International Journal of Computational Methods and Experimental Measurements
https://www.acadlore.com/journals/IJCMEM

Prediction and Anomaly Detection Framework Using Graph Density Time Series for Cyber Network Data

Dua M. Ghadi[1,2*], Basad Al-Sarray[3,4]

[1] Department of Mathematics, College of Science, University of Baghdad, 10011 Baghdad, Iraq
[2] Wasit Education Directorate, Ministry of Education, 10011 Baghdad, Iraq
[3] Department of Computer Science, College of Science, University of Baghdad, 10011 Baghdad, Iraq
[4] Department of Big Data Analytics, College of Artificial Intelligent, University of Baghdad, 10011 Baghdad, Iraq

* Correspondence: Dua M. Ghadi (doaa.kathy2303@sc.uobaghdad.edu.iq)

Received: 04-24-2026 **Revised:** 06-12-2026 **Accepted:** 06-23-2026

Citation: D. M. Ghadi and B. Al-Sarray, "Prediction and anomaly detection framework using graph density time series for cyber network data," *Int. J. Comput. Methods Exp. Meas.*, vol. 14, no. 3, pp. 343–358, 2026. https://doi.org/10.56578/ijcmem140301.

Abstract: This study suggests a hybrid model of prediction and anomaly detection of dynamic network based on graph density time series. The main issue that is being tackled is that traditional linear models cannot explain non-linear structural shocks and volatility clustering that are facts in cyber network data. The methodology proposed implies turning network flows of the UNSW-NB15 dataset into dynamic graph snapshots, deriving graph density as a scalar measure, and stabilizing the series by converting it to log-returns. The existence of the "fat tails" and non-Gaussian shocks which cannot be detected using traditional statistical tools was verified by the use of advanced diagnostic tests, like Kurtosis and Jarque-Bera test. As a result, a hybrid model that was a combination of the autoregressive moving average (ARMA) and exponential generalized autoregressive conditional heteroscedasticity (EGARCH) was applied. This research used the selection of the ARMA (p, q)-EGARCH (u, v) model as the best specification in terms of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The result of the hybrid model had an accuracy with running time spent in predictive and anomaly detection. Compared with two different methods, the methodology of ARMA (p, q)-EGARCH (u, v) has demonstrated the highest level of anomaly detection with a decrease in time processing in prediction and detection processes. This paper shows that structural graph analysis with modeling can be used to increase the resilience and sensitivity of intrusion detection systems.

Keywords: Dynamic network; Graph density time series; ARMA-EGARCH model; Prediction; Volatility clustering; UNSW-NB15; Anomaly detection; Cyber network security

1 Introduction

Security of data systems and monitoring is an important challenge because of the increasing threats that face data security nowadays. As network data grows in complexity, cyberattacks have developed into advanced, changing threats with time. Anomaly network traffic modeling as dynamic graphs has grown as an effective research model to capture these dynamics [1]. The temporal fluctuations of network features serve as an essential indicator of system status, where anomalies often manifest as sudden structural shifts or statistical outliers. The balanced approach between mathematical rigor and computational efficiency is required to detect these anomalies. In cybersecurity data research, machine learning and deep learning have been commonly used in several recent research studies in the intrusion detection field and have undergone critical research these days. This section overviews some of the modern research, including the deep learning frameworks, the statistical time-series frameworks, and the dynamic networks research.

The recent research has placed much focus on neural network structures for intrusion detection. Hussien and Dhannoon [2] have suggested an anomaly detector based on deep neural networks (DNNs) and dropout to avoid overfitting. Their framework which was tested on the NSL-KDD data shows that the application of the concept of Bernoulli noise is effective in increasing the classification accuracy in attacks of more than two classes, such as the denial of service (DoS) and the Probe. Subhi et al. [3] have suggested a complicated hybrid method consisting

of Ribonucleic Acid (RNA) encoding with a ResNet50 model due to the need to enhance the detection of high-dimensional data by emphasizing the feature representation of network packets. Manaa et al. [4] created a distributed denial of service (DDoS) mitigation optimization algorithm in the fog layer, which is part of the Internet of Things (IoT). They show the need to minimize latency in dynamic settings, with Support Vector Machine (SVM) to be used in classification. Moreover, Abdulrahman and Ibrahem [5] responded to the fact that the traffic of a network is dynamic, and it is necessary to implement data stream classification, and they focus on the low computation time and real-time processing. Beyond deep learning approaches, several studies have explored statistical and network-based perspectives for anomaly detection. Peel and Clauset [6] analyzed structural change points in evolving networks, demonstrating the importance of topological dynamics for identifying fundamental shifts in interaction patterns. Furthermore, Soule et al. [7] showed that time-series modeling of traffic measurements can effectively reveal anomalies without heavy computational overhead, utilizing filtering techniques to distinguish normal behavior from deviations. Finally, Ranshous et al. [8] introduced a survey categorizing dynamic network anomaly detection methods, where these methods the graph-level event and change detection, which focuses on detecting structural shifts in global network features.

The superiority of sequential hybridization has been further validated by Purwanto et al. [9], who demonstrated that a hybrid model combining autoregressive integrated moving average (ARIMA) with linear trend components significantly outperforms standalone versions by effectively capturing multi-layered volatility. Furthermore, the efficacy of two-step composite modeling, modeling variance and mean sequentially, was validated by Mudhir [10]. By implementing a mixed autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity (ARMA-EGARCH) strategy to address heteroscedasticity, Muhammad established that partitioning signals into distinct modeling phases is a scientifically sound approach for enhancing predictive fidelity in complex datasets.

According to the surveyed literature, it is found that the research on intrusion detection typically takes two different streams. While one stream concerns itself with graph-based analysis for capturing the evolution of the network structure, another stream uses statistical forecasting techniques for studying the temporal dynamics and volatility of the network. The relatively less emphasis has been placed upon developing a unified approach of incorporating dynamic graphs along with volatility-oriented statistical forecasting for monitoring cyber networks. This is because conventional studies based on graphs seldom consider the possibility of mapping graph-based indicators to time-series and then modeling them through hybrid ARMA–EGARCH modeling.

However, a crucial challenge in using a hybrid ARMA-EGARCH model is the selection of the optimal model and its employment in cybersecurity data. Proposes an automated approach that uses the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [11] to select the best model, and the main aim of this research is to present a hybrid statistical framework that can enhance predictive capabilities and monitor significant volatilities in network dynamic data and classify them as anomalies.

The previous research efforts on the detection of intrusion focus on machine learning methods and deep neural models, this study proposes a statistical graph model that considers both the dynamics and behavior of variance exhibited by cyber networks.

The main contributions of this research are summarized as follows:

• The dynamic network representation is generated by converting the UNSW-NB15 dataset into temporal network snapshots and extracting the graph density.

• The graph density time series framework is proposed, which is making the high dimensionality of the adjacency matrices to be represented as a univariate time series suitable for prediction and anomaly analysis.

• The hybrid ARMA–EGARCH model strategy is presented to capture temporal dependencies and asymmetric volatility clustering in time series, which allows in predicting values and detecting behavioral anomalies in the dynamic network structures.

• The dynamic statistical envelope is constructed, enabling adaptive anomaly detection instead of fixed threshold monitoring.

