MACQUARIE University
SYDNEY · AUSTRALIA

Project Report for COMP7860

Reproduce Paper: *Is decentralized finance actually decentralized? A social network analysis of the Aave protocol on the Ethereum blockchain*

Abrren Chen
Master of Research Student
mengmeng.chen@students.mq.edu.au

Faculty of Science and Engineering,
School of Mathematical and Physical Sciences

November 7, 2022

Contents

1 Introduction

Decentralized finance (DeFi) is a financial ecosystem that uses cryptocurrency (Digital currency) and blockchain (A distributed public ledger that is shared on a decentralized network) to manage financial transactions. DeFi leverages the architecture of peer-to-peer (P2P) networks to allow the participants to validate transactions among each other without any centralized authorities, which challenges the traditional centralized finance system. However, the actual level and effect of decentralization in the current DeFi system are largely unknown [1].

In this research, they applied social network analysis on DeFi token transactions on the Ethereum blockchain to evaluate the true level and impact of decentralization [1].

Firstly, they generated a significant core-periphery structure using 1-year AAVE token (Lending & Borrowing DeFi token) transaction data, which shows the two largest centralized crypto exchanges play central roles. Secondly, the analysis of network features indicated a consistent trend that the AAVE network initially tends to be more decentralized and then reverts to being more centralized. Finally, they found that a higher degree of decentralization is associated with higher returns and higher volatility [1].

In the process of replicating this paper [1], we start with replicating the existing results using the original datasets.Then we further extend this paper by performing the same analysis on new datasets to show whether the conclusion still holds with new datasets.

1.1 Social Network Analysis

Social network analysis is a set technique used to study the exchange of information among actors (individuals, groups or organizations) [3]. In this research, social network analysis is used to measure and evaluate the level of decentralization for token transactions on the AAVE Ethereum blockchain.

1.2 Justification

This paper was recently accepted by the 29th Global Finance Conference. This Global Finance Conference and its journal (Global Finance Journal) are very influential in the finance field. The Cite-Score and Impact Factor of this journal are 4.6 and 2.853 (See Figure 1). Based on both indexes, this journal is ranked as a top 25% journal in this field.

Figure 1: Global Finance Journal Impact Factor

[1] https://arxiv.org/abs/2206.08401

The author of this paper makes the source code and data available in the GitHub repository [2], which makes reproducing this paper becomes much easier.

1.3 Evaluation Framework

This paper has used three main methods to evaluate the actual level of decentralization on the AAVE blockchain and its impact on market return and volatility:

1. Various network features are used to measure the degree of centralization in the token networks of AAVE, such as number of components, giant size ratio, modularity score and std of degree.
2. Network significance test is used to measure the level of decentralization for each observation.
3. Regression analysis between network features and market return and volatility.

2 Original datasets

The original datasets used in the paper are from three open sources: economic feature variables of the AAVE token from Coin-metrics, Total value Locked (TVL) in AAVE from DeFi Pulse and blockchain transaction records of AAVE token from Bigquery public datasets on the Ethereum blockchain.The selected period of these three datasets is from Oct. 10, 2020, to Oct. 9, 2021.

2.1 AAVE Token Transaction Data

The AAVE token transaction data is from a complete live historical Ethereum blockchain dataset [3] on Kaggle (See Figure 2). This dataset can be extracted into CSV files.

timestamp	from_address	to_address	value
10/10/2020	0x00000000000000006f6502b7f2bbac8c30a3f67e9a	0xd75ea151a61d06868e31f8988d28dfe5e9df57b4	1.63E+20
10/10/2020	0x00000000000000006f6502b7f2bbac8c30a3f67e9a	0xdfc14d2af169b0d36c4eff567ada9b2e0cae044f	1.63E+20
10/10/2020	0x000000000025d4386f7fb58984cbe110aee3a4c4	0x6d59cf780d70927f022e3b827f31d6a6235a8d20	3.25E+19
10/10/2020	0x000000000025d4386f7fb58984cbe110aee3a4c4	0xdfc14d2af169b0d36c4eff567ada9b2e0cae044f	3.25E+19
10/10/2020	0x005c17dd5beaad3cb9226bb50a02f066eb125e0b	0x4da27a545c0c5b758a6ba100e3a049001de870f5	1.51E+19
10/10/2020	0x011e7ba61a5e93f3d3e15323bd74c3e6a265c45e	0x037fc8e71445910e1e0bbb2a0896d5e9a7485318	1.14E+20
10/10/2020	0x0122a528c9dcc2cec9c8cbcd7da72beb828d0527	0x4da27a545c0c5b758a6ba100e3a049001de870f5	8.50E+18
10/10/2020	0x0151cc37e580002656bd501d6084e6a2131d8b04	0x317625234562b1526ea2fac4030ea499c5291de4	3.00E+17
10/10/2020	0x0211f3cedbef3143223d3acf0e589747933e8527	0x94c7ec7c03fc568c020ab9eca764eed039531c83	3.31E+19
10/10/2020	0x0211f3cedbef3143223d3acf0e589747933e8527	0xf5ddc4965698be89d0cb769be042004146bed340	1.00E+21
10/10/2020	0x0348be0b0d3f5802ac9d583f9ac000e222ee37b6	0x317625234562b1526ea2fac4030ea499c5291de4	5.50E+18
10/10/2020	0x0348be0b0d3f5802ac9d583f9ac000e222ee37b6	0x4da27a545c0c5b758a6ba100e3a049001de870f5	5.50E+18
10/10/2020	0x037fc8e71445910e1e0bbb2a0896d5e9a7485318	0x4b00296eb3d6261807a6abba7e8244c6cbb8ec7d	1.14E+20
10/10/2020	0x040d99f5bd3b2f90bee7edc3501211d403643489	0xe2ff85b1b4fbdd7f22914086376c6a0657372eb5	5.00E+20
10/10/2020	0x04581b79bf2fac9ed93b41b8023271ecb86b6857	0x317625234562b1526ea2fac4030ea499c5291de4	6.18E+18
10/10/2020	0x04581b79bf2fac9ed93b41b8023271ecb86b6857	0x4da27a545c0c5b758a6ba100e3a049001de870f5	6.18E+18
10/10/2020	0x049d00c84a878fb98069530acaa93042b5e855b3	0x317625234562b1526ea2fac4030ea499c5291de4	5.00E+19
10/10/2020	0x049d00c84a878fb98069530acaa93042b5e855b3	0x4da27a545c0c5b758a6ba100e3a049001de870f5	5.00E+19
10/10/2020	0x054fa0f8cde92f9fe429a514b7f050763178b064	0x317625234562b1526ea2fac4030ea499c5291de4	1.87E+19
10/10/2020	0x054fa0f8cde92f9fe429a514b7f050763178b064	0x4da27a545c0c5b758a6ba100e3a049001de870f5	1.87E+19

