An Advanced Pricing Mechanism for Nonfungible Tokens (NFTs) Based on Rarity and Market Dynamics

Wenze Xiong, Yetong Wang, Wanxin Li, *Member, IEEE*, Yutong Zhang, Jie Zhang, *Member, IEEE*, and Hao Guo, *Member, IEEE*

***Abstract*—The nonfungible tokens (NFTs) are unique cryptocurrencies that exist on a blockchain and cannot be replicated. However, today's NFT market lacks a sensible pricing framework, which causes NFT price fluctuations to interfere with the investment market. The purpose of this research is to build a dynamic pricing mechanism for NFT based on NFT features, validate the improvement of factor analysis on the pricing model, and integrate the rarity-based model with market factors in an online market. We presented and implemented a prototype pricing algorithm through a designed NFT market. This implementation shows that this advanced pricing mechanism could be used in a real-world pricing scenario and could provide a reasonable price change when market factors change. Furthermore, the experimental results demonstrated that the rarity scores of the features and market factors can potentially induce price fluctuations within the range of negative 0.4 to positive 0.4.**

***Index Terms*—Blockchain, factor analysis, market factors, nonfungible token, pricing, rarity.**

I. Introduction

NONFUNGIBLE tokens, often referred to as NFTs, are blockchain-based tokens that have been proven useful in various realms, such as Metaverse [1], [2] and its improvement to the development of distributed autonomous organization (DAO) [3]. NFTs are unique and irreplaceable virtual assets that have garnered substantial attention in recent years, peaking in 2021 [4]. This research focuses on NFT pricing, exploring its significance in economics, markets, and regulations. Several crucial aspects need consideration in NFT pricing research.

Economics, influenced by supply and demand, is a primary pricing factor. Variables such as the number of participants, market competition, and potential transaction revenue contribute to pricing. Art is among the most sought-after categories of NFTs and holds significant value. NFT art can encompass various subcategories such as paintings, music albums, and video clips [2].

Metadata and smart contracts are vital to NFT for the determination of dynamic NFT (dNFT) and static NFT [5]. dNFTs allow for metadata modifications through smart contracts triggered by external events. Smart contracts manage the programming for metadata adjustments, dictating when and how changes occur. The use of dynamic NFTs is beneficial when continuous updates to NFT data are required, particularly for digital assets tied to real-world objects. Owners can manually or automatically update data associated with dynamic NFTs based on predefined conditions. In contrast, a static NFT is characterized by immutable features and data recorded on the blockchain. These NFTs are created with fixed attributes and cannot be modified. Because static NFT cannot change since they have been created, they are more secure [6].

The legal aspect of NFT research is crucial due to its susceptibility to illegal activities such as money laundering, further amplified by price volatility. Compared to traditional art, NFTs have unique qualities that attract money launderers: They offer inherent anonymity as cryptographic assets, making regulatory oversight challenging. Being digital assets, NFTs trade online with high liquidity and minimal expenses, contributing to their appeal. The absence of standardized pricing mechanisms allows price manipulation, such as wash trading. Additionally, NFT has proved to be a solution to the problems of intellectual property. For example, the NFT-based patent framework reduces major administrative disruptions and financial burdens on patent offices or their users [7].

Although some researchers examine the factors that influence the price of NFT [8], [9], [10], [11], [12], few research integrates detailed rarity and NFT market design. In this research, we further explored how rarity factors impact NFT pricing, which gives an exact pricing formula embedded in the market construction. Different collections of NFTs could have different

Manuscript received 11 January 2024; revised 20 May 2024 and 4 July 2024; accepted 15 July 2024. This work was supported in part by XJTLU Research Development Fund under Grant RDF-22-02-106, Grant 2023-2026, and Grant RDF-21-02-014; in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2021A1515110286 and Grant 2021-2024; and in part by the Basic Research Programs of Taicang under Grant TC2022JC23 and Grant 2022-2024. An earlier version of this paper was presented at the 2023 IEEE International Conference on Blockchain [DOI: 10.1109/Blockchain60715.2023.00021]. *(Corresponding authors: Wanxin Li; Jie Zhang; Hao Guo.)*

Wenze Xiong, Yetong Wang, Wanxin Li, Yutong Zhang, and Jie Zhang are with the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China (e-mail: wenze.xiong22@alumni.xjtlu.edu.cn yetong.wang19@alumni.xjtlu.edu.cn; wanxin.li@xjtlu.edu.cn; yutong.zhang22@alumni.xjtlu.edu.cn; jie.zhang01@xjtlu.edu.cn).

Hao Guo is with the School of Software, Northwestern Polytechnical University, Taicang 215400, China (e-mail: haoguo@nwpu.edu.cn).

Digital Object Identifier 10.1109/TCSS.2024.3430846

features based on their smart contracts; thus, we proposed a general and dynamic function to express the prices. Moreover, we introduce a market activity index we designed for our rarity-based model and integrate this composite model into our online market.

To address potential concerns regarding market manipulation or insider trading facilitated by advanced pricing mechanisms, our research emphasizes transparency and openness in the design and operation of pricing mechanisms. By providing participants with clear pricing rules and data sources, we can reduce the potential for manipulation. Additionally, we propose implementing an effective market monitoring and warning system to detect abnormal trading or manipulation based on the pricing model. For example, advanced technology and algorithms can analyze market data to identify potential insider trading or abnormal price behavior in a timely manner.

Focusing on the characteristics of each NFT, the main goal of this study is to develop an NFT pricing model that enhances liquidity for these assets in decentralized finance (DeFi). In summary, this study contributes to the following.

1) We proposed an advanced NFT pricing mechanism based on the hedonic model. Unlike other studies, our hedonic regression model is based on the rarity data of NFTs, which can easily be integrated into the virtual market. In addition, features in our rarity-based model are dynamic, which means that it can be compatible with different kinds of NFT collections.
2) In our experiments, we constructed a rarity-based pricing model for real-world NFT datasets and implemented factor analysis for dimensionality reduction and market activity indexing. We evaluated our model's performance on four datasets, highlighting the improvements from factor analysis. Our findings show that the "rarity + market factor" model effectively tracks price changes and highlights the crucial role of market factors in stabilizing prices and preventing extreme situations.

Here are the remaining sections. In Section II, we discuss related work that focuses on NFT factors, NFT pricing approach, and market factors and compare our works with others. For Section III, we provide details about the proposed advanced pricing model and associate the NFT pricing algorithm with the logic of implementation. Then, in Section IV, we conduct multiple experiments to construct the model with real-world datasets and explain the procedure of market factors experiments. Moreover, we design and implement a model-based NFT market. Following that, in Section V, we detailedly measure and evaluate the performance of our model and test the factor analysis' functionality on model improvement. Last, in Section VI, we conclude this article and suggest potential future research directions.

II. Related Work and Preliminaries

In this section, we explore the existing research efforts in the domain of NFTs, focusing on factors influencing their pricing and the methodologies used to analyze this complex landscape. We delve into market-side, financial, sell-side, and smart contract factors, as well as the use of the related pricing approaches to understand the relationship between NFT features and their valuations. Moreover, we explain what the market factor is in this article. Last, we make a comparative analysis of our works and other studies.

A. NFT Factor Study

Within the domain of NFTs, researchers have approached the subject from diverse perspectives, examining a range of market-side, financial, and sell-side factors that collectively shape the NFT landscape. Market-side factors delve into the unique culture surrounding NFT communities, where consensus on specific NFT collections fosters discussions within online forums. The widespread international platform, Twitter, is leveraged by NFT project owners to promote their offerings, potentially influencing NFT valuation. For instance, Kapoor [13] employed scatter plots and machine learning algorithms to establish correlations between Twitter followers and NFT asset values, akin to Luo's [14] research.

Analysis of financial factors reveals that NFTs, as a nascent investment instrument, are susceptible to volatility emanating from both traditional financial markets and Web3-related factors. Exploring the relationships between traditional financial variables and NFT metrics, such as returns [15], sales volume [16], and NFT attention [17], offers valuable insights into NFT pricing dynamics. Christopher [8], in addition, established a link between NFT value and the metaverse through the application of the hedonic model.

On the sell side, factors encompass NFT characteristics and inter-NFT influences. Rarity assumes a pivotal role, with scarcer NFTs commanding elevated prices and experiencing lower transaction frequencies [9], [18]. An intriguing aspect pertains to the impact of racial color, wherein figures with lighter hues tend to command higher prices [11]. Furthermore, certain studies [19] spotlight the cointegration of NFT submarkets, elucidating how the success of newer projects reverberates across established markets and vice versa.

In addition to the nontechnical factors discussed previously, another crucial determinant of NFT pricing is the design of the associated smart contract. This design fundamentally shapes the NFT's valuation within the transactional marketplace. Notably, the creation of NFTs is initially established according to the guidelines outlined in the ERC-721 standard [20]. This Ethereum Request for Comment document introduces essential prerequisites and solutions governing the transactional aspects of NFTs within the designated Ethereum NFT market.

Developers typically assign an initial price to the NFT, which may be subject to random selection. Over time, this price can fluctuate in response to market dynamics. Furthermore, as the NFT market evolves, an increasing number of ERC standards have been proposed to address diverse concerns. These encompass multitoken transactions, copyright matters, and other pertinent issues [21]. For example, the recent ERC-2981 standard is specifically formulated to address issues concerning royalty payment mechanisms [22]. This standard empowers developers to incorporate a royalty value based on the original price within the contract, thereby facilitating automated NFT pricing.

