patterns, on the basis of which the lower and upper control limits of traffic are determined. For statistical detection of an attack, the primary interest is the appearance of a situation in which the upper control limit is exceeded. The overstepping of the lower control limit can be understood as a statistical anomaly, but in the case of this research, it is only related to the local maximum (and not to the aggregate network traffic) and as such does not endanger the security of the computer network in general. Therefore, the situation in which the value of

network traffic falls below some lower limit is not considered to be a suspicious event or attack, because the initial presumption of this research is the increase of traffic during an external attack. For the observed pattern of traffic, EWMA values are calculated and if these values are outside of the control limits, that situation is interpreted as a statistical anomaly. Emphasis in this work is placed on determining the occurrence of false alarms, as an important security feature of the applied algorithm.

2 Optimized Exponential Smoothing

Calculating the optimal value of parameter λ is based on the study of authentic samples of network traffic. Random variations of network traffic are normal phenomena in the observed sample. In order to decrease or eliminate the influence of individual random variations of network traffic on occurrence of false alarms, the procedure of exponential smoothing is applied, as an aspect of data preprocessing.

For any time period *t*, the smoothed value S_t is determined by computing:

$$S_t = \lambda\, y_{t-1} + (1 - \lambda)\, S_{t-1} \quad \text{where } 0 < \lambda \leq 1 \text{ and } t \geq 3 \tag{5}$$

This is the basic equation of exponential smoothing. The formulation here is given by Hunter [2]. It should be noted that there is an alternative approach, in which, according to Roberts [4], y_t is used instead of y_{t-1}.

This smoothing scheme starts by setting S_2 to y_1 (there is no S_1), where S_i stands for smoothed observation or EWMA, and y_i stands for the original observation. The subscripts refer to the time periods 1, 2, ..., *n*. For example, the third period is $S_3 = \lambda\, y_2 + (1 - \lambda)\, S_2$ and so on.

There is no generally accepted statistical procedure for choosing λ. In that situation, the method of least squares might be adequate to determine the optimal value of λ for which the sum of the squared errors (SSE) $(S_{n-1} - y_{n-1})^2$ is minimized.

The method of least squares represents a standard approach to the approximate solution of over-determined systems (i.e. sets of equations in which there are more equations than unknowns). The most important application is in data fitting. The best fit in the least squares sense minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by a model.

Here is an illustration of this principle through an example. Consider the following data set consisting of *n* observations of data flow over time – for starting $\lambda = 0.1$:

Table 1
Smoothing scheme

Time	Flow (y_t)	S_t	Error ($S_t - y_t$)	Error squared
1	y_1			
2	y_2	y_1	E_2	E_{22}
3	y_3	S_3	E_3	E_{32}
...	...	...	...	...
n	y_n	S_n	E_n	E_{n2}
				SSE_n

The sum of the squared errors (SSE) is then $SSE_{0.1}$. After that, the SSE is calculated for $\lambda = 0.2$. If $SSE_{0.2} < SSE_{0.1}$ then $SSE_{0.2}$ is better value for λ. This iterative procedure is related to the range of λ between 0.1 and 0.9. In this way, the best initial choice for λ is determined and then, for getting more precise value, search optionally continues between $\lambda-\Delta\lambda$ and $\lambda+\Delta\lambda$, where $\Delta\lambda$ is an arbitrarily small interval around λ (for instance, in practical applications, ± 10% around optimal λ).

Table 2
Comparison of smoothing schemes

time	y_t	EWMA	S_t
1	52,00	50,60	
2	47,00	49,52	52,00
3	53,00	50,56	50,50
4	49,30	50,18	51,25
5	50,10	50,16	50,67
6	47,00	49,21	50,50
7	51,00	49,75	49,45
8	50,10	49,85	49,91
9	51,20	50,26	49,97
10	50,50	50,33	50,34
11	49,60	50,11	50,39
12	47,60	49,36	50,15
13	49,90	49,52	49,39
14	51,30	50,05	49,54
15	47,80	49,38	50,07
16	51,20	49,92	49,39
17	52,60	50,73	49,93
18	52,40	51,23	50,73
19	53,60	51,94	51,23
20	52,10	51,99	51,94
			51,99

Comparative analysis of two different approaches (Roberts and Hunter) can be shown using the example of a process (y_t), with adopted values $EWMA_0 = 50$ and $\lambda = 0.3$. EWMA values in the table below correspond to Roberts's and S_t to Hunter's equation.

