A Cognitive Multifractal Approach to Characterize Complexity of Non-Stationary and Malicious DNS Data Traffic Using Adaptive Sliding Window

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—This paper presents a cognitive feature extraction model based on scaling and multifractal dimension trajectory to analyze internet traffic time series. DNS (Domain Naming System) traffic time series is considered that contains tagged DNS Denial of Service attacks. The first step of the analysis involves transforming the DNS time series into a multifractal variance dimension trajectory keeping statistical stationarity of data intact. Then features of the trajectory are extracted to remove high variability noise. The extracted set of features indicates the presence of an attack when the denoised trajectory shows increasing variance fractal dimension. This technique is superior in finding changing patterns of a data series due to the presence of noise and denial of service attack because it is not dependent on integer dimensions and mono-scale measurement of variations in data series. Moreover, this technique provides adaptive and locally stationary windows in a highly non stationary data series.

Keywords— Cognitive Machine Learning, Chaos, Fractal, DNS, DDoS Amplification, Internet, Anomaly Detection, Cyber threats, Data traffic, Multifractal, Variance Fractal Dimension, Wavelet, Haar function, Change Detection, Non Sationary, Trend Analysis, Adaptive Sliding Window, Time Series Analysis.

I. Introduction

The classification of patterns and features of a time series is evolving towards cognitive multifractal analysis. Monofractal analysis is limited in detecting hidden features of a process due to its inability to see the variations at smaller scale of measurement. Time series which arise from chaotic and nonlinear systems have great wealth of details and complexity which cannot be analyzed completely using monoscale analysis [1] [2]. Multifractal analysis measures the complexity of an object in a non-integer dimension that is in contrast to monoscale analysis where dimensions are restricted to integers only. The concept of non-integer dimension arises due to the power law relationship in the data samples. There are various estimates of fractal dimension from a practical time series such as the Box-Counting Method [3], the Katz-Servcik method [4], Detrended Fluctuation Analysis [5], Power Spectral Density Analysis and Critical Exponent Analysis [6]. Variance fractal dimension (VFD) estimate is another method that measures the change in variance at multiple scale for a given data series [7]. In this work, we utilized variance fractal dimension estimates to characterize a DNS count data series for the identification of the presence of an attack.

Estimation of stationarity of a data series is another important aspect, since measurement of fractal features at multiple scales requires that the rate of change of variance remains linear at multiple scale i.e. log-log linear plot. If the data series is not stationary, variance estimates will produce many outliers at multiple scales and the log-log plot will not be linear. There are various methods to measure the stationarity of a time series using sliding window and finding locally stationary data set [8] [9] [10]. In this work, the concept of variance change detection using wavelet transformation is used for a particular data window.

A preliminary analysis is performed to characterize the features of a non-stationary data series using a sliding and adaptive window algorithm. The data series is contaminated by noise, non stationarity and denial of service attacks. A multifractal analysis using variance fractal dimension trajectory is performed and the test of stationarity is done using wavelet based variance detection algorithm. Moreover, a trend analysis of the

Proc. 2015 IEEE 14th Int'l Conf. on Cognitive Informatics & Cognitive Computing (ICCI*CC'15)
N. Ge, J. Lu, Y. Wang, N. Howard, P. Chen, X. Tao, B. Zhang, & L.A. Zadeh (Eds.)
978-1-4673-7290-9/15/$31.00 ©2015 IEEE

calculated variance fractal dimension is also performed to identify the presence of an attack.

II. Multifractals

Internet traffic analysis employs sampling the traffic at network gateway based on various protocols of the data traffic. Converting a network traffic into a time series provides a baseline for the characterization of network traffic based on various features. If the features are independent, then mathematically, we can model the time series as an N dimensional process where N represents the number of independent features. We can classify the analysis of a time series into three types; (1) time domain analysis, (2) frequency domain analysis (3) multiscale analysis. Traditionally, available work on internet traffic characterization is dependent on monoscale analysis in either time or frequency domain. However, recently, multiscale analysis using simultaneous time and frequency parameters has gained attention of the researchers [2]. Multiscale (multifractal) analysis of a time series is based on calculating the non-integer dimension of a data process in discrete time. Multifractal analysis is used to detect self-similarity or self-affinity among signal components which is also known as power law or log-log analysis.

