Anomaly Detection with Ensemble Empirical Mode Decomposition for Secure QUIC Communications: A Simple Use Case

Keyang Gu[1(✉)], Junyi Wu[1], Fan Jiang[1], Ruiwen Ji[1], Lejun Ji[1], and Tao Lei[2]

[1] Jiangxi Normal University, Nanchang 330022, China
gukeyang@jxnu.edu.cn
[2] University College London, London, England
lei.tao.21@ucl.ac.uk

Abstract. QUIC (Quick UDP Internet Connections) proposed by Google is a new secure general-purpose network transport protocol. Compared with TCP and TLS, QUIC combines the advantages of many other protocols and is a new multiplexing and secure transmission protocol. However, with the development of network technology and the gradual expansion of network scale, the network environment has become increasingly complex, and network security has become increasingly severe. QUIC network monitoring faces enormous challenges. Based on the self-similarity of QUIC traffic, an anomaly detection method for QUIC traffic based on Ensemble Empirical Mode Decomposition (EEMD) is proposed in this paper. By decomposing the network traffic, several Intrinsic Mode Functions (IMFs) and a residual trend term are obtained, and then several IMF components with low frequency and low noise are selected for reconstruction. Calculate the Hurst value of the reconstructed signal and judge whether the QUIC network has been attacked by comparing the change of the Hurst value before and after adding abnormal traffic. The simulation experiment verifies the effectiveness and accuracy of the method.

Keywords: QUIC · Ensemble Empirical Mode Decomposition · Anomaly detection · Hurst parameters

1 Introduction

Nowadays, with the continuous development of network technology and the rapid increase of Internet traffic, people are increasingly dependent on the Internet.

Y. Cao and X. Shao (Eds.): MONAMI 2022, LNICST 474, pp. 413–422, 2023.
https://doi.org/10.1007/978-3-031-32443-7_30

Therefore, researchers are exploring and improving new network protocols and seeking to develop new technologies. So Google developed the QUIC protocol (Quick UDP Internet Connections) [1] in 2013. QUIC is a protocol used at the transport layer. Unlike other traditional protocols, the UDP protocol is used instead of the TCP protocol. By transferring the reliability of TCP from the transport layer to the application layer, QUIC enables faster development and even allows multiple variants of the protocol [2].

Google published QUIC-related papers at the SIGCOMM conference in 2017, which has aroused significant repercussions in the industry. Nowadays, more than 50% of Chrome browser traffic and 75% of Facebook traffic use QUIC for transmission. But since QUIC is a new protocol, only a few studies have focused on examining its security. DeConinck et al. [3] and Viernickel et al. [4] introduced the concept of multipath into QUIC so that a QUIC connection can use multiple underlying network links simultaneously. Han Y et al. [5] proposed an enhanced BBR Congestion Control Algorithm (eBCC). Chiariotti et al. [6] combined the congestion control and multi-stream properties of the QUIC transport protocol with an appropriate scheduling algorithm to maximize the value of the information at the receiving end. Shi X et al. [7] proposed a priority-based multipath QUIC stream scheduling mechanism to avoid the blocking problem between multipath QUIC streams. Shi X et al. [8] designed an MPQUIC scheduler by prioritizing the streams that make up the critical rendering path, which can effectively speed up the first rendering time in page loading. There is a severe lack of research on QUIC traffic security.

Statistical analysis of network traffic to detect network anomalies is the focus of research in abnormal traffic detection. In recent years, Leland et al. [9] analyzed different networks and found that network traffic has statistical self-similarity. Numerous studies have shown that normal network traffic has self-similarity [10,11], and abnormal traffic will impact the network's self-similarity. By observing and summarizing network traffic behavior, Barford et al. [12] stipulated that the similarity of statistical features at the network flow level was used to classify network anomalies into three categories: network operation anomalies, network attack anomalies, and sudden congestion anomalies. Pei J et al. [13] proposed a network traffic anomaly detection method based on long-term and short-term memory network self-encoding. Pan Y et al. [14] proposed a custom user abnormal behavior detection model based on a deep neural network. The fine-grained analysis and customized user behavior management settings can be used to achieve anomaly detection in specific network environments. Hu Z et al. [15] proposed various processing methods for network traffic indicators for further evaluation of network information security.

