IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 24, 2020, accepted November 8, 2020, date of publication November 17, 2020,
date of current version December 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3038550

Steem Blockchain: Mining the Inner Structure of the Graph

BARBARA GUIDI, ANDREA MICHIENZI, AND LAURA RICCI
Dipartimento di Informatica, Università di Pisa, 56127 Pisa, Italy
Corresponding author: Barbara Guidi (guidi@di.unipi.it)

This work was supported in part by the European Commission under Contract H2020-825585 HELIOS.

ABSTRACT Since their introduction, Online Social Networks (OSNs) have transformed the way people interact with each other. Lately, a new trend is rising in the development of OSNs, fueled by an increasing interest of the blockchain technology and the benefits it can bring to the world of OSNs. Blockchain Online Social Media (BOSMs) are Social Media applications that are supported by the blockchain technology. Thanks to a blockchain, BOSMs either try to enforce the privacy of the users or try to redistribute with their users the economic wealth generated by the platform through a rewarding system. There are countless BOSMs available which incorporate a rewarding system. Among them, Steemit can be considered the most well-known platform exceeding 1 million registered users. Steemit is supported by the blockchain Steem, which is a blockchain that natively supports the development of social applications by the usage of transactions that model social activity. Even if other important blockchains, such as Ethereum has been widely analysed, at the best of our knowledge, no study exists concerning the topology of the transactions graph of Steem. The main goal of this paper is to study the structure of the Steem transaction graph to understand its characteristics and unveil crucial knowledge concerning their users. More in detail, we build the Interactions Graph and, after its study, we evaluate three subgraphs that capture its social and monetary aspects. The degree distributions of the graphs follow a power-law. Additionally, we detect a substantial number of bots that offer paid services on the platform among the most active users. Lastly, the investigation of the four analysed graphs through a bow-tie structure, suggesting that half of the users have a passive social behaviour and that 80% of the users tend to accrue economic value.

INDEX TERMS Blockchain, bow-tie, graph analysis, online social media, online social networks, Steem, Steemit.

I. INTRODUCTION

People use Online Social Networks (OSNs) to share their personal information, as a daily activity. Today, the number of Internet users is more than 4 billion, and the number of social media users exceeds 3.5 billion, increasing year by year. Current popular OSNs are centralized which means they are based on centralized servers storing all the information of the users. The centralized structure has several drawbacks. From the user's perspective, the major drawback is the problem of privacy because data can be managed, sold, or stolen without active control of the owner of the data. The major scandal which involved Facebook, one of the most used Social Network, is the Cambridge Analytica's scandal.[1] About 87 millions of Facebook users used an application published on Facebook which was able to collect profiles of users and friends. Data were delivered to Cambridge Analytica which analysed them for political goals. This is one of the main examples of privacy disclosure, but it is not the only issue regarding the usage of OSNs. Indeed, censorship has become an important issue during the last years. Facebook has been banned in some countries, such as China, Tunisia, and Iran. The aforementioned problems represent the main motivations that have lead to an evolution of the social services towards decentralised implementations. A Distributed Online Social Network (DOSN) [1] is an Online Social Network implemented on a distributed platform, such as a

The associate editor coordinating the review of this manuscript and approving it for publication was Hocine Cherifi.

[1] https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram

network of trusted servers, P2P systems, or an opportunistic network. During the last ten years, several DOSNs have been proposed [2]–[6]. These platforms represent the first evolution towards a new generation of Online Social Networks and focus on user's privacy preservation.

However, decentralization techniques have radically changed during the last few years, in particular when the Blockchain technology has been taken into account in several research fields, as the main revolution to overcome several issues concerning the centralization. A blockchain is essentially a public distributed ledger of records that are shared among participating parties, and it can be referred to as a chain of blocks. The first major application of the Blockchain technology was Bitcoin [7], which can be considered the motivation of why blockchains are so famous today. The other major application is Ethereum,[2] which was launched in 2015 with the novelty of *smart contracts*, pieces of code describing self-executing contracts with the terms of the agreement between two parties.

