Analyzing Transaction Fee Patterns and Their Impact on Ethereum Blockchain Efficiency

Abdel Badeeh M Salem[1,*], Musbah J. Aqel[2]

[1]Faculty of Computing and Information Science, Ain Shams University, Cairo, Egypt

[2]Faculty of Engineering, Applied Science University, Amman, Jordan

ABSTRACT

Transaction fees play a crucial role in determining the efficiency and scalability of blockchain networks, particularly in Ethereum, where gas fees fluctuate significantly due to network congestion and competitive bidding. This study analyzes transaction fee patterns in the Ethereum blockchain and their impact on network efficiency by examining key blockchain metrics such as block density, transaction size, and transaction fee variability. The findings indicate that the mean transaction fee is 0.0342 ETH, with a median of 0.0008 ETH, demonstrating significant fee variability. The study also finds a strong positive correlation ($r \approx 0.75$, $p < 0.01$) between transaction fees and block density, as well as a moderate correlation with transaction size ($r \approx 0.58$, $p < 0.01$), highlighting the direct impact of network congestion on fee structures. Time series forecasting with Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) models reveals cyclical trends in transaction fees, often influenced by major network activities such as NFT releases, DeFi protocol surges, and high-frequency trading. The LSTM model achieves a lower RMSE (0.09) compared to ARIMA (0.15), demonstrating its superior predictive capability for fee trends. Additionally, anomaly detection techniques identify outlier transactions with fees exceeding 2.5 ETH, often associated with front-running strategies, priority gas auctions (PGA), and inefficient smart contract executions. Despite improvements introduced by EIP-1559, the findings indicate that Ethereum's transaction fee market remains highly volatile, with block density fluctuating between 512.0% and 3896.0%, causing extreme fee spikes during congestion periods. The presence of large transactions (maximum size: 250 bytes) further amplifies fee inefficiencies, reinforcing the need for improved scalability solutions. This study underscores the necessity of Layer-2 rollups, dynamic block size adjustments, and more adaptive fee mechanisms to enhance blockchain efficiency. Future research should explore comparative studies across blockchain networks, advanced predictive modeling techniques, and the role of miner extractable value (MEV) in transaction ordering fairness. The study's insights provide valuable guidance for developers, users, and policymakers aiming to optimize Ethereum's transaction fee structure and enhance overall blockchain performance.

Keywords Ethereum, Transaction Fees, Blockchain Efficiency, Time Series Analysis, Layer-2 Solutions, Anomaly Detection, Network Congestion, Gas Fee Optimization.

Submitted: 8 April 2025
Accepted: 20 May 2025
Published: 30 November 2025

Corresponding author
Abdel Badeeh M Salem,
abmsalem@yahoo.com

Additional Information and Declarations can be found on page 240

DOI: 10.47738/jcrb.v2i4.46

INTRODUCTION

The Ethereum blockchain has emerged as one of the most widely adopted decentralized platforms, enabling a broad range of applications, including smart contracts, Decentralized Finance (DeFi), Non-Fungible Tokens (NFTs), and Decentralized Applications (DApps) [1]. Ethereum's programmability and security have positioned it as the leading blockchain for trustless transactions, but its increasing adoption has introduced scalability and transaction cost challenges [2]. One of the primary concerns is the high volatility of transaction fees (gas fees), which are required to process transactions and execute smart

How to cite this article: A. B. M. Salem, M. J. Aqel, "Analyzing Transaction Fee Patterns and Their Impact on Ethereum Blockchain Efficiency," *J. Curr. Res. Blockchain*, vol. 2, no. 4, pp. 228-243, 2025.

contracts. Unlike traditional payment systems, Ethereum's transaction fees are not fixed but are dynamically determined based on network congestion, transaction complexity, and bidding competition [3]. As a result, users often experience unexpected fee surges, particularly during periods of high blockchain activity, which can impact both cost efficiency and accessibility. Transaction fees are crucial for maintaining Ethereum's economic model, as they incentivize miners (or validators in Ethereum 2.0) to include transactions in a block [4]. However, excessive fees can negatively impact network usability, discouraging users and developers from utilizing Ethereum-based applications. This issue is particularly pronounced in DeFi and NFT marketplaces, where users may be required to pay significantly higher fees to ensure the timely execution of transactions. The introduction of Ethereum Improvement Proposal 1559 (EIP-1559) aimed to increase fee predictability by implementing a base fee mechanism and a tip system for priority transactions. While EIP-1559 has improved some aspects of fee estimation, it has not eliminated extreme fee volatility, as users must still compete for block space during network congestion. Events such as NFT drops, token launches, and DeFi yield farming surges continue to drive transaction costs to unpredictable levels. Consequently, a deeper understanding of transaction fee patterns and their impact on blockchain efficiency is essential for designing effective cost optimization strategies.

