2012 International Conference on Signal Processing Systems (ICSPS 2012)
IPCSIT vol. 58 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V58.8

Self-adaptive Detection Method for DDoS Attack Based on Fractional Fourier Transform and Self-similarity

Chen Shi-wen [+], Wu Jiang-xing , Guo Tong and Lan Ju-long

National Digital Switching System Engineering and Technological R&D Center,
Zhengzhou 450002, China

Abstract. Hurst parameter estimation accuracy is very important in DDoS attack detection model based on traffic self-similarity changes. If the estimation accuracy is poor, the undetected and false detection should happen. This paper proposed the fractional Fourier transform method to estimate Hurst parameter. Combined with the idea of regression diagnosis and change point analysis, an effective, adaptive estimator of Hurst parameter based on weighted least squares was presented. We used FGN data set to verify our idea. The experimental results show that our method has high estimation accuracy, and the ability to select the optimal scale interval adaptively than the existing common method. So the method can improve the detection performance of DdoS attacks effectively.

Keywords: DdoS attack detecting; fractional fourier transform; hurst parameter; self-similarity; weighted least squares

1. Introduction

Distributed denial-of-service attack (DdoS) [1], [2] is an immense threat of the Internet security. Researchers have shown that network traffic has self-similarity and correlation [3], [4]. DDoS attacks would affect the self-similarity of network traffic and change the Hurst parameter value which characterizes the burst nature of the network traffic. Typically, the network traffic Hurst parameter is between 0.5 to 1, the greater the Hurst parameter is, the self-similar (long correlation) of the network is much higher, and the burst is stronger. When DDoS attack occurs, the attack packets reduce the self-similarity of the network and cause Hurst parameter lower. The traffic tends to Poisson distribution if it blocked completely with the Hurst parameter value become 0.5. Typically, many defense mechanisms have been proposed to combat the problem, such as Variance-Time(V-T), Rescaled Range(R/S), Periodogram methods, Whittle methods, wavelet analysis and so on[5],[6].

In recent years, the fractional Fourier transform (FrFT) [7] has been widely applied in the field of signal processing and communications technology with its time-frequency rotation characteristics. In 1996, Ozaktas et al. proposed a discrete algorithm with low computation [8]. Fractional Fourier transform attracted the attention of scholars in the field of signal processing [9-12]. YangQuan Chen et al. analyzed the self-similarity of the network Based on FrFT [10]. This paper proposed a DDoS attack self-adaptive detecting method based on FrFT. The experimental results reveal that our method is reliable and accurate, and can select the optimal scale interval adaptively than the existing common method.

The paper is organized as follows. In Section 2, we derive the definition of the fractional Fourier transform and Hurst parameter estimation based on FrFT. Section 3 presents the self-adaptive detection method in FrFT domain. Section 4 presents some of the experiment to illustrate the benefits of the new methods. Conclusions are drawn in Section 5.

2. FrFT and Hurst Parameter Estimate

[+] Corresponding author. Tel.: +86-13663008915; fax: +86-371-81630754.
E-mail address: chenshiwen2010@gmail.com

The fractional Fourier transform is the promotion of the classical Fourier transform. The fractional Fourier domain can be understood as a unified time-frequency transform domain [7].

The definition of FrFT[7] of a signal $x(t)$ is $X_a(u) = F_a(u) = \int_{-\infty}^{+\infty} x(t)K_a(t,u)dt$,Where,

$K_a(t,u) = \sqrt{(1 - i\cot\alpha)}e^{i\pi(t^2\cot\alpha - 2ut\csc\alpha + u^2\cot\alpha)}$ ($\alpha \neq n\pi$), $K_a(t,u) = \delta(t-u)$ ($\alpha = (2n \pm 1)\pi$)

$n \in \mathbb{Z}$, $\alpha = a\pi/2$, a is the order of the fractional Fourier transform. When a = 1, FrFT is the standard Fourier transform. FrFT is abbreviated as Fa. Its computational complexity is *O(NLogN)* [9,12].