Based on the self-similarity of QUIC traffic, this paper adopts an Ensemble Empirical Mode Decomposition (EEMD) method to detect abnormal traffic. Compared with the problems existing in the traditional network traffic anomaly detection, such as poor adaptive ability, low efficiency, and high energy consumption, this method detects whether an attack occurs by observing the change of the Hurst parameter, which has the advantages of less computation and less resource consumption.

2 Related Technology Introduction

2.1 QUIC Protocol

The QUIC protocol was first designed and proposed by Google as a network protocol applied to the transport layer. It is based on the UDP protocol drafted by the IETF working group. Figure 1 is a structural diagram of the QUIC protocol. The OSI reference architecture shows that the QUIC protocol is above the network layer and involves the transport, session, presentation, and application layers. As shown in the figure, the protocol used at the transport layer is the UDP protocol instead of the TCP protocol, precisely because the QUIC protocol is designed to bypass the TCP protocol.

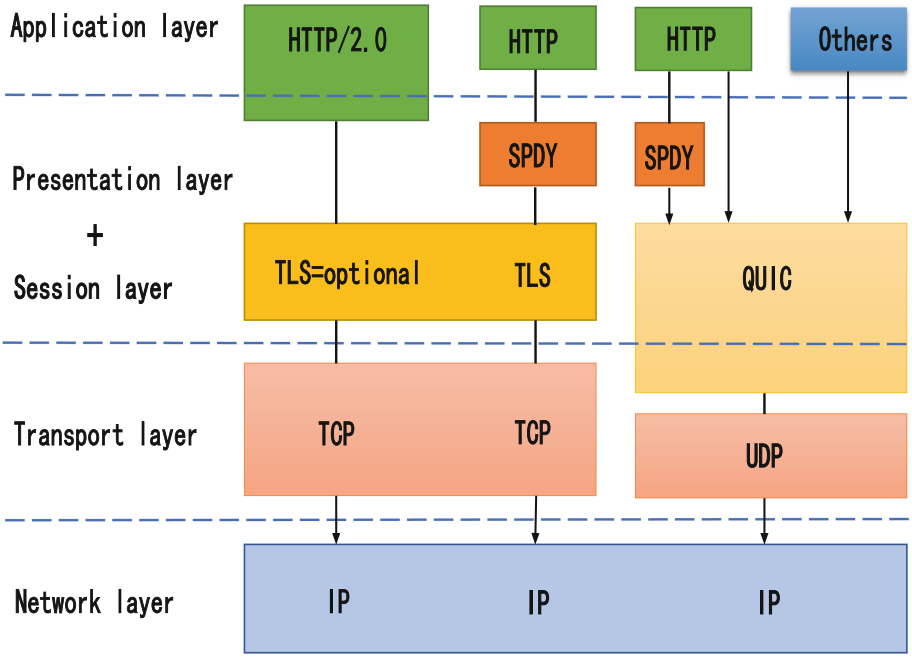

Fig. 1. QUIC protocol stack

QUIC uses the TLS1.3 method implemented in the session layer and presentation layer to encrypt effectively, which enhances the security of the protocol to a certain extent [16]. At the same time, HTTP/3 at the application layer only needs to parse the HTTP protocol so that the QUIC protocol can complete functions such as HTTP/3 multiplexing and link selection, scheduling, and management [2]. Based on the characteristics of UDP, the QUIC protocol does not require ordered packet delivery, which cleverly avoids the HOL blocking problem [17]. QUIC replaces the traditional IP quadruple by adding the connection

ID parameter to all data packets, thereby achieving a zero round-trip time (0-RTT) between the two terminals that have established a connection, reducing the connection cost [18].

At present, multipath transmission protocols such as Multi-path TCP (MPTCP) have excellent deployment resistance, and most of the research is in path management and data scheduling [19,20]. Therefore, Multi-path QUIC (MPQUIC) supporting parallel data transfer has gradually become a new research hotspot. MPQUIC can realize mobile ultra-high-definition video data transmission in 5G signal coverage holes and weak coverage scenarios, which has great practical significance for the development of the new generation of mobile ultra-high-definition video applications [21].

