Hypergraph Based Analysis of ERC 721 Token Transactions

(Mini Project II Report)
(Course Code: CS491)

Prepared by :
Piyush Anand (CS21B1019)
Neeharika Telu (CS21B1029)

Project was done under Dr. Priodyuti Pradhan

Contents

1 Abstract 4

2 Introduction 5

3 Pre-requisites 8
3.1 Blockchain 8
3.2 Ethereum 8
3.3 Non Fungible Token 9
3.4 ERC-721 9
3.5 Hypergraph 10

4 ERC721 Token Transaction Data and Network Modeling 11
4.1 Transaction Data sets - XBLOCK-ETH [43] 11
4.2 Data Preprocessing 11
4.2.1 Data Downloading 12
4.2.2 Unix to Human Readable Timestamp 12
4.2.3 Dividing Data into Daywise Files 13
4.2.4 Address Labelling 13
4.2.5 Hypergraph Format 14
4.3 Analysis Using Traditional Graphs 17
4.4 Limitations of Traditional Graphs 17
4.5 Introduction to Hypergraphs 17
4.5.1 Advantages of Hypergraphs over Traditional Graphs 18
4.6 Token as Hyperedges 18
4.6.1 Daywise Hypergraph Creation 18

4.6.2 Temporal Evolution of Nodes, Edges, and Hyperedges . . . 19
4.6.3 Degree of a Node . . . 19
4.6.4 Node Type Classification . . . 19
4.6.5 Null Address Handling . . . 20
4.6.6 Hyperedge Classification by Node Count . . . 20

5 Results **22**
5.1 Input data . . . 22

6 References **28**

1 Abstract

This paper presents a novel approach to modeling ERC-721 token transactions on the Ethereum blockchain using hypergraphs. By leveraging a comprehensive dataset from X-Block ETH, we explore the dynamics of token transfers and the relationships between wallets. Traditional graph methods, while useful, often overlook complex interactions and temporal changes within the data. In contrast, our hypergraph model provides deeper insights into the network of token transactions, revealing unique patterns and facilitating a better understanding of wallet interactions. Our approach captures the multifaceted relationships where multiple wallets interact with a single token or vice versa, going beyond pairwise connections. Additionally, by incorporating temporal data, the hypergraph model highlights how wallet interactions evolve over time.

2 Introduction

Over the years, networks have been extensively investigated and have successfully provided insights into the structure and dynamics of physical [1], biological [2, 3, 5, 4], social [6, 7] and several other natural systems [8].

The growing complexity of networks has led researchers to analyze optimal structures and strategies for studying the diverse structural properties and dynamic behaviors of complex networks. Prominent examples relevant to this analysis are interconnected networks represented by multilayer networks [9, 10] and higher-order interactions [11, 12, 13, 14], which are investigated using hypergraphs and simplicial complexes.

Higher-order interactions represent the connections among two or more than two nodes simultaneously in complex systems. Such interactions can be modeled as hyperedges analogous to edges in conventional graphs or simplices in simplicial complexes. Simplicial complex is a specific type of hypergraph with an inclusion property: For a triadic interaction (i, j, k) to exist, the corresponding dyadic interactions (i, j), (i, k) and (j, k) must also be present. In this article, we make use of hypergraphs for our analysis.

Various fields have witnessed the existence of higher-order interactions such as the human brain [15], collaboration networks [17], species interactions [16], cellular networks [18], and drug combinations [19]. Notably, in several real-world networks, presence of higher-order interactions has been found to give new insights, for example, in [20], authors developed a novel method called high-order functional connectivity, which captures interactions among three or more brain regions across various spatiotemporal scales. This method was applied to detailed EEG and fMRI characterizations of neurodegenerative conditions.

Faes et al. [21] introduced the O-information rate (OIR) metric to evaluate higher-order interactions in multivariate time series. This approach was applied to physiological networks, such as heart period, respiratory variability, and arterial pressure in healthy subjects, as well as brain networks described by electrocorticographic signals in animal experiments during anesthesia. Moreover, a study proposed STHAN-SR, a neural hypergraph architecture for stock selection that models complex relationships between stocks using a hypergraph and temporal Hawkes attention mechanism, resulting in a new spatiotemporal attention [22]. In [23], authors conducted neuronal studies revealing that while pair-wise interactions between neurons explain population responses in sensory areas, higher-order interactions are necessary to explain responses in executive areas involved in decision-making and actions.

Santoro et al. [24] analyzed multivariate time series to uncover higher-order patterns in brain functional activity, epidemics, and financial markets. Additionally, a multivariate signal processing technique was developed to construct higher-order networks from time series, which was applied to resting fMRI signals to identify higher-order communications between different brain regions.

Also, these group interactions give rise to new collective phenomena, including synchronization [25, 26], contagion dynamics [27, 28], diffusion processes [29], and evolutionary games [30, 31]. For example, in the evolutionary dynamics of public good games on hypergraph, it is found that overlap between hyperedges tend to foster cooperative outcomes and the heterogeneity of hyperedges promotes cooperation [32]. Burgio et. al [33] studied the susceptible-infectious-susceptible (SIS) process on a hypergraph and demonstrated that overlap between three body interactions lowers the critical point and produces smaller spreads.

Cencetti et al. examined the formation and evaluation of higher-order organizations in temporal networks by studying five social network datasets from various social contexts [34].

3 Pre-requisites

3.1 Blockchain

A blockchain is a decentralized ledger that involves blocks of data, usually with several transactions, a date, the previous block's hash, and a nonce for hash verification. Because blocks are encrypted, it means they are linked through their unique and unalterable hash values, thus ensuring integrity. A consensus mechanism validates transactions through a network, where nodes agree on whether the block is valid before adding it to the chain [35]. This process secures the blockchain, preventing any form of tampering, which enables trust in peer-to-peer asset transfers, as seen in Bitcoin, whereby miners validate blocks and are rewarded with cryptocurrency [36].

3.2 Ethereum

Ethereum, proposed in Vitalik Buterin's paper [37], addresses several limitations of Bitcoin's scripting language. Its chief contributions include full Turing-completeness, supporting all kinds of computations, including loops. More importantly, it improves the structure of a blockchain by allowing for transaction states and other such improvements. Abstractly, it provides an additional layer where users can define customized rules for ownership and transaction formats and also state transition functions via smart contracts-cryptographic rules executed when certain conditions are met [38].

3.3 Non Fungible Token

Non Fungible Token (NFTs) are digital tokens issued on the Ethereum blockchain, similar to Bitcoin, in that they're "minted" and sold. But unlike Bitcoin, which is "fungible," that is, indistinguishable from another bitcoin, NFTs are "non-fungible," that is each token unique. This uniqueness assigns a property right over digital assets, for instance Beeple's artwork. Anyone can view Beeple's "Everydays—The First 5,000 Days"; only the owner of the associated NFT can make an ownership claim.

This dynamic introduces exclusivity into digital art. The NFT is recorded on the immutable Ethereum blockchain after being minted, which proves ownership of the art. While digital art can be looked at, copied, and shared, NFTs ensure that ownership cannot be faked, enabling true exclusive ownership of digital art, a phenomenon previously unthinkable [39].