Figure 2: Ethereum Blockchain Dataset on Kaggle

[2] https://github.com/Blockchain-Network-Studies/BNS

[3] https://www.kaggle.com/datasets/bigquery/ethereum-blockchain

2.2 AAVE Token Economic Feature Data

The AAVE token economic feature dataset [4] is also a live dataset, which is provided by Coin Metrics (See Figure 3). It contains relevant economic variables of the AAVE token. (e.g., PriceBTC and PriceUSD) This dataset can also be exported into CSV file.

	time	AdrActCnt	AdrBal1in100KCnt	AdrBal1in100MCnt	AdrBal1in10BCnt	AdrBal1in10KCnt	AdrBal1in10MCnt	AdrBal1in1BCnt	AdrBal1in1KCnt	AdrBal1in1MCnt	AdrBalCnt	AdrBalNtv0.001Cnt	AdrBalNtv0.01Cnt	AdrBalNtv0.1Cnt
0	10/10/2020	458	200	1385	1474	50	1154	1439	18	604	1530	1482	1447	1405
1	11/10/2020	378	201	1432	1528	52	1189	1491	18	627	1590	1536	1499	1456
2	12/10/2020	453	234	1591	1693	56	1334	1651	19	723	1760	1701	1661	1614
3	13/10/2020	538	272	1774	1880	66	1487	1841	22	811	1947	1887	1850	1800
4	14/10/2020	710	301	1989	2106	67	1668	2064	23	920	2192	2113	2073	2019
5	15/10/2020	926	339	2321	2458	87	1943	2415	27	1078	2555	2464	2422	2357
6	16/10/2020	791	368	2575	2730	91	2164	2680	27	1189	2837	2735	2690	2612
7	17/10/2020	670	393	2745	2914	99	2299	2859	27	1273	3032	2919	2871	2784
8	18/10/2020	658	409	2935	3117	102	2445	3060	27	1346	3238	3122	3072	2980
9	19/10/2020	739	424	3093	3275	106	2570	3219	28	1410	3402	3281	3230	3138
10	20/10/2020	865	439	3263	3453	101	2711	3396	25	1473	3593	3459	3406	3311
11	21/10/2020	831	468	3517	3728	105	2929	3658	26	1584	3872	3733	3672	3569
12	22/10/2020	860	512	3757	3974	115	3144	3905	27	1721	4126	3980	3919	3814
13	23/10/2020	811	521	3952	4179	116	3292	4111	26	1787	4339	4185	4125	4009
14	24/10/2020	665	529	4095	4333	116	3410	4261	26	1839	4498	4339	4275	4153
15	25/10/2020	626	533	4219	4467	112	3514	4393	26	1892	4633	4473	4408	4277
16	26/10/2020	628	554	4334	4676	118	3600	4601	30	1944	4847	4682	4616	4477
17	27/10/2020	584	563	4464	4809	121	3700	4730	29	1998	4984	4816	4748	4602
18	28/10/2020	599	570	4565	4918	124	3783	4838	30	2023	5104	4928	4857	4706
19	29/10/2020	593	571	4688	5042	120	3892	4962	30	2073	5226	5052	4980	4829

Figure 3: AAVE Token Economic Feature Data on Coin Metrics

2.3 AAVE Token TVL Data

Total Value Locked (TVL) is the total value of all assets deposited in a decentralized finance (DeFi) protocol that is generating economic activity (for example borrowing & lending). The TVL data of AAVE token is provided by DeFi Pulse [5] via API. The TVL data can be exported into CSV file (See Figure 4).

date	tvlUSD	tvlETH	BTC	ETH	DAI	project
10/10/2020	1196560689	3278068.843	16587.27293	292198.2758	9083363.887	AAVE
11/10/2020	1160655252	3136397.481	16788.9561	296717.4969	10204315.36	AAVE
12/10/2020	1142108406	3057444.535	17040.68656	315503.4578	7258282.313	AAVE
13/10/2020	1173304812	3051428.603	16348.91889	335909.6191	9183986.89	AAVE
14/10/2020	1154299271	3037949.445	16269.25278	341034.1872	10420921.74	AAVE
15/10/2020	1133266097	3004417.012	15436.95581	341693.4311	10101734.68	AAVE
16/10/2020	1107559729	2963687.696	14222.61937	356232.8719	10549981.71	AAVE
17/10/2020	1089202360	2981175.717	15067.05794	361922.5262	9877506.234	AAVE
18/10/2020	1071884902	2908541.78	15029.86747	353762.5307	10954764.01	AAVE
19/10/2020	1087195523	2886104.388	14990.05311	355539.1213	10545879.5	AAVE
20/10/2020	1056228302	2797362.948	15371.1394	356764.3745	10849806.44	AAVE
21/10/2020	961115848	2619129.736	15629.61262	326236.2925	10358794.05	AAVE
22/10/2020	1071902493	2755179.265	15705.52365	328407.8645	11562265.04	AAVE
23/10/2020	1156143925	2801075.53	15206.14978	383595.6249	11249313.45	AAVE
24/10/2020	1266008616	3100302.721	14606.61368	438393.9348	24956251.89	AAVE
25/10/2020	1262037372	3083554.955	12916.6038	434943.7135	24935380.78	AAVE

Figure 4: Total Value Locked (TVL) of AAVE

By confirming with DeFi Pulse, they shut down their Data API services on 18th May 2022. Hence, we will not be able to extract any new TVL datasets from DeFi Pulse. We tried to extract the TVL data from other sources, but they did not give the TVL data which we need for this research. By checking the main results in this paper, the TVL data is not used in producing any of the main results. Therefore, we decided not to construct the new TVL dataset.