Using discrete wavelet transform and multi-resolution analysis method [13, 14], we can get the Hurst parameter estimating equation based on FrFT as

$$G(j) \leftrightarrow (2H+1)j + cons\tan t$$

3. Self-Adaptive Parameter Estimation

Literature [10] use local analysis method in the fractional Fourier domain based on one-dimensional weighted linear regression. We use one-dimensional weighted estimation method of least-squares regression just like literature [15].

As for random variables $Y_j, x_j, j = 1,\ldots,J$, Given regression model $\tilde{Y}_j = \beta_0 + \beta_1\tilde{x}_j + \tilde{\varepsilon}_j$, where $E\left(\tilde{\varepsilon}_j\right) = 0, Var\left(\tilde{\varepsilon}_j\right) = \sigma^2 / N_j$. In matrix form, this equation can be written as

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}} \tag{1}$$

where $\tilde{\boldsymbol{X}}$ is a $J \times 2$ Matrix. Let $W = diag\{w_1,\ldots,w_J\}$, where $w_j = N_j$, Let the following variables: $\boldsymbol{y} = \boldsymbol{W}^{1/2}\tilde{\boldsymbol{y}}$, $\boldsymbol{X} = \boldsymbol{W}^{1/2}\tilde{\boldsymbol{X}}$, $\boldsymbol{\varepsilon} = \boldsymbol{W}^{1/2}\tilde{\boldsymbol{\varepsilon}}$, we get:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

Where $\varepsilon \sim \left(0, \sigma^2 I\right)$. Thus, the weighted least-squares estimation model (1) is equivalent to the ordinary least-squares estimation model (2). Let $\boldsymbol{b}$ as least-squares estimation of $\boldsymbol{\beta}$ in Eq. (2). We get unbiased estimate of $\boldsymbol{b}$ as:

$$\hat{\boldsymbol{b}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y} = \boldsymbol{H}\boldsymbol{y} \tag{3}$$

Let e_j as j Residuals, we get

$$\hat{e}_j = Y_j - \hat{Y}_j = Y_j - x_j^T\,\hat{\boldsymbol{b}} \tag{4}$$

where $\hat{Y}_j$ is the regression value of Y_j .

The choice of the scale interval $[j_1, j_2]$ has a great influence on the fitting result, we use variance analysis method based on the ideas of change point analysis [16], [17], using different scales range to estimate Hurst parameter fitting degree, thus adaptively selects the optimal range of scales.

Theorem 1. x_j represents a j row vectors of X , $\boldsymbol{b}(\boldsymbol{j})$ denotes the least squares estimation of $\boldsymbol{\beta}$ in formula (2) after row j is removed. The relationship between $\boldsymbol{b}$ and $\boldsymbol{b}(\boldsymbol{j})$ are summarized as:

$$\boldsymbol{b}-\boldsymbol{b}(\boldsymbol{j})=\frac{\left(X^{T} X\right)^{-1} x_{j} e_{j}}{1-h_{jj}} \tag{5}$$

It is used as a regression diagnostic tool for the detection variable j , thus we can judge whether it affect the estimated results of the regression model.

Theorem 2. Define the residual sum of squares: $SSR=\sum_{i}\left(Y_{i}-\hat{Y}_{i}\right)^{2}$. Let s^2 as the general estimation of σ^2 , $s^2(j)$ is the estimation of σ^2 in formula (2) after row j is removed. And, define $s^{2}=\frac{SSR}{n-1}=\frac{1}{n-1} \sum_{i=1}^{J}\left(Y_{i}-x_{i}^{T} b\right)^{2}$. Analogously, define $s^{2}(j)=\frac{1}{(n-1)-1} \sum_{i \neq j}^{J}\left\{Y_{i}-x_{i}^{T} b(j)\right\}^{2}$, so the relationship between s^2 and $s^2(j)$ meets:

$$(n-2) s^{2}(j)=(n-1) s^{2}-\frac{e_{j}^{2}}{1-h_{jj}} \tag{6}$$

Theorem 1 and Theorem 2 compose the foundation to examine the statistics when choosing the region of the scale. We may regard the problem as choosing a child model that presents lots of linearity among the regression models. Therefore, with regard to each child model whose index label is j , the problem can turn to checking up the hypothesis testing

