World Wide Web
https://doi.org/10.1007/s11280-022-01077-4

Towards understanding governance tokens in liquidity mining: a case study of decentralized exchanges

Sizheng Fan[1] · Tian Min[1] · Xiao Wu[2] · Cai Wei[1]

Received: 28 January 2022 / Revised: 18 April 2022 / Accepted: 9 June 2022

Abstract

The boom of liquidity mining has attracted enormous attention, which has brought tens of times increment in total value locked (TVL) to decentralized finance (DeFi) community. Meanwhile, governance tokens, as part of the liquidity mining reward, have been adopted by most decentralized applications (DApps) to attract users. However, the effectiveness of this method has not been proven in detail. In this paper, we choose one of the most representative cases where SushiSwap absorbed a significant amount of Uniswap liquidity in no time by forking Uniswap's code and issuing the governance token ahead to understand the governance tokens in liquidity mining. Specifically, we collect transaction records of Uniswap and SushiSwap for over a year and perform a detailed analysis of liquidity providers' (LPs) activities. Moreover, we design a scalable unsupervised clustering method, which uses metrics between transaction flows to build a similarity graph that can capture patterns between LPs with similar behaviour. These LPs range from inactive and cautious LPs, providing tiny liquidity to risk-seeking LPs, focusing on short time-intervals. On this basis, we discuss how the governance token affects liquidity mining, and use its impact on behaviours and decision-making to explain its attractiveness to users.

Keywords Decentralized finance · Liquidity mining · Governance token · Automatic Market Maker-based decentralized exchanges

---

This article belongs to the Topical Collection: *Web-based Intelligent Financial Services*
Guest Editors: Hong-Ning Dai, Xiaohui Haoran, and Miguel Martinez

---

✉ Cai Wei
caiwei@cuhk.edu.cn

Sizheng Fan
sizhengfan@link.cuhk.edu.cn

Tian Min
tianmin@link.cuhk.edu.cn

Xiao Wu
wuxiao@whitematrix.io

[1] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

[2] White Matrix, Nanjing, China

Published online: 12 July 2022

1 Introduction

Liquidity mining was first introduced by one of the largest decentralized exchanges (DEXs), IDEX[1], back in 2017 and was fine-tuned by Synthetix[2] and Chainlink[3] in 2019. After a year in the middle of 2020, liquidity mining flourished and was used at full throttle by Compound[4], SushiSwap[5], and Uniswap[6]. Thus far, liquidity mining has caused a permanent and significant impact on the decentralized finance (DeFi) community. Based on data from DeFi Pulse[7], the market hosted \$1.05 billion in collateralized assets at the start of June 2020. By September, the community had kick-started the market, which led to a 10-times increment in locked assets. This phenomenal growth can even rival the Initial Coin Offering (ICO) boom of 2017 in terms of enthusiasm [1].

Liquidity mining represents a new way of using cryptocurrency, similar to bank deposits [2]. Specifically, liquidity providers (LPs) provide token pairs in an automatic market marker (AMM)-based DEX, a smart contract that serves as a trading venue without an order book mechanism. Traders can effectively swap tokens with the token pairs locked in the DEXs. After completing each trade, DEXs will charge traders a commission, which is 0.3% in both Uniswap [3] and SushiSwap, as an interest reward for the LPs. Besides trading fees, governance tokens, as stakes in the protocol, are also added to the reward to incentive LPs. Governance tokens are cryptocurrencies that represent voting power on a blockchain project. Recently, they are mostly integrated into DeFi projects since they need to distribute rights to their clients to remain decentralized. Moreover, governance tokens have economic value, which holders can trade on centralized exchanges (CEXs) and DEXs. The number of governance tokens issued is regulated by a protocol similar to the ICO white paper, with essentially a decreasing trend in the number issued.

Today, it has been adopted by most decentralized protocols and is considered an innovative and efficient way of allocating governance tokens to achieve decentralization. The protocols can attract more LPs by issuing or increasing the governance tokens reward in liquidity mining, as in SushiSwap's Vampire Attack on Uniswap [4]. To the best of our knowledge, there is no in-depth analysis of the effectiveness of this method.

A straightforward way to answer this question is to survey how the governance tokens affect users' strategies. However, conducting user surveys will incur higher costs, especially on the blockchain. For the survey results to be representative, a relatively large number of user profiles would need to be collected. It would be challenging to send questionnaires to eligible users by their wallet addresses or select users who participated in governance tokens from the blockchain community. At the same time, the cryptocurrency used to pay as the rewards will entail a substantial monetary cost. Based on these considerations, it would be wiser to extract the information we need from open-source data.

Therefore, in this paper, we have selected a particular case where SushiSwap has gained significant liquidity by issuing the governance token, SUSHI, and *permanently* adding it to the rewards of liquidity mining, which forms a Vampire Attack on Uniswap. Uniswap subsequently issued the governance token, UNI, and added it to the rewards of liquidity mining

1 https://idex.io/
2 https://synthetix.io/
3 https://chain.link/
4 https://compound.finance/
5 https://sushi.com/
6 https://uniswap.org/
7 https://defipulse.com/

for, *two months* from September 18, 2020, to November 18, 2020. Based on the differences between Uniswap and SushiSwap in this example in terms of issuance time and reward cycle of governance tokens, we try to understand the impact of governance tokens in liquidity mining in the following two aspects. 1) *Macroscopic Level*: we analyze the difference between Uniswap and SushiSwap through macro data, reflecting the effect of external factors such as ETH price and the initiation of incentive policies on total value locked (TVL), function calls, and the number of users[8]. 2) *Microscopic Level*: we perform unsupervised clustering on LPs and compare their behaviour before and after Vampire Attack, trying to discover the sensitivity of different types of LPs to incentives. Our main contributions can be summarised as follows:

- *Database Construction*: We briefly introduce the significant invents in Uniswap, the most popular DEX on the Ethereum, and SushiSwap, the earliest and most prominent fork of Uniswap, respectively. Although the transparency of the blockchain dictates open-source on-chain data, integrating, pre-processing, and extracting the usable parts is still a considerable undertaking. We collected nearly a year of granular transaction data, including records of around 300,000 addresses from these two DEXs to build an database.
- *Transaction Flow Extraction and Clustering*: To the best of our knowledge, We are the first to propose a method of formatting address's *transaction flow* based on our database. Moreover, we utilize an unsupervised hierarchical clustering method to capture LPs' behaviours and divide LPs into six categories: dispensable, active light, inactive light, risk-averse medium, risk-seeking medium and heavy LP.
- *Result Analysis*: Based on the cluster results, we conclude that governance tokens reward can attract more LPs in a short time, but several days later, attracted LPs tend to remove their liquidity for higher revenue, decreased annual percentage yield (APY)[9], and potential impermanent loss, especially for medium and heavy LPs, which not only deviates from the original participatory purpose of governance tokens but will also lead to a vicious cycle where traders are more likely to suffer slippage, further bringing down LPs returns. Besides, by comparing the overlapping LP ratios of SushiSwap and Uniswap from one year to the next, we can obtain that SushiSwap has gradually developed dedicated fixed LP groups over almost a year through the long-term SUSHI incentive.

The remainder of this paper is organized as follows: Section 2 introduces the preliminary background knowledge. Section 3 presents the related works and analyzes their limitations. Data collection process and a staged summary are presented in Section 4. We perform feature extraction and unsupervised clustering in Section 5. Macroscopic data analysis and cluster results are discussed in Section 6 sequentially. As a conclusion, Section 7 and Section 8 summarize this paper and discuss possible future vision, respectively.

