Characterizing Common Quarterly Behaviors In DeFi Lending Protocols

Aaron Green[1], Michael Giannattasio[1], Keran Wang[1], John S. Erickson[1], Oshani Seneviratne[1], and Kristin P. Bennett[1]

The Rensselaer Institute for Data Exploration and Applications
Rensselaer Polytechnic Institute
http://idea.rpi.edu
{greena12,giannm,wangk16,erickj4,senevo,bennek}@rpi.edu

Abstract. The emerging decentralized financial ecosystem (DeFi) is comprised of numerous protocols, one type being lending protocols. People make transactions in lending protocols, each of which is attributed to a specific blockchain address which could represent an externally-owned account (EOA) or a smart contract. Using Aave, one of the largest lending protocols, we summarize the transactions made by each address in each quarter from January 1, 2021, through December 31, 2022. We cluster these quarterly summaries to identify and name common patterns of quarterly behavior in Aave. We then use these clusters to glean insights into the dominant behaviors in Aave. We show that there are three kinds of keepers, i.e., a specific type of users tasked with the protocol's governance, but only one kind of keeper finds consistent success in making profits from liquidations. We identify the largest-scale accounts in Aave and the highest-risk kinds of behavior on the platform. Additionally, we use the temporal aspect of the clusters to track how common behaviors change through time and how usage has shifted in the wake of major events that impacted the crypto market, and we show that there seem to be problems with user retention in Aave as many of the addresses that perform transactions do not remain in the market for long.

1 Introduction

Decentralized Finance (DeFi) is an emerging economic ecosystem built on blockchain technologies and smart contracts. The primary feature differentiating DeFi from traditional finance is the lack of intermediaries controlling financial services. DeFi's growth in the last half-decade has started with the creation of protocols that seek to replicate the services of traditional financial institutions. For example, a popular kind of DeFi protocol is the lending protocol, which mimics the functionality of a bank from traditional finance.

Lending protocols in DeFi give users the opportunity to lend their crypto assets to a lending pool, effectively acting as a savings account. Users can then borrow crypto assets from the lending pool based on how much they have contributed to the pool themselves. Thus, the lending pool is inherently over-collateralized. Borrowed assets will accrue interest on the loan, and a portion of that interest is paid back to the lenders of the assets. In this way, lenders can accrue "deposit" interest on the money they put into the protocol. If a borrower's collateral loses too much value or the loan accrues too much interest, so their account no longer meets the minimum collateral requirements set by the protocol, their account is subject to being liquidated. In this case, a third party, called a "keeper," can pay off some of the loans to acquire a portion of the borrower's collateral. The amount of collateral purchased through a liquidation often comes with a small "liquidation bonus" that acts as an incentive for keepers to perform these liquidations and keep the lending pool healthy.

Having decentralized protocols that allow for these actions creates new ways for retail and institutional users to try leveraging their assets for profit. Whether someone wants to accrue passive interest on their crypto assets or whether they want to use the collateralized borrowing enabled by lending protocols to pursue more aggressive, riskier positions, lending protocols represent an important part of the emerging DeFi economy. One of the major lending protocols is Aave.[1] [8,9,13] As of February 20, 2023, Aave has over $6.87 billion of assets locked in its lending pools across two versions and seven markets. At its peak in October 2021, Aave had over $18 billion of crypto assets locked in its lending pool.[2]

Using the transaction data of Aave's V2 [9] deployment on the Ethereum blockchain [10] as the primary subject of this work, we summarize quarterly user and smart-contract transactions on the platform. We then cluster these summaries to identify the predominant types of behaviors in Aave. We use these clusters to present novel insights about the behavior of keepers, the trends of the largest accounts (likely owned by financial institutions), and the highest risk behaviors on Aave. We also track how the clusters change over time and in response to some of the major events that had affected the crypto market, such as when China announced an impending crypto ban in May 2021 [1] and the November 2022 crash of major crypto exchange FTX [2]. Finally, we examine related work in the field and discuss the potential impacts of this analysis and how this can be used to help future work.

2 Methods

2.1 Data Sources

We examine DeFi behavior at the address level since addresses are the equivalent of accounts in traditional finance. Every DeFi transaction is associated with an address which may be either an externally-owned account (EOA) or a smart-contract. To understand user behavior in a quarter, we cluster addresses. thus we need data that summarizes the behavior that has been associated with each address in a quarter. This is different from the raw data that we collect, which is transaction-level data. We briefly describe the transaction-level data, and then explain how we convert this data into address-level summaries. To distinguish

[1] `aave.com`

[2] see `https://defillama.com/protocol/aave-v2`

between EOA's and smart-contacts, we obtained blockchain address data from Amberdata[3] for each address present in our data to classify them as either an EOA or a smart contract.

Code used to collect these data can be found on a public GitHub repository.[4] The transaction-level data from TheGraph is freely accessible, but any data acquired from Amberdata requires one's own API key. The code used to query Amberdata is included, but with no API key.

