THE RESPONSE OF THE AUTOREGRESSIVE COEFFICIENT TO A DDOS ATTACK

Juraj SMIEŠKO – Ondrej ŠKVAREK

University of Žilina, Faculty of Management Science and Informatics, Slovak Republic

e-mail: {juraj.smiesko, ondrej.skvarek}@fri.uniza.sk

Abstract

Currently, in the field of cybersecurity, the topic of early DDoS attack recognition is most urgent. This article presents the results of evaluating the reaction of the autoregressive coefficient indicator to the onset of a DDoS attack. From the tested real scenarios, in addition to standard types, in which the coefficient curve was identical, we also selected non-standard attacks, in which the coefficient curve behaved atypically. Based on the findings, we propose detection functions suitable for further exploration of machine-based DDoS attack recognition.

Keywords: *DDoS attack, Autoregressive Coefficient, Detection Function, IP Traffic, prediction tunnel*

1 INTRODUCTION

In general, it can be assumed that the onset of a DDoS attack significantly alters not only the mean intensity of packet flows in standard IP traffic but also substantially affects other probabilistic characteristics of the traffic. For standard traffic comprising the aggregation of a large number of IP services, such as that observed in core networks, it is reasonable to assume a certain degree of stationarity, with a symmetric, unimodal distribution of packet arrivals (e.g., Bernoulli or MMRP models).

The initiation of a synthetically generated DDoS attack introduces significant disruptions to this probabilistic model. It is expected that the onset of a DDoS attack will be reflected significantly in the behaviour of sliding statistical coefficients, such as the coefficient of variation, skewness, kurtosis, entropy, Hurst exponent, autocorrelation and autoregressive coefficients, and others.

In this article, we focus on analysing the response of the autoregressive coefficient to the onset of various selected types of DDoS attacks.

We initially conducted experiments on simulated scenarios, in which we modelled normal and attack traffic using flows generated by 2-state On/Off processes, Poisson and Pareto distributions, as well as samples from the MNIST and CIFAR databases. Based

on these simulated experiments, we obtained a basic understanding of the response of the autoregressive coefficient to sudden changes in the traffic flow.

Subsequently, for testing, we used recordings of real DDoS attacks from datasets such as ISCX 2012 and CIC-IDS 2017, and primarily from our own custom dataset [1]. More detailed information about the attacks can be found in [2, 3, 4].

When using real attacks, the experimental results matched the simulated scenarios in roughly 70% of cases, while in the remaining cases, the coefficient often reacted in surprisingly non-standard ways. We aim to present these different situations in this paper using six selected DDoS attack recordings.

2 AUTOREGRESSION COEFFICIENT

An autoregressive model (AR) is a linear model used in the analysis and prediction of time series. Any observation in the time series depends only on the value of the previous observation and on the random component.

The autoregressive model $AR(n)$ assumes that the considered random process $\{Xt\}_{t\in T}$ has the structure

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_n X_{t-n} + \varepsilon_t, \tag{1}$$

while random variables ε_t represent model errors; we assume that they are uncorrelated with each other and have a normal probability distribution $\varepsilon_t \sim N(0,\sigma^2)$ (random variables ε_t constitute white noise) [5].

Let the values $(a_1, a_2, \ldots, a_N)$ (increments of flow) represent N realisations of the random process $\{Xt\}_{t\in T}$. If the $AR(n)$ model is valid, the realisations form the following structure

$$a_t = \beta_1 a_{t-1} + \beta_2 a_{t-2} + \cdots + \beta_n a_{t-n} + \varepsilon_t, \tag{2}$$

for $t = n+1,\ldots,N$. For all N measurements, we get a matrix notation (3). Suppose we denote the matrix of all N values of the dependent variable at as $\mathbf{y}$ and ε as the vector of measured errors. In that case, we can describe the autoregressive model using the standard multivariate linear regression model (3).

$$\begin{bmatrix} a_{n+1} \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} a_n & \cdots & a_1 \\ \vdots & \ddots & \vdots \\ a_{N-1} & \cdots & a_{N-n} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_{n+1} \\ \vdots \\ \varepsilon_N \end{bmatrix} \quad => \quad \boldsymbol{y = X\beta + \varepsilon} \tag{3}$$

Estimates of unknown parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^T$ are obtained by the least squares method (LSM), minimizing the sum of squared errors (residues)

$$min\ \boldsymbol{\varepsilon\varepsilon^T} = in \sum_{i=n+1}^{N} \varepsilon_i^2 \tag{4}$$

