CCF Transactions on Pervasive Computing and Interaction
https://doi.org/10.1007/s42486-022-00094-6

REGULAR PAPER

Portrait of decentralized application users: an overview based on large-scale Ethereum data

Tian Min[1] · Wei Cai[1]

Received: 19 July 2021 / Accepted: 24 January 2022

Abstract
Decentralized application (DApp) is an emerging technology designed to address distrust, privacy, and security issues. However, we note that the research community in human factors has not conducted in-depth research on user behavior in this unique distributed environment yet. In this paper, unlike a small sample of user interviews, we attempt to profile DApp users through publicly available data. Using Ethereum as an example, we build a series of datasets containing more than 73.8 million transactions generated by 230,000 addresses. By transforming hexadecimal addresses into readable application names, we analyze the behavioral characteristics of the user based on the categories of DApp. Furthermore, we apply an unsupervised clustering method on the 230,000 addresses to distinguish investors and players and analyze their behavioral patterns and sensitivity to blockchain markets, such as ETH prices. In addition, we implement heuristics to demonstrate how blockchain data mining can facilitate practical systems, including anomaly detection and recommend systems. Finally, we discuss future directions for studying human factors in a decentralized context and hope that this work will attract more research attention and support the development of DApp and further Metaverse ecosystems.

Keywords Human factors · Blockchain · Decentralized applications · User behavior

1 Introduction

A contemporary blockchain can be introduced as a combination of three sophisticated technologies (Elsden et al. 2018): **distributed ledger**, which is a public database managed by peer to peer (p2p) network, recording transaction history among the nodes in the form of a series of data blocks linked together by hash; **consensus model**, a set of rules that restricts the permission of modification unless the majority of the network has verified the operations; and **smart contracts**, the programs that can be automatically executed without any centralized control. With the support of these fundamental technologies, decentralized applications (DApps) have developed rapidly on various blockchain platforms. There have been thousands of kinds of DApps that cover the categories as wide as the traditional applications (Apps).

Studying blockchain users can provide valuable insights into human behavior in the p2p environment, provide guidance for decentralized system design, and develop governance standards for this emerging community. However, conducting crowd-sourced surveys is difficult and expensive due to the limited number of blockchain users and privacy concerns, as opportunities for social engineering risks arise once blockchain addresses are associated with the IP addresses and surveyed information. Also, to ensure that the respondents are blockchain users, the reward needs to be in the form of virtual currency, which can cost hundreds of dollars per transaction for a commission depending on the network situation. Therefore, we believe that it is wise to extract as much information as possible from *publicly available open-source data* before surveys and interviews and that the information obtained can help adjust survey strategy and questionnaire design, which will further reduce the opportunity cost of interview research. However, data mining on the blockchain operates in a black-box mode: the ability of supervised learning is quite limited due to the lack of accurate training and testing sets. As Meiklejohn et al.

✉ Wei Cai
caiwei@cuhk.edu.cn

Tian Min
tianmin@link.cuhk.edu.cn

1 The Chinese University of Hong Kong, Shenzhen, China

Published online: 15 February 2022

(2013) pointed out in their paper on Bitcoin re-identification: although we can examine the current gap between actual and potential anonymity, clustering heuristic is not fully robust in the face of changing behavior. As a result, our paper will focus more on the analysis of visual data, statistical metrics, and unsupervised learning. We suggest that using multiple methods to verify each other so that we could approximate the truth a step further. The main contributions of this work are presented as follows:

- We collect more than 73.8 million lines of records generated by 235,420 addresses from Ethereum, arrange it into a user profile dataset through deriving extra attributes to enrich the data dimension.
- We visualize and analyze the user profile dataset from the aspects including *Transaction Number*, *DApp Categories*, *First Interaction* and *Transaction Value*. We discuss the motivation and values of DApp users combing the real-world data and theoretical deduction.
- We suggest an unsupervised Self-Organizing Map (SOM) based classification of addresses to distinguish blockchain investors and players. We compare behavioral differences between two groups and their reactions towards market fluctuations, for example, token price. We find that over 50% of address in the Ethereum sample are static. The remaining active addresses have about ten times as many investors as players as their primary role. Among them, players are more significantly affected by ETH prices and fees, while investors are willing to tolerate increased commission to continue trading. We believe that the large presence of speculators will crowd out space and opportunities for other users from the aspect of congested network.
- We implement heuristic attempts including abnormality detection and recommendation using the datasets. We demonstrate how data-driven methodologies and our datasets can assist user behavior study on blockchain and the development of practical systems.
- We discuss the future directions of the human–computer interaction (HCI) study on DApps. We emphasize that to push forward the understanding of blockchain users, researchers may combine the data approach and user investigation including surveys and interviews.

2 Background

2.1 Ethereum

Ethereum is a blockchain platform that builds on Bitcoin's innovation, but with huge differences. It aims to provide the end-developer a tightly integrated end-to-end system for building software, the DApps, on a hitherto unexplored compute paradigm in the mainstream: a trustful object messaging compute framework (Wood 2020). Like Bitcoin, Ethereum uses Proof-of-Work (Dwork and Naor 1992) as a consensus model, which is responsible for regulating the p2p network and rewarding nodes that contributed to the community. Mining is a basic and typical kind of contribution: transactions on the network need to be confirmed before being packaged into a new block. The packing efficiency will affect the transaction delay, thus having an impact on the quality of experience (QoE). The interval between two blocks on Ethereum is about 10–20 s (Gervais et al. 2016). With the development of technology, this interval will continue to be compressed, so that users can have experience close to centralized Apps.

In order to avoid network abuse and to sidestep the inevitable questions stemming from Turing completeness, all programmable computation in Ethereum is subject to fees (Wood 2020). From the users' aspect, *gas* is a bounty system: the more *gas* put in a transaction, the more likely it is to be packed by miners first. This value fluctuates with the network conditions. Even though it can be a small amount, it lowers the accessibility of DApp, which leads to the loss of users who do not want to take the trouble.

2.2 DApp

Decentralized application (DApp) is a kind of application that hosts parts of their back-end services and databases on p2p networks, as opposed to typical applications supported by centralized servers. As Cai et al. (2018) summarized, a DApp is expected to have the properties of:

- **Open source:** DApp is expected to make their codes open source so that audits from third parties become possible. In reality, however, more DApps tend to only expose their application binary interface (ABI) (Ethereum 2020a) for security reasons and indicating that the source code has not been modified.
- **Internal cryptocurrency support:** Internal currency is the vehicle that runs the ecosystem for a particular DApp. With tokens, it is feasible for a DApp to quantify all credits and transactions among participants of the system, including content providers and consumers.
- **No central point of failure:** A fully decentralized system should have no central point of failure since all components of the applications will be hosted and executed in the blockchain.

Figure 1 shows the structure of a typical web-based DApp. It can be seen as a traditional application architecture with distributed service and databases as extensions. A DApp user will register a virtual wallet on the corresponding blockchain platform as a unique identity. Critical operations

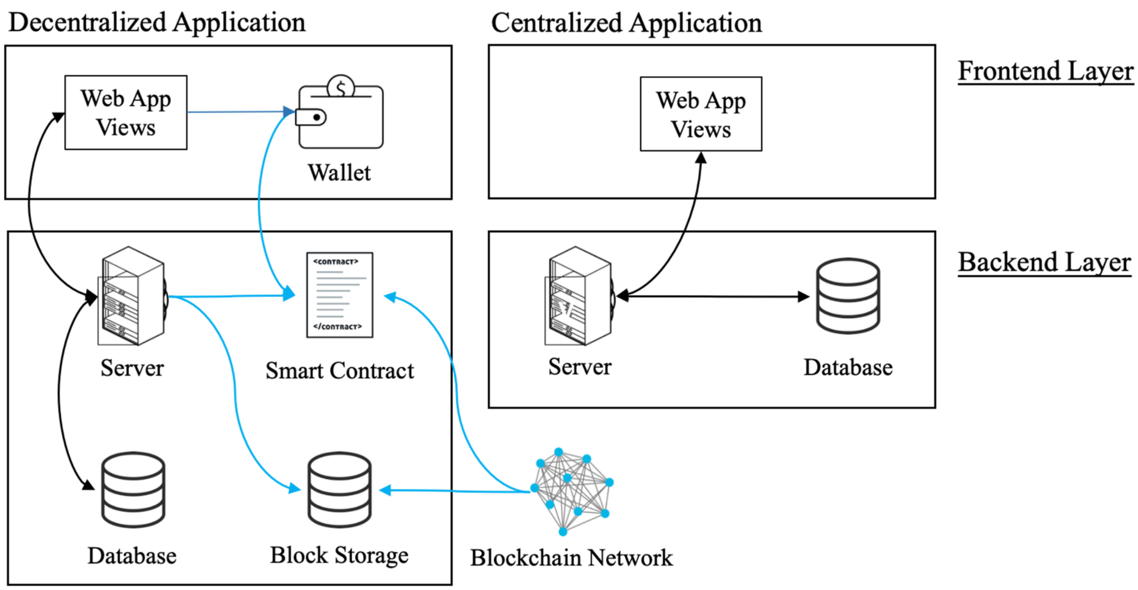

Fig. 1 Structure of a typical web-based DApp and centralized App

including purchase, sale, or random number generation will be performed by smart contracts. Important information, such as in-app deposits, character information, will be preserved in blockchain storage. Traditional servers will be responsible for providing in-app services such as rendering interfaces or handling requests. Moreover, it acts as messengers between users and Ethereum through Web3.js (Ethereum 2020b). Traditional databases will serve as repositories for application resources and other secondary data.

2.3 DApp category

Previous research Werbach (2017) proposed a blockchain application typology of finance, proof-as-a-service, and DApps. However, among the DApps, there has not been an exact agreement on how to classify them. We reviewed DApp indexing sites including DAppRadar,[1] DAppTotal[2] and StateoftheDApps,[3] finding that there is a strong consistency in highly differentiated categories such as games, casinos, and exchanges, while classification standard diverse when it comes to ambiguous or niche categories such as utilities and social media. In this paper, we consolidate and summarize a classification method with fewer categories and greater generality. We list the following nine categories.

