Fractal dimensionality of network traffic as a feature for intrusion detection

Viktoras Bulavas
Institute of Data Science and Digital Technologies
Vilnius University

Introduction

- Cyber threats are an evolving aspect of our daily lives, intrusion detection being one of the remedies to address information security breach.
- Intrusion detection relies on observation of network traffic features and their dynamics in time, which allows intrusion detection systems to prevent certain types of attacks upon detection.
- While rule based systems are following decision trees of prescribed conditions, anomaly recognition systems await for deviation from usual behavior of network users.
- While multiple event counters help rule-based recognition, various aggregates are calculated in order to detect anomalies.
- In this research author studies the use of Minkowski–Bouligand dimension, also known as a box-counting fractal dimension, calculated according to T. Higuchi[3] algorithm, as a possible indicator of cyber attack.

Method

- Use of fractal dimension to describe complexity of an image has been introduced by the French mathematician Benoit Mandelbrot [1]. Of the wide variety of methods for estimating the fractal dimension that have so far been proposed, the box-counting method is one of the more widely used ones, as it can be easily computed and applied to patterns with or without self-similarity.
- Based on the analogy of successful use of fractal features to recognize patterns in other fields of application including signal analysis and pattern recognition, an experiment with network traffic dataset[2] CSE-CIC-IDS 2018 was setup to study fractal dimension features of network traffic as a possible indication of a cyber-attack.
- Network traffic aggregates were represented as two-dimensional images using a method, proposed by S. Kim [3].
- Further two dimensional imagers are presented as an animation, allowing real time observation of the development of an attack (see Figure 1 a and b).

