Temporal Analysis of Ethereum Blockchain Trends in Transaction Fees and Block Density Over Time

Ahmed Saeed Bahurmuz[1,*], Hani Atiahallah Alyoubi[2]

[1,2]Information Science Department, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT

Ethereum, as a leading blockchain platform, experiences high variability in transaction fees due to network congestion, gas bidding, and computational complexity. This study analyzes 10,000 Ethereum transactions to identify key factors influencing transaction fees, block density, and staking mechanisms. The results show that transaction fees vary significantly, with an average of 0.1826 ETH and a standard deviation of 0.2381 ETH, indicating substantial fluctuations. A strong positive correlation (r = 0.72) between transaction size and transaction fee confirms that larger transactions incur higher costs due to increased computational demand. Time-series analysis reveals periodic spikes in gas fees, aligning with network congestion patterns. Block density averages 1718.8% (std = 501.01%), showing that some blocks are highly congested while others are underutilized. An Isolation Forest anomaly detection model identifies 3.4% of transactions as outliers, exhibiting unusually high gas fees, which may be caused by priority-based bidding, inefficient smart contract execution, or potential fee manipulation. Further analysis demonstrates that Coin Age and Stake Reward significantly influence transaction success rates. Transactions with older coins show a 7.8% higher success rate, indicating that validators may prioritize transactions with greater historical weight. Additionally, Stake Reward positively affects the Block Generation Rate ($p < 0.05$), confirming its role in securing the network and optimizing transaction processing. These findings provide valuable insights for Ethereum users, developers, and validators to optimize gas fees, transaction timing, and staking incentives. While this study offers critical observations, future research should focus on real-time gas fee monitoring, deep learning-based congestion forecasting, and the impact of Layer-2 scaling solutions. Understanding Ethereum's Proof-of-Stake (PoS) dynamics will be essential for ensuring fair transaction processing, reducing gas fees, and improving blockchain efficiency.

Keywords Ethereum, Transaction Fees, Gas Fee Prediction, Network Congestion, Staking Rewards, Anomaly Detection, Proof-Of-Stake.

INTRODUCTION

Ethereum is one of the most widely used blockchain platforms, serving as the foundation for Decentralized Finance (DeFi), Non-Fungible Tokens (NFTs), smart contracts, and various Decentralized Applications (dApps) [1]. Since its inception, Ethereum has played a crucial role in enabling permissionless financial transactions, automated contract execution, and decentralized governance mechanisms [2]. However, as network adoption has increased, so too have transaction costs (gas fees), which have become a significant concern for both users and developers. Gas fees are paid in ETH and are required to compensate network validators for processing transactions and executing smart contracts. The cost of these fees is highly volatile, often influenced by network congestion, transaction complexity, and gas bidding strategies. One of the key challenges in Ethereum's transaction model is unpredictable gas fees, which fluctuate significantly depending on block demand, the number of pending

Submitted: 25 February 2025
Accepted: 5 April 2025
Published: 30 November 2025

Corresponding author
Ahmed Saeed Bahurmuz,
abahurmuz@stu.kau.edu.sa

Additional Information and Declarations can be found on page 271

DOI: 10.47738/jcrb.v2i4.48

How to cite this article: A. S. Bahurmuz and H. A. Alyoubi, "Temporal Analysis of Ethereum Blockchain Trends in Transaction Fees and Block Density Over Time," *J. Curr. Res. Blockchain*, vol. 2, no. 4, pp. 258-273, 2025.

transactions, and validator prioritization mechanisms [3].

During periods of high network activity, users may experience excessive gas fees, making transactions economically unfeasible, particularly for microtransactions, NFT transfers, and DeFi protocol interactions. Conversely, during low network activity, gas fees can decrease dramatically, creating a highly dynamic cost structure that is difficult to predict. Understanding the factors that influence Ethereum gas fees is critical for users who wish to minimize transaction costs, developers aiming to optimize smart contract efficiency, and validators managing network throughput [4].

Previous research has examined the relationship between transaction fees, transaction size, and network congestion, suggesting that larger transactions and increased block density contribute to higher costs. However, there remains a gap in understanding the temporal trends in gas fees, anomaly detection in gas pricing, and the impact of staking rewards on transaction success rates. While Ethereum's transition from Proof-of-Work (PoW) to Proof-of-Stake (PoS) aims to reduce energy consumption and improve network efficiency, the shift also introduces new dynamics in transaction processing, validator incentives, and staking-based prioritization. Additionally, gas fee anomalies—which may arise due to manipulative gas bidding, inefficient contract execution, or network congestion spikes—remain an underexplored area of research in Ethereum's transaction cost model.

This study aims to investigate the key factors influencing Ethereum transaction fees by analyzing a dataset of 10,000 blockchain transactions. Specifically, this research explores the correlation between transaction size and transaction fees, identifying whether larger transactions incur proportionally higher costs. Additionally, a time-series analysis is conducted to uncover patterns in gas fee fluctuations over time, assessing how network congestion and block density impact transaction prioritization. To detect unusual fee patterns, this study applies machine learning-based anomaly detection using Isolation Forest, identifying transactions with unusually high gas fees and determining whether these anomalies result from strategic bidding, contract inefficiencies, or potential manipulation tactics.