– *Game*: blockchain games combine blockchain technology with online games so that virtual assets such as props and coins are possibly be turned into cash, or transferred to other games. With the p2p network, user-generated content (UGC) is generally involved.
– Decentralized Finance (*DeFi*): *DeFi* is a decentralized version of traditional finance services including banking and crowdfunding, which minimizes the need for trust and central authorities.
– *Gambling* This category consists of various types of virtual casinos. It utilizes the transparency characteristics of blockchain to ensure fairness.
– *Exchange Exchange* is where blockchain users buy and sell their various kinds of currencies: cryptocurrencies from different platforms or tokens issued by different organizations.
– *Collectible* In *Collectible* DApps, users can trade virtual collections. Rare items tend to sell at high prices. Most of these DApps use UGC to increase the richness of merchandise.
– *Marketplace Marketplace*, the blockchain version of Amazon or eBay, is where people trade or auction for virtual items including game props, collections, and so on.
– *Social Social* DApps are decentralized social networking services (SNS) including messaging, blogs, and file sharing. The purpose of these DApps is to create social networks in which the content will not be censored or manipulated.
– *High-risk High-risk* category is consist of speculation DApps, which are driven by variants of pyramid scheme.
– *Other*: miscellaneous DApps including charity, development tools, storage service, and so on.

1 https://dappradar.com/.
2 https://dapptotal.com/.
3 https://www.stateofthedapps.com/.

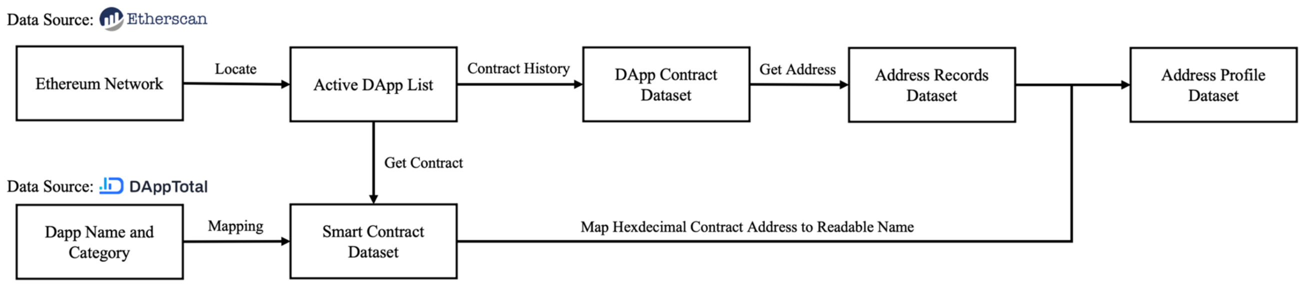

Fig. 2 Process of data collection

3 Related work

3.1 Blockchain and human factors

There have been a few, but emerging discussions on the role of the human factor research community in linking the design and application of blockchain technology towards lived experience and the articulation of human values (Elsden et al. 2018). These discussions are mostly based on the classic topics of human factors with finance, community engagement, and p2p environment (Bellotti et al. 2014; Carroll and Bellotti 2015; Gao et al. 2016; Presthus et al. 2017; Prinz 2018). They made great contributions in revealing the perceptions and motivations of blockchain users. Today, blockchain has transformed from the 1.0 era of Bitcoin into the 2.0 era (Cai et al. 2018): the era of decentralized applications (DApps). The new features and functionalities that DApp brought will make a huge difference to users' perceptions and behavior. Elsden et al. (2018)'s work set up the conceptual and methodological guidelines for human factor study on the blockchain.

3.2 Explore blockchain data

Previous efforts (Meiklejohn et al. 2013; Ranshous et al. 2017; Chan and Olmsted 2017; Jourdan et al. 2018; Chen et al. 2018) focus on mining blockchain data by proposing different schemes for graph analysis, trying to understand whether de-anonymization is feasible to quantify the promise of anonymity. Meiklejohn et al. (2013) suggested a heuristic clustering to group Bitcoin addresses based on evidence of shared authority. Due to the existence of mixing service, many efforts like studies conveyed by Biryukov et al. (2014) and Ranshous et al. (2017) contributed to the improvement of transaction graph analysis. Chen et al. (2018) extended the graph study under the Ethereum context. Their specific work challenged those who use Bitcoin for criminal or euro purposes (Meiklejohn et al. 2013) and made great contributions to the security and robustness of infrastructure. However, the user behavior under the blockchain context has rarely been studied from open-source data directly. We need more generalized investigations to introduce the blockchain to researchers with various backgrounds. Based on their methodologies, our work stands on transaction records but concentrates on exploring the transaction itself and depicting an overview picture.

4 Dataset

In this section, we introduce our data collection strategy and the composition of our datasets. We use Etherscan[4] to access Ethereum transaction logs. Figure 2 shows our process of data collection. We take all addresses that interacted with top-ranked DApp, as a snapshot on May 1, 2021, and arrange them into datasets including *Smart Contract Dataset*, *Address Records Dataset*, and *Address Profile Dataset*. Following Ethereum's open-source license, the *Address Profile Dataset* will be available on Github.[5] There will be no privacy issues with the datasets used in this paper, and all data is sourced from the public on-chain records of the Ethereum mainnet. All user addresses will remain in the form of hexadecimal hashes for analysis. The research in this paper will not attempt to perform re-identification, compare transaction amounts to real-world currency, or any operation with social engineering implications. Any organization or individual using the same collection and processing methods is expected to obtain identical results.

4.1 Smart contract dataset

We collect 10,936 smart contracts belonging to 1903 DApps and tagged the classes to which the DApps belonged, creating an index we call *Smart Contract Dataset*. Each row of the record represents information describing a smart contract. The attributes contain the hexadecimal contract address, the DApp name, the DApp author, and the number of contracts contained in the DApp. DApps belonging to the

4 https://etherscan.io/.

5 https://github.com/Gullintani/BlockchainDataAnalysis.

Table 1 Attributes in *Address Records Dataset*

Attributes	Datatype	Description
Blocknumber	Integer	Height of Ethereum blockchain
Hash	Hex	Unique hash value for each transaction
Timestamp	Integer	Time when the transaction made
From	String	Address where the transaction from
To	String	Address where the transaction to
Value	Integer	Number of value this transaction carries in GWei
Gas	Integer	Number of gas in this transaction

we collect all the users' transaction records, and form them into a *Address Records Dataset*.

4.3 Address profile dataset

The *Address Profile Dataset* contains information extracted from address transaction records. We calculate statistical indicators including the mean, median, and standard deviation of the values (number of transactions, transaction value, gas). In addition, we recorded time-related data, including the frequency of transactions and the dates of the first and last interactions. We use the dictionary type attribute to store

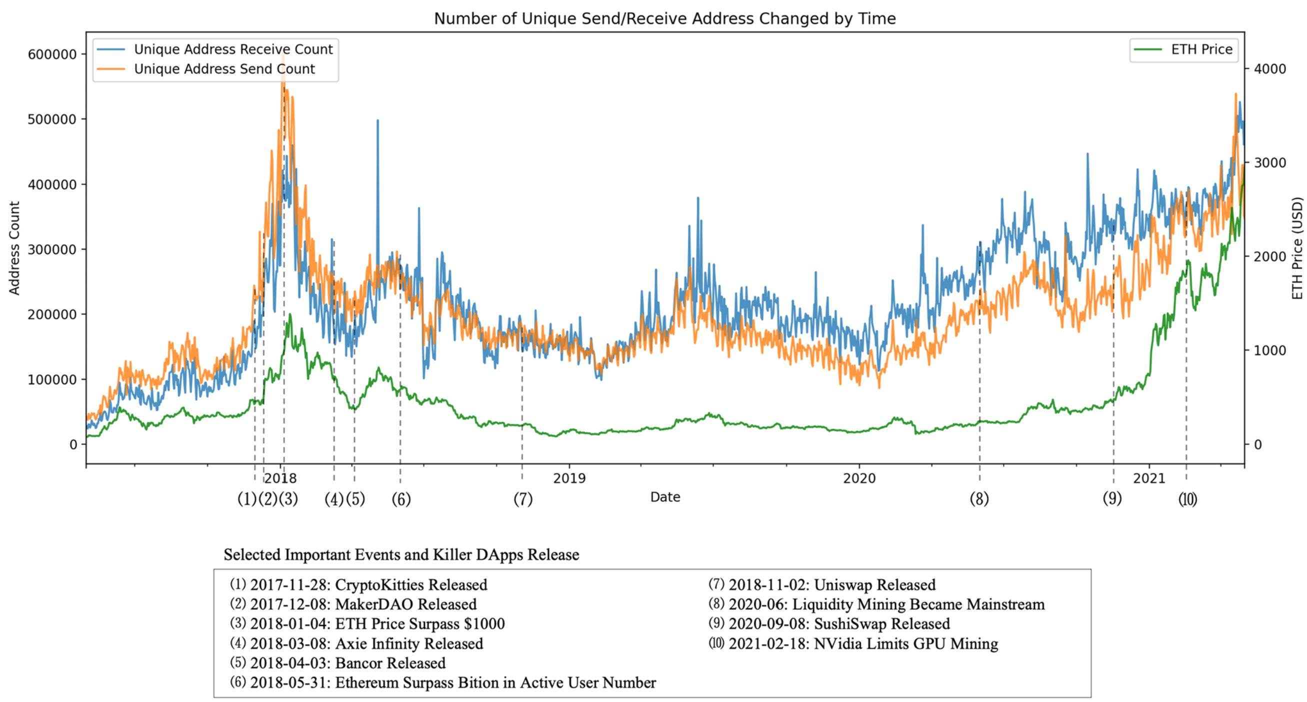

Fig. 3 Unique send/receive address verses ethereum price with important events marked

Game category make up the largest majority of contracts. There are 1229 DApps run by just one contract, while the rest all have more than 5. This dataset will be used as a reference when collecting active users, and as a mapping dictionary when interpreting the transaction logs (Table 1).

4.2 Address records dataset

From the statistics given by DAppTotal (2020), for all blockchain platforms, only 220 out of 2670 DApps are active. Given this situation, we select the top 10 popular DApps in each category, from which we filtered out 3,848 contracts in *Smart Contract Dataset*. We gather the latest transaction records through these contracts and arranged them as a 235,420 in-length list of user address index, through which

the number of times a user interacted with a particular DApp and category. This data will help us visualize and describe the user profiles.