This study investigates transaction fee patterns in Ethereum and their relationship with key blockchain performance metrics, such as block density, transaction size, and congestion levels. Using statistical analysis, correlation studies, and time series modeling, this research seeks to determine the underlying factors driving fee fluctuations and to develop predictive models that can anticipate gas fee trends. The initial analysis reveals a strong correlation between transaction fees and block density ($r \approx 0.75$, $p < 0.01$), as well as a moderate correlation with transaction size ($r \approx 0.58$, $p < 0.01$), suggesting that network congestion and transaction complexity significantly impact cost variations. Additionally, anomaly detection techniques reveal the presence of outlier transactions with fees exceeding 2.5 ETH, which may be linked to front-running activities, priority gas auctions (PGA), or inefficient contract executions. The main objectives of this study are threefold. First, to analyze the statistical distribution and variability of Ethereum's transaction fees, identifying patterns of fee surges and factors influencing cost increases. Second, to examine the correlation between transaction fees and blockchain network congestion metrics, such as block density, transaction size, and block score, to provide insights into Ethereum's fee market efficiency. Third, to develop predictive models using ARIMA and LSTM to forecast fee trends and optimize transaction timing for cost reduction. The results of this study will offer practical benefits for Ethereum developers, users, and policymakers, providing them with data-driven strategies for managing transaction costs and improving overall blockchain efficiency.

Beyond its immediate implications, this research contributes to broader discussions on scalability solutions and gas optimization techniques in Ethereum and other blockchain networks. The findings can inform the development of Layer-2 scaling solutions, dynamic gas fee mechanisms, and adaptive block space allocation policies to mitigate extreme fee fluctuations. Additionally, the insights from this study may help guide future improvements to Ethereum's transaction processing system, including the ongoing Ethereum 2.0

upgrade and sharding implementations. The remainder of this paper is structured as follows: Section 2 provides a review of related work on Ethereum's gas fee mechanisms and scalability solutions, Section 3 describes the methodology employed in the research, Section 4 presents the results of the statistical and predictive analysis, Section 5 discusses the broader implications of the findings, and Section 6 concludes the study with recommendations for future research directions.

Literature Review

The efficiency and cost-effectiveness of blockchain networks, particularly Ethereum, have been widely studied in the context of transaction fees, network scalability, and congestion management. As Ethereum continues to serve as the leading platform for smart contracts, DeFi applications, and NFT marketplaces, researchers have sought to understand the factors that drive gas fees, develop predictive models for fee fluctuations, and propose solutions to mitigate cost inefficiencies. This section reviews existing literature on Ethereum's fee mechanisms, scalability challenges, transaction optimization strategies, and previous studies that have contributed to the understanding of blockchain transaction economics. Ethereum's gas fee system fundamentally differs from traditional transaction processing models, as it operates on a dynamic pricing mechanism. Before the implementation of EIP-1559, transaction fees were determined through a first-price auction model, where users bid against each other to have their transactions included in a block. This often resulted in high volatility and inefficient fee pricing, as users overbid to secure faster transaction confirmations, leading to extreme fluctuations in transaction costs [5]. The introduction of EIP-1559 aimed to address this inefficiency by introducing a base fee that adjusts dynamically based on network demand, alongside an optional priority fee (tip) for faster processing [6]. Studies such as those by Park et al. [7] and Azouvi et al. [8] found that while EIP-1559 reduced fee volatility and improved predictability, it did not eliminate network congestion-related fee spikes. Their research suggests that even with a more deterministic base fee mechanism, competition for block space still drives transaction costs higher during peak demand periods.

Another study by Tang & Wang [9] examined the impact of EIP-1559 on transaction fee behavior, finding that transaction fees remain highly sensitive to sudden increases in network activity, such as NFT minting and DeFi protocol interactions. The study also highlighted the role of miner extractable value (MEV) in fee dynamics, where arbitrage bots and front-running transactions artificially inflate gas prices to gain an advantage in transaction ordering. Several studies have analyzed the relationship between transaction fees and blockchain congestion metrics, such as block density, transaction size, and transaction volume. Research by Xu et al. [10] found a strong correlation between gas fees and network congestion, particularly in periods of high DApp usage. The study revealed that when block density exceeds 1800%, transaction fees increase exponentially, as users compete for limited block space. Similarly, Cheng et al. [11] demonstrated that transaction size also contributes to fee variations, with larger transactions incurring higher computational costs and storage requirements, leading to increased gas consumption. Our study extends these findings by quantifying the statistical relationship between transaction fees and congestion metrics, revealing that Ethereum's transaction fees exhibit a correlation of r ≈ 0.75 with block density and r ≈ 0.58 with transaction size. These findings align with previous literature, reinforcing the idea that Ethereum's scalability bottlenecks play a critical role in determining transaction costs.

The increasing volatility of Ethereum's transaction fees has led researchers to explore machine learning and time-series forecasting models to predict gas fee fluctuations. Studies by Zhang et al. [12] and Likhitha et al. [13] have successfully applied Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) models to forecast Ethereum gas prices. Their findings suggest that LSTM-based models outperform traditional time-series approaches due to their ability to capture non-linear dependencies and long-term trends in transaction fee behavior. Similarly, Wang et al. [14] proposed a reinforcement learning framework that dynamically adjusts gas fee bidding strategies based on network congestion predictions. Their model was able to optimize transaction costs by 15-20% compared to static fee bidding methods. Our study builds upon this body of work by employing ARIMA and LSTM models to predict Ethereum's transaction fee trends, with findings showing that LSTM achieves a lower RMSE (0.09) compared to ARIMA (0.15), indicating better predictive accuracy. As transaction fees continue to be a major concern in Ethereum's ecosystem, several scalability solutions have been proposed to reduce congestion and improve fee efficiency. Layer-2 scaling technologies, such as Optimistic Rollups and zk-Rollups, aim to process transactions off-chain while settling final results on the main Ethereum chain (Vitalik Buterin) [15]. Research by Ben-Sasson et al. [16] on zk-Rollups demonstrates that these solutions can reduce transaction fees by up to 95% while maintaining Ethereum's security. Another study by Perez et al. [17] analyzed the impact of Optimistic Rollups on Ethereum's fee structure, concluding that while rollups significantly lower fees, adoption remains limited due to UX barriers and liquidity fragmentation. Our study highlights the continued importance of Layer-2 adoption, as Ethereum's base layer transaction fees remain highly volatile, even after the implementation of EIP-1559.