According to [7, 8], the sum of residuals $\varepsilon\varepsilon T$ acquires a minimum for the vector $\widehat{\boldsymbol{\beta}}$ which is a solution of the equation

$$\boldsymbol{X^T X\,\widehat{\beta}} = \boldsymbol{X^T y} \quad => \quad \boldsymbol{\widehat{\beta}} = \frac{\boldsymbol{X^T y}}{\boldsymbol{X^T X}}. \tag{5}$$

In the case where n =1, we assume that the random variable X_t depends linearly on the previous random variable X_{t-1},

$$X_t = c\,X_{t-1} + \varepsilon_t, \tag{6}$$

whereby we denote the parameter c as the autoregression coefficient. We denote such a process as AR(1).

For increments of flow (a_1, a_2,…,a_N) in a given time window of length N, the system of Equation (3) has the form (7), and the estimate of the autoregressive coefficient, $\widehat{c}$ is obtained from the formula:

$$\begin{bmatrix} a_{n+1} \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} a_n \\ \vdots \\ a_{N-1} \end{bmatrix} c + \begin{bmatrix} \varepsilon_{n+1} \\ \vdots \\ \varepsilon_N \end{bmatrix} \Rightarrow \boldsymbol{\widehat{c}} = (\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} = \left(\sum_{i=n}^{N-1} \boldsymbol{a_i^2}\right)^{-1} \sum_{i=n+1}^{N} a_i\,a_{i+1}. \tag{7}$$

3 FUNDAMENTAL DESCRIPTION OF IP FLOWS

In both deterministic and stochastic models of IP networks, the cumulative arrival process $A(s,t)$ is employed to characterize the flow of packets. This process represents the total number of packets observed within the time interval $\langle s,t\rangle$.

In discrete-time models, the analysis is typically conducted over predefined time slots or sampling windows, within which the number of arriving packets, denoted by a_i, representing the increment at time i, are recorded.

$$\forall t,s \in R^+;\ A(s,t) = A(0,t) - A(0,s) = \textstyle\sum_{i=s}^{t} a_i, \tag{8}$$

In the case of measuring IP traffic using W-shrike, we have at our disposal a vector of the cumulative time of packet arrivals T_j. After choosing the size of the time slot and using addition, we obtain the values of the increments of the IP flow in the given time slot, see Figure 1. We used a Poisson flow simulation with an average rate λ_{avg} = 2p/ms to demonstrate the method of sampling the time record of the measured traffic from Wireshark into increments of the flow time series. In Figure 1a, we have shown 50 cumulative exponential values of variables T_j, representing the occurrence of packets in time. For the size of the time slot or sample window, we chose t_s = 5ms. After addition, we obtained increments a_i with Poisson distribution with average rate λ_{avg} = 10p/ts, $a_i(\omega)$ ~ Po(10). We displayed the first 30 increment values as a time series in Figure 1.

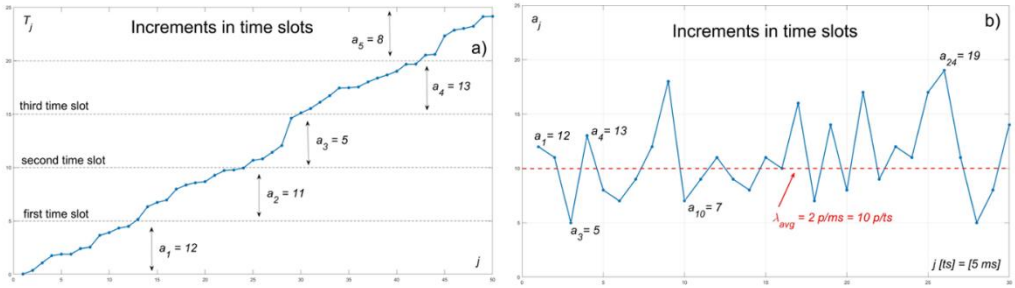


Figure 1a The 50 cumulative exponential values of variables T_j, 1b: first 30 Poisson increment values

To calculate the autocorrelation coefficient, we use wo consecutive vectors of increments $cw_1 = (a_1, \ldots, a_N)$ and $cw_2 = (a_2, \ldots, a_{N+1})$ which we will call compute windows. The increments of the flow a_i in each successive window actually represent realizations of some random variables $X(\omega)$ and $Y(\omega)$:

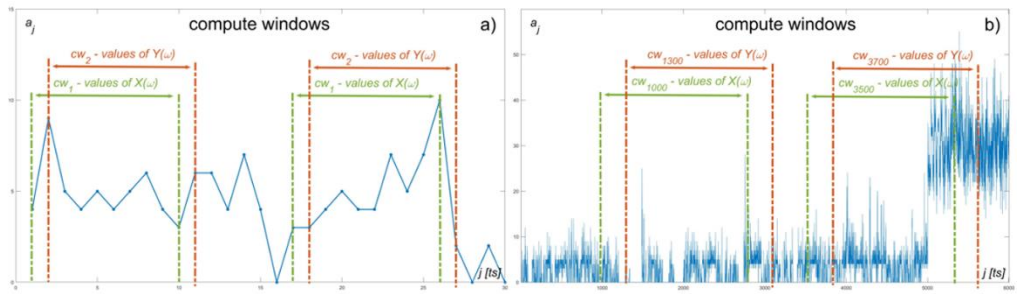


Figure 2a Compute windows cw_1 and cw_2 for Poisson process, 2b: Compute windows cw_1 and cw_2 for real DDoS attack

4 BEHAVIOR OF THE MOVING AUTOREGRESSIVE COEFFICIENT

To detect anomalies in monitored IP flow, we created a time series of the autoregressive coefficient c. We used cw_1 and cw_2 $compute$ windows to calculate these coefficient values – we gradually shift overlapping compute windows by the t_s interval, the flow increments grew by $a_i$values. For the compute window sizes, we successively chose 10 t_s, 50 t_s , 100 t_s, 500 t_s and 1000 t_s.

Standard DDoS Attack S1

Attack S1 is a standard DDoS attack, in which IP traffic suddenly increases several times in volume. The first graph in Figure 3 shows the flow increments along with the moving average rate values. In the second figure, the autocorrelation coefficient, already at the window size $cw = 10\ t_s$ responded to the onset of the attack traffic by significantly reducing its variability, while the coefficient values began to oscillate around $c(t) = 1$. With increasing compute window size, the coefficient stops responding to peaks in traffic, and for window sizes $cw = 500\ t_s$ and $cw = 1000\ t_s$, we get a threshold $c(t) = 0.95$ suitable for machine recognition of the attack.