5 Analysis

5.1 Overall analysis

5.1.1 Transaction

Figure 3 is a double y-axis time series, with the left y-axis showing the number of unique addresses sending or receiving transactions per day on Ethereum, while the right y-axis

Table 2 Interaction frequency and number of address

Category	Frequency	Frequency %	*i>1* address	*i>10* address	*i>20* address
Game	7,215,050	10.80%	45,834	28,214	22,342
Exchange	19,297,395	28.89%	129,535	96,271	80,934
DeFi	1,572,319	2.25%	61,069	22,298	13,904
Other	839,638	1.26%	25,171	6301	3781
Gamble	812,955	1.22%	6378	3141	2416
Marketplace	927,102	1.39%	24,996	9829	6397
High-risk	594,833	0.89%	13,269	6592	4683
Collectible	171,717	0.26%	3660	1476	999
Social	124,897	0.19%	5910	1,507	830
Unknown	35,242,500	52.76%	210,614	154,195	129,195

shows the price in the USD curve for ETH. In addition, we have selected several important events about Ethereum or the release dates of popular applications, marked on the curve with dashed lines and numbers. As a whole, we find that unique sending addresses are generally larger than the receiving ones until the beginning of 2019, while after that, the number of receiving addresses is larger than the number of sending addresses. This can indicate to some extent that the number of contract addresses serving customers, and the number of complex operations where contracts are invoked by contracts are on the rise.

Compared to the real-world stock market or other markets, ETH prices are more susceptible to news, traffic, and media due to the relatively small number of participants and lack of liquidity. The most obvious examples are as the game, CryptoKitties, released on November 28, 2017, which brought explosive user growth and sent the price of ETH soaring at the end of the year; or such as NVIDIA's announcement on February 18, 2021, which stated that it will limit its GPU performance support for mining. The news dropped the price of ETH by more than $500 once it was released. Furthermore, by observing the three curves from 2018 to 2021, we can see that microscopically, the number of active addresses increases with the ETH price, and in some localized fluctuations, the abnormal number of trading addresses on a given day is often accompanied by a slight increase in the ETH price. It is difficult to quantitatively assess the impact of events on user activity or ETH prices, but one can be sure that the impact must exist to a greater or lesser extent. We can conclude that the participants in such a market are mostly motivated by monetary gains and the expectation of short-term returns in this highly fluctuating market.

In our sample, there were 73,801,904 transactions generated by 235,420 addresses, which indicates that on average, 313.49 transactions were made per address. As expected, the distribution of the number of transactions has a long tail: the first, second, and third quartiles are 15, 71, and 266, respectively, suggesting a large standard deviation of 1159.63. Only 6.69% of the addresses have more than 1000 transactions. These results reveal the fact that behind the surge of unique addresses on Ethereum (Etherscan 2020b), the vast majority are inactive. The ratio of addresses sending to receiving transactions is about 12 to 1. Comparing this with the ratio of the overall number of addresses sending and receiving in Fig. 3, we find a much more disparate ratio of individuals in the sample, which also indicates that as far as individual addresses are concerned, more of their activity is focused on actively invoking contracts or sending transfers rather than passively receiving them.

5.1.2 Categories and DApp

To explore the pattern of transactions to different categories, we calculated their frequency and distribution of address number under the condition of 'interacted with category more than *n* times (denoted as *i(cate)>n*)' in Table 2.

The second and third columns show the number of transactions to different categories. *Unknown* stands for the personal wallet or contract addresses that cannot be mapped to any category. *Exchange* holds the largest proportion, which is twice as much as the *Game* category; *Gamble*, as one of the earliest kinds of DApp on the blockchain, does not account for much proportion of transaction as we expected. The fourth to seventh columns of Table 2 represent the number of addresses that at least interacted with the corresponding category for 0, 10, and 20 times. The results show that the number of addresses falls sharply as the threshold increases: from 0 to 20, each category falls 75.73% on average. Only 49,281 addresses (26.81%) interact with at least one of the categories more than 20 times.

In this case, we came up with a research question: *Are there differences in category preference between the active and static address?* We use *transaction = 10* as a boundary to query two groups of addresses. The results show that for addresses with number of transaction smaller or equal to 10, there are 43.03% transactions were sent to unknown

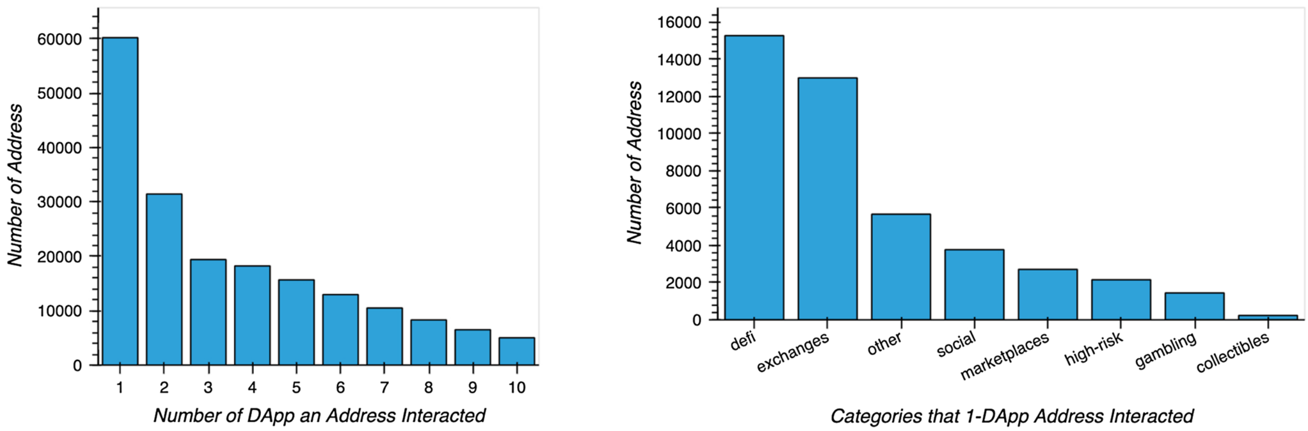

Fig. 4 Distribution of address by the number of different DApps used (left); distribution of categories chose as the only DApp

wallet addresses, 35.16% were received from unknown wallet addresses, 6.49% sent to *Exchange*, 5.75% sent to *DeFi*, 3.28% sent to *Game*, and rest of them are shared by other categories; for the addresses with number of transaction greater than 10, 44.5% transactions sent to unknown wallet, *Game* took 21.61%, *Exchange* took 10.28%, while *DeFi* only took 2.94%. The conclusion is the disparity in transaction number, may account for the difference in interaction times towards different categories. The number of transactions can distinguish the usage of the addresses to some extent.

From the above data observations, we can see that the exchanges, which are a prerequisite for a smooth DApp experience, have the largest share of transactions in the sample. When we use the number of interactions as a filter to increase the volume requirement, the game category has the least percentage of decline. This can indicate that blockchain games can attract users for a longer period of time, which is related to the fact that the nature of the game itself is entertaining and requires long-term commitment, unlike the properties of tools like exchanges. Overall, the vast majority of addresses have a short-term or superficial experience with the app, with few long-term users.

Diversity in DApp usage can be reflected by the number of different DApps an address has ever interacted with. The histogram on the left side of Fig. 4 shows the distribution of the number of unique DApps that addresses have interacted with, with each address having interacted with 3.90 DApps on average (median = 2, SD = 4.64). In the sample, there are 60209 addresses that use only one DApp. We suspect that these addresses may be bonus hunters, light users trying new things, or affiliate addresses used for purposes such as testing. In the case of Uniswap, for example, when it opened its governance token distribution in September 2020, it gave a significant number of tokens (Coindesk 2020) to addresses that had provided liquidity for it. Similar to the case of offering dividends to the old users are commonly seen on Ethereum, and predicting the development trend of certain DApps, participating in them in advance is also a way to invest with substantial returns. This conjecture is verified by the histogram on the right side of Fig. 4, which represents the categories distribution to which those DApps that are used as unique belong. DApps with obvious financial attributes such as DeFi and exchanges make up the majority. We can conclude from this distribution that the diversity of Ethereum DApp usage stays low after two years of development. For traditional application services, games, video, reading, development, or online shopping, the abundance of applications in every category on the Internet now far exceeds that of blockchain, and we will even use several applications or digital games in one category. DApp market might be far less in breadth and competitiveness than the traditional market in the current stage.

5.1.3 First interaction

The first impression is considered to be of great significance in user experience study (Thielsch et al. 2013). Under the blockchain scenario, we can identify the common introductory DApp so that the developers can be targeted to provide extra guidelines; or we can identify which DApp rise the number of registration or evaluate the performance of the advertisement. Figure 5 shows the time of the first transaction of 235,420 samples. The height of the histogram represents the number of addresses that made their first transaction in the specific month. The earliest address was registered in August 2015 while the latest was created in April 2021. The first interaction time of the samples was basically in line with the unique address curve provided by Etherscan (2020b), which showed a sharp increase after mid-2017 and late 2020.