2.2 EEMD Algorithm

EMD is a new time-frequency analysis method proposed by Huang et al.. It has certain research prospects in complex neural networks [22–24], traffic prediction models [25–27], network security [28,29]. EMD is based on the time scale of the signal data and decomposes it into a set of orthogonal components, namely the Intrinsic Mode Function(IMF). In the EMD method, the IMF component is used as a basis function and a natural oscillation mode in the signal. Therefore, the EMD method is a nonlinear and non-stationary adaptive signal processing. However, EMD has limitations. The time scale of the acquired signal is derived from local features, and then the EMD method has a mode mixing problem. Generally speaking, mode mixing is the problem of scale confusion in the intrinsic mode mixing function. The following situations are the specific manifestations of the mode mixing problem: (1) In the same IMF component, the time scale distribution of some signals is large, but the signals based on the component are different; (2) some of the signals in different IMF components have similar scales. The mode mixing problem will lead to the loss of the original single characteristic scale of the IMF component, resulting in oscillation, mixed scales, and loss of the original physical meaning.

To optimize the EMD method, Huang et al. [30] proposed an improved Ensemble Empirical Mode Decomposition method (EEMD). A new noise-assisted analysis method realizes the decomposition and analysis of the signal by defining the intrinsic mode function as the average value of a set of experiments. Each experiment adds a finite amplitude white noise to the decomposed signal result. This method is based on the statistical analysis of white noise characteristics [31,32], which shows that adding auxiliary white noise based on the EMD method can effectively form a particular adaptive binary filter bank. In addition, Flandrin et al. [33] show that white noise is helpful for the EMD method for signal data analysis. The above studies have promoted the generation of the EEMD method to some extent.

The content of the EEMD method is to add white noise to the original signal and then decompose the original signal after adding white noise. The decomposition operation is repeated to obtain the data of multiple decompositions.

Finally, the results of these data are summarized by the formula, and the average value is obtained. The white noise added to the original signal can fill the space of the entire time frequency uniformly by using components of different time scales, thus forming a uniform white noise background. The signal is automatically projected when a new signal uses this background as a reference scale. Because each time the signal with added white noise is decomposed, it will get more detailed results, but in a single experiment, different noises are taken in each experiment, reducing sufficient trajectories in the ensemble average. The ensemble mean is the only correct answer because as more and more trials are added to the ensemble, the only persistently stable part is the signal data. Specific steps are as follows:

1) First, obtain an original signal data $A(t)$, and then add white noise $c_i(t)$ conforming to the normal distribution of the data. The formula is as follows:

$$A_i(t) = A(t) + c_i(t) \tag{1}$$

2) The generated new signal data $A_i(t)$ is decomposed by the EEMD method, thereby obtaining multiple IMF components $imf_{i,j}(t)$. i represents adding i times of white noise, and j represents the jth IMF component.
3) Repeat the operations of steps 1 and 2 of the cycle for N times. You need to add a different white noise that conforms to the normal distribution each time. Form these M IMF components into a set.

$$imf_{i,1}(t), imf_{i,2}(t), ...imf_{i,M}(t), i = 1, 2, 3...N \tag{2}$$

4) After multiple cyclic operations, multiple IMF components are obtained, and an average procedure is performed on these IMF components to get a new IMF, which is expressed as $IMF_i(t)$. The formula is as follows:

$$IMF_i(t) = \frac{1}{M} \sum_{i=1}^{M} imf_{i,j}, i = 1, 2, ...N \tag{3}$$

3 Experiment Analysis

This paper conducts network simulation experiments in the network simulator NS-3 equipped with the QUIC environment to generate the sample sequences required for the experiments. The original data obtained are decomposed and reconstructed by the EEMD method. The amplitude ratio coefficient of adding white noise is preset as 0.2, and the overall average number of times is 100.