This research after this section, organized as follows: Section 2 presents the problem formulation; Section 3, details the hybrid methodology; Section 4 presents the describes and represents the datasets; Section 5 presents the experimental results; and finally, Section 6 is the conclusion.

2 Problem Formulation

The current security dynamic networks such as communication networks represent one of the most challenging nowadays. Securing these networks through structural analysis of a specific behavior for prediction and high sensitivity anomaly detection. The time series of the network traffic for each snapshot of the network such as a graph density time series (G_ρ) represents the essential connectivity patterns. The creation of a monitoring method is complicated because of the volatility clustering and non-stationarity established in the UNSW-NB15 dataset. By Using MATLAB R2024a, our initial diagnostic tests of the squared log-returns series r_t^2 of the data refers a relative

stability in variance that is unacceptable to has a conditional heteroscedasticity as appears in Table 1 of the Ljung-Box Q-test results demonstrated (h = *false*, $p > 0.05$).

Table 1. Ljung-Box Q-test results of squared log-returns series r_t^2

Lag	h	p-Value	Q-Test	Critical Value
1	false	0.12503	2.35318	3.84146
2	false	0.08401	4.95375	5.99146
3	false	0.14307	5.42696	7.81473
4	false	0.24477	5.44315	9.48773
5	false	0.26242	6.47826	11.07050
6	false	0.33542	6.84504	12.59159
7	false	0.43111	6.97863	14.06714
8	false	0.53699	6.99670	15.50731
9	false	0.62666	7.10045	16.91898
10	false	0.67292	7.54772	18.30704
11	false	0.74815	7.60548	19.67514
12	false	0.79161	7.91665	21.02607
13	false	0.82986	8.20693	22.36203
14	false	0.85987	8.53109	23.68479
15	false	0.34454	16.58068	24.99579

Although, the diagnostic test on the squared log-returns series first showed that there was relative stability in variance (h = *false*), it turned out that finding it was still misleading since it failed to capture the localized, high intensity fluctuations when it was further diagnosed using Advanced Distributional Analysis (Table 2 and Figure 1), which showed that the excess kurtosis was very high (7.3312) and the Jarque-Bera null hypothesis was significantly rejected ($p < 0.001$).

Table 2. Distributional analysis of returns before exponential generalized autoregressive conditional heteroscedasticity (EGARCH) modeling

Metric	Value	Standard Value	p-Value
Kurtosis	7.3312	3	–
Jarque-Bera	81.851	0.05	<0.001

Note: (–) = not applicable.

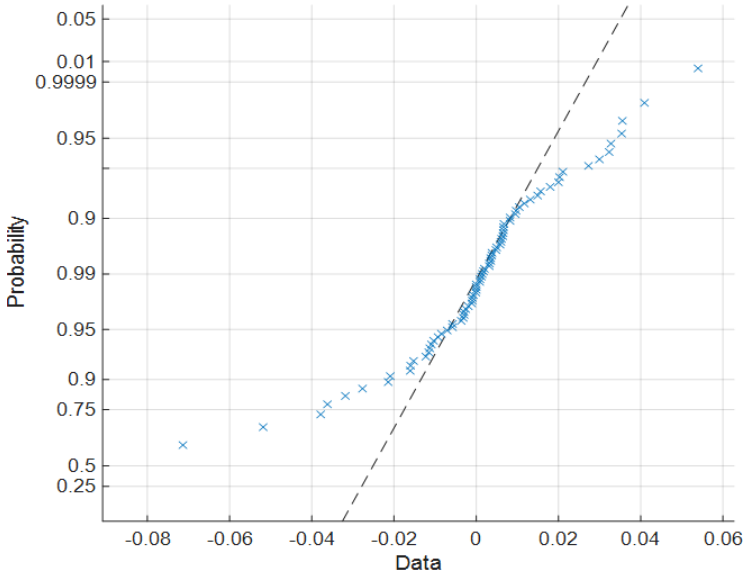

Figure 1. Probability plot (evidence of heavy tails/anomalies)

These results are conclusive statistical data of the existence of fat tails, showing that the network snapshots

are experiencing extreme non-Gaussian structural shocks. The conventional predictive theories tend to use the homoscedasticity hypothesis, which does not reflect the non-linear nature of series with great fluctuations [10]. It means that any successful security model should address the prediction problem first, i.e., by modeling the time varying mean over time and the asymmetric volatility of the network structure correctly. Traditional linear models fail in this predictive task since they cannot explain the shocks in network stability. Consequently, they produce large prediction errors that directly affect the subsequent anomaly detection phase. In security context, large fluctuations reflect structural conflict between normal traffic and attack patterns. Such asymmetric shocks are appropriately modeled by a modeling framework such as EGARCH which is statistically justified and necessary in this context to hold these fat-tailed behavior and security-driven fluctuations. The predictive accuracy of any anomaly detection system is the key to its effectiveness. In the UNSW-NB15 dynamic network dataset, the detection threshold will not be reliable when the prediction model is not capable of producing a low error rate. This requires a resilient security boundary, implemented as a dynamic statistical envelope that adapts to the changing state of the network. The proposed framework provides precise prediction of the future density as a baseline, and detects outliers using a dynamic statistical envelope. An observation is flagged only when it falls outside the K-sigma confidence interval, which dynamically widens and narrows according to the anticipated volatility.

3 Proposed Methodology

The proposed framework is based on a multi-step statistical pipeline that is meant to capture the non-linear, multi-stage dynamics of network stability. This methodology is a combination of an ARMA model to predict the mean and an EGARCH model for asymmetric volatility in order to result in an enhancement model for prediction and anomaly detection, as described in the Figure 2.

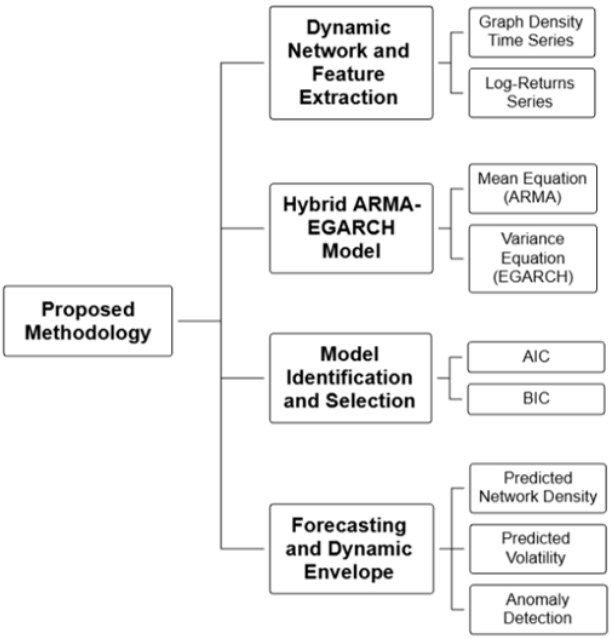

Figure 2. Flowchart of the proposed methodology

Note: ARMA = autoregressive moving average; EGARCH = exponential generalized autoregressive conditional heteroscedasticity; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

