2017 3rd International Conference on Electronic Information Technology and Intellectualization (ICEITI 2017)
ISBN: 978-1-60595-512-4

A Haar Transform-Based Detection Approach to Network Traffic Anomalies in Power Telecommunication Access Networks

Fanbo Meng, Sihang Zhao, Zhuo Di, Zhe Zhang, Liangliang Yu, Wenjing Li and Ping You

ABSTRACT

This paper proposes a new detection approach to find the abnormal parts in network traffic. Firstly, network traffic is regarded as a discrete time series. Then it is normalized and is carried out the feature component decomposed. Secondly, according to mathematical theory, the feature components in network traffic is effectively refined from the normalized series. The network traffic is divided into feature and residual components. Thirdly, the Haar time-frequency decomposition is carried out for these two components. In this case, a quick anomaly detection algorithm is presented. Simulation results show that our approach is feasible.

INTRODUCTION

Network traffic anomalies in power telecommunication access networks have impact on network performance and users' experience quality [1-2]. How to effectively and efficiently detect and diagnose abnormal and anomalous components in network traffic for power communications has become a larger challenge [3-4]. Therefore, network traffic anomaly detections are very significant in current power network operations. This has received attentions from academic and industries [5-8].

The time-frequency method [1,5,9], SOSS and FARIMA models [2], and empirical mode decompositions [3] were used to detect abnormal network traffic. The parameter-based detection method [10] and periodicity features were exploited to extract traffic anomalies [6]. The abnormal part of network traffic for multimedia

Fanbo Meng, Zhuo Di, Liangliang Yu, State Grid Liaoning Electric Power Company Limited, Shenyang 110006, China
Sihang Zhao, State Grid Huludao Electric Power Supply Company, Huludao 125000, China
Zhe Zhang, Wenjing Li, State Grid Information & Telecommunication Group Co., Ltd., Beijing 102211, China
Ping You, State Grid Info-Telecom Great Power Science and Technology Co., Ltd, Fuzhou 311003, China

applications [7] was also studied. The flow template [11] and the spectral kurtosis analysis was proposed to recognize and diagnose abnormal parts in network traffic [12]. Dynamic anomaly detection approaches [4], information theory [13] and compressive sensing [14] could be used to characterize network traffic.

This paper proposes a new quick detection approach to find out the anomaly components in traffic for power telecommunication access networks, which combines the feature component analysis with the Haar time-frequency decomposition method. Firstly, we regard network traffic as a time series of signals, which is used to construct a normalized time series. Secondly, the feature component decomposition is performed for the normalized series. Network traffic is divided into feature and residual components, in which feature components denotes the main features in network traffic while residual components describe the unimportant features in network traffic. Thirdly, we exploit the Haar decomposition to decompose these two components. Thus, we can describe the features of network traffic. Then, a quick anomaly detection algorithm is presented to perform the accurate recognition of anomalous network traffic. Simulation results show that our approach is promising.

PROBLEM STATEMENT

Network traffic in power telecommunication access networks fluctuates with time. Practically, it is a time signal. Network traffic holds obviously time-varying features. Hence, we can exploit the time signal processing method to analyze network traffic. Assume network traffic x_t at time t, and then a time series $x = \{x_t \mid t = 1, 2, ...\}$ represents any network traffic over the time. Without loss of generality, assume a network traffic $\tilde{x} = \{x_t \mid t = 1, 2, ..., n\}$ with the length by n where n is an integer. We normalize the time series $\tilde{x}$ and attain the normalized time series z. z can be converted into the following equation:

$$y_1 = a_{11}z_1 + a_{12}z_2 + \cdots + a_{1n}z_n, ..., y_m = a_{m1}z_1 + a_{m2}z_2 + \cdots + a_{mn}z_n \qquad (1)$$

So the time series z is converted to be m new variables that is linearly independent. And every alternation variable of n original variables, which is multivariable analysis, also called matrix data analysis. According to mathematic theory, the variance sum of series z is equal to new variables $y = (\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_m)$. If the first m principal components' cumulative contribution rate can run up to 90%, then new variables $y_1, y_2, \cdots, y_m$ can represent the characteristics of original variables z. To the end, we let $r_{co} = \lambda_k / \sum_{i=1}^{n} \lambda_i$ and $r_{cc} = \sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{n} \lambda_i$ denote the contribution rate and cumulative contribution rate. Then we calculate the relative matrix for z as follows:

$$R=\begin{bmatrix}1 & r_{12} & \cdots & r_{1n}\\ r_{21} & 1 & \cdots & r_{2n}\\ \vdots & \vdots & & \vdots\\ r_{n1} & r_{n2} & \cdots & 1\end{bmatrix} \quad \text{and} \quad |R-\lambda I|=\begin{bmatrix}1-\lambda & r_{12} & \cdots & r_{1n}\\ r_{21} & 1-\lambda & \cdots & r_{2n}\\ \vdots & \vdots & & \vdots\\ r_{n1} & r_{n2} & \cdots & 1-\lambda\end{bmatrix}=0 \qquad (2)$$