The lack of success of DOSNs, and the increasing problems concerning OSNs, such as fake news or data disclosure, has been the primary motivation to combine social platforms with the blockchain technology. Several Blockchain Online Social Media (BOSMs) have been proposed, motivated by the need to give value to the content generated by the users. Social Networks and Social Media represent a goldmine of data, which are usually used by centralized providers to enrich themselves. Instead, BOSMs provide a way to redistribute the wealth generated by their users by the means of a reward granted to the creators of contents which positively engage other users. These rewarding systems are usually based on the attention economy and/or token economy [8], [9]. In these systems content creators are awarded cryptocurrency whenever they are able to generate a relevant social activity and attract the attention of other users, usually in the form of positive feedback (the equivalent of *likes* in Facebook). All users can express their feedback and this can be awarded additional cryptocurrency. Nevertheless, the real benefit introduced by the blockchain in a social environment is still unclear, because the behaviour of these platforms is unknown due to the lack of a real analysis, and in some cases, the introduction of the rewarding process seems to change the real behaviour of the social platform [10] (for example the usage of bots to retrieve more tokens).

Steemit,[3] with more than 1 million users is the most used BOSM. It is based on the Steem blockchain, which differs from other well-known blockchains, such as Ethereum and Bitcoin, not only for its social nature, but in particular for its performances. Indeed, it is a blockchain that is explicitly designed for the development of social networking platforms, providing 38 different transactions to cover all the most important social actions included in a social media. Moreover, it uses a Delegated Proof of Stake (DPoS) consensus protocol, which makes possible the creation of a new block every 3 seconds. Even if Bitcoin and Ethereum are well-studied [11]–[15], the characteristics of Steem, and how social and economic characteristics can affect each other is still unknown.

The goal of this paper is to study the inner structure of Steem by studying the Steem transactions graph to discover particular characteristics of the economic and social nature of it. A way to study the structure of the graph is the use of a well-known structure: the bow-tie structure [16], which has been used to study the structure of various famous graphs, such as the Web graph [16]–[18]. Our study can be divided into three different phases:

- **Data Collection.** We download the Steem blockchain and parse the blocks to extract the transactions. We downloaded the Steem blockchain from the 8th of March, 2017 to the 1st January 2019;
- **Graph Creation.** We build the Transactions Graph modelling each transaction obtained in the previous step as an edge. In this paper, we study the Transactions graph and three relevant sub-graphs: the Social Interactions Graph, built by considering only the transactions with a social meaning, the Social No Bot Interactions Graph, which is obtained by removing the identified bots from the Social Interactions Graph, and finally the Monetary Interactions Graph which is created considering only the transactions belonging to the economic side of the platform.
- **Bow-tie analysis.** We provide relevant analyses concerning the four graphs, and we study their bow-tie structure.

Our analysis shows that the majority of the users has a strong interest in the social side of the blockchain. This happens mainly because Steemit introduces a rewarding system that offers economic rewards to some social actions, performed by users, and their impact. It is extremely important to understand that, differently from other blockchains, such as Bitcoin, where transactions model economic transfers, in Steem a transaction can be used also to model a social action, such as a post or a comment. But, thanks to the nature of the platform itself, since the social actions are eligible for an economic reward, the social and economic sides of the platform are highly intertwined and equally present.

Finally, in this work, we study and compare the structure of the Steem transaction graph with the most relevant analogous structures studied in previous works. It is a relevant scenario because transactions in Steem cover both the social nature and the economic nature of the blockchain, and, additionally, bots may have a relevant impact on the structure of the graph. Our study reveals that the relative size of the components does not respect the original bow-tie structure but instead shows that the structure is more complex than the original one. Additionally, we show that the monetary transactions make up a very bizarre shape, suggesting that these transactions are used only by a small fragment of users.

[2]https://www.ethereum.org/

[3]https://steemit.com/

The paper is organized as follows. Section II gives the reader background information concerning DOSNs, BOSMs, and the bow-tie graph structure. Section III presents the most important features of the Steem blockchain, while Section IV explains how we model the Steem transactions with graph theory and the rationale behind the need to produce four graphs. Section V presents briefly the collected dataset. We present a relevant set of analyses concerning the Steem Interactions Graphs in Section VI, and we study the bow-tie structure of the graphs in Section VII. A comparison and discussion concerning the identified structure is presented in Section VIII, and finally Section IX concludes the paper, pointing out possible future works.

II. BACKGROUND

In this Section, we present an overview of the proposals in terms of decentralization of Social Media, focusing on the new blockchain proposals. Additionally, we provide an overview of the bow-tie structure and its application to evaluate the characteristics of a graph.