3.1 Dynamic Network and Feature Extraction

A dynamic network can be represented as separable networks of discrete-time, ordered sequences of graph snapshots $G_t = \{G_1, G_2, \ldots, G_T\}$, $t = 1, 2, \ldots, T$. Each $G_t = (V, E_t, A_t)$ represents the network at time t, with V as the set of nodes (vertices), E_t as the set of active edges (interaction) during the interval time t, and A_t as the adjacency matrix. Depending on the nature of the interactions, the network can be represented as either an undirected or a directed graph, which is the adjacency matrix A_t is symmetric ($A_{ij} = A_{ji}$), where $A_{(ij,t)} = 1$ indicating a link (edge) between nodes i and j, while $A_{(ij,t)} = 0$ otherwise, and the other type is the directed graph, A_t represents non-symmetric ($A_{ij} \neq A_{ji}$) which is considered better suited for most cybersecurity contexts (e.g., packet flows or financial transfers, etc.), where the network is modeled as a directed graph [12]. These snapshots are translated into

graph density time series to quantify the structural connectivity over time using the following computation [13]:

$$G_{\rho,t} = \frac{2\,|E_t|}{|V|(|V|-1)} \tag{1}$$

To stabilize the series for heteroscedasticity analysis, and to be able to capture the dynamic change rate of connectivity, $G_{\rho,t}$ is converted into a log-returns series (r_t) :

$$r_t = \ln\left(G_{\rho,t}\right) - \ln\left(G_{\rho,t-1}\right) \tag{2}$$

The transformation quantifies the relative change in structural connectivity between successive time steps, which is critical in the modeling of the volatility of the network environment. Figure 3 illustrates the r_t series of $G_{\rho,t}$ whereas Figure 4 presents both the original graph density series and the corresponding log-returns series.

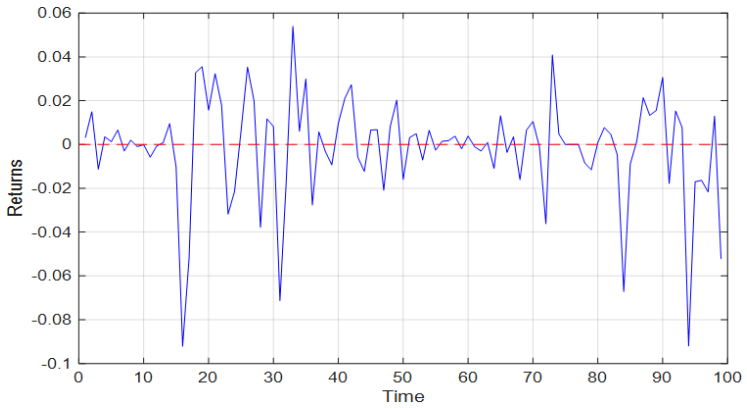

Figure 3. Plot of the return series

Figure 4. Raw of graph density series

In this research, graph density was chosen as the main structural descriptor for generating the time series; however, dynamic networks may also exhibit other localized variations in structure. In this regard, the possibilities of incorporating complementary graph features, including the average degree and clustering coefficient, to describe topological dynamics, which are left open for future expansions of the methods.

3.2 The Hybrid ARMA-EGARCH Model

Dynamic network analysis is based on a series of network snapshots defined by summary statistics or features, which change over time. A high-dimensional network is converted into a univariate time series using a scalar measure

to capture its global structure properties. This transformation allows the stochastic process modeling to analyze, forecast network behavior while identify anomalies. To capture the linear dependencies and non-linear asymmetric volatility inherent in network traffic, the hybrid ARMA (p, q)-EGARCH (u, v) model is used and is described by the following equations [10]:

The Mean Equation is considered the conditional mean of r_t is modeled using ARMA (p, q) to capture linear temporal dependencies:

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \varepsilon_t \sim N\left(0, \sigma^2\right) \tag{3}$$

Eq. (3) can be written by using lag operator as follows:

$$\phi(L) r_t = c + \theta(L) \varepsilon_t$$

Such that

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$
$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

In order to make the ARMA process stationary and invertible, all roots of the polynomials $\phi(L)$ and $\theta(L)$ should lie outside the unit circle. Although ε_t is typically assumed to be distributed as a Gaussian, our empirical diagnostics (Table 2) show that the Excess Kurtosis (7.33) is large and the Jarque-Bera test ($p < 0.001$) rejects the null hypothesis of normality. This proves that network shocks are characterized by fat tails, behavior commonly observed in security-related anomalies. The model therefore employs the Quasi-Maximum likelihood (QML) model, which remains consistent under non-Gaussian disturbances and mild regularity conditions, as reported in the Quasi-Maximum likelihood estimator (QMLE) literature [14, 15].

The EGARCH model, developed by Nelson [16], was used to model the non-linear shocks and volatility clustering observed in the network snapshots by estimating the conditional variance σ_t^2. The EGARCH (u, v) model, in contrast to the usual generalized autoregressive conditional heteroscedasticity (GARCH), takes into account the asymmetry in network stability, and is defined by the following variance equation [10, 17].

$$\varepsilon_t = \sigma_t z_t, z_t \sim N(0,1)$$
$$\ln \sigma_t^2 = w + \sum_{i=1}^{u} \beta_i \ln \sigma_{t-i}^2 + \sum_{j=1}^{v} \gamma_j \left(\frac{\varepsilon_{t-j}}{\sigma_{t-j}}\right) + \sum_{j=1}^{v} \alpha_j \left(\left|\frac{\varepsilon_{t-j}}{\sigma_{t-j}}\right| - E\left|\frac{\varepsilon_{t-j}}{\sigma_{t-j}}\right|\right) \tag{4}$$

where, w is the constant and α_j denotes the magnitude effect, which is measuring the response of volatility to the size of the shock (autoregressive conditional heteroscedasticity (ARCH) effect), β_i denotes volatility persistence of GARCH effect, indicating how long the impact of a network shock lasts, and γ_j asymmetry (leverage effect). The stationary requirement of the EGARCH model is expressed as $\sum_{i=1}^{u} \beta_i < 1$.

This hybrid structure enables the model to estimate the expected future network density while simultaneously predicting the conditional volatility necessary to build the dynamic statistical envelope for anomaly detection.

3.3 Model Identification and Selection

The mathematical framework of the model depends on applying Box-Jenkins modeling. Specifically, after converting the dynamic network into a univariate time series, next step is to determine the best order of ARMA (p, q)-EGARCH (u, v) by employing AIC and BIC. The model achieving the minimum value of AIC and BIC is selected as the optimal model, defined as:

$$AIC = -2\ln(\hat{\mathcal{L}}) + 2k \tag{5}$$

$$BIC = -2\ln(\hat{\mathcal{L}}) + 2k\ln(N) \tag{6}$$

where, $\hat{\mathcal{L}}$ represents the value of likelihood estimation, N is the sample size, and k is the number of the model's parameters in the hybrid ARMA (p, q)-EGARCH (u, v). The AIC tends to put more importance on predictive accuracy, BIC puts a heavier burden on model complexity, thus, making parsimonious specifications and minimizing the risk of over-parameterization. Using the systematic selection process and the grid search findings (refer to Table 3), the ARMA (p^*, q^*)-EGARCH (u^*, v^*) model is revealed as the best specification. This model attains a good balance among the fit of the model and model complexity to give a strong background on which it is possible to capture the dynamics of the network and that of the dynamic statistical envelope for anomaly detection.