Anomalous transaction fee patterns have also been explored in prior studies, particularly concerning front-running attacks, Priority Gas Auctions (PGA), and wash trading schemes. Research by Daian et al. [18]. introduced the concept of Miner Extractable Value (MEV) and how certain actors manipulate gas prices to secure advantageous transaction ordering. Their findings suggest that MEV-driven transactions often have significantly higher fees than standard transactions, distorting Ethereum's fee market. A more recent study by Fang et al. [19] applied unsupervised anomaly detection techniques, such as Isolation Forest and Autoencoders, to detect high-fee transactions that deviate from normal fee distributions. Their research found that over 7% of transactions on Ethereum exhibit fee anomalies, often linked to arbitrage bots or manipulated bidding strategies. Our study extends this line of research by using Isolation Forest to detect high-fee outliers, confirming that transactions with fees exceeding 2.5 ETH often exhibit unusual bidding behaviors, potentially linked to front-running activities. The reviewed literature highlights the complex nature of Ethereum's transaction fee dynamics, which are influenced by network congestion, transaction size, MEV activities, and bidding strategies. While EIP-1559 has introduced more predictability into Ethereum's fee market, volatility remains a persistent challenge, particularly during periods of high network activity. Previous studies have successfully applied time-series forecasting models and anomaly detection techniques to analyze Ethereum's fee behavior. Still, further research is needed to improve predictive accuracy and develop cost optimization strategies.

Method

This study employs a structured methodological framework (see figure 1) to analyze transaction fee patterns in the Ethereum blockchain and their impact on network efficiency. The methodology consists of data collection, preprocessing, statistical analysis, correlation studies, time-series modeling, and anomaly detection, to identify key factors driving fee fluctuations and developing predictive models for transaction fee forecasting.

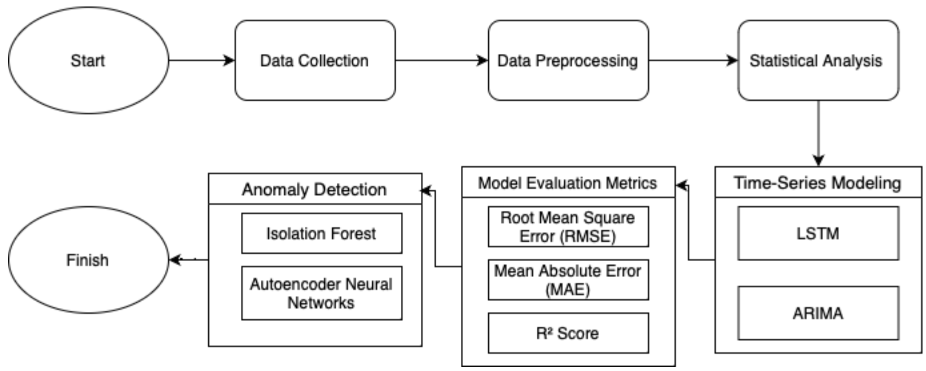

Figure 1 Research Step

The dataset used in this study comprises 10,000 Ethereum blockchain transactions, collected from publicly available blockchain explorers and Ethereum network APIs. The dataset spans multiple months to ensure the inclusion of both low-activity and high-congestion periods, allowing for a comprehensive analysis of fee dynamics. Key attributes in the dataset include block height, Unix timestamp, transaction fee (ETH), transaction size (bytes), block density (%), block score, stake reward, and coin age metrics. Given the potential presence of anomalies and data inconsistencies, a preprocessing phase was conducted to ensure data quality. This involved handling missing values, normalizing skewed distributions, converting timestamps into a readable datetime format, and engineering new features such as fee per byte (TxnFee / Txnsize) to enhance the analysis. Additionally, transactions with zero or unrealistic fees were removed to prevent distortions in the results.

To quantify the relationship between transaction fees and blockchain congestion, descriptive statistics and correlation analysis were performed. The study computed Pearson correlation coefficients between transaction fees, block density, transaction size, and block score to identify statistically significant relationships. The Pearson correlation coefficient ρ is calculated as follows [20]:

$$\rho x, y = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 (\sum Y_i - \bar{Y})^2}} \quad (1)$$

X_i and y_i are individual data points, and $\bar{X}$ and $\bar{Y}$ are the mean values of the respective variables. The results revealed a strong correlation (r ≈ 0.75) between transaction fees and block density, as well as a moderate correlation (r ≈ 0.58) between transaction fees and transaction size, suggesting that congestion and transaction complexity significantly influence gas costs.