A. DECENTRALIZED ONLINE SOCIAL NETWORKS

The privacy issue was the main motivation to propose new solutions to overcome all the several problems of current OSNs [1], [19]. Decentralized Online Social Networks (DOSNs) [19] are Online Social Networks implemented by exploiting decentralized networks, such as P2P networks, that provide the decentralization of social services. In a P2P network, there is no distinction between clients and servers. Every user's device can act as both, making them functionally equivalent as peers. During the last ten years several DOSNs have been proposed [2], [3], [5], [20]–[22]. The first big project in this area is Diaspora [23], founded in 2010. The main characteristic of Diaspora is that it is a Federated Online Social Network, formed by independent and federated servers managed by individual users. Another important platform is Mastodon, a popular federated alternative to Twitter, with more than 2 million users. Diaspora and Mastodon are part of Fediverse,[4] a confederation of federated platforms able to communicate between them.

DOSNs represent the first evolution of centralized OSNs. Their impact in the market was not as disruptive as expected, and OSNs still manage to keep their large pool of users. However, the motivations to persevere in the revolution are still there, and there are more technologies at disposal of the researchers. With the evolution of the decentralization techniques, and in particular, with the introduction of the blockchain technology, another class of distributed solutions has been proposed, and DOSNs have been re-thought by exploiting the blockchain technology.

B. BLOCKCHAIN-BASED SOCIAL MEDIA

During the last ten years, blockchain technology has been applied to several research fields, including social media. A Blockchain Online Social Media (BOSM) [24] is a decentralised social media built upon a blockchain. The most important BOSMs proposals, usually heavily rely on the blockchain, and use it for many aspects, such as the implementation of the functionalities, and the storage of data. A blockchain is one of the implementations of a distributed ledger and is a chain of blocks that contains transactions between users of the peer-to-peer network that manage it. It is not stored in a central site, but each network node has an identical copy of it, hence it can be described as a decentralized and distributed ledger of transactions. Blockchain has key characteristics of decentralization, persistency, anonymity, and auditability [25].

The main aim of all BOSMs is to overcome the problems of current OSNs, in particular Facebook. In [24], four common characteristics are identified:

- *No Single-Point of Failure*. BOSMs eliminate central authority and they do not have a single point of failure.
- *No Censorship*. The decentralization of the contents overcome the problem of censorship, which is an important issue in countries like China, North Korea, and Syria.
- *Rewards for Valuable Contributions*. BOSMs usually reward users for their social contributions. The rewarding mechanism is fully transparent and its effect is verifiable because transactions are tracked and audited by everyone.
- *Content Authenticity*. BOSMs face the problem of fake news thanks to the usage of reward strategies. The usage of blockchain technology is useful to treat this problem by using an economic incentive to both rank and reward content.

All current BOSMs proposals are based on these four common points. In detail, the Single-Point of Failure problem is faced by exploiting the decentralised architecture of the blockchain technology. Its decentralised nature and the fact that the data stored in the blockchain cannot be modified help to face the problem of censorship because data is accessible from countless sources and cannot be artificially modified by anyone. Rewards for valuable contents are implemented through rewarding mechanisms which are naturally supported by the blockchain. They usually take deep inspiration from the attention economy [8] and the token economy [9]. Finally, the authenticity of content can be easily achieved if it is stored in a blockchain, because all the involved actors agree on the data that is put in the blocks.

Currently, there are already numerous active BOSMs platforms. The most famous is **Steemit**, which has reached more than 1 million users in the last years, and represents today an important alternative to centralized OSNs. The description of the relevant system features is detailed in Section III.

Peepeth[5] is a Twitter-like application that runs over Ethereum, and it includes: Peepeth, an open-source smart contract running on the Ethereum blockchain (data storage);

[4]https://fediverse.party/

[5]https://peepeth.com/welcome

and Peepeth.com, the front-end for the smart contract. Data is saved to the Ethereum blockchain, and anyone can monitor Peepeth's public data. Instead of a Like button, Peepeth has the once-per-day like, called Ensō. Its rarity makes it more special to receive, and it encourages the creation of "dignified, beautiful, and timeless content".[6] The idea is that, since the number of Ensō is limited to one per day per user, content creators must give their best and create contents that contain truthful and important information. An Ensō is forever because there is no undo. Peepeth is moderated. Moderation is transparent on Peepeth, because of the open nature of the blockchain.