The Ethereum platform encourages developers to innovate and design their smart contracts and NFTs within the framework of these established ERC standards. Consequently, an array of smart contract variations has emerged to satiate the nuanced demands of the market. An illustrative example is the proposition of an auction protocol grounded in smart contract technology, designed to operate within a cross-chain environment. This innovative offering presents an alternative avenue for NFT creators seeking diverse transactional solutions [23]. However, it has come to light that the aforementioned deliberations do not definitively ascertain an unequivocal valuation for NFTs, potentially leading to discrepancies even among NFTs of similar types or identical instances. Consequently, a compelling need arises to introduce a comprehensive pricing mechanism capable of establishing a judicious and equitable valuation grounded in the intrinsic rarity attributes characteristic of each NFT.

B. NFT Pricing Approach

In this study, we assume the explicit rarity degree of each NFT in the NFT collection might impact their price. Such sort of research is rare, and we are going to evaluate to what extent our assumption is correct. To accomplish this, we have selected the hedonic model as our research method, as it is a well-established approach commonly employed to determine the pricing of NFTs. Additionally, we find it suitable for our objective of integrating the model into the smart contract. Our study involves a comparative analysis of the effectiveness of hedonic models both before and after the incorporation of factor analysis.

The concept of hedonic regression was first proposed by Sherwin Rosen in his article "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" [24]. To be explicit, this theory thinks the value of a good consists of packages of characters. The hedonic model has been applied by some NFT-related research [8], [12], [18], [25], [26], [27], [28], [29]. Among these studies of NFTs based on the hedonic model, some put factors in the metauniverse into the model [8], some research a certain NFT collection based on hedonic model [18], and others even notice the quantity associated with characteristics for NFTs [12]. However, few studies focus on the detailed characteristics of a particular type of NFT collection and thus determine their relationship with pricing models.

Machine learning is an algorithm approach based on the scientific study of statistical models, which allows computer systems to learn the internal pattern of change. After setting the initial parameter, the algorithm can identify the data relationship. There are some researches, attempting to find the internal statistical mechanism of NFT based on machine learning algorithm [30], [31], [32], [33].

Factor analysis is a statistical method used to describe variability among observed, correlated variables. Factor analysis was first proposed by British Psychologist Charles Spearman in his research on 33 students' scores of classical language, French and English. He found the scores of these three subjects are highly correlated. According to this, he proposed that there might be a common element involved that can decide the scores of these three subjects. For example, language ability might be this element. This kind of factor can be called a common factor.

C. Market Activity Index

The variables, such as heterogeneity in access to roads, plazas, and districts, number of attributes, item's name, and algorithmically generated description within NFT pricing models, are predominantly fixed [8], [25], [26]. Unless altered intentionally by the NFT creators, this consideration remains relatively constant in most other scenarios, resulting in relatively fixed prices. However, in practice, the price of any asset traded in the market should exhibit variability influenced by factors within the market, which include investor trading preferences, gas fees, data availability, and the exchange rate in the market [27], [28], [29]. As our study aims to standardize NFT pricing, it naturally includes market fluctuation as one of its considerations.

Within this study, we design *market activity index* to encapsulate the collective impact and intensity of market activities represented by the individual indices. It serves as a single numerical representation that reflects the level and nature of market participation, liquidity, transaction volume, buyer–seller dynamics, price movements, profitability, and supply conditions within the studied market. We introduce the market activity index that influences NFT prices, allowing them to fluctuate within a reasonably defined range. This value can be computed based on data such as daily trading volumes, the number of new project releases, social media mentions, and similar factors for the current day. All these market activity indices are synthesized into a unified market activity index and integrated with the previously mentioned hedonic regression model. Each market activity index is associated with a specific date-time stamp to relate it to other specific dates. The objective is to amalgamate diverse datasets (such as daily trading volumes, new project releases, and social media mentions) into a unified market activity index. A comprehensive scoring method is employed to either weight or normalize these data, resulting in a composite activity value. This value serves as a new variable input into our regression model.

D. Comparison With Other Related Studies

In Table I, we compared our model with other four studies from the perspective of model type, computational cost, scalability, use case, and variable consideration. It is found that the previous four studies [4], [33], [34], [35] are based on different types of machine learning algorithms for predictive usage. Moreover, the data size of these studies is large or complex that causes high-computational costs. In addition, their study lacks a discussion of model implementation and complexity analysis that leads to low scalability. Our proposed scheme is based on the hedonic regression for pricing and predictive purposes at a relatively low-computational cost and high scalability. This is because our model offers flexible implementation options and could use off-chain methods to avoid computational costs

TABLE I
COMPARATIVE ANALYSIS OF THE PROPOSED MODEL WITH THE EXISTING MODELS

References	Model Type	Computational Cost	Scalability	Use Case	Variable Consideration
Costa et al. [34]	Machine Learning	High	Low	Predictive	Image + Text + User Preferences
Nadini et al. [4]	Machine Learning	High	Low	Predictive	Sale History + Visual Features
Julianto et al. [35]	Machine Learning	High	Low	Predictive	Transaction Data
Ma et al. [33]	Machine Learning	High	Low	Predictive	Time Series Data
Proposed Model	Hedonic Regression	Low	High	Pricing + Predictive	NFT Feature Rarities + Market Factors

in a blockchain environment. In terms of variables researched, Costa's [34], Nadini's [4], and our models incorporate both internal (feature-related) variables and external variables, while Julianto's [35] and Ma's [33] only consider external variables.

III. PROPOSED PRICING MECHANISM

This section introduces a novel pricing framework for NFTs utilizing feature rarities and designs our market activity index. By synthesizing multiple variables into composite indicators through factor analysis, factor analysis reduces dimensionality while retaining essential information. We present a rarity-driven model, aligned with the hedonic model, to establish NFT prices based on features. Moreover, we calculate the weight of related market factors and combine these factors with the previously proposed model for dynamic prices in our online market.

A. Explore Multidimensional Factors

In practical research, data collection is crucial for understanding objectives, but it can lead to overlap and excessive computation. Removing variables may result in data loss. Factor analysis synthesizes variables into composite indicators (factors), reducing dimensionality while minimizing data loss.

We assume a random vector $Y = (y_1, \ldots, y_p)'$ with an average of μ, where $l_{p,m}$ represents the loading. $y_p - \mu_p$ centralizes the data, which SPSS handles automatically in our experiments. When factors are uncorrelated, factor loadings are the correlation coefficients between the pth feature rarity and the mth factor. Higher absolute factor loadings indicate stronger correlations. This model assumes that feature rarity Y depends linearly on m unobservable common factors $F = (f_1, \ldots, f_m)'$ and p unobservable special factors $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)'$. The orthogonal factor model is expressed as

$$\begin{aligned} y_1 - \mu_1 &= l_{1,1}f_1 + l_{1,2}f_2 + \ldots + l_{1,m}f_m + \varepsilon_1 \\ y_2 - \mu_2 &= l_{2,1}f_1 + l_{2,2}f_2 + \ldots + l_{2,m}f_m + \varepsilon_2 \\ &\ldots \\ y_p - \mu_p &= l_{p,1}f_1 + l_{p,2}f_2 + \ldots + l_{p,m}f_m + \varepsilon_p. \end{aligned} \tag{1}$$

In the formula above, coefficient $l_{j,k}$ refers to the loading of the ith feature rarity on the kth factor, expressing the characterization of this factor on this feature rarity. If we use matrix signal, the formula above can be rewritten as follows, in which $L_{p\times m}$ is the loading matrix

$$Y_{p\times 1} - \mu_{p\times 1} = L_{p\times m} \times F_{m\times 1} + \varepsilon_{p\times 1}. \tag{2}$$

In our research, one of the key concepts is the factors' variance contribution. It refers to the sum of the square of elements of the mth column in the factor loading matrix, reflecting the ability of the mth factor to explain the total variance of the original variable. The higher this figure is, the more important the factor is

$$S_m^2 = \sum_{p=1}^{w} a_{m,p}^2. \tag{3}$$

In a nutshell, the orthogonal factor model is the core of factor analysis. It can be expressed in two ways. The first one is a matrixlike model

$$Y - \mu = LF + \varepsilon. \tag{4}$$

This model has several prerequisites. f and ε satisfy the assumption as follows:

$$E(f) = 0_{m\times 1}, \mathrm{COV}(f) = E(ff') = I_{m\times m}. \tag{5}$$

The first line of formula describes the assumption to common factors (f). The expected value of each common factor is zero and the covariance matrix is the identity matrix (orthogonality), which means that the variance of each common factor is one. f_j is uncorrelated to f_k if k is not equal to j

$$\begin{aligned} E(\varepsilon) = 0_{p\times 1}, \mathrm{COV}(\varepsilon) &= E(\varepsilon\varepsilon') = \psi_{p\times p} \\ &= \begin{bmatrix} \psi_1 & 0 & \ldots & 0 \\ 0 & \psi_2 & \ldots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \ldots & \psi_1 \end{bmatrix}. \end{aligned} \tag{6}$$

The second line of the formula describes the assumption of special factors, which is a little bit similar to the assumption of common factors. The expected value of each special factor is zero. Each ε belonging to different original variables should be independent, so there can be a diagonal matrix to explain their independence. One difference of assumption here from common factors is the variance of each special factor may not be equal to one.

The last assumption is that f is not correlated to ε: COV(ε, f) $= 0_{p\times m}$.