Figure 2 shows the jitter of the normal QUIC network, and Fig. 3 shows the jitter of the QUIC network after the attack. From the comparison, it can be seen that after adding attack traffic, the QUIC network jitter fluctuates violently and obviously. The experiment also compared the jitter reconstruction graph and Hurst parameters before and after adding attack traffic. Figure 4 shows the reconstructed QUIC network jitter before and after the attack. By comparison, it

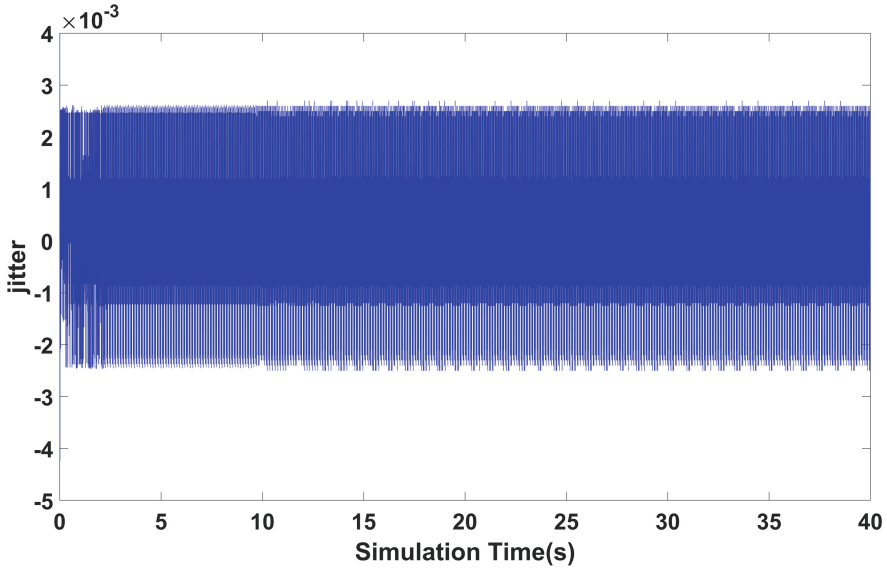

Fig. 2. Jitter in normal QUIC network

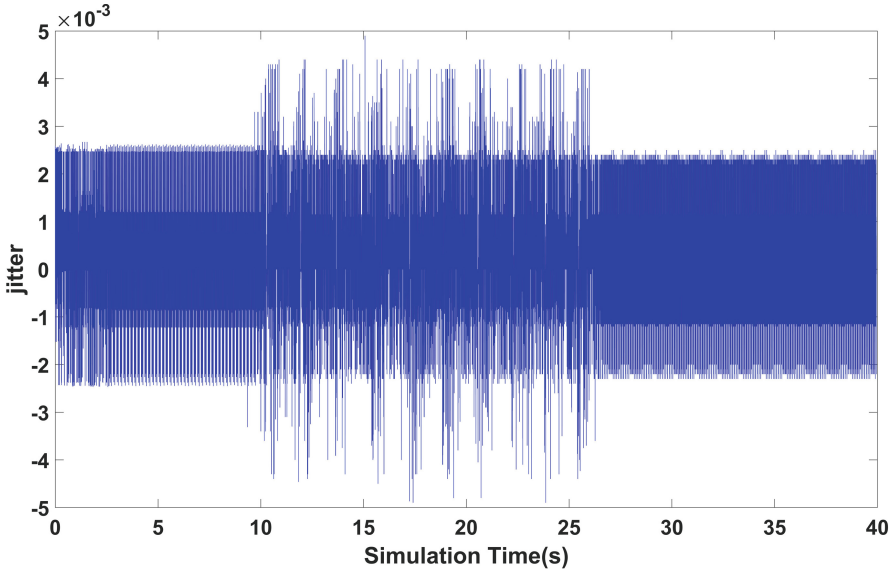

Fig. 3. Jitter in attacked QUIC network

can be seen that the addition of the attack traffic will cause the jitter to oscillate significantly, indicating that the QUIC network has been affected by the attack traffic.

