A polyscale autonomous sliding window for cognitive machine classification of malicious Internet traffic

Conference Paper · July 2015

Muhammad Salman Khan
University of Manitoba, MB, Canada
20 PUBLICATIONS 10 CITATIONS
SEE PROFILE

Ken Ferens
University of Manitoba
13 PUBLICATIONS 26 CITATIONS
SEE PROFILE

Witold Kinsner
University of Manitoba
378 PUBLICATIONS 1,998 CITATIONS
SEE PROFILE

A Polyscale Autonomous Sliding Window for Cognitive Machine Classification of Malicious Internet Traffic

Muhammad Salman Khan, Ken Ferens, and Witold Kinsner
Dept. of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada
muhammadsalman.khan@umanitoba.ca, ken.ferens@umanitoba.ca, witold.kinsner@umanitoba.ca

Abstract*—Features of an Internet traffic time series can be estimated using dynamical systems. Dynamical systems may exhibit chaos and strange attractors [1] [2]. Since Internet traffic shows non stationarity and long term dependence among data samples, a cognitive polyscale approach should be taken to analyze the hidden features in a nonlinear data time series. It is necessary to estimate a reasonable window of time series so that the polyscale analysis can be performed without violating the statistical bounds of the analysis. In this work, a feature extraction algorithm is developed using variance fractal dimension trajectory and the statistical parameters of the calculation are validated using an autonomous varying window of data samples. Our analysis shows promising results since the algorithm is able to capture the presence of DNS denial of service attack and has extracted the bursts of data sample accurately.*

Keywords— Cognitive machine learning, Fractal, Polyscale, DNS DDoS amplification attacks, Anomaly detection, Cyber threats, Variance fractal dimension, Non stationary trend analysis.

I. Introduction

Analysis of internet traffic requires converting the traffic parameters into a time series. With the careful implementation of sampling intervals, the time series of the traffic depicts the estimated behavior of the internet traffic [3]. If the data series is composed of N independent features, then we can represent the time series in N dimensional space. Analysis of time series [4] can be categorized as (i) time analysis, (ii) frequency analysis and, (iii) multiscale time and frequency analysis (wavelet analysis). There is a fourth category introduced by Witold Kinsner in [5] and called as polyscale analysis. Wavelets and Multifractal analysis may be used to illustrate the scale invariance and long-term memory properties of time series and objects. Both wavelets and multifractal analysis characterize the traffic using multiscale analysis; However, multifractal analysis considers the information at multiple scales simultaneously, while wavelets considers information at different scales independently [6] [5]. Therefore, multifractal analysis is also called polyscale analysis. This term was first coined and then conceptualized by Witold Kinsner in [5] [7] to signify the difference from wavelet multiscale analysis. In addition, internet traffic streams are well characterized by self-similarity and long-term memory properties [8] [9] [10]. Polyscale analysis uses various measures of complexity to extract features of the time series. Polyscale analysis is different from traditional mono-scale analysis, such as statistical methods and Fourier analysis, in that polyscale methods not only calculate statistical information at multiple scales, but they also measure the connecting factor, the fractal dimension, which is typically obtained through a log-log plot across these scales simultaneously. [11] [12].

Polyscale analysis of a time series provides significant tools to analyze the non-linearity of the series using non integer fractal dimensions and multifractal analysis. Multifractal analysis is used to detect complexity using the self-similarity or self-affinity features of time series at different scales [13]. Mathematically, we can calculate fractal dimensions by finding the exponent of the power law relationship of the multiscale coordinates over a log-log plot [14].

Contemporary statistical analytical models estimate the statistical characteristics of a time series using probability distributions, hypothesis testing and/or various probabilistic learning techniques. Alternatively, a time series may be modeled and analyzed using dynamical systems where fractal analysis plays a significant role [4] [15]. As multifractal/polyscale analysis considers non-integer dimensions embedded within topological or integer dimensions, the hidden complexities or features can be extracted by estimating the exponent of log-log relationship.