[4]https://raw.githubusercontent.com/coinmetrics/data/master/csv/aave.csv
[5]https://docs.defipulse.com/metrics/tvl

3 Replication of original work

As the author has made the original code and data available, our focus is replicating the results and verifying the core claims of this paper. To reproduce the results, we attempt to run the original code on the given existing datasets.

3.1 Replication Challenges & Solutions

The GitHub repository [6] provided in the original paper has all the original code and data, but it does not have any process instructions and programming environment information. Because of this, we faced several challenges in replicating all the results. The process of discovery and overcoming these issues are summarized below:

Challenge 1: Clear instructions on extracting the AAVE token transaction records using Bigquery are not available. So we had a lot of problems running Bigquery using a local machine.

Solution: To extract the token transaction data using Bigquery, we have tried many different ways. And finally, we found that using Kaggle notebook to connect with the dataset directly works the best.

Challenge 2: The package version number for cpnet (Social Network Analysis Package in Python) is not available. Therefore, some of the given arguments are not valid in the new version which caused us a lot of problems in replicating the network graphs for significant and insignificant days.

Solution: By checking the cpnet GitHub repository commit history and documentation, we found that these given arguments are not available for the latest cpnet version 0.0.21. In the updated notebook, we have provided the version number for each package and also the scripts which work for each of the cpnet package. (0.0.20 & 0.021)

Challenge 3: When we were trying to replicate the result of the network significance test for each day during the selected period (365 days), the original code did not have the script to generate this result.

Solution: To replicate the same result, we had to write a new function script to perform the network significance test on each day in the dataset.

Challenge 4: There are no notes or instructions available on the running time of this network significance test. Due to the large volume of transactions, running the significance test on each transaction day took much longer than we expected.

Solution: To replicate the significance test results on both existing and new

[6]https://github.com/Blockchain-Network-Studies/BNS

datasets, we used two local machines to run part of the dataset simultaneously which still took us over a week.

Challenge 5: In the process of replicating counterfactual impact evaluation, the code was not well-structured and some of the functions were not working.

Solution: To replicate all the regression results on market return and volatility, we had to rewrite some of the regression functions and restructure the code. In the updated notebook, we have provided detailed instructions of each script and notes on how the notebook is structured.

Note: All the updated notebooks and new data are provided in this new GitHub repository[7].

3.2 Results

There are three main results in the original paper, which we need to replicate using the existing datasets.

Result 1: Defining decentralization via network measures

The first main result shows how these network features are calculated and the dynamic of decentralization on the AAVE blockchain. The following time-series plot shows that the level of decentralization increased first, and then it started to decrease, suggesting a high level of centralization. (See Figure 5)

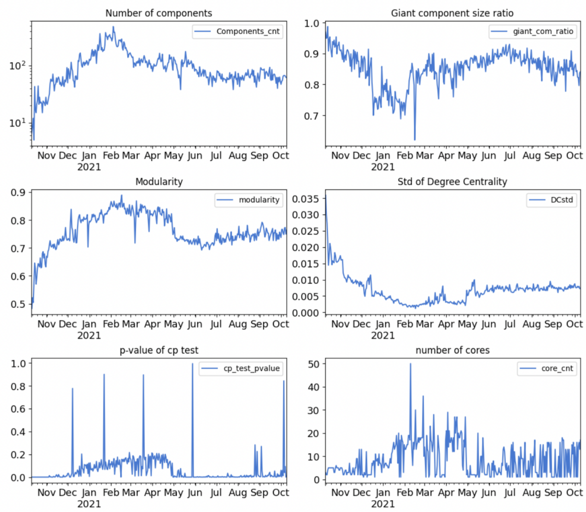

Figure 5: Time-series plots of network features

Note: The correlation heat map of the network feature is attached in the Appendix. (See Figure 11)

[7]https://github.com/AbrrenC/COMP7860$_{P}rojectRepo$

Result 2: Core Periphery Structure

The second main result shows the core-periphery structure outputs, which include the distribution of core nodes, distribution of average number of neighbors and core days count distribution of contract account (CA) and externally owned account (EOA). We applied the Borgatti-Everett (BE) algorithm [2] to test the significance of all 365 observations. The significant day has fewer core nodes and more neighbor nodes, which indicates that the market is more centralized. The insignificant day has more core nodes and fewer neighbor nodes, which suggests that the market is more decentralized. (See Figure 6)

From the core days count distribution of CA and EOA, we can see that the two outliers among EOAs are Binance and Coinbase, which are the top two centralized exchanges in the market. The influence of both centralized exchanges brings a high level of centralization to the market. The two outliers among CAs are decentralized exchanges, which evidence that blockchain can mitigate the dependence on trusted centralized entities [1].

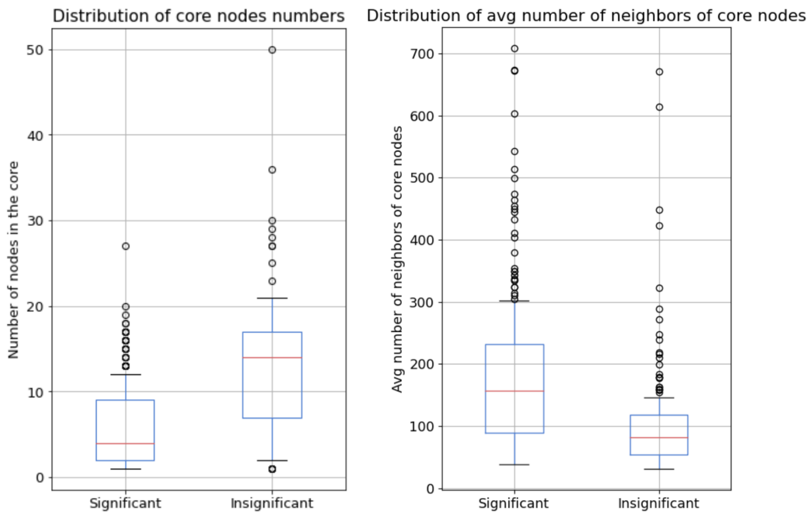

Figure 6: Core-periphery structure features distribution box plots

Note: The core days count distribution of CA and EOA and network graphs are attached in the Appendix. (See Figure 12 & Figure 13)

Result 3: Counterfactual Impact Evaluation

The third main result shows the regression analysis between the level of decentralization and market return and volatility respectively. From both results, we can conclude that in a more decentralized market, the token returns and market volatility will be higher compared with a centralized market.

