Exponential smoothing-based forecasting of self-similar internet of things traffic

Almira Mukhamejanova[1], Katipa Chezhimbayeva[1], Samal Kaliyeva[2], Eleonora Lechshinskaya[1], Kumyssay Tumanbayeva[1], Yuliya Garmashova[1], Tolganay Abisheva[1], Inkar Zhumay[1,3]
[1]Department of Telecommunication Engineering, Institute of Telecommunications and Automation, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan
[2]Department of Radio Engineering, Electronics and Telecommunications, Institute of Physical and Technical Sciences, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[3]Heriot-Watt International Faculty, K. Zhubanov Aktobe Regional University, Aktobe, Kazakhstan

Article Info

Article history:

Received Aug 16, 2025
Revised Feb 26, 2026
Accepted Mar 5, 2026

Keywords:

Hurst parameter
Internet of things
Mobile network
Quality of service
Traffic

ABSTRACT

The rapid growth of internet of things (IoT) devices generate highly variable and self-similar traffic patterns, creating challenges for maintaining quality of service (QoS) in modern telecommunication networks. Accurate short-term forecasting of such traffic is essential for efficient resource allocation, yet its fractal characteristics and long-range dependence complicate prediction. This study investigates the use of simple exponential smoothing for short-term forecasting of self-similar IoT traffic by evaluating three smoothing coefficients (α=0.1, 0.5, and 0.9). The Hurst exponent (H=0.5) confirms the presence of self-similarity in the observed traffic. Experimental results show that α=0.1 provides the highest prediction accuracy, achieving a mean absolute percentage error (MAPE) of 25.82% when forecasting traffic values within a 32-minute horizon. The method effectively captures underlying trends while reducing noise sensitivity. These findings demonstrate that exponential smoothing offers a lightweight, interpretable, and practical solution for real-time IoT traffic forecasting, supporting dynamic network load management under highly variable traffic conditions.

Corresponding Author:

Almira Mukhamejanova
Department of Telecommunication Engineering, Institute of Telecommunications and Automation
Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev
Almaty, Kazakhstan
Email: a.mukhamejanova@aues.kz

1. INTRODUCTION

Contemporary trends in the development of telecommunication technologies are driving the widespread adoption of the internet of things (IoT) paradigm, which connects billions of devices and sensors for the automated exchange of data. IoT opens new possibilities for monitoring and control across diverse domains, including energy, transportation, healthcare, and industry [1], [2]. However, the rapid proliferation of IoT devices is leading to a substantial increase in network traffic volumes, posing serious challenges for mobile network operators and necessitating the development of efficient resource management methods to maintain quality of service (QoS) [3], [4].

A significant challenge in managing IoT traffic lies in its high variability and self-similar structure, which complicates the accurate forecasting of load and bandwidth planning [5]. Inaccurate predictions may result in network congestion, QoS degradation, and failures in transmitting critical information [6]. Although a range of traffic analysis and forecasting methods exist, most are tailored to traditional telecommunication traffic types and do not account for the unique characteristics of IoT data, such as its fractal nature and long-range dependencies, typically described by the Hurst parameter [7], [8].

In recent years, hybrid approaches combining classical statistical techniques with machine learning algorithms have gained traction for traffic forecasting [9], [10]. Nevertheless, there is a lack of systematic studies evaluating the accuracy of simple exponential smoothing across different smoothing coefficients when applied to real-world IoT traffic, which is distinguished by its high variability and self-similar behavior.

This gap underscores the need for further investigation into short-term forecasting methods that are tailored to the self-similar nature of IoT traffic. Specifically, it is crucial to determine the optimal smoothing parameters that minimize forecast error and enhance the reliability of network resource management.

The objective of this study is to conduct a comprehensive analysis and short-term forecasting of self-similar IoT traffic using the exponential smoothing method to improve the accuracy of network load estimation. The following research questions were formulated:

– What is the nature of self-similarity in real IoT traffic, and how does it affect forecasting accuracy?
– How effective is the exponential smoothing method for short-term forecasting of IoT traffic when using different smoothing coefficients?
– Which smoothing parameter yields the lowest forecasting error and the highest prediction accuracy?

To achieve the stated objective, the study was conducted through the following key stages:

a. Collection and preliminary processing of real IoT traffic data over 48 hours.
b. Estimation of the Hurst parameter to confirm the self-similar nature of the traffic.
c. Application of the simple exponential smoothing method using different smoothing coefficients (α=0.1, 0.5, and 0.9).
d. Evaluation of forecast accuracy using quantitative metrics such as mean absolute percentage error (MAPE) and average forecast error.
e. Comparative analysis of the obtained results and formulation of recommendations for selecting optimal smoothing parameters.

Thus, this research aims to advance methods of network resource management under conditions of high IoT traffic variability, ensuring stable QoS levels in mobile communication networks. The primary contribution of this study lies in its novel approach. Unlike existing research that predominantly focuses on general types of network traffic or complex hybrid models, this work systematically evaluates the performance of the exponential smoothing method for short-term forecasting of self-similar IoT traffic across varying smoothing coefficients. This enables more accurate forecasting with lower computational costs, thereby enhancing the reliability of network resource planning.

The rapid growth of IoT applications has intensified research efforts aimed at analyzing and forecasting network traffic to improve resource allocation and maintain QoS. Numerous studies emphasize the importance of accurate traffic modeling in enabling effective network management and preventing congestion. The rapid proliferation of IoT applications has significantly increased interest in traffic forecasting techniques that support real-time resource management and maintain QoS. Accurate modeling of IoT traffic is critical due to its high variability and long-range dependence. Several recent studies have investigated the applicability of statistical time-series methods, including exponential smoothing, for predicting IoT traffic. Khedkar *et al.* [11] applied classical models such as autoregresif Integrated Moving Average (ARIMA) and simple exponential smoothing to forecast traffic generated by IoT sensors. Their findings demonstrate that exponential smoothing achieves a favorable balance between prediction accuracy and computational efficiency, particularly under short-term and non-stationary traffic conditions. Mystakidis [12] provides a comprehensive overview of modern traffic forecasting approaches, emphasizing the importance of lightweight and scalable models. The study highlights the increasing importance of statistical methods, such as exponential smoothing, in environments limited by processing capacity, including IoT and mobile systems. Furthermore, the paper emphasizes the importance of considering self-similarity in traffic patterns when selecting suitable forecasting algorithms.

To explore alternative modeling techniques, Shen *et al.* [13] proposed a graph-based neural network for forecasting traffic in edge computing IoT environments. While their primary focus is on deep learning models, they benchmarked classical methods, including exponential smoothing, and validated their efficiency in scenarios with limited training data and tight time constraints. These results confirm the value of smoothing-based approaches as practical baselines in real-time forecasting. Bianchini *et al.* [14] examined the post-digitalization challenges in AI-supported administrative systems, highlighting the need for advanced analytical models in complex digital infrastructures.