Time series of network traffic can be considered as being generated by a chaotic process in a nonlinear dynamical system [11] [12] [13]. They possess multiscale complexity and can be evaluated using multifractal analysis. Typically, multifractals are characterized by 2 distinct features:

1. They render self-similarity (Isotropic) or self-affinity (anisotropic) at different scales which is due to the fact that whole structure of the object or process is related to the smaller parts of the structure with a non-integer scaling factor.
2. Multifractals cannot be defined using integer dimensions or topological dimensions. Non integer dimensions are used to characterize the multiscaling property of multifractals.

Also, multifractal analysis is divided into morphological fractals and information fractals [2] [14]. This paper utilized Variance Fractal Dimension (VFD) which falls in the category of information based fractal dimension. VFD has certain advantages over the existing information based fractal dimensions as follows [15]:

1) It extracts the complexity of the underlying process.
2) It provides a normalized range for the extracted features/complexity of the underlying process which is the range of embedding dimension.

It is also noted that non integer fractal dimensions are bounded by embedding dimensions which describe the upper and lower bound related to the fractals. For example, a single feature based variance dimension is embedded between dimension 1 (a line) and dimension 2 (an area).

Variance Fractal Dimension

Variance Fractal Dimension (VFD) is an information based fractal dimension that extracts the variance feature of an object. For a single dimension object, VFD is embedded between a line (topological dimension is 1) and an area (topological dimension is 2) [7]. VFD is calculated by the Hurst Exponent (H) which is characterized by fractional Brownian motion process (fBm). A fBm $\{B_t, t \geq 0\}$ is a self-similar stochastic process that has stationary increments. A fBm is governed by Hurst parameter H ∈ [0,1]. The stationary increments have a normal distribution with zero mean and variance that is dependent on the time step t. As $H = \frac{1}{2}$, fractional Brownian motion process is called standard/ordinary Brownian motion process and the stationary increments also become independent [16].

Let $\{B_t, t \geq 0\}$ is defined as a fractional Brownian motion process of Hurst parameter H ∈ [0,1] with zero mean and covariance function as follows:

$$E(B_t B_{t+\tau}) = \frac{1}{2}((t+\tau)^{2H} + t^{2H} - |\tau|^{2H}) \qquad (1)$$

where $H = \frac{1}{2}$, $E(B_t B_{t+\tau}) = \min(t, t+\tau)$ [17], which is the covariance of a zero mean Gaussian process and represents the independence of increments which is the property of standard Brownian motion process. Also, it can be seen that where $H \neq \frac{1}{2}$, the increments are not independent. VFD calculation is performed using power law relationship between the amplitude increments of the time series. It is imperative to note that the time series is required to be stationary in the statistical sense for the valid calculation of VFD. Therefore, a sliding window of data samples is chosen for VFD calculation such that the stationarity is ensured in the weak sense and a trajectory of the VFD is obtained which varies within the embedding dimensions of the time series. This trajectory is called Variance Fractal Dimension Trajectory (VFDT).