In Table 2 the fields with approximately equal values are marked with a lighter color, while fields with equal values are marked with a darker color. From this analysis it can be concluded that after a certain number of samples (in this case about the 16th sample) both schemes give the same smoothed values.

The behavior of both smoothing schemes will be examined also with SSE values. After calculating SSE for different λ, results were as follows:

Table 3
Comparison of values for SSE according to Roberts and Hunter

λ	SSE (Roberts)	SSE (Hunter)
0.1	62.81	75.01
0.2	49.95	55.86
0.3	39.28	42.16
0.4	30.25	31.62
0.5	22.40	23.01
0.6	15.50	15.71
0.7	9.55	9.57
0.8	4.70	4.66
0.9	1.31	1.29

Analysis of the obtained results has shown that approximately similar values were obtained, with greater coincidence at higher values of smoothing factor.

The initial EWMA plays an important role in computing all the subsequent EWMA's. There are several approaches to define this value:

1) Setting S_2 to y_1
2) Setting S_2 to the target of the process
3) Setting S_2 to average of the first four or five observations

It can also be shown that the smaller the value of λ, the more important is the selection of the initial EWMA. The user would be well-advised to try several methods before finalizing the settings.

For different input values of initial parameter S_2, an application in "Matlab" is created which calculates and plots the dependence of SSE and partial value of λ in range of $0 \div 1$, with adjustable step. In addition, the optimal value λ_{opt} is also calculated. For the optimal value, in accordance with the smoothing scheme, that particular value is taken for which the SSE is minimal. The following figure shows an example for calculating the optimal value of the parameter λ for a specific S_2.

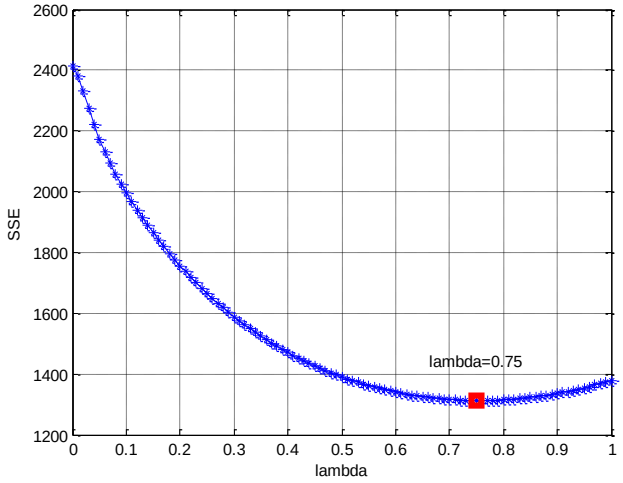

Figure 1
Calculation of λ_{opt}(SSE)

Due to the lack of an exact method of calculation in available publications about the determination of the initial S_2 in the procedure of exponential smoothing, the authors of this paper have dealt with researching the link between selection of $S_2 = y_1$ and λ_{opt}, i.e. $S_2(\lambda_{opt})$. In that sense, the range of S_2 is determined using the lowest to the highest sample value during the period of observation. This research was conducted on an authentic sample of network traffic of an Internet service provider and the segment of observation was the range of values of local maximums (in this concrete case $S_2 = 8 \div 34$ Mb/s), with a large enough number of values, taking into account the generality of conclusions. The period of observation was one month. The next table shows the numerical and graphical dependence $S_2(\lambda_{opt})$.

Since a set of different results has been obtained for partial values of λ_{opt}, in order to determine the overall optimal parameter Λ_{opt}, the measures of central tendency will be examined. There are three most common measures of central tendency:

- Average - the arithmetic mean, calculated by adding a group of numbers and then dividing by the count of those numbers.
- Median - the middle number of a group of numbers; that is, half the numbers have values that are greater than the median, and half the numbers have values that are less than the median.
- Mode - the most frequently occurring number in a group of numbers.