3.4 ERC-721

The ERC-721 Non-Fungible Token (NFT) Standard [40] proposes a structure to represent one-of-a-kind tokens that are not interchangeable within the Ethereum blockchain. In the case of ERC-20 tokens, their uniqueness and interchangeability come to an end, ERC-721 on the other hand has no crops and it plays the role of managing unique asset properties. They afford atomic management for ownership of each unique digital asset, thus finding usefulness with 'one of a kind' digital objects such as NFTs, games, or intellectual property. By default, the standard states that the American Public policy has ten basic functions and three events which should be observed.

ERC-721 standard tokens were created and distributed via smart contracts and placed on the Ethereum network. The standard is free and open by nature, which means that it can be leveraged when creating non-fungible tokens that are aimed at specific digital assets/objects, which can be owned individually and have historical backgrounds. An application of ERC-721 tokens is CryptoKitties which are internet pets available for auction. Thirdly, the ERC 721 standard lays out basic provisions for creating, issuing, managing, transferring and trading of these specific types of tokens, thus introducing new digital forms of scarcity and asset representation [41].

3.5 Hypergraph

A hypergraph H denoted by $H = (V; E = (e_i)_{i \in I})$ on a finite set V is a family $(e_i)_{i \in I}$, where I is a finite set of indexes, of subsets of V called hyperedges. Sometimes V is denoted by $V(H)$ and E by $E(H)$ [42].

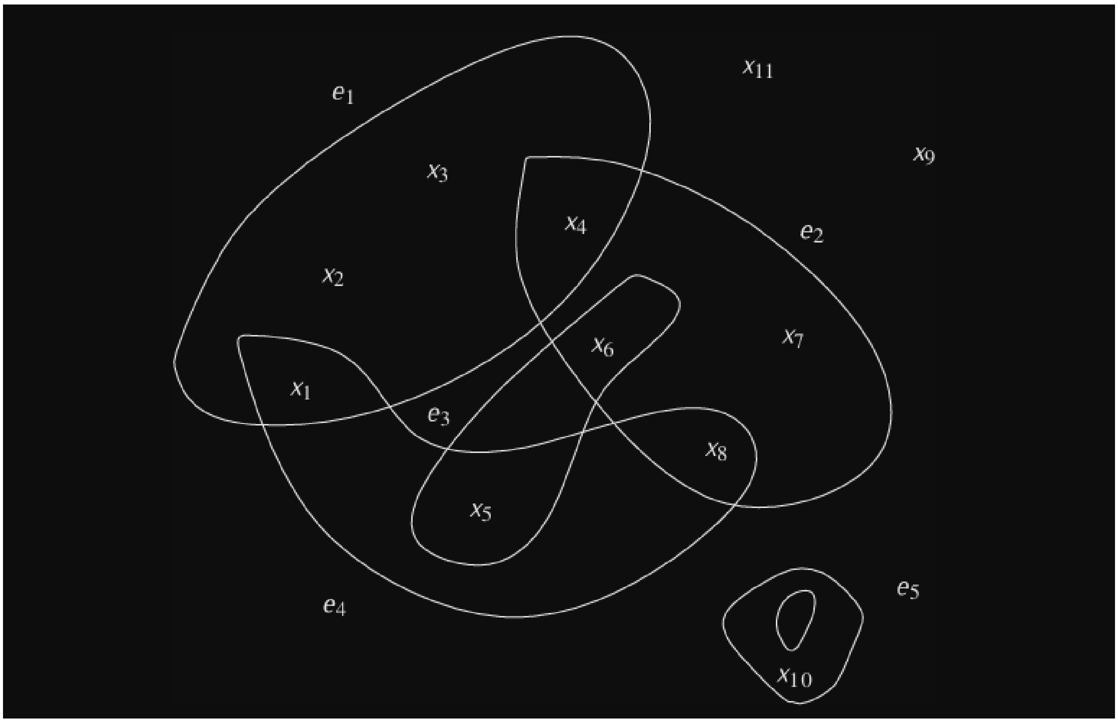

Figure 3.1: The hypergraph H above has 11 vertices, 5 hyperedges e_5, and 2 isolated vertices: x_{11}, x_9. The degree of x_1 is 2 [42].

4 ERC721 Token Transaction Data and Network Modeling

4.1 Transaction Data sets - XBLOCK-ETH [43]

The dataset comprises detailed records of ERC-721 token transactions, which are the foundation for non-fungible tokens (NFTs) on the Ethereum blockchain. Each entry within the dataset provides essential information, including the block number (which indicates where the transaction is recorded on the blockchain), a timestamp (showing when the transaction took place), a unique transaction hash (serving as an identifier), and the addresses of both the sender and recipient.

Additionally, the dataset specifies whether these addresses belong to smart contracts or externally owned accounts, adding another layer of context to the transactions. Spanning from 2015 to 2024, this dataset offers a rich temporal range, enabling analysis of trends and developments in the NFT market over nearly a decade.

4.2 Data Preprocessing

Our analysis began with the preprocessing of the X-Block ETH dataset.

4.2.1 Data Downloading

The initial step in the data preparation procedure was sourcing the needed transaction data found on repository X-Block ETH that contained records of Ethereum transactions. The information contained several key features such as the transaction timing, a respective block to the transaction, the addresses involved and the type of token.

blockNumber	timestamp	transactionHash	tokenAddress	from	to	fromIsContract	toIsContract	tokenId
1001165	1455424860	0x1d11b3ea5593e	0x55b9a11c2e835	0x38150290c18	0x3d2068aeb96	0	0	200000000
1001165	1455424860	0x1d11b3ea5593e	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1003181	1455460299	0x26a30aa663f6e1	0x55b9a11c2e835	0x38150290c18	0x3d2068aeb96	0	0	500000000
1003181	1455460299	0x26a30aa663f6e1	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1003393	1455463847	0xc0db923ac0c6d	0x55b9a11c2e835	0x3d2068aeb96	0xb51446cc429	0	0	100000000
1003393	1455463847	0xc0db923ac0c6d	0x55b9a11c2e835	0x3d2068aeb96	0x5d2c24efac49	0	1	130000
1005878	1455505636	0x113de15964f857	0x55b9a11c2e835	0x38150290c18	0xc0cfc0969d0d	0	0	1000000000
1005878	1455505636	0x113de15964f857	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1005946	1455506683	0x09581f238af3eb	0x55b9a11c2e835	0xc0cfc0969d0d	0xa220568ace9	0	0	100000000
1005946	1455506683	0x09581f238af3eb	0x55b9a11c2e835	0xc0cfc0969d0d	0x5d2c24efac49	0	1	130000
1006034	1455508426	0x637e4263f7825d	0x55b9a11c2e835	0x38150290c18	0xa220568ace9	0	0	200000000
1006034	1455508426	0x637e4263f7825d	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1006046	1455508554	0x5eac9cbd9cae1	0x55b9a11c2e835	0xc0cfc0969d0d	0xa220568ace9	0	0	898000000
1006046	1455508554	0x5eac9cbd9cae1	0x55b9a11c2e835	0xc0cfc0969d0d	0x5d2c24efac49	0	1	1167400
1006059	1455508793	0xc71da93bc6643	0x55b9a11c2e835	0x38150290c18	0xa220568ace9	0	0	4000000000
1006059	1455508793	0xc71da93bc6643	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1007626	1455536210	0x6c6729b405473	0x55b9a11c2e835	0x38150290c18	0xad917335252	0	0	200000000
1007626	1455536210	0x6c6729b405473	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1007645	1455536477	0x0a67fa2e1500f4	0x55b9a11c2e835	0x38150290c18	0xad917335252	0	0	1800000000
1007645	1455536477	0x0a67fa2e1500f4	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1007784	1455539020	0x8d55d21c2e77a	0x55b9a11c2e835	0x38150290c18	0xad917335252	0	0	8000000000
1007784	1455539020	0x8d55d21c2e77a	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0
1008013	1455542892	0xb3a5580191a1c	0x55b9a11c2e835	0x9ac937835f6	0x5d2c24efac49	0	1	411
1008013	1455542892	0xb3a5580191a1c	0x55b9a11c2e835	0x38150290c18	0x9ac937835f6	0	0	1000000000
1008013	1455542892	0xb3a5580191a1c	0x55b9a11c2e835	0x38150290c18	0x5d2c24efac49	0	1	0