2 Background

2.1 Ethereum and modern cryptocurrencies

Ethereum is a blockchain platform that builds on Bitcoin's innovation but provides the end-developer a tightly integrated end-to-end system for building software on a hitherto

[8] The users here represent the LPs and traders of DEXs.

[9] The APY is the rate of return on the investment, considering the effect of compounding interest.

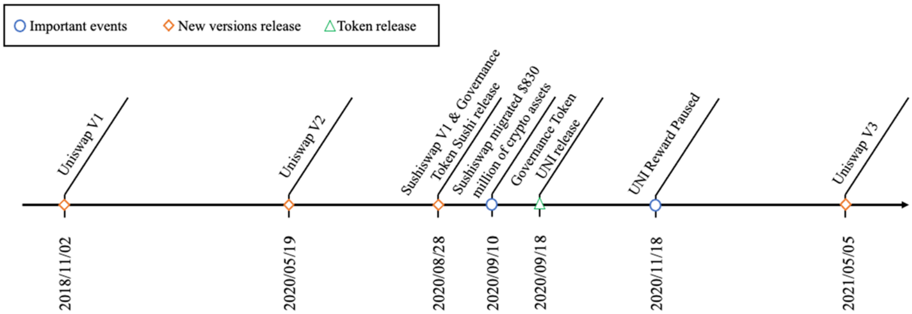

Figure 1 Timeline of the DEXs Development

unexplored compute paradigm in the mainstream: a trustful object messaging compute framework [5]. Smart contracts are scripts that run synchronously on multiple nodes of a distributed ledger without the need of an external trusted authority [6, 7, 8].

Ethereum-based cryptocurrencies are all Ethereum Request for Comments 20 (ERC-20) compliant[10], and they share several key characteristics whose tradability lays the foundation for the motivation of this paper. They each contain a supply of tokens that is both discrete and finite, and use Ethereum that anybody can inspect. Additionally, currency owners are able to transfer custody of the tokens amongst each other. These properties have led to the emergence of markets whereby users exchange tokens for other tokens, either with the assistance of exchanges including DEXs and CEXs or through some peer-to-peer process.

2.2 Impermanent loss

As illustrated in Section. 1, liquidity mining is simply a passive income method that helps cryptocurrency holders profit by utilizing their existing assets rather than leaving them inactive in the wallets. Assets are deposited to a decentralized exchange and in return, the platform distributes fees earned from trading to each LP proportionally. Impermanent loss, also known as divergence loss, refers to the loss that funds are exposed to when they are in a liquidity pool [9]. This loss typically occurs when the ratio of the tokens in the liquidity pool changes, which means the user has suffered negative returns compared with simply holding their tokens outside the pool [10]. In this case, the DeFi protocols tend to use the transaction fee from the trader to compensate the LPs. Some even add additional rewards – governance tokens to attract more liquidity.

2.3 The timeline of uniswap and sushiswap

In this subsection, we will briefly introduce remarkable events in the development of DEXs, and illustrate a brief timeline, as shown in Figure 1. Specifically, we divide the development of Uniswap and SushiSwap into the four stages: *Steady growth period*, *Vampire Attack*, *The counterattack from Uniswap* and *Boom period*.

[10] The ERC-20 introduces a standard for fungible tokens, in other words, they have a property that makes each token be exactly the same (in type and value) of another token.

As shown in Figure 1, the initial version of Uniswap, called Uniswap V1, was published to the Ethereum mainnet on November 2, 2018. The first version of the protocol was launched with $30,000 worth of initial liquidity across three different tokens. However, the Uniswap V1 protocol is only designed to promote automatic exchange transactions between ETH and ERC-20 tokens, which means that each liquidity pool must have ETH as one of the trading pairs. Therefore, trade between ERC-20 tokens must be implemented via ETH, which usually results in higher gas fees[11], commission and slippage. Simply speaking, the exchange between token pairs needs to be done through the intermediary of ETH, which turns the original need of only one transaction into two.

Stage 1. *Steady growth period* (2020/05/19-2020/08/28): In May 2020, Uniswap launched its second version, whose main feature is the adding ERC-20/ERC-20 liquidity pool. The direct swap between ERC-20 tokens significantly reduces transaction fees and waiting time. In addition, providing liquidity to the ERC-20/ERC-20 pool allows LPs to have a lower probability of facing impermanent losses than an ETH-related pool [3]. Moreover, Uniswap V2 implements features such as on-chain price feedback and flash swap.

Stage 2. *Vampire Attack* (2020/08/28-2020/09/17): SushiSwap also came into play at the end of August, aiming at directly competing with Uniswap by forking the project by adding the governance token reward for Uniswap's LPs and eventually stealing Uniswap's liquidity into the SushiSwap platform. Specifically, the first step of a Vampire Attack is to incentivize Uniswap's LPs who stake their UNI-V2[12] by rewards paid in SUSHI. SushiSwap started with an aggressive schedule for the SUSHI token: 1,000 SUSHI per Ethereum block are allocated to Uniswap's LPs across multiple different pools. Once enough liquidity has been transferred, staked UNI-V2 tokens are migrated from Uniswap to SushiSwap. Eventually, SushiSwap stole not only the liquidity but also the trading volume and the users from Uniswap.

Stage 3. *The Counterattack from Uniswap* (2020/9/18 - 2020/11/18): On September 16, to deal with the Vampire Attack from SushiSwap, Uniswap announced the launch of their new token – UNI. The most surprising was that a part of UNI was retrospectively allocated. Addresses who had interacted with Uniswap before September 1 were eligible to claim 400 UNI worth around $1200 at that time. Besides, Uniswap announced four liquidity pools that would incentivize LPs with extra UNI tokens in **the next two months**, which resulted in millions of dollars of increment in liquidity.

Stage 4. *Boom period* (2020/11/18-2021/5/19): After Uniswap stopped rewarding LPs with extra governance tokens, its TVL experienced a brief dip. Then, Uniswap and SushiSwap's trading volume and TVL have risen rapidly with the price of ETH and Bitcoin since November 2020 because more and more funds started to flow into the DeFi. By the way, Uniswap V3 protocol was released on May 5, 2021.

3 Related work

Recently, the explosion of blockchain-related technologies has attracted extensive attention from academia and industry [1112]. Bitcoin is a decentralized digital currency, without a central bank or single administrator, that can be sent from user to user without the need

[11] Everyone who calls smart contracts on the blockchain must pay a gas fee to miners, determined by the gas and gas prices.

[12] UNI-V2 tokens represent supplied liquidity in Uniswap. If an LP provides liquidity in Uniswap, it will receive a certain number of UNI-V2 tokens. When the liquidity is removed by the LP, the UNI-V2 tokens will be destroyed.

for intermediaries [13]. In 2013, Vitalik Buterin proposed a decentralized platform named Ethereum [14, 15], which introduced a smart contract for autonomous and transparent program execution, with thousands of novel decentralized applications (DApps) developed [16], e.g. Metaverse [17], blockchain games [18, 19] and DeFi [20, 21].

There is a growing body of literature on DeFi, especially in contrast to centralized finance (CeFi). Qin et al. [22] systematically analyzed the differences between CeFi and DeFi, covering legal, economic, security, privacy and market manipulation. Besides, they provided a structured methodology to differentiate between a CeFi and a DeFi service. Another reference analyzed the existing database to investigate the development of the CEXs and DEXs, and the difference between the role of governance tokens between these two kinds of exchanges [23]. Chen et al. [24] discussed the benefits of DeFi, identified existing business models[13], and evaluated potential challenges and limits. They concluded that as a new area of financial technology, decentralized finance might reshape the structure of modern finance and create a new landscape for entrepreneurship and innovation. Moreover, they showed the promises and challenges of decentralized business models.