For a symmetrical distribution of a group of numbers, these three measures of central tendency are all the same. For a skewed distribution of a group of numbers, they can be different.

Table 4
Calculation of $S_2(\lambda_{opt})$

	S_2	λ_{opt}
1	8	0,72
2	9	0,72
3	10	0,72
4	11	0,72
5	12	0,71
6	13	0,71
7	14	0,72
8	15	0,72
9	16	0,72
10	17	0,72
11	18	0,72
12	18,5	0,73
13	19	0,73
14	19,5	0,73
15	20	0,73
16	20,5	0,73
17	21	0,74
18	21,5	0,74
19	22	0,74
20	22,5	0,75
21	23	0,75
22	23,5	0,75
23	24	0,75
24	25	0,76
25	26	0,77
26	27	0,77
27	28	0,78
28	29	0,79
29	30	0,8
30	31	0,8
31	32	0,81
32	33	0,82
33	34	0,82

In this particular case the following values are calculated (according to the previous table): average = 0.7482, median = 0.74 and mode = 0.72. Since the values for average and median do not differ significantly, the authors suggest for the overall optimal parameter Λ_{opt} to accept the average of all the partial results. In this particular case it is approximately 0.75, which significantly differs from the usually suggested values (between 0.2 and 0.3).

3 ARL Curves

Using a graphical method, the EWMA chart can be designed to have minimal ARL for the out-of-control situation, for the known shift of the mean δ and given ARL for the in-control situation. This chart has two parameters - λ and k (derives from the definition of control limits).

The figures below show the dependence of λ and k of the mean shift δ, for ARL as parameter. Using appropriate curves, values k = 2.7878 and λ = 0.1417 were determined as the optimal choice for the earliest detection of shift δ = 1σ.

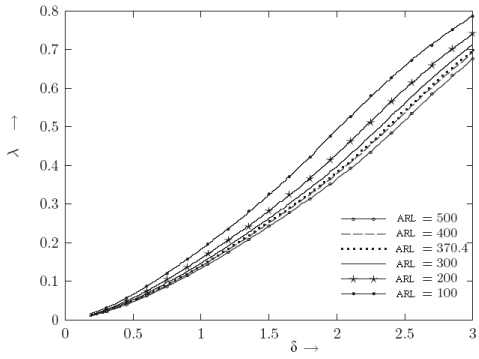

Figure 2

Optimal choice of *λ* in function of the mean shift [18]

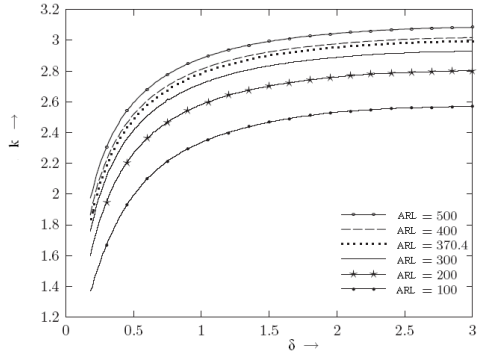

Figure 3

Optimal choice of *k* in function of the mean shift [18]

4 Autocorrelation

Autocorrelation or serial correlation of a time series means that the value of the observed variable in a time unit depends on values which appear prior to or later in the series. In practical situations, autocorrelation of the first order is usually examined, which may be shown by a simple correlation coefficient or so-called autocorrelation coefficient. Let R_t be the time series data, where t = 1, 2, ...,T, then the autocorrelation coefficient of the first order is given by:

$$\rho(R) = \frac{\sum_{t=2}^{T} R_t \cdot R_{t-1}}{\sqrt{\sum_{t=2}^{T} R_t^2 \cdot \sum_{t=2}^{T} R_{t-1}^2}}, \; -1 \leq \rho \leq 1 \tag{6}$$

One of the standard features of traffic time series is that the increasing rates of traffic R_t are not mutually significantly autocorrelated, i.e. the value of the autocorrelation coefficient is near zero. At the same time, this means that the distribution of positive and negative values of increasing rates is random and that does not follow a specific systematic regularity. Positive autocorrelation implies that the positive values are followed by mainly positive values and negative values by negative ones and then $\rho \approx +1$. In the case of negative autocorrelation, there is often a change of sign, i.e. the positive rate in most cases leads to a negative rate and vice versa and then $\rho \approx -1$. Since there is no typical scheme, on the basis of the positive rate in one particular time period there is no way of concluding (it cannot be concluded) with a significant probability that in the next period either a growth or decline will appear. The same applies to the situation for the negative rate.