Note: The regression analysis summary on market return and market volatility is attached in the Appendix. (See Figure 14 & Figure 15)

4 Construction of New Data

To perform the above analysis on new datasets, we need to construct two new datasets and the selected period is 10/07/2021 - 09/07/2022.

1. AAVE token transaction records on Kaggle [8]
2. AAVE token economic feature data on CoinMetrics [9]

4.1 AAVE Token Transaction Data

This is a complete live historical Ethereum blockchain dataset, which can be extracted using Google cloud Bigquery. To extract this dataset, we need to follow below steps:

Step 1: Create a Google project in Google Cloud
Step 2: Link your Google Cloud project with your Kaggle notebook using your project token
Step 3: Set up the Ethereum Blockchain dataset and extract the data using Bigquery
Step 4: Extract the data which you need and save it as CSV (See Figure 7)

token_address	from_address	to_address	block_timestamp	value	timestamp
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0xc559ffc237ace915ae123b41b161df7e2271c5c0	0x0cde991a9562455dab7c77fde88ff4f632e3233a	2021-07-18 07:52:07+00:00	6.81E+16	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x85442367f64f1e4fe9acd22495215366ef107c9f	0x00254d7092da84d87d1b4234f8e5ae28630dd88e	2021-07-18 13:46:55+00:00	1.37E+18	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0xe66b31678d6c16e9ebf358268a790b763c133750	0xa50d8f4f3ca22525ae5aecdef64373c25d97aad9	2021-07-18 15:42:01+00:00	2.75E+17	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x853c2d147a1bd7eda8fe0f58fb3c5294db07220e	0x5ba42facdbc4aeae6befcc567fd62e9242fc69ad	2021-07-18 19:27:31+00:00	1.68E+19	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x4d1ad4a9e61bc0e5529d64f38199ccfca56f5a42	0x18ec8dc1f38512a844c763359bf3f71adc7b9ffd	2021-07-18 01:10:42+00:00	2.60E+18	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x40ec5b33f54e0e8a33a975908c5ba1c14e5bbbdf	0x2a53bc51e4ea1aae0d8d1c2d10443d5a324f8eb9	2021-07-18 02:08:09+00:00	9.25E+17	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0xe66b31678d6c16e9ebf358268a790b763c133750	0xa3a1e2024a90f3b557b89dca2af3c81aa8fe11cc	2021-07-18 00:11:43+00:00	3.70E+18	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x2e280bef16f7ad354d7433781b9cec3ac72d152b	0x4f6742badb049791cd9a37ea913f2bac38d01279	2021-07-18 15:43:06+00:00	1.35E+17	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x6ef5942bdd818fcbcb2864802150985c52309243	0x6cc5f688a315f3dc28a7781717a9a798a59fda7b	2021-07-18 17:35:11+00:00	5.89E+19	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x3b8938867cded2c30e704d81576cde54933ddcd9	0x74de5d4fcbf63e00296fd95d33236b9794016631	2021-07-18 22:28:17+00:00	2.05E+18	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0xc8db353c5e09c5ac4ffb1eea0ea2dfe8fd1a1ae5	0x4674abc5796e1334b5075326b39b748bee9eaa34	2021-07-18 04:57:24+00:00	7.60E+17	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0xbb3be72a5f745a5a055dafa72b13a6e04fc5859e	0x4da27a545c0c5b758a6ba100e3a049001de870f5	2021-07-18 10:57:41+00:00	4.12E+19	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x7ad30d7a631d788eda5f75249f88b7c4de41ec11	0x4da27a545c0c5b758a6ba100e3a049001de870f5	2021-07-18 01:13:43+00:00	4.50E+19	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x7d8551fc420d87b6d0d0e581bf283c6fb3ad37ad	0x5f65f7b609678448494de4c87521cdf6cef1e932	2021-07-18 16:40:34+00:00	3.34E+17	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x492f5e8b40da0b02742b658780483d9dfae13cb3	0x28c6c06298d514db089934071355e5743bf21d60	2021-07-18 09:00:32+00:00	1.50E+20	18/07/2021
0x7fc66500c84a76ad7e9c93437bfc5ac33e2ddae9	0x000000000005117dd3a72e64a705198753fdd54	0x26607ac599266b21d13c7acf7942c7701a8b699c	2021-07-18 06:58:50+00:00	1.61E+19	18/07/2021

Figure 7: New Ethereum Blockchain Dataset on Kaggle

4.2 AAVE Token Economic Feature Data

This dataset contains token economic feature data, which can be extracted directly using pandas "read_csv" function. (See Figure 8)