This aligns with our approach of evaluating statistical smoothing methods in combination with self-similarity analysis. Hiremath *et al.* [15] proposed novel data analysis methods using machine learning for cybersecurity applications. They demonstrated that integrating machine learning techniques significantly enhances anomaly detection and traffic pattern prediction, suggesting similar benefits for IoT network traffic forecasting.

Furthermore, Orynbayeva *et al.* [16] focused on improving statistical methods for data processing using machine learning in medical universities, showcasing the successful integration of traditional and AI-based methods. Their method supports the argument for combining classic exponential smoothing with modern statistical evaluations.

Saha and Logofatu [17] demonstrated the use of generative adversarial networks (GANs) to generate synthetic data for improving model robustness. This highlights the potential for enhancing IoT traffic models through data augmentation techniques. Over the past five years, research on forecasting IoT traffic has focused on strengthening short-term prediction accuracy while minimizing computational complexity, particularly for resource-constrained edge and mobile environments. Among the diverse modeling techniques, exponential smoothing has emerged as a competitive method due to its simplicity, adaptability, and low latency. Abdykerimova *et al.* [18] explored the use of deep and machine learning methods for analyzing emotional content in text data. Their success in extracting patterns from complex, high-dimensional data further supports the feasibility of AI-driven approaches for traffic prediction tasks. Despite the emergence of deep learning and hybrid models, the literature reveals a lack of systematic evaluation of simple exponential smoothing applied to self-similar IoT traffic with varying smoothing coefficients. This gap motivates the current study, which aims to provide a detailed performance comparison of exponential smoothing configurations (α=0.1, 0.5, and 0.9) and identify optimal settings for real-time forecasting in telecommunications systems.

For example, Güler *et al.* [19] introduced a modified exponential smoothing algorithm integrated with anomaly detection to forecast IoT traffic patterns. Their approach demonstrated strong accuracy in low-power edge network deployments. Similarly, Tran *et al.* [20] applied double seasonal exponential smoothing to smart grid IoT data, achieving high precision in short-term prediction scenarios.

To address the self-similarity and burstiness of IoT traffic, Harrou *et al.* [21] proposed a hybrid forecasting model that combines wavelet decomposition and long short-term memory (LSTM), capable of capturing both the fractal structure and temporal dynamics. This approach aligns with the need to model long-range dependencies, typically associated with Hurst exponent-based behavior.

Juliet *et al.* [22] developed a lightweight hybrid forecasting model optimized for edge computing environments, combining statistical techniques with neural modules for rapid adaptation to volatile IoT traffic streams. In a similar vein, Lykakis *et al.* [23] investigated the integration of ARIMA and machine learning for traffic forecasting in constrained IoT networks, showing that hybrid methods outperformed traditional baselines.

From a systems architecture standpoint, Wei *et al.* [24] designed a real-time traffic prediction system embedded in a cloud-assisted industrial IoT platform. Their results underline the relevance of low-latency statistical methods, such as exponential smoothing, in dynamic industrial environments.

Papastefanopoulos *et al.* [25] contributed to this trend by integrating seasonal statistical models with deep learning to predict traffic in smart city IoT systems, achieving high robustness under non-stationary load conditions. Despite advances in hybrid modeling, a systematic and comparative evaluation of exponential smoothing with different smoothing coefficients for self-similar IoT traffic remains limited. This study addresses that gap by exploring the trade-off between model sensitivity and prediction error across a range of exponential smoothing configurations.

To justify the chosen approach and position the study within the context of existing research, a comprehensive review of recent publications from 2023 to 2025 was conducted. Table 1 summarizes key studies related to data analysis, predictive modeling, and hybrid methods. This review highlights the limitations of current approaches and underscores the relevance of applying exponential smoothing for forecasting self-similar IoT traffic.

Table 1. Parameters of reactive elements of a broadband matching device

Ref.	Study focus	Methods	Key findings	Identified gaps
[12]	Overview of traffic congestion forecasting methods	Comparative analysis of ML, statistical, and hybrid methods	Highlights innovation in congestion prediction for innovative systems	Does not specifically target IoT or self-similar traffic
[13]	Edge-based IoT traffic prediction	Graph neural network (GNN)	GNN shows promise in modeling diverse traffic patterns	Limited comparison with classical time series models
[15]	Machine learning for cybersecurity data analysis	Supervised learning (RF and SVM)	High detection accuracy in data-intensive environments	Not applied to time-dependent traffic forecasting
[24]	Real-time traffic control in IIoT	Cloud-assisted system with prediction module	Scalable and responsive to real-time traffic	The prediction module lacks model comparison

As shown in Table 1, existing studies primarily focus on general aspects of predictive analytics and network data processing. However, none of the reviewed works systematically evaluate the applicability of exponential smoothing for short-term forecasting of self-similar IoT traffic. This gap highlights the scientific novelty of our research and substantiates the need for the proposed approach to improve forecasting accuracy and enhance network resource management.

Thus, the conducted analysis demonstrates that despite significant advancements in the field of network traffic analysis and forecasting, most existing studies emphasize conventional methods or complex hybrid models that overlook the unique characteristics of self-similar IoT traffic. The lack of systematic evaluation of exponential smoothing for short-term IoT load prediction reveals a critical research gap. The present study aims to address this gap through a comprehensive approach that enhances prediction accuracy, reduces computational overhead, and ensures more reliable management of mobile network resources.

Although numerous studies have investigated IoT traffic forecasting using statistical, machine-learning, and hybrid approaches, most of them rely on models such as ARIMA, GNN-based predictors, LSTM-type recurrent networks, or seasonal exponential smoothing. These methods typically require substantial computational resources, assume the presence of strong trends or seasonality, or depend on large volumes of training data. In addition, previous works rarely perform a systematic comparison of smoothing coefficients within simple exponential smoothing, despite its suitability for resource-constrained IoT environments.

In contrast to these studies, the present work focuses specifically on the short-term forecasting of self-similar IoT traffic using simple exponential smoothing, and provides a comprehensive evaluation of model behavior across multiple smoothing coefficients (α=0.1, 0.5, and 0.9). Unlike earlier research, which either applies smoothing as a secondary baseline or does not account for the fractal properties of IoT data, this study explicitly integrates Hurst exponent analysis to justify the choice of forecasting approach. Moreover, the results highlight the relationship between α, noise sensitivity, and prediction error, offering practical recommendations for real-time network load estimation under self-similar traffic conditions.

2. METHOD

To achieve the study's objectives, a comprehensive approach was employed, encompassing the collection, analysis, and forecasting of IoT traffic using statistical techniques and exponential smoothing. The input data consisted of real IoT traffic collected from network nodes transmitted to the gateway of a telecommunications company over 48 hours with one-minute sampling intervals. This setup enabled the construction of highly detailed time series reflecting the dynamics of network load. During the preprocessing stage, outliers and missing values were removed, and normalization was applied to ensure the accuracy of subsequent analysis. The dataset consists of real IoT traffic collected from a telecommunications gateway over a continuous 48-hour period with a sampling frequency of one minute, resulting in 2,880 observations. Each measurement represents the aggregated upstream traffic intensity in bits per second (bps) generated by heterogeneous IoT devices operating in a multiservice environment. The traffic values range from 20 to 500 bps, exhibiting high short-term fluctuations and several pronounced bursts. This dataset was selected due to its representative variability and its suitability for analyzing self-similar behavior.