Researchers in [5] dealt with the influence of autocorrelated and uncorrelated data on the behavior of an intrusion detection algorithm. In their work they came to conclusion that the EWMA algorithm for autocorrelated and uncorrelated data works well in the sense of intrusion detection in some information systems. The advantage of the EWMA technique for uncorrelated data is that this technique (as opposed to the case of the autocorrelated data) can detect not only rapid changes in the intensity of events, but also small changes in the mean value realized through the gradual increase or decrease of the intensity of events. However, in EWMA for uncorrelated data, the initial value of smoothed intensity events is to be reset after intrusion detection, in order to avoid the impact of current values of parameters on future results (*carry-over effect*). In the case of EWMA for autocorrelated data this reset is not necessary, because EWMA automatically adjusts the upper and lower control limits. Generally, the smoothing constant should not be too small, so that a short-term trend in the intensity of events in the recent past can be detected. Other publications have also shown the need for taking into account the autocorrelation of input data. As is emphasized in [17], in the case of dynamic systems the autocorrelation in variables is taking into account incorporating time lags of the time series during the modeling stage.

The samples of network traffic were obtained by the network software "MRTG" (Multi Router Traffic Grapher). This software generates three types of graphs:

- Daily – with the calculation of a 5-minute average
- Weekly – with the calculation of a 30-minute average
- Monthly – with the calculation of a 2-hour average

The graphs also enable numerical information on the maximum and average traffic for the appropriate period of time.

Daily, weekly and monthly graphs of the first measurement will be used for the calculation of the initial historical data, while the application of EWMA statistics, with the aim of checking the validity of certain parameters, will be realized on daily, weekly and monthly traffic graphs of the second measurement.

For the application of exponential smoothing method to the network traffic, it is necessary to first determine the historical values: $EWMA_0$ and standard deviation σ_0. For this purpose, it is necessary to collect appropriate traffic samples to perform adequate calculations. This study will use a total of 105 samples of local maximum: 35 samples from the daily traffic graph, 35 samples from the weekly traffic graph and 35 samples from the monthly traffic graph.

On the basis of the data presented in the given table the following can be calculated: $EWMA_0 = 23.10$ and $\sigma_0 = 4.87$.

In accordance with the method described above, and to justify the usage of EWMA statistics, it is important to determine the statistical independence of the samples, which will be examined by checking the existence of correlation between data. For this purpose, Pearson's correlation coefficient will be used, which is supplied as a ratio of covariances of two variables and the product of their standard deviations:

$$\rho_{xy} = Cov\,(X,Y) / (\sigma_x \cdot \sigma_y) \qquad -1 \leq \rho_{xy} \leq 1 \tag{7}$$

Other authors have proposed different interpretations of ways of correlation coefficient. Cohen [1] noted that all the criteria are based on a greater or lesser extent of arbitrariness and should not be kept too strictly. Yet, one often-used interpretation of these coefficients is given below, as described in [15]:

- ρ between 0 and 0.2 – no correlation or it is insignificant
- ρ between 0.2 and 0.4 – low correlation
- ρ between 0.4 and 0.6 – moderate correlation
- ρ between 0.6 and 0.8 – significant correlation
- ρ between 0.8 and 1 – high correlation

Table 5
Network samples

Time	y_t(daily)	y_t(weekly)	y_t(monthly)
1	12	21	23
2	10.5	22.5	30
3	8.5	23	27
4	10.5	20	27
5	18	20.5	25
6	22	23.5	27
7	25.5	24	22
8	20	21	24
9	33.9	23	23
10	25	25	20
11	24	25.5	24.5
12	26.5	24.5	26.5
13	27.5	22	28
14	23	25.5	27
15	25	27	23
16	24	28	22.5
17	23	27	26.5
18	23	28	31
19	22	25.5	22.5
20	23	30	22.5
21	23	29	27
22	23	26.5	25
23	23	29	26
24	16	26.5	28
25	16	27.5	21
26	9	26	24
27	11.5	25	22
28	8.5	24	22
29	8.5	23.5	22
30	14	22	23
31	23	22.5	27
32	23	24	29
33	20	24	25
34	23	25	25
35	23	23	22