Figure 4.1: Downloaded Data From XBlock-Eth

4.2.2 Unix to Human Readable Timestamp

One of the earliest activities in the preprocessing process was to manage the timestamps which varied among other formats supposedly stored in Unix epoch. In order to make the reader's work easier, these timestamps were formatted in an easier to comprehend (YY-MM-DD) thereby ensuring that the information in the dataset would be well understood not only by humans but also by the algorithms which process it.

4.2.3 Dividing Data into Daywise Files

On completion of the conversion of currencies, the data set was arranged in the form of day wise files. This was done to limit the analysis of the transactions to the specific dates, making it easier to analyze the transactional trends and patterns on a daily basis. All these daywise files were also dated meaning that every single file was clearly marked with the date so that the information was organized properly and could be retrieved easily for more work.

Each daywise file contained a few key columns which were designed to capture important details about the transactions. More specifically, these included the date and time stamp of each transaction, the block in which the transaction's records were added, the respective addresses of the different parties in the transaction (from, to), the address of the token which typified the transfer, and whether or not the addresses that transacted were smart contracts. These flags were of utmost importance as they explained how different addresses operate within their blockchain network.

The last two columns-venturing information on which address was or was not a smart contract remained constant over the course of the processing.

4.2.4 Address Labelling

The 64-bit numerical addresses in the dataset presented a challenge that needed to be addressed in a special way. To further enhance the understanding and make the dataset more interpretable, these numerical addresses were converted into appropriate labels with a junction of a dictionary. As such, it aided to comprehend how the addresses were interconnected in addition to creating a rough outline of the distributions or outliers observed in the changes of the data i.e the mentioned occurrences of certain transfers of tokens to particular wallets or smart contracts from the available ones.

blockNumber	timestamp	transactionHash	tokenAddress	from	to	fromIsContract	toIsContract	tokenId
5918558	2018-07-07 0:01	0xeb8b537664ad398	0x06012c8cf97bead5	0x4fabda075e15e9	0x0acb5378a71	0	0	824413
5918560	2018-07-07 0:01	0x8ac506f81242adc	0x06012c8cf97bead5	0x9eec7e6391071	0xfa580941ec47	0	0	556396
5918567	2018-07-07 0:03	0xc37b4a2b5c0a979	0x7fdcd2a1e52f10c2	0x1e74456cc75e4	0x2842a67ce34	0	1	3929
5918570	2018-07-07 0:04	0x2a1cf54f8a812cd3	0x06012c8cf97bead5	0xc7af99fe5513eb	0x4fabda075e15	1	0	825908
5918581	2018-07-07 0:05	0xee2280d9243df7e	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	831457
5918583	2018-07-07 0:06	0x2deb5153f579991	0x7fdcd2a1e52f10c2	0x54133617ad48a	0x92cace836fa2	0	1	3449
5918584	2018-07-07 0:06	0x1a4f17ffa7394388	0x06012c8cf97bead5	0x0000000000000	0x442dccee684	0	0	833692
5918584	2018-07-07 0:06	0x2001c1e251ee963	0x06012c8cf97bead5	0x0000000000000	0x68b42e44079	0	0	833693
5918584	2018-07-07 0:06	0x803b1644d2d38f8	0x06012c8cf97bead5	0x0000000000000	0x7891f796a5d4	0	0	833694
5918584	2018-07-07 0:06	0x5555e5bc64204a3	0x06012c8cf97bead5	0x0000000000000	0x22d1a32a0be	0	0	833695
5918584	2018-07-07 0:06	0x67a495b4f359d78	0x06012c8cf97bead5	0x0000000000000	0x820c0557307	0	0	833696
5918584	2018-07-07 0:06	0x1a001ce852f4dc8	0x06012c8cf97bead5	0x0000000000000	0xc7a024df3876	0	0	833697
5918584	2018-07-07 0:06	0x2023c3fa1d92e7f6	0x06012c8cf97bead5	0x0000000000000	0x06012c8cf97b	0	1	833698
5918584	2018-07-07 0:06	0x2023c3fa1d92e7f6	0x06012c8cf97bead5	0x06012c8cf97bea	0xb1690c08e21	1	1	833698
5918586	2018-07-07 0:06	0xa225f3cef80c37c2	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	830720
5918587	2018-07-07 0:07	0x9c991844105e681	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	830716
5918587	2018-07-07 0:07	0xaf2aeb1b040cb0fc	0x06012c8cf97bead5	0xb1690c08e213a	0x5be7d57e9b1	1	0	833072
5918591	2018-07-07 0:08	0x21aa77639430734	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	829836
5918592	2018-07-07 0:08	0x507e9c6f5150c1a	0x1a94fce7ef36bc90	0x0000000000000	0x2e4ee826597	0	0	9521
5918594	2018-07-07 0:08	0xef0637db5d13a01	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	828948
5918599	2018-07-07 0:10	0xb129c21b597642d	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	828570
5918600	2018-07-07 0:10	0x746a38f26b80be8	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	828196
5918603	2018-07-07 0:10	0x5b1d6fc19252734	0x7fdcd2a1e52f10c2	0x5ab10735fd42ec	0x2842a67ce34	0	1	4065
5918604	2018-07-07 0:10	0x2a2738c45a5fc801	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	828185
5918604	2018-07-07 0:10	0xe3bdda39c984755	0x06012c8cf97bead5	0x4fabda075e15e9	0xb1690c08e21	0	1	828179

Figure 4.2: Daywise Splitting

4.2.5 Hypergraph Format

We stored the Hypergraph in JSON file format, where they Key will be a hyperedge, and the values present in the Key will be the nodes associated with the hyperedge.