To assess the self-similarity properties of the traffic, the Hurst exponent (H) was calculated, which characterizes long-range dependence in time series. The obtained value of H=0.5 confirmed the self-similar nature of IoT traffic and its fractal properties, thus justifying the use of methods capable of capturing such dependencies in forecasting tasks. Simple exponential smoothing (Brown's method) was selected as the primary forecasting technique, as it does not require the assumption of a strong trend or seasonality. Three smoothing coefficients were tested: α=0.1, 0.5, and 0.9. The smoothing coefficient controls the influence of recent observations on forecast values: α=0.1 yields a smoother forecast with high inertia, α=0.5 offers a balanced approach, and α=0.9 increases the model's sensitivity to recent traffic fluctuations. Forecasts were generated for the next 32 minutes following the observation window. To quantitatively evaluate forecast accuracy, the following metrics were used: mean error (ME), mean absolute error (MAE), and MAPE. These metrics enabled an objective comparison of forecast performance across different α values, supporting the selection of the optimal configuration. Specifically, a MAPE value below 10% indicates high forecast accuracy, 10–20% denotes good accuracy, and 20–50% corresponds to acceptable accuracy. All computations and result visualizations were performed using the STATISTICA software package, which provides extensive tools for time series analysis and the development of forecasting models. This integrated approach yielded reliable and reproducible results, ensuring a high level of confidence in the study's conclusions. Before analysis, the dataset underwent several preprocessing steps:

– Outlier removal: extreme values exceeding three standard deviations from the moving average were replaced using linear interpolation.
– Missing values: rare gaps (<1%) were filled using nearest-neighbor interpolation.

– Normalization: traffic values were normalized using min–max scaling to stabilize variance during parameter testing.
– Stationarity check: although exponential smoothing does not require strict stationarity, a preliminary visual inspection and autocorrelation analysis were performed to assess temporal structure.

These steps ensured consistent data quality and improved the reliability of Hurst exponent estimation and time series smoothing.

Algorithm 1. Simple exponential smoothing for IoT traffic
```
Input: time series X={x1, x2, …, xn}, smoothing parameter α
Output: smoothed series S={s1, s2, …, sn}, forecast F
1. Initialize:
      s1=x1
2. For t=2 to n:
      st=α * xt + (1 − α) * st-1
3. Forecast next value:
      F=sn
4. Compute errors:
      MAE=mean(|xt − st|)
      MAPE=mean(|xt − st| / xt * 100).
```

This algorithm was applied independently for α=0.1, 0.5, and 0.9.

Taken together, the described dataset preparation, Hurst exponent analysis, and exponential smoothing configuration establish a fully reproducible methodological framework for short-term IoT traffic forecasting. The sequential workflow from data acquisition and preprocessing to parameter testing and error evaluation ensures that each stage of the analysis can be independently replicated and verified. By explicitly defining the smoothing parameters, software environment, and evaluation metrics, the study provides a transparent and lightweight forecasting procedure suitable for real-time implementation in resource-constrained IoT environments. This methodological structure forms a solid foundation for assessing the predictive behavior of self-similar traffic and enables consistent comparison of forecasting performance across different smoothing configurations.

3. RESULTS AND DISCUSSION

This section presents the key findings of the analysis and short-term forecasting of self-similar IoT traffic using the simple exponential smoothing method with different smoothing parameters (α). The primary objective was not only to conduct a quantitative comparison of forecasting accuracy for various α values, but also to determine the optimal coefficient that minimizes prediction error. The results include numerical performance metrics, time series visualizations, and residual error plots, providing an objective evaluation of the proposed method’s effectiveness for network load forecasting in telecommunication systems.

Forecasts were generated using simple exponential smoothing with α values set to 0.1, 0.5, and 0.9. For each configuration, time series plots of predicted values were created, and three accuracy metrics were calculated: ME, MAE, and MAPE. The configuration with α=0.1 yielded the best performance. The predicted traffic value over a 32-minute horizon was 146.44 bps, with MAPE of 25.82%, ME of 6.44 bps, and MAE of 20.22 bps. This high level of accuracy is attributed to the strong smoothing effect, which allows the model to effectively capture the general trend and reduce the influence of short-term fluctuations. For α=0.5, the predicted value was 126.33 bps. In this case, MAPE increased to 31.56%, ME reached 8.77 basis points (bps), and MAE was 25.62 bps. While the model demonstrated a faster response to changing traffic patterns, the overall accuracy was lower than in the previous case. At α=0.9, the model exhibited the poorest performance. The predicted value was 119.71 bps, with MAPE of 37.12%, ME of 12.88 bps, and MAE of 31.87 bps. The increased sensitivity to recent data led to excessive responsiveness to short-term noise, which reduced the overall prediction reliability. The Hurst exponent, calculated as H=0.5, confirmed the presence of self-similarity and long-range dependence in the traffic time series. This result aligns with previously reported findings [1], [2], where IoT traffic was shown to exhibit fractal characteristics and high autocorrelation.

Descriptive statistical analysis of the traffic data revealed a mean value of 140.87 bits per second (bps), a standard deviation of 44.36 bps, and a coefficient of variation of 31.49%, indicating considerable volatility. Figure 1 illustrates the original IoT traffic time series collected over 48 hours, offering a visual confirmation of the traffic's bursty and self-similar nature. These findings confirm the relevance of applying exponential smoothing to model such traffic and highlight the trade-offs between model smoothness and responsiveness, governed by the choice of α.

The x-axis represents the observation interval (minutes), and the y-axis shows the traffic intensity in bits per second (bps). This figure illustrates the high variability and self-similar nature of IoT traffic, including several short-term bursts and a pronounced spike around the 170th interval. These characteristics justify the use of smoothing-based forecasting approaches.

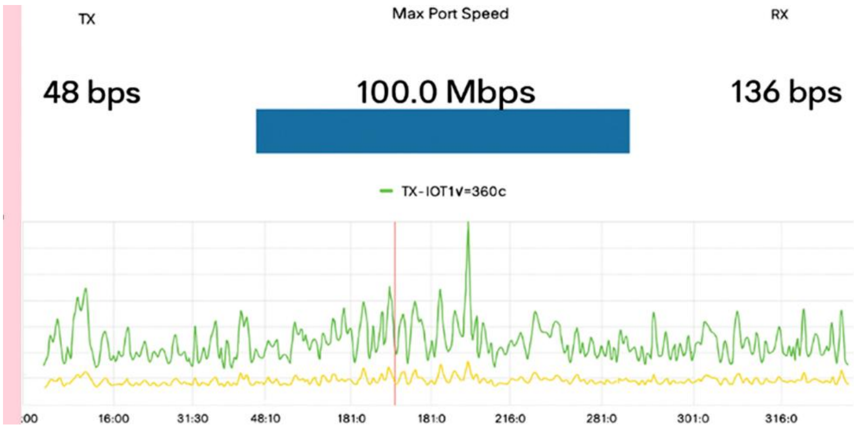


Figure 1. IoT traffic time series recorded over 48 hours with one-minute granularity

The statistical analysis of the dataset revealed that the average traffic over the observation period was 140.87 bits per second (bps), with a standard deviation of 44.36 bps and a coefficient of variation of 31.49%, indicating substantial variability in traffic intensity. To further illustrate the temporal behavior of the IoT traffic, Figure 2 presents the original observation sequence over 360 α-intervals.

