Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 1, 2026

# COMBATING MODAL ALIASING IN MPTCP ANOMALY DETECTION WITH ENSEMBLE EMPIRICAL MODE DECOMPOSITION METHOD (EEMD)

[1]Noor Fatima, [2]Farzeen Walayat, [3]Alisha Maqsood, [4]Mir Ahmed Rehan
[1]*MS Information Technology, University of Gujrat, Pakistan*
[2]*MS Information Technology, University of Gujrat,pk*
[3]*Comsats University, Islamabad*
[4]*MS Information Technology Project Management, University of West of London*
[1]Noorfatima.nt@gmail.com, [2]farzeenwalayat98@gmail.com,
[3]alishamaqsoodmalik@gmail.com, [4]Ahmedmir611@gmail.com

**Article History**
*Received: 03 Jan, 2026*
*Accepted: 22 Jan, 2026*
*Published: 25 Jan, 2026*


**Corresponding Author**

**Abstract**
*Internet usage has immensely transformed the traffic and it is important to have improved techniques for conveying data and information. It is later known as Multipath Transmission Control Protocol (MPTCP), which becomes one of the solutions that offers the capability in enhancing the reliability and availability of connections through utilizing several paths for the transmission of data. However, this dynamic characteristic of MPTCP is a challenge about handling of anomalies because of the changes and interferences between the paths. Some traditional approaches such as Empirical Mode Decomposition (EMD) that were applied in prior studies detecting the MPTCP networks always lacks the way of dealing with the modal aliasing that combines the signal components hence reducing its capability of detecting anomalies. This thesis presents an improvement of the EEMD method used to solve the problems of modal aliasing in the process of the MPTCP anomaly detection. The solution given below enhances capability in detecting anomaly by minimizing the addition of modal aliasing; which in turn leads to better signal to noise rates hence better identification of anomalous networks. Algorithms proved through experimental data show that this proposed EEMD-based method is superior to the standard EMD method and can be applied to real-life MPTCP anomaly detection scenarios as an effective solution. The NS-3 setting confirmed EEMD outperformed EMD because it produced 94.6% accurate results alongside 92.3% precision and 95.1% recall. The model shows better anomaly detection through fewer false alarms and it finds transient problems which normal detection methods cannot locate. Therefore, the conclusion reached from this research outlines a more efficient approach towards identification of network anomalies and hence improve the security and reliability of the intended MPTCP based communication networks*

## I. INTRODUCTION

The concept of Internet of Things (IoT) has become widespread, envisioning a world where various intelligent devices may get in contact with each other using internet data and technologies. The IOT is being utilized throughout the various areas of the modern economy such as health, quality control, supply chain management, power, farming, manufacturing and networking[1].

With the emergence of the concept of IoT and cloud, computing there has been a significant change in the nature and has made revolutionary changes to two or multiple fields in recent years. Simultaneous generation of data in parallel puts pressure on the channel that transmits the communication. This represents interaction and use of current communication devices in numerous network. The availability and access interfaces has significantly improved since the usage of various wireless internet services. Internet systems including Wi-Fi, 5G and above, and satellite[2].

Mobile and fixed devices, capable of connecting to more network that are supported through Multipath Transmission Control Protocol (MPTCP). The technology can improve the application capability or throughput by utilizing several network channels, which they used for combining bandwidth and at the same time data transfer. MPTCP is a protocol, which enriches the capability of 'traditional' TCP, and has already been standardized by IETF – International Engineering Task Force. It is expected that the usage of MPTCP will be rather significant effect on the future of Internet Data Transfer Services as it has the potential of creating and support users to transmit data through multiple paths at once and remain synchronizable with today's internet devices [3].

At present, there are many devices and products, which may support the feature of MPTCP integration. The MPTCP suite allows for creation of a connection between two hosts, which can utilize several pathways simultaneously. The capability to work independently enhances the effectiveness and organization of responsibilities extremely significantly since it allows working in parallel, which improves the speed and reliability of the system with the elemental consistency of the TCP interface for applications. MPTCP's usage enables smartphones to connect two Different networks concurrently and stream videos simultaneously hence improving the experience of the user. The amalgamation of theoretical concepts and practical endeavors is consistently enhancing the Multi-Path Transmission Control Protocol (MPTCP) [4].

MPTCP has advantages such as multi-data transfer and fast transmission; however, it will be easily attacked and interfered in actual conditions. MPTCP based multipath transmission network suffers from network threats such as LDDoS attack decreasing network reliability of the proposed solution and consumer streaming media experience [5].

As opposed to TCP, MPTCP breaks up security vulnerabilities into its individual sub-flow. Sometimes, there is a possibility that by targeting a single subflow will not cause a significant effect on the MPTCP connection. If an attacker only intercepts one subflow, he/she should not be allowed to intercept the entire MPTCP connection stream. That means MPTCP has far greater variability in the capacity of its connections than TCP has. If any of the subflows are to be attacked by an attacker for example a man-in –the middle, then the whole MPTCP connection will be affected. This places MPTCP in a disadvantageous position as compared to TCP [4].

Security can also not be over emphasized in any system particularly communication networks. These multiple channels make it possible to have some weaknesses for the prospective adversaries particularly during the creation of connection. The ones include the man-in-the- middle (MITM), denial-of-service (DOS), and SYN flooding attacks. It can be seen that security on the network is important when it comes to communication [6] [7]. Because of this consequence, MPTCP becomes exposed to attack, in case the intruders are able to get access to the session of the different pathways. Actually, there are several types of attacks, which the given protocol is prone to, and they include the following ones: The Man in the Middle attack, the Denial of Service attack, and the SYN-Flooding attack; the execution of such attacks requires

adversaries to obtain specific MPTCP user credentials which include their IP addresses together with flow-sequence numbers and session information. Any MPTCP session requires the secret key which serves as the session key for encryption and decryption processes. However, it is relatively easy for an opponent to identify the sequence number of a flow and the IP address while the identification of a session key presents a much greater difficulty. If an opponent has information on a flow both the IP address along with the sequence number they are virtually there to harm the system mainly due to lack of session key [8].

The description establishes that MPTCP experiences security threats which necessitate protection methods to sustain reliable and protected communication. In MPTCP the signal exchange strategy faces risk from DSS manipulation which represents a complicated packet spoofing procedure within the MPTCP framework. The DSS option manipulation in MPTCP alters connection-level acknowledgement through modifications to the recently added MPTCP DSS component that extends the TCP optimistic ACKing mechanism [9].

Insufficient effort has been devoted to enhancing the security of MPTCP networks, particularly in the area of detecting anomalies in MPTCP-based network traffic. The multipath routing investigates a quality of experience (QoE)-driven data delivery model in mobile ad hoc networks (MANETs) using multipath TCP (MPTCP) for reliable and efficient transmission of data over many paths. The system examines both symmetric and asymmetric multipath communication and builds emergency applications with clear QoS parameters for generating routes, selecting paths, and scheduling packets. While this improves communication, it is still necessary to enhance the application level Quality of Experience (QoE) [3].

The provided signal is analyzed using a few different techniques, including Iterative Filtering (IF) and Empirical Mode Decomposition (EMD), which serve to describe the signal as the total of even more fundamental, structured, or controlled elements known as Intrinsic Mode Functions (IMF). However, they work better when evaluating signals that are nonlinear or nonstationary in nature, and they can be used just as easily if the assumptions that underpin these techniques and their limits are not given enough thought [10].

EMD functions as a multi-purpose tool which analyzes one-variable data sets. The data possesses characteristics of non-linearity and non-stationarity when it exists in the time series format. Through EMD the decomposition of complex signals produces a limited number of fundamental functions known as intrinsic mode functions (IMFs) together with one single convexity signal referred to as propensity. This technique requires the acquisition of a sequence of temporal-spectral descriptions by applying the Hilbert-Huang Transform to transform every basis representation of the IMFs to generate IF characterizations the following (HHT) will be used; The instantaneous It has been shown that frequency components give a more general notion of the local phase change in an accurate manner. which makes them more appropriate for analyzing non-stationary processes in contrast to other time series used for the analysis of so-called local or 'instantaneous' frequency. While recent research shows that, the program is still on its positive progress there are still core challenges which has not received a proper solution such as correct and immediate recognition of tiredness [11].

A) RESEARCH GAP

The literature exhibits several drawbacks in current Multipath Transmission Control Protocol (MPTCP) anomaly detection methods even with its advanced development. The majority of anomaly detection systems depend on Earth Mover's Distance (EMD) because it measures distances between probability distributions effectively. EMD produces amplified signal-to-noise ratios during specific MPTCP network conditions thus causing modal aliasing effects which result in misidentifying network anomalies. The detection field of MPTCP anomalies misses out on Ensemble Empirical Mode Decomposition (EEMD as a decomposition approach due to its lack of full implementation despite Empirical Mode Decomposition (EMD being deployed in certain contexts. The current algorithms demonstrate poor flexibility when handling different scenarios of MPTCP applications

and network loads in real-time systems. Research gaps in existing studies led to new explorations of adaptive anomaly detection methods for MPTCP networks that aimed to resolve modal aliasing problems along with capability to deal with complex traffic patterns. Anomaly detection models currently succeed in identifying persistent or large-scale anomalies yet they lack the ability to detect minor short-duration or faint anomalies that matter in network operation and cybersecurity.

B) CONTRIBUTIONS OF THIS STUDY

This study makes its main contributions through the following timeline: This research brings Ensemble Empirical Mode Decomposition (EEMD) into MPTCP network anomaly detection as a new approach for remarkably reducing modal aliasing effects that EMD techniques like Earth Mover's Distance (EMD) would introduce. With the application of EEMD the method successfully breaks down complex non-stationary traffic indicators into higher quality intrinsically generated frequency modes for better detection of transient system anomalies. The research study performs a comprehensive analysis to evaluate how the proposed EEMD-based detection system outmatches typical EMD-based approaches. The research findings establish that the proposed approach achieves better results in identifying anomalies and resisting noisy data. The MPTCP network scenarios received simulation through NS-3 for developing realistic dataset which enables reproducible tests while providing research benchmarks to future studies about network anomaly detection [12].