Overall, the preprocessing steps were designed to ensure the dataset was clean, well-structured, and ready for further analysis. This approach allowed for easy extraction of meaningful insights from the transaction data and laid the foundation for more complex analyses, such as network analysis, node degree calculations, and hypergraph modeling.

timestamp	fromLabel	toLabel	tokenAddressLabel	fromIsContract	toIsContract	tokenId
2018-07-07 00:01:25	71594	90964	53	0	0	824413
2018-07-07 00:01:50	98051	101339	53	0	0	556396
2018-07-07 00:03:07	87786	101274	381	0	1	3929
2018-07-07 00:04:04	828	71594	53	1	0	825908
2018-07-07 00:05:22	71594	825	53	0	1	831457
2018-07-07 00:06:00	82513	93145	381	0	1	3449
2018-07-07 00:06:06	1179	825	53	1	1	833698
2018-07-07 00:06:49	71594	825	53	0	1	830720
2018-07-07 00:07:33	71594	825	53	0	1	830716
2018-07-07 00:07:33	825	71036	53	1	0	833072
2018-07-07 00:08:12	71594	825	53	0	1	829836
2018-07-07 00:08:46	71594	825	53	0	1	828948
2018-07-07 00:10:01	71594	825	53	0	1	828570

Figure 4.3: Labelled Data

{"24":[99,86582],"53":[37889,98051,91143,94220,10005,65048,21017,22042,65820,64031,90657,70179,
14118,21544,22314,14894,93743,815,28465,67638,825,828,40514,74564,94539,68684,71758,68686,12879,
85840,33104,90964,50004,74844,66143,80736,70753,90737,65398,75384,83833,72316,71036,12157,87423,
93570,81284,67717,41094,65162,65164,72844,76430,75923,6292,93844,91540,101270,1179,86687,66720,
8609,101281,66725,28070,95144,92842,71594,95918,94894,50862,65455,53681,101044,91067,30142,101311,
101315,40396,16333,41421,92376,72409,101337,55515,101339,69341,101340,101342,101343,70881,101345,
101346,101348,101349,101350,101351,101352,6632,59114,2283,101354,101355,101356,101357,101360,8945,
93431,65021],"69":[95348,56823],"146":[66720,69399],"266":[72547,72572,72947],"381":[87763,100485,
101256,84617,94989,84625,101274,83486,87753,82476,85934,100398,82992,100787,82995,85814,82873,
100028,101309,101310,82882,82630,85703,82631,87752,87754,87755,87756,87757,87758,87759,87760,82513,
87762,87761,87764,87765,93396,87767,87768,93145,87770,87771,87772,87773,87774,82911,87775,87776,
87778,87779,87777,87766,87780,87786,87787,88829,97790],"396":[101330,12157,84269],"440":[94273,
93697,95235,89961,101353,92141,101277,95471,95856,90778,95964,87965,75743],"463":[88512,85840,
90691,91067,83436,12157,88799],"484":[101248,90816,92228,100486,100429,94993,6292,101338,95394,
101347,100646,95529,94961,95541,94775,92092,92093],"493":[94265,93382,85134,93383],"512":[100617,
99062,99599],"519":[100420,58180,88565,74264,65769,94331,93406]}

Figure 4.4: Storing Hypergraphs as JSON Files

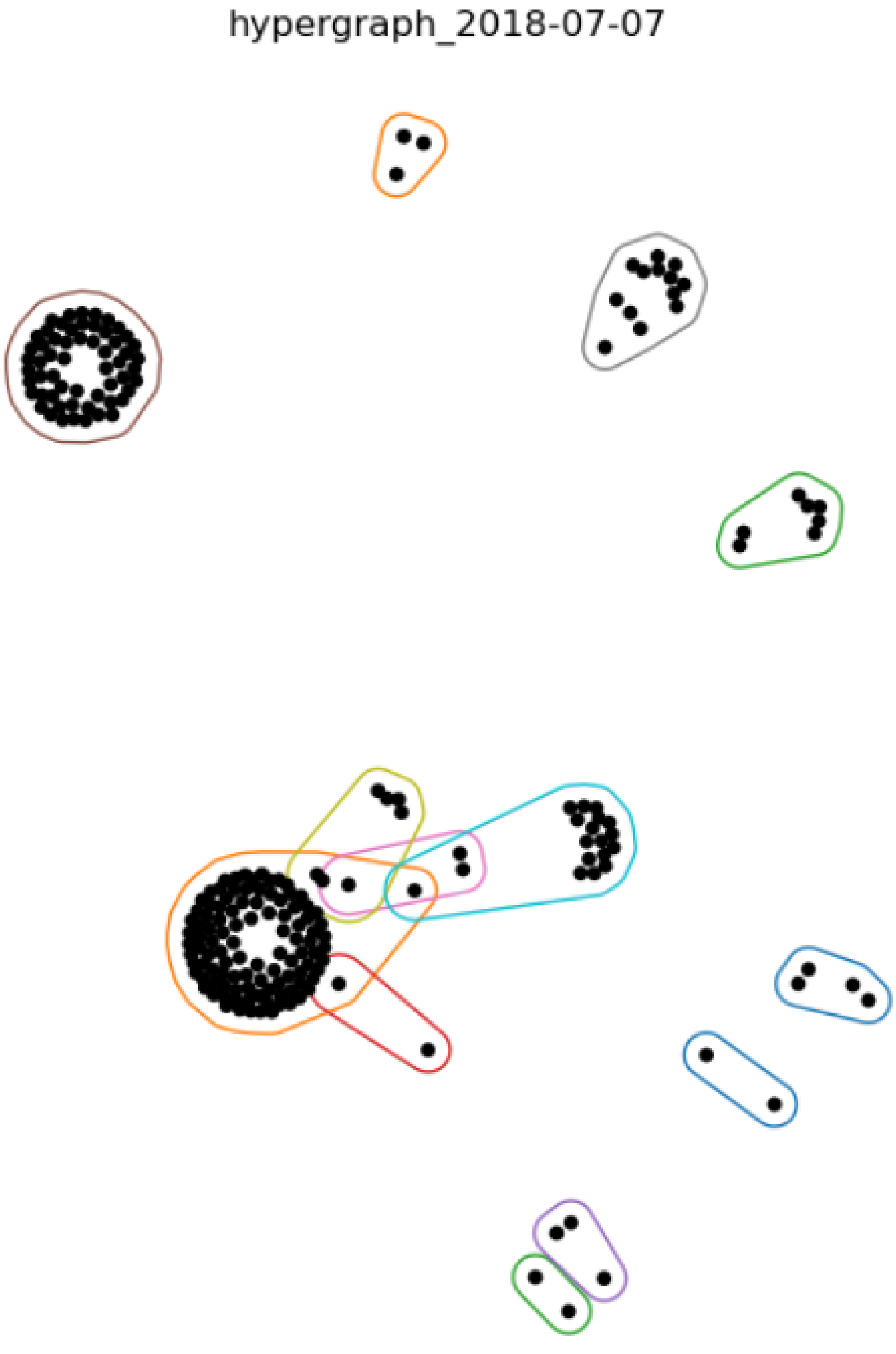

Figure 4.5: Visualisation of A Day's Hypergraph

4.3 Analysis Using Traditional Graphs

To understand the relationships between wallets and transactions, we implemented a series of analyses. We developed a script to calculate the total number of unique wallets (nodes) involved in the transactions, with the results represented on a log scale for better visualization. Following this, we plotted the number of unique nodes on a daywise basis to observe participation trends over time. Additionally, we represented daywise transactions with edges, illustrating the pairwise connections between nodes. To further enhance our analysis, we also plotted daywise unique edges, which allowed us to visualize the distinct transactions occurring each day.