The most successful application of DeFi is DEX, such as Uniswap. Extensive studies have been conducted to analyze the theoretical properties of Uniswap, including the arbitrage model [25], LP risk profile [26] and the improvement of AMM [27]. In contrast, few researchers investigated LPs' behaviours in AMM-based DEXs. The existing reference analyzed LPs' different investment strategies among different liquidity pools and their benefits in Uniswap. However, they might neglect external factors like the opponent of Uniswap and the effect of governance tokens [28]. However, all of them focus on the macroeconomic data like the total number of liquidity pools rather than LPs' behaviour. Hence, in this paper, we fill the blank and present the first open-source database about Uniswap and SushiSwap from May 2020 to July 2021. Moreover, we analyzed the LPs' behaviours during this period systematically.

4 Dataset

As illustrated in Section 2.3, SushiSwap's Vampire Attack on Uniswap, and Uniswap's counterattack provide a particular case where the incentive effect of governance tokens in liquidity mining can be studied. Hence, in this section, we take the two most representative DEXs, Uniswap and SushiSwap, as examples, obtain the open records of addresses who have interacted with their contracts from May 2020 to July 2021 with Etherscan[14]. Based on the overview and analysis and the dataset, we have the following observations.

- **Observation 1.** While Uniswap has ten times the volume of SushiSwap, SushiSwap has twice the LP address percentage of Uniswap. Its perpetual governance token distribution may be the reason behind the higher LP participation.
- **Observation 2.** Although Uniswap and SushiSwap offer the same services, there are differences in the percentage of users using these services. Users tend to make ETH-Token deals at Uniswap, while a greater percentage of Token-Token trades at SushiSwap.

[13] The classified the existing business model as decentralized currencies, decentralized payment services, decentralized fundraising and decentralized contracting.

[14] https://etherscan.io/

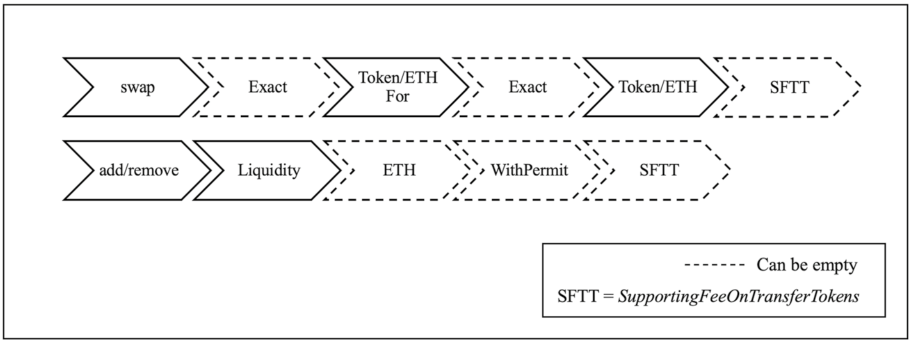

Figure 2 Function Naming Convention

4.1 Contracts and function

DEX's router contracts are integrated interfaces for exchanging token pairs or managing liquidity. Therefore, we can use the router contract as retrieval to record all DEX users' operations.

To implement these features, router contracts contain a variety of token trading and liquidity-related functions that often serve as interfaces to implement a specific set of operations by evoking other contracts. In our dataset, Uniswap and SushiSwap have been called by 26 and 33 functions, respectively, of which 21 are identical. These functions can be broadly classified into three categories in terms of their capabilities: 'swap' functions, which implement trades between ETH-tokens or token-token in various contexts, 'add/remove' functions, which are used to increase or decrease liquidity in a unit of ETH or token pairs, and 'utility' functions, which serve as panels for inquiries, administration or emergency response. To cope with all kinds of possible transaction scenarios, these functions generate variants that follow a certain naming convention, as shown in Figure 2. By filtering keywords, we can select addresses that have used a particular function.

4.2 Data collection

We used a segmentation approach to obtain the relevant external transaction records up to July 2021, block by block. As shown in Figure 3, after collecting router contract records,

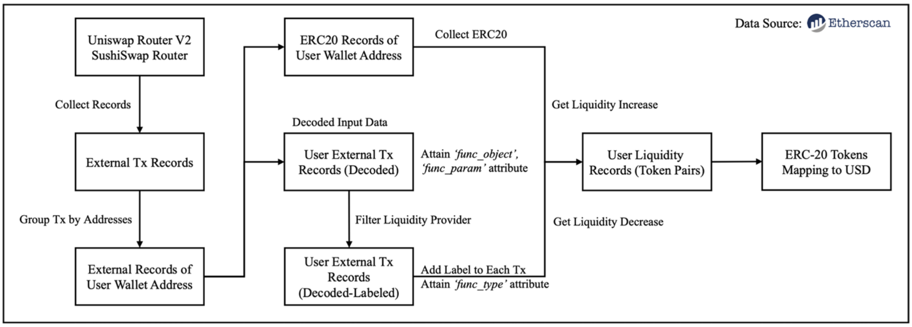

Figure 3 Data collection process

Table 1 Transactions and user statistics for Uniswap and SushiSwap

DEXs	Transaction	Unique Address	Liquidity Participant
Uniswap	46,077,169	2,310,175	297,345
SushiSwap	2,030,355	160,345	43,705

Table 2 Function Type Called and Percentage of Address Called for Uniswap

	Uniswap		
Function Type	# Called	% Called	% Called by Address
ETH-Token	17,504,158	44.18%	84.12%
Token-ETH	10,698,595	27.00%	49.08%
Token-Token	936,842	23.64%	32.13%
Add Liquidity	1,337,593	3.38%	12.69%
Remove Liquidity	715,129	1.80%	9.25%

we grouped the external transactions by wallet address to obtain a list of users who have interacted with the DEXs. With the Application Binary Interface (ABI) of the router contracts, we can decode the input value in each transaction to obtain the corresponding function object (*func_obj*) and parameters (*func_param*). After decoding, LPs can be filtered out by the keywords in the name of the functions. We set up a series of dictionaries to classify the function objects, which are used to generate the *func_type* attribute. Then, we extract the operation of liquidity decrements by the keyword *'add liquidity'* or *'remove liquidity'* in function type to get the corresponding number of token pairs.

On the one hand, we decode, filter, and label the external transactions, while on the other hand, we obtain their ERC-20 transaction records based on the list of addresses that have participated in liquidity activities. From the ERC-20 token transaction log, we can locate the numbers of token pairs acquired or given away by LPs when adjusting liquidity based on the hash value of transactions, which was then converted to USD at a price on the day[15] the transaction was made based on over 9,000 kinds of ERC-20 price lists obtained from CoinGecko[16]. After the above preprocessing steps, we can have the timestamp and quantity of liquidity changes for every LP, which can be formatted into a time series to further research in the subsequent sections.

Up to July 2021, the Uniswap and SushiSwap router contracts have 46,077,169 and 2,030,355[17] trade records, as illustrated in Table 1. Categorization by same address yields 2,310,175 and 160,345 independent addresses, respectively, of which 297,345 and 43,705 addresses are involved in providing liquidity, representing 12.8% and 27.2% of the total number of addresses. After comparing the address lists, Uniswap and SushiSwap have 50,176 overlapping addresses, of which 27,521 are overlapping LPs.

[15] We have taken the price at the time of transactions made in the paper. We believe this value can better reflect LP's decisions and behaviours and is the most feasible and persuasive solution to make the data discrete.