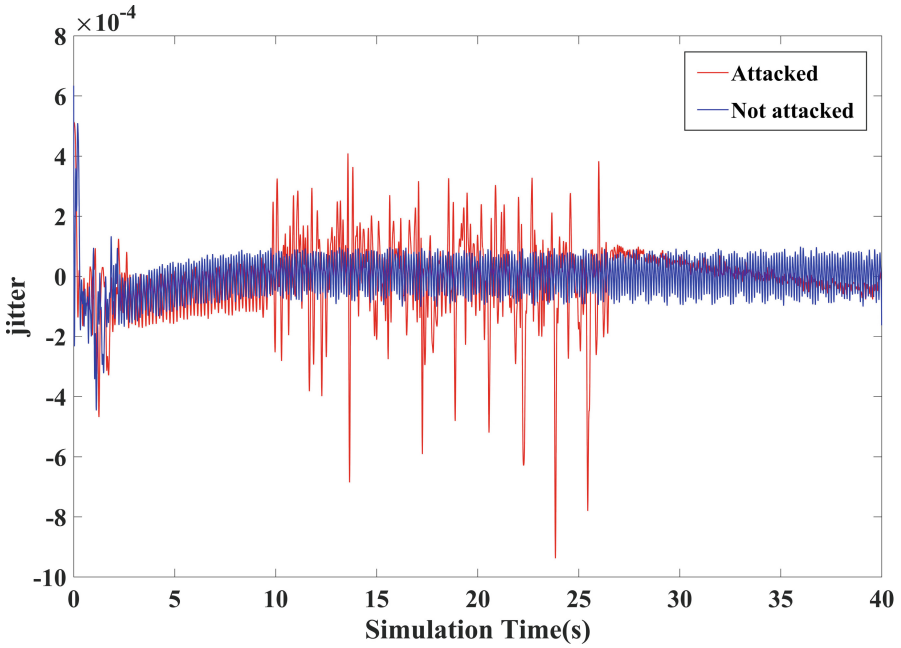

Fig. 4. Reconstruction graph of jitter in the QUIC network before and after the attack

Studies have shown that normal network traffic has self-similarity and Long-Range Dependence (LRD), and abnormal traffic will impact the self-similarity of the network. The Hurst parameter is an important parameter to describe long-range dependence. The higher the Hurst parameter, the higher the degree of self-similarity and the stronger the burst degree of network traffic. In the experiment based on QUIC traffic anomaly detection, the Hurst value before and after the attack is compared to determine whether the QUIC network generates abnormal traffic. The comparison results are shown in Fig. 5.

It can be seen from the comparison results that the Hurst value maintained a stable fluctuation before the attack traffic was added. When the attack traffic is added in the 9s, the Hurst value fluctuates violently, which disappears until the attack stops in the 29s and maintains a relatively stable fluctuation. The comparison results also show that abnormal traffic attacks can destroy or affect the self-similarity of normal network traffic. Therefore, this paper proposes an anomaly detection method for QUIC traffic based on Ensemble Empirical Mode Decomposition (EEMD) to detect abnormal traffic accurately and has a significant detection effect.