Figure 6a shows the cumulative frequency of the number of first interactions with the top five categories. The

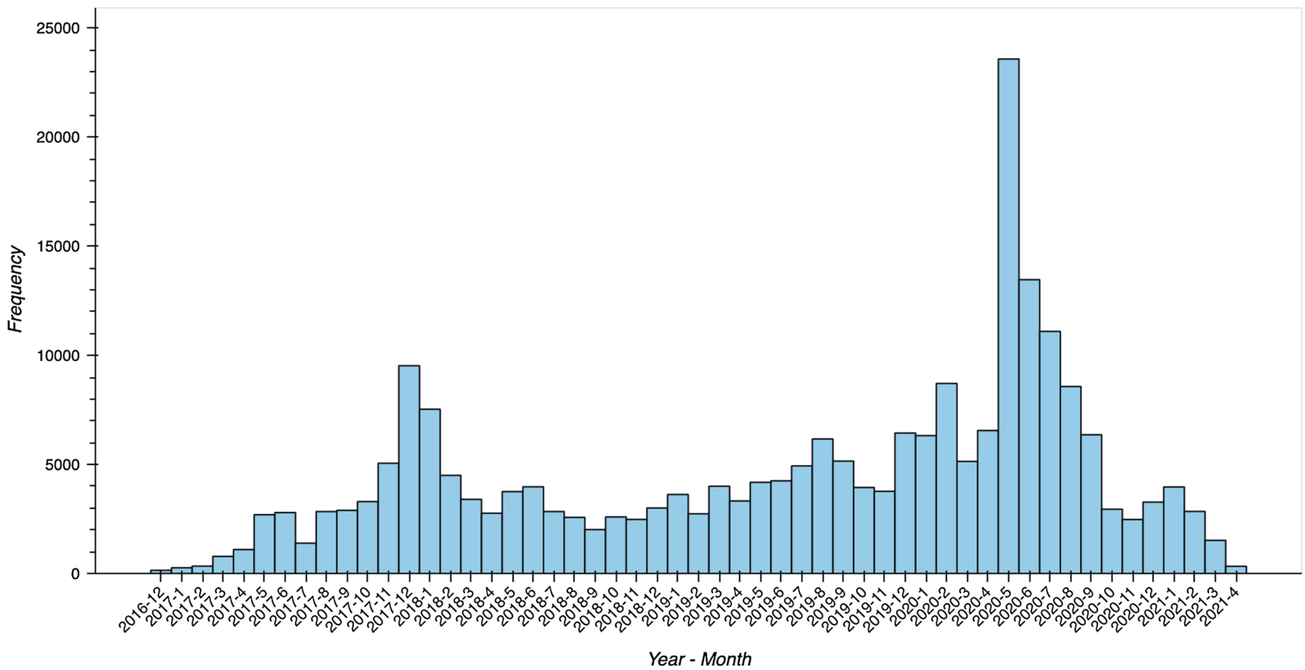

Fig. 5 Time of first interaction

remaining categories, including *Gamble*, *Social*, *Marketplace*, and *Collectible*, are not shown in the figure because the sum of these four categories only accounts for 8% of the overall. In December 2017, we can find a spike in the number of first interactions caused by CryptoKitties[6] which indicated a similar pattern in *Game*. This surge also led to a domino effect, attracting newcomers from other categories. Another high point occurred in May 2020, which may be related to the release of the second version of the famous exchange Uniswap; the price of Ether also reached its highest point ever during the same period (Etherscan 2020a). Figure 6b shows the DApps selected for the most first interactions, with these eight applications accounting for 43% of all DApps. Linking the peak of the curve in Fig. 6b to the trend in Fig. 6a, we can identify the dominant DApps, such as *CryptoKitties (game)*, *My_Crypto_Heroes (game)*, *MakerDAO (DeFi)* and *Uniswap (exchange)*, which have a strong presence in their categories and are attracting new users.

5.2 Group analysis

5.2.1 Address group observation

In the previous section, we have discussed an overview of blockchain data in terms of categories and DApps. In this section, we attempt to look at the DApp category interaction of addresses, explore initial groupings, and identify behavioral differences and overlaps between the groups for a preliminary probe into the parameter setting for unsupervised learning that follows. We grouped the addresses into three groups according to the use of DApps. *Entertainment*, *Finance* and *Utility*. As a first step, we need to initially label the users with the category they interact with and observe the distribution. Based on the number of addresses in Table 2, we decide to use *at least 30 interactions with a category* as the threshold for labelling users. We filter users by removing addresses with a total number of transactions of less than 100. There is no 'reasonable' value for these requirements but rather depends on the purpose of future researchers who may use our dataset or similar methods.

We use *i(cate)>30* and *transaction>100* to filter out a user group for each category. For each group, we calculate the transaction percentage and the average transaction count for each category. A heatmap is used to visualize the results as shown in Fig. 7: x-axis represents the user group; y-axis represent categories. Figure 7a shows the transaction compositions of each group in percentage. Figure 7b shows the average transaction counts of each group in absolute value. The darker the color in a block, the more frequently the group interacts with the corresponding category. By looking into the distribution of colored blocks, we found certain correlations between groups. For example, the *Gamble* column, its users interact almost the same frequency between *Gamble* and *Game* categories.

[6] https://www.cryptokitties.co/.

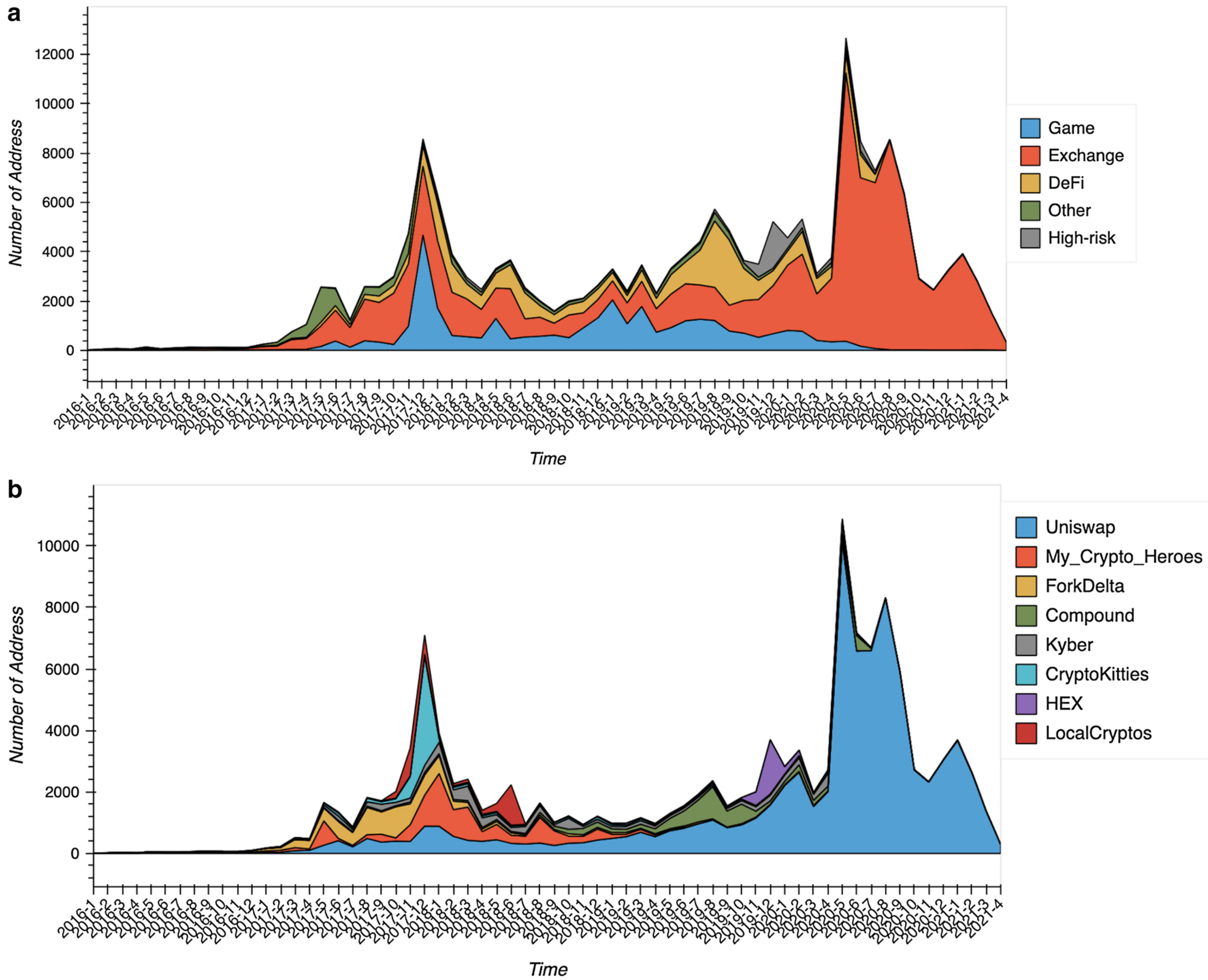

Fig. 6 Frequency of first interaction towards categories (**a**) and DApps (**b**)

However, there are not many transactions into other categories as shown in the *Game* column, which reveals a one-way association: most players are focused on game; a small number of them exist in both *Game* and *Gamble* groups. Due to the difference in size between two groups (*Game* = 13,604, *Gamble* = 1511), there is a significant difference in percentage. Given this situation, we took into account the number of attributes and users within the groups and referred to the typology by function proposed in Elsden et al. (2018). In analyzing the behavioral patterns of the groups, we rearrange the nine categories into three larger groups for better integration. On this expectation, we plan to focus on differentiating between Game and Finance attributes in unsupervised learning and determine the number of clusters in multiples of three. The expected cluster grouping will be based on this DApp classification basis:

– **Entertainment**: includes *Game*, *Gamble* and *Collectible*. The DApps in these categories are designed for entertainment purposes;
– **Finance**: includes *DeFi*, *High-risk* and *Exchange*, which have a clear investment or speculative purposes;
– **Utility**: includes *Marketplace*, *Social*, *Other*. This is the niche groups that DApps mainly act as the role of proof as a service.

5.2.2 Clustering

SOM is an unsupervised artificial neural network (Kohonen 1990). Unlike general neural networks that are trained based on the backward transfer of loss functions, it uses a competitive learning strategy that relies on neurons competing with each other to gradually optimize the network, which also uses the nearest neighbor function to maintain the topology of the input space. As shown in Fig. 8, we normalize

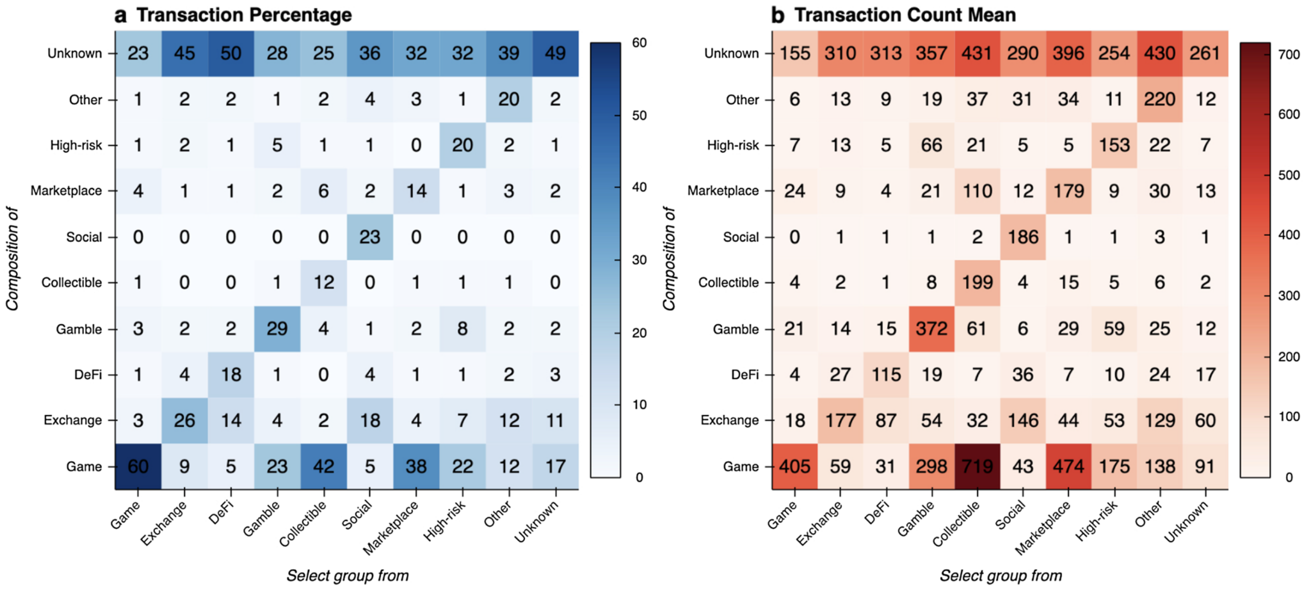

Fig. 7 Interaction distribution for ten categories (unknown included)