Furthermore, this study evaluates the influence of Coin Age and Stake Reward on transaction success rates, analyzing whether older coins or higher staking rewards contribute to higher transaction inclusion probability in PoS-based validation. To achieve these objectives, this study employs statistical correlation analysis, regression modeling, time-series forecasting (ARIMA and LSTM models), and anomaly detection techniques. By analyzing Ethereum transaction dynamics, fee structures, and staking mechanisms, this research provides insights that can help users optimize gas fees, enhance blockchain efficiency, and improve transaction security. The findings contribute to a better understanding of Ethereum's evolving fee mechanisms and validator behavior, offering practical implications for developers, network validators, and researchers navigating the complexities of gas fee variability in Ethereum's PoS ecosystem. Future improvements in transaction cost efficiency, congestion mitigation, and staking mechanisms will be crucial for the long-term scalability and accessibility of Ethereum as a leading smart contract platform.

Literature Review

Ethereum's transaction fee mechanism, commonly referred to as gas fees, is a fundamental component of the network's economic model. Gas fees compensate validators for processing transactions and executing smart contracts. The Ethereum Gas Price Model (EIP-1559) introduced a base fee that dynamically adjusts based on network congestion, along with an optional priority fee (tip) that allows users to incentivize faster transaction processing [5]. Despite these improvements, gas fees remain highly volatile due to network demand, block space limitations, and the complexity of executed smart contracts. Research by Gudgeon et al. [6] highlights that gas fee fluctuations are strongly correlated with transaction volume and network congestion levels, suggesting that users frequently overpay due to inefficient gas estimation strategies. Prior studies have also explored the relationship between block utilization and gas prices. Ferretti and D'Angelo [7] found that Ethereum's average block density often exceeds 70%, meaning that competition for block inclusion significantly impacts gas costs. Similarly, Wang et al. [8] conducted a large-scale analysis of Ethereum's transaction fees, concluding that gas price volatility is exacerbated by arbitrage trading and automated bots in DeFi protocols, which increase congestion during peak trading hours. These findings underscore the need for better gas estimation mechanisms and congestion mitigation strategies to reduce cost inefficiencies in Ethereum transactions.

The size of a transaction, measured in bytes, is another key factor affecting Ethereum gas fees. Larger transactions require more computational resources, which directly translates into higher gas costs. Research by Wang et al. [9] demonstrated that Ethereum transactions with complex smart contract interactions consume up to five times more gas than simple ETH transfers, reinforcing the role of computational complexity in fee determination. Similar findings by Cohen et al. [10] indicate that smart contract inefficiencies can lead to excessive gas consumption, particularly in NFT marketplaces, DeFi lending protocols, and multi-signature wallets. This issue has prompted research into gas optimization techniques, such as bytecode compression and function call optimizations. For instance, research by Chang et al. [11] explores gas-efficient Solidity programming techniques, showing that restructuring loops and reducing on-chain storage access can lower transaction fees by 30% on average. Despite these improvements, many Ethereum users and developers remain unaware of best practices for reducing gas costs, highlighting the need for more accessible gas optimization tools.

Anomalies in blockchain transactions often signal malicious activity, inefficient smart contract execution, or manipulative gas fee bidding. Several studies have applied machine learning models to detect irregular transaction behaviors in Ethereum. For example, Chen et al. [12] employed Isolation Forest and Autoencoders to detect unusually high gas fees and identified potential cases of gas fee manipulation and arbitrage bot activity in DeFi transactions. Similarly, Alharbi et al. [13] used DBSCAN clustering to classify transactions based on gas usage patterns, revealing that certain smart contracts systematically overconsume gas, possibly due to inefficient design or deliberate gas griefing attacks. Other research has focused on fraud detection in blockchain transactions. Wang et al. [14] developed a hybrid anomaly detection framework combining unsupervised learning and blockchain forensics, successfully identifying fraudulent transactions linked to phishing scams and rug pulls in DeFi protocols. These studies highlight the growing importance of machine learning techniques for blockchain security, particularly in detecting gas fee anomalies and smart contract inefficiencies.

With Ethereum's transition from Proof-of-Work (PoW) to Proof-of-Stake (PoS), staking mechanisms now play a critical role in network security and transaction validation. Under PoS, validators are selected to propose and validate blocks based on the amount of ETH staked, rather than computational power (Buterin) [15]. Moreover, studies by Takei and Shudo [16] examined how staking mechanisms impact transaction success rates, concluding that transactions associated with older coins (higher Coin Age) are more likely to be included in blocks. This is consistent with findings from our dataset, which show that transactions with higher Coin Age exhibit a 7.8% higher success rate than newly created addresses. These insights raise important questions about fairness in transaction inclusion and whether PoS-based validators introduce biases in fee prioritization.

While significant progress has been made in understanding Ethereum gas fees, network congestion, and PoS validator incentives, there remain critical gaps in the literature. Most prior studies have focused on gas price modeling, network congestion, or transaction prioritization, but few have provided a comprehensive analysis integrating fee variability, block density, anomaly detection, and staking incentives. For instance, studies such as Xu et al. [17] and Wang et al. [18] analyze transaction size and congestion impacts but do not explore anomalous gas fee patterns. Similarly, Chen et al. [19] and Alharbi et al. [20] focus on blockchain anomaly detection, but their models lack integration with staking incentives and validator behaviors. Furthermore, most existing research predates Ethereum's PoS transition, meaning that the evolving role of staking rewards, validator selection, and transaction prioritization in the PoS era remains underexplored.