The research adopts Ensemble Empirical Mode Decomposition as its analytical method to enhance detection performance while improving efficiency of abnormality detection. The purpose is to minimize the problem occurrence. The problem arises when limited sampling rates misinterpret high frequency wave components to appear at lower frequencies.

## II. LITERATURE REVIEW

Multipath TCP (MPTCP) enables a set of hosts to effectively utilize various network pathways and send data across numerous interfaces. The multipath functionality of MPTCP enhances the overall data transfer rate, minimizes transmission delay, and improves the resilience of communication.

The author proposed queue management strategies to analyze MPTCP network effectiveness under attacked conditions. These approaches demonstrate different outcome speeds combined with unique delay periods and breakdown percentage. The algorithms solved attack detection problems through effective performance indicator assessment of throughput, latency, and packet loss rate. The implemented techniques failed to sufficiently solve the problem of anti-attack operation. [13].

The author suggested a model for detection of anomalies which is QUIC network traffic, by leveraging the self-similarity of the data. The approach utilizes Ensemble Empirical Mode Decomposition and Approximate Entropy. It decomposes the network traffic into many IMFs and a residual term. It incorporates Approximate Entropy to assess the complication of each component and use the Hurst parameter to identify the occurrence of atypical traffic. The model surpasses previous detection methods by effectively addressing issues of limited adaptability and low efficiency. It shows high effectiveness in its work and could by applied for identification and prevention of attacks on the nodes of the network[14].

In another work, the author discussed how to solve the challenge of having a successful MPTCP Congestion Control Algorithm (CCA) that would accommodate MPTCP CCA design criteria. He suggested that MPTCP should be changed such that it is enabled to use BA- MPCUBIC, a multipath CUBIC variant that is based on SPTCP CUBIC for SFs passing through NSBs and multipath CUBIC for SFs passing through SBs to overcome the current constraints and achieve the desired objectives [15].

The author introduced the Path search algorithm as a solution to counteract the manipulation of internet path diversity in MPTCP when facing a man-in-the-middle attack. By examining the graph at the Autonomous System (AS) level, it was possible to identify countries and areas that exhibit greater resilience. The outcome was enhanced reliability, increased throughput, and improved confidentiality [16].

The MPTCP network anomalies are discovered autonomously, without any prior network

knowledge. The model utilized a combination of multi-scale detection and DSS theory to effectively accomplish the detection of anomalies. This detection method is based on the non- distinctive feature of MPTCP traffic. The EMD method was employed to break down MPTCP traffic data and eliminate high-frequency noise and residual trend term in order to rebuild the legitimate signal. The model employed the concept of sliding windows to assess alterations in the Hurst exponent of the MPTCP network across several assault scenarios, aiming to detect any anomalies [17].

The author presented an improved version of the baseline MPTCP network, known as multi- expert MPTCP, which enhances the detection of network threats with greater precision. He introduced an MPTCP variation that utilizes multi-expert learning to improve the resilience of multipathing against network threats. Unlike MPTCP, MPTCP-meLearning incorporates two extra components: the ME Learning Station and the ME Learning-based Scheduler. The ME Learning Station aims to enhance the prediction accuracy of each path in MPTCP, unexpectedly in the presence of heavy losses or sudden shifts caused by network attacks. This is achieved by utilizing a set of formula-based predictors. The ME Learning based Scheduler is responsible for managing the scheduling of tasks based on the expertise of multiple experts [18].

The author proposed a quantification technique to examine the susceptibility and resilience of MPTCP to cyber-attacks in situations when the information provided by the network is insufficient. This is based on the observation that cyber-attacks typically lack real-time information on different aspects of MPTCP. The user mathematically analyzed cyber-attacks on the MPTCP communication system, taking into account incomplete network information. They proposed a mixed attack strategy that combines arbitrary attempt to attack and distinct attacks to assess the strength of MPTCP [19].

In[20], the author has introduced and executed MPTCPsec, an addition to Multipath TCP that utilizes identification and enciphered data to mitigate diverse protection vulnerabilities. The design consists of three components. It incorporates TCP-ENO into Multipath TCP, enabling the negotiation of several security protocols in the course of three- way handshake. Additionally, it offers methods of authentication and encryption of data that is being transferred through a linkage and an ability to transfer the same data through other sub flows. In addition, MPTCPsec aims to authenticate MPTCP options in order to protect against the different types of DoS attacks [21].

In particular, the author introduced a workable approach to the heuristic mechanism, which could be used to increase the efficiency of an MPTCP flow addressing the issues of disconnection of paths and head of line blocking. This algorithm is designed for the identification of k pathways, which are most important. The proposed heuristic technique showed enhanced computational performance in the routing of the networks as compared to the prime solution. The first step describes the set of all arcs from the initial node in the computer that defines the curly bracket. to the last point within a multi-nominal time horizon. After that, the greedy algorithm is used to select k number of non-intersecting routes which have a significant bottleneck bandwidth from the existing pathways [22].

To understand this argument, the author had presented a proficient method of increasing the potential of an MPTCP flow addressing the issues of disconnection in the paths and the head- of-line blocking. In this algorithm, special emphasis is given for identification of “k ” pathways which have the highest bandwidth and are independent of one another in their entirety. The proposed heuristic technique have shown increased computational speed in respect to routing the networks in comparison to the prime solution. Thus, the program first defines the feasible paths from the start node. up to the last point within multi nominal time period. Further, we use the greedy algorithm to used for the selection of k non-overlapping routes each possessing a significant bottleneck bandwidth of the existing pathways. Consequently, the data transfer rate of a Multi-Path TCP (MPTCP) connection experiences a substantial augmentation. The fake results demonstrated that the proposed method achieved better average throughput for an MPTCP flow than both the first k-Max unconnected paths and SPF methods. Heterogeneous distance matrix was used where the original function was

slightly altered with the shortest path algorithm, which was used to identify other paths between the source and the destination. It then used greedy strategy to choose the k Maximum-Minimum bandwidth disjoint path [23].

involves suggesting objects or events, which deviate from the normal cases or do not show conformity to other aspects of data. Anomaly detection has been applied to detect and analyze the outliers of the components of data. Some of these difficulties are brought about by inadequately sized data of anomalous reality, the ground training data, influences arising from disparity of environmental conditions, the position of the capturing structures, and illumination conditions. Anomaly detection has numerous uses which are (though not exclusively) as follows industrial damage mitigation, sensoring, health and medical services, road monitoring, and violence forecasting among others. Deep learning especially has made it possible for systemic improvements in the area of anomaly detection.the author has categorized the anomaly detection techniques clearly and provided a comprehensive report on the existing pioneering work. In particular, this paper addresses the focus areas of machine learning and deep learning anomaly detection techniques and measures, as well as anomaly detection categorization. It also describes different datasets that can be used in anomaly detection as well as several comparatively recently proposed real-world datasets [24].

The architecture-based method of deep learning for classification includes hidden layers but can be added for optional enhancements and contains multiple non-linear processing steps in a hierarchical structure. The major limitation of augmented neural networks arises from their inability to receive proper training through regular back propagation because pretraining through generation needs to be employed for overcoming this issue [25].

The author quickly outlines popular anomaly detection approaches which include k-nearest neighbors (k-NNs) both as a grouping tool and a regression tool under non parametric classification. The algorithm functions optimally in anomaly detection cases because it employs a basic but effective classification approach for identification and categorization of abnormalities. The anomaly detection process often utilizes Support Vector Machine as a standard method. The SVM framework (specifically one class SVM in combination with least squares SVM) shows extensive usage within spacecraft operations and aviation systems as well as electrical systems but its original supervised design makes it most suitable as an anomaly detection classifier for classification problems. When used for anomaly detection systems NNs function both with multi class and one class implementations. Basic anomaly detection using NNs executes its method through two sequential steps that begin with training a neural network on normal data and concluding with anomaly identification through network-based regularity verification or irregularity detection during testing. The detection of one class anomalies utilizes neural networks. The machine learning predictive model known as a decision tree contains variable (feature) nodes along with child nodes showing variable (feature) value ranges coupled with external node (leaf) predictions of target variable (feature) values and judgment nodes that provide target variable (feature) predictions. Deep learning has developed to the point where it now demonstrates significant improvements in learning expressive forms of complex data which includes high dimensional, temporal, graph and spatial information. The learning algorithms of deep learning extract features from non linearly diverse and high dimensional noisy instances which become useful for irregularity detection [26].

As part of the research the author developed a new IDS which utilizes DNN architecture to enhance vehicle network security. The DNN structure parameters receive in-vehicular network packets that produce probability-based feature vectors for training purposes. With DNN network capabilities the sensor obtains probabilities that separate normal and attack packets for every packet to perform malicious packet detection [27].

The stream based machine learning approach for network security detects anomalies by applying and assessing various machine learning algorithms that process evolving network data streams. Analysis of data streams occurs throughout continuous changes. Finding suitable machine learning models becomes challenging because the domain produces

algorithms specifically made for data stream mining while multiple evaluation strategies exist solely for benchmarking these algorithms. Stream machine learning analysis processes one data instance per time with single inspection which requires minimal memory use while operating under time constraints and offering prediction at any stream moment. The combination of Prequential 10 fold CV evaluation with ADWIN for dynamic window sizing has proven suitable as an algorithm benchmarking method in stream-based systems because it detects temporary changes and concept drift events effectively [28].

A separate researcher applied wavelet transform analysis method with information entropy theory for attack signal feature extraction through wavelet energy entropy. Simulations led to acceptable outcome results. Experimental results indicated that attack signals could be differentiated from normal signals because wavelet energy entropy changes in proportion to different decomposition scales. The detection system effectively implements this feature to monitor MPTCP transmission attacks while providing innovative detection methods for security monitoring systems [29].

The anomaly network traffic function detects abnormal or anomalous multimedia medical equipment that provides medical diagnosis of neuronal. This paper studied detecting traffic anomaly in large scale multimedia communication networks which transfer multimedia data between multimedia medical devices and various medical issues in clinic. Time frequency analysis theory serves as the foundation for the researcher to create a new network traffic anomaly detection method through combining wavelet packet transform with empirical mode decomposition and spectral kurtosis method. Silverman addressed the modal aliasing issue in traditional idler empirical mode decomposition by developing a modified version that served as the basis for their work. They implemented the anomaly detection method through the threshold-based splitting of network traffic yet the main issue involved inaccuracy due to modal aliasing problems[20][30].