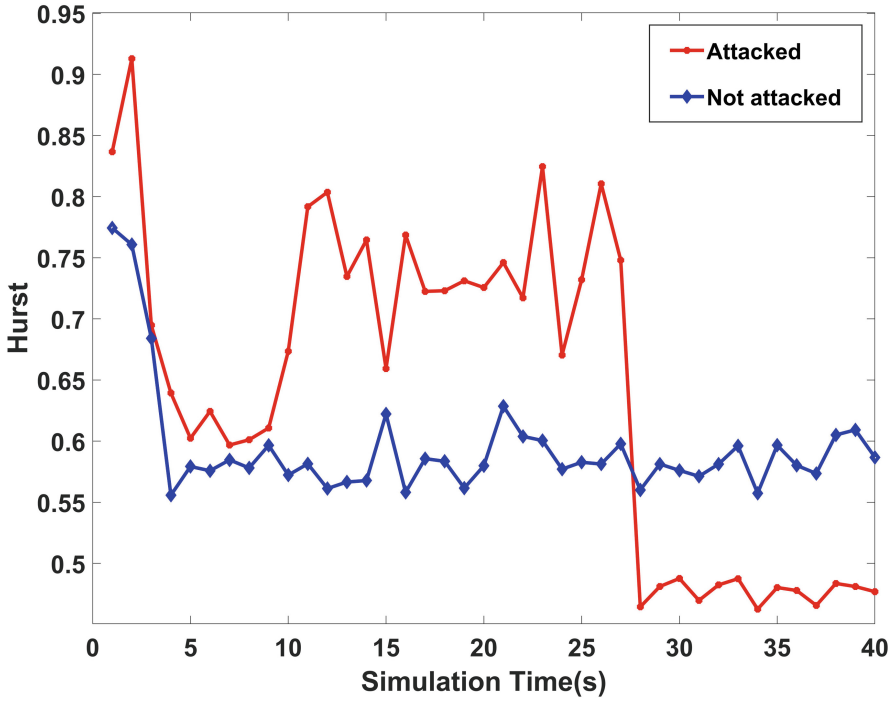

Fig. 5. Hurst parameter comparison chart

Conclusion

This paper uses an EEMD-based method to detect traffic anomalies in the QUIC network. This method overcomes the mode mixing problem in the decomposition process. The complex signal containing noise is decomposed into several relatively stable IMF components and a residual trend term, and then reconstructing the IMF component, and the high-frequency noise is removed to achieve effective denoising of the original signal. This method has good adaptability. The reconstructed signal can maintain many original data characteristics and reduce the experimental data's computational complexity. In the simulation experiment, the Hurst parameter is used as the judgment basis to test the self-similarity of the QUIC traffic to detect abnormal traffic in the QUIC network. The simulation results show that abnormal traffic will significantly affect the Hurst value, and continuous attacks will make the Hurst parameter vibrate intensely. In addition, based on the observation of the change of the Hurst parameter, the occurrence of the attack can also be judged, which has a good detection effect. Future work must consider other feature extraction and noise reduction methods to analyze the intrinsic mode components, search for the accurate criteria of attack occurrence, and study new detection mechanisms.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61962026.

References

1. Iyengar, J., Thomson, M.: QUIC: A UDP-based multiplexed and secure transport. Internet Engineering Task Force, Internet-Draft draft-ietf-quic-transport-27 (2020)
2. Nithya, B., Prakash. L.M., Kishore, J.N., Akash, M.N.: Performance analysis of pluggable congestion control in QUIC protocol. In: AIP Conference Proceedings, p. 02003. AIP Publishing LLC (2022)
3. De Quentin, C., Bonaventure, O.: Multipath QUIC: design and evaluation. In: Proceedings of the 13th International Conference on Emerging Networking Experiments and Technologies, pp. 160–166 (2017)
4. Viernickel, T., Froemmgen, A., Rizk, A., Koldehofe, B., Steinmetz, R.: Multipath QUIC: a deployable multipath transport protocol. In: 2018 IEEE International Conference on Communications (ICC), pp. 1–7. IEEE (2018)
5. Han, Y., Zuo, M., Yuan, H., Zhong, Y., Yuan, Z., Bi, T.: A QoS-based fairness-aware BBR congestion control algorithm using QUIC. Wirele. Commun. Mob. Comput. (2022)
6. Chiariotti, F., Deshpande, A.A., Giordani, M., Antonakoglou, K., Mahmoodi, T., Zanella, A.: QUIC-EST: a QUIC-enabled scheduling and transmission scheme to maximize VoI with correlated data flows. IEEE Commun. Mag. **59**(4), 30–36 (2021)
7. Shi, X., Wang, L., Zhang, F., Zhou, B., Liu, Z.: PStream: priority-based stream scheduling for heterogeneous paths in multipath-QUIC. In: 29th International Conference on Computer Communications and Networks, pp. 1–8. IEEE (2020)
8. Shi, X., Zhang, F., Liu, Z.: PriorityBucket: a multipath-QUIC scheduler on accelerating first rendering time in page loading. In: Proceedings of the Eleventh ACM International Conference on Future Energy Systems, pp. 572–577 (2020)
9. Leland, W.E., Taqqu, M.S., Willinger, W., Wilson, D.V.: On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM Trans. Netw. **2**(1), 1–15 (1994)
10. Giorgi, G., Narduzzi, C.: A study of measurement-based traffic models for network diagnostics. IEEE Trans. Instrum. Meas. **57**(8), 1642–1650 (2008)
11. Lemeshko, O., Mersni, A., Nevzorova, O.: Analysis of influence of network architecture nonuniformity and traffic self-similarity properties to load balancing and average end-to-end delay. In: Radivilova, T., Ageyev, D., Kryvinska, N. (eds.) Data-Centric Business and Applications. LNDECT, vol. 48, pp. 767–787. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-43070-2_33
12. Barford, P., Plonka, D.: Characteristics of network traffic flow anomalies. In: Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement, pp. 69–73 (2001)
13. Pei, J., Zhong, K., Jan, M.A., Li, J.: Personalized federated learning framework for network traffic anomaly detection. Comput. Netw. **209**, 108906 (2022)
14. Pan, Y.: Network security and user abnormal behavior detection by using deep neural network. Internet Technol. Lett. **4**(3), e260 (2021)
15. Hu, Z., et al.: Statistical techniques for detecting cyberattacks on computer networks based on an analysis of abnormal traffic behavior. Int. J. Comput. Netw. Inf. Sec. **12**, 1 (2020)
16. Kumari, N., Mohapatra, A.K.: A comprehensive and critical analysis of TLS 1.3. J. Inf. Optim. Sci., 1–15 (2022)