The distributed denial of service (DDoS) DNS amplification attack exploits the DNS protocol to amplify the payload of DNS packets. These packets do not contain any useful information and thus reduce the available and useful bandwidth of the network. In this type of attack, the attacker broadcasts a control message (a DNS request) to authentic computing nodes over the network in the disguise of originating from an authentic DNS server. It manipulates the source and destination IP address such that the victim node does not send the request towards the DNS server and the attacker. Rather, the DNS server traffic is directed towards the victim's computer [16] [17]. Now there is a one way route from a group of authentic nodes towards the victim node in such a way that the DNS recursive server sends responses towards the victim's node. Since these requests come from many authentic nodes continuously, the response of the DNS server directed towards the victim node is overwhelming and the victim node faces reduction in availability of the bandwidth and ultimately faces denial of service state and becomes unable to communicate to any network request.

As the attacker cannot be traced because the attack is launched using authentic nodes and the attacker remains anonymous, it is important to analyze the traffic continuously and extract the features based on varying characteristics of the traffic. There are various methods to detect DNS DDoS amplification attacks. The authors in [17] describe a method of mapping and monitoring the DNS mechanism of requests and responses to detect anomaly in the packet flows. This method shows better results in detection, but is limited due to scaling issues in a large network. Moreover, it is useful for local DNS servers only. In [18], the authors utilized hardware based Bloom filters to analyze DNS packets to detect DNS amplification attacks. Also, as mentioned in [16] [19], there are location based and time based methods to detect DNS DDoS amplification attacks. There are various methods to detect the attacks and include packet based signature analysis and node based collaborative techniques.

In this work, we applied variance based polyscale feature extraction mechanism to detect DNS attacks in a nonlinear internet time series with one selected and observed feature i.e. samples of DNS packet count. This technique is effective as it is scalable and works accurately for long duration attacks. Also it provides a unique measure of complexity introduced by the attack i.e. variance fractal dimension. Moreover, this work provides an important application of polyscale analysis in detecting malicious anomalies (attacks) in a nonlinear time series. Also, this work validates that a non-stationary data time series can be analyzed by autonomously estimating a set of subsample window with weak sense of stationarity.

II. VARIANCE FRACTAL DIMENSION AND TRAJECTORY

Variance fractal dimension analysis is a class of information based fractal analysis where second order statistics of the data samples at multiple scales are used simultaneously to estimate the power law relationship among the scales. For a single parameter/attribute/feature data time series, the variance fractal dimension is embedded within the topological dimension of 1 (a line) and 2 (an area) [12] [14]. Variance fractal dimension is calculated by estimating the Hurst exponent which is characterized by the fractal Brownian motion process [20]. Hurst exponent of 0.5 represents standard Brownian motion process. It is mandatory to ensure stationarity of time series before applying variance fractal dimension analysis [21].

Let $x(t)$ represents a periodically sampled data time series. It is important to note that the sampling frequency should be considered such that it follows the Nyquist sampling criterion [22] [23]. Moreover, since internet data time series consists of digital information packets therefore, we should consider sampling the data time series such that the original information should be preserved i.e. Nyquist sampling criterion. Moreover, as internet data time series already contains analog information i.e. speech signal represented by VoIP packets, we should consider the characteristics of digital process in such a way that the packet flow information should be preserved inside a digital sample [24]. In the current data set, we have considered DNS based packets which require a maximum round trip time (RTT) of 100ms. Although, we can configure higher round trip time since DNS packets are UDP packets, but our analysis of the data set provided us the maximum figure of 100ms for the DNS packets in the data set.

For the process $x(t)$, the variance of the data samples is:

$$var[f(x(t))] = E[(f(x(t)) - \overline{f(x(t))})^2] \qquad (1)$$

where E[.] is the statistical expectation operator and

$$\overline{f(x)} = E[f(x(t))] \qquad (2)$$

The $f(x(t))$ represents any function to calculate magnitude of samples in the multiscale calculations. For example, [12] considered $f(x(t)) = x(t_1) - x(t_2)$, where $x(t_1)$ is the first sample of the sub-window and $x(t_2)$ is the last sample of the sub-window. For our work,

$$f(x(t)) = range(x(t)) \quad (3)$$

$$range(x(t)) = \max(sample\ sub - window) - \min(ample\ sub - window) \quad (4)$$

According to power law [25],

$$var(f(x(t))) \sim |t_2 - t_1|^{2H} \quad (5)$$

where H is the Hurst parameter and is bounded between 0 and 1. If H=0, the process exhibits long range negative autocorrelation i.e. if current sample is low valued then future sample will have high value with high probability. If H=1, the process shows long range dependence and exhibits persistence of the trend i.e. if current sample is high valued then the future sample will have high value with high probability. If H=0.5, then the process exhibits no autocorrelation and samples.