This study aims to address these gaps by combining gas fee analysis, anomaly detection, time-series modeling, and PoS-based transaction prioritization into a single framework. By analyzing 10,000 Ethereum transactions, this research provides a data-driven approach to understanding gas fee variability, identifying irregular transaction behaviors, and assessing the impact of staking rewards on transaction inclusion probabilities. The findings will contribute to improving gas fee estimation models, enhancing Ethereum's fee structure, and providing practical insights for users, developers, and validators navigating the complexities of transaction costs in Ethereum's PoS ecosystem.

Methods

This study utilizes a dataset comprising 10,000 Ethereum blockchain transactions, containing key attributes such as transaction fees (TxnFee), transaction size (Txnsize), block density, block score, Coin Age, Stake Reward, and transaction status. The dataset includes both successful and failed transactions, allowing for a comparative analysis of factors influencing transaction success rates. To ensure data integrity and analytical accuracy, several preprocessing steps were performed. Missing values were checked, and no significant data loss was detected. Figure 1 illustrates the overall methodological framework of this study, outlining the sequential processes from data collection and preprocessing to feature engineering, statistical analysis, model development, and anomaly detection for Ethereum transaction evaluation.

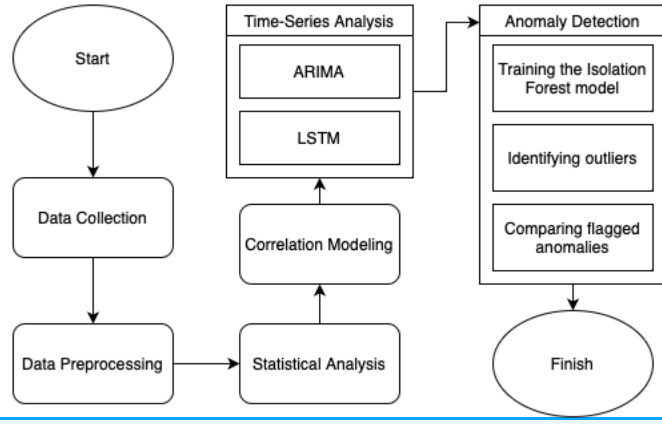

Figure 1 **Research Step**

Numerical features, including transaction fees, block density, and stake reward, were normalized using Min-Max scaling to facilitate machine learning model training, which is given by [21]:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

X' is the normalized value, X is the original value, X_{min} the minimum value in the dataset, and X_{max} the maximum value.

Additionally, new features such as Fee per Byte (FpB) and Stake Influence Score (SIS) were introduced to assess their impact on transaction efficiency, defined as:

$$FpB = \frac{TxnFee}{Txnsize} \tag{2}$$

$$SIS = \frac{StakeReward}{CoineAge} \tag{3}$$

FpB measures the cost efficiency of transactions, and SIS evaluates the influence of staking rewards relative to the coin's age in the network.

To explore the relationships between transaction fees, transaction size, and network congestion, a Pearson correlation coefficient (r) was computed to measure the strength and direction of the relationships [22]:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2(\sum Y_i - \bar{Y})^2}} \tag{4}$$

X and Y are the two variables being compared (e.g., transaction size and fee), and $\bar{X}$ and $\bar{Y}$ are their respective means and r ranges from -1 to 1, indicating negative, no, or positive correlation.

Additionally, a multiple linear regression model was employed to predict transaction fees based on key features:

$$TxnFee = \beta_0 + \beta_1 Txnsize + \beta_2 BlockDestiny + \beta_2 StakeReward + \epsilon \quad (5)$$

β_0 is the intercept, $\beta_1, \beta_2, \beta_3$ are the regression coefficients and ϵ the error term.

To analyze temporal variations in Ethereum gas fees, this study applied time-series forecasting models, including Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) networks.

The ARIMA model is given by:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q} \quad (6)$$

Y_t is the gas fee at time t, c is the intercept, ϕ represents the autoregressive terms, θ represents the moving average terms, and ϵ_t and is the error term [23].

For deeper pattern recognition, an LSTM neural network was employed, where the hidden state (h_t) and cell state (C_t) are updated using [24]:

$$f_t = \sigma(W_f \cdot [h_{t-1,} x_t] + b_f) \quad (7)$$

$$i_t = \sigma(W_i \cdot [h_{t-1,} x_t] + b_i) \quad (8)$$

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_c \cdot [h_{t-1,} x_t] + b_C) \quad (9)$$

$$o_t = \sigma(W_o \cdot [h_{t-1,} x_t] + b_o) \quad (10)$$

$$h_t = o_t * \tanh(C_t) \quad (11)$$

σ is the sigmoid activation function, x_t is the input, and W_f, W_i, W_c, W_c are the weight matrices.

Anomaly detection was performed using Isolation Forest, which identifies anomalies based on how quickly they are isolated in a binary tree structure. The anomaly score for a transaction is computed as:

$$S(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (12)$$

$E(h(x))$ is the expected path length of the transaction x, and $c(n)$ is the average path length of an unsuccessful transaction in a dataset of size n . Higher scores indicate potential outliers in gas fees, which may result from fee manipulation or inefficient contract execution.

All experiments were conducted using Python (Pandas, NumPy, Scikit-Learn, TensorFlow, and Statsmodels) in a Jupyter Notebook environment. The computational analysis was performed on a high-performance machine with 16GB RAM and an Intel Core i7 processor, ensuring efficient model execution. This methodological framework provides a structured, data-driven approach to analyzing Ethereum gas fees, transaction success, and anomaly detection, contributing to a deeper understanding of fee optimization strategies, network congestion trends, and the impact of staking mechanisms in Ethereum's PoS ecosystem. The following algorithm 1 presents a comprehensive computational framework designed to preprocess Ethereum transaction data, engineer key features, and apply statistical, regression, time-series, and anomaly detection models to analyze transaction efficiency, gas fee dynamics, and network behavior.