2.2 Transaction-Level Data

The transaction-level data we use in this analysis comes from The Graph.[5] We use data from the lending protocol Aave, which consists of six transaction types: deposits, collaterals, withdraws, borrows, repays, and liquidations. Liquidation transactions involve two addresses: one associated with the keeper and the other with the liquidatee. Because performing liquidations and being liquidated are both interesting, we duplicate the liquidation transactions and for the two copies, treat the liquidator as the subject in one and the liquidatee as the subject in the other. This effectively turns the liquidation transactions into "liquidations performed" transactions and "liquidated" transactions. Each transaction has data regarding the time the transaction occurred, the address that initiated the transaction, the asset(s) involved in the transaction, the amount of the asset(s) used in the transaction (in both native amounts and USD amounts), and any relevant third-party addresses like the liquidatee. We use all transactions that occurred between January 1, 2021 and December 31, 2022, for a total of two years of transaction data. This amounts to 1,665,737 total transactions involving 172,872 unique blockchain addresses.

2.3 Address-Level Summaries

From this transaction data, we group the transactions by the address that initiated the transaction and the quarter during which the transaction took place. For each groups of address-quarters, we create the summary features in Table 1.

Feature Name	Description
Smart Contract?	Binary value with 0 indicating that the address is an EOA and 1 indicating that the address is a smart contract.
Number of Withdrawals	Number of withdraw transactions this account performed during the quarter.

[3] amberdata.io
[4] https://github.com/aaronmicahgreen/Characterizing-Common-Quarterly-Behaviors-In-DeFi-Lending-Protocols
[5] thegraph.com

Value of Withdrawals	Value (in USD) of assets this address withdrew from its account during the quarter.
Number of Repayments	Number of repay transactions this address performed during the quarter.
Repaid Value	Value (in USD) of the assets this address repaid during the quarter.
Percentage of Stable Borrows	Percentage of this address' borrows that used the stable borrow rate during the quarter.
Number of Borrows	Number of borrow transactions this account performed during the quarter.
Borrowed Value	Value (in USD) of the assets this address borrowed during the quarter.
Mean Value per Transaction	Mean amount (in USD) of each transaction made by this address during the quarter.
Number of Deposits	Number of deposit transactions this account performed during the quarter.
Deposited Value	Value (in USD) of assets this address deposited into its account during the quarter.
Number of Collateral Transactions	Number of collateral transactions this account performed during the quarter.
Number of Days Active	Number of days in the quarter during which the address posted at least one transaction.
Number of Transactions	Number of transactions performed by this address during the quarter.
Number of Liquidations Performed	Number of liquidation transactions that this address performed in the quarter.
Liquidations Performed Value	Value (in USD) of assets that this address liquidated during the quarter.
Number of Times Liquidated	Number of transactions that liquidated this address during the quarter.
Liquidated Value	Value (in USD) of assets that were owned by this address and that were liquidated during the quarter.
New User?	Binary value with 0 indicating that the address has made transactions in a prior quarter and 1 indicating that the address is new this quarter.
Number of Active Collateral Assets	Number of unique assets posted as collateral by this address at the end of the quarter.

Table 1: Address-quarter summary features derived from transaction-level data used for clustering.

2.4 Computation of Clusters

To cluster this data, we use a fuzzy c-means algorithm [7]. We use the R Programming language [22], and compute the clusters using the `ppclust` package [11]. To select the number of clusters, we combined the results from two methods: the elbow method [27] and the Silhouette scores [24]. The elbow method indicated that 5 or 8 clusters would work well, and the silhouette score for 2 and 8 clusters were the highest. For these reasons, we chose to use 8 clusters. Since the transaction data tends to be heavy tailed, we scaled the data using the `lambertW0` function from the "lamW" package [5] prior to clustering.

3 Results

The discovered clusters of quarterly behavior show quite distinct traits. A heatmap of the features within the clusters can be seen in Figure 1. This row-scaled heatmap was created in R using the `pheatmap` package [17]. In this heatmap, darker red features indicate that the feature has higher values within that cluster compared with other clusters, and vice versa with darker blue features. For clarity, the "Smart Contract?" feature will appear darker red when the given cluster is composed of a higher proportion of smart contract addresses, and the "newUser?" feature will appear darker red when the cluster is composed of a higher proportion of addresses that are new in the quarter.

3.1 Interpretations of Clusters

Inspecting the properties of each cluster, we provide names and interpretations of each cluster in Table 2. The number of address-quarter behaviors that fit into each cluster is also given, which shows a fairly balanced spread of behavior across Aave's history. We give justification of these names and interpretations here, making use of Figures 1, 3, and 2, as well as some numerical data when we feel that raw numbers help clarify a point.

We use the term "retail" to describe behavior that we believe is performed by individual users or their smart contracts on a smaller scale. We use this term in a similar way to the term "retail investor" in traditional finance, which is used to describe non-professional investors trading through brokerage accounts or their own savings. This is in contrast to "institution" accounts, which we hypothesize to be addresses that are run by professional organizations such as banks or hedge-funds and are transacting with significantly higher amounts than retail addresses.

We chose the name "Whales" for C1 because this cluster has the highest proportion of smart contracts (19.38%), and the average amount per transaction in this cluster is the highest ($1.19 million per transaction, compared to the overall average of $564,129). To pair with huge transactions, the addresses in this cluster also make the most transactions, making an average of 33.12 transactions per quarter, which is more than double the average of the next highest cluster.