$$H_0(j): EY=cons\tan t, \qquad H_1(j): EY=linear.$$

Assuming $SSR_j(old)$ represents the residual squares sum of null hypothesis, and $SSR_j(new)$ represents the residual squares sum of the alternative hypothesis, we may carry through an optimal examination using the statistical parameter $T(j)=\frac{\left(SSR_{j}(old)-SSR_{j}(new)\right)/1}{SSR_{j}(new)/\left(N(j)-2\right)}=\frac{\sum_{i=j}^{J}\left(\bar{Y}-\hat{Y}_{i}\right)^{2}}{\sum_{i=j}^{J}\left(Y_{i}-\hat{Y}_{i}\right)^{2}/\left(N(j)-2\right)} \sim F\left(1, N(j)-2\right)$.

Thus, we can get the optimal region of the scale $[j_1, j_2]$, where $j_1=\arg\min_{j \geq 1}\{T(j)\}$, $j_2=J$.

4. Experimental Results and Analysis

We use FGN sequence as the original data to estimate Hurst parameter. With sequence generation algorithm based on FFT from Stilian Stoev[18], we get FGN sample sequences with different Hurst index values from 0.55 to 0.95,interval of 0.05. The length is N = 216 = 65536.

4.1. Optimal transform order selection

In order to obtain the optimal order of the transform, we use different FrFT order to estimate the Hurst value of the FGN sequence, and calculate each estimated deviation compared to the actual value of the statistical variance, then select a as the most optimal FrFT order where the mean of statistical variance is least. Figure 1 shows the variances of Hurst parameter estimations with different FrFT orders for FGN. Let

$\hat{H}_i$ as the estimated Hurst value, H_i as the real Hurst value, and $\widehat{a^*} = \arg\min_{i \geq 1} \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\hat{H}_i - H_i}{H_i} \right)^2 \right\}$, the statistical variance is the least when $\widehat{a^*} = 0.6$. So, we can select 0.6 as the optimal transform order.

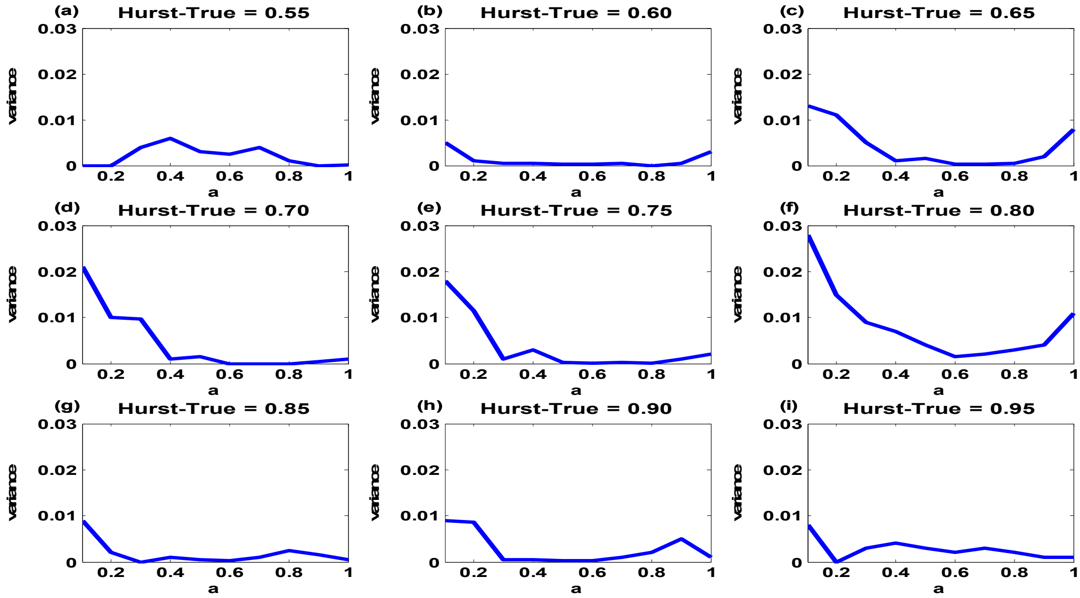

Fig. 1: The variances of Hurst parameter estimations with different orders of FrFT for FGN (a)H=0.55, (b) H=0.6, (c) H=0.65, (d) H=0.7, (e) H=0.75, (f) H=0.8, (g) H=0.85, (h) H=0.9 and (i) H=0.95