[16] https://www.coingecko.com/

[17] Considering the most crucial anonymity feature of blockchain, we follow the assumption made by Lee [29], which also takes an address as the basic unit

Table 3 Function Type Called and Percentage of Address Called for SushiSwap

Function Type	SushiSwap		
	# Called	% Called	% Called by Address
ETH-Token	486,885	28.27%	62.39%
Token-ETH	415,598	24.13%	40.96%
Token-Token	544,125	31.59%	34.69%
Add Liquidity	170,363	9.89%	29.23%
Remove Liquidity	105,549	6.13%	23.97%

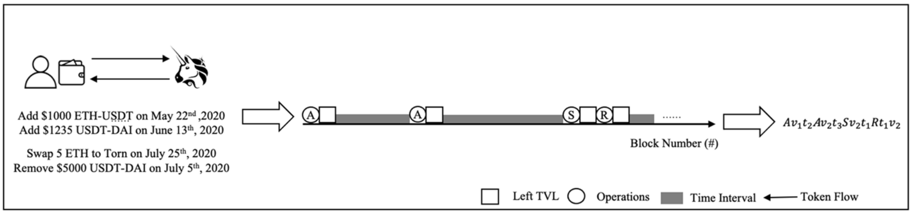

Figure 4 The Process of Formatting Transaction Flow

Tables 2 and 3 shows the function called of Uniswap and SushiSwap. We classify the functions from router contracts into five categories based on the naming convention in Figure 2, using their functionalities as the basis for categorization. Among the categories, there are three 'Swap' classes: 'ETH-Token', 'Token-ETH', and 'Token-Token', each of which represents an exchange pair that can be traded on DEXs, and two liquidity classes: 'Add Liquidity' and 'Remove Liquidity'. The **# Called** in the header of the table represents the total number of times the corresponding function has been called; the **% Called** represents its percentage of the total number of calls, and the **% Called by Address** indicates what percentage of all addresses have called the corresponding function at least once.

As we can see from the statistics, Uniswap and SushiSwap offer the same services, but there is a difference in their usage. Uniswap has tens of times the users of SushiSwap, but the percentage of its users using the five categories of functions may not increase proportionally in terms of **% Called by Address**. The functions used in Uniswap are focused on 'Swap', where 'ETH-Token' trade have a definite advantage. In contrast, in SushiSwap, 'Token-Token' trade is more popular, with 34.69% of users providing 31.59% of the number of its transactions. Moreover, the liquidity function of SushiSwap is higher than that of Uniswap in terms of both percentages of times and percentage of users.

5 Methodology

In the following section, we first introduce the definition of *transaction flow* and users' *similarity graph*. Then, we describe the unsupervised clustering algorithm to capture user groups with similar behaviours.

5.1 Formatting transaction flows

Based on the data obtained from Section 4.2, for each LP address, we form its records into a specific transaction flow: a discrete sequence of events, which describes the function called by an address, the time interval of the operations, and the liquidity remains after the operation. Figure 4 illustrates the transaction flow of address *0x72*, for example, representing this address added \$1,000 worth of liquidity in ETH-USDT pool in Uniswap; then, it added another \$12,365 worth of liquidity in USDT-DAI pool on June 13; after that, it swapped 5 ETH, and removed \$5,000 worth of liquidity from USDT-DAI on July 25.

We use *A, R* and *S* to denote the behaviour of *Add Liquidity*, *Remove Liquidity* and *Swap*, respectively. At the same time, we calculated the TVL remaining at the address after each operation and discretized the result with the following rule: *[0, \$100), [\$100, \$1000), [\$1000, \$10,000), ≥ \$10,000*, i.e., data binning. The same data binning method is applied to generate discrete time gap: *[0, 1day), [1day, 1week), [1week, 1month), ≥ 1month*. After these processes, the above transaction flow are further formatted to $Av_2t_3Av_3t_4Sv_3t_1Rv_2$, as shown in Figure 4.

5.2 Transaction flow similarity graph

Our Clustering algorithm is based on a similarity graph, where each node represents an address, and each edge represents similarity weight between two addresses' transaction flows [30]. We identify the address behavioural clusters by partitioning the similarity graph. To do so, we need a metric to measure the similarity degree between any two transaction flows.

Our method is to extract subsequences from the transaction flow as features to compare similarities. Specifically, as shown in Figure 4, we formalize the transaction flow of the m th address as a sequence $S_m = \left\{I_1,\ I_2, \dots, I_q\right\}$, where I_i is the i th item in the transaction flow (either a operation, left value or a time gap), and q is the total number of items in the sequence. We use $\mathcal{S}$ to denote the set of all sequences and M to represent the total number of sequences in $\mathcal{S}$. However, the difference in the number of interactions between users and smart contracts will lead to the difference in the length of transaction flows, which will make it difficult to measure the similarity of vectors. To solve this problem, we define k as the number of consecutive elements, p to represent the padding number, and t as the distance between two k consecutive elements. On the basis, we formulate all possible combinations of k consecutive elements with distance t after padding as $\mathcal{V}_{k,p,t}$. Next, we count the normalized frequency of each $\gamma_n \in \mathcal{V}_{k,p,t}, n \in 1, 2, \cdots, N$ within each sequence S_m as array $[c_{m,1},\ c_{m,2},\ \cdots\ ,\ c_{m,n}]$. After extracting the features of LPs' behaviours to format the vector of a particular address, we choose Polar distance over other alternatives (e.g., Euclidean distance) because Polar distance is more suitable to handle the highly sparse matrix [3132].

$$d(S_1, S_2) = \frac{1}{\pi} cos^{-1} \frac{\sum_{j=1}^{n} c_{1,j} \times c_{2,j}}{\sqrt{\sum_{j=1}^{n} (c_{1,j}^2)} \times \sqrt{\sum_{j=1}^{n} (c_{2,j}^2)}} \tag{1}$$

$d(S_1,\ S_2)$ ranges from 0 to 1, and a small distance value indicates a high similarity between two transaction flows. The parameters chosen in our method are fundamental to the results of the cluster. Intuitively, if we choose a larger k, we can obtain longer operation sequences, which are unlikely to repeat as a feature. Furthermore, the length of the feature vector increases exponentially with k. Similarly, the chosen t will lead to various features resulting in the length of the feature vector. Moreover, we use the padding method to paddle the operation sequence S_m because the last operation does not have the time gap with the next operation. Hence, we use the end date of the data collection as the date of the next operation. To ensure the integrity of the operation sequence, we choose $k = 3$, $t = 3$, and paddling number $p = 1$ in our method.

5.3 Clustering

The algorithm used is divisive hierarchical clustering (DHC) [33], which is suitable for arbitrary metric space to find clusters of arbitrary shapes. Specifically, DHC starts with one group containing all the items and repeatedly splits aggregates until a specific level of granularity has been reached, i.e., clustering quality reaches a minimal threshold, and format a tree hierarchy of behavioural clusters. We use the average link to measure cluster results in DHC. After obtaining the cluster results, we can infer the meaning of the clusters based on the LPs' features. The base for the selection of features is whether they contribute to the differentiation of the clusters in terms of their corresponding aspects. It can therefore be used as a reason for the formation of clustering, as well as an explanation for the content of clusters.

6 Dataset analysis

This section investigates the effect of governance tokens in liquidity mining from macroscopic temporal data and LPs' behaviours. We first statistically analyzed the TVL of Uniswap and SushiSwap, the call of different functions and the address activity in combination with significant events in the timeline mentioned in Section 2.3. Moreover, we show the result of the unsupervised clustering method suggested in Section 5, investigate the LPs' behaviours in different clusters, and obtain the following observations.