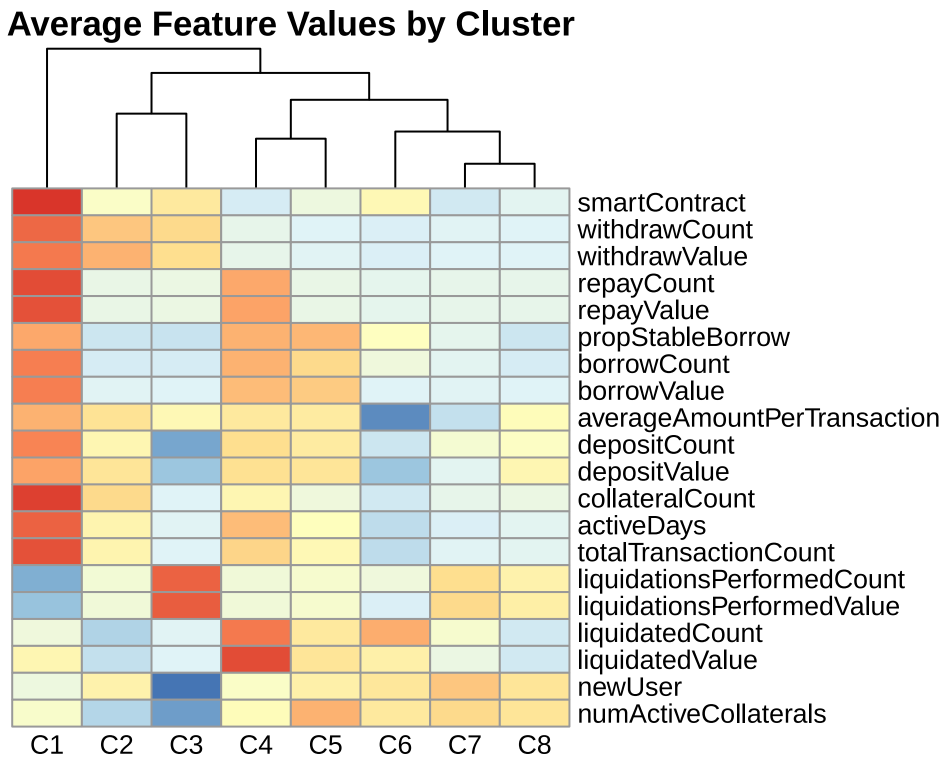

Fig. 1: Heatmap showing the relative values of features across clusters of behavior associated with individual addresses within quarters.

We choose the name "Retail Savers" for C2. Figure 1 shows this cluster is dominated by its deposits and withdrawals. These accounts perform few-to-no borrows (and likewise, few-to-no repays), and also have the lowest likelihood of any cluster for being liquidated. These factors all indicate that this cluster contains behaviors reminiscent of retail accounts that are either seeking to deposit their crypto-assets into the lending pool to accrue deposit interest, or withdraw their assets to exit the lending pool.

We chose the name "Experienced Keepers" for C3. This cluster accounts for the vast majority of the liquidations performed and the high-value liquidations, and the C3 addresses have done little else. This cluster has the second-highest proportion of smart contracts (13.05%), and 83.77% of the transactions performed by this cluster were by smart contracts. which was expected for keepers because performing liquidations needs to be done through contracts (see [8]). This cluster contains very few new addresses, so most of these keepers have been active prior to the quarter in which they fall into this cluster.

We choose the name "Highest Risk Behavior" for C4 because this cluster's most notable characteristic is the number of times its accounts are liquidated,

Table 2: Names, counts, and descriptions of clusters of quarterly behaviors in Aave from January 1, 2021 - December 31, 2022.

Cluster Number	Count	Name	Description
C1	15,291	Whales	High activity, high-value contracts that emphasize creating and re-balancing complex but safe positions.
C2	16,842	Retail Savers	Medium volume and value deposits and withdraws within a single quarter.
C3	18,435	Experienced Keepers	Knowledgeable, high-value users and their contracts that find profitable liquidation opportunities
C4	14,060	Highest-Risk Behavior	EOAs with high amounts of borrows and repays and whose accounts were liquidated the most.
C5	14,024	Yield Farming	A mix of contracts and EOAs depositing and borrowing mid-value positions.
C6	34,082	Inactives	Low-to-no-activity EOAs closing their lingering positions through repayments or being liquidated.
C7	33,803	New Keepers	New accounts with low activity overall but higher-than-average liquidations performed.
C8	20,093	Retail Keepers	EOAs who try to find available liquidations for small profits

as well as the value of those liquidations. Addresses in this cluster are far more likely to be liquidated than in any other cluster. They also tend to borrow high amounts without depositing very much, which is exactly the kind of behavior that is expected to lead to liquidations.

We chose the name "Yield Farming" for C5 because this cluster has high proportions of borrows and deposits, which is reminiscent of how yield farmers leverage deposited assets to repeatedly borrow and deposit assets in order to accrue higher amounts of interest on their deposits. Leveraging assets in this way also increases the riskiness of these accounts' positions, making them more likely to be liquidated in the event of higher-than-expected price volatility with the currencies they have used. Predictability is key when creating high-risk positions, so borrowing with stable borrow rates allows for better control over their positions, and indeed this cluster has a high proportion of stable-interest-rate borrows.