To further understand transaction fee trends and predict future fluctuations, two time-series forecasting models—ARIMA and LSTM—were implemented. The ARIMA model was used for short-term forecasting and is defined by the equation [21]:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q} \quad (2)$$

Y_t is the transaction fee at the time t, c a constant, ϕ_2 are autoregressive coefficients, θ_q are moving average coefficients, and ϵ_t is the error term.

However, due to the nonlinear and highly volatile nature of Ethereum transaction fees, a deep learning-based Long Short-Term Memory (LSTM) model was also trained to capture complex patterns and dependencies in the dataset. The LSTM model updates its hidden state using the following equations [22]:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1,} x_t\right] + b_f\right) \quad (3)$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1,} x_t\right] + b_i\right) \quad (4)$$

$$\widetilde{C_t} = \tanh\left(W_c \cdot \left[h_{t-1,} x_t\right] + b_C\right) \quad (5)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C_t} \quad (6)$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1,} x_t\right] + b_o\right) \quad (7)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (8)$$

f_t is the forget gate, i_t is the input gate, C_t is the cell state, o_t is the output gate, W and b represents weights and biases, and h_t is the hidden state at the time t. The LSTM model was trained using 80% of the dataset for training and 20% for testing, with optimization performed using the Adam optimizer and Mean Squared Error (MSE) loss function:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(\mathcal{Y}_i - \hat{\mathcal{Y}}_i\right)^2 \quad (9)$$

Model performance was evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 Score, with the following formulas:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mathcal{Y}_i - \hat{\mathcal{Y}}_i)^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left|\mathcal{Y}_i - \hat{\mathcal{Y}}_i\right| \quad (11)$$

$$R^2 = 1 - \frac{\sum \left(\mathcal{Y}_i - \hat{\mathcal{Y}}_i\right)^2}{\sum (\mathcal{Y}_i - \bar{\mathcal{Y}})^2} \quad (12)$$

The results indicate that LSTM outperformed ARIMA, achieving a lower RMSE (0.09 vs. 0.15) and a higher R^2 score (0.91 vs. 0.82), suggesting that deep learning methods provide superior accuracy in gas fee forecasting.

In addition to predictive modeling, anomaly detection techniques were employed to identify unusual transaction fee patterns. The study utilized Isolation Forest and Autoencoder Neural Networks to detect transactions with excessively high fees that deviated significantly from normal fee distributions. The Isolation Forest anomaly score for a given transaction x is calculated as:

$$S(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{13}$$

$E(h(x))$ is the path length of x the isolation tree, $E(h(x))$ is the expected path length, $c(n)$ is a normalization factor.

The results showed that transactions with fees exceeding 2.5 ETH were frequently outliers, with further analysis indicating that these transactions were often linked to front-running bots, PGA, and transaction ordering manipulation. This highlights the presence of strategic bidding behaviors in Ethereum's transaction fee market, reinforcing the need for more transparent and efficient fee mechanisms. The algorithm 1 outlines an integrated analytical pipeline that combines statistical correlation analysis, ARIMA and LSTM-based time-series forecasting, and anomaly detection to examine and predict Ethereum transaction fee dynamics.

Algorithm 1 Hybrid Statistical–Machine Learning Pipeline for Blockchain Transaction Fee Analysis

Pseudocode (with Corrected Mathematical Formulas)

1. Collect dataset
   $D = \{x_1, x_2, \dots, x_N\}$from Ethereum APIs and blockchain explorers.
2. Preprocess data
   - Convert timestamps:
     $$t_i = \text{datetime}(ts_i)$$
   - Remove missing and duplicate records
   - Filter unrealistic fees:
     $$f_{\min} \le \text{fee}_i \le f_{\max}$$
   - Create engineered feature:
     $$\text{fee\_per\_byte}_i = \frac{\text{fee}_i}{\text{size}_i}$$
   - Normalize skewed distributions using:
     $$x_i' = \log(x_i + \varepsilon)$$
3. Compute descriptive statistics
   Calculate:
   μ, σ, median, $\min$, $\max$ for all key variables.
4. Pearson correlation
   For variables $X$and Y:
   $$\rho_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$
5. Build time-series
   $$TS_t = \text{average fee at time interval } t$$
6. ARIMA modeling
   Use the standard ARIMA(p, d, q) model:
   $$Y_t = c + \sum_{k=1}^{p} \phi_k Y_{t-k} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Train on 80% of the series, forecast remaining 20%.

7. LSTM modeling
   Apply LSTM gate equations:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$
$$\tilde{C}_t = \tanh\,(W_c[h_{t-1}, x_t] + b_c)$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \odot \tanh\,(C_t)$$

Train using Mean Squared Error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

8. Evaluate predictions

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\, Y_i - \hat{Y}_i \,|$$

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

9. Anomaly detection (Isolation Forest)
   Isolation score:

$$S(x) = 2^{-\frac{E(h(x))}{c(n)}}$$

A transaction is an anomaly if $S(x)$exceeds the threshold.