Sapien provides a platform for users to publish, access, and view content. Instead of using different Social Media platforms for different forms of news and media, Sapien provides a common platform for everything: articles, videos, images, and much more. Content can be public or private, which means that the platform guarantees a level of visibility of social data. Social services offered are: add friends, form groups and tribes, public profiles, share and comment on posts. Sapien cuts out the intermediary by rewarding content creators directly through peer-to-peer (P2P) transactions. Other important projects are: **Minds**, **FORESTING**,[7] and **SocialX**. In terms of academic proposals, **BCOSN** [26] represents a BOSM focused on privacy issues, when the blockchain is used as a trusted server to provide central control services.

At the best of our knowledge, these platforms have not been studied in depth yet, in particular, to evaluate their social properties. The main problem is that they are new platforms, and the amount of users is limited or it is not enough for a general evaluation. In other cases like Minds, data are not easily collectable due to the lack of APIs. A platform with a huge amount of users and with APIs to retrieve data is Steemit. In [27] authors presented an empirical study of the witness election process on the platform Steemit. However, this study does not provide any social characteristics of Steemit.

C. THE BOW-TIE STRUCTURE

The name bow-tie was used for the first time to describe the macroscopic structure of the Web [16]. An example of the structure is shown in Figure 1. There are several reasons behind the study of this structure in graphs: from designing new strategies for indexing pages, to understanding the sociology of content creation on the Web, from the detection of roles of the nodes in the network's domain, to estimate the robustness of the topology. The original work shows that the web graph consists of a giant connected component and many disconnected components. However, the bow-tie structure is not exclusive to the graph of the Web. Indeed, we can find many other studies concerning the bow-tie structure in several other applications, such as the Bitcoin transactions graph [28], ownership networks [29], [30], social media [31], and others [32]. In particular, in [28] authors study how users transfer the cryptocurrency in Bitcoin to understand its economic dynamics. Given a directed graph G, and following the definition given in [28], nodes in the graph can be assigned to one of the following sets:

- **DISCONNECTED:** nodes not connected to the giant weakly connected component of G.
- **SCC:** nodes in the giant (largest) strongly connected component of G.
- **IN:** nodes not in SCC and able to reach the nodes in SCC.
- **OUT:** nodes not in SCC and reachable by the nodes in SCC.
- **TENDRIL IN:** nodes that can be reached from IN but cannot reach SCC nor OUT.
- **TENDRIL OUT:** nodes that can reach OUT but cannot be reached from IN.
- **TUBE:** nodes not in SCC, that can reach at least a node in OUT, and can be reached by at least one node in IN.
- **FRINGE:** nodes not in any of the previous categories.

In the original structure [16], was found that the components SCC, IN, OUT, TENDRILS, and DISCONNECTED are the most important ones and that almost the totality of the nodes belong to them. Indeed SCC contained 28% of the nodes, IN and OUT were made of 21% of the nodes each, and the combined size of TENDRILS plus DISCONNECTED is 29% of the nodes. This structure was criticised in [33] where authors claim that the structure depends on the crawling process. Indeed in their study, they obtain components with a different size proportion: SCC is much larger and contains 51% of the nodes, IN contains 32% of the nodes, OUT 6% and DISCONNECTED + TENDRILS is 10%. A number of similar studies found results which support this claim [17], [18], [34], [35]. In particular, in [17] the authors prove that the graph of the web is not *self-similar*, meaning that the IN and OUT components do not show a bow-tie-like structure, mainly because of the absence of an SCC. Additional analyses concerning the structure of IN and OUT brought the authors to conclude that the two components are highly fragmented and shallow. They call the structure they find a *daisy*, where the pistil is the SCC, and the IN and OUT components are the petals. An alternative structure, called *teapot*, is presented in [18], where the authors present the shape of the Chinese Web as a teapot. The Chinese Web is studied at three levels, page level, host level, and domain level, and in the case the Web pages, the corresponding graph has SCC equal to 44% of the nodes, IN makes 25% of the nodes, OUT 15%, and DISCONNECTED + TENDRILS is 16% of the nodes.