We choose the name "Inactives" for C6 because the number of transactions that these addresses perform in a quarter is the lowest across all clusters. Addresses in this cluster only perform an average of 2.67 transactions in the quarter. Additionally, these addresses are making deposits the least of any transactions, and their accounts are liquidated with relatively high frequency. This cluster contains an above-average number of new users, too. It likely represents both addresses of users who just did a couple of small transactions to test out the functionality of the platform and also accounts that are passively accruing interest. With the higher number of times being liquidated, some of these accounts

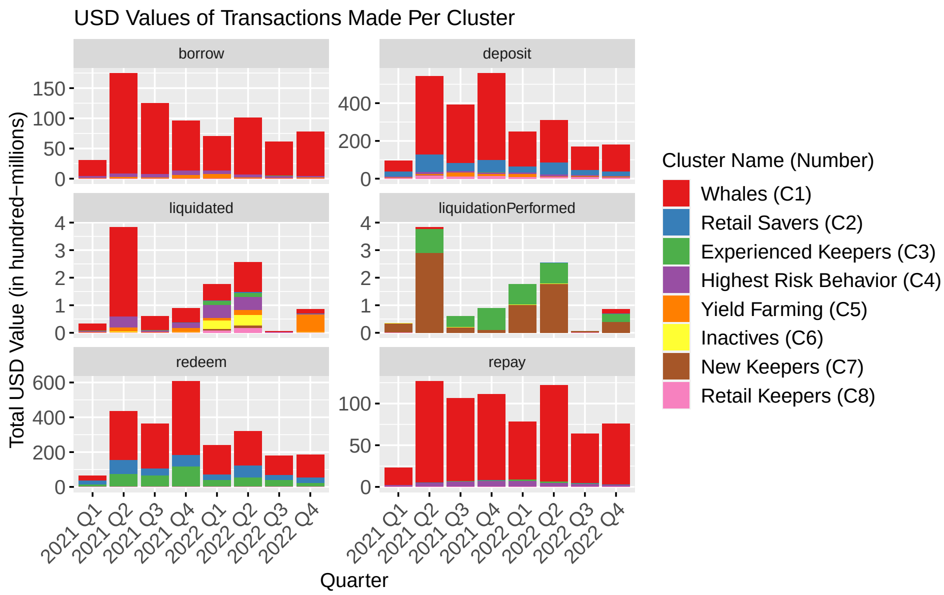

Fig. 2: Set of bar plots showing the USD value of the transactions each cluster made in each quarter, with a separate bar plot for each transaction type.

are likely holding onto positions that have become unhealthy while the account owners either did not care to rebalance their positions, have been priced out of rebalancing due to high transaction fees [20], or have just done poor jobs of monitoring their account health. Despite this cluster being characterized by low transaction counts, we see from Table 2 that this is the largest cluster over the course of the eight quarters. Because there are so many addresses that fall into this cluster, the number of transactions performed by this cluster is non-negligible, as seen in Figure 3, but these transactions are of such low value that they do not make any impact on the combined transaction values in Figure 2.

We choose the name "New Keepers" for C7 because this cluster's most notable features are the number of liquidations they perform and the value of those liquidations, which leads to the "keeper" label. These addresses also have the highest tendency to be new in the quarter when they are assigned this cluster, leading to the label "new."

Finally, we choose the name "Retail Keepers" for C8 because the most notable features of this cluster are the number of deposits, the value of deposits and the liquidations performed. This cluster is very similar to C7, but the amount and value of the liquidations performed by this cluster are lower, and the number and value of deposited assets are higher. We hypothesize that these accounts are retail accounts that are trying to make some deposit interest and who are occasionally trying to perform some liquidations, but at low success rates. They do not tend to be smart contracts (only 5.18% are smart contracts), which indicates

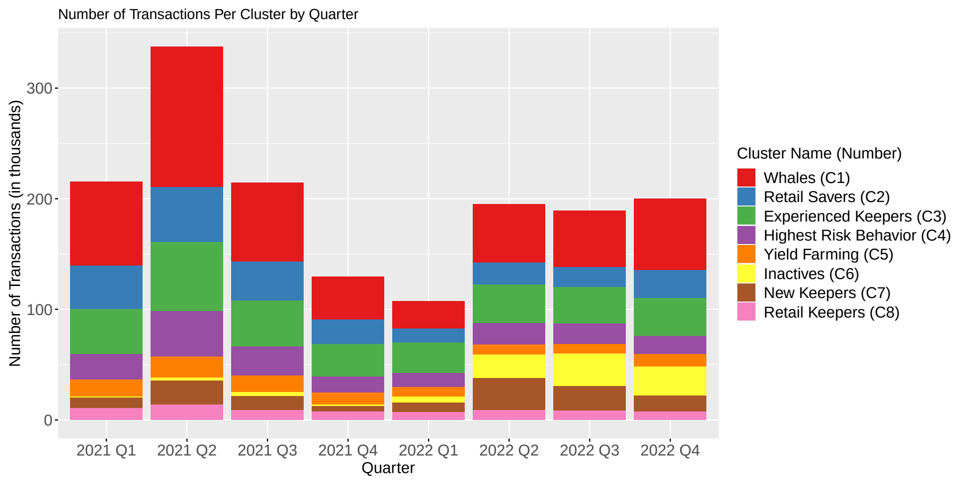

Fig. 3: Bar plot showing the amounts of transactions that addresses in each cluster performed in each quarter from 2021 Q1 through 2022 Q4.

that these are likely users who are searching for potential accounts to liquidate manually instead of in an automated way, and perhaps cannot afford to liquidate larger accounts in their entirety since they do not tend to have high-value accounts.

3.2 Insights Derived From Clusters

Now that we have an understanding of what the clusters mean, we show how these clusters can be used to gain meaningful insights into the DeFi market. We focus on three main insights from these clusters. First, we discuss the "Whales," which are the addresses that account for the largest portion of money spent in Aave. Then we discuss the different clusters of keepers and how they have shifted over time. We conclude with a discussion of the overall trend in Aave usage and how the platform seems to have poor user retention overall.