Now using log,

$$\log[var(f(x(t_1) - x(t_2)))] \sim 2Hlog[\Delta t] \quad (6)$$

Therefore,

$$H = \frac{1}{2} \lim_{\Delta t \to 0} \frac{\log[var(f(\Delta x_{\Delta t}))]}{\log[\Delta t]} \quad (7)$$

The variance fractal dimension is related to H as follows [12] [25]:

$$D_v = E + 1 - H \quad (8)$$

where D_v is the variance fractal dimension embedded between integer dimension 1 and $E - 1$. For a single feature time series (i.e. DNS packet count time series), $E = 1$. Therefore,

$$D_v = 2 - H \quad (9)$$

Variance Fractal Dimension Trajectory

In order to calculate variance fractal dimension of a non-stationary time series in a continuous fashion, it is required to divide the time series in windows of data samples where each window should be chosen such that stationarity of the chosen samples is preserved. Then, VFD is calculated over each window continuously. The plot of VFD is termed as Variance Fractal Dimension Trajectory (VFDT).

Following are the important considerations for the correct calculation of variance fractal dimension in a given internet data time series:

1) Data series over which variance fractal dimension calculation is considered must show stationarity in the weak sense of second order statistics.
2) Sampling interval must be equal and should contain information of the samples reasonably i.e. round trip time (RTT) of the DNS packets. In general, the necessary condition for sampling is to know that Nyquist criterion is fulfilled. Moreover, the sufficient condition is to ensure that details of protocols and applications are analyzed properly and embedded in the data sampling interval.
3) If variance fractal dimension calculation violates the bounds of embedding dimensions i.e. calculations show negative dimension or values greater than the upper bound, then it is a sign of non stationarity and the window of samples should be varied accordingly.
4) Log-Log plot should not have saturation points in the calculations since saturation points introduces bias in Hurst value and do not contribute any polyscale information. If there are such points, consider removing them first.
5) Outliers in the calculations of variance fractal dimension should be considered as noise and the window size should be varied to remove those outliers. As the window progresses, these outliers will eventually be included in the overall calculations of VFD.
6) Only steady state samples should be considered in the data time series. Initial transient samples should be removed before applying VFDT.
7) The number of samples should be sufficiently large.

III. DATA SET

The data used in this work was the PREDICT ID USC-Lander/ DoS_DNS_amplification-20130617 (2013-06-17) to (2013-06-17) [26]. There are 19 ERF packet capture files with anonymized IPs. There are total 59,928,920 (~ 60 million) packet counts out of which there was a total of 358019 DNS packets. Out of 358019 DNS packets, 340865 packets were DNS attack packets. The total capture file size was 5.3 GB. The first packet in the file started at June 17, 2013, 21:52:45.395326000 and the last packet ended at June 17, 2013, 22:25:32.859674000. The first DNS attack packet arrived at 22:00:12 and the last DNS attack packet arrived at

22:15:34. According to the USC-Lander, this data set was composed of one DNS Denial of Service Amplification attack staged between USC/ISI, Marina del Rey, California to CSU, Fort Collins, and Colorado. The attack was performed on a single destination IP. The attacker IP was not present in the data set which used 6 DNS servers to generate a botnet network.

IV. Algorithms

Data Parsing and Generating Time Series

1) Collect the PCAP or ERF capture file using Wireshark or Tshark.

2) Break the file into small chunks of 100,000 packets per chunk.

3) Read Timestamp, Source IP, Destination IP, Packet size, packet info fields into Matlab data structures.

4) Create a DNS time series of DNS packet flow.

Adaptive Window Algorithm

1) Set the following parameters:
    a. Data pointer: d_p
    b. Window size: lag
    c. Window= d_p+ lag

2) Initialize $\boldsymbol{d_p}$ at first sample of the data series.

3) Run a loop till the end of data series.

4) Pass this window through **VFDT()** function and get estimated variance.

5) Check if estimated **variance** falls within embedding integer dimensions of 1 and 2 (check for stationarity):

    a. If **variance** > 2, increase **lag** by 64 and recalculate variance. Do not increment $\boldsymbol{d_p}$. Do this till valid variance is returned from **VFDT()**.

    b. If **variance** < 0, decrease **lag** by 64 and recalculate variance. Do not increment $\boldsymbol{d_p}$. Do this till valid variance is returned from **VFDT()**.