III. STEEM: THE SOCIAL BLOCKCHAIN

Steem is a blockchain that supports the construction of a community and social interactions through the reward. As stated in [36], it is the first blockchain that focuses on the social issue, which tries to reward the contribution given

[6] https://peepeth.com/about
[7] https://foresting.io/

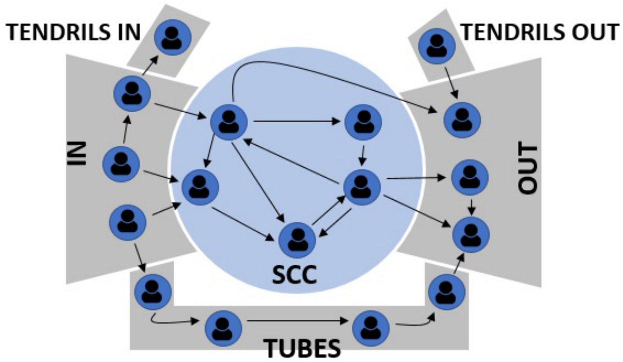

FIGURE 1. Example showcasing the bow-tie structure.

by the community. As described in Section II the economic incentives granted through the rewarding system can help the growth of a new social media platform. The cooperation between cryptomoney and social media thus gives the blockchain Steem a big advantage in the market. The challenge launched by the blockchain developers is that of an algorithm able to score users' contributions, which the choice of community members considers fair. It is a publicly accessible distributed database, which stores the transaction data in text format and distributes the rewards through the network.

One important aspect of the blockchain is that it is based on a multi-token economy, indeed multiple cryptocurrencies coexist in the platform. The three currencies are: `STEEM`, `VEST`, and `STEEM Blockchain Dollars` (SBD) [10]. STEEM is the liquid currency of the system, indeed it can be freely transferred. VEST is a measure of the wealth staked in the platform. It cannot be directly traded, but it gives more power and importance in the platform. Finally, SBD is a stablecoin [37] designed to help newcomers invest in the Steem economic system.

Steem is a blockchain-based on Graphene, the same technology that powers BitShares.[8] It produces new tokens whenever a block is produced. While in the Bitcoin network, and other similar blockchains, everyone can be a miner without anyone's consent, in Steem only a small number of users can do it. Block producers are chosen among a pool of the so-called witnesses. Any user of Steem can become a witness, they just need to set up a witness node. Witnesses are special nodes indeed they can approve forks of the blockchain and they have to publish a price feed, which is the value of the cryptocurrency STEEM. Witnesses can be voted by users, and each vote is based on the stake vested in the platform by the voter. This way, it is possible to define a ranking of the witnesses. The top 20 witnesses, plus one extra witness chosen at random, are also in charge of producing new blocks. The consensus protocol used by Steemit is the Delegated Proof of Stake (DPoS) [38], which enables the possibility to mine blocks at a fixed rate. Blocks are produced once each 3 seconds, which is a crucial property for social applications. Table 1 shows the main properties of Steem.

[8]https://bitshares.org/

TABLE 1. Steem: An overview of the main characteristics.

	Steem
Cryptovalue	STEEM, VEST, SBD
Type of Blockchain	Permissioned
Consensus Algorithm	Delegated Proof of Stake (DPoS)
Block production	every 3 seconds
Fees	No
Block producers	21 witnesses
Release date	24th of March, 2016
1 STEEM value	$0.169

Steem provides a set of 38 transaction types to support its operations and the development of social applications. These transaction types cover all the aspects of the platform, indeed the only way to do something in the platform is through transactions on the blockchain. We categorise the transactions in three sets by taking into account their meaning (see Table 3):

- **Social.** Social transactions are the transactions used to model social actions.
- **Monetary.** Monetary transactions are the transactions used to model economic actions. For example, *transfer* is used to transfer cryptocurrency between two users.
- **Management.** This category includes transactions for creating or recovering accounts, and transactions that can be used only by witnesses as part of their role in the network.

It is through these transactions that users can interact and perform all kind of operations, including the social ones. For instance, through a *vote* transactions, a user can express a positive or negative feedback to a post or a comment which is going to influence the reward that post or comment is going to get. The positive feedback is called up-vote, while the negative feedback is called down-vote. The *comment* transactions can be used to publish a post or to create a comment to an existing post. A special mention goes to *custom_json*, which is a multi-purpose transaction that is used to implement several operations, such as the follow and resteem operations. A follow operation can be used by a user to express the interest to see another user's contents in their feed. A resteem operation is instead used to grant additional visibility to a content, in a similar way to the share function in Facebook. A user that resteems a content, will make that content appear in the feed of its followers possibly increasing the audience of that post.