Risk-Averse Whales: Cluster 1 contains the addresses that are often known as "Whales." This is the cluster that performs the most transactions, and also makes transactions of the highest amounts. We hypothesize that this cluster consists predominantly of accounts owned by various financial institutions. It seems likely that, should an institution such as a bank or hedge fund decide to try building a profitable position in DeFi, they would have a tendency to do so through smart contracts they have written. Additionally, we would expect such an institution to have much more capital than individual users. We might also expect these institutions to be more risk-averse, and this cluster accounts for a low proportion of the total number of times liquidated despite having the largest open positions in the market.

Mueller discusses in [20] the effects of transaction fees on markets in DeFi, and how periods of time with higher transaction fees can affect some users' abilities to react to changing market conditions. Since institutions have access to more capital and are holding onto larger market positions, they are not going to be priced out of re-balancing or closing positions when the market shifts. Additionally, these institutions likely have the most knowledge of how to safely act in financial markets, and have better access to financial data that will allow them to preemptively change their positions to avoid getting liquidated. We see these features represented in this cluster, as the addresses in this cluster are consistently making the highest number of transactions and accounting for the bulk of the USD amount of all transactions in Aave, showing their propensity to be undeterred by transaction fees. However, because these accounts have such large positions, the few liquidations that these accounts experience account for a large portion of the total amount liquidated.

Keepers' Behaviors in Aave: One of the most interesting facets of this clustering is that it gives some key insights about keepers in Aave. With over $69 million in total profits made through liquidations in these two years, it makes sense that there would be many people seeing this opportunity and trying to insert themselves into the pool of keepers who turn big profits. This is why the second-most-populous cluster is the "new keepers" cluster, and the "retail keepers" cluster is the third largest. However, these clusters have a high propensity of only being active for a single quarter. 75.94% of "new keepers" and 55.19% of "retail keepers" are only active for a single quarter. The successful accounts among these two clusters do often end up in the "experienced keepers" cluster in later quarters, with over 15% of, both "new keepers" and "retail keepers" eventually becoming "experienced keepers".

Of these three clusters, the "new keepers" and "experienced keepers" account for the vast majority of the amount of USD that is actually liquidated. The 33,803 accounts in the "new keepers" cluster liquidated accounts worth a combined total of $676 million, and the 18,435 accounts in liquidated accounts worth a total of $387 million total. These numbers are in stark contrast to the "retail keepers" cluster, which does contain accounts that perform liquidations more frequently than most other clusters, but which only total about $2,782 worth of liquidations performed. Despite the large amount of total money that has been available through liquidations in this two-year period in Aave, retail users have not tended to make much money out of liquidations, especially if they are reliant on manually monitoring Aave accounts to find potential liquidations. Combined with the fact that the three clusters we classify as containing keepers contain over 43.4% of all addresses across these two years, this indicates that many DeFi users are more interested in trying to profit off of the risky behavior of other users rather than use DeFi platforms for their other functionality. Since the profits of keepers are not distributed very evenly across those trying to perform liquidations, this is likely a contributing factor to the high amount of user churn we discuss more in the next section.

Cluster Changes Over Time and Issues With User Retention: Through this eight-quarter stretch, the number of addresses in each cluster of behavior per quarter can be visualized with the Sankey plot in Figure 4. Two additional "clusters" have been added to this visualization to help show the flow of address behavior from one quarter to the next: the "future active addresses" cluster (shown in orange) and the "addresses with no new activity" cluster (shown in pink). "Future active addresses" is a strictly-decreasing cluster that contains the addresses that were observed making transactions in Aave during any of the eight quarters, but that has not done so yet. Since every address in our dataset makes at least one transaction, this cluster contains no addresses in 2022 Q4. "Addresses with no new activity" contains addresses that have made transactions in a past quarter, but made no transactions in the quarter in question. Understandably, there are no addresses present in this cluster in 2021 Q1.

Each of the 8 computed clusters shows fairly consistent sizes throughout these eight quarters, with the exception of 2022 Q4. It is interesting to see through this two-year span of Aave how many of the addresses fall into the "addresses with no new activity" group, which would seem to indicate a problem with user retention on the platform. This user-retention problem stems from at least a couple of factors. First, as discussed earlier when examining the various clusters of keepers, there are many accounts that surface in Aave which seem to solely be attempting to break into the liquidation market. The market appears fairly saturated, though, as many of the new keepers do not turn significant profits and end up only acting in a single quarter. The largest liquidation spikes in Aave's history came from two events: the China announcement of an impending ban on cryptocurrencies in May 2021 and the failure of the Terra Luna blockchain in May 2022. We see in figure 2 that the quarters when these events took place (2021 Q2 and 2022 Q2, resp.) correspond to the largest amounts of liquidations. Looking at figure 4, we also see increases in "new keepers" and "retail keepers" cluster sizes in both of these quarters, indicating that the events which cause high amounts of liquidations bring with them an influx of new accounts trying to profit from the liquidations, but that most do so unsuccessfully. Since these accounts make up over 40% of accounts that have transacted within Aave over this time span, these are certainly one driving factor for apparent user-retention problems.