time	AdrActCnt	AdrBal1in100KCnt	AdrBal1in100MCnt	AdrBal1in10BCnt	AdrBal1in10KCnt	AdrBal1in10MCnt	AdrBal1in1BCnt	AdrBal1in1KCnt	AdrBal1in1MCnt	AdrBalCnt	AdrBalNtv0.001Cnt
30/10/2022	940	1441	89355	116110	334	49263	104932	79	10665	126819	117961
29/10/2022	718	1444	89242	115765	335	49244	104644	79	10674	126443	117603
28/10/2022	679	1439	89165	115571	334	49215	104478	81	10668	126244	117406
27/10/2022	686	1445	89093	115418	331	49185	104340	81	10670	126085	117255
26/10/2022	818	1442	89020	115252	332	49153	104208	82	10659	125916	117089
25/10/2022	903	1451	88911	115017	331	49115	104007	82	10657	125674	116850
24/10/2022	627	1430	88783	114640	332	49038	103681	83	10624	125289	116472
23/10/2022	695	1443	88750	114555	339	49034	103610	82	10648	125194	116387
22/10/2022	464	1434	88698	114386	336	49008	103472	82	10633	125011	116215
21/10/2022	592	1436	88664	114319	339	48991	103410	82	10638	124934	116145
20/10/2022	533	1444	88598	114200	329	48961	103306	84	10638	124811	116028
19/10/2022	658	1445	88543	114122	332	48925	103231	84	10636	124727	115948
18/10/2022	639	1440	88484	114031	344	48908	103149	84	10644	124623	115854
17/10/2022	542	1433	88424	113943	341	48866	103070	83	10627	124533	115766
16/10/2022	637	1430	88341	113831	342	48817	102974	83	10623	124423	115652
15/10/2022	456	1432	88282	113648	344	48799	102830	83	10623	124213	115450
14/10/2022	587	1432	88228	113529	345	48790	102745	83	10629	124077	115325
13/10/2022	662	1435	88158	113407	346	48746	102641	83	10620	123947	115200
12/10/2022	526	1434	88060	113279	346	48702	102517	84	10607	123809	115072
11/10/2022	449	1431	88080	113292	346	48691	102535	84	10607	123818	115084
10/10/2022	401	1431	88030	113231	343	48663	102479	85	10590	123750	115023
9/10/2022	328	1427	87967	113158	342	48622	102408	85	10585	123680	114950
8/10/2022	612	1430	87899	113080	340	48582	102331	85	10582	123602	114874

Figure 8: New AAVE Token Economic Feature Data on Coin Metrics

[8] https://www.kaggle.com/datasets/bigquery/ethereum-blockchain
[9] https://raw.githubusercontent.com/coinmetrics/data/master/csv/aave.csv

5 Results on New Data

By running the original code on the new datasets, we are able to produce the three main results which are shown in the original paper. We will apply the same evaluation framework to evaluate the results on the new datasets.

Result 1: Defining decentralization via network measures

In the mid of 2021, the crypto market entered a bear market which is shown in the transaction volume and cryptocurrency price. Lower trading volume may be one of the key reasons that some of the results are not as obvious as shown in the original paper.

From the following time-series plots of network features (See Figure 9), we can tell that the level of decentralization is relatively consistent throughout the selected period.

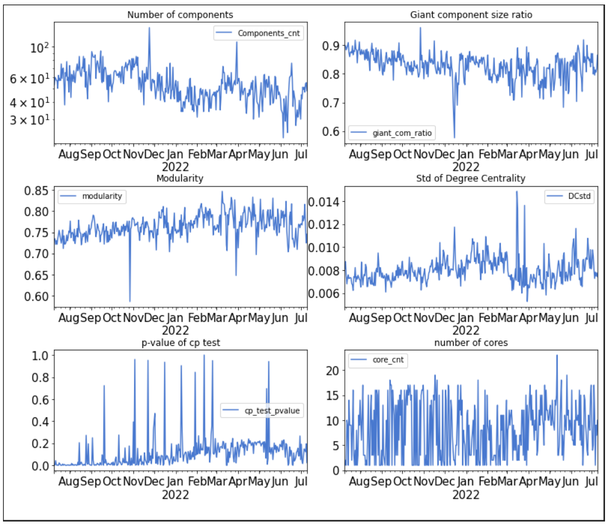

Figure 9: Time-series plots of network features

Note: The correlation heat map of network feature is attached in the Appendix. (See Figure 16)

Result 2: Core Periphery Structure

From the following feature distribution box plots (See Figure 10), we can see that the significant day (centralization) has fewer core nodes and more neighbor nodes, and the insignificant day has more core nodes and fewer neighbor nodes. This result is consistent with the claim in the original paper.

From the core days count distribution of CA and EOA, we can tell that the two centralized exchanges (Binance & Coinbase) dominate the EOA transactions, and the two decentralized exchanges (AAVE & Uniswap) have the largest contribution to CA transactions. This result is also consistent with the original result.

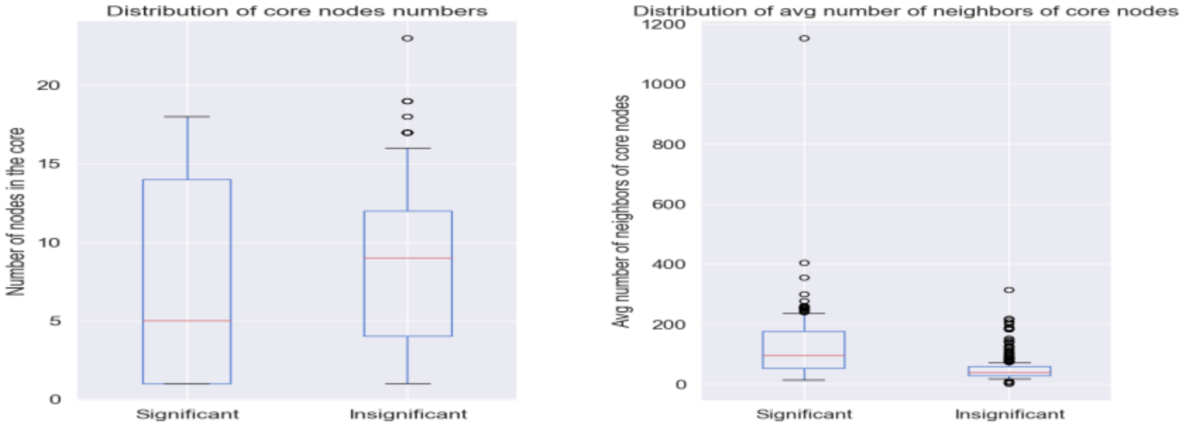

Figure 10: Core-periphery structure features distribution box plots

Note: The core days count distribution of CA and EOA and network graphs are attached in the Appendix. (See Figure 17 & Figure 18)

Result 3: Counterfactual Impact Evaluation

From the regression analysis summary of market return, we can tell that there is a positive relationship between the level of decentralization and market return, but the relationship is not as strong as the original result. From the regression analysis summary of market volatility, there is also a weak positive linear relationship between the level of decentralization and the market volatility.