Table 3. Description of some symbols

Symbol	Description
G_t	Sequences of graph snapshots
A_t	Adjacency matrix
V	Set of the nodes (vertices)
E_t	Set of the active edges
$G_{\rho,t}$	Graph density series
r_t	Log-returns series
ϕ_i	Autoregressive (AR) coefficients
θ_j	Moving average (MA) coefficients
σ_t^2	Conditional variance
γ_j	Leverage effect
β_i	Volatility persistence of GARCH effect
α_j	The magnitude effect
$\widehat{r}_{t+h}$	Forecast of the log-returns value
$\widehat{G}_{\rho,t+h}$	Forecast of the graph density series value
AD_t	Anomaly detection at time t

3.4 Forecasting and Dynamic Envelope for Anomaly Detection

After the optimal ARMA (p^*, q^*)-EGARCH (u^*, v^*) model is identified and its parameters are estimated, the framework proceeds to the forecasting phase. Unlike traditional models that predict only the conditional mean, The proposed approach provides a two-fold forecast by predicting both the network density and its conditional volatility (conditional variance). The conditional mean equation is then used to obtain the h-step-ahead forecast of the log-returns $\hat{r}_{t+h}$. To reconstruct the original network density scale, the predicted values are recursively transformed.

$$\widehat{G}_{\rho,t+h} = \widehat{G}_{\rho,t+h-1} \cdot e^{\hat{r}_{t+h}} \tag{7}$$

This transformation guarantees the predicted traffic patterns maintain the network density metrics. At the same time, EGARCH model predicts the conditional variance $\hat{\sigma}^2_{t+h}$. This represents the "uncertainty" or "instability" expected in the network at time t. When there is a security event or a traffic burst, $\hat{\sigma}_t$ rises, indicating the asymmetric shocks of the leverage effect (γ_j). Importantly, asymmetric reaction to shocks on networks is naturally included in the volatility forecasts with the EGARCH leverage term (γ_j). This enables the model to respond to positive and negative structural changes differently; this increases the sensitivity of the dynamic envelope to abnormal network behaviors. The dynamic statistical envelope is built on the original scale of the network density by a volatility-adjusted exponential transformation:

$$\text{Upper}_{\ t} = \widehat{G}_{\rho,t} * e^{K \cdot \hat{\sigma}_t} \tag{8}$$

$$\text{Lower}_{\ t} = \hat{G}_{\rho,t} * e^{-K \cdot \hat{\sigma}_t} \tag{9}$$

where, $\hat{G}_{\rho,t}$ denotes the forecast value of the network density obtained from ARMA (p^*, q^*)-EGARCH (u^*, v^*) model and $\hat{\sigma}_t$ is the forecasted conditional standard deviation. This process is consequence of the log-return representation, which has the network density changing multiplicatively as $G_{\rho,t} = G_{\rho,t-1}e^{r_t}$. Therefore, the volatility in the exponent guarantee's consistency between the forecasting model and the creation of the dynamic envelope, with preserving the multiplicative structure of network evolution.

Although the Gaussian assumption is violated (as confirmed by the Jarque-Bera test), the scaling factor, $K = 3$, has been chosen to be employed as a heuristic threshold that is strong and empirically reflects the empirical distribution of standardized residuals.

The process of the anomaly detection (AD_t) is unsupervised, meaning the flagged is raised at time t when measured network density $G_{\rho,t}$ observes a density not within the dynamic range:

$$AD_t = \begin{cases} 1, & \text{if } G_{\rho,t} > \text{ Upper}_{\ t} \text{ or } G_{\rho,t} < \text{ Lower}_{\ t} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where, $G_{\rho,t}$ represents the observed value of network density. This reasoning finds Structural Anomalies, which are major violations of the learnt temporal and heteroscedastic dependencies of the network.

4 Data Description and Representation

All experiments in this research were performed using MATLAB R2024a. This research uses the UNSW-NB15 dataset [18], one of the most common intrusion detection benchmarks, which includes normal and malicious network traffic. The data consists of both numerical and categorical values, and binary and multi-class. The original training and testing subsets are combined to create a single dataset to provide consistency. Categorical attributes are converts into numerical representations employing label encoding. A two-step normalization approach is employed, i.e. min-max scaling of all features is performed, followed by an 80/20 train-test split. The training data are then normalized using Z-score standardization with the same parameters applied to normalize the testing data to avoid data leakage; this preprocessing was adapted to MATLAB based on the framework developed by Rahman [19].

The normalized training data are divided into fixed-size snapshots (2048 flows each). In each snapshot, a line graph is created in which each flow is considered a node and edges will be created between flows with a common source or destination IP address. This model is inspired by dynamic graph-based models like DIGNN-A [20]. This graph representation was selected as it is more flow-centric, rather than the static host level connectivity and allows for a focus on behavioral dynamics at the interaction level. Typically, a time of events relationship can be observed between network flows rather than the scanning of individual IPs, communication with other IPs, moving laterally, and the launching of a distributed attack in the context of malicious activity in a cybersecurity environment. It flows as nodes allowing the method to maintain the fine-grained pattern of communication and temporal correlation among traffic events. The edges created based on common source or destination IP addresses enable the graph to model coordinated or repeated communication patterns that are typically considered intrusion propagation or unusual traffic changes.

This approach is consistent with recent graph intrusion detection dynamic frameworks [20] that also models the temporal dependency in cybersecurity networks, based on interaction-driven graph-based structures. This interaction-based graph formulation enables to represent temporal dynamics in coordinated communications and structural connectivity patterns in cybersecurity traffic by generating the graph density series.

It is applied with an incidence matrix S_t and the adjacency matrix is calculated as $A_t = S_t S_t^\top$. Subsequently, self-loop is removed from A_t to ensure a valid and simple graph structure for each snapshot. Each snapshot is denoted by its adjacency matrix, after finishing this process, the UNSW_NB15 dataset's each snapshot is denoted by its adjacency matrix, containing 100 undirected network snapshots with 2048 nodes and a total of 201,857,695 edges across all snapshots. A sliding window is used to group sequences of 10 consecutive snapshots into dynamic samples in order to capture temporal dependencies, where the labels are assigned with respect to the final snapshot. Subsequently, each snapshot is then extracted to produce structural graph measures, such as graph density, to create a time-series description of network evolution.

5 Discussion and Results

This section presents the experimental results of the ARMA (p^*, q^*)-EGARCH (u^*, v^*) model and discusses their statistical and practical consequences for dynamic network security analysis. As discussed in Section 3, the initial problem statement, especially, is the existence of non-stationarity, heavy-tailed behavior, and the clustering of volatility in the UNSW-NB15 dataset. This section, discusses the results of the model selection, diagnostic validation, and the appropriateness of the proposed methodology. The model is evaluated in terms of its effectiveness in terms of forecasting accuracy and supporting a dynamic statistical envelope to detect anomalies.

5.1 Justification for Model Selection and Mathematical Representation

By searching across the different orders, the ARMA (p^*, q^*)-EGARCH (u^*, v^*) model was found as the best fitting, with the lowest values AIC of BIC as shown in Table 4 which presents the most 15 best models in this project.

Based on sensitivity analysis, the chosen ARMA (4, 3)-EGARCH (2, 2) model was still the best overall model. It provided the lowest AIC/BIC values as shown in Table 4 and the minimum Root Mean Squared Error (RMSE) (0.0309), whereas the marginally smaller Mean Absolute Error (MAE) difference of the model from the alternative best model as shown in Table 5. It follows that the model choice is reliable and robust to small changes in parameters.