B. Rarity-Driven Model for NFT Pricing

In the preceding section, we elucidated the application of factor analysis as a sophisticated technique for dimensionality reduction. In this section, we shall examine the conventional iteration of the dynamic hedonic model in NFT pricing based

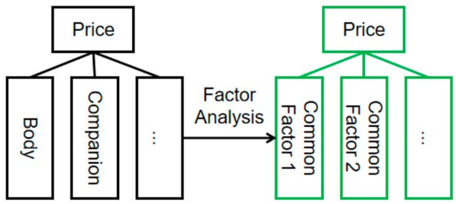


Fig. 1. Before and after using factor analysis on hedonic model.

TABLE II
FEATURES OF SELECTED NFT COLLECTIONS

HV-MTL	Sappy Seals	Lazy Lions	Meta Mansions by KEYS
Back Attachment	Background	Background	Accent Lighting
Body	Body	Body	Biome
Companion	Extra	Bodygear	Building Material
Crest	Face	Earring	Level
Faceplate	Head	Eyes	Light Color
Head	Skin	Headgear	Mansion
HV Type	N/A	Mane	Weather
Weapon System	N/A	Mouth	Windows

on rarity data from OpenSea.[1] Subsequently, we undertake the intricate task of model improvement through factor analysis.

The hedonic model is used to estimate the impact that external factors and internal characteristics have on the prices of assets and properties [36]. Compared to other research based on the hedonic model, our hedonic model focuses on NFT features. In our research, the implementation of the above concept can be expressed as follows:

$$P_i = a + \sum_{j=1}^{J} \alpha_j x_{j,i} + e_i \tag{7}$$

where P_i represents the sale price of an NFT, a is the regression intercept, and $\sum_{j=1}^{J} \alpha_j x_{j,i}$ denotes the features that single NFT has. $x_{j,i}$ refers to variable j of the ith NFT has. The coefficient α_j reflects the attribution of a relative shadow price to each of the j characteristics. e_i is the difference between the predictive value and the real value.

Before using a regression model, it is necessary to make assumptions about the data, which are the prerequisites for using this model. There are three main points: sample independence diagnosis, multicollinearity diagnosis, and finally, the normality of residuals.

The first assumption is that each observation in the dataset is independent. The Durbin–Watson (DW) test is one of the methods to deal with the problem of sample independence testing, detecting the presence of autocorrelation in the residuals of a regression. Generally, if the DW result is around 2, it will be acceptable, meaning that the degree of autocorrelation is low.

Second, multiple linear regression assumes that none of the predictor variables are highly correlated with each other. One way to detect multicollinearity is by using a metric known as the variance inflation factor (VIF). Generally, if it is lower than 5, it can be recognized that there is low multicollinearity. Third, residuals are normally distributed.

When considering factor analysis, the construction of the model has a slight alteration. Fig. 1 shows before and after using factor analysis on hedonic model. Table II shows four NFT collections' features. The rarity of these features is selected for the following model fitting.

C. A Method to Calculate Weight

In this experiment, we opted for the CRITIC method as the weighting approach to compute the market activity index. While there are other common weighting methods such as the coefficient of variation method, entropy weighting method, and independence weight method, each carries certain biases. The coefficient of variation method computes the variance S_i of data under each indicator and divides each S_i by the sum of all S_i to establish the weight of the ith indicator. The weight increases with a higher variance. The entropy weighting method calculates entropy weights based on the dispersion of data for each indicator, followed by adjustments for relatively objective indicator weights. The independence weight method determines weights based on the collinearity strength between indicators. If an indicator exhibits a high correlation with other indicators, indicating significant information overlap, its weight is lower; conversely, weak correlations suggest a higher information content for the indicator, warranting a higher weight. By contrast, the CRITIC weighting method comprehensively considers both the comparative strength and conflict between indicators [37].

D. Identification of Market Activity Index

We identified potential market activity index considerations from CryptoSlam. The reason for choosing the CryptoSlam platform is that compared to OpenSea, it has some more complete macrodata indicators (eight indicators in the article). These factors cover important aspects of the digital asset market, including but not limited to market liquidity, investor sentiment, and industry development trends. The comprehensive market indicator derived from these indicators combined with the previous Rarity formula can be used to form the pricing formula in the virtual market we have built. The detailed terminology is in Table III.

Through individual analyses of these factors between 23 June 2017 and 14 October 2023 (daily and weekly analyses), as demonstrated in Tables IV and V, it can be observed a high degree of correlation among most of the indicators. This further validates the suitability of using the CRITIC weighting method. As a considerable portion of the wash data is zero in daily data, causing inconvenience in weight calculation, we opted to perform computations using weekly data for analysis.

Out of the nine market activity indicators considered in this study, not all indicators have a positive impact on NFT prices. Indicators such as wash sales USD, wash sales transactions, and wash sales percentage reflect wash trading behavior in NFT transactions. This behavior undoubtedly undermines market and buyer confidence, resulting in a negative impact on NFT transaction prices. Therefore, these three indicators have

[1] https://opensea.io/

TABLE III
EXPLANATION OF EACH MARKET FACTOR

Name	Explanation
TotalTransactions	The overall volume of buying and selling activities related to the NFT
UniqueBuyers	An individual or entity that has acquired the NFT at least once
UniqueSellers	An individual or entity that has listed and successfully sold the NFT at least once
TotalPriceUSD	The cumulative amount of money exchanged through buying and selling the NFT
TradeProfitUsd	Calculated based on the difference between the selling price and the buying price of the NFT across all transactions
Supply	The total quantity or number of units of the specific NFT that exist or are available
WashSalesUsd	Reflecting the total transaction amount of wash sales
WashTransactions	Reflecting the total count of wash sales
WashSalesPercentage	Reflecting the percentage of wash sales in total transaction volume, i.e., the extent of wash sales

TABLE IV
DAILY SITUATION OF CORRELATION

Market Factors	Total Transactions	Unique Buyers	Unique Sellers	Total Price	Wash Sales	Wash Transactions	Trade Profit	Supply
Total Transactions	1	0.976	0.978	0.941	0.901	0.889	0.039	0.905
Unique Buyers	0.976	1	0.982	0.971	0.904	0.872	0.085	0.887
Unique Sellers	0.978	0.982	1	0.954	0.913	0.888	0.077	0.889
Total Price Usd	0.941	0.971	0.954	1	0.887	0.837	0.156	0.854
Wash Sales Usd	0.901	0.904	0.913	0.887	1	0.95	−0.069	0.884
Wash Transactions	0.889	0.872	0.888	0.837	0.95	1	−0.176	0.901
Trade Profit Usd	0.039	0.085	0.077	0.156	−0.069	−0.176	1	−0.174
Supply	0.905	0.887	0.889	0.854	0.884	0.901	−0.174	1

TABLE V
MONTHLY SITUATION OF CORRELATION

Market Factors	Total Transactions	Unique Buyers	Unique Sellers	Total Price	Wash Sales	Wash Transactions	Wash Sale Percentage	Trade Profit	Supply
Total Transactions	1	0.978	0.978	0.945	0.921	0.898	0.688	0.065	0.913
Unique Buyers	0.978	1	0.979	0.974	0.926	0.885	0.676	0.11	0.898
Unique Sellers	0.978	0.979	1	0.95	0.934	0.905	0.704	0.091	0.904
Total Price Usd	0.945	0.974	0.95	1	0.907	0.847	0.622	0.19	0.862
Wash Sales Usd	0.921	0.926	0.934	0.907	1	0.952	0.849	−0.037	0.909
Wash Transactions	0.898	0.885	0.905	0.847	0.952	1	0.823	−0.165	0.913
Wash Sales Percentage	0.688	0.676	0.704	0.622	0.849	0.823	1	−0.226	0.709
Trade Profit Usd	0.065	0.11	0.091	0.19	−0.037	−0.165	−0.226	1	−0.146
Supply	0.913	0.898	0.904	0.862	0.909	0.913	0.709	−0.146	1

a negative impact on establishing the composite market activity indicator.

However, the other five market activity indicators, to some extent, reflect the situation and prospects of the NFT market and have a positive impact on the composite market activity indicator. Thus, the formula for the composite market activity indicator is as follows, where M_1 denotes total transactions, M_2 denotes unique buyers, M_3 denotes unique sellers, M_4 denotes total price USD, M_5 denotes wash sales USD, M_6 denotes wash transactions, M_7 denotes wash sales percentage, M_8 denotes trade profit USD, and M_9 denotes supply

$$\begin{aligned} M = {} & \alpha_1 M_1 + \alpha_2 M_2 + \alpha_3 M_3 \\ & \alpha_4 M_4 + \alpha_5 M_5 + \alpha_6 M_6 \\ & - \alpha_7 M_7 - \alpha_8 M_8 - \alpha_9 M_9 \quad (p > 0). \end{aligned} \tag{8}$$

E. The Determination of Pricing Model Based on Feature Rarity and Market Factors

In the prior sections, we addressed the creation of dynamic hedonic pricing model grounded in rarity and the development of the comprehensive market activity index. In this study, the calculated comprehensive market activity, expressed as an ETH value, directly quantifies the influence of market activity on pricing. This value can be directly added to the predicted results from the hedonic model, culminating in the final forecasted outcome

$$T = P + M \tag{9}$$

where P is the predicted outcome of the model based on rarity, and M is the comprehensive market activity factor. Algorithm 1 provides the detailed logic to show how to implement this advanced pricing mechanism. Moreover, this algorithm could be used in different conditions such as smart contract and JavaScript based on real-world requirements.

IV. EXPERIMENTS

In this section, we conduct relevant experiments and corresponding analyses to build and implement a pricing model. Initially, we introduce how factor analysis operates and perform necessary processing and analyses. Subsequently, we select the NFT datasets and calculate precise model parameters for pricing the designated NFT category. Finally, we design a "market factor + rarity"-based NFT market with multiple entities and successfully implement it on the computer.