Both regression analysis results on new datasets are consistent with the results in the original paper, but both relationships are not as strong as the original claim.

Note: The regression analysis summary on market return and market volatility is attached in the Appendix. (See Figure 19 & Figure 20)

6 Reflections

From the replication of this research, we realize that several areas that could be explored in the future:

1. Extract new datasets from other popular DeFi protocols (payment, assets and derivatives).
2. Extend the analysis to study the interplay of other network features and economic variables.
3. Further investigate the network feature algorithm and produce different features of transaction network.
4. Extend the analysis on core days count for all exchanges and investigate the proportion of centralized and decentralized exchanges.

7 Appendix

7.1 Results on Existing Data

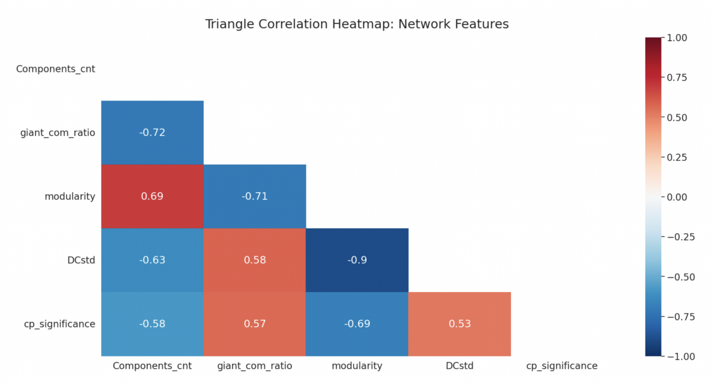

Figure 11: Correlation heat map of network features

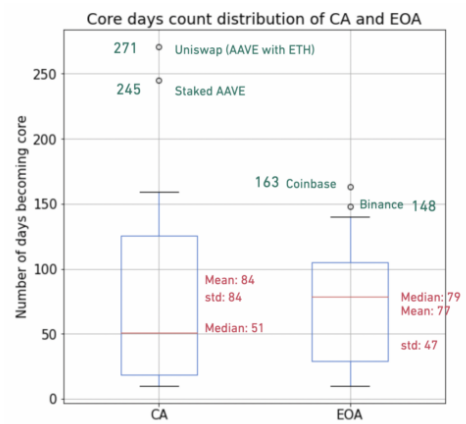

Figure 12: Core days count distribution of CA and EOA

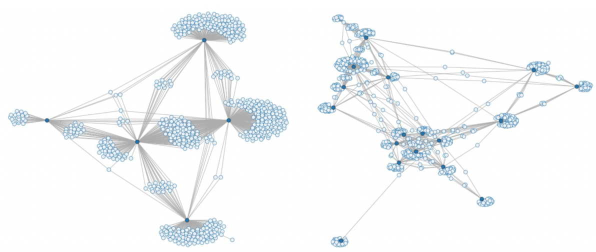

Figure 13: Network graphs on 2020-10-12 (left panel) and 2021-02-22 (right panel)

Table 1: Results of the token market returns (USD)

Time horizon	t, t+1	t, t+7	t, t+14	t, t+21	t, t+28	t, t+35	t, t+42	t, t+49	t, t+56	t, t+90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ component cnt	-0.015	0.113	0.281**	0.423***	0.342***	0.318***	0.261**	0.293**	0.320**	0.258**
R^2	0	0.007	0.029	0.047	0.037	0.033	0.021	0.025	0.025	0.017
Residual Std. Error	(0.129)	(0.132)	(0.163)	(0.192)	(0.178)	(0.177)	(0.185)	(0.192)	(0.213)	(0.221)
Δ giant com ratio	0.024	-0.028	-0.086	-0.188*	-0.212*	-0.144	-0.176	-0.377**	-0.442***	-0.332***
R^2	0	0.001	0.003	0.012	0.017	0.008	0.011	0.049	0.056	0.032
Residual Std. Error	(0.129)	(0.132)	(0.165)	(0.195)	(0.180)	(0.179)	(0.186)	(0.190)	(0.209)	(0.219)
Δ log(modularity)	0.011	0.059	0.089	0.278***	0.288***	0.212**	0.239**	0.351***	0.368**	0.592***
R^2	0	0.003	0.004	0.027	0.034	0.019	0.023	0.046	0.041	0.108
Residual Std. Error	(0.129)	(0.132)	(0.165)	(0.194)	(0.179)	(0.178)	(0.185)	(0.190)	(0.211)	(0.211)
Δ log(DCstd)	0.018	0.005	-0.082	-0.217***	-0.268***	-0.220***	-0.220***	-0.267**	-0.257*	-0.337***
R^2	0	0	0.004	0.021	0.037	0.026	0.025	0.034	0.026	0.046
Residual Std. Error	(0.129)	(0.132)	(0.165)	(0.195)	(0.178)	(0.178)	(0.185)	(0.191)	(0.212)	(0.218)
cp significance	-0.014	-0.090**	-0.163***	-0.278***	-0.322***	-0.324**	-0.314*	-0.188	0.056	1.834***
R^2	0.007	0.031	0.039	0.061	0.046	0.028	0.018	0.005	0	0.124
Residual Std. Error	(0.080)	(0.242)	(0.391)	(0.528)	(0.718)	(0.931)	(1.138)	(1.269)	(1.351)	(2.432)
PCA component1	0.024	0.047*	0.063*	0.092***	0.076**	0.059*	0.045	0.018	-0.033	-0.236***
PCA component2	-0.015	-0.117	-0.196	-0.401***	-0.332***	-0.279***	-0.266***	-0.249**	-0.256**	-0.651***
PCA component3	0.056	0.228***	0.400***	0.575***	0.487***	0.438***	0.435***	0.459***	0.469***	0.561***
R^2	0.011	0.107	0.194	0.3	0.252	0.202	0.177	0.172	0.146	0.42
Residual Std. Error	(0.129)	(0.126)	(0.149)	(0.165)	(0.158)	(0.161)	(0.170)	(0.178)	(0.200)	(0.170)

Note: This table reports the results of predicting the future market return (USD) using the 7-day moving average of network variables (except cp significance). Columns (1)-(10) represent one day, one week to eight weeks, and 90 days respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The data frequency is daily. The residual standard errors are reported in parentheses.