4.2. Adaptive selection of the scale interval

Table 1 and Table 2 show the estimated Hurst parameter values and corresponding goodness-of-fit value R of the FGN sample sequence with H=0.55 and 0.95, respectively. That with "__" and bold logothe interval is our adaptive selected result. We can see, different scale interval has a great influence on the accuracy of the estimated Hurst parameter, the optimal scaling interval corresponding to the H estimate closest to the true value of H, and the corresponding goodness-of-fit value R minimum. It shows a linear relationship between the random variable regression models to achieve the best in the optimal scale interval, so the scale interval selection method is in line with the principle of optimality.

Table 1: Hurst parameter estimation and goodness of fit for different scale interval(H=0.55)

[j1,j2]	[1,16]	[2,16]	[3,16]	[4,16]	[5,16]	[6,16]	[7,16]
H	0.529	0.597	0.513	0.519	0.613	0.605	0.552
R	0.00137	0.00126	0.00098	0.00331	0.00015	0.00026	0.00009

Table 2: Hurst parameter estimation and goodness of fit for different scale interval(H=0.95)

[j1,j2]	[1,16]	[2,16]	[3,16]	[4,16]	[5,16]	[6,16]	[7,16]
H	0.961	0.973	0.984	0.958	0.953	0.967	0.935
R	0.01351	0.03786	0.00926	0.00508	0.00002	0.00496	0.00665

4.3. Accuracy analysis

Table 3 shows the Hurst parameter estimation results of the FGN sequence by the commonly used estimation method, Literature [10] method and the adaptive method proposed in this paper. It also gives the optimal scale interval [J1, J2] and the deviation between the actual H value and the estimated value. Because sample sequence length N = 216 = 65536, so the largest scale in the adaptive method order is 16. As can be seen from the table 3, in FrFT adaptive method, and in the range of optimal scaling, the estimated precision of the results is better than the other methods.

Table 3: Hurst parameter estimation with various methods

H	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
V-T	0.481	0.540	0.596	0.715	0.767	0.808	0.837	0.844	0.923
R/S	0.556	0.614	0.649	0.698	0.729	0.811	0.844	0.897	0.907
Wavelet	0.566	0.607	0.664	0.695	0.756	0.806	0.878	0.942	0.984
FrFT	0.584	0.628	0.671	0.682	0.732	0.809	0.835	0.873	0.939
FrFT Adaptive	0.552	0.604	0.652	0.698	0.750	0.805	0.847	0.895	0.953
[j1,j2]	[7,16]	[7,16]	[5,16]	[6,16]	[6,16]	[5,16]	[7,16]	[7,16]	[5,16]
|Htrue-Hestimate|	0.002	0.004	0.002	0.002	0.000	0.005	0.003	0.005	0.003

5. Summary

This paper introduced the fractional Fourier transform to DDoS attack detection, it make use of the high precision of the Hurst parameter estimation in fractional Fourier transform domain. Combined with the idea of regression diagnosis and change point analysis, the weighted least-squares regression model in fractional Fourier domain was presented. The experimental results show that the method has high estimation accuracy, and the ability to select the optimal scale interval adaptively than the existing common method. So it can effectively reduce the DDoS attacks detection of false negative and false positive rate. We will focus on the fast FrFT domain detection algorithm for DDoS attack in future studies.

6. References

[1] B. Hakem and D. Geert. Analyzing well-known countermeasures against distributed denial of service attacks[J]. *Computer Communications*, 2012, 35(11): 1312-1332.