Algorithm 1 NFT Pricing Algorithm

Require$(V_1, V_2, V_3, \ldots, V_j)$
  $P(Wight) \leftarrow a + \sum_{j=1}^{J} \beta_j V_j + e$
Input$(x_1, x_2, x_3, \ldots, x_8, M_1, M_2, M_3, \ldots, M_9)$
  $P(Price) \leftarrow V_1 + V_2 x_1 - V_3 x_2 - V_4 x_3 - V_5 x_4 + V_6 x_5 - V_7 x_6 - V_8 x_7 + V_9 x_8$

  $M(Market\ Factor) \leftarrow \alpha_1 M_1 + \alpha_2 M_2 + \alpha_3 M_3 + \alpha_4 M_4 + \alpha_5 M_5 + \alpha_6 M_6 - \alpha_6 M_7 - \alpha_8 M_8 - \alpha_9 M_9$
if $Price < 0$ **then**
  return Error: Price cannot be negative
end if
$T(TotalPrice) \leftarrow P(x_1, x_2, x_3, \ldots, x_8) + M(M_1, M_2, M_3, \ldots, M_9)$
return $T(TotalPrice)$

A. Data Collection

We collected the data described in this section. The NFT data were derived from OpenSea, one of the greatest NFT online markets.

In this study, we selected Sappy Seals,[2] HV-MTL,[3] Lazy Lions,[4] and Meta Mansions by KEYS[5] as our research subjects for two reasons. First, their inclusion from different NFT categories helps to generalize the usability of our model. HV-MTL is a game that begins as a blend of a pet simulation game and a casual world builder, evolving into a competitive dungeon crawler. Lazy Lions NFTs are profile-picture (PFP)-style NFTs that owners can use as digital avatars online. Meta Mansions by KEYS consists of 8888 digital luxury mansions, forming the main residency of the KEYS Metaverse. Sappy Seals, as implied by the name, features seal-themed NFT PFPs, each with unique attributes.

The second reason for our selection is feature set suitability. We observed that the number of features in these NFT collections is neither too high nor too low, making them ideal for regression experiments. Another crucial factor is the consistency of feature categories. For instance, in an NFT collection, there might be a situation where the first NFT has blue hair and no wings, the second has purple hair and no wings, and the third has no hair but white wings. In this scenario, the first and second NFTs have the hair feature category but lack the wing category, while the third NFT is the opposite. If such inconsistencies occur frequently, it hinders the construction of a reliable regression model for rarity and its market integration in the later stage of our article. Therefore, we aimed to select NFT collections where such discrepancies are minimal.

The data of 1808 NFTs are from Sappy Seals, 2153 NFTs are from HV-MTL, 1371 NFTs are from Lazy Lions, and 594 NFTs are from Meta Mansions by KEYS. Feature rarity data on each NFT are calculated as

$$\text{rarity} = \frac{N_f}{N_t} \tag{10}$$

where N_f refers to the number of NFTs having this feature, and N_t refers to the total number of NFTs in the collection. Before using the hedonic model, we first do a linear correlation analysis, the reason of which is that it is beneficial to analyze how factor analysis improves models based on different NFT collections. Correlation analysis is mainly help to analyze the effect of factor analysis on the model, but factor analysis is not in the step of building a model suitable for virtual markets, so there is no complexity problem in practical application. Correlation is absolutized

$$C_1 = \text{abs}(C_0) \tag{11}$$

where C_1 is the absolutized correlation and C_0 is the original correlation. In Figs. 2 and 3, color legends from left to right indicate an increasing correlation. From these two graphs, it can be seen that there is some correlation between the features of HV-MTL, while the correlation between the features of Sappy Seals is not as strong. Therefore, the hedonic model, as one of the linear regression models, is predicted to perform better in HV-MTL.

The reason why there is no correlation analysis of Lazy Lions and Meta Mansions by KEYS is that their KMO value is not higher than 0.6, meaning they are not suitable for factor analysis, so there is no need for correlation analysis. But at the

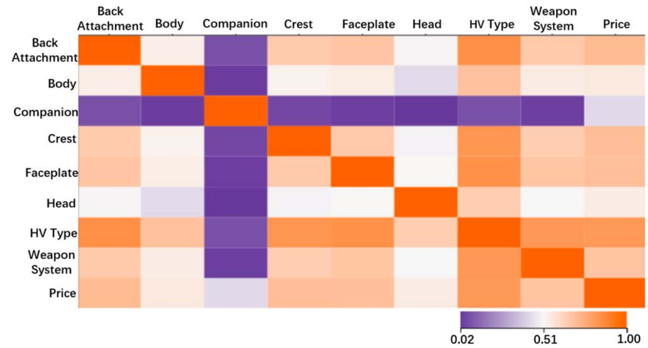


Fig. 2. Heatmap of feature correlations in the HV-MTL dataset.

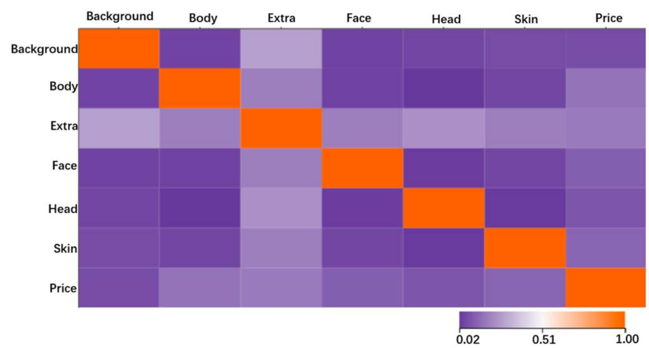


Fig. 3. Heatmap of feature correlations in the Sappy Seals dataset.

[2]https://opensea.io/collection/sappy-seals
[3]https://opensea.io/collection/hv-mtl
[4]https://opensea.io/collection/lazy-lions
[5]https://opensea.io/collection/metamansionsbykeys

TABLE VI
ADAPTIVE ANALYSIS OF TWO DATASETS

	HV-MTL	Sappy Seals	Lazy Lions	Meta Mansions by KEYS
KMO	0.731	0.635	0.490	0.471
Significant	0.000	0.000	0.222	0.000

same time, it is feasible to fit the relationship between price and rarity in these two datasets based on a regression model. The KMO results and linear fitting results for these two datasets can be seen in Sections IV-B and V, respectively.

B. Experimental Results

Factor analysis can amalgamate original variables and construct novel shared factors. In the upcoming experiments, we intend to delve into this technique to ascertain whether its application can potentially yield favorable alterations in the outcomes of our model. Through this investigation, we aim to evaluate the potential of factor analysis to enhance the model's performance and influence the results.

1) Data Preprocessing: Since different indicators have different quantitative outlines and may not be comparable, it is necessary to standardize the raw data to eliminate the effect of quantitative outlines. The standardization formula is as follows, in which Z_i is the original data and Z_i is the standardized data:

$$Z_i = X_i - \mu. \tag{12}$$

2) Adaptive Analysis: Adaptive analysis is used to justify whether the research method is suitable for the data. So before we carry out factor analysis, it is necessary to justify whether there is enough correlation between original variables, which is the precondition of implementing factor analysis. Generally, the variable should be subject to Bartlett's and Kaiser–Meyer–Olkin (KMO) testing. Taking the correlation coefficient matrix of the original variables as a starting point, and assuming that the correlation coefficients are unit matrices, if the p-value corresponding to this test is less than the given level of significance a, the original hypothesis should be rejected, and the original variables should be considered suitable for factor analysis.

The KMO test takes values between 0 and 1, the closer to 1 the stronger the correlation of the variables, indicating the more suitable for factor analysis. If the KMO value is greater than 0.6 and Bartlett's test of sphericity is significant [38], and then the dataset is suitable for factor analysis. Table VI shows that both HV-MTL and Sappy Seals are suitable for factor analysis, while Lazy Lions and Meta Mansions by KEYS are not.

3) Number of Common Factors: Scree Plotto can help to select the number of common factors through factor analysis, which illustrates eigenvalues on the y-axis against the number of factors on the x-axis. The plot typically shows a descending curve. The "elbow" of the curve, where the slope noticeably levels off, indicates the optimal number of factors to be derived from the analysis. But the decision to "elbow" is subjective.

In our article, we use another important metric namely the total amount of variability of the original variables explained by each factor solution. In factor analysis, the number of factors corresponds to the number of variables, and these factors are

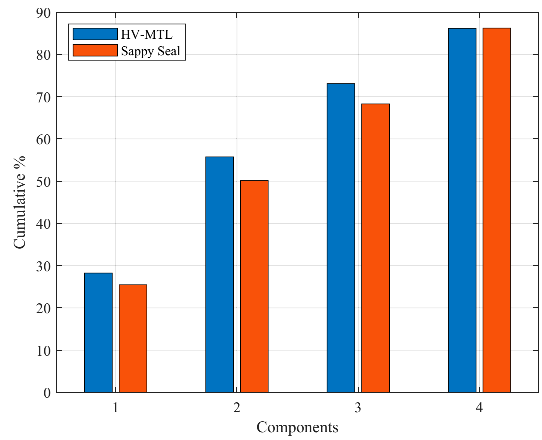


Fig. 4. Total variance of growing components in the HV-MTL and Sappy Seals datasets.

arranged based on the variance they account for. While retaining more factors can capture a greater overall variance, additional factors beyond the initial few minimally explain the variability. For instance, if the first three factors collectively account for a substantial portion of the variance originally explained by ten variables, these factors can effectively replace the complexity of all ten variables. Consequently, dropping the remaining factors would result in a minimal loss of the original variability. The accumulative variance contribution can illustrate how these new common factors explain those original ones.