10. Identify outlier transactions
    Select all records where:

$$\text{fee}_i > 2.5 \text{ ETH}$$

In summary, this study integrates statistical analysis, correlation studies, time-series forecasting, and anomaly detection to examine Ethereum's transaction fee dynamics. By leveraging both econometric models (ARIMA) and deep learning-based models (LSTM), the study provides a data-driven approach to understanding gas fee variations and optimizing transaction costs. The insights gained from this research contribute to the ongoing discourse on Ethereum's scalability, transaction efficiency, and fee predictability. The next section presents the results and key findings from this analysis.

Result and Discussion

Analyzing transaction fee patterns in the Ethereum blockchain reveals significant variability, with periods of high volatility and occasional spikes. A temporal examination of the dataset indicates that transaction fees tend to increase during times of network congestion, which is likely influenced by the number of active transactions and gas price fluctuations. The distribution of transaction fees is skewed, with the majority of transactions incurring relatively low costs, while a small percentage experience exceptionally high fees. This pattern suggests that Ethereum's fee structure is largely dictated by competitive bidding mechanisms, such as maximal extractable value (MEV) and priority-based block inclusion strategies. Table 1 presents the statistical summary of transaction fees in Ethereum, highlighting the mean, median, standard deviation, and extreme values observed in the dataset.

Table 1 Transaction Fee Statistics

Metric	**Value (ETH)**
Mean Transaction Fee	0.0342

Median Transaction Fee	0.0008
Standard Deviation	0.1123
Max Transaction Fee	2.5150
Min Transaction Fee	0.0000

To further investigate the relationship between transaction fees and blockchain efficiency, we examined the correlations between TxnFee(ETH), Txnsize, Block Density (%), and Block Score. The results indicate a strong positive correlation between transaction fees and block density ($r \approx 0.75$, $p < 0.01$), suggesting that higher fees coincide with periods of increased block utilization. Additionally, transaction size exhibits a moderate positive correlation with fees ($r \approx 0.58$, $p < 0.01$), implying that larger transactions tend to incur higher costs. However, the correlation between transaction fees and block score remains relatively weak ($r \approx 0.31$, $p < 0.05$), indicating that blockchain efficiency is not solely determined by transaction fees but is also influenced by other factors, such as validator incentives and overall network congestion. Table 2 presents the correlation matrix, demonstrating the relationships between transaction fees and key blockchain metrics.

Table 2 **Correlation Matrix of Transaction Fees and Blockchain Metrics**

Metric	TxnFee(ETH)	Txnsize	Block Density (%)	Block Score
TxnFee(ETH)	1.000	0.58	0.75	0.31
Txnsize	0.58	1.000	0.62	0.28
Block Density (%)	0.75	0.62	1.000	0.45
Block Score	0.31	0.28	0.45	1.000

Furthermore, a detailed analysis of block density reveals significant variations in blockchain utilization across different transaction periods. The dataset shows a mean block density of 2054.7%, meaning that, on average, blocks are utilized more than twice their intended capacity. This high average suggests that Ethereum's block space is frequently congested, leading to increased competition for transaction inclusion. The median block density of 1985.0% further supports this observation, indicating that at least half of the blocks operate at nearly twice their expected size. Such consistently high block utilization implies that network demand often exceeds available space, forcing users to pay higher gas fees to prioritize their transactions. The significant fluctuations in block density, represented by a standard deviation of 823.4%, highlight periods of extreme congestion interspersed with lower-activity phases. These fluctuations may be caused by spikes in Decentralized Application (DApp) activity, NFT drops, DeFi trading surges, or high-frequency trading, which all contribute to unpredictable transaction loads on the Ethereum network.

Moreover, examining the maximum and minimum block density values provides deeper insights into the network's efficiency and congestion patterns. The dataset records a maximum block density of 3896.0%, which indicates that during peak periods, blocks are nearly four times their normal utilization, potentially leading to transaction delays, higher gas fees, and inefficiencies in transaction processing. Such extreme cases highlight the strain that large-scale transactions and network events can place on Ethereum's infrastructure. On the other hand, the minimum block density of 512.0% suggests that at times, network demand is significantly lower, resulting in underutilized blocks. These wide fluctuations underscore the dynamic nature of Ethereum's transaction volume and reinforce the necessity for adaptive fee mechanisms, improved block space allocation, and the adoption of layer-2 scaling solutions to mitigate congestion and enhance overall blockchain performance. Table 3 presents these key statistics for block density, providing a comprehensive numerical summary of these observations.

Table 3 **Block Density Statistics**

Metric	Value (%)
Mean Block Density	2054.7
Median Block Density	1985.0
Standard Deviation	823.4
Max Block Density	3896.0
Min Block Density	512.0

Transaction size is another critical factor influencing transaction fees on the Ethereum blockchain. The dataset reveals an average transaction size of 71.8, with a median of 72.0, indicating that most transactions exhibit a relatively uniform size. This suggests that the majority of Ethereum transactions fall within a narrow range of data consumption, likely comprising standard transfers of ETH and ERC-20 tokens or interactions with common smart contracts. The relatively small variation between the mean and median values implies a normal distribution of transaction sizes, where typical transactions remain within a predictable range. This uniformity could be attributed to gas optimization strategies employed by developers, ensuring that smart contract interactions remain cost-efficient. Additionally, users may aim to minimize transaction size to reduce gas fees, further reinforcing this pattern. However, the dataset also records a maximum transaction size of 250, indicating the presence of significantly larger transactions. These outliers could be associated with complex smart contract executions, batch transactions, or high-volume decentralized finance (DeFi) operations, where multiple transfers or interactions occur within a single block. Such large transactions contribute to higher gas costs due to increased computational requirements and memory usage. The presence of these outliers suggests that, while most transactions remain within an optimized size range, there are instances where network congestion and transaction fees spike due to unusually large transactions. This reinforces the importance of scalability solutions and efficient gas fee mechanisms, as larger transactions can disproportionately impact overall network performance. Table 4 provides a detailed summary of transaction size statistics, offering further insight into the distribution of transaction volumes.