*Variance Fractal Dimension Trajectory – **VFDT()***

1) Let N samples are considered in a given window of data. Let's call it *main window*.

2) Select the largest size of cover (samples per cover) that corresponds to the largest scale. Let's number it scale 1. It should be chosen such that the main window of samples should provide at least 30 covers in the first scale i.e. scale 1.

$$K_H \geq \left\lceil \frac{log30}{logb} \right\rceil$$

Therefore, we chose the following relation to calculate K_H as the total number of scaling levels:

$$K_H = \left\lceil \frac{logN}{logb} \right\rceil - \left\lceil \frac{log30}{logb} \right\rceil$$

3) Select the lowest size of cover such that at least 2 samples per cover are available.

$$K_L \geq 1$$

Therefore, we chose the following relation as the lower bound of scaling level:

$$K_L = 1$$

4) Run main loop for K from K_H till K_L.

5) Now set the following parameters:

    a. Total number of samples per cover at level K

$$n_K = b^K$$

where b is a generic number base. We use b=2.

    b. Total number of cover at K- level are:

$$N_K = \left\lfloor \frac{N}{n_K} \right\rfloor$$

6) In each iteration of the main loop, run second loop from 1 till N_{K_H}. This loop calculates the difference in maximum variation in each cover as follows:

$$\Delta y_{K_{H_i}} = \max(sample\ sub - window) - \min(ample\ sub - window)$$

where ***i*** will run from 1 till n_{K_H}.

7) When the second loop is completed, calculate the variance at level K_H as follows:

$$var(\Delta y_{K_H}) = \frac{1}{N_k - 1}\left[\sum_{j=1}^{N_k}(\Delta y_{K_H})^2 - \frac{1}{N_k}(\sum_{j=1}^{N_k}\Delta y_{K_H})^2\right]$$

8) Calculate the coordinates on log-log plot as follows:
$x_k = \log(n_{K_H})$
$y_k = \log(var(\Delta y_{K_H}))$

9) Reduce the value of K by 1 till it reaches K_L. Return to step 5.

10) From the set of coordinates, calculate slope of the log-log plot and Hurst parameter of the *main window*.

V. Experimental Results

In this work, an autonomous sliding window algorithm is developed to characterize a non-stationary data time series using a range based variance fractal dimension trajectory. The PREDICT data set [26] was used to perform the analysis. This dataset contains 19 ERF capture files with anonymized IPs. Moreover, the attack was recorded for 10 minutes and the packets were captured for 32 minutes and 47 seconds. Each file has more than 3.5 million packets. One target IP and six DNS server IPs are known a-priori. The files contain both attack and legitimate packets, including DNS packets. The algorithm was implemented using Matlab, and the data parsing was done by breaking a file into multiple parts, where each part contained 100,000 packets. From this dataset, we generated a time series plot of the DNS traffic.

The experiment investigated the effect of applying different window sizes and lag values over the data set. As shown in Fig. 1, the time series of DNS packet counts contains both normal and attack packets. The attack start time and end time are shown by a blue arrow. Moreover, we observe a large spike at the start of the time series (sample number 75); this represents the start of the system where nodes broadcast to DNS server for their query resolution. It is a normal process and as expected; this spike is resolved within 100ms of the time sample and the series settles down with small values. Fig. 2 shows an edited version of original time series in Fig. 1, where the early spike is removed. We can observe very high varying data series having multiple bursts of DNS packets. The attack started at sample number 729 and ended at sample number 9068. The series climbed up in the vicinity of 729 and settled down again in the neighborhood of 9068. As shown in Fig. 3, the variance fractal dimension trajectory (VFDT) of this data series (without removing the early spikes) is generated when the window size of 256 samples are chosen with a single sample sliding window.

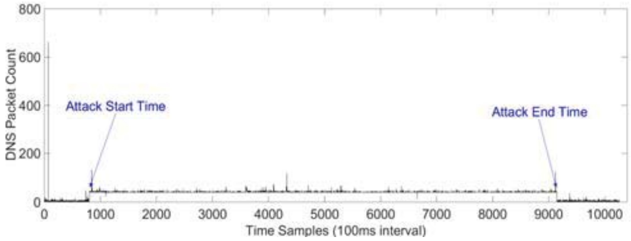

Fig. 1 DNS packet count time series.