The value of correlation coefficient ρ_{xy} can be calculated using the statistical function CORREL (array1, array2) in MS Excel. When examining the table above, it is possible to identify three series of data (daily, weekly and monthly) and in this sense three different correlation coefficients can be calculated:

- correlation coefficient for daily – weekly series: $\rho_1 = 0.28 \rightarrow$ low correlation
- correlation coefficient for daily – monthly series: $\rho_2 = 0.04$
- correlation coefficient for weekly – monthly series: $\rho_3 = -0.04$

Besides testing the correlation coefficient within a single measurement, it is important to check the existence of correlation between corresponding periods

from different measurements. For that purpose, values of correlation coefficient of two daily (weekly, monthly) intervals are checked and the following results are obtained:

- correlation coefficient for daily – daily series: $\rho_4 = -0.15$
- correlation coefficient for weekly – weekly series: $\rho_5 = 0.11$
- correlation coefficient for monthly – monthly series: $\rho_6 = -0.02$

As all calculated coefficients are with low degree of correlation, or without it, it can be concluded that the used data are statistically independent and that the application of EWMA statistics is justified.

5 Illustration of Results

The influence of the appropriate lambda value is illustrated by the following figure. Based on samples of authentic network traffic (daily traffic), a set of correspondent EWMA values are calculated. The curve in the first diagram was obtained for the optimized value of lambda (in this case, $\lambda = 0.91$), while the second diagram relates to the often proposed value of 0.25.

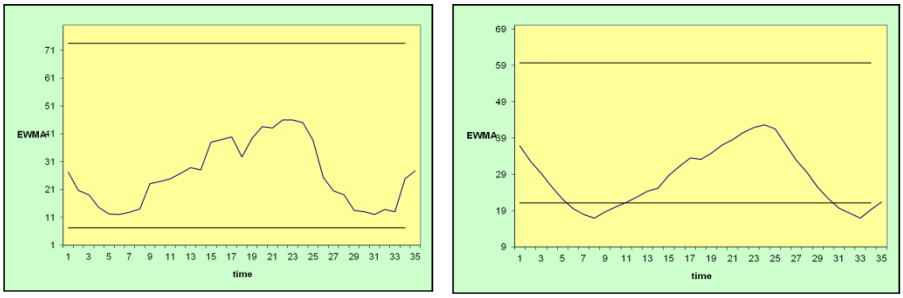

Figure 4
The effect of optimization

It is important to emphasize the multiple exceeding of the lower control limit in the second diagram, which represents a situation of statistical anomaly causing false alarms.

Conclusions

The aim of this research was to examine the possibility of applying EWMA statistics in intrusion detection in network traffic.

The research has shown that direct application of this algorithm on computer network traffic, as applied in industrial processes, does not provide acceptable results. Namely, often proposed values for exponential smoothing factor in case of

network application of the algorithm, may in some circumstances lead to the creation of false alarms, thus endangering the security level of the system. Due to the lack of an acceptable precise method for the determination of the initial value of the coefficient in exponential smoothing procedure in available publications, this research has been directed towards establishing a relation between the choice of initial ratio and optimal value for smoothing. By creating the appropriate application, the practical way was presented for testing the impact of different values of parameters on the level of anomaly detection. This enabled the establishment of graphical presentation of input depending on output sizes, which all contributed to the creation of the proposed method for calculating the optimal value of smoothing factor.

Before the start of the implementation of statistical analysis of traffic, the extent of autocorrelation between the used data must be examined, by calculating the correlation coefficients. One of the important results is that it is shown that analysis of network traffic properties based on individual patterns of daily traffic exclusively is not recommended, because of the increased level of autocorrelation. For this reason, when calculating the historical parameters, network traffic must be viewed in a wider context of time, taking into account the weekly and monthly periods. Using the network monitoring software, it is also necessary to determine the maximum variations of basic traffic characteristics (average and maximum).