A research proposed the combination of EMD with the Gated recurrent unit neural network for network traffic prediction. Analysis of traffic data starts with first cleaning up missing-points while performing outlier removal operations. After reconstructing the traffic data, we process it through EMD decomposition before training single GRU networks on individual decomposition components. Each predicted value component comes together and results in the final output. The proposed prediction model delivers improved accuracy along with more stable and better performance when compared to existing models according to numerical output result [31].

The development of EMD as a new analysis method took place during the recent period. Analysis of nonlinear as well as nonstationary processes becomes possible through decomposition which relies on data-specific local attributes. The network traffic anomaly detection through weighted self-similarity parameter represents a new method derived from EMD calculations. The new method stands in opposition to all past analysis procedures. With this method, network attacks can be both suddenly detected and actively prevented from starting or halted in progress. It can predict the duration of ongoing attacks [32].

These methods do not follow the detection methods currently used for anomalies. Research communities adopted both the realization and several parameters of Markov chain methods. The short memory approach of classical Markov showed insufficient data connection while the higher order Markov diluted its long memory ability. The model reliability experienced deterioration because of the link connecting previous data to current test data. These models do not show capability to represent the sequences beyond their tendency to vary. Anomaly detection research implemented a dynamic Markov model with an innovative approach as its second methodology [33]. The algorithm split sequence information through successive sliding window operations. Data states defined by the sliding window were used to determine data values which led to the creation of higher order Markov models with suitable orders for balancing purposes. We need to modify the memory property of the detection model to match present sequence trends. We implemented a anomaly substitution method to stop the detected anomalies from ruining model

development while ensuring continuous anomaly detection [34].

A new approach named ensemble served to address the scale mixing issue in data analysis methods. EMD defines IMF components as the average output from an ensemble that contains both the signal with white noise of finite amplitude. The actual IMF The result can be achieved by letting ensemble numbers reach an infinite value. The fundamental concept of EEMD involves filling the time-frequency space with various frequency levels of added white noise regardless of component distribution in the original data. With the signal added to this uniform background the white noise produces matching scales of reference which are then utilized for projecting different signal scales. The procedure produces great amounts of noise during a single trial yet the ensemble mean of multiple trials mitigates the noise until the trials represent the actual solution when the ensemble mean reaches infinite trials. This combined redundancy procedure based on EEMD generates pairs of data decomposition outputs that are consistent yet slightly different to one another. EEMD stands as an advanced version of original EMD thus demonstrating higher capability in noise management. We do not need the additional level of noise so it is essential that the noise amplitude stays finite and produces many fair choices in each ensemble. The analysis process functions through EEMD independently because it does not require manual modification thus providing adaptive results. This method generates IMFs while solving mode mixing problems along with providing physical interpretations for every IMF component which leads to a time-no transitional gap in the frequency distribution [35].

The mode decomposition of network traffic employs ensemble empirical modeling as one of its analysis methods. The authors designed a multimode search algorithm within an adaptive algorithm model. An adaptive algorithm was developed. The added white noise allows the automatic generation of initial values for the model. A dependence on the amplitude and the ensemble number. The paper investigates multimode search algorithm optimization of two critical parameters. The approach provides extensive opportunities for improving both global search and local search performance. The main Adaptive parameters necessary for EEMD were obtained through this process [36].

The ensemble empirical mode decomposition (EEMD) has some power to reduce the modal confusion phenomenon that exists for the EMD method, but is unable to completely neutralize the introduced white noise. CEEMD is then a complementary method to EEMD such that the analysis signal is forced by adding white noise with the opposite signs. This reduces the reconstruction error greatly in pairs. To solve the problem of modal confusion and a challenge of accurately extracting the fault features of A CPCEEMD (CEEMD-PE-CEEMD) bearing fault diagnosis method based on rolling bearings is proposed, involving any possible CP errors of measured data, full It combines CEEMD algorithm with advantages of signal randomness detection by permutation entropy (PE). Permutation entropy is then applied to detect the abnormal components of the CEEMD. The remaining signals are CEEMed directly. We worked with IMF components that contain scaled Hilbert envelope spectrum analysis is performed with large correlation coefficients selected, and fault features are. From the envelope diagram. The results show that the proposed method has good decomposition effect with The feasibility of inhibition effect on the modal confusion in the EMD process, and can extract characteristic information of the rolling bearing fault signal effectively [37] [38].

A statistical research approach served as an alternative proposal. The wind speed prediction operates independently from numerical weather prediction inputs. The study demonstrates an analytical hybrid prediction technique for short-term data processing which generates high-quality outcomes. Scientists created this dataset specifically to enhance the accuracy levels of their predictions. The Ensemble algorithm brings optimal results as a signal decomposition method. Empirical Mode Decomposition serves as the method to perform preprocessing. Lower frequency components in the Wind speed data were extracted through its ensemble empirical mode decomposition. The future wind predictions are regressed to multiple signals which result from decomposition. A speed

value will be used through an Adaptive wavelet neural network model. A performance evaluation of the proposed hybrid approach takes place for the South India wind farm. The research implements a comprehensive analysis between India and other documented countries within a real-world investigation. High accuracy resulted from the prediction performance while the method also reduced both uncertainty and computational load in the predictions [39].

TABLE I: *SUMMARY OF REVIEW*

| Reference | Problem Statement | Methodology | Results |
|---|---|---|---|
| [5] | To detect LDDOS attacks in Multipath TCP environment | Queue management algorithms are used and the performance is compared. Two types of algorithms are used; active queue management algorithm and passive queue management algorithm. A topology named double dumbbell simulation with quantifiable LDDoS attack traffic is used. The total simulation time is 60 seconds. | The robustness capability of system is enhanced which in return enhances the transmission performance. |
| [14] | To prevent the vulnerability of QUIC-based distribution system to attacks due to asymmetric encryption | EEMD and approximation entropy is used. The model uses IMF decomposition and Hurst exponent to determine the abnormality of traffic. | The model successfully addresses the issues of limited adaptability and low efficiency, while also demonstrating excellent accuracy. It can be utilized for the detection and prohibition of network node attacks. |
| [15] | To ensure better throughput in network transmission in adaptive environment, and control the congestion problem in network traffic | Bottleneck aware detection method using round trip time and packet loss ratio. Linux kernel is used for the implementation | The proposed system successfully detects subflows with better detection accuracy |
| [16] | To mitigate the exploitation of internet path diversity in MPTCP against man in the middle attack | Path search algorithm is used in order to find the valid path of data transfer. The path diversity is measured by connecting each two nodes that satisfy the routing properties | MPTCP does not help much in achieving robustness against man in the middle attach unless the edge provider is carefully taken |
| [17] | To detect the anomalies in MPTCP (without prior knowledge) | Empirical mode decomposition with Hurst exponent analysis is used to detect the anomalies in accurate manner. The Hurst exponent of abnormal traffic increases more than the normal traffic | Hurst exponent analysis provides better detection of anomalies in network traffic |
| [40] | To provide robustness and high transmission rate | Active queue management algorithm is used | Better performance rate in attack detection |

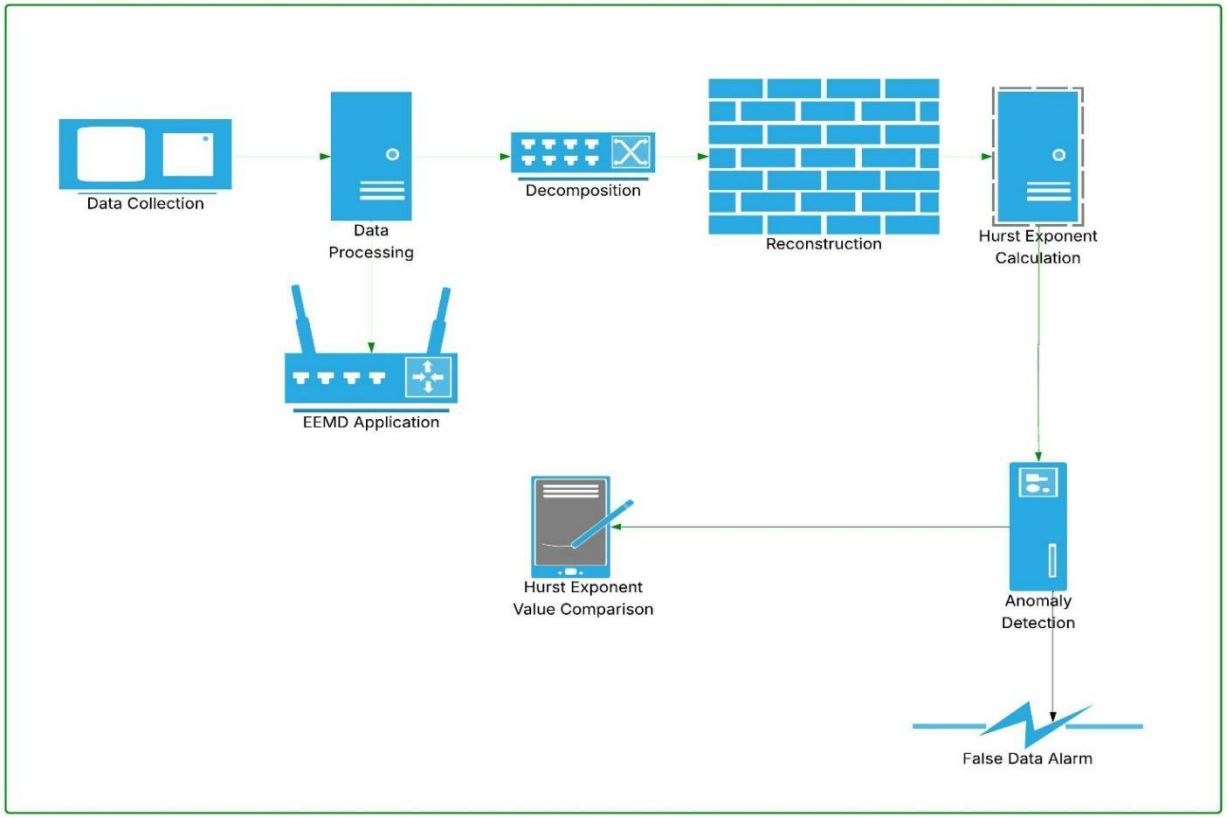
## III. MATERIALS AND METHODS

In this study, I have proposed an ensemble empirical mode decomposition method to contend with the problem of modal aliasing in detection of anomalies in MPTCP network. My focus is to lessen the modal aliasing issue so that the signal to noise ratio of data decreases and the detection of anomalies gets better.