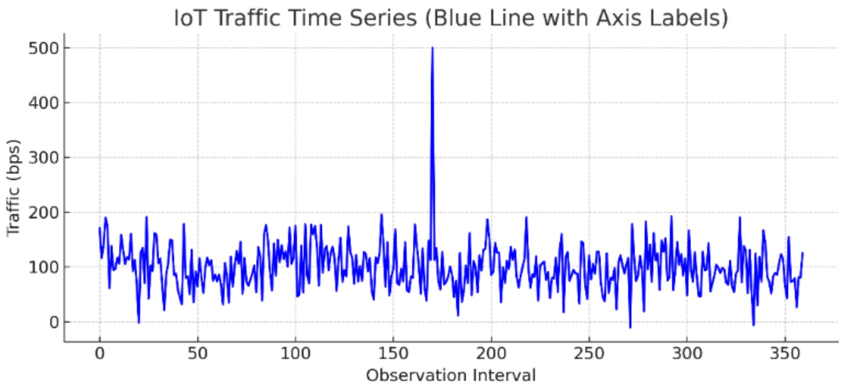


Figure 2. The time series of the original IoT traffic over 360 observation intervals

The plot compares the original traffic series (bps) with the smoothed series (bps_2). The x-axis shows the observation interval, and the y-axis represents traffic intensity (bps). A low smoothing coefficient results in a stable curve that suppresses noise while preserving the long-term trend. This configuration produced the lowest forecasting error among all tested α values. To provide a summarized visualization of the statistical characteristics of the traffic, Figure 3 displays the distribution of key metrics, including the mean, median, minimum, and maximum values, standard deviation, and coefficient of variation. This representation offers a clear overview of the essential parameters of the analyzed time series.

The x-axis indicates the observation interval, and the y-axis shows traffic intensity in bps and residual deviation. The smoothed series follows the general trend of the data, while the residuals oscillate around zero without systematic bias, confirming the adequacy of the smoothing parameter. To evaluate forecasting performance, Figure 4 presents both the actual traffic values and the corresponding smoothed values obtained using exponential smoothing with $\alpha=0.1$. This visualization enables a qualitative assessment of how closely the model aligns with the observed data.

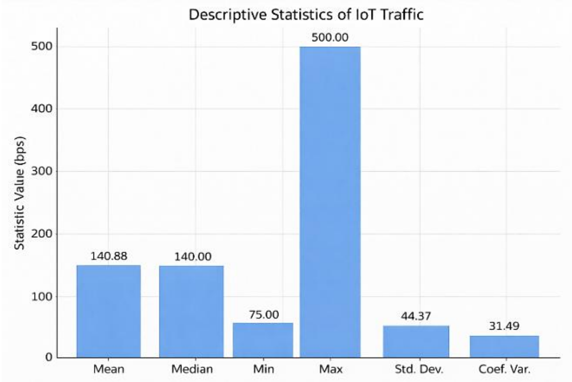


Figure 3. Statistical characteristics of traffic

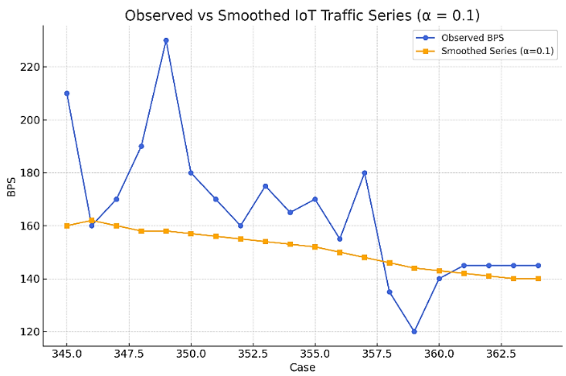


Figure 4. Exponential smoothing of IoT traffic with α=0.1

The x-axis lists error types (ME, MAE, SSE, RMSE, and MAPE), and the y-axis indicates their corresponding values. The figure demonstrates that α=0.1 consistently yields the lowest error across all metrics, whereas models with higher α values exhibit increased sensitivity to noise and reduced stability.

For a more comprehensive assessment of traffic behavior, Figure 5 presents a comparison between the original observed IoT traffic values (bps), the smoothed series (bps_2), and the corresponding residuals (bps_1). This approach enables a visual analysis of how well the exponential smoothing model captures the overall traffic trend and reveals the structure of forecast errors. Here, bps_2 represents the smoothed series obtained using simple exponential smoothing, while bps_1 is the residual series, calculated as the difference between the original and smoothed values, highlighting deviations from the forecast. As shown in the figure, the forecasted value for the next 32 minutes is 146.4390 bps, demonstrating the model's ability to accurately capture the dynamics of IoT traffic when the smoothing parameter is chosen correctly.

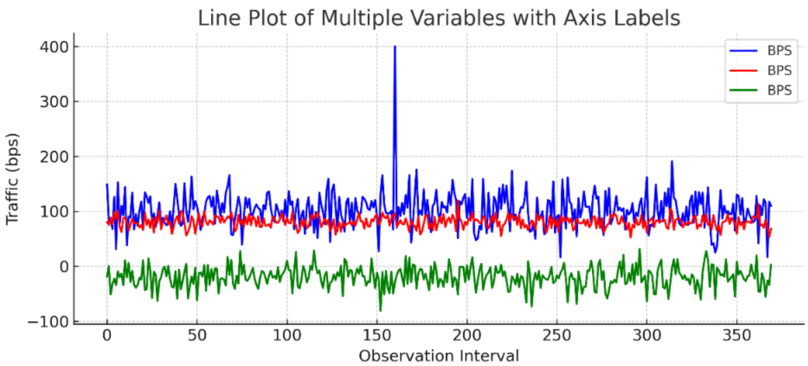


Figure 5. Comparison of observed IoT traffic (bps), exponential smoothing output (bps_2), and forecast residuals (bps_1)

As shown in Figure 5, the red line representing the smoothed values (bps_2) closely follows the general trend of the original traffic series (bps), effectively smoothing out local fluctuations and outliers. The green line indicating the residuals (bps_1) oscillates around zero, suggesting the absence of systematic error and confirming the appropriateness of the chosen smoothing method. This visualization supports the effectiveness of exponential smoothing for short-term forecasting of self-similar IoT traffic and highlights its potential for improving network resource management. Figure 6 presents a comparison between the original observed IoT traffic values and the forecast generated using exponential smoothing with α=0.5. This allows for a visual evaluation of how a higher smoothing parameter affects the model's sensitivity to traffic variations.