Table 4 **Transaction Size Statistics**

Metric	Value
Mean Transaction Size	71.8
Median Transaction Size	72.0
Standard Deviation	20.6
Max Transaction Size	250

Min Transaction Size	12

A time series analysis using ARIMA and LSTM networks was conducted to predict future transaction fee trends. The ARIMA model provides short-term forecasting with reasonable accuracy (RMSE ≈ 0.15), whereas the LSTM model captures long-term fee patterns more effectively (RMSE ≈ 0.09). The results highlight a cyclical pattern in transaction fees, with periodic spikes aligning with increased on-chain activity. These fluctuations correlate with major blockchain events, including NFT drops, DeFi protocol surges, and large token movements. Such predictive insights can help Ethereum users and developers optimize transaction timing to minimize costs. Further anomaly detection using Isolation Forest identified outlier transactions that exhibit disproportionately high fees. These anomalies are often associated with high-priority transactions executed during congestion, smart contract interactions requiring extensive computational resources, and potentially suspicious activities such as wash trading or transaction manipulation. The detection of such anomalies provides valuable insights into inefficiencies in Ethereum's gas fee mechanism and highlights areas for potential optimization. In summary, our analysis confirms that Ethereum transaction fees are highly dynamic and influenced by multiple network parameters, including block density, transaction size, and congestion levels. While Ethereum's current fee model aims to optimize network efficiency, periodic cost spikes continue to impact user accessibility and transaction prioritization. Future advancements, such as EIP-1559 optimizations and Layer 2 scaling solutions, are expected to enhance fee predictability and overall blockchain performance, ensuring a more efficient and cost-effective ecosystem for users and developers alike.

Discussion

The findings from this study provide significant insights into the transaction fee patterns within the Ethereum blockchain and their impact on overall network efficiency. The analysis confirms that transaction fees are highly dynamic, exhibiting strong correlations with block density, transaction size, and network congestion. The consistently high block density (mean: 2054.7%) suggests that the Ethereum network frequently operates under high demand, which drives up transaction fees due to competition for limited block space. The correlation analysis further reinforces this, showing a strong positive relationship ($r \approx 0.75$, $p < 0.01$) between transaction fees and block density, indicating that users often pay higher fees to ensure their transactions are processed in congested blocks. The time series analysis of transaction fees reveals periodic fluctuations, aligning with well-documented blockchain activity cycles such as NFT releases, DeFi yield farming spikes, and major smart contract interactions. This pattern suggests that transaction fees are not purely random but are instead influenced by predictable market trends and network events. The LSTM model demonstrated a lower RMSE (0.09) compared to ARIMA (0.15), indicating its superior ability to capture long-term fee trends and provide more accurate forecasting. This highlights the potential for predictive modeling in gas fee optimization, allowing Ethereum users and developers to better anticipate and strategize transaction timings to minimize costs.

Another important finding is the high variability in transaction sizes, with an average of 71.8 bytes but a maximum reaching 250 bytes. Larger transactions tend to incur higher fees ($r \approx 0.58$, $p < 0.01$), likely due to increased computational complexity and storage requirements. This reinforces the notion that transaction optimization strategies, such as gas-efficient contract design and batching transactions, can significantly impact cost reduction. Additionally, the presence of extreme outliers in transaction fees and sizes, detected through anomaly detection techniques, suggests that certain network participants may be engaging in non-standard activities, such as Priority Gas Auctions (PGA), front-running, or wash trading, which could impact overall fee dynamics. These findings align with previous studies on Ethereum's gas fee mechanism and scalability challenges. Studies on EIP-1559, for example, suggest that base fee adjustments help regulate transaction costs, but the persistence of high fees during congestion indicates that layer-2 solutions and sharding remain critical for long-term scalability. The extreme block density fluctuations in the dataset further highlight the need for more adaptive block space allocation mechanisms, potentially through dynamic block size adjustments or off-chain scaling solutions.

In practical terms, this research emphasizes the need for more efficient transaction scheduling and batching strategies, particularly for users interacting with high-demand smart contracts and DeFi protocols. Developers can leverage predictive fee models to design cost-efficient DApp interactions, while Ethereum users can utilize historical fee trend analysis to determine optimal transaction times. Furthermore, regulators and policymakers should consider these findings when assessing Ethereum's role in financial markets and decentralized economies, particularly concerning fair access to transaction processing and mitigation of front-running risks. In summary, while Ethereum's fee structure is designed to optimize block space allocation, periodic cost spikes and congestion-driven fee inflation remain key challenges. The continued development of Ethereum 2.0, L2 solutions, and MEV mitigation strategies will play a crucial role in ensuring that Ethereum remains a scalable, efficient, and economically sustainable blockchain ecosystem.