Figure 14: Results of the token market returns (USD)

Table 2: Results of the 30-day volatility growth rate

Time horizon	t, t+1	t, t+7	t, t+14	t, t+21	t, t+28	t, t+35	t, t+42	t, t+49	t, t+56	t, t+90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ component cnt	0.062	0.300***	0.294**	0.349**	0.348**	0.223	0.164	0.155	0.045	-0.279**
R^2	0.005	0.04	0.026	0.025	0.021	0.008	0.005	0.005	0	0.022
Residual Std. Error	(0.085)	(0.151)	(0.185)	(0.226)	(0.253)	(0.264)	(0.237)	(0.237)	(0.246)	(0.219)
Δ giant com ratio	-0.022	-0.041	-0.017	0.01	-0.043	-0.034	-0.061	-0.106	-0.079	0.049
R^2	0.001	0.001	0	0	0	0	0.001	0.003	0.001	0.001
Residual Std. Error	(0.085)	(0.154)	(0.187)	(0.229)	(0.255)	(0.265)	(0.238)	(0.237)	(0.246)	(0.221)
Δ log(modularity)	-0.013	-0.036	-0.128	-0.277**	-0.280*	-0.338**	-0.249**	-0.137	-0.131	-0.082
R^2	0	0.001	0.008	0.024	0.02	0.027	0.018	0.006	0.005	0.003
Residual Std. Error	(0.085)	(0.151)	(0.187)	(0.226)	(0.253)	(0.262)	(0.238)	(0.237)	(0.245)	(0.221)
Δ log(DCstd)	-0.041	-0.162*	0.016	0.128	0.16	0.232	0.153	-0.036	-0.022	0.2
R^2	0.004	0.018	0	0.005	0.007	0.013	0.007	0	0	0.017
Residual Std. Error	(0.085)	(0.152)	(0.187)	(0.228)	(0.254)	(0.263)	(0.237)	(0.237)	(0.246)	(0.219)
cp significance	-0.002	-0.028	-0.068**	-0.122***	-0.210***	-0.294***	-0.364***	-0.412***	-0.425***	-0.431***
R^2	0.001	0.01	0.02	0.035	0.068	0.103	0.131	0.152	0.152	0.191
Residual Std. Error	(0.043)	(0.137)	(0.233)	(0.315)	(0.384)	(0.432)	(0.467)	(0.485)	(0.500)	(0.444)
PCA component1	0.007	0.03	0.036	0.072*	0.136***	0.191***	0.217***	0.252***	0.267***	0.307***
PCA component2	0.035	-0.004	-0.196	-0.271	-0.286	-0.280*	-0.141	-0.019	0.009	0.441***
PCA component3	0.025	0.132*	0.220**	0.201*	0.066	-0.109	-0.291**	-0.468***	-0.573***	-0.831***
R^2	0.008	0.031	0.044	0.05	0.076	0.124	0.184	0.257	0.299	0.508
Residual Std. Error	(0.085)	(0.152)	(0.184)	(0.224)	(0.246)	(0.249)	(0.215)	(0.205)	(0.207)	(0.156)

Note: This table reports the results of predicting the 30-day volatility growth rate using the 7-days moving average of network variables (except cp significance). Columns(1)-(10) represent one day, one week to eight weeks, and 90 days respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The data frequency is daily. The residual standard errors are reported in parentheses.

Figure 15: Results of the 30-day volatility growth rate

7.2 Results on New Data

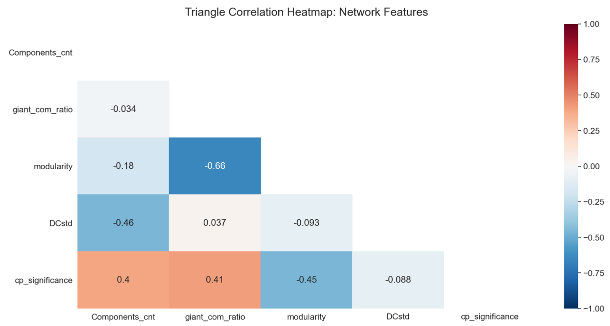

Figure 16: Correlation heat map of network features

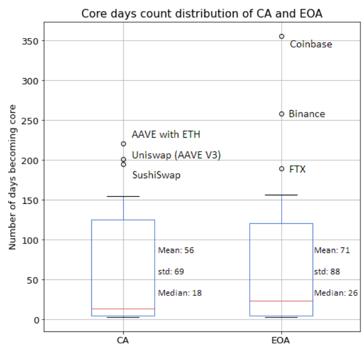

Figure 17: Core days count distribution of CA and EOA

Figure 18: Network graphs on 2021-08-08 (left panel) and 2022-03-14 (right panel)

Table 1: Results of the token market returns (USD)