Where $r_{ij}=\frac{1}{n-1}\sum_{k=1}^{n}x_{ik}x_{kj}=r_{ji}$; $i,j=1,2,\cdots,n$; I denote the identity matrix. We use Equation (2) to attain the eigenvalue λ. Then, let $r_{cc}>90\%$ to decide parameter m. Hence, we attain the below equation:

$$\lambda_1\geq\lambda_2\geq\ldots\lambda_{m-1}\geq\lambda_m,\ Ra_i=\lambda_i a_i,\ i=1,2,\ldots,m, \text{ and } a_i=(a_{i1},a_{i2},\ldots,a_{in}) \qquad (3)$$

By (2)-(3), we obtain parameters a_{ij} (where $i=1,2,\ldots,m$; $j=1,2,\ldots,n$) in (1). Then, we attain $y=(y_1,y_2,\ldots,y_m)$. For discrete time series y, we carry the inverse process of normalizing transform. A new time series $u=(u_1,u_2,\ldots,u_m)$ with m principal features can be attained, which can stand for the features of network traffic $\tilde{x}$. Now we use the Harr transform $H_{Har}(\cdot)$ to further analyze u as:

$$U_n=\frac{1}{m}\sum_{k=0}^{m-1}H_{Har}(n,\frac{k}{m})u_k,\quad u_k=\sum_{n=0}^{m-1}H_{Har}(n,\frac{k}{m})U_n \qquad (4)$$

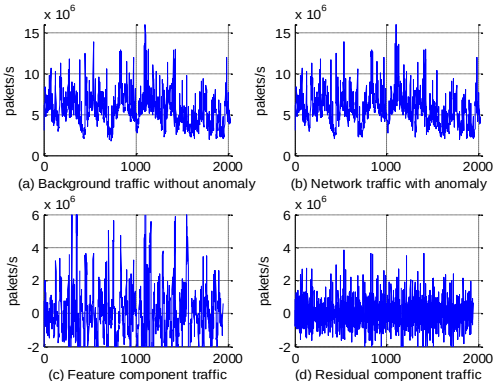

Figure 1. Network traffic and feature component decompositions.

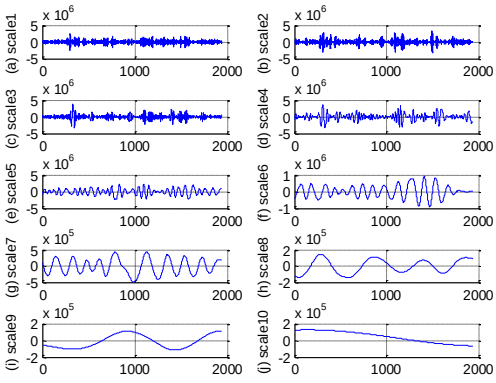

Figure 2. Haar decompositions for feature component traffic.

$$H_{Har}(2^n+k,t)=\sqrt{2^n},\ k/2^n\leq t<\frac{2k+1}{2^{n+1}};\ -\sqrt{2^n},\ \frac{2k+1}{2^{n+1}}\leq t<\frac{k+1}{2^n};\ 0,\ t<k/2^n, t\geq\frac{k+1}{2^n} \qquad (5)$$

$$U_n=\frac{1}{m}H_{Har}(n)u_k,\ u_k=H_{Har}^T(n)U_n,\ Har(n)=(h_{ij})_{(2^n-1)\times 2^n},\ h_{ij}=Har(i,j/2^n),\ i,j=0,\ldots,2^n-1 \qquad (6)$$

Then, we attain series U_n. The hidden features are refined. The 3δ method is used to label the abnormal network traffic. Our detection algorithm steps are as follows:

Step 1: Give network traffic $\tilde{x} = \{x_t \mid t = 1, 2, ..., n\}$, the number m of top principal components in network traffic.

Step 2: According to normalized transform, attain series z.

Step 3: According to Equations (2)-(3), calculate the relative matrix R about series z.

Step 4: By Equation (2), compute the eigenvalue λ about series z.

Step 5: Decide the principal components' number m.

Step 6: Through Equation (3), find the feature vectors $a_i = (a_{i1}, a_{i2}, ..., a_{in})$.

Step 7: By (1) and $a_i = (a_{i1}, a_{i2}, ..., a_{in})$, attain new time series $y = (\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_m)$.

Step 8: By the inverse normalizing transform, attain time series $u = (u_1, u_2, ..., u_m)$.

Step 9: Carry Haar transform for u, and attain Harr series U_n by (4)-(6).

Step 10: Perform the feature a extraction for U_n. Find abnormal components.

Step 11: If the detection process is over, exit and save the detection results to the file.