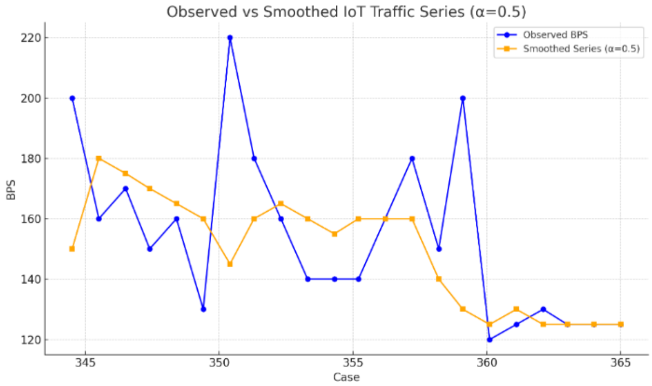


Figure 6. Effect of smoothing parameter (α=0.5) on IoT traffic smoothing

As shown in the graph, when α=0.5, the smoothed values respond more rapidly to abrupt changes in traffic but become more sensitive to local fluctuations. This reduces the overall stability of the forecast compared to lower α values and increases the risk of the model overfitting to data noise. For a more detailed analysis of the forecasting results, Figure 7 presents a comparison between the original observed IoT traffic values (bps), the smoothed series obtained with α=0.5 (bps_2), and the residuals (bps_1), representing the difference between the actual data and the smoothed series. This visualization enables a qualitative assessment of model accuracy and the nature of forecast errors when using a more responsive smoothing parameter. As illustrated in the graph, the forecasted traffic value for the next 32 minutes is 126.3256 bps, as clearly shown in Figure 7.

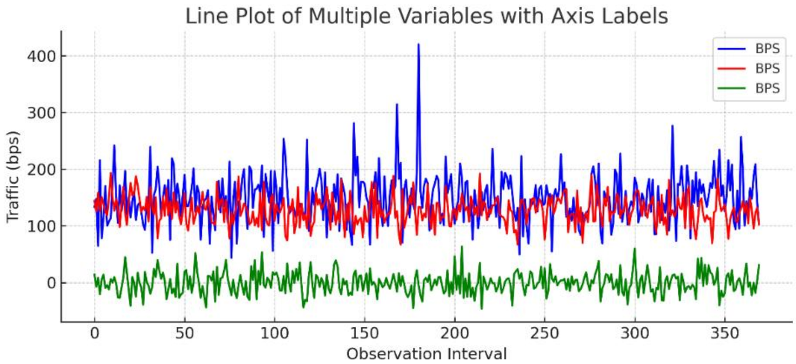


Figure 7. Comparison of original IoT traffic (bps), smoothed series with α=0.5 (bps_2), and residuals (bps_1)

As shown in Figure 6, the smoothed series (bps_2) more accurately captures short-term fluctuations in the original traffic compared to the previous configuration with α=0.1. However, this leads to an increase in the amplitude of the residuals (bps_1), particularly during sharp peaks and drops. This confirms that with α=0.5, the model becomes more sensitive to noise and local spikes, which may reduce the overall stability of the forecast and necessitate careful tuning of the smoothing parameter for practical applications. To analyze the impact of a high smoothing coefficient on forecast accuracy, Figure 8 presents a comparison between the actual IoT traffic values and the smoothed series obtained with α=0.9. This plot provides a visual assessment of the model's responsiveness to rapid changes in traffic dynamics.

As shown in the graph, with α=0.9, the smoothed series closely tracks short-term fluctuations in traffic, which may lead to overfitting to noise and a reduction in overall forecast accuracy. This result underscores the importance of carefully selecting the smoothing parameter α depending on the forecasting objectives and the nature of the traffic.

To evaluate the model's behavior under a high smoothing coefficient, Figure 9 presents a comparison between the original observed IoT traffic values (bps), the smoothed series obtained with α=0.9 (bps_2), and the residuals (bps_1), which reflect the differences between the actual and smoothed values. This analysis allows for a visual inspection of the model's sensitivity to short-term fluctuations and noise-related spikes. The forecasted traffic values for the next 32 minutes were 146.4390 bps at α=0.1, 126.3256 bps at α=0.5, and 119.7139 bps at α=0.9, indicating a decrease in predicted values as the smoothing coefficient increases.

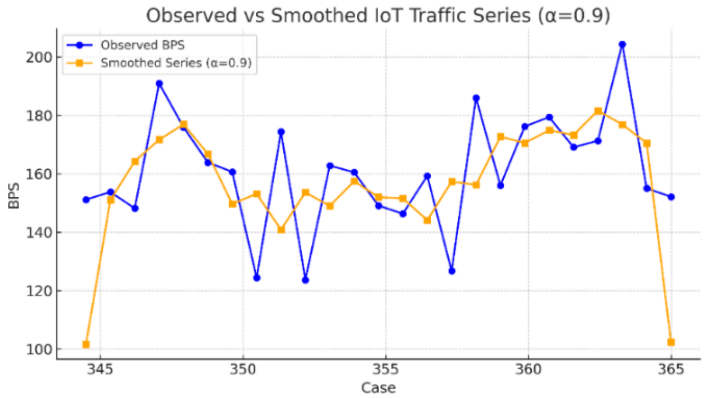


Figure 8. Exponential smoothing of IoT traffic with α=0.9

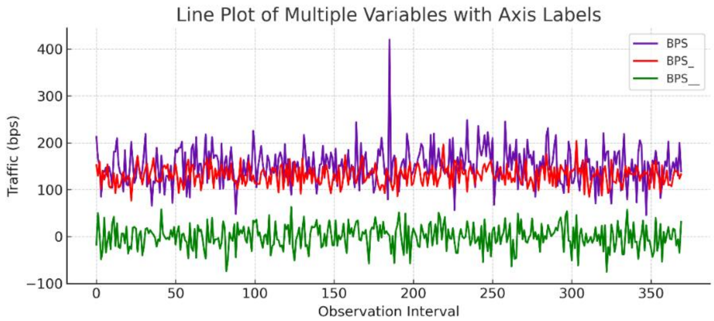


Figure 9. Exponential smoothing of IoT traffic with α=0.9

As shown in Figure 9, the smoothed series (bps_2) with α=0.9 more closely follows the fluctuations of the original traffic, including sharp peaks and drops. However, this results in significant variations in the residuals (bps_1), which exhibit increased amplitude and indicate the model's heightened sensitivity to noise. This outcome confirms that with higher values of α, the model becomes less stable and more prone to overfitting, thereby reducing the overall reliability of short-term IoT traffic forecasting.

To conduct a detailed comparative analysis of forecast accuracy, key error metrics were calculated for each smoothing coefficient (α=0.1, 0.5, and 0.9). Specifically, the following metrics were evaluated: ME, MAE, sum of squared errors (SSE), root mean squared error (RMSE), mean percentage error (MPE), and MAPE. A comparison of these metrics is presented in Figure 10 as a line chart, clearly illustrating how forecasting errors vary with the selected smoothing parameter α.