To determine the number of common factors, it is important to observe the result of the total variance. The calculation of the total variance is as follows:

1) Assume S_1^2, S_2^2,..., S_p^2 are p common factor variance.
2) The first variance is $C_1 = S_1^2 \,/\, p$.
3) The accumulative variance contribution of the first k common factors are $C_k = \sum_{i=1}^{k} S_i^2 \,/\, p$.

The higher the accumulative variance is, the better common factors can explain the original variables. Generally, 85% is a critical value. As such, according to Fig. 4, we can select four common factors in both datasets.

4) Regression Result: To conduct this study, it is necessary to assess the efficacy of the proposed dynamic pricing model within the context of an NFT market. To achieve this, we must utilize data from previous datasets to derive a specific formula that yields regression results. Upon reviewing Table VII and analyzing the regression outcomes for both datasets, the HV-MTL result has been selected as the ideal candidate for demonstrating the corresponding model formula. The formula of it is presented as follows where p refers to price, x_1–x_8 refers to from features 1 to 8:

$$\begin{aligned} p = 3.638 &+ 0.206x_1 - 1.202x_2 - 0.007x_3 \\ &- 0.322x_4 + 0.208x_5 - 2.035x_6 \\ &- 0.116x_7 + 0.143x_8 \quad (p > 0). \end{aligned} \tag{13}$$

C. Market Factor

The selected indicators in the analytical framework each possess their respective dimensions and variations, resulting

TABLE VII
REGRESSION RESULT OF HV-MTL AND SAPPY SEALS

	HV-MTL	Sappy Seals	Lazy Lions	Meta Mansion by Keys
Constant	3.638	2.032	0.453	0.225
Feature 1	0.206	0.015	0.000485	0.000018
Feature 2	–1.202	–0.01	–2.612E–5	–0.01
Feature 3	–0.007	–0.015	–0.014	–0.000369
Feature 4	–0.322	–0.001	–0.000184	0.00026
Feature 5	0.208	–0.01	–0.000222	–0.014
Feature 6	–2.035	–0.005	–0.000164	0.000126
Feature 7	–0.116	N/A	–0.001	–0.000151
Feature 8	0.143	N/A	–0.000279	N/A

TABLE VIII
WEIGHT CALCULATION

Market Factors	Sigma	Sum	C_j	W_j
Total Transactions	0.287014154	6.17194784	1.77143639	0.144502469
Unique Buyers	0.270968096	6.053079508	1.64019143	0.133796343
Unique Sellers	0.30699404	6.084144772	1.867796183	0.152362886
Total Price Usd	0.15395727	6.122289845	0.942571033	0.076888926
Trade Profit Usd	0.084793948	6.737298233	0.571282119	0.046601547
Supply	0.197617083	7.757913981	1.533096334	0.125060209
Wash Sales Usd	0.102693003	9.481864169	0.97372111	0.07942995
Wash Transactions	0.178657597	9.832671393	1.756681447	0.143298855
Wash Sales Percentage	0.128047807	9.387820665	1.202089848	0.098058814

in inconvenience for composite analysis modeling. Therefore, preprocessing of the collected data is necessary to eliminate the effects of dimensions and variations. The CRITIC weight method typically involves either positive or reverse handling. Standardization is deemed inappropriate here because using this process would render all standard deviations as 1, meaning all indicators would have identical standard deviations, thus rendering volatility indicators meaningless.

1) Nondimensionalization: If the higher values of the indicators are preferable, such indicators are considered positive indicators. In this study, indicators such as total transactions, unique buyers, unique sellers, total price USD, trade profit USD, and supply belong to this category. The formula for positive handling is as follows, where $x_{\min}$ represents the minimum value, $x_{\max}$ denotes the maximum value, and x_{ij} represents the numerical value of the *j*th evaluation indicator for the *i*th sample

$$x_{ij}{}' = \frac{x_j - x_{\min}}{x_{\max} - x_{\min}}. \tag{14}$$

If lower values of the indicators are preferable, such indicators are considered negative indicators. In this study, they include wash sales usd, wash transactions, and wash sales percentage. The formula for negative handling is similar in structure to the positive handling formula, as follows:

$$x_{ij}{}' = \frac{x_{\max} - x_j}{x_{\max} - x_{\min}}. \tag{15}$$

2) Conflict Measurement: The correlation coefficient is used to illustrate the relationship between indicators. The stronger the correlation between an indicator and others, the less conflicting it is with other indicators. This implies that it reflects more redundant information, thereby weakening the evaluative intensity of that particular indicator. Accordingly, the weight allocated to such an indicator should be reduced. The formula is as follows, where r_{ij} represents the correlation coefficient between evaluation indicators i and j, and R_j signifies the sum of the correlation coefficients involving j

$$R_j = \sum_{i=1}^{p}(1 - r_{ij}). \tag{16}$$

3) Contrast Intensity: In the CRITIC method, the standard deviation is utilized to represent the internal variation in the values of each indicator. A larger standard deviation suggests greater variability in the indicator's values, reflecting more information and stronger evaluative intensity for that specific indicator. The formula for expressing the variability of indicators in terms of standard deviation is as follows, where S_j is the standard deviation of the jth indicator:

$$S_j = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2}{n-1}}. \tag{17}$$

4) Information Content: The information content determines the significance of each indicator. A higher value suggests that the indicator plays a more substantial role in the entire evaluation system and should, therefore, be allocated more weight. The formula for the information content of indicator j equals the product of its variability and conflict, given by

$$C_j = S_j \times R_j. \tag{18}$$

5) The Determination of Weight:

$$W_j = \frac{C_j}{\sum_{k=1}^{m} C_j}. \tag{19}$$

where W_j represents the proportion of indicator j in the total information content, i.e., the weight. In this experiment, the proportion of each market activity in the total information content is illustrated in the W_j column in Table VIII. This table shows each variable's weight is decided by both standard deviation and correlation.

Therefore, the formula for the comprehensive market factor is as follows:

$$\begin{aligned} M = {} & 0.145M_1 + 0.134M_2 + 0.152M_3 \\ & + 0.077M_4 + 0.047M_5 + 0.125M_6 \\ & - 0.079M_7 - 0.143M_8 - 0.098M_9. \end{aligned} \tag{20}$$

D. Implementation of the NFT Pricing Model

To implement the proposed pricing model based on market factors and to evaluate its performance in practice, we designed and established an NFT market with the following entities shown in Fig. 5, and the webpage of the NFT market is shown in Fig. 6. The code files of implementing the prototype market can be found at our GitHub repository.[6]

Our marketplace architecture integrates several key components to ensure seamless and efficient operations. The digital wallet, specifically MetaMask, facilitates user interactions by accepting or rejecting requests from both the server and users, enhancing fault tolerance and stability. The front-end provides an intuitive interface, allowing users to engage with the marketplace effortlessly. It communicates user actions to the back-end,

[6]https://github.com/Mastin-z/NFTmarketplace

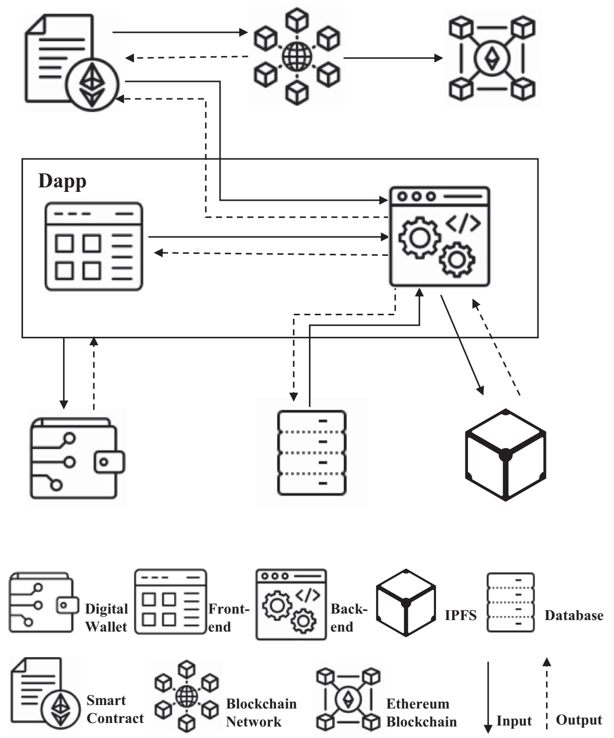


Fig. 5. System architecture of the proposed NFT market based on the pricing model.

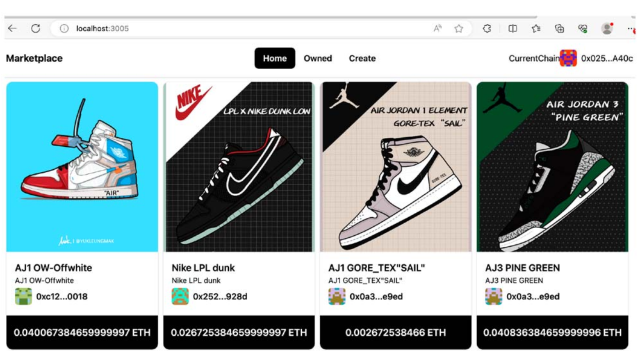


Fig. 6. Webpage of the proposed NFT market based on the pricing model.

which is built on a NodeJS architecture. The back-end processes these actions, dynamically calculates NFT price fluctuations using a JavaScript-based pricing algorithm, and stores relevant data.