As shown in Table 4, the best model with minimum values of AIC/BIC was ARMA (4, 3)-EGARCH (2, 2) model and its formula will be as follows:

Mean Equation:

$$r_t = -0.0005 + 0.9477 r_{t-1} - 1.2850 r_{t-2} + 0.7855 r_{t-3} - 0.4487 r_{t-4} \\ + \varepsilon_t - 1.0326 \varepsilon_{t-1} + 1.0385 \varepsilon_{t-2} - 0.6437 \varepsilon_{t-3}$$

Variance Equation:

$$\ln \sigma_t^2 = -1.7981 + 1.0114 \ln \sigma_{t-1}^2 - 0.1913 \ln \sigma_{t-2}^2 - 0.2459 \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) + \left(\frac{\varepsilon_{t-2}}{\sigma_{t-2}} \right)$$
$$+0.5308 \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - E \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| \right) + \left(\left| \frac{\varepsilon_{t-2}}{\sigma_{t-2}} \right| - E \left| \frac{\varepsilon_{t-2}}{\sigma_{t-2}} \right| \right)$$

Table 4. The model selection of the 15 best ARMA (p^*, q^*)-EGARCH (u^*, v^*) models and values of AIC and BIC

Model	AIC	BIC
ARMA (0, 2)-EGARCH (2, 1)	-428.999	-412.413
ARMA (0, 4)-EGARCH (2, 1)	-431.811	-410.486
ARMA (1, 2)-EGARCH (1, 1)	-429.432	-412.846
ARMA (1, 4)-EGARCH (2, 1)	-427.876	-404.181
ARMA (1, 4)-EGARCH (2, 2)	-438.927	-412.863
ARMA (2, 3)-EGARCH (2, 2)	-428.996	-402.932
ARMA (2, 4)-EGARCH (2, 1)	-433.163	-407.099
ARMA (2, 4)-EGARCH (2, 2)	-441.227	-412.793
ARMA (3, 3)-EGARCH (2, 2)	-436.239	-407.805
ARMA (3, 4)-EGARCH (2, 1)	-428.910	-400.477
ARMA (3, 4)-EGARCH (2, 2)	-445.240	-414.437
ARMA (4, 3)-EGARCH (2, 2)	-451.013	-420.210
ARMA (4, 4)-EGARCH (1, 2)	-431.421	-400.619
ARMA (4, 4)-EGARCH (2, 1)	-427.609	-396.807
ARMA (4, 4)-EGARCH (2, 2)	-443.088	-409.915

Note: ARMA = autoregressive moving average; EGARCH = exponential generalized autoregressive conditional heteroscedasticity; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

Table 5. Parameter sensitivity analysis of the 15 best ARMA (p^*, q^*)-EGARCH (u^*, v^*) models based on RMSE and MAE

Model	RMSE	MAE
ARMA (4, 3)-EGARCH (2, 2)	0.030923619	0.021113482
ARMA (3, 4)-EGARCH (2, 2)	0.031777419	0.020939513
ARMA (4, 4)-EGARCH (2, 2)	0.031472581	0.020767893
ARMA (2, 4)-EGARCH (2, 2)	0.03093976	0.02074136
ARMA (1, 4)-EGARCH (2, 2)	0.031200828	0.021266759
ARMA (3, 3)-EGARCH (2, 2)	0.030975655	0.021424234
ARMA (2, 4)-EGARCH (2, 1)	0.031526604	0.021624556
ARMA (0, 4)-EGARCH (2, 1)	0.031463613	0.021582033
ARMA (4, 4)-EGARCH (1, 2)	0.031247178	0.02157273
ARMA (1, 2)-EGARCH (1, 1)	0.031481427	0.021301352
ARMA (0, 2)-EGARCH (2, 1)	0.031472603	0.021497737
ARMA (2, 3)-EGARCH (2, 2)	0.031102412	0.021941315
ARMA (3, 4)-EGARCH (2, 1)	0.031957636	0.022079428
ARMA (1, 4)-EGARCH (2, 1)	0.031556955	0.021818166
ARMA (4, 4)-EGARCH (2, 1)	0.031675226	0.021060602

Note: ARMA = autoregressive moving average; EGARCH = exponential generalized autoregressive conditional heteroscedasticity; RMSE = Root Mean Squared Error; MAE = Mean Absolute Error.

Such that the mean equation defined as the ARMA (4, 3) model and it is the stationary requirement has been verified by the following:

The roots of the polynomial for Autoregressive model Autoregressive (AR) (4), which is:

$$1 - 0.9477L + 1.2850L^2 - 0.7855L^3 + 0.4487L^4 = 0$$

Then, the value of the roots is:

$$a_1 = 0.9227 + 0.9749i, \text{ then } |a_1| = 1.3423 > 1$$
$$a_2 = 0.9227 - 0.9749i, \text{ then } |a_2| = 1.3423 > 1$$
$$a_3 = -0.0473 + 1.1112i, \text{ then } |a_3| = 1.1122 > 1$$
$$a_4 = -0.0473 - 1.1112i, \text{ then } |a_4| = 1.1122 > 1$$

Hence, all roots lie outside the unit circle. While, the roots of the polynomial for Moving Average model Moving average (MA) (3), which is:

$$1 + 1.0326L + 1.0385L^2 + 0.6437L^3 = 0$$

Then, the value of the roots is:

$$b_1 = 1.2990 + 0i, \text{ then } |b_1| = 1.2990 > 1$$
$$b_2 = 0.1571 + 1.0822i, \text{ then } |b_2| = 1.0936 > 1$$
$$b_3 = 0.1571 - 1.0822i, \text{ then } |b_3| = 1.0936 > 1$$

Hence, all of these roots lie outside the unit circle, which is mean its invertible.

The variance equation corresponds to the EGARCH (2, 2) model, and it is the stationary condition is verified as follows:

$$\sum_{i=1}^{2} \beta_i = 1.0114 - 0.1913 = 0.8201 < 1$$

Table 6. Distributional analysis of standardized residuals after exponential generalized autoregressive conditional heteroscedasticity (EGARCH) modeling

Metric	Value	Standard Value	*p*-Value
Kurtosis	3.0782	3	–
Jarque-Bera	2.2566	0.05	0.2167

Note: (–) = not applicable.