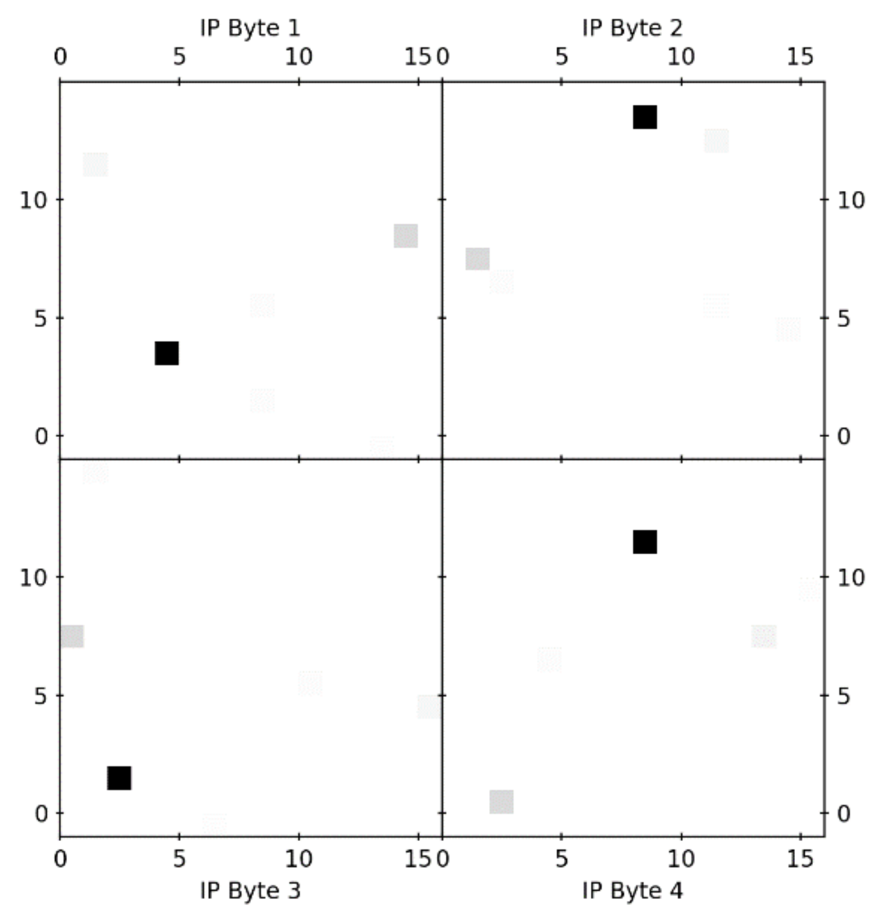

Figure.1 a. At the start of the DOS attack

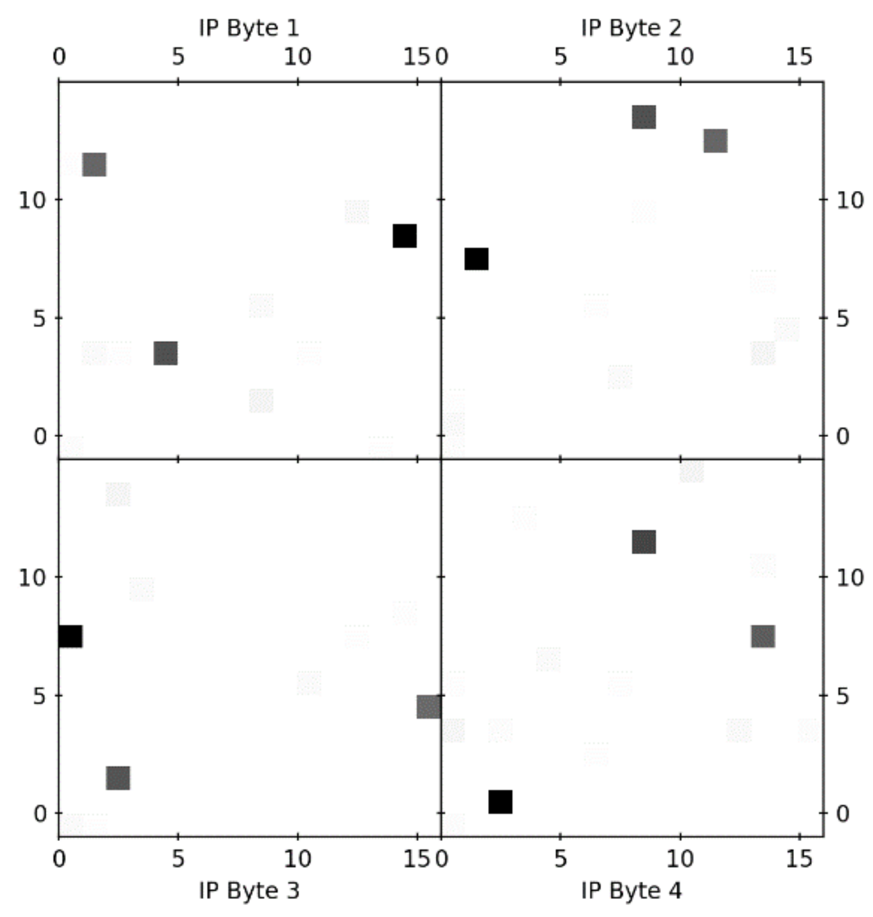

Figure 1 b. At the peak of the DOS attack

Box-counting dimension

- Definition of Box-Counting dimension was first introduced by Benoit Mandelbrot [1].
- Let $H(R^m)$ be the space of all nonempty compact subsets of the Euclidean space R^m. A compact set A in R^m is said to be self-similar if A is the union of distinct closed balls of radius r needed to cover A.
- For simplicity, as proposed by Higuchi [4], R^m is further covered by square boxes of side length r..
- For each r > 0, let $N_r(A)$ be the smallest number of boxes of side r which intersect A.
- If $D_f(A) = \lim_{r\to 0} \frac{\ln N_r(A)}{\ln\left(\frac{1}{r}\right)}$, exists, then $D_f(A)$ is called the fractal dimension or box dimension of A.