We show in Table 2 the most important features of Bitcoin, Ethereum and Steem. Ethereum is the only blockchain that natively supports the development of smart contracts, however it must be noted that smart contracts can be written for both Bitcoin and Steem using side-chains that support them such as RSK[9] or Steem Engine.[10] Bitcoin has only one transaction type that corresponds to the transfer of cryptocurrency from a set of sender accounts to a set of receivers account, also Ethereum has only one transaction but whether is a cryptocurrency exchange, a smart contract deployment

[9]https://www.rsk.co/rsk-blockchain/
[10]https://steem-engine.com/?p=faq

TABLE 2. The most important features of Bitcoin, Ehereum and Steem.

Blockchain	Year	Smart contracts	Number of transactions	Consensus protocol
Bitcoin	2009	NO	1	Proof of Work
Ethereum	2015	YES	1	Ethash, a variant of Proof of Work
Steem	2016	NO	38	Delegated Proof of Stake

or a smart contract activation depends on the content of the transaction. As we described earlier in this Section, in Steem there are 38 transaction types, and each transaction covers a different action that can be performed. Concerning the consensus protocols, both Bitcoin and Ethereum use variants of the Proof of Work, which consist in the computation of the inverse of a hash function, a `sha-256` based function for Bitcoin and a `KEC512` based function for Ethereum. The idea, in both cases, is that inverting a hash function is a computationally hard task and succeeding in doing so should be awarded the possibility to create a new block for the blockchain, therefore claiming a block creation reward. Steem instead uses a completely different approach which is based on a set of witnesses and the stake of the users, as we explained earlier in this Section.

IV. STEEM INTERACTIONS GRAPH DEFINITION

The first step of our work is the creation of the Interactions graph. In the Interactions Graph, each node represents a user and each edge represents the fact that the two connected users interacted at least once. We use the collected blocks of Steem to retrieve the transactions. The transactions provide us with the information about the interactions between users, which are considered to create the Interactions Graph. We clarify that edges are directed and that they follow the same direction of the transaction, meaning that if a transaction has the user u as source user and user v as destination user, the resulting edge will have u as a source node and the user v as the destination node. Therefore, if in the dataset we find a transaction from u to v, in the Interactions Graph we only add a directed edge from u to v (i.e. edges have the same directions of transactions). Lastly, we remove the nodes whose only edges are self-loops, while we keep nodes that have also connections with other nodes. The decision to remove nodes with only self-loop edges was made because they are disconnected from the rest of the graph and therefore they will have no impact on the structure we aim to find.

Our analysis consists of the study of the topology of the Interactions Graph. To evaluate the specific characteristics of the Steem Blockchain, we decide to evaluate the whole Interactions Graph, and then, to build and study three different subgraphs: Monetary Interactions Graph, Social Interactions Graph, and Social No Bot Interactions Graph (Figure 2). The need to study the different subgraphs reflects the various natures of Steem. Being a blockchain specifically designed for social applications, Steem has an intrinsic social nature, but, considering the cryptocurrencies and the rewarding system, it has also a monetary nature.

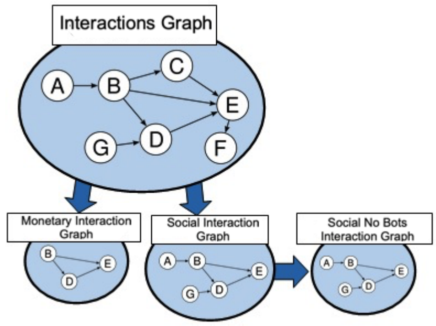

FIGURE 2. Graph construction recap.

In order to create the three subgraphs, we use the transactions categories listed in Table 3.

We explain the rationale behind this differentiation:

- In the Social Interactions Graph we consider only social interactions, such as up/down-votes, comments, follow and so on.
- The Social No Bot Interactions Graph is obtained by removing the bots from the Social Interactions Graph. Bots are identified by the means of centrality measures.
- In the Monetary Interactions Graph we consider only economic interactions, such as cryptocurrency transfers or conversions.