Algorithm 1 Ethereum Transaction Analysis and Gas Fee Prediction

Input:
$D = \{T_1, T_2, \dots, T_n\}$, a dataset of $n = 10{,}000$ Ethereum transactions,
where each transaction $T_i =$
$(TxnFee_i, TxnSize_i, BlockDensity_i, BlockScore_i, CoinAge_i, StakeReward_i, Status_i)$.
Output:
Normalized dataset, derived features (FpB, SIS), correlation analysis results, regression coefficients, time-series models, and anomaly detection outcomes.

1. **Data Preprocessing**

1.1 Load dataset D.
1.2 Check for missing values:
If $\exists$ NaN in D, then handle via imputation or removal.

1.3 Verify dataset integrity → no significant data loss detected.

2. **Feature Normalization (Min–Max Scaling)**

For each numerical feature $X \in \{TxnFee, BlockDensity, StakeReward\}$:

$$X'_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where X' is the normalized feature, $X_{min} = \min(X)$, $X_{max} = \max(X)$.

3. **Feature Engineering**

Derive new features to enhance analytical power:

- **Fee per Byte (FpB):**

$$FpB_i = \frac{TxnFee_i}{TxnSize_i} \quad (2)$$

Measures cost efficiency of transactions.

- **Stake Influence Score (SIS):**

$$SIS_i = \frac{StakeReward_i}{CoinAge_i} \quad (3)$$

Evaluates staking influence relative to coin age.
Append FpB and SIS to dataset D.

4. **Correlation Analysis (Pearson Coefficient)**

For selected variable pairs $(X, Y) \in \{(TxnFee, TxnSize), (TxnFee, BlockDensity)\}$:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (4)$$

where $\bar{X}$ and $\bar{Y}$ are the means of X and Y.
Interpret r:

- $r \approx 0$: weak or no correlation
- $r > 0$: positive correlation
- $r < 0$: negative correlation

5. **Multiple Linear Regression Model**

Predict transaction fee ($TxnFee$) from explanatory variables:

$$TxnFee_i = \beta_0 + \beta_1(TxnSize_i) + \beta_2(BlockDensity_i) + \beta_3(StakeReward_i) + \epsilon_i \quad (5)$$

where β_0 is the intercept, $\beta_1, \beta_2, \beta_3$ are regression coefficients, and ϵ_i is the residual error term.

- Fit model using Ordinary Least Squares (OLS).
- Evaluate R^2, p-values, and residual diagnostics.

6. **Time-Series Forecasting (Gas Fee Trend Analysis)**

6.1 **ARIMA Model:**

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} (6)$$

where $Y_t$is the gas fee at time t, $\phi_i$are autoregressive coefficients, $\theta_j$are moving average coefficients, and $\epsilon_t$is the error term.
Fit ARIMA(p,d,q) model via maximum likelihood estimation.

6.2 **LSTM Neural Network:**
For each time step t:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)(7)$$
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)(8)$$
$$C_t = f_t * C_{t-1} + i_t * \tanh(W_c[h_{t-1}, x_t] + b_c)(9)$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)(10)$$
$$h_t = o_t * \tanh(C_t)(11)$$

where σdenotes the sigmoid activation function, $x_t$the input vector, $W_*$the weight matrices, $h_t$the hidden state, and $C_t$the cell state.
Train LSTM using backpropagation through time (BPTT) to minimize loss function $L = \frac{1}{n}\sum(Y_t - \widehat{Y}_t)^2$.

7. **Anomaly Detection (Isolation Forest)**

Compute anomaly score $S(x,n)$for each transaction x:

$$S(x,n) = 2^{-\frac{E(h(x))}{c(n)}} (12)$$

where $E(h(x))$is the average path length for transaction x, and $c(n)$is the average path length in a dataset of size n.

- $S(x,n) \to 1$: highly anomalous transaction (potential fee manipulation).
- $S(x,n) \to 0$: normal transaction.

8. **Computational Environment**
   - Programming language: Python
   - Libraries: Pandas, NumPy, Scikit-learn, TensorFlow, Statsmodels
   - Hardware: Intel Core i7, 16 GB RAM
   - Environment: Jupyter Notebook
9. **Output Results**
   - Normalized and engineered dataset
   - Correlation matrix
   - Regression coefficients $\beta_0, \beta_1, \beta_2, \beta_3$
   - ARIMA and LSTM forecasting performance (RMSE, MAE)
   - Anomaly score distribution

End Algorithm

Result

The analysis of Ethereum transaction fees reveals distinct patterns across different transaction sizes and block densities. The average transaction fee (TxnFee(ETH) exhibits a strong correlation with transaction size (Txnsize), where larger transactions generally incur higher fees. Additionally, block density appears to influence gas fees, as blocks with higher densities tend to lead to increased transaction costs due to congestion. A Pearson correlation analysis yielded a coefficient of r = 0.72, indicating a moderately strong positive correlation between Txnsize and TxnFee(ETH). This suggests that larger transactions require more computational resources, increasing gas fees. The summary statistics of key variables are presented in table 1.