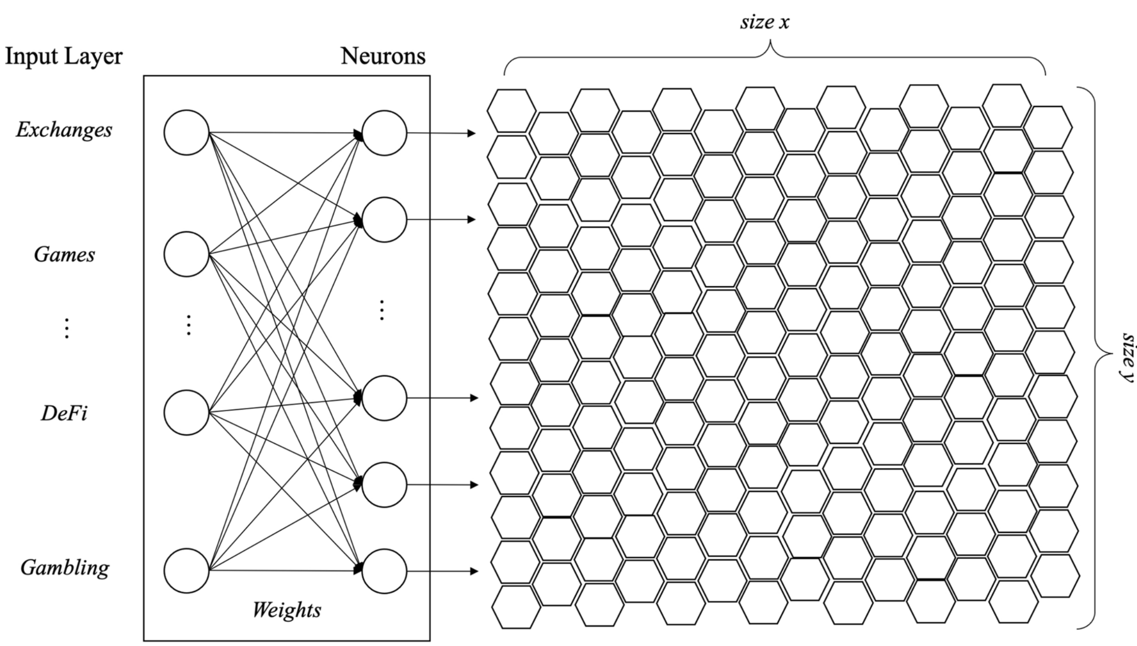

Fig. 8 SOM's architecture applied to unsupervised address clustering

the number of interactions between addresses and the nine DApp categories into a matrix as SOM input. We construct and traverse over 500 *sigma*, a parameter that controls the competitiveness of the nodes, and finally selected the clustering results at *sigma* equal to 1.4796 by adjusting the iteration and learning ratio.

After removing the 132,837 almost static addresses with less than 10 transactions, we performed SOM-based clustering on the remaining 102,583 active addresses and obtained four groups of addresses based on the classification criteria of category and frequency. We named them as *Finance High Frequency*, *Finance Low Frequency*, *Entertainment High Frequency*, and *Entertainment High Frequency*, and *Entertainment Low Frequency*, which have 5875, 85,814, 1298 and 9596 addresses respectively. We created box plots for four groups shown in Fig. 9, using

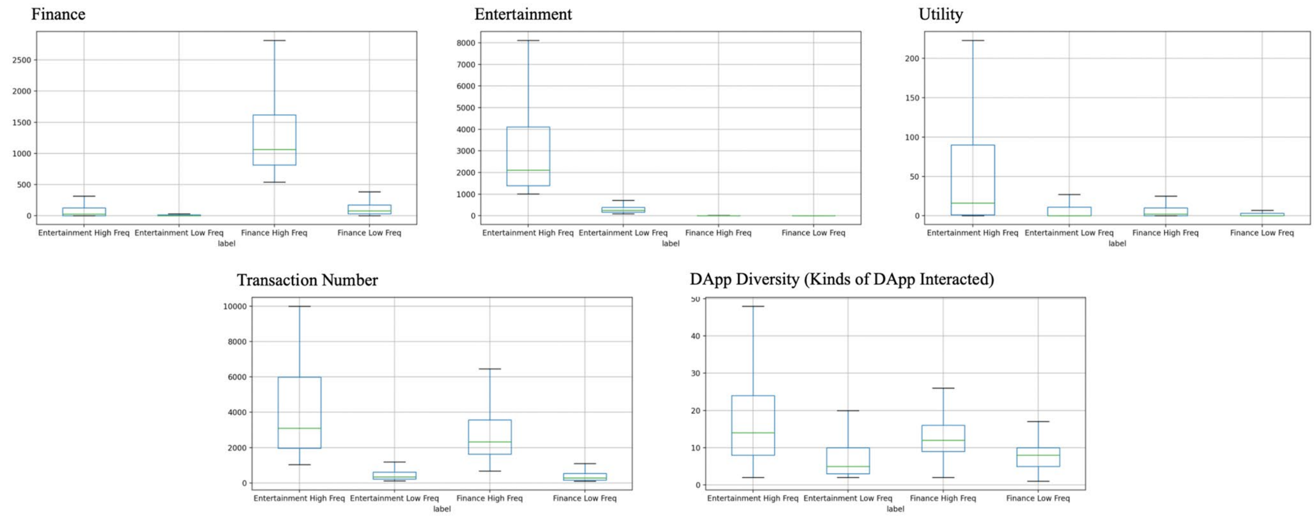

Fig. 9 Box plots for groups

the number of transactions, interacted app diversity, and the three categories we integrated into the previous section as indicators to see how well the groups are differentiated.

From the top three charts in Fig. 9, we can find that both the high-frequency groups of finance and entertainment have significantly higher activity than the other groups in the box plot of the corresponding indicators. And in Utility, containing Marketplace, Social, and Other DApp, the *Entertainment High Frequency* group occupies the largest majority, which should be explained by the fact that game or collectible DApp players need to rely on Marketplace to find the props they want, or to sell the NFT items they acquire in the game. In the bottom two graphs, the left graph shows the distribution of the transactions number, while the right graph shows the distribution of the number of DApp used. Of course, because they have more transactions, thus are marked as *High Frequency* addresses, but combined with the chart on the right we can find that the richness of interacted DApps for *High Frequency* addresses is much higher, in which the DApp diversity for *Entertainment* group is higher than that of *Finance*. From the aspect of addresses number in groups, *Finance*is about 10 times in quantity of *Entertainment*, which also proves from the side that most DApp users are sensible, mainly for profit as the purpose, rather than as leisure.

We generated word clouds for the four groups in the Fig. 10. Among them, purple and dark blue represent DeFi and exchanges, sky blue represents games, orange represents marketplace, and dark blue represents DeFi. Each word represents DApp names or contract names, and the text size reflects how often this DApp appears in the corresponding group. We can find differences in the diversity of DApps across the groups: there are significantly more DApps in the *Entertainment* groups than in the *Finance* groups, which can be explained by the fact that entertainment software is a creativity-centric business, while the financial business is more strictly delineated: borrowing, exchanging or managing money, and users will tend to rely on more stable and established providers. From the aspect of clustering results, in the *Finance High Frequency* group, exchanges such as Uniswap occupy too much weight, so that the larger variance leads to smaller word fonts for other DApps, while in the *Finance Low-Frequency group*, the variance in usage frequency is smaller. From the two groups of Entertainment, the variance of its High-Frequency group is smaller than that of the Low-Frequency group, which can indicate that the enthusiastic players are more willing to widely explore. In addition, among the four groups, *Entertainment High Frequency* has the most gambling DApps, where the reason behind this can be inferred from the number of frequent bets.