Time horizon	t, t+1	t, t+7	t, t+14	t, t+21	t, t+28	t, t+35	t, t+42	t, t+49	t, t+56	t, t+90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ component cnt	-0.066	0.141	0.092	0.172	0.139	0.126	0.015	-0.015	0.107	0
R^2	0.003	0.007	0.003	0.01	0.006	0.004	0	0	0.003	0
Residual Std. Error	(0.126)	(0.172)	(0.181)	(0.177)	(0.185)	(0.202)	(0.195)	(0.204)	(0.191)	(0.201)
Δ giant com ratio	0.065	-0.107	-0.162	-0.132	-0.124	-0.17	-0.021	-0.13	-0.216**	-0.199*
R^2	0.003	0.005	0.01	0.007	0.005	0.008	0	0.005	0.015	0.011
Residual Std. Error	(0.126)	(0.173)	(0.181)	(0.177)	(0.185)	(0.202)	(0.195)	(0.203)	(0.190)	(0.200)
Δ log(modularity)	-0.076	0.109	0.152	0.164	0.052	0.025	0.045	0.053	0.144	0.223
R^2	0.003	0.003	0.006	0.007	0.001	0	0	0.001	0.005	0.011
Residual Std. Error	(0.126)	(0.173)	(0.181)	(0.177)	(0.185)	(0.203)	(0.195)	(0.204)	(0.191)	(0.200)
Δ log(DCstd)	0.073	-0.063	-0.034	-0.068	-0.237	-0.042	0.002	-0.01	-0.087	-0.015
R^2	0.003	0.001	0	0.001	0.016	0	0	0	0.002	0
Residual Std. Error	(0.126)	(0.173)	(0.181)	(0.178)	(0.184)	(0.203)	(0.195)	(0.204)	(0.192)	(0.201)
cp significance	0.007	0.038*	0.072**	0.099***	0.103**	0.099**	0.100**	0.071*	0.051	0.056
R^2	0.003	0.011	0.022	0.028	0.027	0.026	0.027	0.015	0.008	0.016
Residual Std. Error	(0.064)	(0.176)	(0.237)	(0.291)	(0.304)	(0.301)	(0.296)	(0.285)	(0.278)	(0.221)
PCA component1	0.019	0.046*	0.060**	0.063**	0.060**	0.060*	0.059*	0.038	0.023	0.046
PCA component2	0.04	-0.103	-0.089	-0.079	-0.018	0.124	0.164	0.127	0.103	0.358***
PCA component3	0.017	0.13	0.152	0.074	0.004	0.143	0.174	0.199	0.156	-0.042
R^2	0.006	0.025	0.031	0.028	0.019	0.028	0.038	0.024	0.015	0.07
Residual Std. Error	(0.127)	(0.171)	(0.179)	(0.176)	(0.184)	(0.201)	(0.192)	(0.202)	(0.191)	(0.195)

Note: This table reports the results of predicting the future market return (USD) using the 7-day moving average of network variables (except cp significance). Columns (1)-(10) represent one day, one week to eight weeks, and 90 days respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The data frequency is daily. The residual standard errors are reported in

Figure 19: Results of the token market returns (USD)

Table 2: Results of the 30-day volatility growth rate

Time horizon	t, t+1	t, t+7	t, t+14	t, t+21	t, t+28	t, t+35	t, t+42	t, t+49	t, t+56	t, t+90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ component cnt	-0.001	0.075	-0.076	0.021	0.07	0.024	0.088	-0.028	-0.025	-0.061
R^2	0	0.002	0.002	0	0.001	0	0.002	0	0	0.001
Residual Std. Error	(0.094)	(0.183)	(0.192)	(0.227)	(0.220)	(0.217)	(0.206)	(0.202)	(0.209)	(0.223)
Δ giant com ratio	0.041	-0.109	0.032	0.066	0.018	-0.021	-0.026	-0.019	0.05	0.066
R^2	0.002	0.004	0	0.001	0	0	0	0	0.001	0.001
Residual Std. Error	(0.094)	(0.183)	(0.192)	(0.227)	(0.221)	(0.217)	(0.207)	(0.202)	(0.209)	(0.223)
Δ log(modularity)	-0.015	0.057	-0.026	-0.003	-0.007	0.077	0.055	0.012	-0.034	-0.023
R^2	0	0.001	0	0	0	0.001	0.001	0	0	0
Residual Std. Error	(0.094)	(0.183)	(0.192)	(0.227)	(0.221)	(0.217)	(0.207)	(0.202)	(0.209)	(0.223)
Δ log(DCstd)	-0.093	-0.231**	-0.033	-0.051	-0.049	0.008	0.042	-0.017	0.001	-0.061
R^2	0.009	0.015	0	0	0	0	0	0	0	0.001
Residual Std. Error	(0.094)	(0.182)	(0.192)	(0.227)	(0.220)	(0.217)	(0.207)	(0.202)	(0.209)	(0.223)
cp significance	0	-0.030*	-0.051*	-0.055	-0.059	-0.068	-0.068	-0.083	-0.095*	-0.211***
R^2	0	0.012	0.014	0.01	0.008	0.008	0.007	0.009	0.012	0.065
Residual Std. Error	(0.047)	(0.133)	(0.210)	(0.275)	(0.331)	(0.367)	(0.395)	(0.421)	(0.427)	(0.401)
PCA component1	0	-0.044	-0.044	-0.039	-0.035	-0.035	-0.027	-0.034	-0.041	-0.130***
PCA component2	-0.01	-0.159**	-0.105	-0.226**	-0.251**	-0.247**	-0.205**	-0.161*	-0.151	-0.137
PCA component3	0.016	0.03	0.017	0.04	0.096	0.133	0.079	0.039	0.039	0.342*
R^2	0.001	0.025	0.015	0.025	0.031	0.032	0.023	0.017	0.017	0.088
Residual Std. Error	(0.095)	(0.181)	(0.191)	(0.225)	(0.218)	(0.214)	(0.205)	(0.200)	(0.208)	(0.214)

Note: This table reports the results of predicting the 30-day volatility growth rate using the 7-days moving average of network variables (except cp significance). Columns(1)-(10) represent one day, one week to eight weeks, and 90 days respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The data frequency is daily. The residual standard errors are reported in parentheses.

Figure 20: Results of the 30-day volatility growth rate

8 References

References

[1] Ziqiao Ao, Gergely Horvath, and Luyao Zhang. Is decentralized finance actually decentralized? a social network analysis of the aave protocol on the ethereum blockchain, 2022.

[2] Stephen P Borgatti and Martin G Everett. Models of core/periphery structures. *Social Networks*, 21(4):375–395, 2000.

[3] Caroline Haythornthwaite. Social network analysis: An approach and technique for the study of information exchange. *Library Information Science Research*, 18(4):323–342, 1996.