To make this algorithm properly applicable in the network environment, it is necessary to perform previous processing of historical data, in order to obtain initial values of key parameters.

Based on the proof lent by the obtained results it can be concluded that the choice of EWMA parameters significantly affects the operation of this algorithm in a network environment. Therefore, the optimization process of parameters before the application of the algorithm is of particular importance.

References

[1] J. Cohen, Statistical Power Analysis for the Behavioral Sciences (2nd ed.), *Lawrence Erlbaum Associates*, Hillsdale, New Jersey, 1998

[2] J. S. Hunter, The Exponentially Weighted Moving Average, *Journal of Quality Technology* 18: 203-210, 1986

[3] J. M. Lucas, M. S. Saccucci, Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements, *Technometrics 32*, 1-29, 1990

[4] S. W. Roberts, Control Chart Tests Based on Geometric Moving Averages, *Technometrics*, 1959, Vol. 42, No. 1, Special 40th Anniversary Issue, pp. 97-101, 2000

[5] Ye et al., Computer Intrusion Detection through EWMA for Autocorrelated and Uncorrelated Data, *IEEE Transactions on Reliability*, Vol. 52, No. 1, 2003

[6] G. Fengmin, Deciphering Detection Techniques: Part II Anomaly-based Intrusion Detection, *White Paper, McAfee Security*, 2003

[7] S. Sorensen, Competitive Overview of Statistical Anomaly Detection, *White Paper, Juniper Networks*, 2004

[8] X. Wu, V. A. Mahadik, D. S. Reeves, A Summary of Detection of Denial-of-QoS Attacks on DiffServ Networks, DARPA Information Survivability Conference and Exposition, 2003, Proceedings, Vol. 2, pp. 277-282

[9] A. S. Neubauer, The EWMA Control Chart: Properties and Comparison with other Quality-Control Procedures by Computer Simulation, *Clinical Chemistry*, Vol. 43, pp. 594-601, 1997

[10] D. Seibold, Enterprise Campus Security–Addressing the Imploding Perimeter, http://www.itsa.ufl.edu/2003/presentations/IntSec.ppt

[11] S. Vasilios, F. Papagalou, Application of Anomaly Detection Algorithms for Detecting SYN Flooding Attacks, Global Telecommunications Conference, 2004, GLOBECOM 04 IEEE, Vol. 4, pp. 2050-2054

[12] J. Viinikka, H. Debar, Monitoring IDS Background Noise Using EWMA Control Charts and Alert Information, Recent Advances in Intrusion Detection, Lecture Notes in Computer Science, 2004, Volume 3224/2004, pp. 166-187

[13] Y. Zhao, F. Tsung, Z. Wang, Dual CUSUM Control Schemes for Detecting a Range of Mean Shifts, *IIE Transactions* 2005 (37), pp. 1047-1057

[14] Engineering Statistics Handbook–EWMA Control Charts, http://www.itl.nist.gov/div898/handbook/pmc/

[15] Savannah State University, Office of Institutional Research & Planning, http://irp.savstate.edu/irp/glossary/correlation.html

[16] P. Čisar, S. Maravić Čisar, A First Derivate-based Algorithm for Anomaly Detection, *International Journal of Computers, Communications & Control*, Volume III (2008), Supplementary Issue – Proceedings of ICCCC 2008, pp. 238-242

[17] J. Mina, C. Verde, Fault Detection for Large Scale Systems Using Dynamic Principal Components Analysis with Adaptation, *International Journal of Computers, Communications & Control*, Volume II (2007), No. 2, pp. 185-194

[18] S. B. Vardeman, J. M. Jobe, Statistical Quality Assurance Methods for Engineers, *John Wiley & Sons*, New York 1999

[19] P. Čisar, S. Maravić Čisar, Skewness and Kurtosis in Function of Selection of Network Traffic Distribution, *Acta Polytechnica Hungarica*, Vol. 7, No. 2, 2010, pp. 95-106