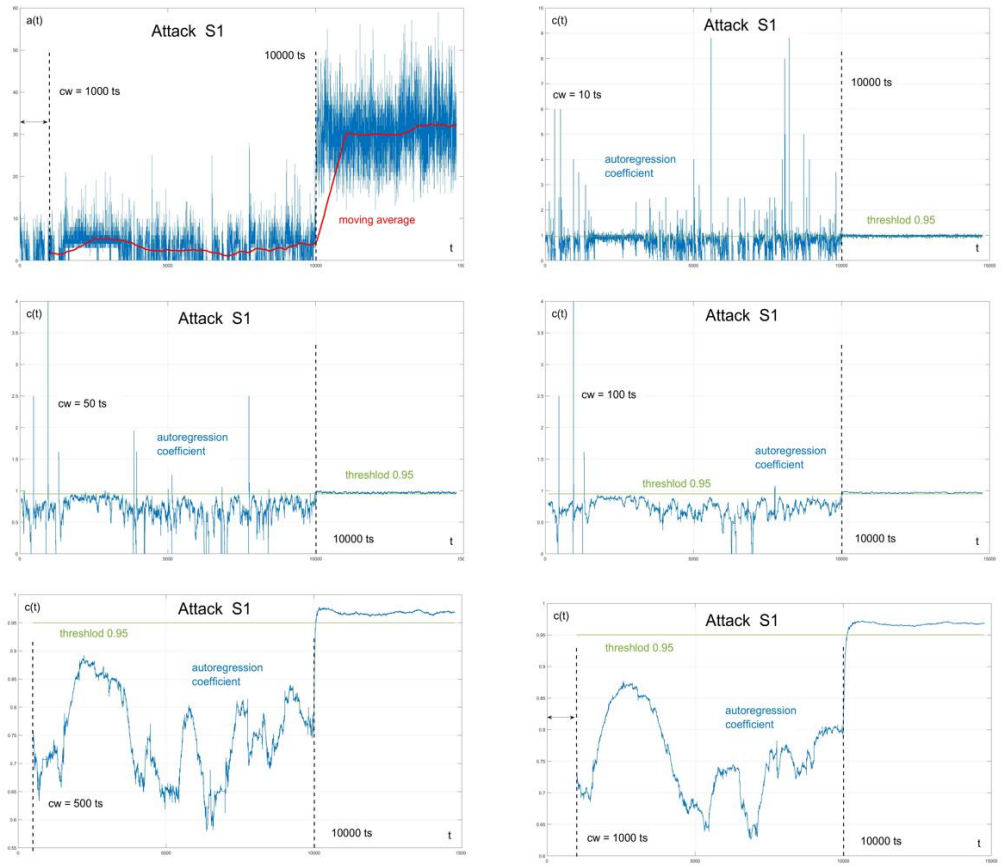


Figure 3 Standard DDoS Attack S1 detection

Standard DDoS Attack S2

Similarly, S2 is a standard type of DDoS attack, in which the volume of IP traffic suddenly increases multiple times. The autocorrelation coefficient also responded to the onset of attack traffic at cw=10 t_s by reducing variability and oscillating around $c(t)$ = 1. At cw=500 t_s and cw = 1000 t_s, the coefficient exceeds the 0.95 slope already during the first attack time slots. In both cases, 500 t_s and 1000 t_s, there were several false positives.

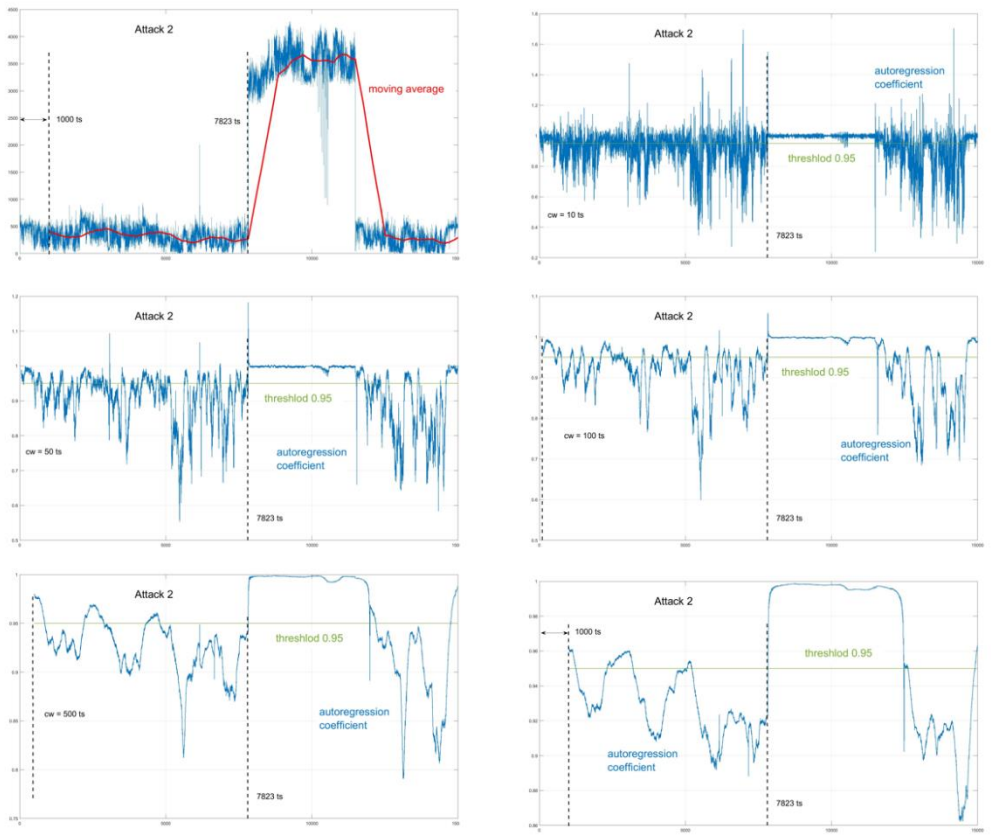


Figure 4 Standard DDoS Attack S2 detection

Two-phase DDoS Attack N7

The N7 attack differs from the previous scenarios. The attack took place in two escalating phases. The autoregressive coefficient initially behaves similarly to previous scenarios. With small compute windows, the coefficient's variance decreases significantly after the onset of the DDoS attack, and with increasing cw values, the c(t) values increase towards 1. However, probably due to the two phases of the attack, there is a significant delay in reaching the threshold $c(t) = 0.95$. At $cw = 1000$ t_s, peaks in normal network traffic were suppressed, and the evident subsequent increase in the coefficient values is recognizable by flow variability mapping methods.