Algorithm

- To support cyber-attack detection, Box-Counting calculation according to T. Higuchi [4] algorithm was performed to extract fractal dimension of a given timeframe traffic block.
- The steps for calculating the fractal dimension are as follows:
- first, the rectangular image of the selected time period traffic is prepared from traffic data. This image is divided into 256 pixels, thus with a square side size of 16;
- second, a straight line is fitted to the points using the least-squares method, and the slope D of this fitted line is calculated. Slope D indicates the degree of complexity, or the Box-counting dimension measure of the image;
- Finally, fractal dimension values are calculated over a fixed time interval.
- The more pixels in the image, the more complex it becomes.
- After normalization of the image according to numbers of traffic records, a threshold is selected equal to average value of the grayscale.
- Following the development of this image, growing dimension indicates possible attack, and sharp decline in the dimension value indicates the end of the attack.

Results

- Data from CSE-CIC-IDS 2018 containing DOS attack records, was analyzed using time interval of 5 minutes (see Figure 2).
- With DOS attack, significant numbers of network flows are generated, therefore complexity of image increases. Due to normalization, immediately following the attack, rest of traffic is filtered by threshold, compared to the attack traffic.
- Maximum values are observed at the time of an attack and minimum following the successful attack. These results are in line with dataset events, confirming a possibility to use this feature for supporting a real time detection of the cyber-attack.