Let x(t) represents a data time series which is sampled at equal intervals. Theoretically, x(t) can be represented as the discrete and sampled output of a stochastic process. It is also imperative to note that the sampling frequency should be chosen such that it preserves the information content of the process. In speech and analog processing systems, Nyquist sampling criterion is used to ensure the preservation of information [18] [19]. However, internet data series are essentially a digital process where packets are formed using bits of the information from a computing platform. These bits in turn are produced by an Analog to Digital Convertor (ADC) and contain statistical information. For example, voice packets using a Voice over IP protocol (VoIP) are sampled and quantized and contain information that has least squared error within the bound of the ADC. For a process x(t);

$$\text{Variance} = \text{var}[\text{x(t)}] = E[(x - \bar{x})^2] \tag{2}$$

where E(.) is the statistical expectation operator and $\bar{x}$ is the statistical mean (first moment) of the processes x(t). Therefore, according to power law:

$$var[x(t_2) - x(t_1)] \backsim |t_{2-}\, t_1|^{2H} \tag{3}$$

Let $$var[x(t_2) - x(t_1)] = var\ [\Delta\text{x}_{\Delta t}] \tag{4}$$

Let $$x(t_2) - x(t_1) = \Delta\text{x}_{\Delta t} \tag{5}$$

Then $$\log(var[x(t_2) - x(t_1)]) \backsim 2Hlog\ [\Delta t] \tag{6}$$

We can plot the relationship between $log(var[x(t_2) - x(t_1)])$ and $log\ [\Delta t]$. The half of the slope of the linear interpolation of the plot gives Hurst Exponent (H) which is mathematically equivalent to the following:

$$H = \frac{1}{2} \lim_{\Delta t \to 0} \frac{\log[var(\Delta\text{x}_{\Delta t})]}{\log(\Delta t)} \tag{7}$$

The variance dimension (D_σ) is calculated using H as:

$$D_\sigma = E + 1 - H \tag{8}$$

where E is the embedded Euclidean dimension.

In the case of single Euclidean dimension, i.e. single independent variable, we will have $E = 1$. Therefore,

$$D_\sigma = 2 - H \tag{9}$$

So for a data time series with one measurable parameter (feature), D_σ varies between 1 and 2. If D_σ=1.5, the process will represent standard fractional Brownian motion (fBm) [17]. When D_σ=1, the process is not showing any multiscale complexity and can be referred as a monofractal. When, D_σ = 2, the process is equivalent to a white noise process.

In order to compute variance fractal dimension, following parameters are required to be considered [7]:

1. Data time series over which the computation is performed must be stationary.
2. The sampling interval Δt should be fixed and is obtained following Nyquist sampling criterion.
3. For a single dimensional data series, the Embedding cover should be within the range of 1 and 2.
4. The saturation points in a log-log plot should be removed before calculating the slope of the log-log plot.
5. Outliers in a log-log plot should be considered in the context of the noise and non-linearity in the time series. Window size for computation should be adjusted to ensure that outliers do not contribute towards noise in the computation of variance fractal dimension.
6. The number of samples should be sufficient enough.

Now assuming that the sampling interval is a fixed Δt and the total time of the data series is T, then the points on a log-log plot are represented as follows:

$$(x_i, y_i) = (\log(\Delta t_i), \log(\Delta \text{B}_i)) \tag{10}$$

where ΔB_i is the amplitude of the first and last samples of the interval Δt.

Therefore, the Linear Least Square regression of the log-log plot is done as follows [7] [1]:

$$\text{slope}= 2H = \frac{K\sum_{i=1}^{i=K} x_i y_i - (\sum_{i=1}^{i=K} x_i)(\sum_{i=1}^{i=K} y_i)}{K\sum_{i=1}^{i=K} {x_i}^2 - ((\sum_{i=1}^{i=K} x_i)^2)} \tag{11}$$

$$\text{slope} = \frac{Cov(X,Y)}{Var(Y)} \tag{12}$$

where H is the Hurst parameters and K is the largest level of fractal cover.

We can write the interpolation equation as follows:

$$y = c + (slope)\bar{x} \tag{13}$$

where c is the intercept of the interpolated:

$$c = \frac{1}{K}(\sum_{i=1}^{i=K} y_i - (slope)\sum_{i=1}^{i=K} x_i) \tag{14}$$

It is important to note here that in order to calculate finite sequence of time increments $[\Delta t_1, \Delta t_2, \dots \Delta t_T]$, the time interval T is divided into a sub-window sizes of N_w each in a dyadic sequence.