As shown in the chart, the model with α=0.1 yields the lowest error values across most metrics: the ME is 0.15 bps, the MAE is 34.4 bps, the SSE is 701,154, and the RMSE is 1,948. With α=0.5, these metrics increase: ME is –0.08 bps, MAE rises to 38.1 bps, SSE reaches 883,083, and RMSE increases to 2,468. The highest errors are observed with α=0.9, where ME is –0.07 bps, MAE is 47.2 bps, SSE is 1,334,328, and RMSE reaches 3,706. Additionally, the MAPE is 25.83% for α=0.1, increases to 28.34% for α=0.5, and reaches 35.02% for α=0.9. These results confirm that a lower α value provides more stable and accurate forecasts by minimizing error and smoothing the influence of short-term fluctuations. In contrast, higher α values make the model more sensitive to noise, resulting in increased errors and reduced overall forecast reliability.

Thus, the comparative analysis demonstrates that the choice of an optimal smoothing coefficient is a critical factor in improving the quality of short-term forecasting for self-similar IoT traffic. To provide a detailed quantitative assessment of forecast accuracy using α=0.1, Table 2 presents the actual IoT traffic values, the corresponding forecasted values, and their absolute and relative errors. This format enables a precise evaluation of how closely the model reproduces real traffic behavior over a 32-minute forecast horizon.

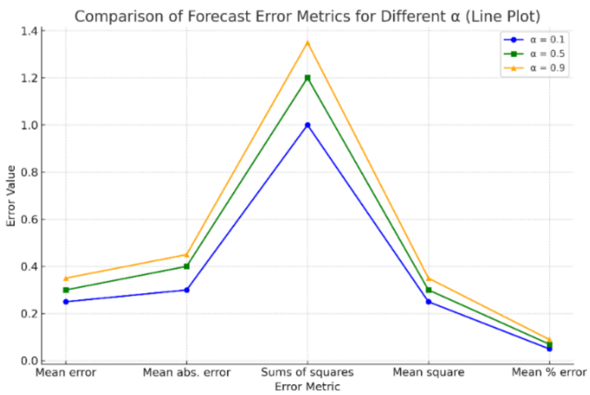


Figure 10. Comparison of forecast error metrics across different smoothing coefficients (α) using a line chart

Table 2. Forecast accuracy evaluation using exponential smoothing with A=0.1, MAPE: 31.1%

Time (min)	Actual value (X_a)	Forecast value ($\hat{X}$)	Absolute error ($\lvert X_a - \hat{X} \rvert$)	Relative error ($\lvert X_a - \hat{X} \rvert / X_a$)
8	100	146.44	46.44	0.464
16	190	146.44	43.56	0.229
24	110	146.44	36.44	0.331
32	120	146.44	26.44	0.22

These results confirm that with a low smoothing coefficient α, the model focuses more on the general trend while minimizing the influence of short-term spikes and noise, making it preferable for controlled load forecasting in IoT networks. Thus, the analysis demonstrates that the choice of smoothing coefficient has a significant impact on forecast behavior: with a low α, the model produces lower errors and more stable tracking of the overall traffic trend, whereas higher α values increase sensitivity to short-term fluctuations and noise, thereby reducing overall accuracy. These findings provide a crucial foundation for further research and practical applications, as they enable network operators to more effectively plan capacity and maintain a high level of QoS under the highly variable conditions of IoT traffic.

4. DISCUSSION

The experimental results confirm that exponential smoothing can serve as a computationally efficient method for predicting short-term IoT traffic, especially in resource-constrained environments. The findings show a clear relationship between the smoothing coefficient (α) and forecasting accuracy. The best performance was achieved when α was set to 0.1, resulting in the lowest MAPE (25.82%) and stable forecast curves, which indicates that lower smoothing coefficients effectively suppress noise and highlight long-term dependencies in self-similar traffic. These results align with the conclusions of Güler *et al.* [19], who applied a modified exponential smoothing method for anomaly-aware traffic forecasting in IoT edge networks, demonstrating its simplicity and reliability. Similarly, Tran *et al.* [20] employed double seasonal exponential smoothing to predict smart grid IoT traffic, observing high precision in handling periodic fluctuations.

The behavior observed in this study, where increased α values (0.5 and 0.9) led to a decline in forecast stability and accuracy, is consistent with the sensitivity analysis reported by Lykakis *et al.* [23], who compared statistical and machine learning methods for bandwidth-constrained IoT networks. While hybrid models often offer higher accuracy, their complexity makes them less suitable for edge devices. This was further emphasized by Juliet *et al.* [22], who designed a lightweight hybrid approach combining neural elements with smoothing techniques, aiming to preserve both adaptability and low computational cost.

The self-similar nature of the IoT traffic, indicated by a Hurst exponent of H=0.5, underscores the need for models that can capture long-range dependencies. Harrou *et al.* [21] addressed this challenge by integrating wavelet decomposition and LSTM, showing robustness in modeling fractal behavior, though at a higher computational cost. In contrast, exponential smoothing—despite its simplicity—offers a practical trade-off between efficiency and predictive performance, particularly when used with an appropriately tuned α. Lastly, the broader relevance of this work is supported by the survey of Mystakidis [12], who emphasized the growing role of interpretable and scalable forecasting models in real-world traffic management, including in IoT and bright environments. The practical implications of the proposed approach are particularly relevant for edge computing and contemporary mobile networks. Since exponential smoothing requires minimal computational resources and operates on a single recursive update, it can be efficiently executed on IoT gateways, microcontrollers, and multi-access edge computing (MEC) nodes without introducing latency or excessive energy consumption. This makes the method suitable for deployment in resource-constrained environments where traditional machine learning models may be impractical. In mobile and 5G/6G networks, short-term traffic forecasting can support dynamic bandwidth allocation, adaptive scheduling, and congestion avoidance, enabling base stations to anticipate traffic bursts generated by dense IoT deployments. Furthermore, accurate minute-ahead predictions can enhance QoS management, optimize handover processes, and assist in real-time anomaly detection. These practical benefits demonstrate that even a lightweight forecasting technique can significantly improve operational efficiency across diverse IoT-driven network infrastructures. In summary, while exponential smoothing cannot fully replace more complex models in all scenarios, it provides a lightweight and interpretable baseline suitable for real-time prediction in self-similar IoT data streams. The results also suggest that careful parameter selection is essential to maximize the model's effectiveness under varying network dynamics.

5. CONCLUSION

In this study, a statistical approach using simple exponential smoothing was developed and applied to analyze and forecast short-term self-similar IoT traffic in telecommunication networks. The proposed methodology aimed to improve the accuracy of traffic prediction and support real-time resource management under dynamic conditions. The research involved statistical analysis of real traffic data, calculation of the Hurst exponent, and evaluation of different smoothing coefficients (α=0.1, 0.5, and 0.9) using the STATISTICA software environment.