[2] D. Christos and M. Aikaterini, DDoS attacks and defense mechanisms: classification and state-of-the-art[J]. *Computer Networks*, 2004, 44(5): 643-666.

[3] H. Gupta, V. J. Ribeiro, and A. Mahanti. A Longitudinal Study of Small-Time Scaling Behavior of Internet Traffic [J]. NETWORKING 2010, LNCS 6091: 83-95.

[4] W. E. Leland, M. S. Taqqu, and W. Willinger, et al. On the self-similar nature of Ethernet traffic[J]. ACM SIGCOMM Computer Communication Review, 1993, 23(4):183-193.

[5] A. BĂrulescu, C. Serban, and C. Maftel. Evaluation of Hurst exponent for precipitation time series[J]. LATEST TRENDS on COMPUTERS, 2010, 2: 590-595.

[6] W. Rea, L. Oxley, M. Reale, et al. Estimators for Long Range Dependence: An Empirical Study[J]. *Electronic Journal of Statistics*, 2009, vol.0: 1-16.

[7] L. B. Almeida. The Fractional Fourier Transform and Time-frequency Representations[J]. *IEEE Trans. Signal Processing*, 1994, 42(11): 3084-3091.

[8] H. M. Ozaktas and O. Arikan, et al. Digital Computation of the Fractional Fourier Transform[J]. *IEEE Trans. Signal Processing*, 1996, 44(9): 2141-2150.

[9] A. Bultheel and H. E. M. Sulbaran. Computation of the Fractional Fourier Transform[J]. *Applied and Computational Harmonic Analysis*, 2004, 16(3): 182-202.

[10] Y. Chen, R. Sun, A. Zhou. An improved Hurst parameter estimator based on fractional Fourier transform[J]. *Telecommun System*, 2010, 43: 197-206.

[11] C. Ciflikli and A. Gezer. Self similarity analysis via fractional Fourier transform [J]. *Simulation Modeling Practice and Theory*, 2011, 19: 986-995.

[12] R. G. Campos, J. R. Melgoza, E. Chávez. A New Formulation of the Fast Fractional Fourier Transform [J]. *SIAM Journal on Scientific Computing*, 2012, 34(2): A1110-A1125.

[13] P. Ably, P. Flandrin, M. S. Taqqu, et al. Self-Similarity and long-range dependence through the wavelet lens [J]. In: *Theory and Applications of Long Range Dependence*, Boston: Birkhauser Press, 2002: 345-379.

[14] J. Park and C. Park, Robust estimation of the Hurst parameter and selection of an onset scaling [J]. Statistica Sinica, 2009, 19 (4): 1531-1555.

[15] P. Tarrio, A. M. Bernardos, J. R. Casar. Weighted Least Squares Techniques for Improved Received Signal Strength Based Localization[J]. Sensors 2011, 11: 8569-8592.

[16] K. C. Chang, C. L. Chiang, and C. B. Lee. The Comparison of Algorithms in Change-Points Problem[J]. *Journal of Applied Science and Engineering*, 2012, 15(1): 11-19.

[17] P. R. Bertrand, M. Fhima, and A. Guillin. Fast change point analysis on the Hurst index of piecewise fractional Brownian motion[J]. *Journée de Statistiques (JDS 2011)*, Tunis 2011: 1-6.

[18] Stilian Stoev's new web page. http://www.stat.lsa.umich.edu/~sstoev.

Shiwen Chen is an engineer in China National Digital Switching System Engineering and Technological R&D Center. He was born in 1974, and received his master degree in Signal Processing from Zhengzhou Information and Science & Technology Institute, China, in 2004. He is mainly working on Broad-Band information network, signal processing and embedded system design.

Jiangxing Wu is a professor of China National Digital Switching System Engineering and Technological R&D Center. His intresting fields include Telecommunications switching, Broad-Band information network.

Tong Guo is a PhD student of China National Digital Switching System Engineering and Technological R&D Center. His intresting fields include Network Security, Routing Technology and etc.

Julong Lan is a professor of China National Digital Switching System Engineering and Technological R&D Center. His intresting fields include Broad-Band information network, Routing technology and etc.