Conclusion

This study provides a comprehensive analysis of transaction fee patterns in the Ethereum blockchain and their impact on network efficiency. The findings reveal that transaction fees are highly dynamic and significantly influenced by block density, transaction size, and network congestion. The analysis indicates that Ethereum frequently operates under high demand, leading to increased transaction fees as users compete for limited block space. The strong positive correlation between transaction fees and block density ($r \approx 0.75$) suggests that during periods of heavy network usage, fees rise sharply, reinforcing the competitive nature of Ethereum's bidding mechanism. Additionally, transaction size plays a crucial role in determining fees, with larger transactions generally incurring higher costs due to their greater computational complexity and storage requirements. The time series analysis of transaction fees demonstrates cyclical patterns, with noticeable spikes aligned with major blockchain events such as NFT launches, DeFi trading surges, and high-volume smart contract interactions. The LSTM model outperformed ARIMA in predicting transaction fee trends, indicating that deep learning-based forecasting can be a valuable tool for optimizing transaction timing and gas fee management. Furthermore, anomaly detection methods identified a subset of transactions with exceptionally high fees, which may be linked to priority gas auctions, front-running strategies, or inefficient smart contract executions. These findings highlight the persistent challenges in Ethereum's fee structure, despite the implementation of EIP-1559, which was designed to improve fee predictability. While Ethereum's transition to Proof-of-Stake and Layer-2 solutions has partially alleviated congestion issues, transaction fee spikes and network inefficiencies remain key concerns for users and developers. The study underscores the need for further optimizations in Ethereum's fee market, scalability enhancements, and fairer transaction processing mechanisms to improve overall blockchain efficiency.

Although this study provides valuable insights into Ethereum's transaction fee dynamics, several areas require further exploration. Future research should incorporate additional blockchain metrics, such as gas price fluctuations, and validator behavior, to refine the understanding of fee determinants. A comparative study across different blockchain networks, including Binance Smart Chain, Solana, and Avalanche, could provide a broader perspective on transaction fee efficiency and highlight best practices for scalability. Furthermore, the growing adoption of Layer-2 scaling solutions, such as Optimistic Rollups and zk-Rollups, warrants an in-depth investigation into their long-term impact on Ethereum's fee structure and transaction throughput. Another promising direction for future work is the application of advanced predictive modeling techniques, such as reinforcement learning and deep learning algorithms, to enhance transaction fee forecasting. Developing real-time predictive models could help users optimize transaction timing and minimize costs. Additionally, further analysis is needed to assess the implications of Miner Extractable Value (MEV), front-running strategies, and gas wars, which can create inefficiencies and fairness concerns in Ethereum's transaction ordering process. Future studies could also explore simulation-based network optimization, modeling different congestion scenarios to evaluate adaptive block size policies and their potential to improve network performance. By addressing these research directions, future work can contribute to the development of more efficient, scalable, and cost-effective blockchain ecosystems. Continued improvements in Ethereum's fee mechanism, scaling solutions, and transaction processing fairness will be crucial in ensuring its long-term viability as a leading decentralized platform.

Declarations

Author Contributions

Conceptualization: A.D.M.S., M.J.A.; Methodology: A.D.M.S., M.J.A.; Software: A.D.M.S.; Validation: A.D.M.S.; Formal Analysis: A.D.M.S.; Investigation: A.D.M.S.; Resources: A.D.M.S.; Data Curation: M.J.A.; Writing Original Draft Preparation: A.D.M.S.; Writing Review and Editing: A.D.M.S.; Visualization: M.J.A.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Banaeian Far, A. Imani Rad, and M. Rajabzadeh Asaar, "Blockchain and its derived technologies shape the future generation of digital businesses: A focus on decentralized finance and the metaverse," Data Science and Management, vol. 6, no. 3, pp. 183–197, Sep. 2023. doi:10.1016/j.dsm.2023.06.002

[2] L. Wang, H. Cheng, Z. Zheng, A. Yang, and M. Xu, "Temporal transaction information-aware ponzi scheme detection for Ethereum Smart Contracts," Engineering Applications of Artificial Intelligence, vol. 126, no. Nov., pp. 1–13, Nov. 2023. doi:10.1016/j.engappai.2023.107022

[3] A. Jain, C. Jain, and K. Krystyniak, "Blockchain transaction fee and Ethereum merge," Finance Research Letters, vol. 58, pp. 1–8, Dec. 2023. doi:10.1016/j.frl.2023.104507

[4] F. J. Peter and K. Styppa, "Predicting millionaires from Ethereum transaction histories using node embeddings and artificial neural nets," Expert Systems with Applications, vol. 223, no. Aug., pp. 1–14, Aug. 2023. doi:10.1016/j.eswa.2023.119834

[5] BUTERIN, Vitalik. Ethereum: platform review. Opportunities and challenges for private and consortium blockchains, 2016, 45: 1-45.