- **Observation 3.** Adding governance tokens as a reward in liquidity mining can spike the number of TVL and LP in a short period. However, it is not a particularly effective measure in the long run.
- **Observation 4.** LPs who offer less liquidity tend to be less active in DEXs, while most active LPs are not those providing the heaviest liquidity in DEXs but some medium LPs who frequently add and remove liquidity to participate in liquidity mining of multiple protocols to earn governance tokens. In contrast, Heavy LPs tend to pursue a long-term trading fee return and are less affected by other external factors.

6.1 TVL

Figure 5 shows the TVL in USD of Uniswap and SushiSwap from August 2020 to 2021. There are several noteworthy features in the graph, which are listed in chronological order as follows: the first being the period marked as *Vampire Attack*. The Vampire

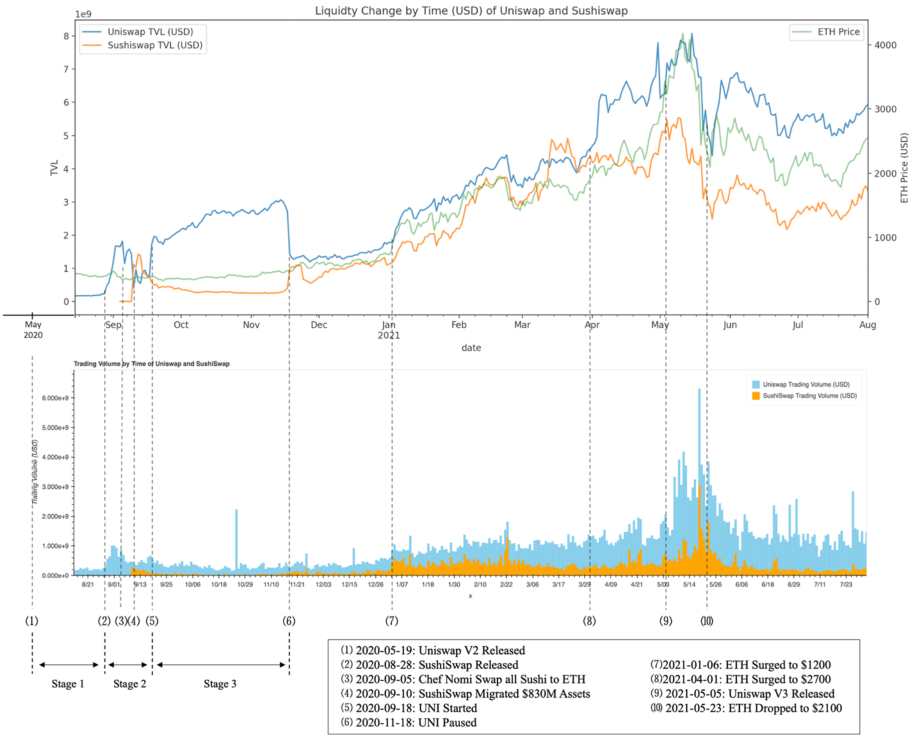

Figure 5 TVL and Daily Trading Volume in Uniswap and SushiSwap

Attack of SushiSwap resulted in Uniswap's liquidity dropping from around $300 million to almost $2 billion and reducing to around $500 million in a matter of days, which is still at a higher point than a week ago. The trading volume also remained strong at around $300-800 million per day. Therefore, we can infer that, despite the incentive of governance tokens in liquidity mining from SushiSwap, some LPs still have confidence in the trading volume of Uniswap, so they choose to provide liquidity in pursuit of commission income.

On September 18, 2020, when Uniswap launched its counterattack, its TVL quickly surpassed and continued to overwhelmingly outperform SushiSwap until November 18, 2020, when Uniswap announced that UNI was discontinued and a large number of Uniswap's LPs switched to SushiSwap. On the day of the UNI launch alone, this event brought in $1.65 billion TVL for Uniswap, while on the same day, SushiSwap lost $159 million TVL comparatively. Over two months, Uniswap's TVL continued to climb until it peaked at $3.06 billion on November 14. After that, with only three days left till the scheduled closure date, some LPs began to remove TVL, causing a slight and sustained decline. On November 18, Uniswap's TVL plummeted by $1.29 billion, while SushiSwap's TVL rose by $578 million, creating another 'Vampire Attack'-liked pattern.

After November 18, 2020, the TVL of Uniswap and SushiSwap tend to follow the ETH price fluctuations. The increase in TVL at (7) *2021-01-06* and (8) *2021-04-01* is accompanied by a surge in ETH price, which can be explained as the ETH price as the dominant currency and its exchange rate directly affects the TVL settled in USD; while, from a market perspective, the appreciation of ETH may lead to a more active blockchain market, implying more frequent token trades, from which liquidity providers can then reap more benefits.

As the trading volume graph in the bottom half of Figure 5 illustrates, the fluctuations in the trading volume are not especially and precisely related to selected events but have a certain correlation with ETH price. However, the change in trading volume implies a certain correlated consequence in the LP's revenue from the commission, and thus, the governance token plays a crucial role in the incentive for SushiSwap to be so comparable to Uniswap in terms of TVL, even though its trading volume is consistently lower than that of Uniswap.

6.2 Function type called

We obtain the time series of the five functions, 'ETH-Token', 'Token-ETH', 'Token-Token', 'Add Liquidity' and 'Remove Liquidity', by counting the number of calls in days from the labeled data. As shown in Figure 6, each cell of the heat map represents a day, and the shade of the color represents the number of times the corresponding function was called that day. The curves overlay on the heat map represents the average number of calls to the five functions.

From Figure 6, we notice that LPs' activity is followed by governance token issuance tends to last 3-4 days. Specifically, the announcement of the SUSHI reward increased the average number of function calls in Uniswap, especially the 'add liquidity' function. After Vampire Attack, the call of the 'add liquidity' function maintains a high value for several days, but the call of 'remove liquidity' function also increases. There are two possible reasons behind this phenomenon: First, the SUSHI reward decreases over time, which leads to some LPs removing their liquidity; Second, Uniswap's counterattack played out. This pattern is consistent with their intense competition policy in early September.

Besides, the five functions tend to surge simultaneously in particular periods, such as the Uniswap heat map in September 2020 or the SushiSwap heat map in January 2021, where we can reach a vertical dark pattern across the five rows. From the rows, there are fewer function calls related to Liquidity than the Swap class functions. For Uniswap, the function to exchange Token from ETH is generally darker than the other lines, which to some extent indicates the primary purpose of the user's interaction with Uniswap.

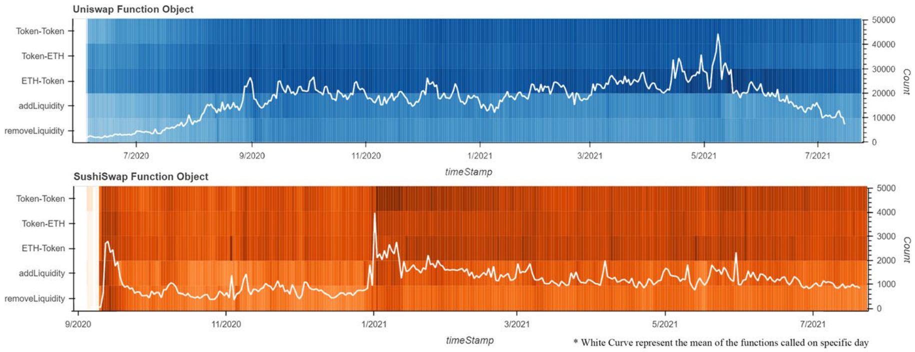

Figure 6 Function Time Series of Uniswap and SushiSwap

6.3 New user and LPs

We extract from the transaction records of both DEXs addresses the timestamp of their first interaction with the DEXs and their first operation on liquidity. We aggregate these timestamps to plot the results in Figure 7. Although there are cases where an address first interacts as a trader and liquidity are provided sometime later, this does not affect our ability to observe the macroscopic temporal characteristics of new users or LPs by grouping them by date. It is worth noting that Uniswap and SushiSwap possess a huge difference in volume, which leads to a nearly tenfold discrepancy in the vertical coordinates in Figure 7, but this does not prevent us from obtaining useful information from the trend.