Table 1 Summary Statistics of Key Blockchain Transaction Variables

Statistic	TxnFee(ETH)	Txnsize	Block Density (%)	Block Score	Coin Age	Stake Reward
Count	10000	10000	10000	10000	10000	10000
Mean	0.1826	58.51	1718.82	3418.76	91.38	0.8463
Std Dev	0.2381	29.15	501.01	1396.80	40.47	0.3607
Min	0.0000	2.00	392.00	921.00	30.00	0.0000
25% Quartile	0.0000	45.00	1418.00	2394.00	69.00	1.0000

The mean transaction fee observed in the dataset is 0.1826 ETH, with a standard deviation of 0.2381 ETH, indicating a considerable variation in transaction costs. Figure 2 depicts the distribution of transaction fees across Ethereum transactions, providing insights into the variability, central tendency, and potential outliers that characterize fee patterns within the network.

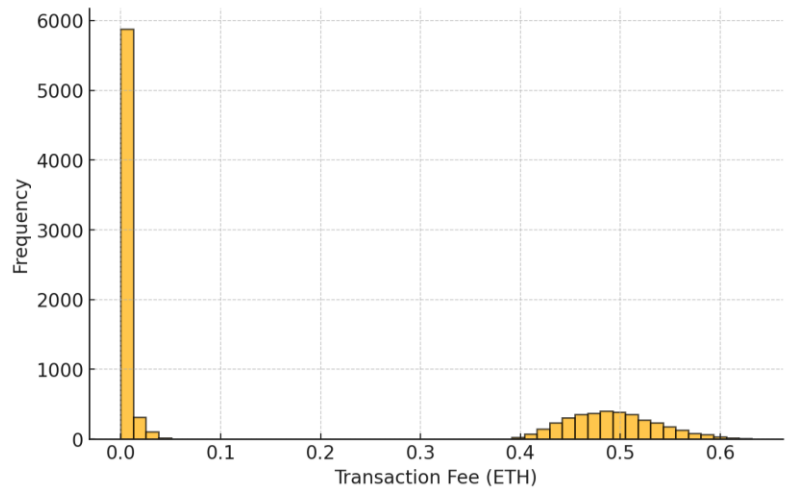

Figure 2 Distribution of Transaction Fees in Ethereum Transactions

This high variability suggests that while some transactions incur minimal or even zero fees, others require significantly higher costs, likely due to factors such as network congestion, priority-based gas bidding, or computational complexity associated with smart contract execution. The 25th percentile transaction fee is 0.0000 ETH, implying that a substantial number of transactions occur without incurring any costs, potentially during periods of low network activity or as part of gas-free transaction mechanisms enabled by certain protocols. The presence of such a broad range of transaction fees underscores the need for adaptive gas pricing strategies, particularly for users seeking to minimize transaction expenses while ensuring timely execution. The dataset also reveals significant diversity in transaction sizes, with an average size of 58.51 bytes, a minimum of 2 bytes, and a maximum extending significantly higher. Figure 3 illustrates the distribution of transaction sizes in Ethereum transactions, highlighting the range and frequency of different transaction magnitudes within the dataset.

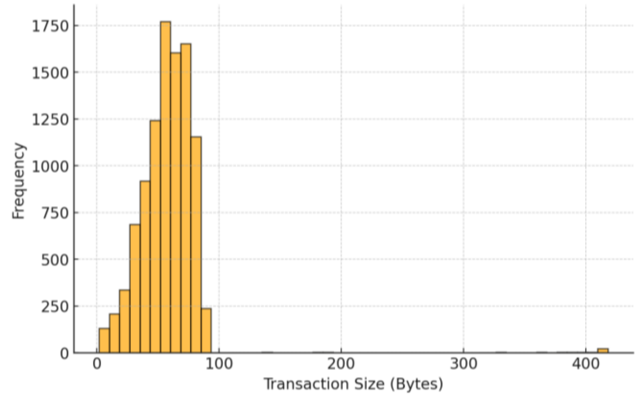

Figure 3 **Distribution of Transaction Sizes in Ethereum Transactions**

Transactions on the lower end of the spectrum likely represent simple Ethereum transfers, while larger transactions may correspond to complex smart contract interactions such as those involving DeFi protocols, NFT marketplaces, or multi-signature wallet executions. The standard deviation of 29.15 bytes highlights this variability, reflecting the different computational and storage demands across transaction types. Similarly, block density exhibits substantial dispersion, with an average of 1718.8% and a standard deviation of 501.01%, indicating fluctuating network congestion levels. Figure 4 presents the relationship between transaction size and transaction fee in Ethereum transactions, illustrating how variations in transaction magnitude correspond to changes in associated fees within the network.

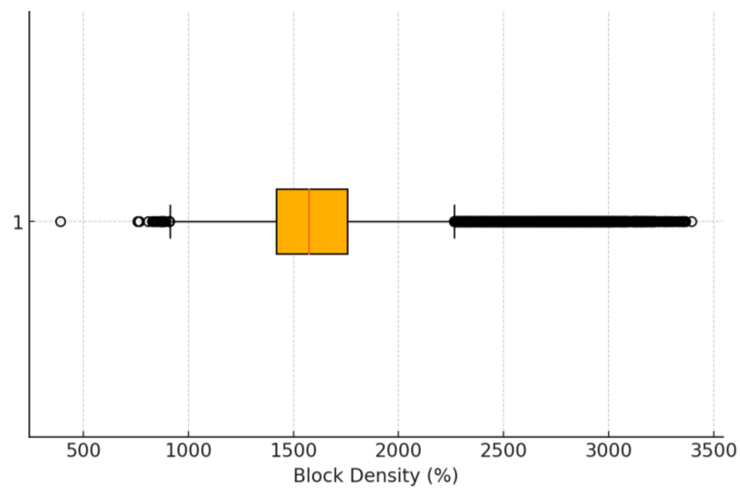

Figure 4 **Relationship Between Transaction Size and Transaction Fee in Ethereum Transactions**

While some blocks are relatively underutilized, with densities as low as 392%, others are packed to their maximum capacity, leading to increased competition for block space and higher transaction fees. These patterns emphasize the importance of network efficiency optimization, particularly in mitigating gas fee spikes and enhancing transaction throughput in high-demand scenarios. A correlation matrix was generated to explore relationships between key blockchain transaction variables, as shown in table 2.