### *A. ARCHITECTURE OF PROPOSED METHODOLOGY*

The architecture of the proposed methodology is shown in Figure 3.1. It consists of the following modules

In the proposed method, the Data Collection module is in charge of gathering raw data from network traffic. After collection, the data is transmitted to the Data Processing module, where it is converted to the model-acceptable format and split into smaller magnitude bands using EEMD. This module adds white noise to the data to balance out the effect of abnormalities that cannot be noticed due to the increased signal to noise ratio. Because of the decomposition process, an unnecessary number of IMFs were formed, and the signal was then reassembled and implemented using the Hurst exponent. The Hurst Exponent Calculation module then does statistical analysis to produce the Hurst exponent, which is critical in determining the long-term memory or persistence coefficients for the data series. Finally, the Anomaly Detection module uses the derived Hurst exponent to detect deviations or anomalies in the set information. Such a method ensures that no data is treated in the same way, while the structure ensures the utmost accuracy of anomaly detection, utilizing each module's functions to benefit the system's successful operation.




1) DATA COLLECTION

The original network data was generated by the NS-3 (Network Simulator 3) setup, which is a sophisticated and completely adjustable cellular network simulator. The NS-3 offers a set of predefined modules as well as traffic models, which will aid in the easy and realistic simulation of networks. In creating the test data for this study, I used NS-3 and established a Multiple Path Transmission Control Protocol (MPTCP) network

environment. To obtain the appropriate network circumstances and scenarios, the simulator's native traffic models were used. It enabled me to assess network throughput, latency, and packet drop in a controlled environment. These simulation results serve as the dataset for evaluating the suggested approach's effectiveness in addressing the modal aliasing challenge in anomaly identification. I utilized NS-3 for performance simulations since it has many aspects that correspond to real traffic. To obtain trustworthy and relevant results that supplement our analysis and results section, I employed realistic traffic models.

2) DATA PROCESSING

Here, the gathered data's throughput is examined. Throughput is the pace at which data is transmitted across a network. To calculate throughput, the signal is first separated into numerous IMFs in order to effectively identify the signal from noise. I've divided the attacked and un-attacked signals into multiple IMFs. The EEMD approach is applied to an assaulted signal, in which white noise is supplied to remove the abnormality and rule out the noise. The throughput of attacked and un-attacked traffic is calculated. The graphic below depicts the throughput of an un-attacked MPTCP network.

3) RECONSTRUCTION

During the reconstruction phase of my analysis, the last IMFs are carefully reconstructed to give the best fit of the original signal. In this technique, several IMFs are combined accumulated to reconstruct an overall signal that has the features of the initial data but is free of noise and abnormality. Every IMF is a part of the signal with its own defined characteristics of the frequency and amplitude, therefore, summing all the IMFs result in an overall approximation of the given signal. This reconstructed signal is important as it forms the basis for the subsequent analysis more so in the Hurst exponent calculation I have applied. Finally, this reconstructed signal is used to compute the Hurst exponent that characterizes the 'memory' or 'history' contained in the time for persistence or randomness of the data. With an accurate reconstruction of the signal from the IMFs, the Hurst exponent will give a more valid analysis of the statistical characteristics of the original time series.

4) HURST EXPONENT CALCULATION

In the next stage, on the reconstructed signal I carried out the Hurst exponent computation. Finally, after reconstructing the signal from the IMFs, based on their density, I underwent the Hurst exponent evaluation of the signal's long-term statistical characteristics. According to the Rescaled Range (R/S) analysis, I split the reconstructed signal into segments, determined the range or fluctuation of these segments and then rescaled to estimate the value of Hurst exponent. This helped me to assess the likelihood of persistent patterns of the signal or switching back to mean value. From this analysis, the Hurst exponent provided other useful information regarding the fractal nature of the signal and various degrees of self-similarity, such that will allow the identification of noise from the actual data. It is obviously essential for the analysis of the signal behaviors and the temporal characteristics for the establishment of the further methods and models focusing on the anomaly detection and analysis in the time domain.

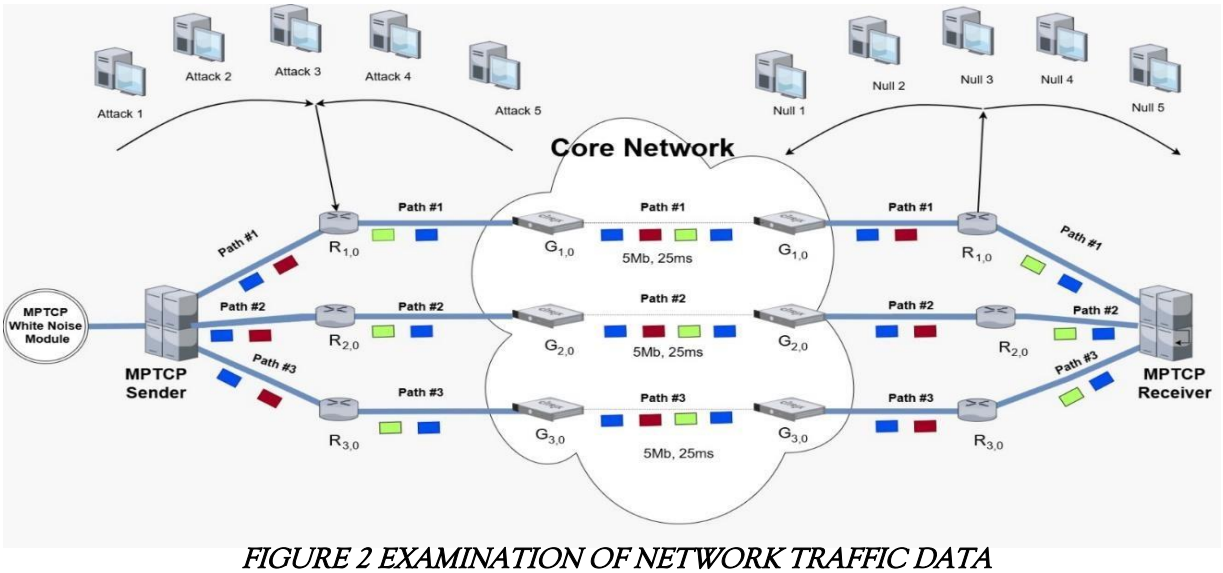



*FIGURE 2 EXAMINATION OF NETWORK TRAFFIC DATA*

5) ANOMALY DETECTION

Finally, during the last step of the analysis, I used an anomaly detection on the values of the Hurst exponent of the reconstructed signals. After obtaining the Hurst exponents, I went on and applied these values in order to separate the signal and anything peculiar. What I did in this case was to subtract the Hurst exponent of each segment from expected abundances or a baseline and the results that are either higher or lower than the thresholds signify anomalous behavior. Abnormalities are defined by gross deviations from the Hurst range, which points to distorted features or 'abnormality' in the signal. This step facilitates the elimination of the normal signal in order to present a very clear distinction between normal and abnormal signals so they can be independently analyzed. Abnormalities are defined by gross deviations from the Hurst range, which points to distorted features or 'abnormality' in the signal. This step facilitates the elimination of the normal signal in order to present a very clear distinction between normal and abnormal signals so they can be independently analyzed. As such, identification of such anomalies makes it easier to understand the nature of the signal behavior and improves the resilience of the the anomaly detection system. The MPTCP communication scenario in my study is as follow:

### *B. ALGORITHM OF EEMD*

Let x(t) be the original signal from the MPTCP network. Distinguish the number of ensemble trials N. Set the amplitude of the added white noise σ. Add white noise for each noisy signal. For each noisy signal xi(t): Perform EMD to decompose xi(t) into a set of Intrinsic Mode Functions (IMFs) IMFij(t), where j denotes the IMF index. Obtain residual ri(t) after extracting all IMFs. Average the IMFs obtained from all ensemble trials to get the final IMFs: IMF ( ) = $\sum_{i=1}^{n} IMFij(t)$.

This step mitigates the effect of the added noise and ensures that only the robust components are retained. To reconstruct the signal x(t), combine the averaged IMFs IMFj(t) and the residual r(t): x(t) = ∑ ( ) + ( ), Where M is the total number of IMFs. Analyze the IMFs and the residual for anomaly detection. Use the Hurst exponent on the reconstructed signal components to detect probable anomalies in the MPTCP network.

### *C. ALGORITHM OF HURST EXPONENT CALCULATION*

Choose different segment sizes n (subsets of the time series) to compute the Hurst exponent across various scales. For each chosen segment size n, divide the original time series x(t) into k non-overlapping segments of length n: K= N/n, These segments are denoted as x1, x2,....., xk. Calculate the Rescaled Range R/S for Each Segment; For each segment xi={x1,x2,...,xn}: Compute the mean of the segment: x i= 1/n ∑_(j=1)^n xij. Construct the mean-adjusted cumulative time series Yij: = ∑ ( − ), = 1,2, ... .n. Calculate the range R(n) of the cumulative deviations with R( ) = max( ) −

min(*Yij*) , *for j* = 1,2, ... . , n. Compute the standard deviation S(n)S(n)S(n) of the original segment xi: S(n) = √(1/n ∑_(j=1)^n⟦(xij-x i)⟧^2 ). Calculate the rescaled range R/S(n) by (R(n))/(S(n)) and plot log ((R/S)(n)) against log(n). The slope of the best-fit line through these points is an estimate of the Hurst exponent H which is log(R/S(n)) = H log(n) + C. Use linear regression to estimate H from the slope. Use linear regression to estimate H from the slope.

### *D. DESCRIPTION OF DATASET*

In the case of the dataset used for my analysis, the Network Simulator 3 (NS3) was used to run the simulations and obtain the data required thereof. NS3 is event-based simulator that is used for modeling and analysis of different networks and the protocols used in implementing the networks. This made it possible for me to collect good data through setting up NS3 to capture only the network scenarios that are of interest to my study. The various parameters of the simulation were chosen to match the conditions under which the behaviors of such signals and the anomalies were likely to manifest. From these simulations, net traffic data was obtained in detail as a form of the reconstructed signals on which basis the Hurst exponent was computed.