4.4 Limitations of Traditional Graphs

While the traditional graph methods provided valuable insights into the number of wallets and transactions, they fell short in capturing several critical aspects. For example, they could not adequately reveal how many wallets were involved in token transfers or how these interactions changed over time. Additionally, traditional graphs often struggle to represent higher-order relationships, where a single transaction can involve multiple participants simultaneously. This limitation highlighted the need for a more sophisticated approach.

4.5 Introduction to Hypergraphs

Hypergraphs extend the concept of traditional graphs by allowing edges to connect any number of vertices, rather than just pairs. In a hypergraph, a hyperedge can link multiple nodes, enabling a more nuanced representation of complex relationships. This framework is particularly advantageous in the context of ERC-721 transactions, where a single token transfer may involve several participants.

4.5.1 Advantages of Hypergraphs over Traditional Graphs

Utilizing hypergraphs in our analysis offers a more comprehensive representation of interactions. First, hypergraphs can effectively capture multi-way relationships, showcasing situations where multiple wallets collaborate to transfer a single token. Additionally, they allow us to observe the evolution of these relationships over time, unveiling patterns that traditional graphs may overlook. Finally, by treating tokens as hyperedges, we can assess the collective influence of multiple wallets on token transfers, yielding valuable insights into collaboration and community dynamics within the NFT ecosystem.

4.6 Token as Hyperedges

In our model, each token was treated as a hyperedge, with the wallets involved in the transfer of that token represented as nodes connected to the hyperedge. We conducted the following analyses: Tokens vs. Days: We plotted the number of tokens transferred against days to visualize trends over time. Unique Tokens: We examined the unique tokens traded daily. Node Degree Calculation: The degree of each node (the number of tokens traded by that wallet on a particular day) was calculated daily. Degree Centrality: We calculated the degree centrality for nodes each day, providing insights into the most active wallets. 3.4 Analysis of Hyperedges We explored the number of nodes involved in each token transfer and calculated the number of 1-hyperedges, 2-hyperedges, and so on, on a daywise basis. Over the period from 2015 to 2024, we identified a total of 4,000 hyperedges, with nearly 130,000 nodes involved in at least one hyperedge.

4.6.1 Daywise Hypergraph Creation

To track and analyze the evolution of ERC-721 transactions, daywise hypergraphs were created based on transaction data, with each hyperedge representing a unique token. For each day, the total number of hyperedges (tokens) was computed, and the nodes involved in token transfers were identified.

4.6.2 Temporal Evolution of Nodes, Edges, and Hyperedges

Daily changes in the counts of nodes, edges, and hyperedges were analyzed to observe shifts in transaction activity over time. From 2015 to October 2017, ERC-721 transactions and token transfers between wallets were minimal. This low-activity period provided limited information value and was thus excluded from further analysis. The focus was placed on the more active period from 2017 to 2024, resulting in a dataset more relevant for hypergraph-based analysis.

4.6.3 Degree of a Node

The degree of a node, representing the number of tokens a wallet transfers in a given day, was calculated for each node. For each day, the node with the highest degree was identified as the most active, highlighting the account with the most token transfers for that day. Analyzing daily maximum degrees helped monitor activity peaks and provided insights into particularly active wallets. When extracting maximum-degree information, efforts were made to classify the node type as either a smart contract or an externally owned account. During this process, a few nodes were found to exhibit both types of behavior, an anomaly in the Ethereum system.

4.6.4 Node Type Classification

To better understand node roles, nodes were categorized as either externally owned accounts (EOAs) or smart contracts. EOAs represent individual users, while smart contracts automate complex transactions. Some nodes appeared as both EOAs and smart contracts while observing the node behavior which has maximum degree, which is not possible within Ethereum's system. Removing these misclassified nodes improved data quality, clarifying the distinction between regular user accounts and contract-driven addresses.

4.6.5 Null Address Handling

The null address, commonly used as a placeholder or for testing purposes, frequently appeared in transactions when extracting information on nodes with maximum degrees. Its presence affected maximum degree calculations by introducing a non-participant entity. Removing the null address allowed the dataset to more accurately reflect real user interactions. After removing the null address and nodes misidentified as both smart contracts and regular accounts, the dataset became cleaner and more representative of actual interactions. Stats?

4.6.6 Hyperedge Classification by Node Count

Hyperedges were classified by the number of nodes involved. For example, a 1-hyperedge involved a single node, while a 2-hyperedge represented token transfers between two nodes. Analyzing these hyperedges revealed a range of transaction types, from single-wallet actions to multi-wallet interactions. After data cleaning, the maximum hyperedge count reached 45,833, whereas it was initially as high as 126,455 when the null address was included.

Hyperedge Classification by Node Count Hyperedges in this network capture token transactions, with the node count indicating transaction types.

1-Hyperedge (Self-Transaction)

This is a single-node transaction, where a node transfers tokens to itself, usually for organizational purposes or automated tasks. Example: Imagine Ravi transferring tokens from one of his wallet accounts to another for internal management.

2-Hyperedge (Standard Two-Node Transfer)

This is the typical transaction where one node sends tokens directly to another node, involving two participants. Example: If Priya sends tokens to Chintu, this forms a 2-hyperedge, with a simple sender-receiver dynamic.

3-Hyperedge (Three-Node Chain)

A three-node hyperedge represents transactions involving three distinct nodes, often seen in multi-step transactions. Example: In a scenario where Priya sends tokens to Bhuvan, who then forwards them to Chintu, this forms a 3-hyperedge.

n-Hyperedge (Multi-Node Transaction)

Transactions with four or more nodes represent complex, multi-party transactions, common in pooled investments or decentralized finance. Example: For instance, if multiple investors pool funds into a shared fund, this can form an n-hyperedge, indicating large-scale, multi-participant involvement.

5 Results

5.1 Input data

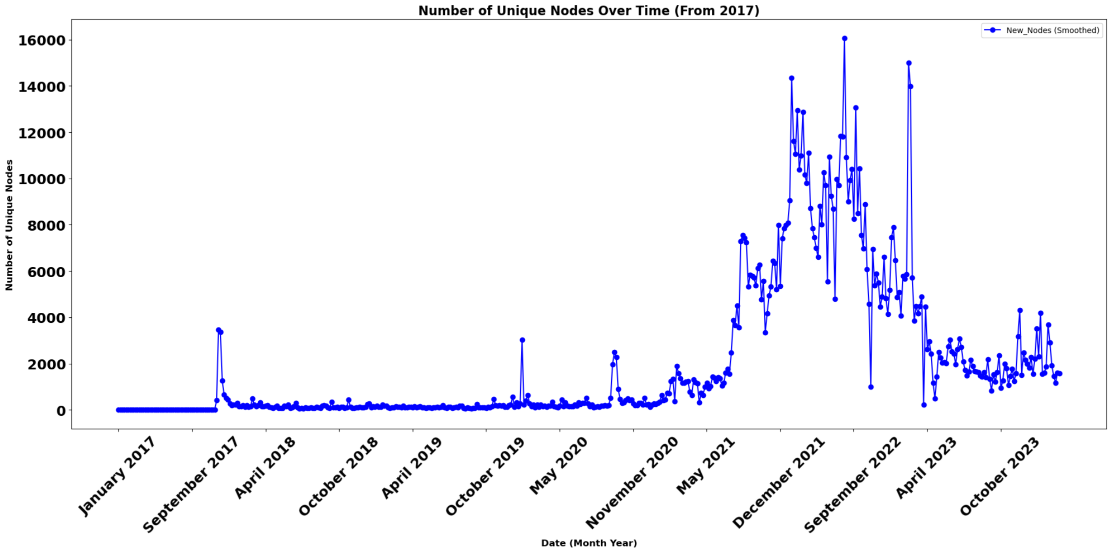

Figure 5.1: The plot illustrates the daily addition of new unique nodes, reflecting the network's growth of unique wallets over time. Each data point represents the number of wallets engaging in transactions for the first time on that specific day. By examining this plot, we can gauge how quickly the network is growing, spot trends in how new users are joining, and identify times of rapid growth or slowdowns. This information helps us understand how the network is evolving and how its adoption is changing over time.