The algorithm to calculate VFD is defined in detail in [7] and [1]. Following is a summarized version:

1) Select the lowest size of fractal cover K_l following a dyadic sequence.
2) Select the largest of iteration level K_m following dyadic sequence and ensure that it should not fall below minimum cover count of 30. This is required to maintain the statistical validity of the computation. In Box counting approach, this is equivalent of considering a minimum box count of 30.
3) Now, set total number of samples N_T.
4) Start the largest cover and calculate the variance of x and covariance of (x, y). Then, select the next count of cover following the dyadic sequence and calculate the x and y points of the log-log plot.
5) Find the slope using equation 10. If the slope is greater than positive 2, decrease the window size in a dyadic sequence. If the slope is negative, increase the window size in a dyadic sequence.
6) If there is an increment or decrement in the window size, go back to step 2 and recalculate all values.
7) Compute variance fractal dimension using equation 8.
8) Step up the window pointer with default window size and go back to step 2.
9) Keep on calculating till the end of total data samples. If the last few data samples are not counted in the calculations, ignore them.

Non Stationary Data Series

There are 2 ways to define a stationary process; strict sense stationary and stationary in the sense of second order statistics. A strict sense stationary stochastic process is defined as a process whose joint probability distribution function does not change with time. Formally,

$$f(x_{t_1}, x_{t_2} \dots x_{t_k}) = f(x_{t_1+\tau}, x_{t_2+\tau}, \dots x_{t_k+\tau}) \quad (15)$$

where f(.) is the joint probability distribution function of k processes and τ represents an arbitrary time delay. Moreover, all the moments of a strict sense stationary process are stationary as well i.e. they do not depend on time and are constant.

A wide/weak sense stationary process or second order stationary process is defined as follows:

1. Mean of the process is constant.
2. Autocovariance function is only dependent on the arbitrary time delay τ only.

If none or any of the above properties are satisfied, then the process is called as a non-stationary process.

Detection of Change in the Stationary Characteristics

Natural data time series does not exhibit stationary characteristics. Internet data traffic is highly non stationary. In order to apply meaningful analysis, it is required to apply stationary tools in a sliding window fashion where window slides when there is a detection of the statistical variations. Variance based stationary change detection models provide estimates of the change in variance using least square Euclidean measures. There are 2 methods of detecting a changes; (1) a-posterior detection when the data is collected offline, and (2) online detection [20].

Literature on detecting changes in the statistical characteristics of a time series categorizes change detection over offline data (complete batch of the data) and online data (small segments of the data) [20]. In this work online data is considered to detect change in variance as an indicator of change in stationarity characteristics. For this work, we choose a variance based online change detection model as described in [20], The method is based on minimizing the following cost function:

$$J(\tau, \theta) = \frac{1}{n}\sum_{i=1}^{i=n} \frac{\|x_i - \bar{x}\|^2}{var(x_i)} + n_i \log[var(x_i)] \quad (16)$$

where:

1. J(.) is the cost function
2. x_i are the samples of the stochastic process x
3. $\bar{x}$ is the first moment/mean of x
4. $var(x_i)$ is the second moment or variance of x
5. $\theta = var(x_i)$
6. n is the data window size

By minimizing the above cost function, the value of θ obtained represents the change in variance in the data series.

The main advantage of using this model is that it is detecting changes in the second order statistical sense and also provides estimated number of changes in a window of samples. Moreover, it is works online data set without requiring a batch of data samples completely.

Wavelet Transformation

Wavelets represent a set of nonlinear basis. Wavelet transformation projects the function over the nonlinear basis function according the time-frequency features of the function being projected. Hence, unlike other linear transformations, where linear basis are used for every input function, wavelets provide an adaptive set of basis functions to represent the input function in terms of its features and scale and time-frequency characteristics.