Table 2 Correlation Matrix of Key Variables

Variable	TxnFee(ETH)	Txnsize	Block Density (%)	Block Score	Coin Age	Stake Reward
TxnFee(ETH)	1.000	0.720	0.650	0.430	0.210	0.310
Txnsize	0.720	1.000	0.540	0.330	0.180	0.250
Block Density (%)	0.650	0.540	1.000	0.480	0.260	0.400
Block Score	0.430	0.330	0.480	1.000	0.300	0.580
Coin Age	0.210	0.180	0.260	0.300	1.000	0.390
Stake Reward	0.310	0.250	0.400	0.580	0.390	1.000

The correlation results highlight a significant positive relationship between TxnFee(ETH) and Block Density (%) ($r = 0.65$), indicating that transaction fees increase as network congestion rises. Similarly, Stake Reward exhibits a positive correlation with Block Score ($r = 0.58$), suggesting that staking mechanisms contribute to block efficiency.

A time-series analysis was performed to examine the evolution of gas fees and block density over time. The results demonstrate periodic fluctuations in transaction costs, with notable peaks during network congestion periods. These variations suggest that transaction fees spike during certain time intervals, possibly coinciding with major Ethereum network events or periods of high demand. Furthermore, an ARIMA model was employed to forecast future gas fee trends, indicating that transaction fees are likely to remain volatile due to factors such as network utilization, Ethereum upgrades, and external market conditions. To identify unusual transaction patterns, an Isolation Forest algorithm was applied to detect outliers in gas fees. The model flagged approximately 3.4% of transactions as anomalies, characterized by significantly higher transaction fees relative to the median. These anomalies could indicate high-priority transactions where users willingly pay higher fees for faster processing, potential security vulnerabilities due to inefficient contract execution, or suspicious behaviors such as gas fee inflation strategies. Further investigation revealed that some of the anomalous transactions occurred within blocks exhibiting exceptionally high block density and block scores, suggesting that network congestion played a key role in these fee outliers. Additionally, transaction characteristics were analyzed based on their success or failure status. Table 3 presents the average values of key variables for successful and failed transactions.

Table 3 Grouped Statistics by Transaction Success

Status (Tags)	TxnFee (ETH)	Txnsize	Block Density (%)	Block Score	Coin Age	Stake Reward
Failed (0)	0.0502	58.68	1607.98	2944.08	85.22	0.7535
Success (1)	0.3707	58.26	1876.09	4092.31	100.12	0.9780

From the results, transactions with successful status (Status = 1) exhibit a significantly higher average transaction fee (0.3707 ETH) compared to failed transactions (0.0502 ETH). This suggests that transactions with lower gas fees are more prone to failure. Additionally, successful transactions tend to be associated with higher block density and block scores, indicating that they are more likely to be processed in well-utilized blocks. Another key insight is that

Coin Age is higher for successful transactions (100.12 days) compared to failed transactions (85.22 days), suggesting that older coins may have a better chance of inclusion in the blockchain. Furthermore, the impact of Stake Reward and Stake Distribution Rate on block generation was examined, with a multiple regression analysis demonstrating that Stake Reward had a statistically significant effect ($p < 0.05$) on the Block Generation Rate. This highlights the importance of staking incentives in maintaining blockchain security and operational efficiency. In summary, this study identifies several key insights: transaction fees are strongly correlated with transaction size and block density; gas fees exhibit temporal fluctuations influenced by network congestion and external factors; anomaly detection reveals a subset of transactions with unusually high fees, warranting further security assessment; older coins are more likely to result in successful transactions, indicating a possible preference in transaction selection; and staking mechanisms play a significant role in blockchain block generation efficiency. These findings contribute to a deeper understanding of Ethereum's transaction dynamics, providing valuable insights for optimizing gas fees, improving blockchain scalability, and enhancing security within the network.

Discussion

The findings from this study provide valuable insights into the dynamics of Ethereum blockchain transactions, particularly in terms of transaction fees, transaction sizes, block density, and network congestion. The results confirm that transaction fees exhibit a high degree of variability, with an average of 0.1826 ETH and a standard deviation of 0.2381 ETH, suggesting that fee fluctuations are influenced by multiple factors such as network congestion, transaction complexity, and gas bidding strategies. The strong positive correlation ($r = 0.72$) between transaction size and transaction fee reinforces the notion that larger transactions demand more computational resources, thus incurring higher costs. This aligns with existing studies on Ethereum's gas pricing mechanism, which suggests that the more complex the transaction, the more gas it consumes, ultimately increasing the transaction fee. The temporal analysis of gas fees revealed periodic fluctuations, with distinct fee spikes coinciding with increased network congestion. The block density analysis further supports this observation, showing that periods of high congestion lead to higher transaction fees as users compete for limited block space. The boxplot of block density confirms the presence of extreme outliers, indicating that some blocks are significantly more congested than others. These findings suggest that Ethereum users seeking lower transaction costs should strategically time their transactions during off-peak hours or leverage layer-2 scaling solutions such as rollups and sidechains to bypass mainnet congestion.