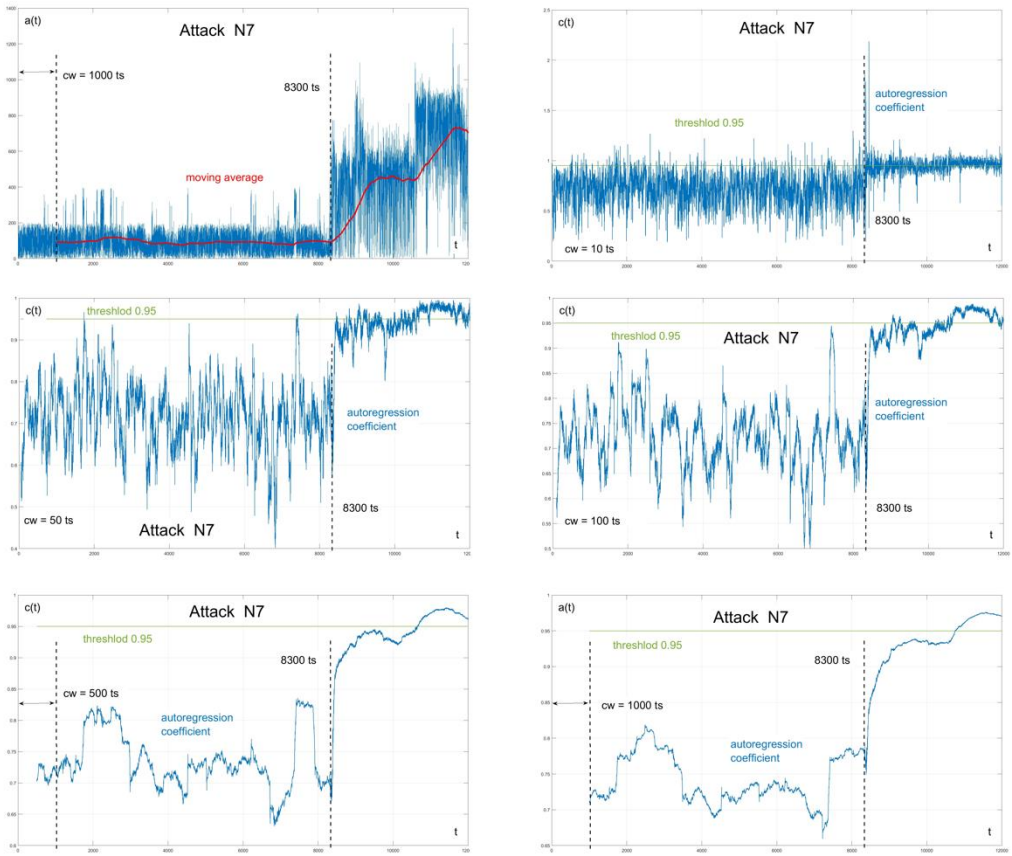


Figure 5 Two-phase DDoS Attack N7 detection

Periodic DDoS Attack N6

The N7 attack belongs to non-standard DDoS attacks. In this case, the attack is carried out with short bursts of impulses. The autoregressive coefficient reacted in the same way as in the previous standard attacks; its variability decreased significantly during the attack operation with an increasing computer window size, and at values of 500 t_s and 1000 t_s the autoregressive coefficient exceeded $c(t) = 0.95$.