With the Social and the Monetary Interactions Graphs we model the two identified natures of the Steem ecosystem. With the Social No Bot Interactions Graph we aim at analysing only the social activity made by actual users. Since bots are an important component of the Steem ecosystem, just consider that some system accounts are bots as well, we expect that the sociality may be extremely conditioned by these users. For all the graphs, we performed some preliminary analyses, along with the main study to assess if the structure is analogous to the bow-tie observed in the literature.

V. STEEM DATASET

The dataset we gathered is made of 29,021,473 blocks. The first block was produced on the 8th of March 2017 (at 17:34), while the last block was mined on the 1st of January 2019 (at 00:40). The number of transactions per transaction type in our dataset is summarized in Table 3 where the number of each transaction type is put side by side its name.

Figure 3 shows the most used transactions. Social transactions are represented with a blue bar, Monetary transactions are grouped and shown with a green bar, Management transactions are grouped and shown with a red bar. The study reveals that the most used transaction is the *vote*

TABLE 3. Transaction types considered in the Social Interactions Graph and in the Monetary Interactions Graph.

Social	Monetary	Management
vote (338,515,801)	convert (65,227)	claim_account (192,006)
comment (75,015,998)	transfer_to_vesting (910,736)	create_claimed_account (38,328)
custom (0)	transfer_to_savings (86,738)	account_create (18,376)
custom_json (158,638,347)	transfer_from_savings (119,084)	account_create_with_delegation (992,173)
delete_comment (420,153)	cancel_transfer_from_savings (88,777)	account_update (2,540,250)
comment_options (7,838,355)	transfer (23,040,994)	change_recovery_account (414)
claim_reward_balance (19,908,316)	limit_order_create (2,861,353)	request_account_recovery (2,231)
	limit_order_create2 (158)	recover_account (1,778)
	limit_order_cancel (731,014)	witness_update (17,602)
	set_withdraw_vesting_route (53,685)	witness_set_properties (9,292)
	delegate_vesting_shares (3,427,459)	account_witness_proxy (13,396)
	withdraw_vesting (256,567)	account_witness_vote (664,497)
	escrow_approve (166)	feed_publish (1,044,777)
	escrow_release (64)	decline_voting_rights (14)
	escrow_dispute (21)	pow2 (22,991)
	escrow_transfer (332)	

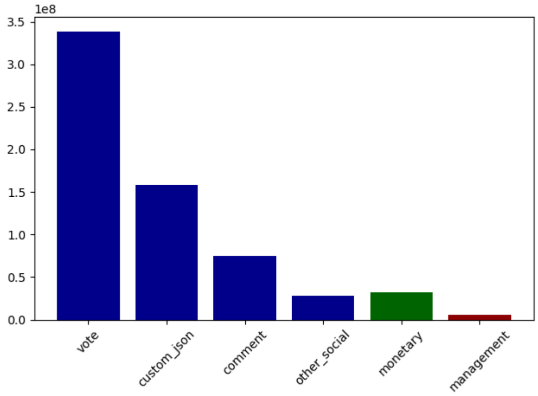

FIGURE 3. Number of transactions per transaction type.

transaction. More than one million users have used this transaction between March 2017 and December 2018, with a total of more than 300 million of transactions. The second most used transaction is *custom_json* that is used in particular for "follow" and "resteem". The third transaction by the number of occurrences is the *comment* transaction which represents not only the real comment action, but also the creation of a post. By distinguishing these two social actions, 22.373.126 transactions are used to publish a post, the 30% of the total, instead the 70% to publish a comment. This analysis is just an expression of the social nature of Steem. Indeed, the three most used transactions concern social actions. Afterwards, we can find *transfer* and *claim_reward_balance*. This reveals the second aspect of Steem concerning the rewarding. Indeed, these two transactions are used to reward users for their content.

VI. STEEM: INTERACTIONS GRAPHS ANALYSIS

The most important features of the Interactions Graph are presented in Table 4. The Interactions Graph is a directed unweighted multigraph with 1.24 million nodes, which represent the users of Steem, and 191.2 million edges, which represent the transactions contained in the blockchain. Approximately 37% of the nodes are sink nodes, with outdegree equals to 0. These nodes represent the users that have signed up for the platform but then never used any of its functionalities. Only 0.0033% of the edges are self-loops, thus we do not expect them to have a large impact on our study. The density is very low, indeed only the 0.012% of the possible pair of nodes has a corresponding edge, which is a result in line with most complex networks. The assortativity coefficient, which measures how much is similar the degree of neighbour nodes, can give information whether the graph has a hierarchical structure or not. For this graph the assortativity coefficient is -0.065, typical of a slightly disassortative network. The observed average local clustering coefficient is 0.395, meaning that the network is highly clusterized, despite its low density and assortativity.