By comparing the positions of the two curves of *New User* and *New LPs*, we can infer the motivation of new users to use either Uniswap or SushiSwap during a specific period. Looking at the curve of Uniswap, we can see that the two curves do not tend to increase or decrease in the same way, which can be explained by the considerable volume of Uniswap in Figure 6: most users use Uniswap for trading purposes. In early September 2020, there are two significant peaks in the New LPs interval, which is in line with *Vampire Attack* stage. In sharp contrast to Uniswap, the image of SushiSwap shows two curves that converge until January 2021, which suggests that most users of SushiSwap are motivated by gaining SUSHI by providing liquidity rather than trading.

In addition, by comparing the two exchange curve peaks with the corresponding event time points in Figure 5, we can find the fluctuating patterns in September,

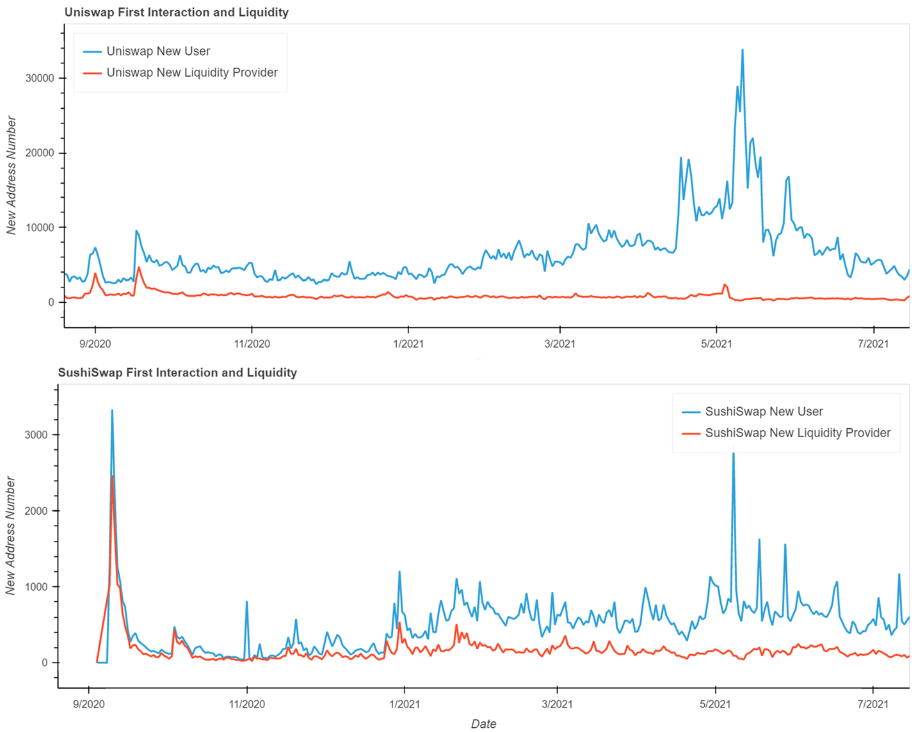

Figure 7 New Users and LPs of Both DEXs by Date

November 2020, and January, May 2021. Thus, special events like a surge in Ethereum price or policies have an incentive effect on users' behavioural patterns and motivations, which can be reflected in the attraction of new participants. For example, after January 2012, as the price of ETH went up, the number of new users per day for Uniswap and SushiSwap also went up, peaking in May 2021 and then gradually declining and leveling off. This type of macro blockchain market-related fluctuation is not highly correlated with the DEXs' own service policies and user benefits.

6.4 Cluster results analysis

The on-chain activity of address suggests that LPs interact with DEXs with different aims. We generated 6 clusters for Uniswap and SushiSwap using the method presented in Section 4.3, respectively.

Dispensable LPs refer to those addresses that have provided liquidity but lack the enthusiasm to participate, addressing 58.1% and 17.2% in Uniswap and SushiSwap, respectively, who maintain a low level of TVL less than $1,000 for a long time, with larger intervals between operations. We presume this group of addresses only wants to try liquidity mining but do not want to invest too much.

Light LPs represent the addresses that have provided more liquidity than dispensable LPs, which account for 18.0% and 42.2% in Uniswap and SushiSwap, respectively. Light LPs can be classified by two clusters according to their operation frequency, namely *inactive light LP* and *active light LP*. Active light LPs interact more frequently with DEXs, especially in Uniswap. Although there is a difference between the proportions of light LP and dispensable LP in two DEXs, the sum of their proportions is both more than 59%, which shows that LPs providing a small amount of liquidity make up the majority in both DEXs. On top of that, the higher proportions of SushiSwap light LPs show that providing long-term governance token rewards incentivizes more sandy users to participate in liquidity mining.

Medium LPs are the addresses that provide higher liquidity in DEXs, addressing 19.9% and 29.7% in Uniswap and SushiSwap, respectively. Medium LP is further divided into *risk-seeking medium LP* and *risk-averse medium LP*. As shown in Figure 8, two kinds of medium LPs have similar liquidity distribution values but are diverse in the number of operations. Specifically, risk-averse medium LPs are more cautious or lack enthusiasm in participation because of the potential loss. However, risk-seeking medium LPs are keen to provide liquidity to obtain high APY governance token rewards in more DEXs, though they may suffer from impermanent loss even the risk of stolen funds.

Heavy LPs take up the fewest proportion in two DEXs, addressing 4.6% and 11% in Uniswap and SushiSwap, respectively. They provide high liquidity in DEXs, for example, address *0xf0fc* has provided more than 14,000,000 USDC and 7,861 ETH in Uniswap, which was worth nearly $ 30 million at that time. Besides, the operation frequency of this group is smaller to risk-seeking medium LPs, which shows that they tend to pursue a long-term commission return, and their decisions are less affected by other external factors.

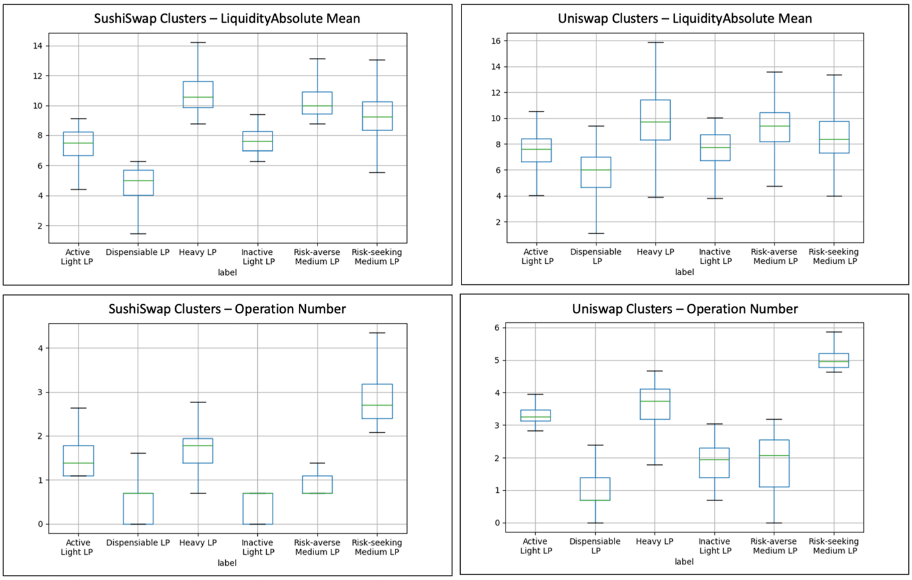

Figure 8 Logged Liquidity Absolute Means and Liquidity Operation Times for SushiSwap (Left Column) and Uniswap (Right Column) Cluster Results

6.5 Behaviour of LPs in uniswap

In this subsection, we collect the list of Uniswap LP addresses that operated liquidity during specific periods or invoked special contracts, starting on August 28 and ending on November 18 when Uniswap stopped UNI rewards, and plot the data in the flow diagram as shown in Figure 9.