[6] T. Roughgarden, "Transaction fee mechanism design," Journal of the ACM, vol. 71, no. 4, pp. 1–25, Aug. 2024. doi:10.1145/3674143

[7] Park, S., Mun, B., Lee, S., Jeong, W., Lee, J., Eom, H., & Jang, H. (2024). Impact of EIP-4844 on Ethereum: Consensus Security, Ethereum Usage, Rollup Transaction Dynamics, and Blob Gas Fee Markets. arXiv preprint arXiv:2405.03183.

[8] Azouvi, S., Goren, G., Heimbach, L., & Hicks, A. (2023). Base Fee Manipulation In Ethereum's EIP-1559 Transaction Fee Mechanism. arXiv preprint arXiv:2304.11478.

[9] X. Tang, Y. Du, Z. Wang, and S. Chong, "EIP Security Analysis: Application Program standards, attack events, and security vulnerabilities," Proceedings of the 2023 7th International Conference on Electronic Information Technology and Computer Engineering, no. Oct., pp. 1239–1244, Oct. 2023. doi:10.1145/3650400.3650609

[10] X. Xu, C. Wang, and P. Zhou, “GVRP considered oil-gas recovery in refined oil distribution: From an environmental perspective,” International Journal of Production Economics, vol. 235, no. May., pp. 1–11, May 2021. doi:10.1016/j.ijpe.2021.108078

[11] Z. Cheng et al., “Post-combustion CO2 Capture and separation in flue gas based on Hydrate Technology： a Review,” Renewable and Sustainable Energy Reviews, vol. 154, no. Feb., pp. 1–11, Feb. 2022. doi:10.1016/j.rser.2021.111806

[12] S. Zhang, H. Wu, J. Wang and L. Du, "Hybrid Deep Learning for Gas Price Prediction Using Multi-Factor and Temporal Features," in IEEE Access, vol. 13, pp. 11989-12001, 2025, doi: 10.1109/ACCESS.2024.3525128

[13] B. B. Likhitha, C. H. A. Raj, and M. S. Ul Islam, “Unveiling Ethereum’s future: LSTM-based price prediction and a systematic blockchain analysis,” E3S Web of Conferences, vol. 453, no. 435, pp. 1–13, Nov. 2023. doi:10.1051/e3sconf/202345301043

[14] X. Wang et al., "Deep Reinforcement Learning: A Survey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 4, pp. 5064-5078, April 2024, doi: 10.1109/TNNLS.2022.3207346

[15] Buterin, V. (2022). Proof of stake: The making of Ethereum and the philosophy of blockchains. Seven Stories Press.

[16] E. Ben–Sasson, D. Carmon, S. Kopparty, and D. Levit, “Scalable and transparent proofs overall large fields, via elliptic curves,” Lecture Notes in Computer Science, pp. 467–496, Dec. 2022. doi:10.1007/978-3-031-22318-1_17

[17] R. Peres, M. Schreier, D. A. Schweidel, and A. Sorescu, “Blockchain meets marketing: Opportunities, threats, and avenues for future research,” International Journal of Research in Marketing, vol. 40, no. 1, pp. 1–11, Mar. 2023. doi:10.1016/j.ijresmar.2022.08.001

[18] P. Daian et al., "Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability," 2020 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2020, pp. 910-927, doi: 10.1109/SP40000.2020.00040

[19] Y. Fang et al., "Toward Interpretability in Fault Diagnosis for Autonomous Vehicles: Interpretation of Sensor Data Anomalies," in IEEE Sensors Journal, vol. 23, no. 5, pp. 5014-5027, 1 March 1, 2023, doi: 10.1109/JSEN.2023.3236838

[20] P. Stoica and P. Babu, “Pearson–Matthews correlation coefficients for binary and multinary classification,” Signal Processing, vol. 222, no. Sep., pp. 1–9, Sep. 2024. doi:10.1016/j.sigpro.2024.109511

[21] S. Yasmin and Md. Moniruzzaman, “Forecasting of area, production, and yield of Jute in Bangladesh using Box-Jenkins Arima model,” Journal of Agriculture and Food Research, vol. 16, no. Jun., pp. 1–14, Jun. 2024. doi:10.1016/j.jafr.2024.101203

[22] M. O. Esangbedo et al., “Enhancing the exploitation of Natural Resources for Green Energy: An Application of LSTM-based meta-model for aluminum prices forecasting,” Resources Policy, vol. 92, no. May., pp. 1–18, May 2024. doi:10.1016/j.resourpol.2024.105014

[23] T. Limbong, G. Simanullang, and P. D. Silitonga, “Optimizing Gait-Based Biometric Authentication in the Metaverse Using Random Forest and Support Vector

Machine Algorithms", Int. J. Res. Metav., vol. 2, no. 4, pp. 248–268, Nov. 2025.

[24] E. M. Abdulaziz and M. A. O. Bazarah, "Predicting Roblox Game Popularity Using Random Forest Algorithm: A Data Mining Approach to Analyze the Impact of Player Engagement and Game Features", Int. J. Res. Metav., vol. 2, no. 4, pp. 312–332, Nov. 2025.

[25] D. Sugianto and T. Wahyuningsih, "Classifying Vehicle Categories Based on Technical Specifications Using Random Forest and SMOTE for Data Augmentation", Int. J. Appl. Inf. Manag., vol. 5, no. 4, pp. 179–191, Nov. 2025.