Wavelet transformation utilizes the concept of multiscale transformation [22] to compute the wavelet coefficients. Let

$$H_{1,k} = \sum_{l=0}^{l=N-1} h_{1,l}\, e^{-\frac{i2\pi lk}{N}} \tag{17}$$

where k = 0, 1, … , N-1 and $H_{1,k}$ is called the discrete transform of h_1.

Let following is the zero padded scaling filter coefficients:

$$g_1 = (g_{1,0}, \dots g_{1,L-1-l}, 0, \dots 0)^T \tag{18}$$

1. $l = 0,\dots,$L-1
2. $g_{1,l} = (-1)^{l+1} h_{1,L-1-l}$

Let $G_{1,k}$denotes its Discrete Fourier transform (DFT).

Now we can define the length N wavelet filter h_j for scale τ_j=2^{j-1} as the inverse DFT:

$$H_{1,k} = H_{1,2^{j-1}kmodN} \prod_{l=0}^{j-2} G_{1,2^{l}kmodN} \tag{19}$$

where k = 0,…,L-1.

For a sampled time series x(t):

$$\mathrm{X} = (x_0, x_1, \dots x_{N-1})^T \tag{20}$$

Following could be coefficients of the discrete wavelet filter based on any wavelet function:

$$W_{j,t} = \frac{1}{2^{j/2}} \phi_{j,2^j(t+1)-1} \tag{21}$$

where

$$\phi_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{l=L_j-1} h_{j,l} X_{t-l} \tag{22}$$

$$t = L_j - 1, \dots, N-1$$

$\phi_{j,t}$ coefficients represent the multiscalar changes of length τ_j and can be computed by subsampling/subscaling every 2^j point of the coefficient $\phi_{j,t}$.

III. Data Set

The data used was the PREDICT ID USC-Lander/ DoS_DNS_amplification-20130617 (2013-06-17) to (2013-06-17) [23]. 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 since it used 6 DNS servers to generate a botnet network.

IV. ALGORITHM

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. Minimum data window size: window_lag.
    b. Maximum data window size: window_max.

```
        c. Data pointer: dp.
    2) Initialize window_lag = 128.
    3) Initialize dp.
    4) Pass the window through a Haar wavelet based
       variance change detection function.
        a. If there is a change detected, vary the
           window size accordingly and jump back to
           step 3.
        b. If there is change within 95% confidence
           interval, jump to step 5.
    5) Pass the window through a VFD calculation
       function
    6) Test if the calculated VFD is within the range of
       dimension 1 and 2.
        a. If No, lag=lag+16 and go back to step 5.
        b. If Yes, lag=128 and go to step 7.
    7) Increment dp=dp+1 and go to step 6.
    8) Pass the time series through an interval dependent,
       Haar based wavelet filter that denoise the data and
       extract the trend of the series.
```

It is noted that default lag size of 128 and increment of 16 is selected after empirical measurements on the data set.

Variance Fractal Dimension Trajectory (VFDT) Computation

```
    1) Set the maximum and minimum scaling levels.
    2) Calculate the step size at each level according a
       dyadic criterion.
    3) Loop through each level and
        a. Calculate points on the log-log plot using
           equation 9.
        b. Calculate slope using equation 11.
```

Post Processing of VFDT

```
    1) Process the VFDT through another wavelet based
       denoising using Haar wavelets to extract the mean
       trend of VFDT.
```

V. Experimental Results

In this work, we characterized the variance fractal dimension trajectory (VFDT) of a non-stationary internet data series that contains DNS DDoS amplification attacks. PREDICT data set [23] 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. It contained 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 that was captured by the network hosts.

An adaptive window size was chosen between 128 time samples and 1024 time samples. The adaptive window was set if the data samples show a fixed variance within a confidence interval of 95%. The minimum window range of 128 was chosen so that we can have at least 2 points on the log-log plot for the calculation of variance fractal dimension. The maximum window size of 1024 was chosen so that the change in variance due to DNS attack was detected. This corresponds to the detection of DNS attack in burst of 1024 packets.