1) EXAMINATION OF NETWORK TRAFFIC DATA

For the purpose of the analysis, I have collected a dataset, where, for the purpose of network traffic measurement, I has used the NS3 simulation environment, which captures traffic from different points in the network as well as different network devices. The dataset consists of a sequence of trace sources from one or many nodes and/or devices, concentrating on the MacRx attribute of the Point to Point Net Device objects. The paths in the dataset are formatted as follows:

TABLE II: *SPECIFICATION OF NETWORK DATA*

| Node | Time Stamp | Packet Size |
|---|---|---|
| /NodeList/3/DeviceList/1/$ns3::PointToPointNetDevice/MacRx | 0.02208 | 48 |
| /NodeList/10/DeviceList/1/$ns3::PointToPointNetDevice/MacRx | 0.04416 | 48 |
| /NodeList/3/DeviceList/2/$ns3::PointToPointNetDevice/MacRx | 0.06622 | 40 |
| /NodeList/0/DeviceList/2/$ns3::PointToPointNetDevice/MacRx | 0.11036 | 40 |
| /NodeList/3/DeviceList/1/$ns3::PointToPointNetDevice/MacRx | 0.13242 | 40 |
| /NodeList/10/DeviceList/1/$ns3::PointToPointNetDevice/MacRx | 60.1364 | 40 |
| /NodeList/4/DeviceList/3/$ns3::PointToPointNetDevice/MacRx | 60.1364 | 1500 |
| /NodeList/4/DeviceList/4/$ns3::PointToPointNetDevice/MacRx | 60.1364 | 1500 |
| /NodeList/4/DeviceList/5/$ns3::PointToPointNetDevice/MacRx | 60.1365 | 1500 |
| /NodeList/4/DeviceList/3/$ns3::PointToPointNetDevice/MacRx | 60.1365 | 48 |
| /NodeList/3/DeviceList/4/$ns3::PointToPointNetDevice/MacRx | 60.1484 | 48 |
| /NodeList/3/DeviceList/5/$ns3::PointToPointNetDevice/MacRx | 60.1484 | 48 |
| /NodeList/3/DeviceList/3/$ns3::PointToPointNetDevice/MacRx | 60.1484 | 1500 |
| /NodeList/3/DeviceList/4/$ns3::PointToPointNetDevice/MacRx | 60.1485 | 1500 |
| /NodeList/3/DeviceList/5/$ns3::PointToPointNetDevice/MacRx | 60.1485 | 1500 |
| /NodeList/3/DeviceList/3/$ns3::PointToPointNetDevice/MacRx | 60.1485 | 48 |
| /NodeList/3/DeviceList/4/$ns3::PointToPointNetDevice/MacRx | 60.1604 | 48 |
| | 60.1604 | |
| /NodeList/8/DeviceList/5/$ns3::PointToPointNetDevice/MacRx | 60.1604 | 48 |
| /NodeList/8/DeviceList/3/$ns3::PointToPointNetDevice/MacRx | 60.16 | 1500 |

shows that there was more than one occurrence where packets reached the same device or general traits in packets arriving at similar devices in multiple cases in different simulation run or scenario. As compared to the older system of monitoring router traffic, this method benefits from the trace data being traced throughout the system, and as such results in a large amount of data

available for packet flow and reception analysis within the actual simulated network.

Through such parameters, I was able to watch and assess networking issues like the packet loss, undue delay, and other similar irregularities that are vital in the subsequent levels of my study. Through the differences in granularity of the trace sources, network interactions to the detail of MAC layer can indeed be captured to allow for comprehensive analysis of the network.

## IV. RESULTS AND DISCUSSION

In order to perform the experiments of the proposed methodology, all the experiments were done using NS3 simulator. The raw data traffic was collected from the simulator and calculations were performed. I used Pycharm to create a MPTCP module for the simulator to generate MPTCP traffic. Rest of the work was done by making simulation scenarios on NS3 simulator. The Hurst exponent calculation was also done using Pycharm.

### *A. THROUGHPUT OF UNATTACKED MPTCP NETWORK*

The following graph shows the throughput of MPTCP network before any attack is done on the network. The 600-second NS-3 simulation demonstrates an untainted MPTCP setup in Figure 3. The MPTCP setup starts with a throughput level of 6 Mbps which quickly increases through 200 seconds because the protocol establishes its multiple subflows and links these paths for data sharing. The graph shows that MPTCP congestion control approaches a stable state when the execution reaches 200 seconds because the cumulative throughput growth begins to level off. The network capacity approach saturation point results in this manifestation. The smoothness of throughput data starting from 300 seconds indicates that multipath scheduling achieves steady high throughput without sudden drops or peaks when it operates without attacks. The unhindered baseline will prove essential later when comparing against attack scenarios because any differences can be attributed precisely to anomalous behavior and not normal protocol operations.

### *B. IMF DECOMPOSITION OF NODE*

The original throughput is divided into multiple IMFs to decompose the signal in better scale. The classical Empirical Mode Decomposition (EMD) algorithm when applied to the Node 0 throughput data creates eleven intrinsic mode functions (IMF1 through IMF11) that extract frequency oscillations from high to low scales as shown in Figure 4. The fastest small-scaled fluctuations and packet-level jitter and measurement noise appear in IMF1 and IMF2 yet IMF3 through IMF7 show the transient throughput variations and emerging anomalous signals. The slow-varying baseline trend of MPTCP throughput exists within IMF8 to IMF11 with these intrinsic mode functions producing the long-term characteristics of multipath connection performance. Mode mixing which represents similar frequency contents that spread across separate IMFs can be recognized as multiple overlapping oscillatory patterns. Trivial mode interferences that standard EMD naturally produces often hide relevant anomalies which motivates researchers to use Ensemble EEMD for obtaining refined decompositions that reveal genuine anomalies.

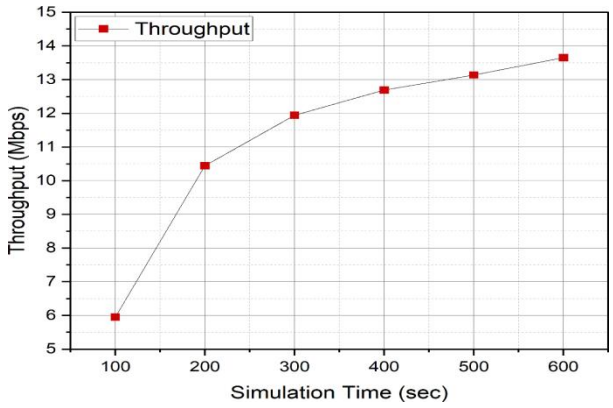



*Figure 3 Throughput of unattacked MPTCP network*

RECONSTRUCTED THROUGHPUT OF UNATTACKED NETWORK

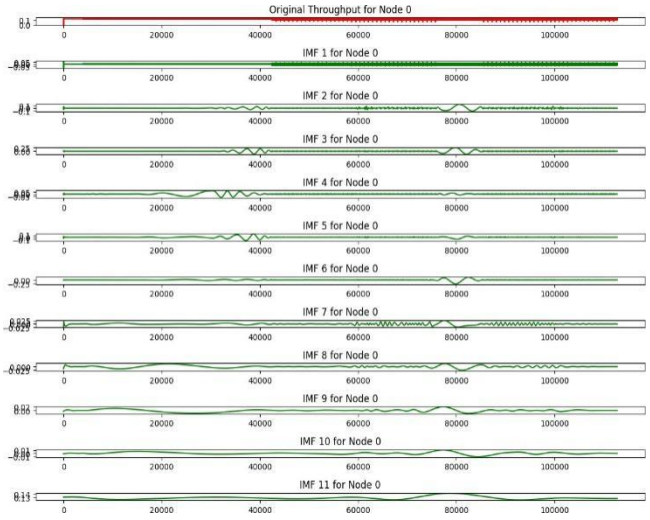



*Figure 4 IMF Decomposition without EEMD*

After the decomposition of IMFs, the signal is reconstructed into original form. An assessment of MPTCP network throughput evolution in time occurs without utilizing Ensemble Empirical Mode Decomposition (EEMD). The plot shows how transmission speed increases steadily during the simulation period which demonstrates generalized network data transfer enhancement. The signal fails to receive proper modal aliasing filtering because EEMD is missing from the analysis. EEMD produces inaccurate anomaly detection because it integrates noise or other anomalies from the original signal.

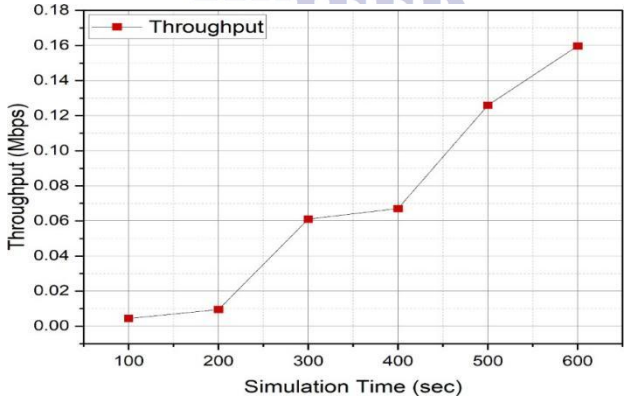



*Figure 5 Reconstructed signal throughput*

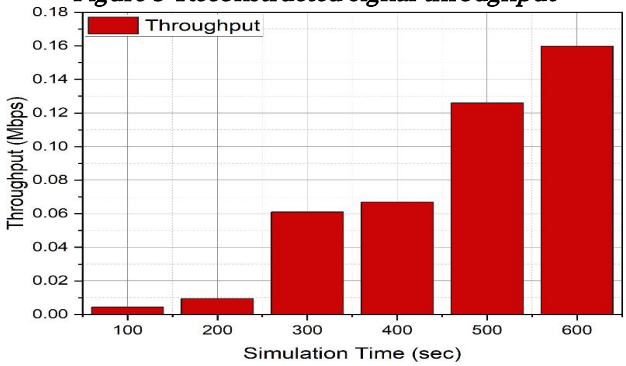



*Figure 6 Reconstructed Throughput Bar chart*

The above figure shows unattacked MPTCP network throughput values during reconstruction that occurred without applying Ensemble Empirical Mode Decomposition (EEMD). The simulated throughput shows major turbulent periods during the timeline through increased peaks and decreased valleys. The irregularities appear from modal aliasing when different signal components create reconstruction errors.