The database, implemented as a CSV file, manages weekly market data updates for real-time price calculations, while the interplanetary file system (IPFS) stores and distributes NFT data, ensuring data immutability [39]. Smart contracts, redesigned to integrate market elements, interact with both the blockchain network and the back-end to facilitate transactions, solving floating-point number issues and reducing costs. The Sepolia Ethereum test network evaluates the marketplace's performance, recording transaction information that users can verify through the front-end. The Ethereum Blockchain, structured on the proof of stake (PoS) consensus algorithm [40],

TABLE IX
HEDONIC MODEL PERFORMANCE WITH FACTOR ANALYSIS

	HV-MTL	HV-MTL With Common Factors	Sappy Seals	Sappy Seals With Common Factors
R^2	0.251	0.27	0.487	0.46
MAE	0.535	0.518	0.056	0.067
MSE	2.149	1.076	0.02	0.021
RMSE	1.466	1.037	0.141	0.144
MSD	−0.014	0.064	0.006	0.002
MAPE	0.254	0.244	0.094	0.122
Adjusted R^2	0.251	0.27	0.487	0.46

TABLE X
HEDONIC MODEL PERFORMANCE

	Lazy Lions	Meta Mansions by KEYS
R^2	0.171	0.426
MAE	0.029	0.03
MSE	0.002	0.005
RMSE	0.044	0.072
MSD	0.001	−0.002
MAPE	0.074	0.257
Adjusted R^2	0.171	0.426

offers resistance to information tampering and decentralization, enhancing transparency and security. This robust blockchain infrastructure allows users to trace transaction details seamlessly, ensuring the integrity of transaction history. Overall, the architecture emphasizes transparency, reliability, and user-friendliness, providing a robust environment for NFT issuance, selling, and trading.

In addition to the previously implemented pricing model in our designed NFT market, the model demonstrates scalability, building on the established logic of the pricing algorithm. Our proposed pricing mechanism combines feature-based pricing with market factor-based pricing, offering flexible implementation options. Specifically, the feature calculation process can be executed in an off-chain environment, such as JavaScript. This approach reduces costs and computing resource demands, as it merely transfers the computed price to the smart contract, thus minimizing direct interaction with the blockchain. Regarding the market factor-based pricing, it has been determined that the values required for the calculation can be stored either in a database or within a smart contract. These values need only be updated on a weekly or monthly basis, depending on specific requirements, thereby enhancing efficiency.

V. EVALUATION AND COMPARISON ANALYSIS

We employed the hedonic model to assess its effectiveness in predicting the price of NFTs, focusing solely on the feature of rarity. Tables IX and X illustrate the statistical methods used to assess the performance of the hedonic model, including the coefficient of determination or the coefficient of multiple determination in case of multiple regression (R^2), mean absolute error (MAE), mean square error (MSE), root mean square error

(RMSE), mean signed difference (MSD), and mean absolute percentage error (MAPE), adjusted R^2.

As data presented in Table IX, it is evident that the majority of indicators show the improvement of factor analysis in the model. The hedonic model can account for more than 25 % of HV-MTL data. Notably, after incorporating common factors, the model exhibits improvements compared to its initial state. This improvement is evident not only in the R^2 value, which increases from 0.251 to 0.27, but also in other key indices. For most of them, the lower the value is, the smaller the error is, and the better the prediction effect of the model is, such as MAE, which decreases from 0.535 to 0.518, MSE, which drops by half, from 2.149 to 1.076, or RMSE, going down from 1.466 to 1.037.

We also incorporated the Sappy Seals dataset into our experiment. Table IX reveals that introducing common factors has a detrimental impact on the results in this case because the values of R^2 and adjusted R^2 have been reduced, and almost all of the remaining values have been increased. Consequently, the performance disparities between HV-MTL and Sappy Seals datasets suggest that applying factor analysis is likely to yield positive effects on the model when there is a high correlation between variables. In cases where such a high correlation is absent, caution should be exercised when considering its implementation.

It is worth noting that the MSD shows a different signal from other indicators. MSD considers the sign difference between the predicted value and the actual value, including the positive and negative directions of the forecast bias. When the value of MSD is positive, it means that the model as a whole overestimates the true value. When the value of MSD is negative, it means that the model as a whole underestimates the true value. Therefore, the value of MSD not only takes into account the accuracy of the prediction but also reflects the biased nature of the model. Although the overall prediction accuracy (e.g., R^2, MAE, and MSE) has improved after the model has been applied to the dimensionality reduction method, there may still be some deviations in the model, resulting in the MSD value deviating from expectations. For example, the model may overestimate or underestimate true values in certain data ranges, resulting in positive or negative MSD values.

Since the datasets of Lazy Lions and Meta Mansions by KEYS were evaluated as unsuitable for factor analysis, only their regression results were calculated. Table X illustrates that the rarity of Lazy Lions-themed NFTs can explain 17.1% of their price, while Meta Mansions by KEYS is much higher, where they can explain 42.6% of price determination.

A. The Assessment of the Model Based on Rarity and Market Factor

The last experiment aims to elucidate the assessment of rarity's influence on NFT pricing by examining the difference between historical and predicted data. However, relying solely on rarity may not comprehensively account for the fluctuation in NFT prices within a specific range. Thus, we introduced market factors to augment this analysis. Given the volatile nature of current NFT prices, an alternative approach was sought to evaluate the model, leading to the utilization of Monte Carlo testing with 10 000 simulated samples. As depicted in Figs. 7 and 8, comprehensive market factors demonstrate a degree of variability, with the majority registering above 0. This variance suggests a necessity for model adjustments and refinements, hinting at potential adaptability requirements.

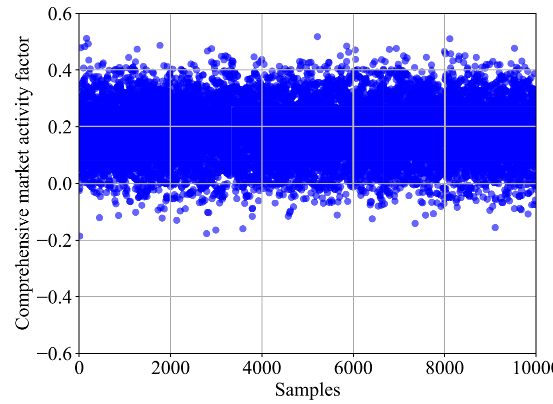


Fig. 7. Scatter plot of Monte Carlo testing before adjustment.

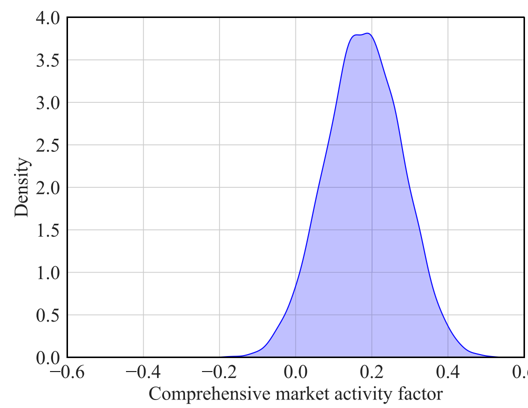


Fig. 8. Distribution of the outcome of Monte Carlo testing before adjustment.

Sensitivity analysis is a method of studying and analyzing the sensitivity of a model's state or output changes to changes in system parameters or surrounding conditions. The general method is to control the change of the value of an important parameter in the model while other parameters remain unchanged and then observe the change in the result. As shown in Fig. 9, in addition to coefficients 4 and 8, the pricing model is more sensitive to other coefficient parameters. Among them, the sensitivity curves of coefficients 1 and 6, and coefficients 2 and 5 partially overlap. In addition, changes in coefficients 3 and 7 have a greater impact on pricing.

We used sensitivity analysis, another type of assessment analysis, in this case. Sensitivity analysis places a different emphasis than error analysis, which focuses on analyzing the cause of the error involved in the projected model and genuinely is utilized for predicted problems. It is a technique for determining how sensitive a model's output or state change is to changes in the

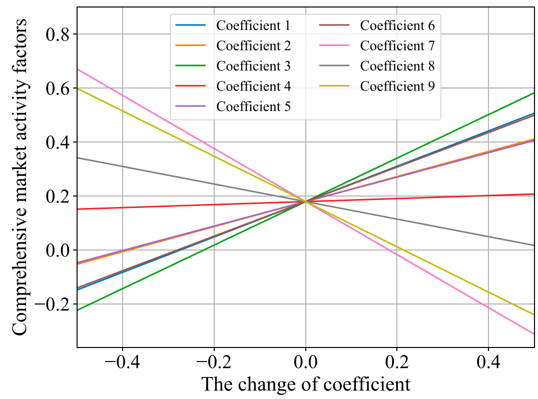


Fig. 9. Sensitivity analysis of market factors.

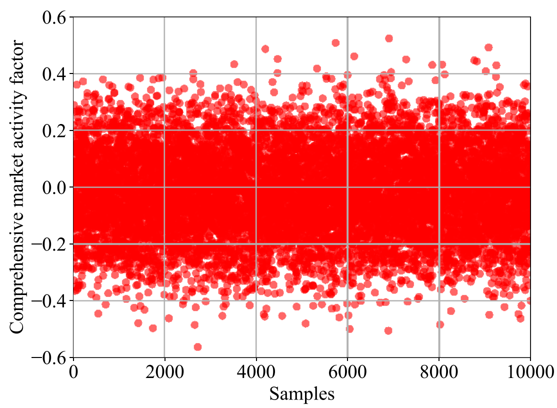


Fig. 10. Scatter plot of Monte Carlo testing after adjustment.

coefficient. Stated differently, hold other coefficients constant while adjusting a particular coefficient and see how the situation changes. Fig. 9 illustrates how each market factor intersects with a zero line.