Interestingly, despite the liquidation spike events, they have not caused any significant changes in the quarterly number of addresses that have transacted on Aave. However, there was a significant decline in the number of active addresses in the final quarter of 2022. Every cluster fell in size except for "retail savers" (C2) in 2022 Q2, and the largest cluster is also the least active (C6). This decline in the final quarter of our data is likely due to the sudden collapse of what was one of the largest crypto exchanges, FTX, in November 2022. It will be interesting to continue this analysis in 2023 to see whether the usage of Aave has recovered following this event.

It is also worth noting the general trend of the cryptocurrency market over these two years. For much of 2021, the crypto market was increasing. It reached

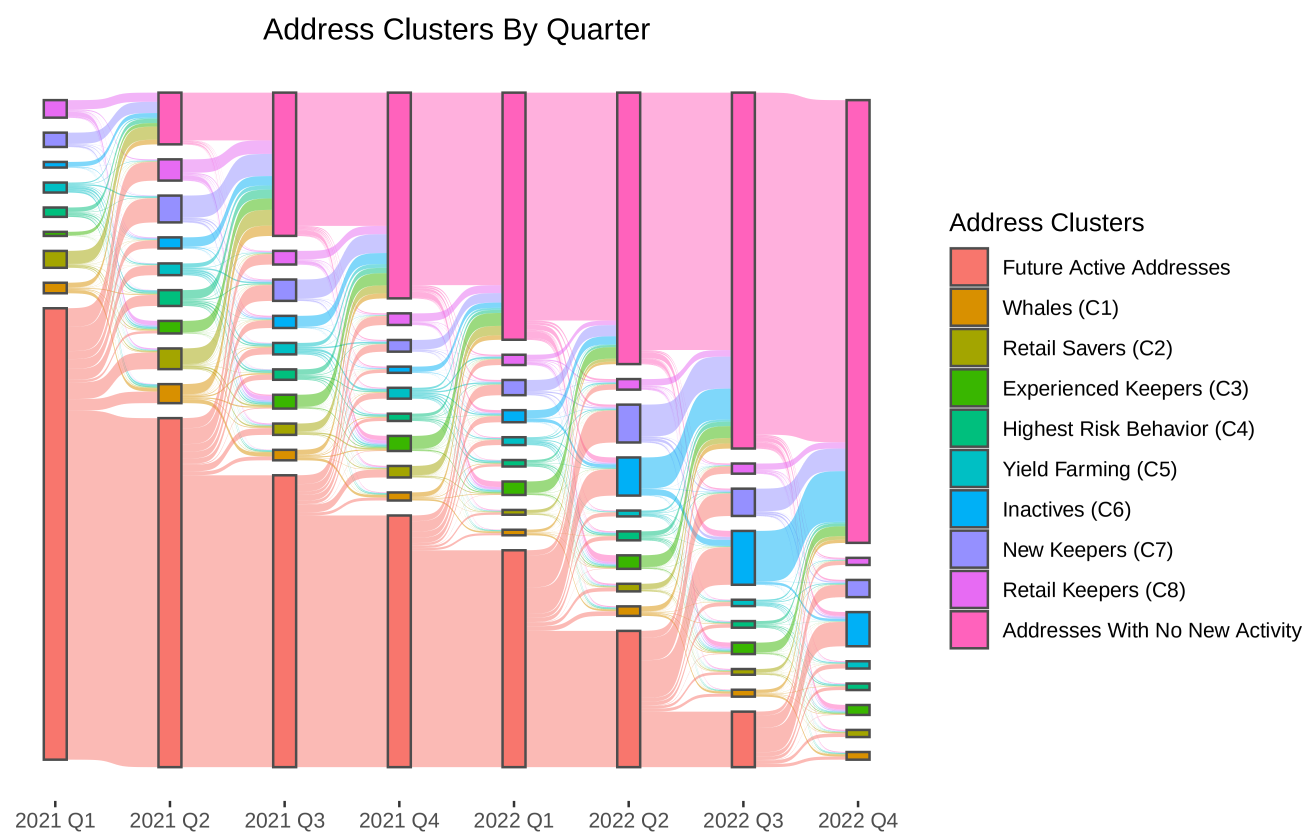

Fig. 4: Sankey diagram showing how addresses changed clusters from quarter to quarter colored by cluster. Orange represents address with activity in later quarters. Pink represents addresses with no activity in that quarter.

an all-time high in early November 2021, which is near when Aave had its highest value locked of over $18 billion. Since then, however, the crypto market as a whole has dipped significantly. By July of 2022, the total market was down about 75% from its peak, and for the remainder of the year, it remained steady around that same level. Considering the usage of Aave did not drop proportionally with the value in the crypto market, this could indicate that there is a market for users who see value in crypto and DeFi beyond simply trying to make a profit, which would be a very positive note.

4 Related Work

There has been other work seeking to characterize user behavior in lending protocols and DeFi. For instance, in Green et. al [14] a process is created for converting the transaction data into a form suitable for the application of survival-analysis methods. This allows for the macro-level analysis of micro-level events, showing how different covariates like whether a borrowed coin is a stable coin affect the time it takes for users to repay the borrowed coin. Such a focus on transaction sequences and the time between events could prove interesting in conjunction with the clustering analysis presented in this paper.