As can be seen from the data series in Fig. 1, the data set shows many spikes and lot of variations. The attack started at time step 729 and ended at time unit 9068. This time window corresponds to an interval of approximate 13 minutes. The additional 3 minutes are added in order to count every packet in the 100ms interval.

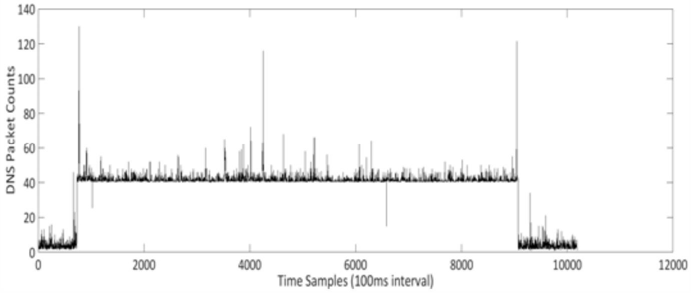

Fig. 1 Time series plot of DNS counts.

Fig. 2 represents the variance fractal dimension trajectory of the DNS count series. The computation is performed using an adaptive window based on wavelet based change detection algorithm. Moreover, the adaptive window slides with a count of 1. This is done in order to computer the long range correlation effect on the current window samples. Moreover, it is also observed that the VFDT calculation falls below the valid range of dimension 1. This happens due to the availability of outliers in the data series which rotates the regression of log-log plot from a positive slope line to the negative one. Therefore, the log-log plot shows negative slope and thus VFD calculation falls below the invalid topological dimension of a line.

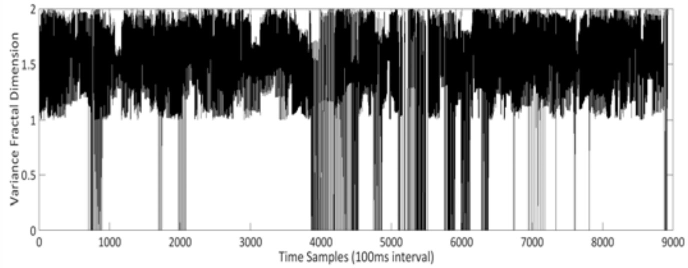

Fig. 2 Variance fractal dimension trajectory.

Also, it is observed that the variance fractal dimension trajectory (VFDT) is showing very fast variations within valid topological dimension range of [1,2]. This is due to the high correlation effect introduced by the sliding window with a lag of 1 sample.

Fig. 3 shows the trend of DNS counts series. Trending is performed after removing the noise from the plot using wavelet based denoising process. It clearly shows that the major points of changes in variance are the bursty spikes.

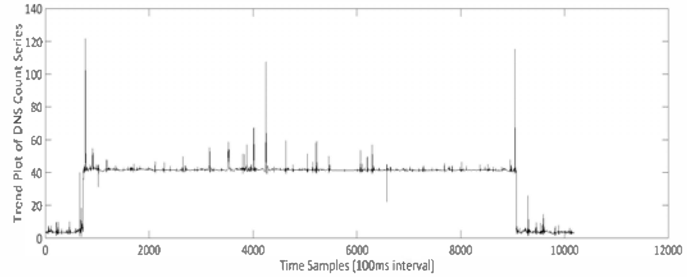

Fig. 3 Trend plot of DNS counts series

Fig. 4 shows the trend plot of variance fractal dimension trajectory. As can be seen, high varying spikes within the vicinity of sample 729 and sample 9068 are clearly shown. Moreover, within the range of 729 and 9068, there are a few such high varying occurrences. However, we can also observe that the trend is showing invalid fractal dimensions which is attributed to the negative slopes of the log-log plot and multifractality of the time series. As Fig 2 did exhibit lot of high frequency variations while Fig. 4 shows the hidden trend in the VFDT, we can deduce that the VFDT calculation window is effected with multiple changes in variances within a window size due to the presence of outliers.