The main conclusions drawn from the research are as follows: i) the calculated Hurst exponent (H=0.5) confirmed the fractal and long-term dependent nature of the IoT traffic, ii) the highest prediction accuracy was achieved with α=0.1, resulting in a MAPE of 25.82%, a ME of 6.44 bps, and a MAE of 20.22 bps, iii) higher values of α (0.5 and 0.9) led to reduced stability and increased forecast error due to greater sensitivity to local fluctuations and noise, iv) the predicted traffic over a 32-minute horizon decreased with increasing α, highlighting the trade-off between responsiveness and robustness, v) comparative analysis and visualizations confirmed that choosing a low α value ensures more accurate and stable forecasting for self-similar traffic, and vi) future work will focus on hybrid models integrating machine learning algorithms and adaptive smoothing techniques further to enhance prediction performance in highly variable IoT environments.

This analytical framework demonstrates strong potential for efficient network traffic forecasting and QoS assurance in modern telecommunication systems. However, this study also has several limitations that should be acknowledged. The forecasting results were obtained using a single IoT traffic dataset collected from one network environment, which restricts the generalizability of the findings. Since traffic patterns may vary significantly across different IoT platforms, device types, communication protocols, and network loads, the performance of the exponential smoothing method cannot be assumed to be universally applicable. Furthermore, the study does not evaluate the impact of different sampling frequencies, data aggregation levels, or noise conditions, which may influence the robustness of the model. Future work should therefore include experiments on multiple heterogeneous datasets and diverse operational settings to better assess the scalability and generalization potential of the proposed approach.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Almira Mukhamejanova	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Katipa Chezhimbayeva	✓		✓		✓		✓	✓		✓			✓	✓
Samal Kaliyeva	✓		✓	✓			✓			✓	✓		✓	✓
Eleonora Lechshinskaya	✓		✓	✓			✓			✓	✓		✓	✓
Kumyssay Tumanbayeva		✓				✓		✓	✓	✓	✓	✓		
Yuliya Garmashova		✓			✓		✓			✓		✓		✓
Tolganay Abisheva		✓				✓		✓	✓	✓	✓	✓		
Inkar Zhumay	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, Almira Mukhamejanova, upon reasonable request. Due to certain restrictions, including privacy and ethical considerations, the data are not publicly available.

REFERENCES

[1] B. Saxena and V. Saxena, "Hurst exponent based approach for influence maximization in social networks," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 5, pp. 2218–2230, May 2022, doi: 10.1016/j.jksuci.2019.12.010.

[2] Y. Feng *et al.*, "Unmasking the internet: A survey of fine-grained network traffic analysis," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 6, pp. 3672–3709, Dec. 2025, doi: 10.1109/COMST.2025.3545541.

[3] I. Ullah, H. Y. Youn, and Y.-H. Han, "Integration of type-2 fuzzy logic and dempster–shafer theory for accurate inference of IoT-based health-care system," *Future Generation Computer Systems*, vol. 124, pp. 369–380, Nov. 2021, doi: 10.1016/j.future.2021.06.012.

[4] R. Fontugne, P. Abry, K. Fukuda, K. Cho, G. Didier, and H. Wendt, "Eigen-value based multivariate selfsimilarity analysis in internet traffic: A case study," in *2025 33rd European Signal Processing Conference (EUSIPCO)*, Sep. 2025, pp. 2587–2591, doi: 10.23919/EUSIPCO63237.2025.11226771.

[5] L. Gunasekara *et al.*, "A systematic review of responsible artificial intelligence principles and practice," *Applied System Innovation*, vol. 8, no. 4, Jul. 2025, doi: 10.3390/asi8040097.

[6] S. Radhakrishnan and W. Gueaieb, "A state-of-the-art review on topology and differential geometry-based robotic path planning—part I: planning under static constraints," *International Journal of Intelligent Robotics and Applications*, vol. 8, no. 2, pp. 435–454, Jun. 2024, doi: 10.1007/s41315-024-00330-5.

[7] X. B. Zhai, S. Fu, C. Yi, Z. Liu, C. Dong, and C. W. Tan, "Deep Reinforcement Learning-Based Task Offloading With Collaborative Inference in UAV-Assisted Mobile Edge Computing Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 27, no. 1, pp. 472-482, Jan. 2026, doi: 10.1109/TITS.2025.3629117.

[8] J. Tang, G. Wu, M. M. Jalalzai, L. Wang, B. Zhang, and Y. Zhou, "Energy-optimal DNN model placement in UAV-enabled edge computing networks," *Digital Communications and Networks*, vol. 10, no. 4, pp. 827–836, Aug. 2024, doi: 10.1016/j.dcan.2023.02.003.

[9] R. Sil, "Random forest based legal prediction system," in *Proceedings of International Joint Conference on Advances in Computational Intelligence*, 2022, pp. 623–633, doi: 10.1007/978-981-19-0332-8_46.

[10] A. Tripathi, I. Pathak, and D. P. Vidyarthi, "Integration of analytic network process with service measurement index framework for cloud service provider selection," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 12, pp. 1–16, Jun. 2017, doi: 10.1002/cpe.4144.

[11] S. P. Khedkar, R. A. Canessane, and M. L. Najafi, "Prediction of traffic generated by IoT devices using statistical learning time series algorithms," *Wireless Communications and Mobile Computing*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/5366222.

[12] A. Mystakidis, P. Koukaras, and C. Tjortjis, "Advances in Traffic Congestion Prediction: An Overview of Emerging Techniques and Methods," *Smart Cities*, vol. 8, no. 1, Feb. 2025, doi: 10.3390/smartcities8010025.

[13] T. Shen, L. Zhang, R. Geng, S. Li, and B. Sun, "Traffic prediction for diverse edge IoT data using graph network," *Journal of Cloud Computing*, vol. 13, no. 1, Apr. 2024, doi: 10.1186/s13677-023-00543-2.

[14] D. Bianchini *et al.*, "Challenges in AI-supported process analysis in the italian judicial system: what after digitalization?," *Digital Government: Research and Practice*, vol. 5, no. 1, pp. 1–10, Mar. 2024, doi: 10.1145/3630025.