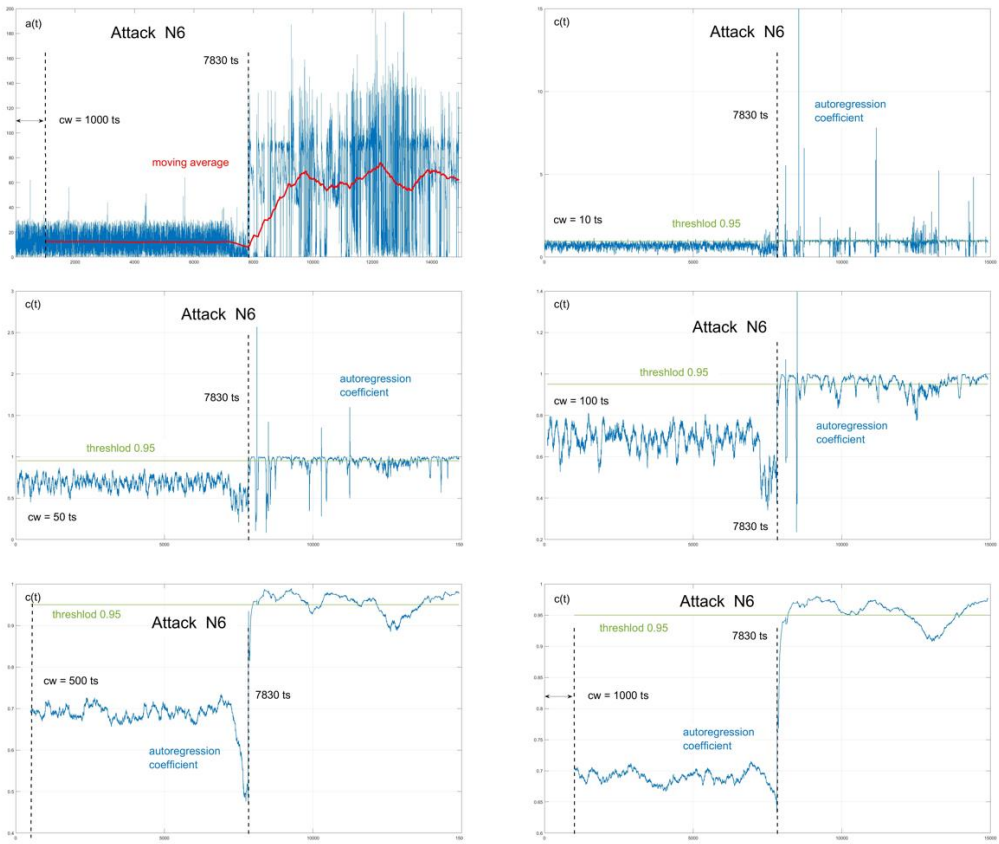


Figure 6 Periodic DDoS Attack N6 detection

Periodic DDoS Attack N8

The autoregressive coefficient in the N8 attack behaves the same as in the previous scenarios. However, in this case, the coefficient values do not reach the threshold of 0.95 during the entire attack. Machine-based attack detection methods must be used that can detect a significant impulse in the trend of the coefficient at the onset of the attack.