Some work has been done on a small scale to compute other features of account-level data, such as their end-of-day market positions and their end-of-day overall health factor [20]. DeFi lending requires over-collateralized loans, so at any point, a borrower should have collateral in their account that exceeds their debt. The ratio of their collateral to their debt makes up the "health factor" of their account, and if the health factor drops below a certain threshold, this is when the debt position is available to be liquidated. Knowing an account's health factor would be useful for characterizing the risk that an account is willing to take on. For instance, maybe one account will re-balance its positions to consistently keep its health factor near 1.5, whereas a second account aims to keep its health factor near 2.0. In this case, we could characterize the first account as riskier than the second, because they intentionally operate with a lower health factor. Computing or acquiring the health factors of accounts through time could be a very useful feature for more informative clustering. Similarly, Qin et al. [21] have analyzed risk management provided by keepers that act on accounts within lending protocols. They have measured various risks that liquidation participants are exposed to on four major Ethereum lending pools (i.e., MakerDAO [19], Aave [8], Compound [23], and dYdX [6]), notably including how borrowers ought to monitor their loan-to-value ratios in order to make timely changes to their account positions in order to try and avoid being liquidated.

Another facet of lending that could be useful for further understanding behavioral patterns in DeFi is the account usage of stablecoins. Stablecoins behave much differently in the crypto market than non-stable coins, as they have nearly constant value (typically they are pegged to the US dollar, and so coins like USDC, USDT, and DAI have held steady right around $1 for years). This property of stablecoins can be exploited in lending protocols to help create positions whose health factors are more predictable. For example, if an account takes out a loan using stablecoins as collateral, then the health of the account should only vary with the relative price of the principal asset as opposed to a more complicated relationship between the prices of the principal and collateral assets. Kozhan and Viswanath-Natraj [18] provide some early empirical evidence on the effects of stablecoin-backed loans in DeFi, and have found relationships between the loan risk and price volatility of the DAI. Quantifying how accounts use stablecoins in their borrowing and lending patterns could be another interesting feature (or set of features) to help more accurately characterize behaviors in DeFi lending.

5 Discussion and Future Work

The crypto market and the associated DeFi ecosystem have been extraordinarily volatile for as long as it has existed. Our clustering of address-level behavior on a quarterly basis helps better understand how usage has changed in the wake of major events that shake up the markets, such as when China announced a ban on cryptocurrencies in May 2021[1], when the Terra Luna blockchain crash in May 2022 [3], and the FTX fraud discovery and subsequent collapse in November 2022[2].

Every time one of these events occurs, we see sizable shocks in the market. Besides watching prices drop, it is interesting to see how people have actually reacted to shock events in their patterns of usage. The May 2021 shock from China's announcement of an impending crypto ban caused the largest spike in liquidations, which can be seen in Figure 2 2021 Q2, and there has been a mostly steady decline in the amount that has been borrowed quarter-over-quarter since then. Similarly, the next largest spike in liquidations occurred in May 2022 after the Terra Luna crash. However, the most visible change in usage patterns seems to have occurred in 2022 Q4, which is when FTX crashed. This crash seems to have many DeFi enthusiasts and institutions just waiting to see what happens next in the market. Due to the huge financial losses suffered by investors through the FTX crash (estimated at over $8 billion [4] and the likelihood that this event leads to new regulations for DeFi, it seems many former DeFi users are more pensive regarding how or whether to engage in DeFi for the time being. Correlating these shock events to changes in observed behavior proves interesting as it can validate our own intuitions and also reveal surprising trends that deviate from our expectations.

In addition to external shock events, many DeFi protocols are built with internal governance mechanisms that allow their own user base to propose and make changes to the operations of the protocol as a group. In Aave, these proposals tend to involve setting protocol-level values for things like individual cryptocurrency loan-to-value ratios, liquidation thresholds, or which coins are allowed to be used as collateral (see [8,9,13]). These governance changes are also enacted in the form of blockchain transactions, and currently, some of them are available on Amberdata. Incorporating these transactions into our existing stream and seeing how behavior within protocols changes following governance changes would also be interesting. Whether such changes in a protocol lead to significant, or even noticeable changes to behavior could help DeFi developers build more effective governance mechanisms in new protocols.

Our work in this paper focused solely on the Aave V2 Ethereum market, but this is just one of many Aave markets. Aave has been deployed on Avalanche [25], Polygon [15], Optimism [28], Fantom [12], Harmony [26], and Arbitrum [16], and on some of these it has been deployed with multiple versions. One next step for this work is to see how well the clusters examined in this paper hold up across the other Aave markets. Each market of Aave likely appeals to different groups for one reason or another. For instance, Polygon has significantly lower transaction fees than the Ethereum market, and thus we would expect to see higher transaction volumes among retail users who would be penalized less for making higher-frequency transactions. Should the clustering hold up well in other Aave markets, it would then be interesting to see how well they apply to the other large lending protocols like Compound [23] and MakerDAO [19]. Expanding the scope of data into decentralized exchanges (DEXes) would also be useful. DEXes account for a large portion of transactions in DeFi, and platforms like Uniswap and Sushiswap are some of the most popular in the DeFi ecosystem. Classifying common behavioral patterns in DEX usage would help in seeing the bigger picture of overall DeFi usage, and it may even be possible to find addresses that are present in more than one platform and cluster their behavior across multiple platforms.