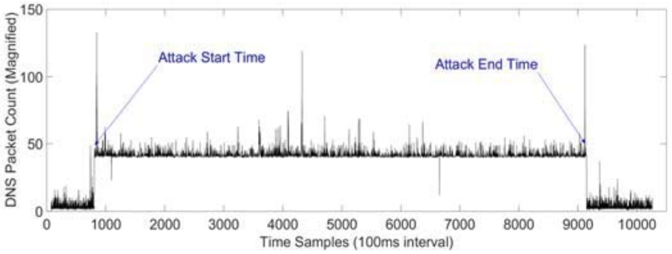

Fig. 2 DNS packet count time series (w/o first spike).

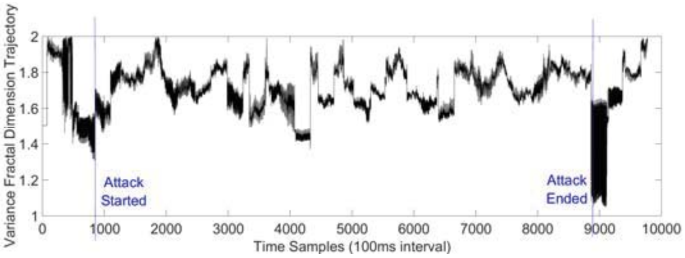

Fig. 3 VFDT, Window size=256, Lag=64.

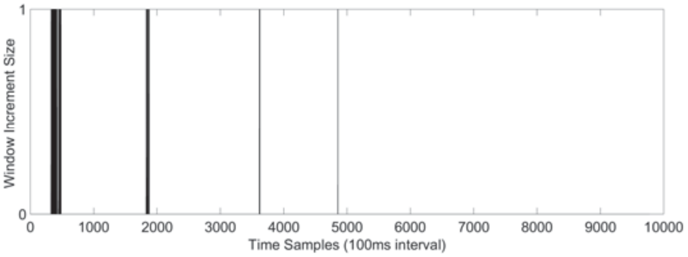

Fig. 4 Window size=256, Window breathing pattern.

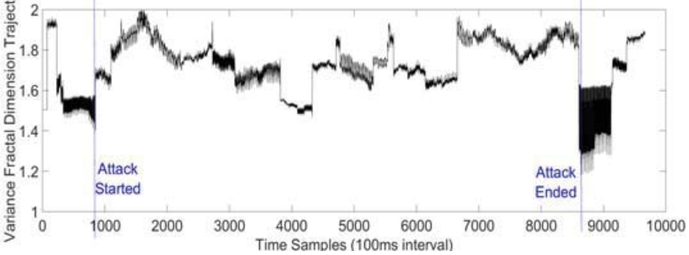

Fig. 5 VFDT, Window size=512, Lag=64.

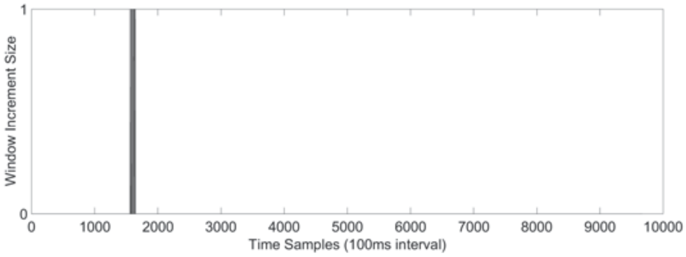

Fig. 6 Window size=512, Window breathing pattern.

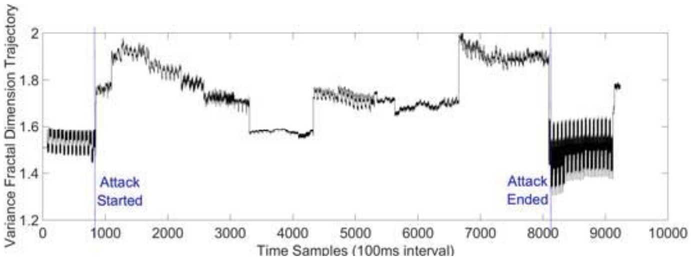

Fig. 7 VFDT, Window size=1024, Lag=64.