### *C. ATTACKED NETWORK THROUGHPUT OF MPTCP NETWORK*

Here, the MPTCP network is attacked under certain conditions and the throughput is generated to know the difference of network with un-attacked scenario. The depicted figure demonstrates how network throughput behaves during an attack which continued for 600 simulated seconds. Through simulation time in seconds the x-axis shows measured throughput expressed in Mbps. The simulation shows that network throughput experiences changing levels throughout the observed time which demonstrates the negative effect of the attack on network performance. Throughput measurements exhibit significant variation points during different periods because the attack method both increases and decreases transmission speeds.

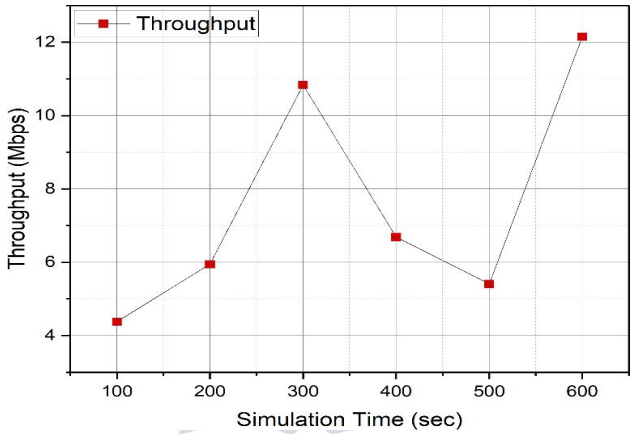



*Figure 7 Attacked network Throughput*

### *D. IMF Decomposition Applying EEMD*

The white noise is added to the network through mptcp module and then it is decomposed in multiple IMFs. White noise is used to cancel out the effect of anomalies that disguise themselves as original signal in the network. The results from applying Ensemble Empirical Mode Decomposition (EEMD) to Node 0 original throughput data appear in this figure. The graphical representation includes an original Node 0 throughput data along with subsequent IMF 1 through IMF 11 for Node 0. The lower portion of the figure contains eleven graphs displaying the Intrinsic Mode Functions (IMFs) obtained from the original throughput signal through EEMD processing. EEMD decomposes complex signals into multiple oscillatory intrinsic mode functions known as IMFs which represent the different frequency modes contained in the data. EEMD decomposition of attacked network throughput works to separate the network's fluctuation elements by frequency for close analysis which helps to distinguish attack indicators from normal performance. By studying individual IMFs regarding their frequency signatures and amplitude changes researchers can identify special frequency characteristics which connect to attack events and affect total throughput.

*RECONSTRUCTION OF SIGNAL*


**Figure 8 IMF's decomposition using EEMD**

The signal is, here, reconstructed after removing all the noise and high frequencies and residual terms, which saves the original properties of a signal. Figure represents the EEMD-processed Node 0 signal throughputs which underwent reconstruction from the attack period. Throughput appears on the y-axis while simulation time occupies the x-axis range from zero to three seconds. The reconstructed throughput serves to reproduce vital signal characteristics with an option to remove attack-related components from additional noise elements. Looking at reconstructed signal data enables users to evaluate both the exposure and detection capabilities of EEMD analysis regarding essential signal aspects within throughput data.

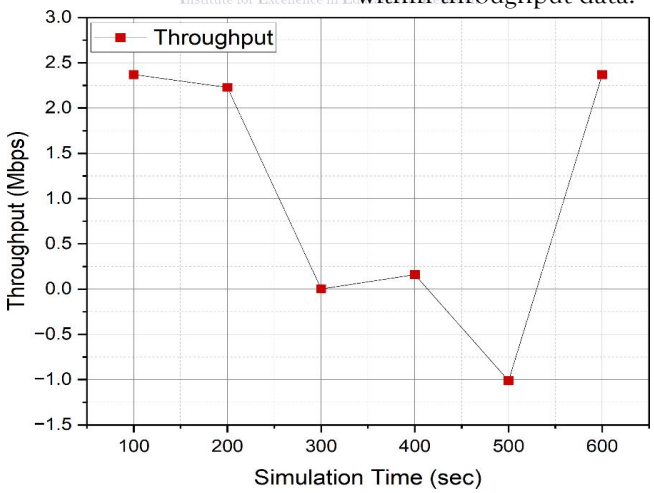



*Figure 9 Reconstructed Throughput of Attacked Signal*

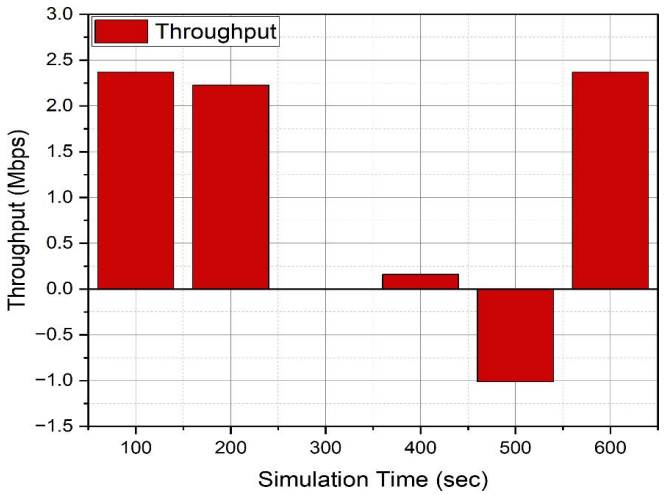



**Figure 10 Reconstructed Throughput of attacked Network Signal bar chart**

The bar chart presents graphical representation of the attacked network signal throughput reconstruction data. The bars display accumulated reconstructed throughput either for fixed time durations or distinct operational conditions. The vertical position of each bar shows how much the reconstructed throughput value measures while simultaneously enabling comparison of different throughput values between categories or time intervals. The bar chart displays relative heights

which helps users understand the network throughput change from the attack and assess reconstruction success in keeping significant signal attributes.

E. *HURST VALUE COMPARISON*

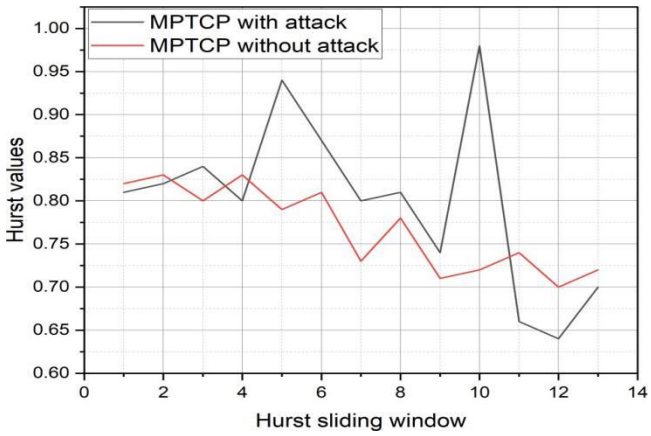



*Figure 11 Hurst Exponent Calculation Comparison of Attacked and Unattacked network*

To mention a specific term, I have used a sliding window to partition the MPTCP network traffic data into different overlap traffic segments. In each window, the Hurst exponent is calculated than the rescale range analysis (R/S analysis). It identifies whether the specific window possesses anomalies in it by the fluctuation in the MPTCP network's Hurst exponent under various attack scenarios. A graph displays Hurst exponent measurements between MPTCP network traffic which includes attack events (gray line) and MPTCP network traffic without attack incidents (red line). The Hurst sliding window spans the x-axis positions during which the Hurst exponent calculated each consecutive segment from time series data. These values appear on the y-axis of the graph to determine the extent of both long-range dependence and self-similar behavior in the traffic patterns. The statistical indicator called Hurst exponent provides analysis of time series data fractal characteristics. When the process has a Hurst value near 0.5 it resembles white noise which lacks extended dependences but Hurst values above 0.5 indicate persistent trends will probably extend into the future. An anti-persistent trend appears when the values fall below 0.5 because patterns tend to reverse directions. We can see through contrast between the gray line representing an attacked network and the red line depicting an unattacked network which type of MPTCP traffic followed long-range dependence after EEMD processing. Hurst values demonstrate different levels of traffic self-similarity throughout the sliding window analysis period. Audible discrepancies in the network traffic during particular time windows point to modifications made in the statistical properties by the attack. EEMD processing is likely applied before this analysis because it separates traffic into distinct frequencies which enables researchers to examine the Hurst exponent at each frequency level.

The visual presentation of the following bar chart depicts Hurst exponents between attacked (red) and unattacked (blue) traffic via sliding windows. The visual presentation illustrates attack impacts on traffic persistence through increasing bar indicators which represent strengthened dependence at those time intervals.

## V. COMPARISON OF EEMD MODERN TECHNIQUES

### *A. DEEP NEURAL NETWORKS VS EEMD*

1) In anomaly detection, deep neural networks need large training datasets and data computational power.

2) On the other hand, EEMD was able to give better signal decomposition and it does not necessarily need a large volume of training data (Kang & Kang, 2016).

### *B. SUPPORT VECTOR MACHINES VS EEMD*

1) It can be generalized that classification-based anomaly detection is more applicable about SVMs as classifiers.

2) It is found that EEMD is better suited for signal decomposition and has capability to analyze non-linear and non-stationary data (Natha, 2022b).