Additionally, the trend analysis of VFD reveals that there is a very long memory effect in the time series which is adding noise in the calculations of VFD. More precisely, the long memory effect is generating outliers and negative slope of the log-log plot where at low scale, very high variances are observed and at large scales, variances are reduced. Also, the slope of log-log plot should remain within 0 and positive 2. If there is a high change in variance as represented by slope greater than 2, the variance fractal dimension will not be within the embedding dimension and the results are meaningless.

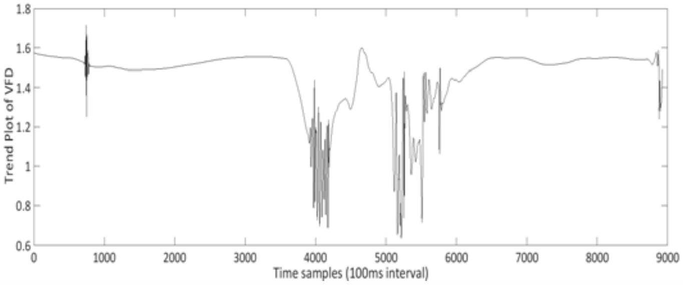

Fig. 4 Trend plot of VFD.

VI. Discussion

Variance fractal dimension trajectory is a methodology that estimates/extracts complexity of a data series in a sliding window fashion. If the time series is multifractal in nature, then over a stationary window size, a single variance fractal dimension can be calculated. However, if the time series is non stationary then over a window size having second order stationarity, it is required to check if the log-log plot is showing positive slope within the range of its embedding dimension. If not, then it may occur due to the presence of outliers (multiple high or low variances over the log-log plot) or the presence of saturation points. Through our preliminary analysis over a non-stationary DNS data series along with DDoS attacks, we show that variance fractal dimension calculation is very sensitive to the outliers and saturation points. However, the trend analysis of the VFDT using multiscale wavelet analysis shows a pattern or trend of the variance fractal dimension trajectory. Moreover, it shows that the attack can be detected and a long memory effect can be identified due to the presence of highly variable traffic characteristics. In this work, we have characterized a DNS time series to extract the changing characteristics of a non-stationary time series. In the future work, we will develop a detection model for attacks in a network time series to measure the effectiveness of the detection with the available detection methodologies.

VII. Conclusions

This paper presents a cognitive multifractal algorithm based on variance fractal dimension to detect changing trends from a non-stationary data series. The algorithm computes the variance fractal dimension and then performs detrending to extract the trend patterns of the variance fractal dimension trajectory (VFDT). The algorithm is able to identify the start and end time of attack in a DNS packet time series.

ACKNOWLEDGEMENTS

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

REFERENCES

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

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

[3] K. Falconer, Fractal Geometry: Mathematical Foundations and Applications, Wiley, New York, 1990.

[4] C. Sevcik, "A procedure to estimate the fracal dimension of waveforms," *Complexity International,* vol. 5, 1998.

[5] C. Peng, S. Havlin, H. Stanely and A. Goldberger, "Quantification of scaling exponents and crossover phenomenon in non stationary heart beat times eries," *Chaos,* vol. 5, 1995.

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

[7] W. Kinsner and W. Grieder, "Amplification of Signal Features Using Variance Fractal Dimension Trajectory," *International Journal of Cognitive Informatics and Natural Intelligence,* vol. 4, no. 4, pp. 1-17, Oct. 2010.

[8] R. W. Dahlhaus, "Empirical spectral processes for locally stationary time series," *Journal of Bernoulli Society for Mathematical Statistics and Probability,* vol. 15, 2009.

[9] M. D., N. James, W. Nicholson and L. Segalini, "Locally Stationary Vector Processes and Adaptive Multivariate Modeling," in *ICASSP*, 2013.

[10] A. Cardinali and G. Nason, "Costationarity of Locally Stationary Time Series," in *Journal of Time Series Econometrics*, 2010.