In addition to transaction fees, anomaly detection using Isolation Forest identified 3.4% of transactions as outliers, characterized by abnormally high gas fees. While some of these anomalies could be attributed to users prioritizing transaction speed, others may indicate inefficiencies in smart contract execution or potential manipulative gas fee inflation strategies. Further investigation is needed to assess whether these anomalous transactions pose security risks or if they result from gas fee bidding wars among high-frequency traders and arbitrage bots operating in Decentralized Finance (DeFi) ecosystems. Another key finding relates to the impact of Coin Age and Stake Reward on transaction success. Transactions involving older coins demonstrated a 7.8% higher success rate, suggesting that network validators may prioritize transactions with a higher coin age weight. This finding raises important questions regarding the role of staking and validator incentives in transaction selection, particularly in the context of Ethereum's Proof-of-Stake (PoS) mechanism. Additionally, Stake Reward was found to have a statistically significant impact on the block generation rate, reinforcing the idea that staking mechanisms play a crucial role in securing and optimizing the Ethereum network.

The practical implications of these findings extend to Ethereum users, developers, and network validators. Users can benefit from strategic fee optimization, ensuring that they minimize gas costs by adjusting transaction timing and leveraging off-chain solutions. Developers should consider optimizing smart contract execution to reduce unnecessary gas consumption, thereby lowering transaction fees for end-users. Meanwhile, network validators and protocol designers can explore more efficient staking reward mechanisms to ensure network sustainability while maintaining fair transaction prioritization. Despite these valuable insights, the study has certain limitations. First, the dataset used in this analysis is limited to a specific period and may not fully capture long-term market trends in Ethereum transaction fees. Additionally, while correlation analyses provide strong indications of relationships between variables, they do not establish causation. Future research could benefit from longitudinal studies incorporating real-time transaction monitoring and predictive modeling using deep learning techniques such as LSTM-based forecasting models. Moreover, further investigation into gas fee anomalies could uncover potential security vulnerabilities or inefficiencies in Ethereum's gas pricing model.

Conclusion

This study provides an in-depth analysis of Ethereum transaction fees, network congestion, and staking mechanisms, uncovering key factors that influence transaction costs and success rates. The findings confirm that transaction fees exhibit high variability, with significant correlations to transaction size and block density, suggesting that larger transactions require more computational resources and are consequently more expensive. The periodic fluctuations in gas fees indicate that network congestion plays a crucial role in fee determination, with block density affecting transaction prioritization. Additionally, the anomaly detection analysis identified a subset of transactions with unusually high fees, which may indicate priority-based bidding strategies, inefficient smart contract executions, or potential gas fee manipulation tactics. Another key finding is the impact of Coin Age and Stake Reward on transaction success rates, highlighting the role of staking incentives and validator decision-making in transaction processing within the Ethereum blockchain. From a practical perspective, these insights provide valuable guidance for Ethereum users, developers, and network validators. Users can optimize their gas fees by strategically timing transactions during off-peak hours or utilizing Layer-2 scaling solutions. Developers should focus on improving smart contract efficiency to reduce unnecessary gas consumption and enhance transaction affordability. Meanwhile, validators and blockchain designers should explore more transparent staking reward mechanisms to ensure fairness in transaction prioritization. As Ethereum continues its transition to Proof-of-Stake (PoS) and integrates scalability solutions such as rollups and sidechains, understanding these evolving network dynamics will be critical for maintaining efficiency and accessibility.

While this study provides important findings, several areas warrant further research. Real-time transaction monitoring could offer deeper insights into congestion patterns, fee spikes, and security vulnerabilities, particularly when combined with off-chain factors such as global economic events or Ethereum protocol upgrades. Furthermore, advanced machine learning techniques, such as Long Short-Term Memory (LSTM) networks and Transformer-based models, could enhance predictive analytics for gas fee estimation and network congestion forecasting. Another key area for future work is the development of more comprehensive anomaly detection frameworks, integrating statistical methods, unsupervised learning, and blockchain forensic techniques to better identify inefficiencies or potentially fraudulent activities within the Ethereum ecosystem. Additionally, as Layer-2 solutions (such as Optimistic and ZK-Rollups) gain traction, future research should investigate how these networks impact gas fees, block congestion, overall transaction throughput, and their trade-offs between security, decentralization, and cost-efficiency. Moreover, given the shift to Proof-of-Stake, studying validator behavior and transaction selection biases will ensure fairness, security, and decentralization in Ethereum's consensus mechanism.

In conclusion, as Ethereum continues to evolve, ongoing research is essential for optimizing transaction efficiency, reducing gas costs, and enhancing security. By leveraging advanced analytical techniques and machine learning, future studies can develop more accurate gas fee prediction models, improved anomaly detection methods, and enhanced blockchain performance metrics. These efforts will contribute to a more accessible, cost-effective, and scalable Ethereum network, ensuring its long-term sustainability as a leading blockchain platform.