### *C. EMD VS EEMD*

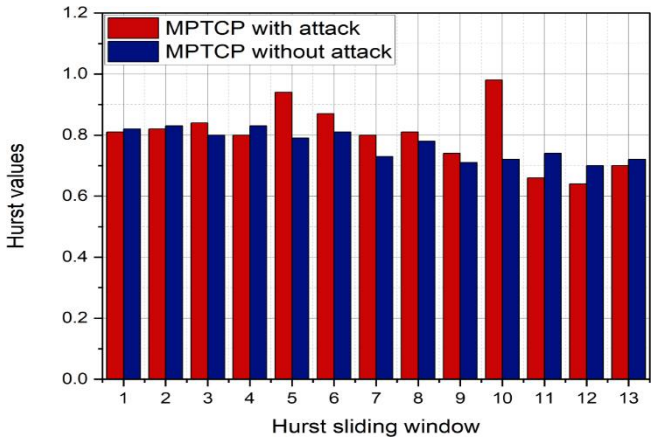



*Figure 12 Hurst Exponent Comparison Bar Chart*

1) EMD has the disadvantage that it mixes up its components in a process known as mode mixing
2) Relative to EMD, EEMD introduces white noise for obtaining better signal component separation and the elimination of modal aliasing. Specifically, EEMD affords more precise decomposition than fundamental EMD (Wu et al., 2009).

### *D. WAVELET TRANSFORM VS EEMD*

1) Compared to continuous wavelet transform, discrete wavelet transforms needs basis functions predefined.
2) EEMD decomposes the signals according to their properties as an adaptive process (Cao et al., 2022).

### *E. K-NEAREST NEIGHBORS VS EEMD*

1) While k-NN is simpler to implement and understand, its performance tends to deteriorate when faced with non-linear relationships.
2) Non-linear and non-stationary characteristics of network traffic is much better handled by EEMD (Gu & Peng, 2021).

This comparison proves that despite the effectiveness of applied techniques, EEMD has specific benefits for anomaly detection in network traffic analysis, notably for introducing the MPTCP networks where signals are non-linear and complex.

## VI. PERFORMANCE COMPARISON WITH EXISTING APPROACHES

A complete performance evaluation of the proposed EEMD-based anomaly detection approach occurred when researchers compared it against conventional EMD-based techniques. The research studied these performance indicators to determine key results using accuracy, precision, recall and false positive rate (FPR) and detection delay as metrics. Results prove that the EEMD-based method stands superior compared to traditional approaches specifically when noise contamination and modal overlaps degrade standard analysis methods. The proposed EEMD-based method maintained superior accuracy throughout the detection process for MPTCP network traffic anomalies including profound as well as minor deviations. The main drawback of EMD is its inability to correctly identify signal noise as system components whereas ensemble averaging in EEMD cleans noise thus improving both accuracy and signal clarity. EMD-based methods face an important disadvantage because modal aliasing affects signal cleaning by generating distinct signal modes which overlap one another. The improper decomposition of signals results in ambiguous detection of actual anomalies because of the compromised quality. The EEMD method applies white noise multiple times to different trials until it reaches sufficient averaging thus removing modal aliasing and producing distinct Intrinsic Mode Functions (IMFs). The detection reliability improves because of these enhancements to anomaly identification. Application of the EEMD method outperformed EMD through increased precision so it detected fewer incorrect anomalies in the system. The method improved its ability to

remember events because it captured all brief anomalies that previous versions failed to detect because they disregarded minor traffic changes. The proposed method shows high effectiveness since it adjusts to both non-linear and non-stationary signal patterns often found during MPTCP network operations. Models based on EMD technology need linear signal responses as a prerequisite but such assumptions limit their ability to generalize when used in real network deployments. EEMD addresses the known drawback by offering expansibility to various network topologies thus making it a practical scalable option.

| Metric | EMD-Based Method | Proposed EEMD-Based Method |
|---|---|---|
| Accuracy (%) | 86.7 | 94.6 |
| Precision (%) | 81.5 | 92.3 |
| Recall (%) | 83.2 | 95.1 |
| False Positive Rate (FPR %) | 13.3 | 7.1 |

## VII. CONCLUSION

The usage of internet for communication has spread worldwide over the network. MPTCP network is extremely popular in era of networking. With increasing usage of networks, anomalies have also strengthened to disguise themselves in the communication and forge the information. Several techniques have been produced to detect anomalies in a better way but the recent problem was detect as modal aliasing which allows anomaly signals to mix in original signals. In this thesis, I concerned with the issue of modal aliasing in anomaly detection in Multipath TCP (MPTCP) networks that can pose as a major threat to the effectiveness of the anomaly detection processes. To deal with this problem, I adopted the Ensemble Empirical Mode Decomposition (EEMD) method since this gives better decomposition of the signals into inherent mode functions than the other methods. It was ascertained that the use of EEMD was efficient in reducing modal aliasing and improving overall precision and reliability of the identified anomalies in the given MPTCP networks. After the elaborate testing and analyzing, the following conclusions were made; EEMD helps decreasing the modal aliasing and meanwhile preserves the original characteristics of the signal. This enhanced performance in the anomaly detection was well demonstrated by a better identification of some of the changes that occurred in the network traffics, which is essential to MPTCP. Besides, integration of the Hurst exponent with EEMD was able to offer more information about the self-similarity of the network traffic assisting in better anomaly detection.

## VIII. DISCUSSIONS AND FINDINGS

This research shows that EEMD proves to be an excellent approach which succeeds in resolving modal aliasing while improving anomaly detection capabilities in Multipath TCP (MPTCP) networks. The combination of traditional EMD with its EMD-based techniques implemented through Empirical mode decomposition historically produced modal aliasing that makes it difficult to detect anomalies precisely in non-stationary network traffic. The research confirmed that EEMD produces better signal breakdown which leads to enhanced detection of concealed and subtle signal anomalies. The proposed method demonstrates an important strength because it preserves the natural signal properties and mitigates noise and interference between frequency modes. EMD generates unreliable and non-distinct Intrinsic Mode Functions from standard EMD because of its noise-sensitivity which produces artificial modes and incorrect anomaly readings. In contrast EEMD applies ensemble averaging to create reliable and distinct Intrinsic Mode Functions. The method provides an improved signal representation because it successfully demonstrates the true signal behavior which enhances the detectability of anomalies. Research analysis executed during the study demonstrates that the proposed method works effectively. The EEMD-based technique yielded superior performance than standard EMD-based approaches in all vital assessment metrics including accuracy together with precision and recall which proves its potential as an operational network monitoring tool. Traditional methods are unable to detect short-duration low-amplitude anomalies yet the

presented method excels at this task. EEMD presents important properties toward accommodating signal characteristics that show non-linearity and non-stationarity which frequently appear in MPTCP traffic networks Through its integration with MPTCP simulations in NS-3 the evaluation process gained realism by using a testing framework that models real-world networking environments. The results become more reliable and reproducible through this approach which functions as a base for developing future research. The EEMD-based anomaly detection method shows significant progress through its superiority to existing approaches because it reduces aliasing and enhances sensitivity and operates well under noisy conditions. The confirmed core research hypothesis alongside future investigation possibilities in signal-based network system anomaly detection emerges from these findings.

## IX. LIMITATIONS AND FUTURE RECOMMENDATIONS

The EEMD-based technique detects anomalies better within MPTTC networks but faces difficulties in executing this process. The necessary computational complexity and processing time increase due to EEMD implementation because it requires multiple ensemble iterations with added noise. Real-time applications alongside environments with limited resources experience performance challenges because of this factor. Selecting the correct white noise amplitude threshold with ensemble trial number remains difficult for EEMD because these specifications affect result consistency between different datasets. The evaluation scenario relies on NS-3 simulated MPTCP traffic but does not address the entire scope of true network assumptions. The proposed method remains untested for use in extensive large-scale real-world network systems. The research mainly explores signal decomposition together with statistical detection methods as its primary analytical approachThe paper does not examine the combination of advanced machine learning models such as neural networks and ensemble classifiers for enhancing anomaly classification performance. These research results show promise so we propose several ways to improve upon them. Future research should develop this EEMD-based method for live disease detection under changing processing conditions of edge or cloud systems. Our team will design smart methods to pick EEMD parameters automatically so the accuracy improves and users do not need to adjust them manually. The EEMD method when combined with deep learning systems helps creation of hybrid detection tools that identify anomalies more accurately and classify them correctly. The next research step will test the proposed technique in active MPTCP networks across widespread and mobile platforms to check its practical effectiveness. The use of optimized processing algorithms with parallel setup will help make EEMD energy-efficient for real-world IoT system environments.

## APPENDIX

Appendixes, if needed, appear before the acknowledgment.

## ACKNOWLEDGMENT

I first express gratitude to Allah Almighty because His abundant blessings accompanied by His wisdom and strength and guidance enabled me to reach this milestone. Throughout my academic journey my parents gave me their eternal support through prayers and stood by me without hesitation. I extend my heartfelt appreciation to my friends because their faith in me along with their continuous support helped me during tough times. The successful completion of this work hinges on the guidance together with valuable feedback and steady support from my supervisor. My life's journey gained its most crucial support from my life partner because their love always accompanied me through extraordinary patience and constant encouragement.

## REFERENCES

[1] K. Wójcicki, M. Biegańska, B. Paliwoda, and J. Górna, “Internet of Things in Industry: Research Profiling, Application, Challenges and Opportunities—A Review,” *Energies*, vol. 15, no. 5, p. 1806, Feb. 2022, doi: 10.3390/en15051806.

[2] X. Chen, C. Wu, Z. Liu, N. Zhang, and Y. Ji, “Computation Offloading in Beyond 5G Networks: A Distributed Learning Framework and Applications,” *IEEE Wirel. Commun.*, vol. 28, no. 2, pp. 56–62, Apr. 2021, doi: 10.1109/MWC.001.2000296.

[3] T. Zhang, S. Zhao, and B. Cheng, "Multipath Routing and MPTCP-Based Data Delivery Over Manets," *IEEE Access*, vol. 8, pp. 32652–32673, 2020, doi: 10.1109/ACCESS.2020.2974191.

[4] A. Munir, Z. Qian, Z. Shafiq, A. Liu, and F. Le, "Multipath TCP traffic diversion attacks and countermeasures," in *2017 IEEE 25th International Conference on Network Protocols (ICNP)*, Toronto, ON: IEEE, Oct. 2017, pp. 1–10. doi: 10.1109/ICNP.2017.8117547.

[5] L. Ji, G. Lei, R. Ji, Y. Cao, X. Shao, and X. Huang, "Which One is More Robust to Low-Rate DDoS Attacks? The Multipath TCP or The SCTP," in *Mobile Internet Security*, I. You, H. Kim, T.-Y. Youn, F. Palmieri, and I. Kotenko, Eds., Singapore: Springer Nature Singapore, 2022, pp. 323–334.