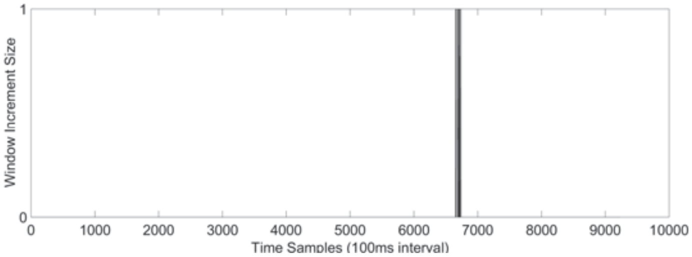

Fig. 8 Window Size=1024, Window breathing pattern.

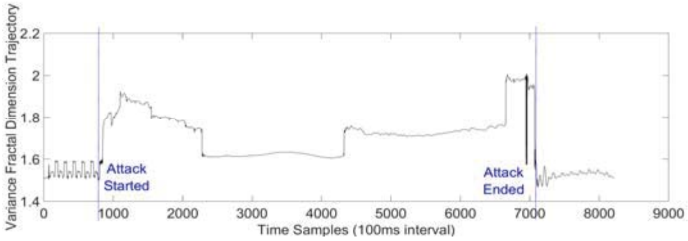

Fig. 9 VFDT, Window size=2048, Lag=64.

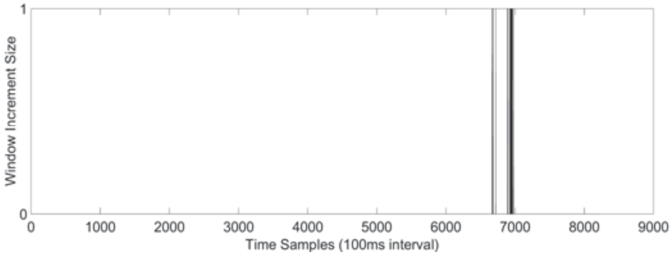

Fig. 10 Window size=2048, Window breathing pattern.

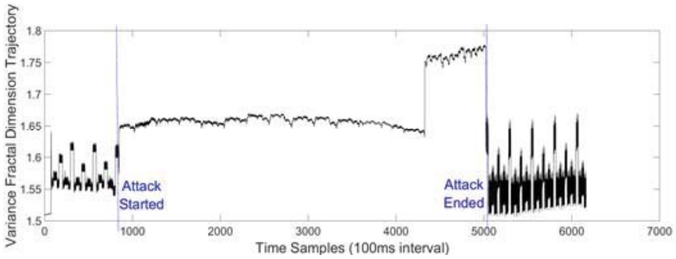

Fig. 11 VFDT, Window size=4096, Lag=64.

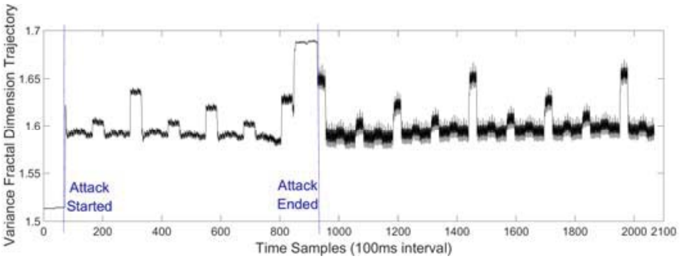

Fig. 12 VFDT, Window size=8192, Lag=64.

The VFDT remains within valid topological dimension of 1 and 2. As shown in Fig. 4, the autonomous increment of 64 samples in window size is performed in order to ensure stationarity in the weak sense of second order statistics. Moreover, it is observed that there are lot of variations in the VFDT that show the independence of calculation of variance among adjacent window sizes. With this trajectory of 256 samples per window and sliding window of single sample, the features of the data series are not only extracted but amplified. If we visually compare this trajectory with the original data series in Fig. 1, the time interval before the start of attack shows very highly varying samples with one outlier peak. Range calculations of variance based multifractal with small window size of 256 samples takes into account the presence of this outlier peak and as a result we see that VFDT stays close to the topological dimension of 2. However, after the start of the attack, the range of peaks remain comparatively low varying and therefore, we see that the trajectory falls below the fractal dimension of 1.4. Further, we see that the data series variance becomes very low varying and then starts having relatively highly varying data series. Therefore, within the time samples of 1000 and 2000, we see an increase in the VFDT trend so that it tends to be close to the topological dimension of 2. Same argument can be extended to the rest of the time series. Moreover, Fig. 4 shows the band of window size increments by 64 samples when the variance fractal dimension calculation shows invalid value i.e. greater than topological dimension of 2. We see that this autonomous breathing of window size is prominent when the data series shows sudden increase in variance that in turn increases the trajectory close to the topological dimension of 2.