As Figs. 7 and 8 illustrate, the model needs to be modified, and based on Fig. 9, we chose to adjust the absolute values of coefficient of market factors 7 and 9 up by 0.2 each. The modified efficacy can be shown in Figs. 10 and 11. This graph demonstrates that comprehensive market factors can float within a certain range, rather than being heavily skewed toward certain intervals.

B. NFT Market Evaluation

In the actual market application, the nine indicators in the market activity index are dynamic according to the actual market situation, and the fixed coefficients in the previous experiments are only to facilitate the verification of the feasibility of the method. The experiment's outcomes of the actual market are depicted in Fig. 12. The graph showcases the price trend of the NFT over a five-week period, with the y-axis indicating the price value and the x-axis indicating the corresponding phase date. It is evident from the data that the price has an increasing trend, primarily due to the continuous and significant increase in all dynamically fluctuating single-market factors such as $Supply$, $TotalTransactions$,

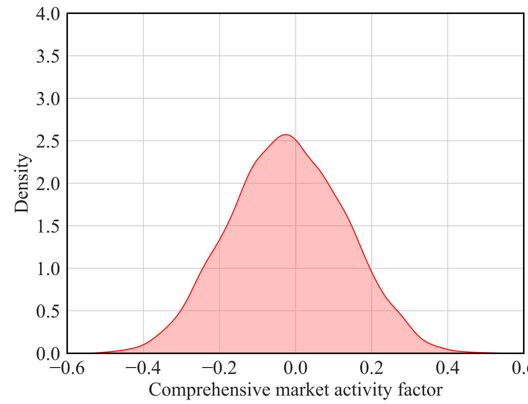


Fig. 11. Distribution of the outcome of Monte Carlo testing after adjustment.

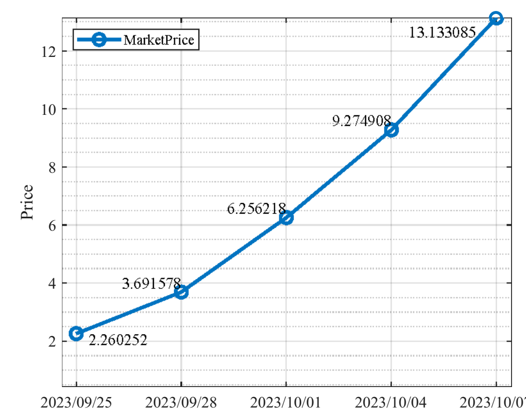


Fig. 12. Price of individual NFT changes on a weekly basis, depending on market factors' based hedonic model.

$UniqueBuyers$, $UniqueSellers$, $TotalPriceUsers$, and $TotalPriceUsd$. Meanwhile, the three factors with a downward trend in price, namely $WashSalesUSD$, $WashSalesTransactions$, $WashSalesPercent$, and $WashSalesPercentage$ are not fluctuating significantly. Hence, the NFT's price trend is generally positively correlated to the number of times it is traded in the market by the public. This has led to a corresponding increase in its price. Additionally, the NFT's trading volume growth over consecutive weeks is relatively high. The trend is consistent with the results shown in Fig. 12, which confirms the reasonableness of our pricing model based on market activity and NFT rarity. Thus, it verifies the feasibility of our proposed pricing mechanism for real-world applications.

C. Model Implement Evaluation

The model for our NFT marketplace is tested on the Ethereum Sepolia testnet, assumed to mirror mainnet performance. Transaction monitoring via EtherScan indicates a typical processing time of approximately 12 s. We analyzed the gas costs of five transactions under various network conditions (see Fig. 13), finding them reasonable and low for blockchain interaction, suggesting effective implementation on Ethereum. Moreover, when considering adaptation to other platforms or

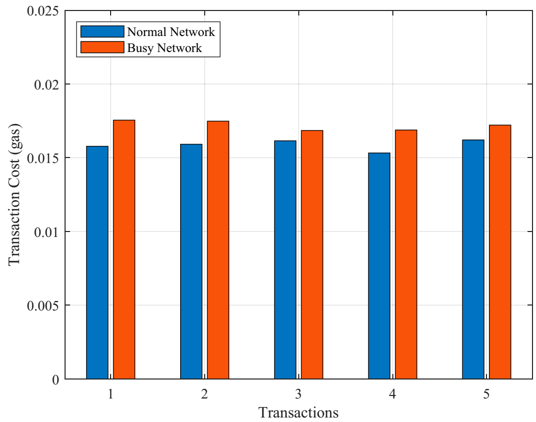


Fig. 13. Costs of five NFT transactions vary across different network conditions.

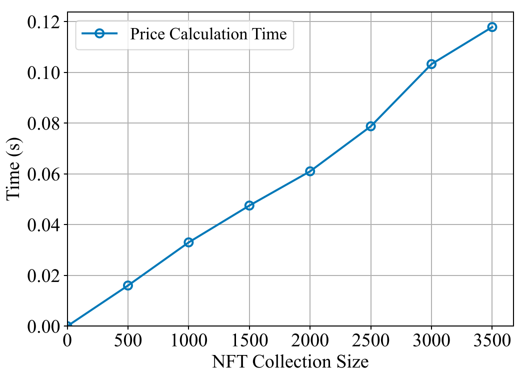


Fig. 14. Time required to calculate prices increases with the size of the NFT collection.

real-world applications, it is necessary to follow the logic of the proposed pricing model and adjust it to meet the specific requirements of each platform. The challenge lies in the time and effort required to learn about different platforms and to make tradeoffs among various tools within the rapidly evolving blockchain technology landscape. Regarding the previously discussed off-chain price calculation, Fig. 14 illustrates the relationship between NFT collection size and price calculation time. The collection size varies from 0 to 3500, while the calculation time ranges from 0.00 to 0.12 s. The data trend indicates that larger NFT collections generally require longer calculation times; however, the time cost remains relatively low, conserving computing resources and ensuring both scalability and efficiency.

VI. Conclusion

This article introduces a dynamic NFT pricing mechanism based on rarity scores from various collections. Our pricing model, grounded in the hedonic framework and enhanced with market factors, uses regression outcomes to establish an NFT market as a pricing solution, demonstrating the NFT market's real-world usability. Our solution offers a clear formula for efficient NFT pricing, enhancing transparency and comprehension. The evaluation shows that factor analysis can improve the model. Our pricing mechanism has potential applications in the collectibles market, requiring adaptation to unique market dynamics, asset types, and valuation factors such as rarity, historical significance, and demand trends. Future research will focus on expanding datasets and categorizing NFTs by type, such as art and games, and developing flexible models adaptable to varying features and data distributions. Techniques such as ensemble methods, neural networks, and deep learning could address these challenges.

References

[1] S. B. Far, A. I. Rad, S. M. H. Bamakan, and M. R. Asaar, "Toward metaverse of everything: Opportunities, challenges, and future directions of the next generation of visual/virtual communications," *J. Netw. Comput. Appl.*, vol. 217, 2023, Art. no. 103675.

[2] S. B. Far, S. M. H. Bamakan, Q. Qu, and Q. Jiang, "A review of non-fungible tokens applications in the real-world and metaverse," *Procedia Comput. Sci.*, vol. 214, pp. 755–762, 2022.

[3] S. B. Far and S. M. H. Bamakan, "Blockchain-based reporting protocols as a collective monitoring mechanism in DAOs," *Data Sci. Manage.*, vol. 5, no. 1, pp. 11–12, 2022.

[4] M. Nadini, L. Alessandretti, F. Giacinto, M. Martino, L. M. Aiello, and A. Baronchelli, "Mapping the NFT revolution: Market trends, trade networks, and visual features," *Sci. Rep.*, vol. 11, no. 1, p. 20902, 2021.

[5] B. Guidi and A. Michienzi, "From NFT 1.0 to NFT 2.0: A review of the evolution of non-fungible tokens," *Future Internet*, vol. 15, no. 6, p. 189, 2023.

[6] M. Solouki and S. M. H. Bamakan, "An in-depth insight at digital ownership through dynamic NFTs," *Procedia Comput. Sci.*, vol. 214, pp. 875–882, 2022.

[7] S. M. H. Bamakan, N. Nezhadsistani, O. Bodaghi, and Q. Qu, "Patents and intellectual property assets as non-fungible tokens; key technologies and challenges," *Sci. Rep.*, vol. 12, no. 1, Feb. 2022.

[8] C. Yencha, "Spatial heterogeneity and non-fungible token sales: Evidence from decentraland LAND sales," *Finance Res. Lett.*, vol. 58, Part A, 2023, Art. no. 103628.

[9] A. Mekacher et al., "Heterogeneous rarity patterns drive price dynamics in NFT collections," *Sci. Rep.*, vol. 12, no. 1, p. 13890, 2022.

[10] W. Xiong, Y. Wang, W. Li, J. Zhang, and H. Guo, "Pricing mechanism of non-fungible token (NFT) driven by rarity design," in *Proc. IEEE Int. Conf. Blockchain (Blockchain)*, Piscataway, NJ, USA: IEEE Press, 2023, pp. 74–79.