For the two high-frequency subgroups, we counted the activities of addresses within the groups and summed the daily interactions to the three categories in chronological order, which are presented in Fig. 11. From the distribution of curves, we can tell the result retains sufficient specificity. The variance of the Finance group is greater than that of the Entertainment group. Furthermore, comparing the red curve in each sub-graph, the ETH price reached a peak in early 2018, and at the same time, the corresponding Finance and Entertainment interactions of both groups reached a considerable high spike. Between early 2018 and early 2021, the ETH price has been fluctuating in a small range below \$1000. During this period, the number of trades in the *Finance High Frequency* group in the sample did not significantly trend in a certain direction but fluctuated around an average of 3,000 trades per day. It is not until January 2020 that its trading frequency shows a ragged upward trend, which

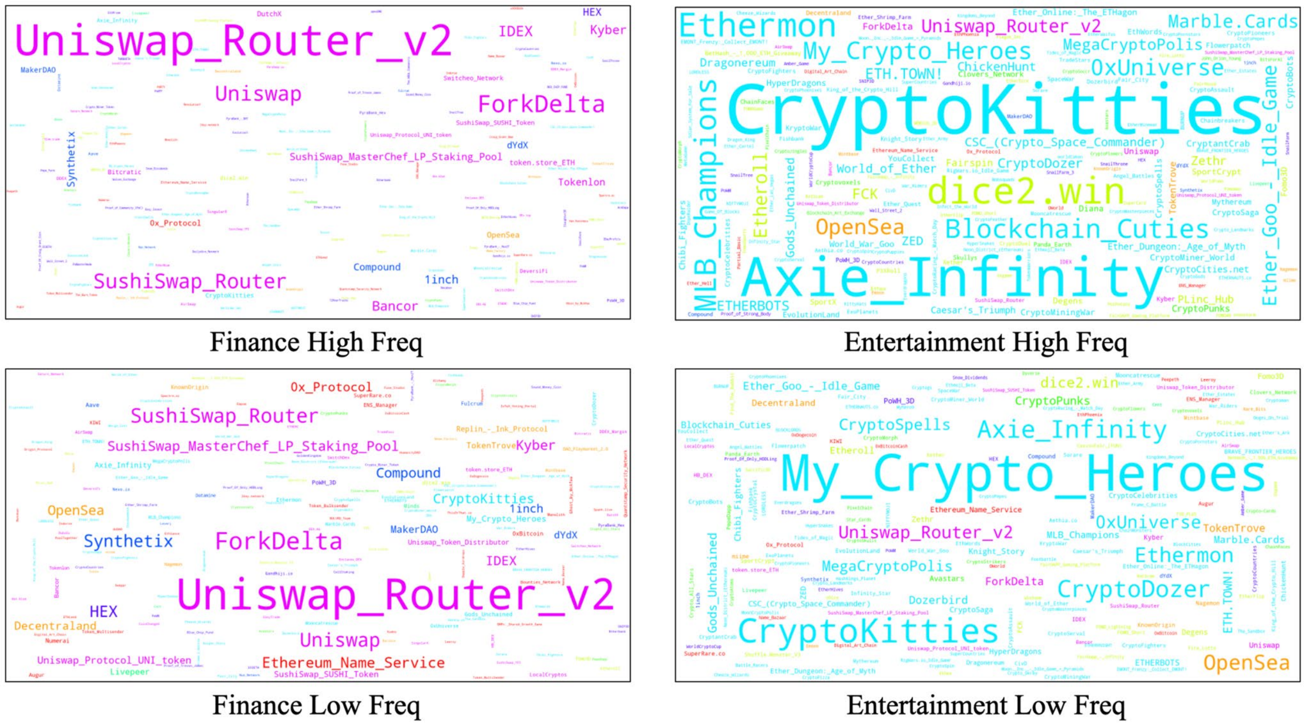

Fig. 10 DApp wordclouds generated by four groups

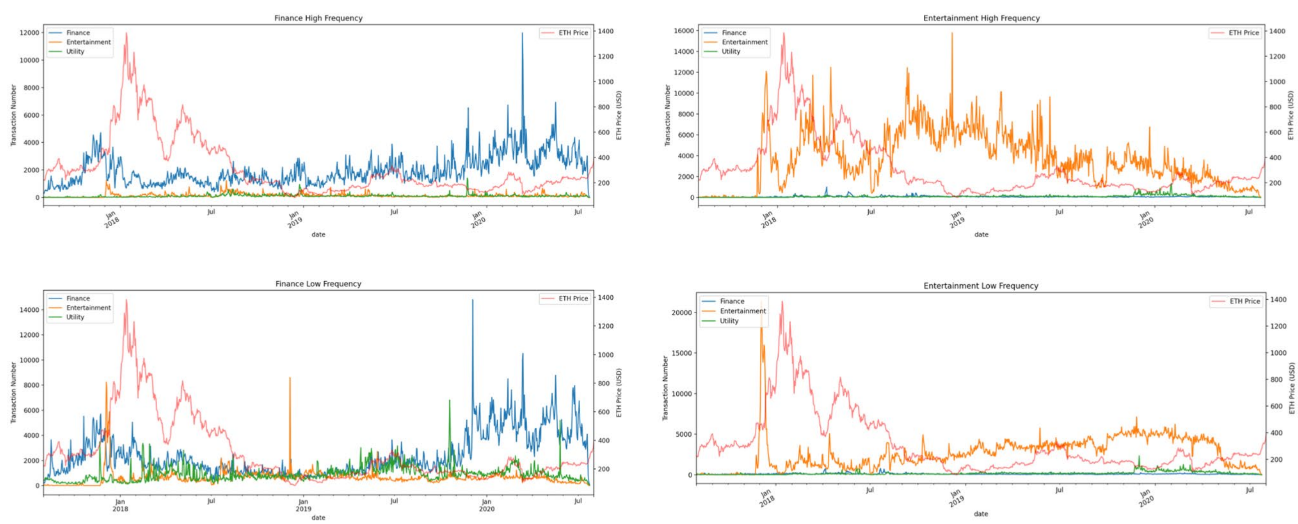

Fig. 11 Time series of interaction towards categories of finance high frequency (top left), entertainment high frequency (top right), finance low frequency (bottom left) and entertainment low frequency (bottom right)

is in line with the number of active unique addresses in Fig. 3. The *Entertainment High Frequency* group fluctuated even more sharply during this period, with a high of over 15,000 transactions per day and a low of under 1000. Unlike the previous group, the number of interactions with Entertainment DApps in this group has been gradually decreasing since January 2020, which can be explained by the higher transaction fees associated with the growth in transactions, causing these addresses to abandon or stop their blockchain games.

6 Discussion

Based on the above analysis, we would like to combine our work with previous blockchain literature in an attempt to provide a conclusive discussion of the nature of Ethereum activity, user behavior, and user groups. If there is one word to describe the activity on Ethereum, we believe '*Disparity*' is well deserved. We try to explain this disparity in terms of transactions, categories, and DApp.

– **Disparity of transactions:** We found that 50.67% of addresses have less than 50 transactions and on average, less than 1 transaction per day. we consider this address group to be static users who tend to use *Game* or *Exchange* ephemerally. This group of addresses may be new users who are only experimenting on the blockchain, or DApp developers for testing purposes. Furthermore, given the very large number of these addresses, they may also be considered non-human operated bots or affiliate accounts hunting for bonuses, since there's almost no cost of registering a wallet on Ether. While the number of addresses has exploded over the past few years, the existence of static accounts makes us suspicious of prosperity.
– **Disparity of categories:** There is a significant divergence in the popularity of DApps in different categories. We can observe this phenomenon visually in table 2, i.e. the traffic distribution. Of the transactions that can be interpreted, 44.89% of them go to the *Game* and *Exchange* categories. This reveals that the current primary use of DApps is for entertainment and investment, but between the popularity of NFT and the rise of GameFi, we can to some extent suggest that the primary use of DApps is to make profits. For non-interpretable transactions, as we have collected almost all of the thousands of major DApps that have some traffic, we infer that they are either p2p payments or functional transactions that are automatically initiated by contracts, with only a small chance that they are major DApps that have been missed. This extreme category distribution is not in line with researchers' expectations of various blockchain usage, such as better autonomous communities or as infrastructure serving distributed systems in the background (Elsden et al. 2018). We interpret this situation as a possible lack of user appreciation of DApps, the profit-seeking nature of the market due to the financial attributes of blockchain, and the development of other categories of DApps that may fail to meet user needs.
– **Disparity of DApps:** We observe a Matthew effect among DApps: the dominant DApps tend to maintain their leading position, and these DApps, like the existing Internet oligopolies, are able to monopolize a particular business, as Uniswap does in the Exchange business. Based on advertising, media coverage, and ranking sites, new users are more likely to choose the most popular DApps. Behind these phenomena is an interesting theoretical paradox: DApps were originally intended to be decentralized and anti-monopolistic (Cai et al. 2018), yet from a data perspective a new form of oligopoly still exists.

Gao et al. (2016) in a study of Bitcoin users suggest that not understanding how Bitcoin works may not be a barrier to using it, which is consistent with our results on the number of large numbers of light users. Krombholz et al. (2017) in a work examining how users experience the Bitcoin ecosystem in terms of security, privacy and anonymity found that most participants used Bitcoin for donations, virtual goods, online shopping, and gambling. While there are analogous similarities between Ethereum and Bitcoin, the dramatic fluctuations in the exchange rates of virtual and actual currencies over time have made the market inclined to treat them as commodities or stocks rather than circulating equivalents. The rise of DApps and decentralized finance has exacerbated this phenomenon. The likelihood of using ETH as payment for material commodities tends to be decreasing, as it is likely to yield greater returns investing in the virtual world. From the aspect of DApps, we conclude that users need to pay a higher cost for decentralization than for using centralized services. The use of a DApps is accompanied by a capital investment as commission, and although it is a small amount, the act suggests that the user has sufficient wealth and leisure. The entertainment value, service value, or financial return from the transaction are expected to cover the cost. In addition to the capital cost, users must also pay higher hidden costs such as knowledge acquisition to understand the blockchain and find valuable information in the clutter of emerging services. As this stands, DApps are more of a pioneering experiment than a service with the user experience as the first goal. Building on these discussions, we try to capture the motivation behinds the fact that users are willing to tolerate the inconvenience and investment risk of using DApp:

– **Curiosity:** In Presthus et al. (2017)'s study, almost all participants concluded the reason for using blockchain as curiosity and interest for new technology. Due to the media reports of various DApps and the gradual development of blockchain technology, we believe that there will be more users of this motivation in the Ethereum scenario.
– **Interest-driven:** Cryptocurrency has been a strong financial factor since the day it was invented. Because of the unique crowdfunding method of Initial Coin Offering

(ICO) (Wiki 2020), financial activities such as investment, speculation, and margin trading are prevalent in DApps (Elsden et al. 2018). The large number of deals with *Exchange* and *DeFi* show that such users make up considerable proportion.
- **Altruism:** Inspired by studies on communities and p2p systems (Bellotti et al. 2014, 2013), we suggested several DApp users are driven by Altruism or belief in the potential of Ethereum. We generalize this altruism into an open-source spirit, anti-censorship, community activism, emphasizing privacy and sharing. Users with this kind of motivation are likely to have a technical background and use the DApps from the perspective of the developer.