As indicated in Fig. 3, the start of attack is obvious when there is an increase in the variance fractal dimension trajectory greater than 1.4. Moreover, it indicates the end of attack when the variance fractal dimension trajectory gets below 1.2. The high change in variance occurs since the window increments are done using a single sample increment (fractal amplification). When the window pointer reaches to the normal sample, it again starts increasing since the variations in the normal data series are quite high.

As shown in figures, 0, Fig. 7, Fig. 9, Fig. 11 and Fig. 12, increasing the window size (512, 1024, 2048, 4096 and 8192 samples respectively) significantly change the envelope of variance fractal dimension trajectory. As observed, the wiggling of VFDT decreases with increasing window size and then increases in Fig. 11 and Fig. 12. Moreover, the start of attack time i.e. change of variance fractal dimension trajectory from approximate 1.5 to 2 shows dependence on the size of window and the range of variability in each window size. Also, as indicated, Fig. 12 shows start of attack at very early stage of the data series which is due to the large window size of 8192 samples. Likewise, the end time is indicated earlier following the same argument. In addition, it is also observed that increasing window size also flattens the VFDT to the extent that it no more reaches the topological dimension of 2 in the vicinity of start time of attack. This happens due to the averaging of the variance calculations at multiple scales and the presence of multiple coordinates on the log-log linear plot. Also, since we are using linear regression to estimate the slope of the log-log plot, therefore, at large window sizes, there are various points on the log-log plot and the estimation of slope is biased towards higher values. This in turn, reduces the fractal dimension calculations away from topological dimension of 2. Carrying on the same argument, window size of 4096 and 8192 samples shows no autonomous variation of window size that is another validation of our argument.

Therefore, it can be deduced that for this data set, a window size of 2048 is sufficient to show variance fractal dimension trajectory while it is able to indicate the presence of attack accurately i.e. dimension increases towards topological dimension of 2. However, the end of attack is indicated quite earlier which is also attributed to the termination of data time series (Fig. 1) that flattens the variance calculations. Moreover, during the duration of attack, the variance dimension trajectory shows lower dimensions but still remains above the topological dimension of normal data series.

VI. Conclusions

This paper has described a new variance fractal dimension trajectory calculation for internet data time series that shows non stationarity and contains malicious attacks. A new range based variance fractal dimension calculation method is presented to generate variance fractal dimension trajectory. If the variance fractal dimension trajectory reaches to the topological dimension of 2 and/or shows greater dimension than the dimension of normal traffic, we can predict the presence of an attack quite accurately. Also, we have described the method of varying data window size autonomously based on testing stationarity using weak-sense second order statistics. Our algorithm is capable of capturing highly varying data samples and is prone to correlation effects of data samples in previous windows.

VII. Acknowledgement

This work is supported in part through a research fellowship from Mitacs-Accelerate Canada. Authors are also thankful to PREDICT USA for providing state-of-the-art data sets.

References

[1] Muhammad Salman Khan, Ken Ferens and Witold Kinsner, "A chaotic complexity measure for cognitive machine classification of cyber-attacks on computer networks," *International Journal of Cognitive Informatics and Natural Intelligence (in print),* vol. 9, 2015.

[2] Muhammad Salman Khan, Ken Ferens and Witold Kinsner, "A chaotic measure for cognitive machine classification of distributed denial of service attacks," in *Proc. 13th IEEE Intern. Conf. Cognitive Informatics and Cognitive Computing, ICCI*CC 2014*, London, UK, 2014.

[3] Thanasis Vafeiadis, Alexandros Papanikolaou, Christos Ilioudis and Stefanos Charchalakis, "Real time network data analysis using time series models," *Simulation Modelling Practice and Theory,* vol. 29, pp. 173-180, 2012.

[4] Holger Kantz and Thomas Schreiber, Non Linear Time Series Analysis, Cambridge University Press, UK, 2004, pp. 87-100.

[5] Witold Kinsner, "It's time for polyscale analysis and synthesis in cognitive systems," in *IEEE 10th Intl. Conf. Cognitive Informatics & Cognitive Computing (ICCI*CC11)*, Banff, AB, 2011.

[6] Witold Kinsner, "It's time for multiscale analysis and synthesis in cognitive systems," in *IEEE 10th Intl. Conf. Cognitive Informatics & Cognitive Computing (ICCI*CC11)*, Banff, AB, 2011.