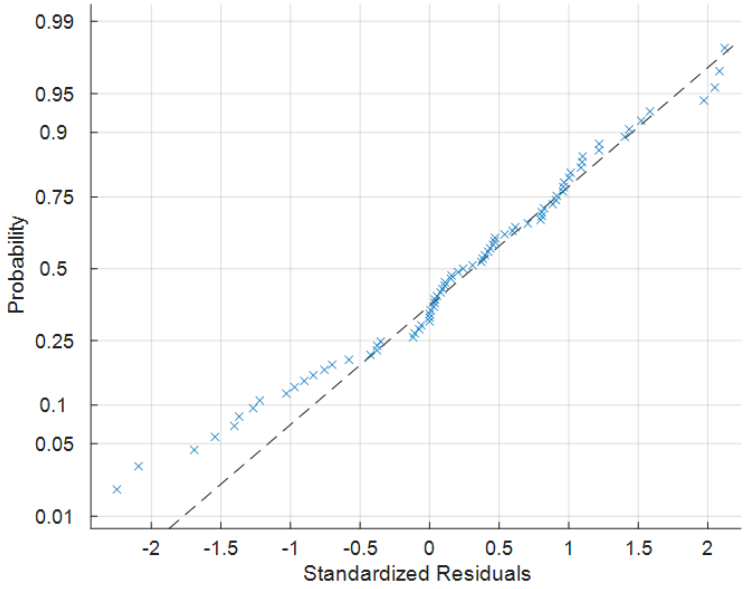

Figure 5. Probability plot of standardized residuals (after processing)

For all these mathematical verifications above, we get that the ARMA (4, 3)-EGARCH (2, 2) model is strictly stationary and reliable for forecasting. Table 6 and Figure 5 show a significant improvement in distributional properties after model estimation. The kurtosis is decreased to 3.0782, which is near the usual standard benchmark, which means that the EGARCH model successfully captured and accommodated the fat-tails, and the residuals are now significantly closer to the normal distribution, while the Jarque-Bera test also shows that it is not rejected (p =

0.2167). Following the validation that the standardized residuals behave as white noise and approximate normality. This change validates that the ARMA-EGARCH model is effective at capturing the conditional heteroskedasticity and purges the data into about white noise residuals. Consequently, the model is a good basis for prediction and anomaly detection.

5.2 Evaluation Metrics

To evaluate the forecasting accuracy, the standard statistical error metrics has been used including Mean Squared Error (MSE), RMSE, MAE, and Mean Absolute Percentage Error (MAPE), which are commonly adopted in time series forecasting research [21]. In addition, Theil's U inequality coefficient was used to assess the relative predictive performance [22]. Although the ARMA–EGARCH model is estimated using the return series to ensure stationarity, the forecasting performance is evaluated on the reconstructed original series $G_{\rho,t}$ to reflect the model's practical predictive capability. These metrics are defined as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(G_{\rho,t} - \hat{G}_{\rho,t}\right)^2 \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(G_{\rho,t} - \hat{G}_{\rho,t}\right)^2} \tag{12}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|G_{\rho,t} - \hat{G}_{\rho,t}\right| \tag{13}$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{G_{\rho,t} - \hat{G}_{\rho,t}}{G_{\rho,t}}\right| \tag{14}$$

$$\text{Theil's } U = \frac{RMSE}{\sqrt{\frac{1}{n}\sum_{t=1}^{n} G_{\rho,t}^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n} \hat{G}_{\rho,t}^2}} \tag{15}$$

In the absence of labeled ground truth data labeled, this research used unsupervised metrics of evaluation of anomaly detection. In accordance with Aggarwal [23], calculate the outlier ratio, which is the ratio of detected anomalies to the total observations:

$$\text{Outlier Ratio } = \frac{N_{\text{anomaly}}}{N} \times 100\% \tag{16}$$

where, N_{anomaly} represents the number of detected anomalies AD_t and N represents the total number of observations in the test set. Following established practices in robust statistics [24], we compute severity score as the average normalized deviation of detected anomalies, which confirms the statistical significance of the identified outliers.

$$\text{Severity Score } = \frac{1}{|N_{\text{anomaly}}|}\sum_{t \in N_{\text{anomaly}}} \frac{\left|G_{\rho,t} - \hat{G}_{\rho,t}\right|}{\sigma_t} \tag{17}$$

where, N_{anomaly} represents the number of detected anomalies AD_t and σ_t is the conditional standard deviation estimated by the EGARCH model.

Additionally, the performance of the detection method can be evaluated by adopted confusion matrix-based evaluation metrics, including the Accuracy, Precision, Recall, and F1-Score metrics. These metrics have been used widely to quantify classification and detection performance methods by converting the series into a binary decision task, where a label of 0 refers to the before-change interval and 1 to the after-change interval. The predicted Change point (CP) splits the timelines into two classes, resulting in a set of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics are defined as follows [25]:

$$\text{Accuracy } = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{19}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{20}$$

$$F1 - \text{Score} = \frac{2TP}{2TP + FP + FN} \tag{21}$$

5.3 Forecasting and Anomaly Detection

The forecasting and anomaly detection are essential steps in the proposed methodology. In this research, the data was partitioned into 80% for training and 20% for testing. The ARMA (4, 3)-EGARCH (2, 2) model trained on the r_t to model both the conditional mean and time-varying volatility of the network time series. The estimated model was then used to forecast the future values of the network density using the by recovering the predicted returns back to the original data. The results indicate that the model is a good predictive with low error measures.

These results refer that the model is effective in capturing the temporal dependencies and nonlinear volatility structure of the data series while keeping high precision in reconstructing the network's actual behavior. To further illustrate the model's performance, Table 7 shows the actual and predicted values of the network density series comparison with ARMA model for the first 20 future steps starting in the test region, while, the Figures 6 and 7 illustrate the forecasting data vs. observed data of ARMA (4, 3)-EGARCH (2, 2) and ARMA (4, 3) models, respectively.

The evaluation prediction performance results of the ARMA-EGARCH and ARMA are shown in the Table 8.

The model also introduces an unsupervised approach to anomaly detection through time-varying volatility. As mentioned in section 3, the dynamic statistical envelope was built using a threshold K = 3 (that is, a 99.7% confidence level). The top of Figure 8 clearly shows how the red marks are located outside the pink area; this is considered the visual realization of the model's success in separating normal behavior from attacks, while the bottom figure shows deviation magnitude, which is a security severity index that measures the magnitude of each anomaly found. This dual representation not only detects an event of an attack but also offers a prioritized view of potential network threats in terms of their statistical significance.

Table 7. Real values vs. predicted values of the ARMA-EGARCH and ARMA model

Time	Actual	ARMA-EGARCH Predicted Value	ARMA Predicted Value
1	0.98107	0.98171	0.98011
2	0.98868	0.99172	0.99275
3	0.99331	0.98915	0.99200
4	0.98860	0.98587	0.98557
5	0.92438	0.98818	0.98767
6	0.91617	0.92836	0.93144
7	0.91746	0.91436	0.91590
8	0.93730	0.91303	0.91007
9	0.94987	0.93828	0.93749
10	0.96474	0.95295	0.95708
11	0.99474	0.96314	0.96428
12	0.97723	0.99153	0.98781
13	0.99229	0.97788	0.97816
14	1.00000	0.99383	0.99890
15	0.91210	0.99833	0.99871
16	0.89671	0.91021	0.90637
17	0.88213	0.89725	0.89808
18	0.86323	0.88280	0.88734
19	0.87451	0.86200	0.86169
20	0.83012	0.87328	0.86970

Note: ARMA-EGARCH = autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity; ARMA = autoregressive moving average.

The results in the Table 9 indicates that the proposed methodology was able to identify 6 anomalies with time spent of 0.0009985 seconds as opposed to the 2 anomalies identified by the median absolute deviation (MAD)

represents the statistical approach and the single anomaly point identified by the one-class support vector machine (OCSVM) method, with time spent of 0.021857 seconds and 0.032116 seconds, respectively. This represents a great improvement in both sensitivity and running time, as the ARMA-EGARCH model is significantly faster than these baselines.