Due to the launch of SushiSwap's Masterchef contract for pledging UNI-V2 tokens to earn SUSHI, new addresses were entered into Uniswap to provide liquidity to acquire UNI-V2 tokens. In just ten days, Uniswap gained as much LP increment as the previous three months combined. Among these new LPs, Heavy and Medium LP have a more significant proportion than the return LPs. We found that more than half of the LPs that provided liquidity between August 28 and September 8 pledged the UNI-V2 tokens they received to MasterChef to earn additional SUSHI, while less than 5% of the return LPs made this move. We speculate that the reason for this phenomenon is that there is a more significant proportion of Dispensable LPs among the return LPs, which tend to be inactive and provide less liquidity in DEXs. As a result, they may be less sensitive to market information and not interested in SUSHI because they invest less capital and do not have access to a great amount of SUSHI.

Moving to the next phase, we note that nearly half of the addresses earning SUSHI removed liquidity in the following two months. Based on the timing of the removal, we divided it into two periods: *A. September 9 to September 17*; and *B. September 18 to November 18*. First, 26.7% of LPs choose to remove liquidity in *Period A*. 57.7% of them are Medium and Heavy LPs. We believe that this group is the most market-sensitive and penny-pinching about liquidity return, and they only provide liquidity during the period with higher APY. Secondly, 21.2% of addresses chose to remove liquidity from SushiSwap

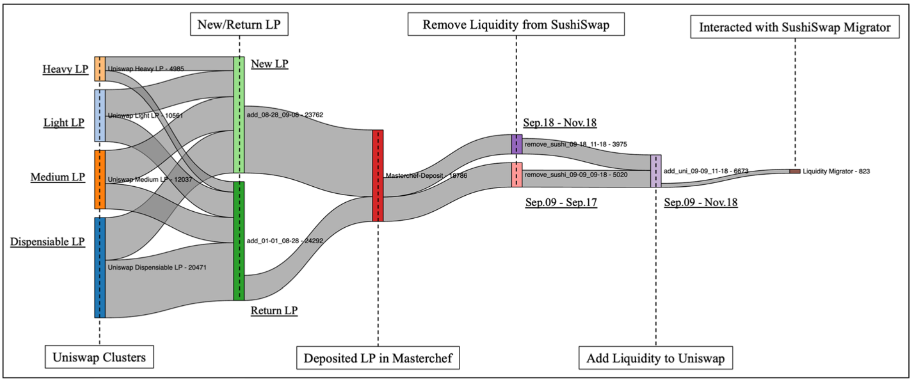

Figure 9 The Flow Chart of LPs in Uniswap from Sep 09, 2020 to Nov 18, 2020

at *Period B*. We hypothesize that the reason behind this is that Uniswap released UNI and turned on liquidity mining on September 18, when the trading fee revenue from Uniswap's enormous daily volume and higher APY quickly took back part of the LP from SushiSwap. Finally, after UNI stopped issuing on November 18, only a very few addresses in this group chose to migrate liquidity to SushiSwap by calling the Migrator contract.

6.6 Overlap address

As mentioned in Section 4.2, Uniswap and SushiSwap have 297,345 and 43,705 independent addresses that have provided liquidity, respectively, from which we can get 27,521 overlapping addresses that provide liquidity to both exchanges. This section will discuss the composition of the overlapping addresses in the context of the cluster results.

Figure 10 depicts the percentage of overlapping addresses in the SushiSwap cluster in the short and long term, respectively. Recall from Figure 5 that in early September 2020, SushiSwap performed the Vampire Attacks, but this advantage quickly disappeared after

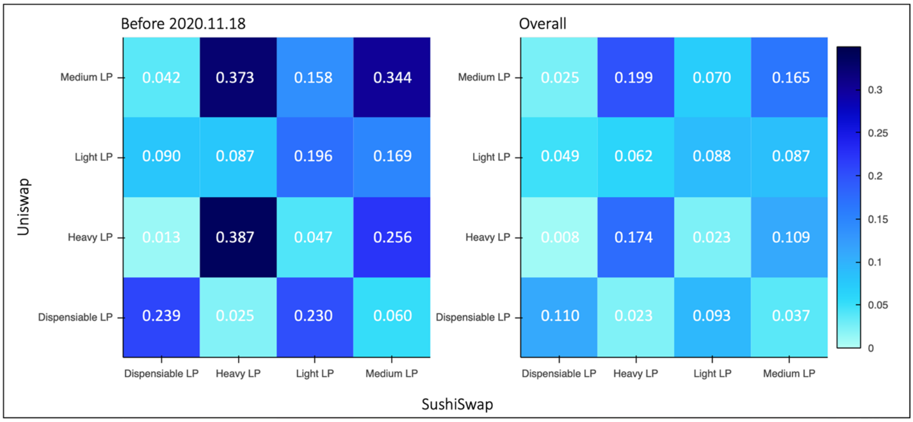

Figure 10 Percentage of overlapping clusters in the corresponding SushiSwap-clusters from 2020.09.08 to 2020.11.18, and 2020.09.08 to 2021.07.18

Uniswap turned on liquidity mining and UNI. Combining this with the heat map shown on the left side of Figure 6 through November 2020, we can see considerable overlap between Uniswap and SushiSwap liquidity providers during the first two months of SushiSwap's existence. Specifically, 87.2% of SushiSwap's heavy LPs had overlapping addresses, in which 76% were classified as medium and heavy LPs for Uniswap. However, almost a year after the release of SushiSwap, the data as of July 18, 2021, shows that the share of overlapping addresses fell, in which SushiSwap heavy LP's overlap rate dropped by 41.4%. Therefore, we can obtain that SushiSwap has gradually developed dedicated fixed LP groups over almost a year through the long-term SUSHI incentive.

7 Conclusion

- We confirm that adding governance token in liquidity mining for different types of LPs have different appeals.
- Based on the clustering results, with the SushiSwap example, we can see that more than fifty percent of Heavy and Medium LPs will withdraw their funds in a short period. We conclude two reasons behind this phenomenon: 1) high rewards from competitors (Uniswap) and 2) the APY of the rewards decreases over time. Therefore, the reward in the form of adding governance tokens to the liquidity mining does not work well in the early stages of the protocol.
- We reveal that the most active participants in liquidity mining are not those with the most capital but the addresses with moderate funding. This phenomenon is because this group of addresses has the strongest subjective motivation for pursuing benefits and rewards, so we infer that incentives such as governance tokens should have the greatest appeal to them.

To wrap up from a high-level and macroscopic perspective, all these raised concerns about the impact of the governance token show that it has not served the governance value for which it was designed but has been used as an arbitrage tool by relatively well-funded speculators without a positive impact on the development of the protocol. However, at the same time, we note that the dereliction of duty in governance capabilities has not undermined its attractiveness to users. By comparing the overlapping LP ratios of SushiSwap and Uniswap in two phases, we find that SushiSwap has gradually gained its stable users and scale through long-term governance tokens issuance and continuous expansion of business innovation.