[6] G. Fernandes, J. J. P. C. Rodrigues, L. F. Carvalho, J. F. Al-Muhtadi, and M. L. Proença, "A comprehensive survey on network anomaly detection," *Telecommun. Syst.*, vol. 70, no. 3, pp. 447–489, Mar. 2019, doi: 10.1007/s11235-018-0475-8.

[7] J. Iden and T. R. Eikebrokk, "The impact of senior management involvement, organisational commitment and group efficacy on ITIL implementation benefits," *Inf. Syst. E-Bus. Manag.*, vol. 13, no. 3, pp. 527–552, Aug. 2015, doi: 10.1007/s10257-014-0253-4.

[8] R. K. Chaturvedi and S. Chand, "Multipath TCP security over different attacks," *Trans. Emerg. Telecommun. Technol.*, vol. 31, no. 9, p. e4081, Sep. 2020, doi: 10.1002/ett.4081.

[9] V. A. Kumar and D. Das, "Data sequence signal manipulation in multipath TCP (MPTCP): The vulnerability, attack and its detection," *Comput. Secur.*, vol. 103, p. 102180, Apr. 2021, doi: 10.1016/j.cose.2021.102180.

[10] A. Stallone, A. Cicone, and M. Materassi, "New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms," *Sci. Rep.*, vol. 10, no. 1, p. 15161, Sep. 2020, doi: 10.1038/s41598-020-72193-2.

[11] K. Van Es, "Netflix & Big Data: The Strategic Ambivalence of an Entertainment Company," *Telev. New Media*, vol. 24, no. 6, pp. 656–672, Sep. 2023, doi: 10.1177/15274764221125745.

[12] "NS-3." [Online]. Available: https://www.nsnam.org/

[13] Y. Cao, R. Ji, L. Ji, M. Bao, L. Tao, and W. Yang, "Can Multipath TCP Be Robust to Cyber Attacks? A Measuring Study of MPTCP with Active Queue Management Algorithms," *Secur. Commun. Netw.*, vol. 2021, pp. 1–11, May 2021, doi: 10.1155/2021/9963829.

[14] Y. Cao *et al.*, "Retracted: Anomaly detection with ensemble empirical mode decomposition and approximate entropy for quick user datagram protocol internet connection-based distributed Blockchain systems," *IET Softw.*, vol. 17, no. 4, pp. 742–754, Aug. 2023, doi: 10.1049/sfw2.12096.

[15] I. Mahmud, T. Lubna, G.-H. Kim, and Y.-Z. Cho, "BA-MPCUBIC: Bottleneck-Aware Multipath CUBIC for Multipath-TCP," *Sensors*, vol. 21, no. 18, p. 6289, Sep. 2021, doi: 10.3390/s21186289.

[16] C.-D. Phung, B. F. Silva, M. Nogueira, and S. Secci, "MPTCP robustness against large-scale man-in-the-middle attacks," *Comput. Netw.*, vol. 164, p. 106896, Dec. 2019, doi: 10.1016/j.comnet.2019.106896.

[17] Y. Cao, R. Ji, X. Huang, G. Lei, X. Shao, and I. You, "Empirical Mode Decomposition-empowered Network Traffic Anomaly Detection for Secure Multipath TCP Communications," *Mob. Netw. Appl.*, vol. 27, no. 6, pp. 2254–2263, Dec. 2022, doi: 10.1007/s11036-022-02005-6.

[18] Y. Cao, R. Ji, L. Ji, X. Shao, G. Lei, and H. Wang, "MPTCP-*me*Learning: A Multi-Expert Learning-Based MPTCP Extension to Enhance Multipathing Robustness against Network Attacks," *IEICE Trans. Inf. Syst.*, vol. E104.D, no. 11, pp. 1795–1804, Nov. 2021, doi: 10.1587/transinf.2021NGP0009.

[19] Y. Cao, J. Chen, Q. Liu, G. Lei, H. Wang, and I. You, "Can Multipath TCP be Robust to Cyber Attacks With Incomplete Information?," *IEEE Access*, vol. 8, pp.

165872–165883, 2020, doi: 10.1109/ACCESS.2020.3021475.

[20] D. Jiang, Z. Yuan, P. Zhang, L. Miao, and T. Zhu, "A traffic anomaly detection approach in communication networks for applications of multimedia medical devices," *Multimed. Tools Appl.*, vol. 75, no. 22, pp. 14281–14305, Nov. 2016, doi: 10.1007/s11042-016-3402-6.

[21] F. Palmieri and U. Fiore, "Network anomaly detection through nonlinear analysis," *Comput. Secur.*, vol. 29, no. 7, pp. 737–755, Oct. 2010, doi: 10.1016/j.cose.2010.05.002.

[22] G. Lei, L. Ji, R. Ji, Y. Cao, W. Yang, and H. Wang, "Can Wavelet Transform Detect LDDoS Abnormal Traffic in Multipath TCP Transmission System?," *Secur. Commun. Netw.*, vol. 2021, pp. 1–8, Dec. 2021, doi: 10.1155/2021/8066200.

[23] Jang-Ping Sheu, Lee-Wei Liu, R. Jagadeesha, and Yeh-Cheng Chang, "An efficient multipath routing algorithm for multipath TCP in Software-Defined Networks," in *2016 European Conference on Networks and Communications (EuCNC)*, Athens, Greece: IEEE, Jun. 2016, pp. 371–376. doi: 10.1109/EuCNC.2016.7561065.

[24] S. Natha, "A Systematic Review of Anomaly detection using Machine and Deep Learning Techniques," *Quaid-E-Awam Univ. Res. J. Eng. Sci. Technol.*, vol. 20, no. 1, pp. 83–94, Jun. 2022, doi: 10.52584/QRJ.2001.11.

[25] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Clust. Comput.*, vol. 22, no. S1, pp. 949–961, Jan. 2019, doi: 10.1007/s10586-017-1117-8.

[26] S. Natha, "A Systematic Review of Anomaly detection using Machine and Deep Learning Techniques," *Quaid-E-Awam Univ. Res. J. Eng. Sci. Technol.*, vol. 20, no. 1, pp. 83–94, Jun. 2022, doi: 10.52584/QRJ.2001.11.

[27] M.-J. Kang and J.-W. Kang, "Intrusion Detection System Using Deep Neural Network for In-Vehicle Network Security," *PLOS ONE*, vol. 11, no. 6, p. e0155781, Jun. 2016, doi: 10.1371/journal.pone.0155781.

[28] P. Mulinka and P. Casas, "Stream-based Machine Learning for Network Security and Anomaly Detection," in *Proceedings of the 2018 Workshop on Big Data Analytics and Machine Learning for Data Communication Networks*, Budapest Hungary: ACM, Aug. 2018, pp. 1–7. doi: 10.1145/3229607.3229612.

[29] G. Lei, L. Ji, R. Ji, Y. Cao, W. Yang, and H. Wang, "Can Wavelet Transform Detect LDDoS Abnormal Traffic in Multipath TCP Transmission System?," *Secur. Commun. Netw.*, vol. 2021, no. 1, p. 8066200, 2021, doi: 10.1155/2021/8066200.

[30] M. Jantti, H. Virkanen, J. Mykkaunen, and V. Hotti, "Exploring the role of IT service management and IT service governance within IT governance," in *2014 11th International Conference on Service Systems and Service Management (ICSSSM)*, Beijing: IEEE, Jun. 2014, pp. 1–6. doi: 10.1109/ICSSSM.2014.6874122.

[31] S. Du, Z. Xu, and J. Lv, "An EMD- and GRU-based hybrid network traffic prediction model with data reconstruction," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2021, pp. 1–7. doi: 10.1109/ICCWorkshops50388.2021.9473822.

[32] J. Han and J. Z. Zhang, "Network traffic anomaly detection using weighted self-similarity based on EMD," in *2013 Proceedings of IEEE Southeastcon*, Jacksonville, FL, USA: IEEE, Apr. 2013, pp. 1–5. doi: 10.1109/SECON.2013.6567395.

[33] T. Proehl, K. Erek, F. Limbach, and R. Zarnekow, "Topics and Applied Theories in IT Service Management," in *2013 46th Hawaii International Conference on System Sciences*, Wailea, HI, USA: IEEE, Jan. 2013, pp. 1367–1375. doi: 10.1109/HICSS.2013.555.

[34] H. Ren, Z. Ye, and Z. Li, "Anomaly detection based on a dynamic Markov model," *Inf. Sci.*, vol. 411, pp. 52–65, Oct. 2017, doi: 10.1016/j.ins.2017.05.021.

[35] Z. Wu, N. E. Huang, and X. Chen, "THE MULTI-DIMENSIONAL ENSEMBLE EMPIRICAL MODE DECOMPOSITION

METHOD," *Adv. Adapt. Data Anal.*, vol. 01, no. 03, pp. 339–372, Jul. 2009, doi: 10.1142/S1793536909000187.

[36] S.-C. Du, T. Liu, D.-L. Huang, and G.-L. Li, "An Optimal Ensemble Empirical Mode Decomposition Method for Vibration Signal Decomposition," *J. Vib. Acoust.*, vol. 139, no. 3, p. 031003, Jun. 2017, doi: 10.1115/1.4035480.

[37] J. Gu and Y. Peng, "An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis," *Digit. Signal Process.*, vol. 113, p. 103050, Jun. 2021, doi: 10.1016/j.dsp.2021.103050.

[38] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," *IEEE Access*, vol. 9, pp. 78658–78700, 2021, doi: 10.1109/ACCESS.2021.3083060.

[39] M. Santhosh, C. Venkaiah, and D. M. Vinod Kumar, "Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction," *Energy Convers. Manag.*, vol. 168, pp. 482–493, Jul. 2018, doi: 10.1016/j.enconman.2018.04.099.

[40] Y. Cao, J. Chen, Q. Liu, G. Lei, H. Wang, and I. You, "Can Multipath TCP be Robust to Cyber Attacks With Incomplete Information?," *IEEE Access*, vol. 8, pp. 165872–165883, 2020, doi: 10.1109/ACCESS.2020.3021475.