Furthermore, these measures demonstrate the model's strength in detecting anomalies, while the MAD and OCSVM methods appeared the higher in precision, they are overly conservative, missing 50% to 75% of the actual threats as shown in their low Recall. In contract, the ARMA-EGARCH model achieved a perfect Recall and the highest F1-Score, indicating a superior balance between sensitivity and accuracy.

Based on these results and in spite of the fact that the baselines maintain high precision by only flagging extreme outliers, the hybrid ARMA-EGARCH model offers a more in-depth description of the data as it simultaneously models the conditional mean and conditional variance. More importantly, the hybrid model proves to be more sensitive in the structural anomaly detection process, with an ability to capture the structural deviations that are overlooked by non-temporal detectors. Thus, the hybrid model has not only the contribution of the accuracy of the prediction but also the ability to detect volatility changes.

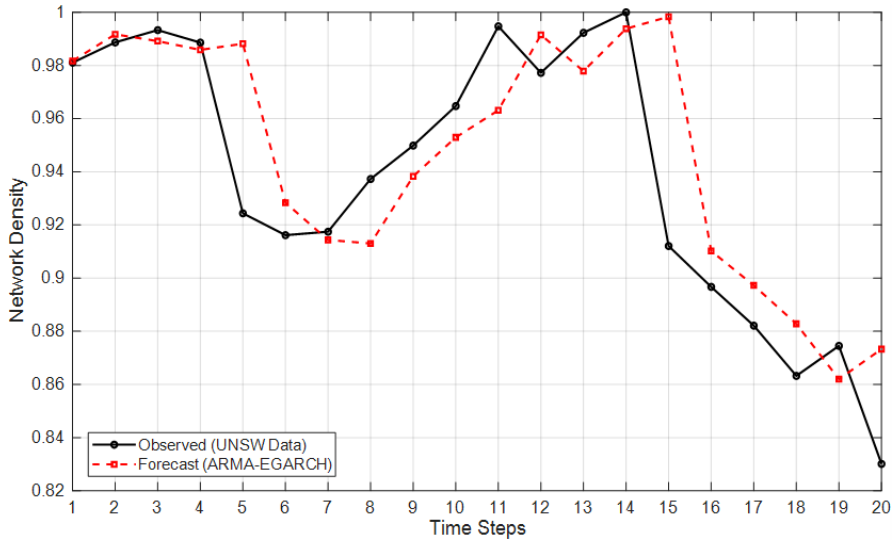

Figure 6. Illustrates the forecasting data vs. observed data of ARMA (4, 3)-EGARCH (2, 2) model

Note: ARMA-EGARCH = autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity.

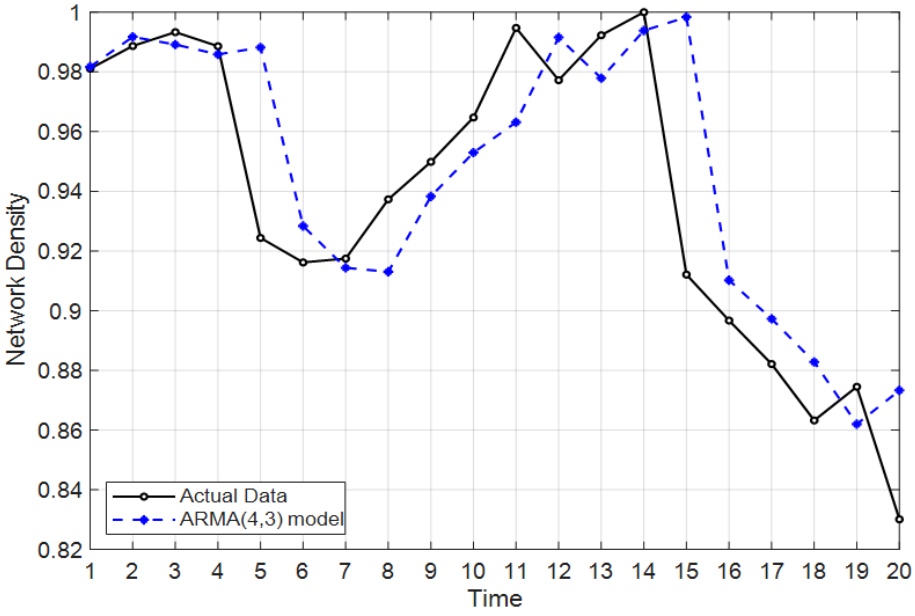

Figure 7. Illustrates the forecasting data vs. observed data of ARMA (4, 3) model

Note: ARMA = autoregressive moving average.

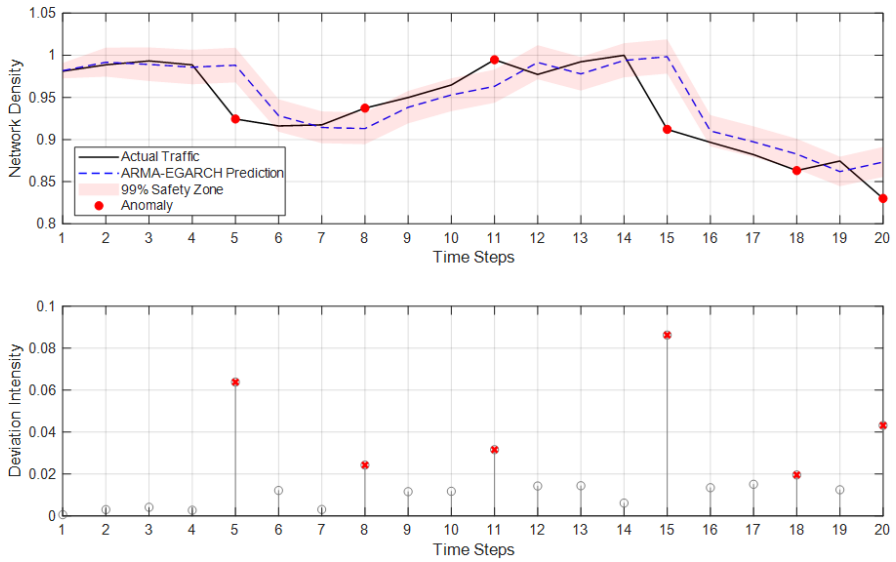

Figure 8. Illustrates the network anomaly detection of the proposed methodology (ARMA-EGARCH)
Note: ARMA-EGARCH = autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity.

Table 8. Comparative performance metrics of the ARMA-EGARCH and ARMA model

Model	MSE	RMSE	MAE	MAPE	Theil's U	Time Processing (s)
ARMA	0.00083	0.02883	0.01909	2.086	0.01527	0.12022
ARMA-EGARCH	0.00084	0.0290	0.01970	2.1476	0.01536	0.04270

Note: ARMA-EGARCH = autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity; ARMA = autoregressive moving average; MSE = Mean Squared Error; RMSE = Root Mean Squared Error; MAE = Mean Absolute Error; MAPE = Mean Absolute Percentage Error.

Table 9. Performance metrics comparison of anomaly detection processes

Metrics	ARMA-EGARCH	MAD	OCSVM
Anomalies detected	6	2	1
Outlier ratio	30%	10%	5%
Severity score	6.56	3.82	34.57
Accuracy	90%	90%	85%
Precision	66.67%	100%	100%
Recall	100%	50%	25%
F1-Score	80%	66.67%	40%
Time processing	0.0009985 s	0.021857 s	0.032116 s

Note: ARMA-EGARCH = autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity; MAD = median absolute deviation; OCSVM = one-class support vector machine.