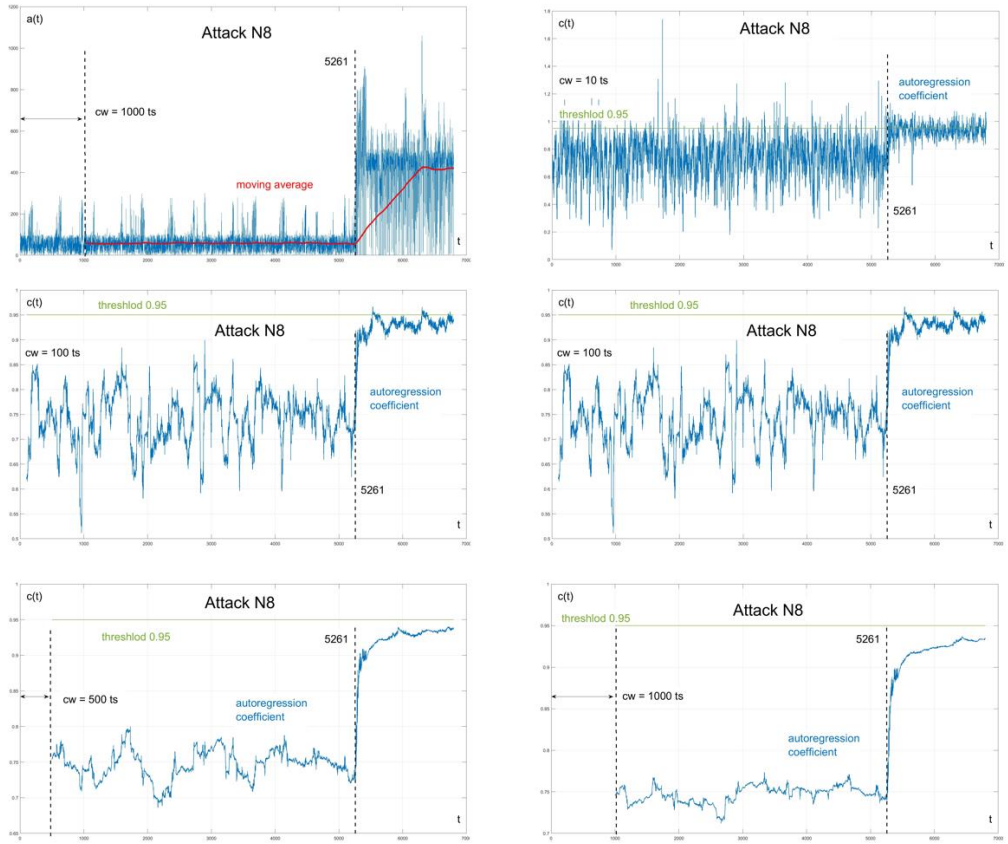


Figure 7 Periodic DDoS Attack N8 detection

Atypical periodic DDoS Attack S5

The S8 attack belongs to non-standard DDoS attacks. Again, the attack is conducted with short bursts of pulses; however, the autoregressive coefficient reacts in the opposite way compared to the previous attacks. The variability of the coefficient during the attack significantly increased, compared to standard traffic. This trend was maintained even when the compute window size was increased; however, at values of 500 and 1000 t_s there was a significant decrease in the *c(t)* values. For the attack recognition, it is again necessary to use methods that evaluate the trend of variability in time series.

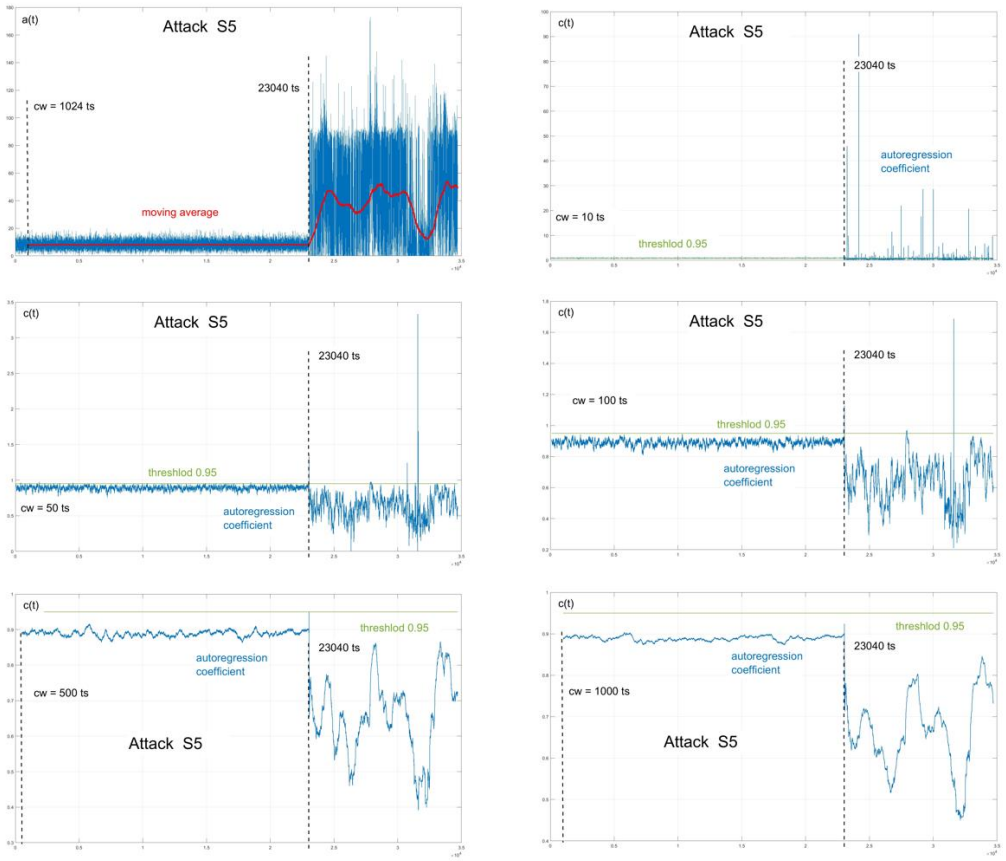


Figure 8 Atypical periodic DDoS Attack S5 detection

5 RECOMMENDATIONS FOR MACHINE-BASED ATTACK DETECTION

When testing the response of autoregressive coefficient values on various DDoS attack records, we noted three significant changes in values during various time slots of the upcoming attack:

1. With a larger number of tested attacks, there was a significant decrease in the variability of the coefficient values,
2. With increasing compute windows, the coefficient values tended to exceed the threshold c(t) = 0.95,
3. In all cases, the coefficient values responded to the onset of the attack with a significant impulse.