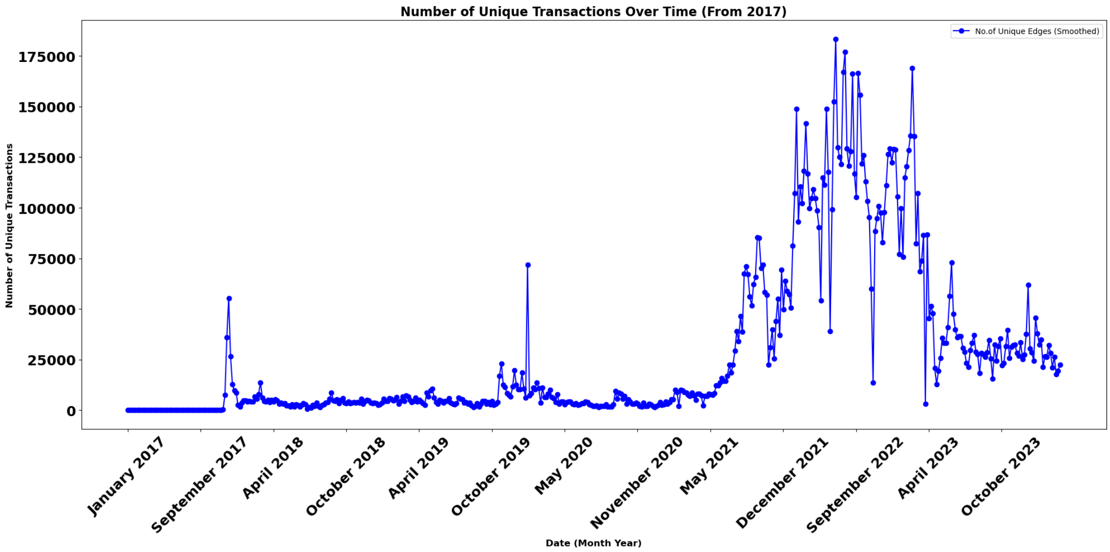

Figure 5.2: The plot depicts the number of new transactions occurring each day, providing valuable insights into transaction activity over time. By analyzing these daily changes, we can observe how transaction volumes vary, such as identifying significant spikes or drops in activity. For instance, a sudden increase in transactions on a specific day may indicate a major event or shift in user behavior. This information is crucial for understanding how the system behaves and helps in making smart decisions based on how transaction patterns change over time.

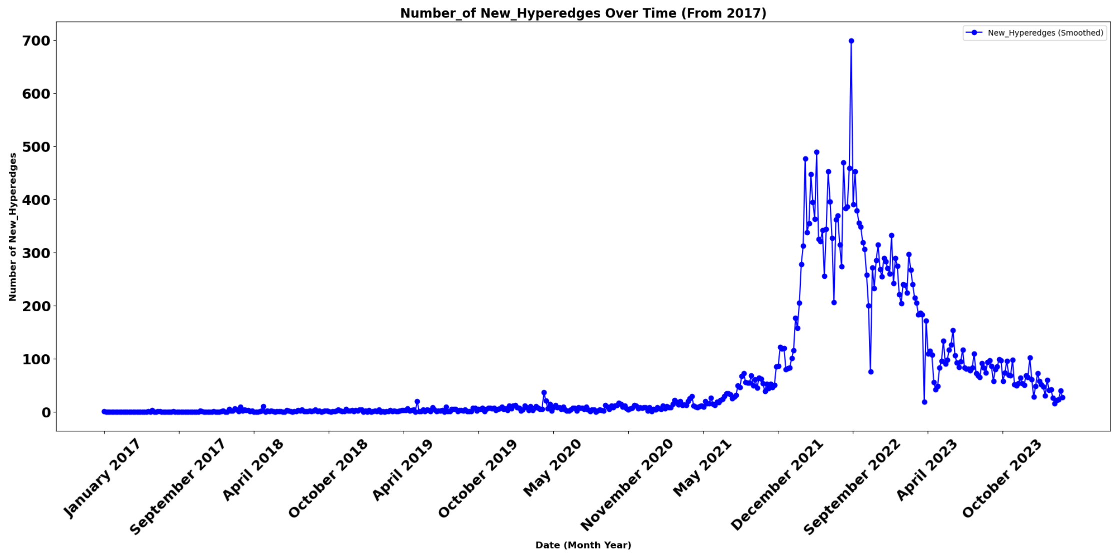

Figure 5.3: By plotting the graph between days and the number of new unique elements being added over time, we can easily track how many new tokens are emerging each day. This visualization helps identify trends in the introduction of new tokens, revealing periods of increased activity or significant events