[11] H. Zhong and M. Hamilton, "Exploring gender and race biases in the NFT market," *Finance Res. Lett.*, vol. 53, no. 103651, 2023, Art. no. 103651.

[12] N. Borri, Y. Liu, and A. Tsyvinski, "The economics of non-fungible tokens," Mar. 2022. Accessed: Feb. 26, 2024. [Online]. Available: https://ssrn.com/abstract=4052045.

[13] A. Kapoor, D. Guhathakurta, M. Mathur, R. Yadav, M. Gupta, and P. Kumaraguru, "TweetBoost: Influence of social media on NFT valuation," in *Proc. Companion Proc. Web Conf.*, 2022, pp. 621–629.

[14] J. Luo, Y. Jia, and X. Liu, "Understanding NFT price moves through social media keywords analysis," 2022, *arXiv:2209.07706*.

[15] C. Urom, G. Ndubuisi, and K. Guesmi, "Dynamic dependence and predictability between volume and return of non-fungible tokens (NFTs): The roles of market factors and geopolitical risks," *Finance Res. Lett.*, vol. 50, no. 103188, 2022.

[16] S. Osivand and H. Abolhasani, "Effect of bitcoin and Etherium on non-fungible token (NFT)," *IOSR J. Bus. Manage.*, vol. 23, no. 9, pp. 49–51, 2021.

[17] C. Pinto-Gutiérrez, S. Gaitán, D. Jaramillo, and S. Velasquez, "The NFT hype: What draws attention to non-fungible tokens?" *Mathematics*, vol. 10, no. 3, 2022.

[18] D.-R. Kong and T.-C. Lin, "Alternative investments in the Fintech era: The risk and return of non-fungible token (NFT)," Aug. 2021. Accessed: Feb. 26, 2024. [Online]. Available: https://ssrn.com/abstract=3914085.

[19] L. Ante, "Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations," *Econ. Innov. New Technol.*, vol. 32, pp. 1–19, Nov. 2022.

[20] W. Entriken, D. Shirley, J. Evans, and N. Sachs, "EIP-721: ERC-721 non-fungible token standard," Jan. 2018. Accessed: Feb. 26, 2024. [Online]. Available: https://eips.ethereum.org/EIPS/eip-721.

[21] M. Angelo and G. Salzer, "Tokens, types, and standards: Identification and utilization in Ethereum," in *Proc. IEEE Int. Conf. Decentralized Appl. Infrastructures (DAPPS)*, IEEE, 2020, pp. 1–10.

[22] B. S. Mondoh, S. M. Johnson, M. Green, and A. Georgopoulos, "NFT legal and regulatory compliance: Connoisseurship and critique." Nov. 2022. Accessed: Feb. 26, 2024. [Online]. Available: https://ssrn.com/abstract=4275613.

[23] H. Guo, M. Chen, and W. Ou, "A lightweight NFT auction protocol for cross-chain environment," in *Proc. Int. Conf. Mach. Learn. Cyber Secur.*, Springer, 2022, pp. 133–146.

[24] S. Rosen, "Hedonic prices and implicit markets: Product differentiation in pure competition," *J. Political Economy*, vol. 82, no. 1, pp. 34–55, 1974.

[25] L. Schaar and S. Kampakis, "Non-fungible tokens as an alternative investment: Evidence from cryptopunks," *J. Brit. Blockchain Assoc.*, vol. 5, pp. 1–12, Apr. 2022.

[26] P. Kireyev and R. Lin, "Infinite but rare: Valuation and pricing in marketplaces for blockchain-based nonfungible tokens," INSEAD Working Paper No. 2021/60/MKT, Oct. 2021. Accessed: Feb. 26, 2024. [Online]. Available: https://ssrn.com/abstract=3737514.

[27] R. Cellini, T. Cuccia, and J. Lyrvall, "On the price dynamics of non-fungible tokens: The "Bored Apes" case," *PLoS One*, vol. 18, no. 11, 2023, Art. no. e0287881.

[28] F. Horky, C. Rachel, and J. Fidrmuc, "Price dynamics of non-fungible tokens: The case of the digital arts market," *Electron. J.*, 2022, pp. 1556–5068.

[29] Y. Wang, "Anchoring effect and loss aversion: Evidence from the non-fungible token market," Jan. 2022. Accessed: Jan. 29, 2024. [Online]. Available: https://doi.org/10.2139/ssrn.4097185.

[30] K.-H. Ho, Y. Hou, T.-T. Chan, and H. Pan, "Analysis of non-fungible token pricing factors with machine learning," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 1161–1166.

[31] X. Zheng, "Multimodal learning for improved NFT price prediction," in *Proc. IEEE Int. Conf. e-Bus. Eng. (ICEBE)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 74–79.

[32] J. Branny, R. Dornberger, and T. Hanne, "Non-fungible token price prediction with multivariate LSTM neural networks," in *Proc. 9th Int. Conf. Soft Comput. Mach. Intell. (ISCMI)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 56–61.

[33] R. Ma, X. Yang, and F. Gao, "Non-fungible token forecast based on LSTM," in *Proc. 4th Int. Conf. Appl. Mach. Learn. (ICAML)*, Piscataway, NJ, USA: IEEE Press, 2022, pp. 81–84.

[34] D. Costa, L. Cava, and A. Tagarelli, "Show me your NFT and I tell you how it will perform: Multimodal representation learning for NFT selling price prediction," in *Proc. ACM Web Conf.*, 2023, pp. 1875–1885.

[35] I. T. Julianto, D. Kurniadi, and F. M. Khoiriyyah, "Price prediction of non-fungible tokens (NFTs) using data mining prediction algorithm," in *Proc. Int. Conf. Comput. Sci., Inf. Technol. Eng. (ICCoSITE)*, Piscataway, NJ, USA: IEEE Press, 2023, pp. 633–637.

[36] R. Kräussl and A. Tugnetti, "Non-fungible tokens (NFTs): A review of pricing determinants, applications and opportunities," *Journal of Economic Surveys*, vol. 38, no. 2, pp. 555–574, Apr. 2024.

[37] D. Diakoulaki, G. Mavrotas, and L. Papayannakis, "Determining objective weights in multiple criteria problems: The critic method," *Comput. & Operations Res.*, vol. 22, no. 7, pp. 763–770, 1995.

[38] J. F. Hair, "Multivariate data analysis," Pearson, 7th Ed., 2009.

[39] B. Guidi, A. Michienzi, and L. Ricci, "Data persistence in decentralized social applications: The IPFS approach," in *IEEE 18th Ann. Consumer. Commun. Network. Conf. (CCNC)*, 2021, pp. 1–4.

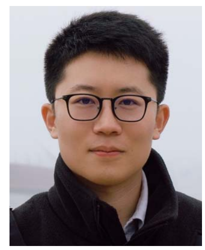
[40] S. M. H. Bamakan, A. Motavali, and A. B. Bondarti, "A survey of blockchain consensus algorithms performance evaluation criteria," *Expert Syst. Appl.*, vol. 154, no. 113385, 2020, Art. no. 113385.

Wenze Xiong received the M.Sc. degree in financial computing from the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China, in 2024. He is about to enroll in the Ph.D. degree in accounting with the University of Auckland, Auckland, New Zealand, with a full scholarship.

His research interests include blockchain, cryptographic assets, decentralized finance, and fintech.

Yetong Wang received the B.S. degree in information and computing science from Xi'an Jiaotong-Liverpool University, Suzhou, China, in 2023. He is currently working toward the M.Sc. degree in statistical data analysis with Australian National University, Canberra, Australia.

His research interests include blockchain, cryptography, and artificial intelligence.



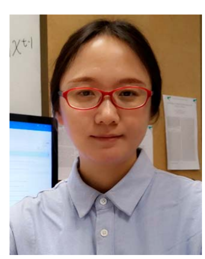
Wanxin Li (Member, IEEE) received the B.S. degree in computer science from Chongqing University, Chongqing, China, in 2015, and the M.S. degree in computer science, and the Ph.D. degree in transportation engineering from the University of Delaware, Newark, DE, USA, in 2017 and 2022, respectively.

He is an Assistant Professor with the Department of Communications and Networking, Xi'an Jiaotong-Liverpool University, Suzhou, China. His research interests include blockchain, cryptography, distributed AI, and smart transportation.

Dr. Li was a recipient of the IEEE Intelligent Transportation Systems Society Best Dissertation Award and the IEEE Technology and Engineering Management Society Outstanding Ph.D. Dissertation Award in 2022. He is a fellow of U.K. Higher Education Academy and a Member of ACM and IET.

Yutong Zhang received the M.Sc. degree in applied informatics from the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China, in 2024.

His research interests include blockchain, cryptographic assets, and decentralized application.



Jie Zhang (Member, IEEE) received the Ph.D. degree in computer science from the University of Liverpool, Liverpool, The United Kingdom, in 2018.

She is currently an Associate Professor with Xi'an Jiaotong-Liverpool University, Suzhou, China. Her research interests include applied cryptography, IoT security, and blockchain.

Hao Guo (Member, IEEE) received the B.S. degree from Northwest University, Xi'an, China, in 2012, and the M.S. degree from Illinois Institute of Technology, Chicago, IL, USA, in 2014, and the Ph.D. degree from the University of Delaware, Newark, DE, USA, in 2020, all in computer science.

Currently, he is an Assistant Professor with the School of Software, Northwestern Polytechnical University, Taicang, China. His research interests include blockchain and distributed ledger technology, privacy-preserving computation, applied cryptography, and Internet of Things (IoT) security.

Dr. Guo is a Member of ACM.