[7] Witold Kinsner, "Polyscale analysis and fractional operators for cognitive systems," in *IEEE 13th Intl. Conf. Cognitive Informatics & Cognitive Computing (ICCI*CC14)*, London, UK, 2014.

[8] Will E. Leland, Murad S. Taqqu, Walter Willinger and Daniel V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking (TON),* vol. 2, no. 1, February 1994.

[9] Mark E. Crovella and Azer Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Transactions on Networking,* vol. 5, no. 6, pp. 835-846, December 1997.

[10] Patrice Abry, Richard Baraniuk, Patrick Flandrin , Rudolf Riedi and Darryl Veitch, "Multiscale nature of network traffic," in *IEEE Signal Proc. Mag.*, 2002.

[11] Changzheng Chen, Zhong Wang, Yi Gou, Xinguang Zha and Hailing Miao, "Wavelet based multifractal analysis to periodic time series," *Journal of Computational and Nonlinear Dynamics,* vol. 10, no. 1, September 2014.

[12] Witold Kinsner and Warren Grieder, "Amplification of signal features using variance Ffactal dimension trajectory," *International Journal of Cognitive Informatics and Natural Intelligence,* vol. 4, no. 4, pp. 1-17, Oct. 2010.

[13] Yingxu Wang, Jean-Claude Latombe, Du Zhang and Witold Kinsner, "Advances in Cognitive Informatics and Cognitive Computing," *International Journal of Cognitive Informatics and Natural Intelligence,* vol. 3, no. 4, pp. 91-95, 2009.

[14] Witold Kinsner, "A unified approach to fractal dimensions," *Int'l Journal of Cognitive Informatics and Natural Intelligence,* vol. 1, no. 4, pp. 26-46, 2007.

[15] Robert L. V. Taylor, "Attractors: Nonstrange to Chaotic," *Society for Industrial and Applied Mathematics, Undergraduate Research Online,* pp. 72-80, 21 6 2011.

[16] Saman Taghavi Zargar, James Joshi and David Tipper,, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," in *IEEE Communications Surveys & Tutorials*, 2013.

[17] Georgios Kambourakis, Tassos Moschos, Dimitris Geneiatakis and Stefanos Gritzalis, "Detecting DNS amplification attacks," *Lecture Notes in Computer Science,* vol. 5141, pp. 185-196, 2008.

[18] Changhua Sun, Bin Liu and Lei Shi, "Efficient and low-cost hardware defense against DNS amplification attacks," in *IEEE GLOBECOM*, 2008.

[19] Tao Peng, Christopher Leckie and Kotagiri Ramamohanarao, "Survey of network-based defense mechanisms countering the DoS and DDoS problems," *Journal ACM Computing Surveys (CSUR),* vol. 39, no. 1, 2007.

[20] Masahiro Nakagawa, "A critical exponent method to evaluate fractal dimensions of self-affine data," *Journal of Physic Society Japan,* vol. 62, 1993.

[21] Angkoon Phinyomark, Pornchai Phukpattaranont and Chusak Limsakul, "Fractals, Applications of Variance Fractal Dimension: A Survey," *Complex Geometry, Patterns, and Scaling in Nature and Society,* vol. 22, no. 1, 2014.

[22] Athanasios Papoulis and S. Unnikrishna Pillai, Probability, Random Variables and Stochastic Processes, 4 ed., McGraw-Hill, 2002.

[23] Alan V. Oppenheim, Ronald W. Schafer and John R. Buck, Discrete-Time Signal Processing, 2 ed., Upper Saddle River, NJ: Prentice-Hall, Inc., 1999.

[24] Peng CK, Havlin S, Stanley HE and Goldberger AL., "Quantification of scaling exponents and crossover phenomenon in non stationary heart beat times eries," *Chaos,* vol. 5, no. 1, 1995.

[25] Brandon J. Whitcher, Simon D. Byers, Peter Guttorp and Donald B. Percival, "Testing for homogeneity of variance in time series: Long memory, wavelets, and the Nile river," *Water Resources Research,* vol. 38, no. 5, 2002.

[26] PREDICT-USC-Lander-DoS_DNS_amplification, "Scrambled Internet Measurement, PREDICT ID USC-Lander/ DoS_DNS_amplification-20130617 (2013-06-17) to (2013-06-17) provided by the USC/Lander Project.," 2013.