[11] H. Kantz and T. Schreiber, Non Linear Time Series Analysis, Cambridge University Press, 2004.

[12] M. Khan, K. Ferens and W. Kinsner, "A Chaotic Complexity Measure for Cognitive Machine Classification of Cyber-attacks on Computer Networks," *IJSSCI,* 2014.

[13] M. Khan, K. Ferens and W. 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*, Southbank University, London, UK, 2014.

[14] Y. Wang, D. Zhang and W. Kinsner, Advances in Cognitive Informatics and Cognitive Computing, vol. SCI 323, Berlin: Springer Verlag, 2010, pp. 265-295.

[15] V. Cheung, K. Cannons, W. Kinsner and J. Pear, "Signal classification through multifractal analysis and complex domain neural networks," in *IEEE Canadian Conference on Electrical and Computer Engineerin*, Montreal, Canada, 2003.

[16] P. Embrechts and M. Maejima, Selfsimilar Processes, Princeton University Press, 2002.

[17] P. Zhang, "Fractal dimension estimation of fractional Brownian motion," in *IEEE proceedings of Southeastcon*, 1990.

[18] A. Papoulis and S. Pillai, Probability, Random Variables and Stochastic Processes, 2 ed., McGraw-Hill, 1984.

[19] A. Oppenheim, R. Schafer and J. Buck, Discrete-Time Signal Processing, 2 ed., Prentice Hall, Feb. 1999.

[20] M. Lavielle, "Detection of multiple changes in a sequence of dependent variables," *Stochastic Processes and their Applications,* vol. 83, no. 79, 1999.

[21] M. Messer, M. Kirchner, J. Schiemann, J. Roeper, R. Neininger and G. Schneider, "A multiple filter test for the detection of rate changes in renewal processes with varying variance," *Annals of Applied Statistics,* vol. 8, no. 4, pp. 2027-2067, 2014.

[22] B. Whitcher, S. Byers, P. Guttorp and D. Percival, "Testing for homogeneity of variance in time series: Long memory, wavelets, and the Nile River," *Water Resources Research,* vol. 38, no. 5, 2002.

[23] D. PREDICT USC-Lander, "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.

[24] Y. Wang, "On cognitive informatics," in *Proc. 1st IEEE Intern. Conf. Cognitive Informatics*, Calgary, 2002.

[25] P. Winter, H. Lamperberger, M. Zelinger and E. Hermann, "On Detecting Abrupt Changes in Network Entropy Time Series," *Communication and Multimedia Security lecture Notes in Computer Science,* vol. 7025, pp. 194-205, 2011.

[26] S. Xiaonan Wu and W. Banzhaf, "The use of computational intelligence in intrusion detection systems: A review," *Applied Soft Computing,* vol. 10, no. 1, pp. 1-35, 2010.

[27] M. Rubio, A. Dougherty and J. Gollub, "Characterization of Irregular Interfaces: Roughness and Self Affine Fractals," *Measures of Complexity and Chaos, NATO ASI Series,* vol. 208, pp. 461-464, 1989.

[28] T. Vafeiadis, A. Papankikolau, C. Ilioudis and S. Charchalakis, "Real time network data analysis using time series models," in *Simulation Modelling Practice and Theory*, 2012.

[29] R. Taylor, "Attractors: Nonstrange to Chaotic," 2005. [Online]. Available: https://www.siam.org/students/siuro/vol4/S01079.pdf. [Accessed 30 May 2014].

[30] D. Shaw and W. Kinsner, "Multifractal Modelling of Radio Transmitter Transients for Classification," in *WESCANEX 97: IEEE Communications, Power and Computing. Conference Proceedings*, May 1997.

[31] J. Minor, "Attractor," Scholarpedia.

[32] M. Mahmoud, M. Dessouky and e. al., "Comparison between Haar and Daubechies Wavelet Transformions on FPGA Technology," in *Proceedings of World Academy of Science, Engineering and Technology*, 2007.