The code for computing the clusters used in this paper, as well as for creating the figures, can be found in a public github repository.[6]

6 Acknowledgements

The authors acknowledge the support from NSF IUCRC CRAFT center research grants (CRAFT Grants #22003, #22006) for this research. The opinions expressed in this publication and its accompanying code base do not necessarily represent the views of NSF IUCRC CRAFT. This work was supported by the Rensselaer Institute for Data Exploration and Applications (IDEA).[7]

References

1. https://fortune.com/2022/01/04/crypto-banned-china-other-countries/
2. https://www.investopedia.com/what-went-wrong-with-ftx-6828447
3. https://www.coindesk.com/learn/the-fall-of-terra-a-timeline-of-the-meteoric-rise-and-crash-of-ust-and-luna/
4. https://www.theguardian.com/business/2023/jan/11/ftx-fraud-value-crypto-sbf

[6] https://github.com/aaronmicahgreen/Characterizing-Common-Quarterly-Behaviors-In-DeFi-Lending-Protocols

[7] https://idea.rpi.edu/

5. Adler, A.: lamW: Lambert-W Function (2015). https://doi.org/10.5281/zenodo.5874874, https://CRAN.R-project.org/package=lamW, r package version 2.1.1
6. Antonio Juliano: dYdX: A Standard for Decentralized Margin Trading and Derivatives. Tech. rep. (09 2017), https://whitepaper.dydx.exchange/
7. Bezdek, J.C., Ehrlich, R., Full, W.: Fcm: The fuzzy c-means clustering algorithm. Computers & Geosciences **10**(2), 191–203 (1984). https://doi.org/https://doi.org/10.1016/0098-3004(84)90020-7, https://www.sciencedirect.com/science/article/pii/0098300484900207
8. Boado, Ernesto: AAVE Protocol Whitepaper. Tech. rep. (01 2020), https://www.cryptocompare.com/media/38553941/aave_protocol_whitepaper_v1_0.pdf
9. Boado, Ernesto: AAVE Protocol Whitepaper V2.0. Tech. rep. (12 2020), https://cryptorating.eu/whitepapers/Aave/aave-v2-whitepaper.pdf
10. Buterin, V.: Ethereum: A next-generation smart contract and decentralized application platform (2014), https://github.com/ethereum/wiki/wiki/White-Paper
11. Cebeci, Z.: Comparison of internal validity indices for fuzzy clustering. Journal of Agricultural Informatics (2), 1–14 (2019). https://doi.org/10.17700/jai.2019.10.2.537
12. Choi, S.M., Park, J., Nguyen, Q., Cronje, A.: Fantom: A scalable framework for asynchronous distributed systems (2018). https://doi.org/10.48550/ARXIV.1810.10360, https://arxiv.org/abs/1810.10360
13. Emilio Frangella, Lasse Herskind: AAVE V3 Technical Paper. Tech. rep. (01 2022), https://github.com/aave/aave-v3-core/blob/master/techpaper/Aave_V3_Technical_Paper.pdf
14. Green, A., Cammilleri, C., Erickson, J.S., Oshani, Seneviratne, Bennett, K.P.: Defi survival analysis: Insights into risks and user behaviors (2022), https://www.marble-conference.org/marble2022-cfp
15. Jaynti Kanani, Sandeep Nailwal, A.A.: Matic network whitepaper. https://github.com/maticnetwork/whitepaper (2020)
16. Kalodner, H.A., Goldfeder, S., Chen, X., Weinberg, S.M., Felten, E.W.: Arbitrum: Scalable, private smart contracts. In: USENIX Security Symposium (2018)
17. Kolde, R.: pheatmap: Pretty Heatmaps (2019), https://CRAN.R-project.org/package=pheatmap, r package version 1.0.12
18. Kozhan, R., Viswanath-Natraj, G.: Decentralized stablecoins and collateral risk. WBS Finance Group Research Paper Forthcoming (2021)
19. MakerDAO: The Maker Protocol: MakerDAO's Multi-Collateral Dai (MCD) System. Tech. rep., https://makerdao.com/en/whitepaper/#abstract
20. Mueller, P.: Defi leveraged trading: Inequitable costs of decentralization (2022), http://dx.doi.org/10.2139/ssrn.4241356
21. Qin, K., Zhou, L., Gamito, P., Jovanovic, P., Gervais, A.: An empirical study of defi liquidations: Incentives, risks, and instabilities. In: Proceedings of the 21st ACM Internet Measurement Conference. pp. 336–350 (2021)
22. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2022), https://www.R-project.org/
23. Robert Leshner, Geoffrey Hayes: Compound: The Money Market Protocol. Tech. rep. (02 2019), https://compound.finance/documents/Compound.Whitepaper.pdf

24. Rousseeuw, P.J.: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics **20**, 53–65 (1987). https://doi.org/https://doi.org/10.1016/0377-0427(87)90125-7, https://www.sciencedirect.com/science/article/pii/0377042787901257
25. Sekniqi, K., Laine, D., Buttolph, S., Sirer, E.G.: Avalanche platform (2020), https://assets.website-files.com/5d80307810123f5ffbb34d6e/6008d7bbf8b10d1eb01e7e16_Avalanche%20Platform%20Whitepaper.pdf
26. Team, H.: Harmony technical whitepaper (2020), https://harmony.one/whitepaper.pdf
27. Thorndike, R.: Who belongs in the family? Psychometrika **18**(4), 267–276 (December 1953). https://doi.org/10.1007/BF02289263, https://ideas.repec.org/a/spr/psycho/v18y1953i4p267-276.html
28. Tyneway, M.: Optimism (2020), https://github.com/ethereum-optimism/optimism