SIMULATION RESULT AND ANALYSIS

Now we carry out some simulations to prove our detection approach for access network traffic in power communications. Fig. 1(a) and (b) indicate that normal and abnormal traffic have no distinct difference. Fig. 1(c) and (d) state that the feature features and residual features of abnormal network traffic can be correctly extracted via our algorithm. This indicates that our algorithm is effective and promising. Fig. 2 formulates the Haar decompositions for feature component traffic. From Fig. 2, we can clearly see that the feature component traffic can be characterized by 10 time-frequency scale Haar functions accurately.

From Fig. 3, we can also clearly see that the residual component traffic can be described by 10 time-frequency scale functions. Fig. 4 shows the traffic anomaly detection results via our algorithm. This further shows that our algorithm can effectively find out the anomalous traffic.

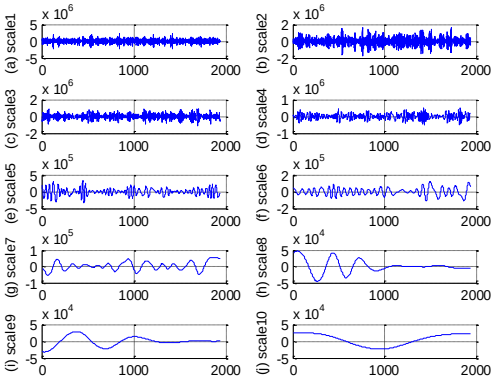

Figure 3. Haar decompositions for residual component traffic.

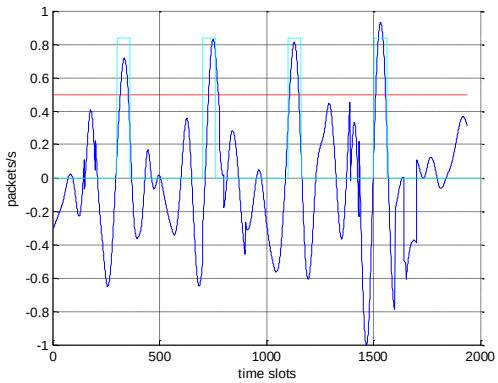

Figure 4. Anomaly detection results for access network traffic.

CONCLUSIONS

This paper proposes a new detection method to traffic for power telecommunication access network, which combines the feature component analysis with the Haar time-frequency decomposition method. The Haar decomposition is used to decompose feature components and esidual components. Simulation results show that our approach is promising.

REFERENCES

1. D. Jiang, Z. Xu, P. Zhang, et al. 2014. A transform domain-based anomaly detection approach to network-wide traffic. *Journal of Network and Computer Applications*, 40(2): 292-306.
2. B. AsSadhan, K. Zeb, J. Al-Muhtadi, et al. 2017. Anomaly detection based on LRD behavior analysis of decomposed control and data planes network traffic using SOSS and FARIMA models, *IEEE Access*, 5: 13501-13519.
3. D. Jiang, W. Li, H. Lv. 2017. An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. *Neurocomputing*, 220(2017): 160-169.
4. W. Xiong, et al. 2014. Anomaly secure detection methods by analyzing dynamic characteristics of the network traffic in cloud communications. *Information Sciences*, 258(2014): 403-415.
5. D. Jiang, C. Yao, Z. Xu, et al. 2015. Multi-scale anomaly detection for high-speed network traffic. *Transactions on Emerging Telecommunications Technologies*, 26(3): 308-317.
6. A. Kortebi, Z. Aouini, M. Juren, et al. 2016. Home networks traffic monitoring case study: Anomaly detection, in Proc. ICC'16, pp. 1-6.
7. D. Jiang, Z. Yuan, P. Zhang, et al. 2016. A traffic anomaly detection approach in communication networks for applications of multimedia medical devices. Multimedia Tools and Applications, online available, pp. 1-25.
8. G. Thatte, U. Mitra, J. Heidemann. Parametric methods for anomaly detection in aggregate traffic. IEEE Transactions on Networking, 2011, 19(2): 512-525.
9. D. Jiang, W. Qin, L. Nie, et al. 2012. Time-frequency detection algorithm of network traffic anomalies, in Proc. ICIIM'12, pp. 1-4.
10. G. Thatte, U. Mitra, J. Heidemann. 2011. Parametric methods for anomaly detection in aggregate traffic, *IEEE Transactions on Networking*, 19(2): 512-525.
11. Y. Wang, R. Jin, W. Han. 2016. An anomaly traffic detection method based on the flow template for the controlled network, in Proc. ICOCN'16, pp. 1-3.
12. D. Jiang, C. Yao, W. Zhang, et al. 2013. A detection algorithm to anomaly network traffic based on spectral kurtosis analysis, in Proc. ITSim'13, pp. 980-983.
13. P. Tune, D. Veitch. 2011. Sampling vs sketching: An information theoretic comparison, in Proc. INFOCOM, pp: 2105-2113.
14. D. Jiang, L. Nie, Z. Lv, et al. 2016. Spatio-temporal Kronecker compressive sensing for traffic matrix recovery. *IEEE Access*, 4: 3046-3053.