6 Conclusions

The model was able to detect finer grained of structural transitions and connectivity patterns that define network evolution by converting raw network traffic into graph density time series, which considered an aggregate measure of the nodes' degree. The use of the hybrid ARMA-EGARCH model is not only a preference but also a statistical necessity due to non-linear characteristics and asymmetric volatility shocks. This ability to model conditional variance made it possible to generate the dynamic statistical envelope that expands and contracts in response to real-time volatility. The results demonstrate that the ARMA-EGARCH provides a desirable trade-off between prediction rigor and detection efficiency, which makes it very appropriate in real-time monitoring and anomaly detection, and its high speed of execution in time processing compared to other methods. It is recommended that future studies should expand on these findings by including other structural graph measures, including centrality and clustering coefficients, to further enhance the prediction abilities in response to new and advanced cyber threats.

Author Contributions

Conceptualization, D.G.; methodology, D.G.; validation, B.S.; formal analysis, D.G.; investigation, D.G.; resources, B.S.; data curation, D.G.; writing—original draft preparation, D.G.; writing—review and editing, D.G.; visualization, D.G.; supervision, B.S.; project administration, D.G. All authors have read and agreed to the published version of the manuscript.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgements

This work was supported by the Department of Mathematics, College of Science, University of Baghdad, Baghdad, Iraq.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey," *Data Min. Knowl. Discov.*, vol. 29, pp. 626–688, 2015. https://doi.org/10.1007/s10618-014-0365-y

[2] Z. K. Hussien and N. B. Dhannoon, "Anomaly detection approach based on deep neural network and dropout," *Baghdad Sci. J.*, vol. 17, no. 2, pp. 701–709, 2020. https://doi.org/10.21123/bsj.2020.17.2(SI).0701

[3] M. Subhi, O. F. Rashid, S. A. Abdulsahib, and M. K. Hussein, "Anomaly intrusion detection method based on RNA encoding and ResNet50 model," *Mesopotamian J. CyberSecurity*, vol. 4, no. 2, pp. 120–128, 2024. https://doi.org/10.58496/MJCS/2024/011

[4] M. E. Manaa, F. J. A. Al-Razaq, and H. A. Al-Khamees, "Enhancing IoT security: An optimization algorithm for fog layer-based DDoS attack mitigation framework," *Iraqi J. Sci.*, vol. 66, no. 2, pp. 765–787, 2025. https://doi.org/10.24996/ijs.2025.66.2.19

[5] A. A. Abdualrahman and M. K. Ibrahem, "Intrusion detection system using data stream classification," *Iraqi J. Sci.*, vol. 62, no. 1, pp. 319–328, 2021. https://doi.org/10.24996/ijs.2021.62.1.30

[6] L. Peel and A. Clauset, "Detecting change points in the large-scale structure of evolving networks," *Proc. AAAI Conf. Artif. Intell.*, vol. 29, no. 1, pp. 2914–2920, 2015. https://doi.org/10.1609/aaai.v29i1.9574

[7] A. Soule, K. Salamatian, and N. Taft, "Combining filtering and statistical methods for anomaly detection," in *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*, Berkeley, California, USA, 2005, pp. 331–344.

[8] S. Ranshous, S. Shen, D. Koutra, S. Harenberg, C. Faloutsos, and N. F. Samatova, "Anomaly detection in dynamic networks: A survey," *WIREs Comput. Stat.*, vol. 7, no. 3, pp. 223–247, 2015. https://doi.org/10.1002/wics.1347

[9] Purwanto, Sunardi, F. T. Julfia, and A. Paramananda, "Hybrid model of ARIMA-linear trend model for tourist arrivals prediction model in Surakarta City, Indonesia," *AIP Conf. Proc.*, vol. 2114, no. 1, p. 060010, 2019. https://doi.org/10.1063/1.5112481

[10] A. A. Mudhir, "Mixing ARMA models with EGARCH models and using it in modeling and analyzing the time series of temperature," *Iraqi J. Sci.*, vol. 62, no. 7, pp. 2307–2326, 2021. https://doi.org/10.24996/ijs.2021.62.7.19

[11] J. Kasali and A. A. Adeyemi, "Model-data fit using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the sample-size-adjusted BIC," *J. Math. Math. Educ.*, vol. 4, no. 1, pp. 43–51, 2022. https://doi.org/10.21580/square.2022.4.1.11297

[12] P. Mei and Y. H. Zhao, "Dynamic network link prediction with node representation learning from graph convolutional networks," *Sci. Rep.*, vol. 14, p. 538, 2024. https://doi.org/10.1038/s41598-023-50977-6

[13] S. Racma and C. Isagani, Jr., "On density of some graphs and their Mycielski graph," *Eur. J. Pure Appl. Math.*, vol. 18, no. 3, p. 5974, 2025. https://doi.org/10.29020/nybg.ejpam.v18i3.5974

[14] S. Ling, "Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models," *J. Econom.*, vol. 140, no. 2, pp. 849–873, 2007. https://doi.org/10.1016/j.jeconom.2006.07.016

[15] Y. Bao, "The asymptotic covariance matrix of the QMLE in ARMA models," *Econom. Rev.*, vol. 37, no. 4, pp. 309–324, 2018. https://doi.org/10.1080/07474938.2016.1140287

[16] D. B. Nelson, "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, vol. 59, no. 2, pp. 347–370, 1991. https://doi.org/10.2307/2938260

[17] I. U. Moffat, E. A. Akpan, and U. A. Abasiekwere, "A time series evaluation of the asymmetric nature of heteroscedasticity: An EGARCH approach," *Int. J. Stat. Appl. Math.*, vol. 2, no. 6, pp. 111–117, 2017.

[18] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in *Proceedings of 2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, Australia, 2015, pp. 1–6. https://doi.org/10.1109/MilCIS.2015.7348942

[19] S. Rahman, "UNSW_NB15 preprocessing implementation," GitHub Repository, 2021. https://github.com/rahmanshiddiqur/UNSW_NB15/blob/main/UNSW_Preprocess.ipynb

[20] J. Liu and M. Guo, "DIGNN-A: Real-time network intrusion detection with integrated neural networks based on dynamic graph," *Comput. Mater. Contin.*, vol. 82, no. 1, pp. 817–842, 2025. https://doi.org/10.32604/cmc.2024.057660

[21] A. Jadon, A. Patil, and S. Jadon, "A comprehensive survey of regression-based loss functions for time series forecasting," in *Proceedings of International Conference on Data Management, Analytics and Innovation*, Vellore, India, 2024, pp. 117–147. https://doi.org/10.1007/978-981-97-3245-6_9

[22] K. Gohain, "Evaluation of Theils U: A naïve forecast application," *Quantum J. Eng. Sci. Technol.*, vol. 2, no. 5, pp. 26–31, 2021.

[23] C. C. Aggarwal, *Outlier Analysis*. Springer, 2017.

[24] P. J. Rousseeuw and M. Hubert, "Anomaly detection by robust statistics," *WIREs Data Min. Knowl. Discov.*, vol. 8, no. 2, p. e1236, 2018. https://doi.org/10.1002/widm.1236

[25] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, p. 6, 2020. https://doi.org/10.1186/s12864-019-6413-7