17. Morawski, M., Karbowańczyk, M.: Multipath QUIC – directions of the improvements. In: Xiang, W., Han, F., Phan, T.K. (eds.) BROADNETS 2021. LNICST, vol. 413, pp. 193–207. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93479-8_13
18. Sharma, A., Kamthania, D.: QUIC protocol based monitoring probes for network devices monitor and alerts. In: Singh, U., Abraham, A., Kaklauskas, A., Hong, T.-P. (eds.) Smart Sensor Networks. SBD, vol. 92, pp. 127–150. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-77214-7_6
19. Wu, B., Li, H., Wu, Q., Jiang, Z., Liu, J.: TMPTCP: a lightweight trust extension for multipath-TCP. In: 2020 International Conference on Networking and Network Applications, pp. 342–347. IEEE (2020)
20. Cao, Y., Ji, R., Ji, L., Lei, G., Wang, H., Shao, X.: l^2-MPTCP: a learning-driven latency-aware multipath transport scheme for industrial internet applications. IEEE Trans. Ind. Inf. **18**, 8456–8466 (2022)
21. Zheng, Z., et al.: Xlink: Qoe-driven multi-path QUIC transport in large-scale video services. In: Proceedings of the 2021 ACM SIGCOMM 2021 Conference 2021, pp. 418–432 (2021)
22. Xu, R., Joshi, J., Li, C.: NN-EMD: efficiently training neural networks using encrypted multi-sourced datasets. IEEE Trans. Dependable Sec. Comput. **19**, 2807–2820 (2021)
23. Du, S., Xu, Z., Lv, J.: An EMD-and GRU-based hybrid network traffic prediction model with data reconstruction. In: 2021 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–7. IEEE (2021)
24. Malik, H., Alotaibi, M.A., Almutairi, A.: A new hybrid model combining EMD and neural network for multi-step ahead load forecasting. J. Intell. Fuzzy Syst. **42**(2), 1099–1114 (2022)
25. Tian, M., Sun, C., Wu, S.: An EMD and ARMA-based network traffic prediction approach in SDN-based internet of vehicles. Wirel. Netw., 1–3 (2021)
26. Tian, Z., Song, P.: A novel network traffic combination prediction model. Int. J. Commun. Syst. **35**(7), e5097 (2022)
27. Zhang, L., et al.: LNTP: an end-to-end online prediction model for network traffic. IEEE Netw. **35**(1), 226–233 (2020)
28. Cao, Y., Ji, R., Huang, X., Lei, G., Shao, X., You, I.: Empirical mode decomposition-empowered network traffic anomaly detection for secure multipath TCP communications. Mob. Netw. Appl. **27**, 2254–2263 (2022)
29. Tao, X., Peng, Y., Zhao, F., Wang, S.F., Liu, Z.: An improved parallel network traffic anomaly detection method based on bagging and GRU. In: Yu, D., Dressler, F., Yu, J. (eds.) WASA 2020. LNCS, vol. 12384, pp. 420–431. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59016-1_35
30. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal. **1**(01), 1–41 (2009)
31. Flandrin, P., Rilling, G., Goncalves, P.: Empirical mode decomposition as a filter bank. IEEE Sign. Process. Lett. **11**(2), 112–114 (2004)
32. Wu, Z., Huang, N.E.: A study of the characteristics of white noise using the empirical mode decomposition method. In: Proceedings of the Royal Society of London, pp. 1597–1611 (2004)
33. Flandrin, P., Gonçalves, P., Rilling, G.: EMD equivalent filter banks, from interpretation to applications. In: Hilbert-Huang Transform and its Applications, pp. 57–74 (2005)