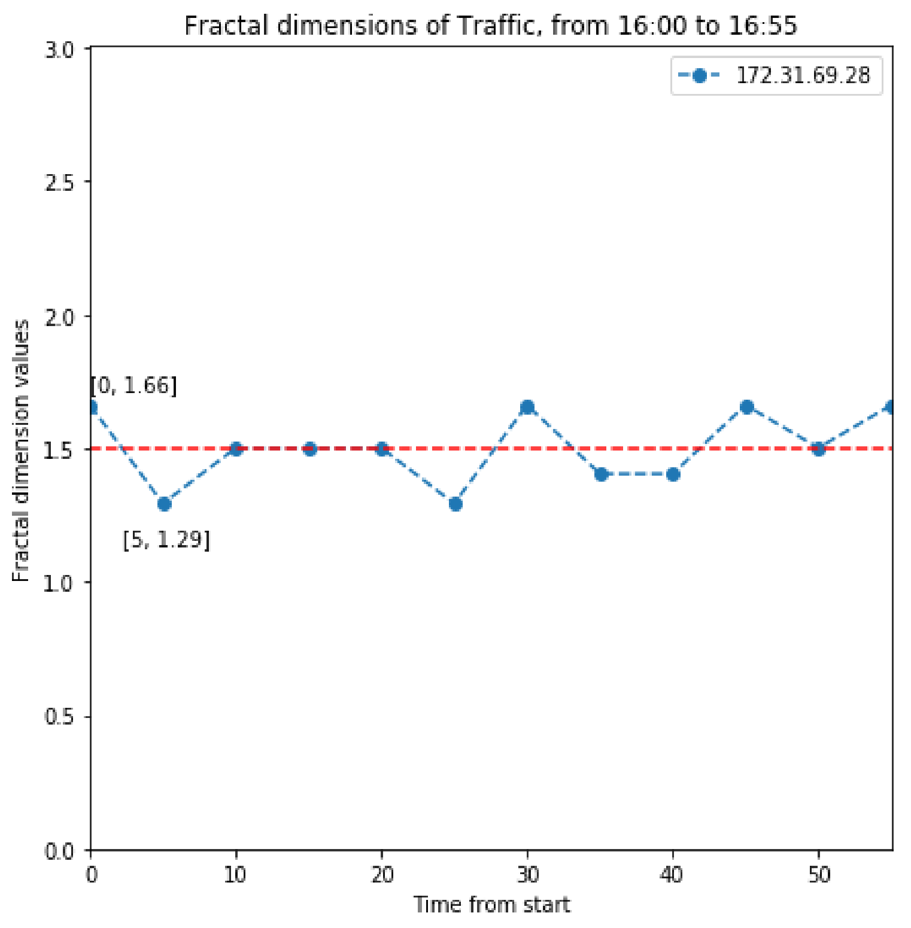

Figure 2 a. Representation of an ordinary traffic

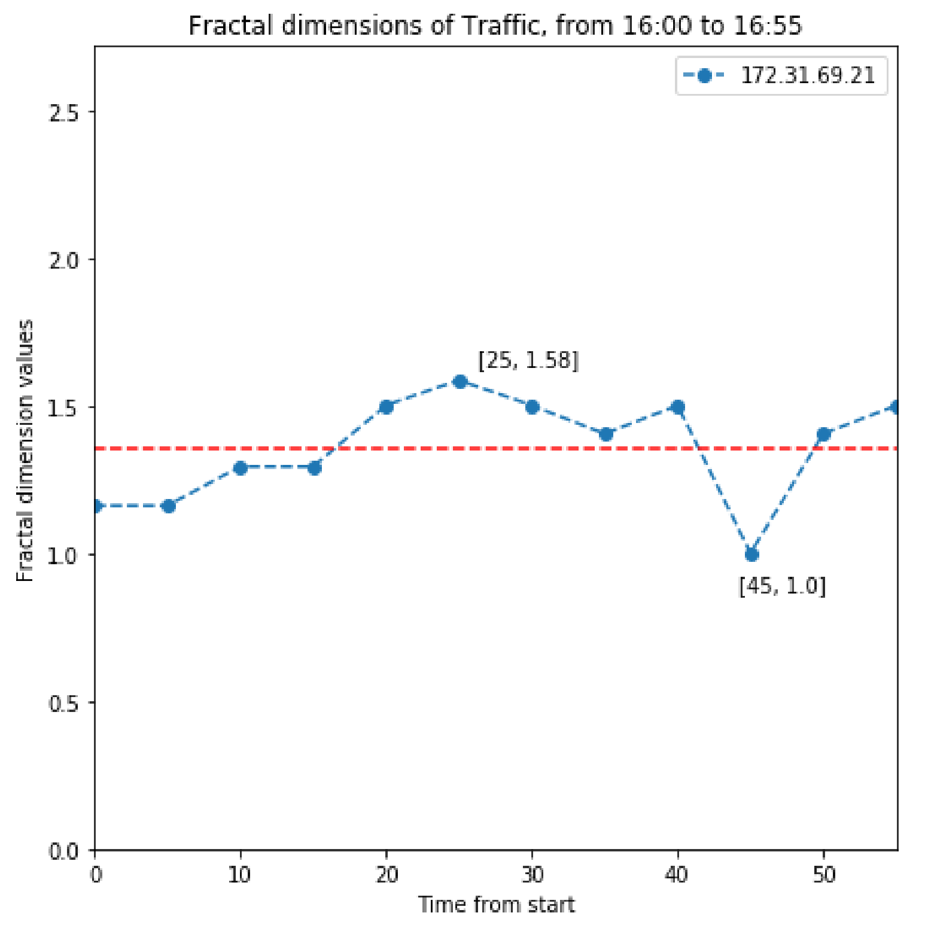

Figure 2 b. Minimum after the attack at 16:45

References

[1] Mandelbrot, B. B. (1983). The fractal geometry of nature. Henry Holt and Company.

[2] Sharafaldin, I., Lashkari, A. H. and Ghorbani, A. A. (2018) ‘Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization’, in Proceedings of the 4th International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and Technology Publications, pp. 108–116. doi: 10.5220/0006639801080116.

[3] Kim, S. S., Reddy, A. L. N. and Vannucci, M. (2004) ‘Detecting traffic anomalies using discrete wavelet transform’, Information Networking, (July), pp. 951–961. doi: 10.1007/978-3-540-25978-7.

[4] Higuchi, T. (1988) ‘Approach to an irregular time series on the basis of the fractal theory’, Physica D: Nonlinear Phenomena, 31(2), pp. 277–283. doi: 10.1016/0167-2789(88)90081-4.

Viktoras Bulavas
viktoras.bulavas@itpc.vu.lt