To further test the possibility of using the moving autoregressive coefficient, we propose three different detection functions that use the previous knowledge from the performed tests, using $cw = 1000\ t_s$ as the computer window size:

1. Change in variability of a random variable describing the coefficient of variability V, which, in case of a significant change in the standard deviation from the mean value, exceeds the value V=1. Detection function e.g., $F[V(t) > 1]$
2. In case of using a threshold, the shape of the detection function is $G[c(t) > 0.95]$
3. To record a significant impulse in a time series, the so-called prediction tunnel can be used, the width and position of which depend on the standard deviation and the average of the previous values of the coefficient $c(t)$, and then tests whether the upper limit hr of the calculated confidence interval is exceeded, condition $H[c(t) > hr]$.

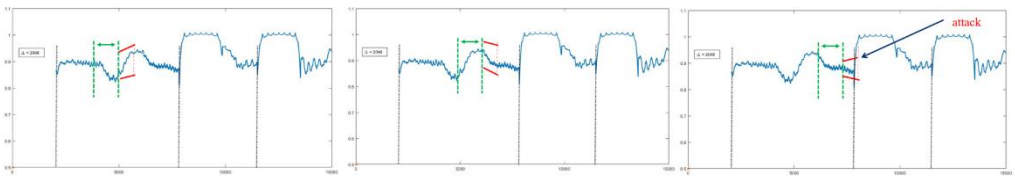


Figure 9 Example of a prediction tunnel for pulse recognition in a time series

To evaluate the reliable attack detection, we can construct different detection functions $D(t)$ using combinations of functions $F(.)$, $G(.)$ and $H(.)$:

$D1(t) = F(t) + G(t) + H(t)$

$D2(t) = F(t).G(t) + F(t).H(t) + G(t).H(t)$

$D3(t) = F(t).G(t).H(t)$

Based on further experiments, we will decide which function is the most accurate with respect to the ideal ratio of false positives to early attack detection.

6 CONCLUSION

In the article, we presented the reactions of the autoregressive coefficient to the onset of a DDoS attack. In selected scenarios, we presented three main responses, which represent a significant change in the variability of the coefficient values, exceeding the reaction threshold with a significant peak-impulse in the data. Based on the findings, we proposed several detection functions for further testing of machine recognition of DDoS using autoregression.

Acknowledgement: Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00010.

REFERENCES

[1] URAMOVA, J. Infrastructure for Generating New IDS Dataset. In Proceedings of the 2018 16th International Conference on Emerging eLearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia, 15–16 November 2018. https://ieeexplore.ieee.org/abstract/document/8572201 (accessed Nov. 27, 2025)

[2] HAJTMANEK, R.; KONTŠEK, M.; SMIEŠKO, J.; URAMOVÁ, J. One-Parameter Statistical Methods to Recognize DDoS Attacks. Symmetry 2022, 14, 2388. One-Parameter Statistical Methods to Recognize DDoS Attacks (accessed Nov. 27, 2025)

[3] SHARAFALDIN, I.; GHARIB, A.; LASHKARI, A.H.; GHORBANI, A.A. Towards a Reliable Intrusion Detection Benchmark Dataset. Softw. Netw. 2017, 2017, 177–200. https://www.researchgate.net/publication/318286637_Towards_a_Reliable_Intrusion_Detection_Benchmark_Dataset (accessed Nov. 27, 2025)

[4] SHARAFALDIN, I.; LASHKARI, A.H.; GHORBANI, A.A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. ICISSp 2018, 1, 108–116. https://www.scitepress.org/papers/2018/66398/66398.pdf (accessed Nov. 27, 2025)

[5] FELLER, W. An Introduction to Probability Theory and Its Applications, 2nd ed.; John Wiley & Sons Inc.: New York, NY, USA, 1971; Volume II. [Google Scholar]

[6] MONTGOMERY, D.; RUNGER, G. Applied Statistics and Probability for Engineers; John Wiley & Sons: Hoboken, NJ, USA, 2014; ISBN 978-1-118-74412-3. [Google Scholar]

[7] CRYER, J.D.; CHAN, K. Time Series Analysis with Applications in R; Springer: New York, NY, USA, 2009; ISBN 978-0-387-75958-6. e-ISBN 978-0-387-75959-3. [Google Scholar]