In the case of SOM-based clustering, using the number of transactions of addresses to different categories as input data yields sufficiently differentiated results. This may be a valid approach to address role classification. We generated four groups of addresses with financial and entertainment, high frequency, and low frequency. We observe that the financial group addresses use DApps with generally lower diversity, but their daily transaction frequency is more stable; the difference between the DApps used in the high-frequency and low-frequency entertainment groups is, we speculate, due to the fact that a large number of players who are attracted to play a game at a particular period are grouped in the low-frequency group, while the players in the high-frequency group choose a game that is popular over time. In addition, we find a marked discrepancy between investors and players facing changes in trading fees. Investors ignore the increase in commission and intensify the frequency of transactions, while players have a decreasing trend. Even those players who makeup only 0.55% of the sample and are considered as the most deserved blockchain game zealot still inevitably reduce their activity when the cost of usage increases. Extrapolating from the data and analysis we currently hold, the DApp usage on Ethereum remains finance-centric. This market is extremely sensitive to news and prices as participants are overwhelmingly composed of purposeful investors. The high volatility of virtual currencies and asset values has brought about a large number of arbitrageurs, creating an impact on blockchain gaming or autonomous communities.

7 Practical heuristics

During our research, we noticed the datasets generated by the transaction logs can do much more practical works than pure data mining. These systems can be built as the infrastructure of DApp ecosystem. We demonstrated two examples in this section trying to inspire future research.

7.1 Abnormality detection

To study a reference-missing domain, we can start by looking for abnormal patterns in data. As proposed by Elsden et al. (2018), a recurring critique of AI and ML is their tendencies to conform human activities to a machine-like view of the world. We can run into the same problems when applying ML methods to blockchain data. In this heuristic, we try to find addresses that are suspected to be non-human controlled.

We used the following filter: (1) *standard deviation of time intervals between transactions smaller than 1*, to filter stable transaction or multi-sending behavior; (2) *total transaction number larger than 1000*, to make sure that the interval between trading interval is small; (3) *transaction number to wallet addresses smaller than 10*, which means most actions are operations on DApps.

From all the samples, we selected 28 addresses that met the criteria above. Take the address, *0x68BXXXX*, as an example, it made more than 10,000 high-frequency interactions with CryptoKitties within 22 days. We plotted its transaction graph in Fig. 12a. The color of the nodes represents the category of the DApps. Nodes are linked with edges if the transaction number between them is greater than 3, and the width of edges shows the interaction frequency between them. We noticed that the only wallet address interacted with *0x68BXXXX* is *0x7ECXXXX*. This can be interpreted as *0x68BXXXX* is likely to be an affiliation of *0x7ECXXXX*, which acts as the representative to perform specific interactions with certain apps. Figure 12b showed that we can continually look for the relationship between nodes by setting the center on *0x7ECXXXX*. This can be an approach to analyze the composition of high-density light user addresses.

7.2 Recommendation system

To the best of our knowledge, none of the DApp indexing websites provide the recommendation list, which is a common service in Google Play or Apple Store. Recommendation methods can be roughly divided into collaborative filtering and content-based filtering. Since there is no centralized user system under the blockchain context, content-based filtering would be a more feasible choice. DApp recommendations can be implemented through the User Group.

We use the classic algorithm, Apriori, which is an algorithm for frequent itemset mining and association rule learning over relational databases (Agrawal and Srikant 1994), to generate a recommendation list. We selected *Entertainment* group from the analysis section as an example. In the group, each address interacted with 5.57 DApps on average (median = 3, SD = 8.23). Because the length of the frequent itemset needs to be greater than one, the length of the input lists should be far greater. We use a

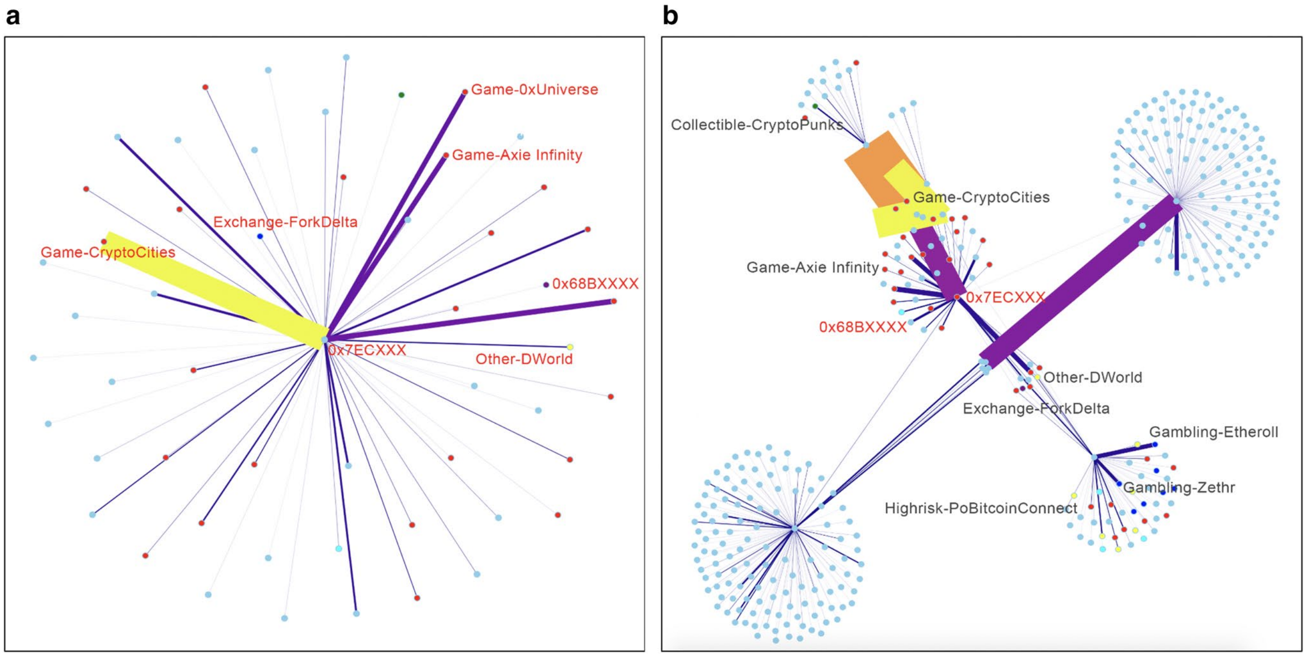

Fig. 12 2 degrees transaction graph centered on *0x68BXXXX* (**a**) and *0x7ECXXXX* (**b**)

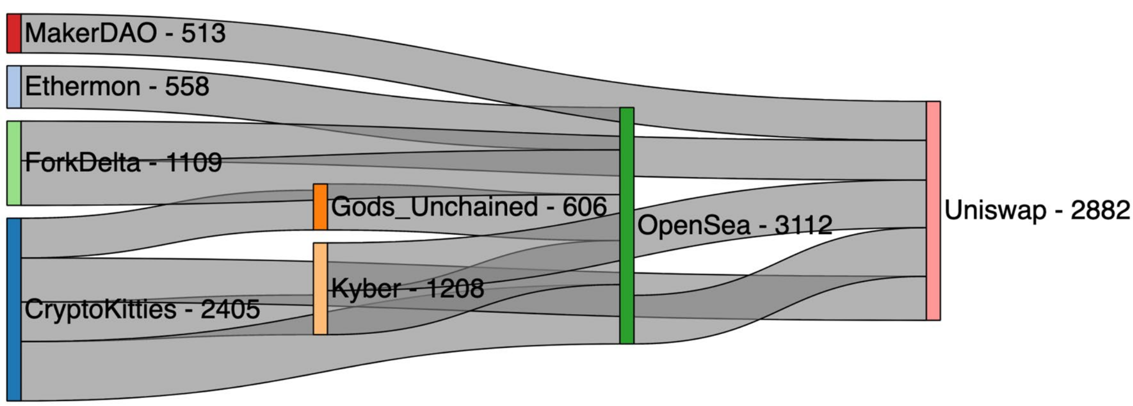

Fig. 13 Frequent item sets in entertainment group with min support = 0.3, min confidence = 0.8

filter of *'at least used 15 different DApps'* to ensure the performance of the Apriori algorithm.

Figure 13 shows the result of Apriori with minimum support of 0.3, minimum confidence of 0.8, and the frequent sets with a length of 2. Each interconnected DApps is considered to be related, and the number next to the name represents the number of times this frequent itemset appears in the input. The recommendation results by this heuristic are only generated by samples in the same user group. In practical implementation, more detailed filters, such as *'users with the same registration time'*, or *'users played CryptoKitties before'* and can be added. The recommendation list can be displayed on the DApp indexing sites and evaluated by counting clicks or analyzing the advertisement feedback.

8 Future direction

Through the analysis and demonstration above, we hope to have presented an overall impression of DApp user behavior to expand the research topics for researchers. At the same time, we hope that our dataset and the methodology to explore blockchain data can contribute to future studies. We concluded the limitation of this work and the future direction of a user study on blockchain as follows.

8.1 Expand the ways of knowing

We see our work as an attempt to understand DApp users from the publicly available data perspective. We are still far away from fully understanding the DApp user habits. One of the most critical issues is that addresses have limited bind with people in the real world. We need more DApp users surveys and interviews to get demographic or psychographic information to feature this population and understand their attitudes, interaction patterns with DApp (Olson and Kellogg 2014). This information will act as a bridge that links together the data and behavioral patterns and improve the accuracy of conclusions. For example, if there is a reference about how many addresses a typical user would register, we can calculate a more precise number of actual users, or use it as a basis to determine subsidiary accounts. Furthermore, the survey and data analysis results can verify the consistency with each other.

It will take great efforts to carry out these investigations. Crowdsourcing might not be an ideal solution, because the blockchain user is a minority group. Researchers can look for respondents through communities such as BitTalk,[7] or only reward participants with cryptocurrencies to ensure they have blockchain addresses (Krombholz et al. 2017). Separate studies for different platforms, such as Bitcoin, Ethereum, or EOS,[8] are expected to be conducted. Nowadays, the new consensus models like *'Proof-of-Stake'* (King and Nadal 2012), have a huge improvement in performance than PoW, which may have a great impact on DApp design ideas and user experience.

8.1.1 Measure the QoE of DApp and blockchain

There is very little research on evaluating the QoE of DApps or blockchains. We can only judge the popularity of a DApp by how often it is used. More internal factors behind a popular DApp need to be found. Similarly, for Blockchain platforms with different consensus models or policies, previous evaluations stop at describing the user experience using words such as *'efficiency safety, and convenience'* (Zheng et al. 2017). The research community can act as the role of standardizing these measurements, and present the evaluations to DApps developers, giving them a more profound understanding of user experience.