8 Future Vision

Adding governance tokens in liquidity mining is often used by protocols to attract users. Based on our analysis results, we conclude that this method can attract more users in a short time but cannot retain them. Hence, the protocols should actively seek other ways to attract users, such as reducing the transaction cost and providing more convenient functions.

Besides, governance tokens are the cornerstone of community self-governance, a crucial foundation for any Web 3.0 protocol. As one of the earliest pioneers in distributing governance tokens, liquidity mining encourages LPs to engage in passive participation rather than active contribution. We acknowledge that higher participation requirements such as growing gas fees and adequate liquidity discourage LPs with insufficient funds. Hence, a new way of governance token distribution is critical for the Web 3.0 community. Decentralized communities need to explore new incentive mechanisms that can proactively motivate members to participate in governance. For example, participants initiate proposals related to protocol development, and if the community adopts these suggestions, proposers are rewarded with governance tokens. Future research could explore, model, and evaluate refined governance mechanisms to build a better-decentralized community.

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

Declarations

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

References

1. Fenu, G., Marchesi, L., Marchesi, M., Tonelli, R.: The ico phenomenon and its relationships with ethereum smart contract environment. In: 2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE), pp 26–32. IEEE (2018)
2. Xu, J., Vavryk, N., Paruch, K., Cousaert, S.: Sok: Decentralized exchanges (dex) with automated market maker (amm) protocols. arXiv:2103.12732 (2021)
3. Adams, H., Zinsmeister, N., Robinson, D.: Uniswap v2 core (2020)
4. Lo, Y.C., Medda, F.: Uniswap and the rise of the decentralized exchange Available at SSRN 3715398 (2020)
5. Wood, G.: Ethereum: A secure decentralised generalised transaction ledger (2020)
6. Zou, W., Lo, D., Kochhar, P.S., Le, X.-B.D., Xia, X., Feng, Y., Chen, Z., Xu, B.: Smart contract development: Challenges and opportunities. IEEE Transactions on Software Engineering **47**(10), 2084–2106 (2021). https://doi.org/10.1109/TSE.2019.2942301
7. Swan, M.: Blockchain: Blueprint for a New Economy. O'Reilly Media Inc (2015)
8. Fan, S., Zhang, H., Wang, Z., Cai, W.: Mobile devices strategies in blockchain-based federated learning: a dynamic game perspective. In: IEEE Transactions on Network Science and Engineering (2022). https://doi.org/10.1109/TNSE.2022.3163791
9. Wang, Z., Qin, K., Minh, D.V., Gervais, A.: Speculative multipliers on defi: Quantifying on-chain leverage risks. Financial Cryptography and Data Security (2022)
10. Tolmach, P., Li, Y., Lin, S.-W., Liu, Y.: Formal analysis of composable defi protocols. In: International Conference on Financial Cryptography and Data Security, pp. 149–161. Springer (2021)
11. Fan, S., Zhang, H., Zeng, Y., Cai, W.: Hybrid blockchain-based resource trading system for federated learning in edge computing. IEEE Internet of Things Journal **8**(4), 2252–2264 (2020)
12. Zhang, H., Fan, S., Cai, W.: Decentralized resource sharing platform for mobile edge computing. In: International Conference on 5G for Future Wireless Networks, pp. 101–113. Springer (2020)
13. Nakamoto, S.: Bitcoin: a Peer-To-Peer Electronic Cash System. Technical report, Manubot (2019)
14. Buterin, V., et al.: A next-generation smart contract and decentralized application platform. White Paper 3(37) (2014)
15. Min. T., Cai W.: Portrait of decentralized application users: an overview based on large-scale Ethereum data[J]. CCF Trans. Pervasive Comput. Interact. 1–18 (2022)
16. Cai, W., Wang, Z., Ernst, J.B., Hong, Z., Feng, C., Leung, V.C.: Decentralized applications: The blockchain-empowered software system. IEEE Access **6**, 53019–53033 (2018)
17. Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., Cai, W.: Metaverse for social good: a university campus prototype. In: Proceedings of the 29th ACM International Conference on Multimedia, pp. 153–161 (2021)

18. Min, T., Wang, H., Guo, Y., Cai, W.: Blockchain games: A survey. In: 2019 IEEE Conference on Games (CoG), pp. 1–8. IEEE (2019)
19. Jiang Y, Fan S, Cai W. Economic Analysis of Loot Box Market in Blockchain Games[C]. In: Proceedings of the Fourth ACM International Symposium on Blockchain and Secure Critical Infrastructure, pp. 35–46 (2022)
20. Zetzsche, D.A., Arner, D.W., Buckley, R.P.: Decentralized finance. Journal of Financial Regulation **6**(2), 172–203 (2020)
21. Zhang, H., Fan, S., Fang, Z., Cai, W.: Economic Analysis of Decentralized Exchange Market with Transaction Fee Mining. In: Proceedings of the Fourth ACM International Symposium on Blockchain and Secure Critical Infrastructure (BSCI '22), pp. 59–70. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3494106.3528677
22. Qin, K., Zhou, L., Afonin, Y., Lazzaretti, L., Gervais, A.: Cefi vs DeFi – comparing centralized to decentralized finance (2021)
23. Froewis, M., Sridhar, K., Makridis, C., Böhme, R.: The rise of decentralized cryptocurrency exchanges: Evaluating the role of airdrops and governance tokens Available at SSRN (2021)
24. Chen, Y., Bellavitis, C.: Blockchain disruption and decentralized finance: The rise of decentralized business models. Journal of Business Venturing Insights **13**, 00151 (2020)
25. Angeris, G., Kao, H. -T., Chiang, R., Noyes, C., Chitra, T.: An analysis of uniswap markets. arXiv:1911.03380 (2019)
26. Aigner, A.A., Dhaliwal, G.: Uniswap: Impermanent loss and risk profile of a liquidity provider. arXiv:2106.14404 (2021)
27. Krishnamachari, B., Feng, Q., Grippo, E.: Dynamic Automated Market Makers for Decentralized Cryptocurrency Exchange. In: 2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 1–2. https://doi.org/10.1109/ICBC51069.2021.9461100 (2021)
28. Wang, Y., Heimbach, L., Wattenhofer, R.: Behavior of liquidity providers in decentralized exchanges. arXiv:2105.13822 (2021)
29. Lee, X.T., Khan, A., Sen Gupta, S., Ong, Y.H., Liu, X.: Measurements, analyses, and insights on the entire ethereum blockchain network. In: Proceedings of The Web Conference 2020, pp. 155–166 (2020)
30. Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H., Zhao, B.Y.: You are How You Click: Clickstream Analysis for Sybil Detection. In: 22Nd {USENIX} Security Symposium ({*USENIX*} Security 13), pp. 241–256 (2013)
31. Houle, M.E., Kriegel, H.-P., Kröger, P., Schubert, E., Zimek, A.: Can shared-neighbor distances defeat the curse of dimensionality?. In: International Conference on Scientific and Statistical Database Management, pp. 482–500. Springer (2010)
32. Wang, G., Zhang, X., Tang, S., Zheng, H., Zhao, B.Y.: Unsupervised clickstream clustering for user behavior analysis. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 225–236 (2016)
33. Elmqvist, N., Fekete, J. -D.: Hierarchical aggregation for information visualization: overview, techniques, and design guidelines. IEEE Trans. Vis. Comput. Graph. **16**(3), 439–454 (2010). https://doi.org/10.1109/TVCG.2009.84