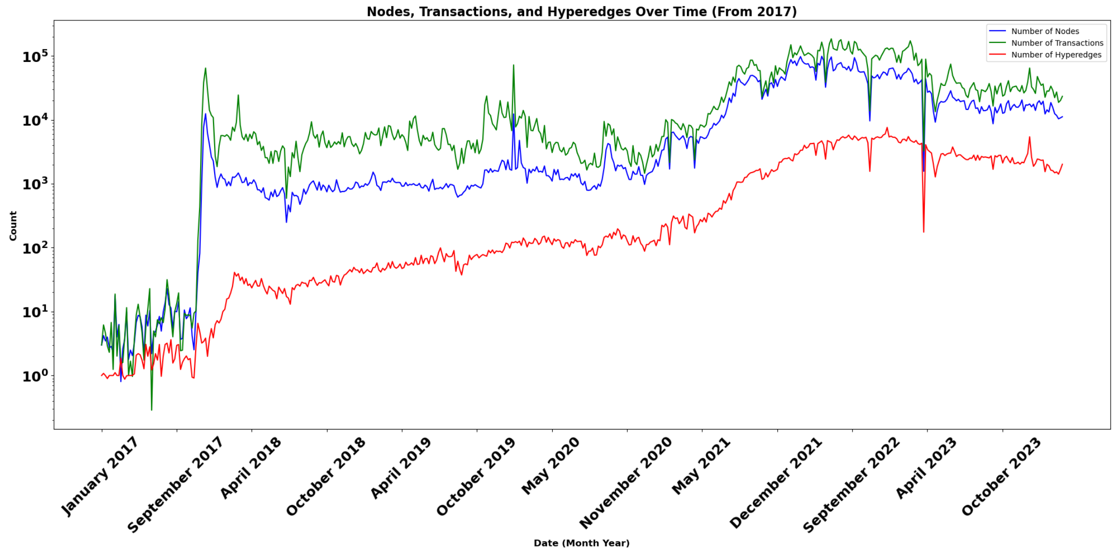

Figure 5.4: The plot tracks the daily changes in the number of active wallets (nodes) involved in transactions, offering a clear view of how network activity fluctuates over time on daily basis. By observing these variations, we can identify trends in user engagement, such as whether the network is expanding or contracting. The plot also helps to detect patterns, revealing consistent increases or decreases in node activity, which provide valuable insights into the overall dynamics and health of the network. The log scale on the y-axis, which enhances the visualization of the plot by compressing large ranges of data, allowing both small and large changes in the number of nodes (wallets) to be observed more clearly. On a normal scale, significant variations might overshadow smaller fluctuations, making it difficult to detect trends or changes in network activity.The log scale, however, evenly spaces out the data, making it easier to see both minor and major shifts in node participation over daily basis. Plotting the graph between days and the number of nodes, transactions and hyperedges(tokens) provides the insights of how the wallets, transactions and tokens are being varied over time . The plot shows the number of transactions (edges) on a daily basis, where each edge connects two nodes. This detailed view allows us to track the volume of transactions occurring each day, offering insights into daily activity levels within the network. By examining these daily transactions, we can gain a clearer picture of how transaction activity evolves over time, identify trends in user interactions, and understand how network engagement changes from day to day. The day vs transactions graph is plotted on a logarithmic scale to better analyze the increase in transactions and reduce any confusion or ambiguity. A logarithmic scale is used to compress large ranges of data, making it easier to visualize both small and large changes in transaction volume. The plot illustrates the relationship between days and the number of tokens, The tokens are represented as hyperedges in this context providing insight into the number of unique tokens being transferred by the nodes each day. Plotting the relationship between days and the number of tokens on a log scale for the y-axis enhances the visualization of data with large variations, making it easier to detect trends, outliers, and exponential growth or decay in token activity. It provides a clearer understanding of token transfer dynamics, highlights periods of significant change, and offers better insights into the distribution and volatility of token transfers over time.

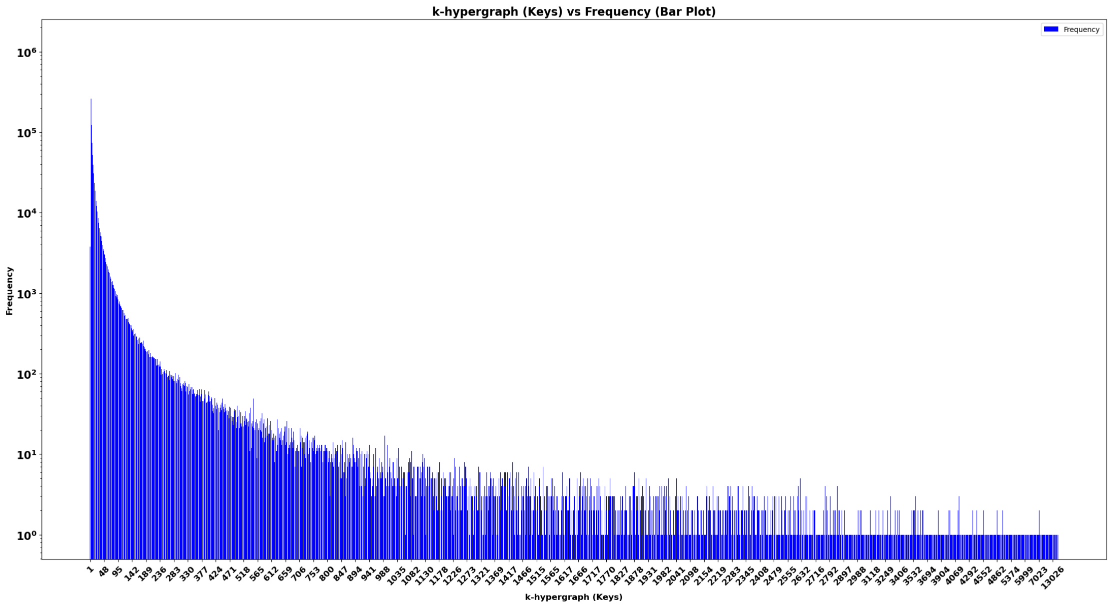

Figure 5.5: The plot shows the frequency distribution of hyperedges of different sizes within the dataset. Each data point represents the count of hyperedges containing a specified number of nodes, where a 1-hyperedge indicates hyperedges with a single node, 2-hyperedges include two nodes, and so forth. This distribution helps to illustrate the overall connectivity represented in the hypergraph

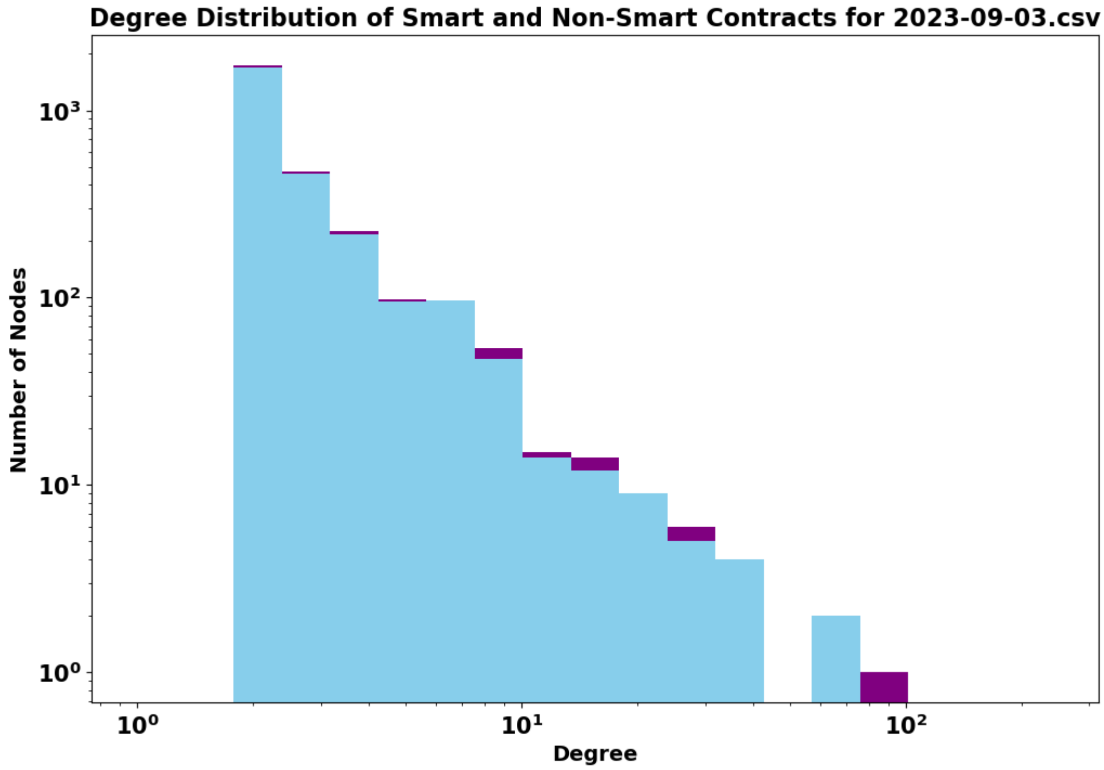

Figure 5.6: Degree of the Node represents the no.of Tokens it is transacting on that particular day. The degree distribution is a probability distribution that shows the frequency of nodes for each degree value. Here the nodes with smart contract is represented in Blue color and the externally owned accounts are represented in sky blue color to differentiate the wallet type that is involved in the transaction of various tokens.