8.1.2 Guide the design of DApps

Krombholz et al. (2017)'s study highlighted the necessity of user interface designs that are targeted at improving the usability of blockchain. Elsden et al. (2018) pointed out that researchers might look to working with industry to uncover unique use cases, and look beyond the investor-driven applications prevalent. We suggest that for DApp scenario, is similarly in need of methodologies that quantify DApp user's experience DApps usability, which can provide valuable design guidelines. Questions like *'Do monopolistic DApps have unique interaction designs?'*, *'Which part is the most critical to users during interaction with DApps?'* count a great deal of research value.

In addition to interaction design support, the research community in human factors can also work on policy aspects of autonomous communities. Currently, centralized communities adopt a variety of policies to maintain decentralized power, a typical example is governance token. members receive governance token by participating in community activities, and these tokens represent voting power for decision making, meaning that active community members will receive a larger number of tokens, i.e., have a greater voice. How to allocate the voting power reasonably becomes the key to decentralize the community. A simple example is the largest exchange on Ethereum, Uniswap, which uses liquidity mining to distribute governance tokens, but because it has ERC20 standard tradability, most holders use them for arbitrage rather than participating in community governance. Testing the effectiveness of the policy and proposing a more reasonable policy will be research community in human factors. community to explore.

9 Conclusion

In this paper, we attempt to portray DApp users through large-scale Ethereum data, seeking to present an understanding of the data aspects of users in the blockchain scenario beyond surveys and interviews. We built a series of datasets labeled with DApp name and extracted information about each address to enrich the data dimension. We then visualized and analyzed the user profile dataset from several aspects and explored methodologies to divide user groups. We propose a way to use the number of interactions by the categories of DApp as cluster input and classify the addresses with SOM network. After that, we separate active addresses into four groups according to the purpose and frequency of use and discuss the differences between them and their sensitivity to the market. In addition, we give examples of how to use transaction data. We combine the results of our analysis with previous research studies to summarize the profile of DApp users and conclude by speculating on their motivations, values, and the current DApp market. We conclude the paper by discussing future directions in this area and encourage more researchers to explore user patterns and participate in the design of decentralized applications.

[7] https://bitcointalk.org/.

[8] https://eos.io/.

Acknowledgements This work was supported by Shenzhen Science and Technology Program (Grant No. JCYJ20210324124205016).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94, p. 487–499. Morgan Kaufmann Publishers Inc., San Francisco(1994)

Bellotti, V., Carroll, J.M., Han, K.: Random acts of kindness: the intelligent and context-aware future of reciprocal altruism and community collaboration. In: 2013 International Conference on Collaboration Technologies and Systems (CTS), pp. 1–12 (2013)

Bellotti, V.M., Cambridge, S., Hoy, K., Shih, P.C., Handalian, L.R., Han, K., Carroll, J.M.: Towards community-centered support for peer-to-peer service exchange: Rethinking the timebanking metaphor. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, p. 2975–2984. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2556288.2557061

Biryukov, A., Khovratovich, D., Pustogarov, I.: Deanonymisation of clients in a bitcoin p2p network.In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, CCS '14, p. 15–29. Association for ComputingMachinery, New York, NY, USA (2014).

Cai, W., Wang, Z., Ernst, J., Hong, Z., Feng, C.: Decentralized applications: the blockchain-empowered software system. IEEE Access **6**, 53019–53033 (2018). https://doi.org/10.1109/ACCESS.2018.2870644

Carroll, J.M., Bellotti, V.: Creating value together: the emerging design space of peer-to-peer currency and exchange. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '15, p. 1500–1510. Association for Computing Machinery, New York (2015). https://doi.org/10.1145/2675133.2675270

Chan, W., Olmsted, A.: Ethereum transaction graph analysis, 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST), pp. 498–500, (2017)

Chen, T., Zhu, Y., Li, Z., Chen, J., Li, X., Luo, X., Lin, X., Zhange, X.: Understanding Ethereum via Graph Analysis, IEEE Conference on Computer Communications, pp. 1484–1492, (2018)

Coindesk.: A David vs. Goliath battle is brewing in ethereum decentralized exchange race. (2019). https://www.coindesk.com/bancor-uniswap-dex-competition

DAppTotal.: Eth-all time-active dapps. (2020). https://dapptotal.com/analytics

Dwork, C., Naor, M.: Pricing via processing or combatting junk mail. In: Proceedings of the 12th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '92, p. 139–147. Springer, Berlin (1992)

Elsden, C., Manohar, A., Briggs, J., Harding, M., Speed, C., Vines, J.: Making sense of blockchain applications: a typology for HCI. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18, p. 1–14. Association for Computing Machinery, New York (2018). https://doi.org/10.1145/3173574.3174032

Ethereum.: Contract ABI specification. (2020a). https://solidity.readthedocs.io/en/v0.5.3/abi-spec.html

Ethereum.: Web3.js documentation. (2020b). https://web3js.readthedocs.io/en/v1.2.6/

Etherscan.: Ether daily price (USD) chart. (2020a). https://etherscan.io/chart/etherprice

Etherscan.: Ethereum unique addresses chart. (2020b). https://etherscan.io/chart/address

Gao, X., Clark, G.D., Lindqvist, J.: Of two minds, multiple addresses, and one ledger: characterizing opinions, knowledge, and perceptions of bitcoin across users and non-users. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, p. 1656–1668. Association for Computing Machinery, New York (2016). https://doi.org/10.1145/2858036.2858049

Gervais, A., Karame, G.O., Wüst, K., Glykantzis, V., Ritzdorf, H., Capkun, S.: On the security and performance of proof of work blockchains. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS '16, p. 3–16. Association for Computing Machinery, New York (2016). https://doi.org/10.1145/2976749.2978341

Jourdan, M., Blandin, S., Wynter, L., Deshpande, P.: Characterizing entities in the bitcoin blockchain. 55–62. (2018). https://doi.org/10.1109/ICDMW.2018.00016

King, S., and S. Nadal, Ppcoin: Peer-to-peer crypto-currency with proof-of-stake, pp. 1, August 19, (2012).

Kohonen, T.: The self-organizing map. Proc. IEEE **78**(9), 1464–1480 (1990). https://doi.org/10.1109/5.58325

Krombholz, K., Judmayer, A., Gusenbauer, M., Weippl, E.: The other side of the coin: user experiences with bitcoin security and privacy. In: Grossklags, J., Preneel, B. (eds.) Financial Cryptography and Data Security, pp. 555–580. Springer, Berlin (2017)

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G.M., Savage, S.: A fistful of bitcoins: Characterizing payments among men with no names. In: Proceedings of the 2013 Conference on Internet Measurement Conference, IMC '13, p. 127–140. Association for Computing Machinery, New York, NY, USA (2013).

Olson, J.S., Kellogg, W.A.: Ways of Knowing in HCI. Springer Publishing Company, Incorporated, Berlin (2014)

Presthus, W., O'Malley, N.O.: Motivations and barriers for end-user adoption of bitcoin as digital currency. Procedia Comput. Sci. **121**, 89–97 (2017). https://doi.org/10.1016/j.procs.2017.11.013. CENTERIS 2017—International Conference on ENTERprise Information Systems/ProjMAN 2017—International Conference on Project MANagement/HCist 2017—International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2017

Prinz, W.: Blockchain and CSCW—shall we care? (2018). https://doi.org/10.18420/ecscw2018_13

Ranshous, S., Joslyn, C., Kreyling, S., Nowak, K., Samatova, N., West, C., Winters, S.: Exchange pattern mining in the bitcoin transaction directed hypergraph, Financial Cryptography and Data Security (FC 2017), Springer, pp. 248–263 (2017). https://doi.org/10.1007/978-3-319-70278-0_16

Thielsch, M.T., Blotenberg, I., Jaron, R.: User evaluation of websites: from first impression to recommendation. Interact. Comput. **26**(1), 89–102 (2013). https://doi.org/10.1093/iwc/iwt033

Werbach, K.: Trust, but verify: why the blockchain needs the law. Berkeley Technol. Law J. **33**, 487 (2017)

Wiki.: Initial coin offering. (2020). https://en.wikipedia.org/wiki/Initial_coin_offering

Wood, D.: Ethereum: a secure decentralised generalised transaction ledger, Yellow Paper (2014). http://gavwood.com/paper.pdf.

Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H.: An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE International Congress on Big Data (BigData Congress), pp. 557–564 (2017)

Tian Min received B.Eng. in Computer Science and Engineering from The Chinese University of Hong Kong, Shenzhen in 2020. His article on blockchain gaming and security was published in 2019 in IEEE Conference on Games and IEEE Games Entertainment Media, and has received over 60 citations to date. His main research interests lie in human-computer interaction, social computing, blockchain technology, data mining, and machine learning.

Wei Cai is currently an assistant professor of Computer Engineering, School of Science and Engineering at The Chinese University of Hong Kong, Shenzhen. He is serving as the director of the Human-Cloud Systems Laboratory, as well as the director of the CUHK(SZ)-White Matrix Joint Metaverse Laboratory. Wei received Ph.D., M.Sc. and B.Eng. from The University of British Columbia (UBC), Seoul National University and Xiamen University in 2016, 2011 and 2008, respectively. Before joining CUHK-Shenzhen, he was a postdoctoral research fellow in Wireless Networks and Mobile Systems (WiNMoS) Laboratory at UBC. He has completed research visits at Academia Sinica (Taiwan), The Hong Kong Polytechnic University and National Institute of Informatics, Japan. Dr. Cai has co-authored more than 70 journal and conference papers in the area of interactive multimedia, distributed computing, and decentralized systems. His recent research interests are mainly in the topic of human factors in metaverse, including blockchain, digital game, human-computer interaction, Web 3.0, DeFi/GameFi, UGC/AIGC, and computational art. He is serving as an associate editor of IEEE Transactions on Cloud Computing, Committee Members of China Computer Federation (CCF) Computational Art Committee, Technical Committee on Human-Computer Interaction, and Technical Committee on Blockchain. He was a recipient of the 2015 Chinese Government Award for the Outstanding Self-Financed Students Abroad, the UBC Doctoral Four-Year-Fellowship from 2011 to 2015, and the Brain Korea 21 Scholarship. He also received the best student paper award from ACM BSCI2019 and the best paper awards from CCF CBC2018, IEEE CloudCom2014, SmartComp2014, and CloudComp2013.