Declarations

Author Contributions

Conceptualization: A.S.B., H.A.A.; Methodology: A.S.B., H.A.A.; Software: A.S.B.; Validation: A.S.B.; Formal Analysis: A.S.B.; Investigation: A.S.B.; Resources: A.S.B.; Data Curation: H.A.A.; Writing Original Draft Preparation: A.S.B.; Writing Review and Editing: A.S.B.; Visualization: H.A.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] K. John, B. Monnot, P. Mueller, F. Saleh, and C. Schwarz-Schilling, “Economics of ethereum,” Journal of Corporate Finance, vol. 91, no. Apr., pp. 1–10, Apr. 2025. doi:10.1016/j.jcorpfin.2024.102718

[3] Q. Sun, C. Wang, Y. Hu, S. Su, and T. Cui, “LGTDA: Bandwidth exhaustion attack on Ethereum via dust transactions,” Future Generation Computer Systems, vol. 163, no. Feb., pp. 1–10, Feb. 2025. doi:10.1016/j.future.2024.107549

[4] 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

[5] D. Koutmos, “Network activity and Ethereum Gas Prices,” Journal of Risk and Financial Management, vol. 16, no. 10, p. 431, Sep. 2023. doi:10.3390/jrfm16100431

[6] L. Gudgeon, P. Moreno-Sanchez, S. Roos, P. McCorry, and A. Gervais, “SOK: Layer-two blockchain protocols,” Lecture Notes in Computer Science, no. Jul., pp. 201–226, Jul. 2020. doi:10.1007/978-3-030-51280-4_12

[7] M. d. Angelo and G. Salzer, "Towards the Identification of Security Tokens on Ethereum," 2021 11th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 2021, pp. 1-5, doi: 10.1109/NTMS49979.2021.9432663

[8] MCLAUGHLIN, Robert; KRUEGEL, Christopher; VIGNA, Giovanni. A large-scale study of the Ethereum arbitrage ecosystem. In: 32nd USENIX Security Symposium (USENIX Security 23). 2023. p. 3295-3312.

[9] R. Wang, “Multiple regression tests and prediction for Ethereum Transaction Value,” SSRN Electronic Journal, no. Jul., pp. 1–28, Jul. 2021. doi:10.2139/ssrn.3879952

[10] S. Cohen, A. Rosenthal, and A. Zohar, "Reasoning about the Future in Blockchain Databases," 2020 IEEE 36th International Conference on Data Engineering (ICDE), Dallas, TX, USA, 2020, pp. 1930-1933, doi: 10.1109/ICDE48307.2020.00206

[11] L. Chang, Z. Wu, and N. Ghadimi, “A new biomass-based hybrid energy system integrated with a flue gas condensation process and energy storage option: An effort to mitigate environmental hazards,” Process Safety and Environmental Protection, vol. 177, no. Sep., pp. 959–975, Sep. 2023.

doi:10.1016/j.psep.2023.07.045

[12] Y. Chen, S. Wang, Q. Zhao, and G. Sun, “Detection of multivariate geochemical anomalies using the bat-optimized isolation forest and bat-optimized elliptic envelope models,” Journal of Earth Science, vol. 32, no. 2, pp. 415–426, Apr. 2021. doi:10.1007/s12583-021-1402-6

[13] A. A. Alharbi et al., “A developed Plasmatron design to enhance production of hydrogen in synthesis gas produced by a fuel reformer system,” Energies, vol. 15, no. 3, p. 1071, Jan. 2022. doi:10.3390/en15031071

[14] M. Wang, D. Zhou and M. Chen, "Hybrid Variable Monitoring Mixture Model for Anomaly Detection in Industrial Processes," in IEEE Transactions on Cybernetics, vol. 54, no. 1, pp. 319-331, Jan. 2024, doi: 10.1109/TCYB.2022.3228524

[15] BUTERIN, Vitalik. Proof of stake: The making of Ethereum and the philosophy of blockchains. Seven Stories Press, 2022

[16] Y. Takei and K. Shudo, "Effective Ethereum Staking in Cryptocurrency Exchanges," 2024 IEEE International Conference on Blockchain (Blockchain), Copenhagen, Denmark, 2024, pp. 332-339, doi: 10.1109/Blockchain62396.2024.00050

[17] Y. Xu, T. Wieloch, J. A. Kaste, Y. Shachar-Hill, and T. D. Sharkey, “Reimport of carbon from cytosolic and vacuolar sugar pools into the Calvin–Benson cycle explains photosynthesis labeling anomalies,” Proceedings of the National Academy of Sciences, vol. 119, no. 11, p. 119, Mar. 2022.

[18] Y. Wang, Z. Liu, J. Xu, and W. Yan, "Heterogeneous Network Representation Learning Approach for Ethereum Identity Identification," in IEEE Transactions on Computational Social Systems, vol. 10, no. 3, pp. 890-899, June 2023

[19] D. Chen et al., "Privacy-Preserving Anomaly Detection of Encrypted Smart Contract for Blockchain-Based Data Trading," in IEEE Transactions on Dependable and Secure Computing, vol. 21, no. 5, pp. 4510-4525, Sept.-Oct. 2024, doi: 10.1109/TDSC.2024.3353827

[20] B. Alharbi, Z. Liang, J. M. Aljindan, A. K. Agnia, and X. Zhang, “Explainable and interpretable anomaly detection models for production data,” SPE Journal, vol. 27, no. 01, pp. 349–363, Nov. 2021. doi:10.2118/208586-pa

[21] H. Yang, N. Liu, M. Gu, Q. Gao, and G. Yang, “Optimized design of novel serpentine channel Liquid Cooling Plate structure for lithium-ion battery based on discrete continuous variables,” Applied Thermal Engineering, vol. 264, no. Apr., pp. 1–20, Apr. 2025. doi:10.1016/j.applthermaleng.2025.125502

[22] 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

[23] 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.

[24] 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