6 References

Bibliography

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D. U. Hwang. Complex networks: Structure and dynamics. Physics Reports. 424, 175-308, (2006).

[2] C. Koch , G. Laurent. Complexity and the nervous system. Science. 1999, **284**(5411) 96-98

[3] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai , A.-L. Barabási. The large-scale organization of metabolic networks. Nature

[4] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai. Lethality and centrality in protein networks. Nature. 2001, **411**, 41–42

[5] G. Weng, U. S. Bhalla, R. Iyengar. Complexity in biological signaling systems. Science. 1999, **284**(5411), 92-96

[6] M. E. J. Newman. The structure of scientific collaboration networks. PNAS. 2001, **98**(2), 404-409

[7] S. Redner. How popular is your paper? An empirical study of the citation distribution. Eur. Phys. J. B. 1998, **4**, 131–134

[8] M.E.J. Newman. The Structure and Function of Complex Networks. SIAM Review. **45**(2), (2003).

[9] M. De Domenico, C. Granell, M. A. Porter and A. Arenas. The physics of spreading processes in multilayer networks. Nature Phys 12, 901–906, (2016).

[10] M. Kivel, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno and M A Porter. Multilayer networks, Journal of Complex Networks, (3), 203–271, (2014).

[11] S. Boccaletti, P. De Lellis, C.I. del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, M. Romance. The structure and dynamics of networks with higher order interactions. Physics Reports, 1018, 1-64, (2023).

[12] F. Battiston et al. Networks beyond pairwise interactions: structure and dynamics. Phys. Rep. 874, 1 (2020).

[13] R. Lambiotte, M. Rosvall, and I. Scholtes. From networks to optimal higher- order models of complex systems. Nat. Phys. 15, 313–320 (2019).

[14] A. R. Benson, D. F. Gleich, J. and Leskovec. Higher-order organization of complex networks. Science. 353, 163–6, (2016).

[15] S. Gu, M. Yang, J.D. Medaglia, R.C. Gur, R.E. Gur, T.D. Satterthwaite and D.S. Bassett. Functional hypergraph uncovers novel covariant structures over neurodevelopment. Human brain mapping, 38, (2017).

[16] J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina, Higher-order interactions stabilize dynamics in competitive network models, Nature (London) 548, 210 (2017).

[17] A. Patania, G. Petri, and F. Vaccarino, EPJ Data Sci. 6, 1 (2017).

[18] S. Klamt, U.-U. Haus, and F. Theis, PLoS Comput. Biol. 5, (2009).

[19] A. Zimmer, I. Katzir, E. Dekel, A. E. Mayo, and U. Alon, Proc. Natl. Acad. Sci. U.S.A. 113, 10442 (2016).

[20] R. Herzog, F. E. Rosas, R. Whelan et. al., Genuine high-order interactions in brain networks and neurodegeneration. Neurobiology of Disease. 175, 105918 (2022).

[21] L. Faes et al., A New Framework for the Time- and Frequency-Domain Assessment of High-Order Interactions in Networks of

Random Processes. IEEE Transactions on Signal Processing, 70, 5766-5777, (2022)

[22] R. Sawhney, S. Agarwal, A. Wadhwa, T. Derr, and R. R. Shah. Stock Selection via Spatiotemporal Hypergraph Attention Network: A Learning to Rank Approach. Proceedings of the AAAI Conference on Artificial Intelligence, 35(1), 497-504, (2021).

[23] M. I. Chelaru, S. Eagleman, A. R. Andrei, R. Milton, N. Kharas and V. Dragoi. High-order interactions explain the collective behavior of cortical populations in executive but not sensory areas. Neuron, 109, 3954-3961, (2021).

[24] A. Santoro F. Battiston, G. Petri and E. Amico. Unveiling the higher-order structure of multivariate time series. arXiv:2203.10702.

[25] P. S. Skardal and A. Arenas, Commun. Phys. 3, 1 (2020).

[26] A. P. Millán, J. J. Torres, and G. Bianconi, Phys. Rev. Lett. 124, 218301 (2020).

[27] S. Chowdhary, A. Kumar, G. Cencetti, I. Iacopini, and F. Battiston, J. Phys. 2, 035019 (2021).

[28] L. Neuhäuser, A. Mellor, and R. Lambiotte, Phys. Rev. E. 101, 032310 (2020).

[29] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, SIAM Rev. 62, 353 (2020).

[30] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latora, Nat. Hum. Behav. 5, 586 (2021).

[31] A. Civilini, N. Anbarci, and V. Latora, Phys. Rev. Lett. 127, 268301 (2021).

[32] D. Wang, P. Yi, Y. Hong, J. Chen, and G. Yan. Evolutionary game on any hypergraph. arXiv:2404.03305.

[33] G. Burgio, S. Gómez, A. and Arenas. Triadic Approximation Reveals the Role of Interaction Overlap on the Spread of Complex Contagions on Higher-Order Networks. Phys. Rev. Lett. 132, 077401, (2024).

[34] G. Cencetti, F. Battiston, B. Lepri, et al. Temporal properties of higher-order interactions in social networks. Sci Rep 11, 7028 (2021).

[35] Glaser, Florian and Glaser, Florian and Bezzenberger, Luis, Beyond Cryptocurrencies - A Taxonomy of Decentralized Consensus Systems (March 1, 2015). 23rd European Conference on Information Systems (ECIS), Münster, Germany, 2015, Available at SSRN: https://ssrn.com/abstract=2605803

[36] Nofer, M., Gomber, P., Hinz, O. et al. Blockchain. Bus Inf Syst Eng 59, 183–187 (2017). https://doi.org/10.1007/s12599-017-0467-3

[37] Buterin, Vitalik. "Ethereum white paper." GitHub repository 1 (2013): 22-23.

[38] D. Vujičić, D. Jagodić and S. Ranđić, "Blockchain technology, bitcoin, and Ethereum: A brief overview," 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina, 2018, pp. 1-6, doi: 10.1109/INFOTEH.2018.8345547.

[39] Kugler, Logan. "Non-fungible tokens and the future of art." Communications of the ACM 64, no. 9 (2021): 19-20.

[40] W. Entriken, D. Shirley, E. Evans, and N. Sachs, "ERC-721 non-fungible token standard," 2018, accessed 2019-10-12.

[41] M. d. Angelo andG. Salzer, "Tokens, Types, andStandards: Identification andUtilization inEthereum," in 2020 IEEE International ConferenceonDecentralizedApplicationsandInfrastructures(DAPPS), 2020,pp.1–10.

[42] Bretto, A. (2013). Hypergraph Theory. Mathematical Engineering. doi:10.1007/978-3-319-00080-0

[43] Zheng, Peilin, Zibin Zheng, Jiajing Wu, and Hong-Ning Dai. "Xblock-eth: Extracting and exploring blockchain data from ethereum." IEEE Open Journal of the Computer Society 1 (2020): 95-106.