[15] S. Hiremath *et al.*, "A new approach to data analysis using machine learning for cybersecurity," *Big Data and Cognitive Computing*,

vol. 7, no. 4, Nov. 2023, doi: 10.3390/bdcc7040176.
[16] A. Orynbayeva, N. Shyndaliyev, and A. Aripbayeva, “Improving statistical methods of data processing in medical universities using machine learning,” *World Transactions on Engineering and Technology Education*, vol. 21, no. 1, pp. 58–63, 2023.
[17] P. K. Saha and D. Logofatu, “Efficient approaches for data augmentation by using generative adversarial networks,” in *International Conference on Engineering Applications of Neural Networks*, 2022, pp. 386-399, doi: 10.1007/978-3-031-08223-8_32.
[18] L. Abdykerimova *et al.*, “Analysis of the emotional coloring of text using machine and deep learning methods,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 3, pp. 3055–3063, Jun. 2024, doi: 10.11591/ijece.v14i3.pp3055-3063.
[19] A. K. Güler, H. F. Alsan, and T. Arsan, “Anomaly Detection and Performance Analysis With Exponential Smoothing Model Powered by Genetic Algorithms and Meta Optimization,” *IEEE Access*, vol. 13, pp. 33361-33378, 2025, doi: 10.1109/ACCESS.2025.3543030.
[20] Q. T. Tran, L. Hao, and Q. K. Trinh, “Cellular network traffic prediction using exponential smoothing methods,” *Journal of Information and Communication Technology*, vol. 18, no. 1, 1-18, 2019, doi: 10.32890/jict2019.18.1.1.
[21] F. Harrou, A. Zeroual, F. Kadri, and Y. Sun, “Enhancing road traffic flow prediction with improved deep learning using wavelet transforms,” *Results in Engineering*, vol. 23, 2024, doi: 10.1016/j.rineng.2024.102342.
[22] A. H. Juliet, P. Malathi, B. Geetha, S. S. Kumar, and S. V, “Edge-AI Enabled IoT Framework for Real-Time Urban Traffic Flow Prediction and Congestion Avoidance,” in *2026 9th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, India, 2026, pp. 1-6, doi: 10.1109/IEMENTech202669403.2026.11434391.
[23] E. Lykakis, I. O. Vardiambasis, and E. Kokkinos, “Data Traffic Prediction for 5G and Beyond: Emerging Trends, Challenges, and Future Directions: A Scoping Review,” *Electronics*, vol. 14, no. 23, 4611, 2025, doi: 10.3390/electronics14234611.
[24] D. Wei, N. Xi, X. Ma, M. Shojafar, S. Kumari, and J. Ma, “Personalized Privacy-Aware Task Offloading for Edge-Cloud-Assisted Industrial Internet of Things in Automated Manufacturing,” in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 7935-7945, Nov. 2022, doi: 10.1109/TII.2022.3159822.
[25] V. Papastefanopoulos, P. Linardatos, T. Panagiotakopoulos, and S. Kotsiantis, “Multivariate time-series forecasting: A review of deep learning methods in internet of things applications to smart cities,” *Smart Cities*, vol. 6, no. 5, pp. 2519-2552, 2023, doi: 10.3390/smartcities6050114.

BIOGRAPHIES OF AUTHORS

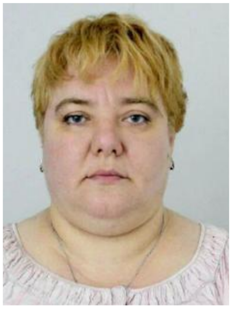
Almira Mukhamejanova Graduated from the Almaty Institute of Power Engineering and Telecommunications (1998) with a degree in automatic telecommunications. In 2023, she defended her doctoral thesis on the profile 6D071900 – radio engineering, electronics and telecommunications, Ph.D., she is an Associate Professor at the Department of Telecommunication Engineering at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan. She is the author of more than 50 scientific papers, including 4 textbooks, 3 articles in the Scopus database. Her research interests include analysis of telecommunication network traffic, and tele traffic theory. She can be contacted at email: a.mukhamejanova@aues.kz.

Katipa Chezhimbayeva graduated from the Almaty Power Engineering Institute in 1996, specializing in Automatic Telecommunications. In 2008, she defended her dissertation in the specialty 05.12.17-Radio Engineering and Television Systems and Devices" and received a candidate of technical sciences degree. Currently, she is a professor at the Department of Telecommunication Engineering, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. Author of more than 200 scientific papers, including 5 textbooks and more than 5 teaching aids, 18 articles in the Scopus database. Research interests include signal processing, application of sensors, IoT/M2M technologies, IP telephony, wireless communications, blockchain technologies, transport optical networks, and systems. She can be contacted at email: k.chezhimbayeva@aues.kz.

Samal Kaliyeva Ph.D., Senior Lecturer. She works at the Department of Radio Engineering, Electronics and Telecommunications at L.N. Gumilyov Eurasian National University. She has over 25 years of scientific and teaching experience. Author of more than 45 scientific publications, including 10 articles indexed in the Scopus database. Research interests: quality of service in multiservice networks, monitoring and management of transport communication networks, and internet of things (IoT). She can be contacted at email: s.kaliyeva@mail.ru.

Eleonora Lechshinskaya graduated from the Kazakh Polytechnic Institute with a degree in Automation and Telemechanics. In 1993 she defended her dissertation and received a Ph.D. Currently, she is an Associate Professor at the Department of Telecommunication Engineering at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. She is the author of more than 120 publications. Research interests: mathematical modeling of telecommunication systems, forecasting of processes in telecommunications, and statistical analysis of network traffic. She can be contacted at email: e.leshinskaya@aues.kz.

Kumyssay Tumanbayeva graduated from the Kazakh State University in 1971, specializing in Mathematics. In 1991, she defended her dissertation in the specialty "05.13.07 - Automation of technological processes and production", and received a candidate of technical sciences degree. Currently, she holds the position of professor at the Department of Telecommunication Engineering at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. She is the author of more than 180 scientific and scientific-methodical works, including 3 articles in the Scopus database. Scientific interests - mathematical modeling of telecommunication systems, and tele traffic theory. She can be contacted at email: k.tumanbayeva@aues.kz.



Yuliya Garmashova graduated from the Almaty Power Engineering Institute in 1995, specializing in Automatic Telecommunications. In 2013, she defended her master's thesis in the specialty "6M071900-Radio Engineering, Electronics and Telecommunications" and received an academic degree of Master of Technical Sciences. Currently, she is a practicing professor at the Telecommunications Engineering Department of the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. Author of more than 16 scientific papers, including 3 articles in the Scopus database. Research interests include video surveillance systems, breg arrays, optical networks, mobile networks, the internet of things, IoT networks, and the industrial internet of things (IIoT). She can be contacted at email: j.garmashova@aues.kz.

Tolganay Abisheva graduated from the Almaty Power Engineering Institute in 1995, and received a master's degree in Radio Engineering, Electronics and Telecommunications from the Humanitarian University of Transport and Law named after D.A. Kunaev in 2013. She is a senior lecturer at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. She is also a Bolashak fellow at the Technical University of Berlin, Germany. Her research interests include optical communication systems, wireless communication technologies, and measurement methods, and tools in telecommunications. She can be contacted at email: t.abisheva@aues.kz.

Inkar Zhumay in 2016 she graduated from the Almaty University of power engineering and telecommunications named after Gumarbek Daukeyev with a Master degree of technical science. Since 2019, a doctoral student at the Department of Telecommunication Engineering, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. Currently, she is a Lecturer at the Heriot-Watt University Branch at the K. Zhubanov Aktobe Regional University in Kazakhstan. Scientific interests - anomaly detection, pattern recognition theory, image processing, data mining, natural language processing, network security, and intrusion detection systems. She can be contacted at email: I.Zhumay@hw.ac.uk.