The 1.244.889 nodes are organized in a total of 543.911 strongly connected components. The large number of components, compared to the number of nodes of the graph, suggests us that the structure of the graph is not trivial and that additional analyses are needed to better understand this result. In Table 5 we see the distribution of the components size. In detail, we notice that the vast majority of components is made of just one node, and that 43.69% of the nodes are in their own component. The largest component rallies over 56.3% of the nodes, while the remaining nodes are organised in 102 components most of which are of size 2. This encouraging result is a first hint that the structure of the network may resemble a bow-tie.

Figure 4 shows the in-degree and out-degree distributions of the Interactions Graph. Both distributions follow a Zipf's law, suggesting us that the network we are considering is a scale-free network. The out-degree has a much longer tail, indeed the plot shows that there are user with hundreds of thousands of out-degree. On the other hand, the in-degree distribution has a much shorter tail, but the distribution is noisier. Indeed we can see that there are in particular two in-degree values for which the distribution shows partially outlining values: 10, and 150. We recognise that the anomaly around 10 may be caused by mechanisms used by the platform to

TABLE 4. Properties of the four considered graphs.

Feature	Interactions Graph	Monetary Interactions Graph	Social Interactions Graph	Social no bot Interactions Graph
Nodes	1,244,889	1,074,722	1,230,568	1,230,554
Edges	191,205,168	4,204,473	187,953,267	182,884,418
Sink nodes	465,552	868,379	524,209	524,728
Density	0.000123	0.000004	0.000124	0.000120
Self-loops	626,187	130,637	469,444	469,431
min/max/avg node degree	0 / 1,005,838 / 153.59	0 / 999,576 / 3.9	0 / 765,879 / 152.73	0 / 765,866 / 148.61
Assortativity	-0.065	-0.10	-0,068263	-0.063540
Strongly Connected components	543,911	913,506	559,852	564,540
Largest SCC size	700,842 (56.30 %)	160,949 (14.98 %)	670,580 (54.49 %)	665,869 (51.11 %)
Average local clustering coefficient	0.395			

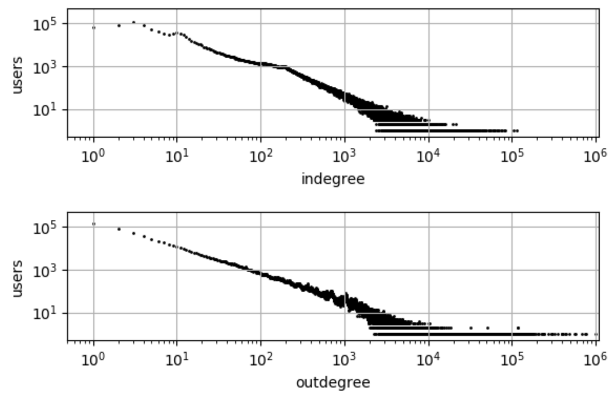

FIGURE 4. In-degree and out-degree distributions of the Interactions Graph.

please a newcomer of the platform: a post to thank the user for signing up in the platform, the delegation of a small amount of VESTs, or other very easily accomplished tasks which would trigger gratification mechanisms. The second anomaly is instead caused by the engagement and dedication required for a user to be successful. Probably a certain amount of users, not seeing a concrete revenue for their actions and failing to engage more people, decided to abandon the platform.

A. MONETARY INTERACTIONS GRAPH

In this section we present a more detailed study about the Monetary Interactions Graph. The transactions considered for building this graph are listed in Table 3.

The Monetary Interactions Graph is much smaller with respect to the Interactions Graph, indeed it has 200k less nodes and 186M less edges (see Table 4). It is interesting to notice that the majority of the nodes are still included in the graph, meaning that many users had the chance to experiment the economic aspect of the social platform. Indeed users can obtain cryptocurrency of the systems either by rewards, or by buying them from external exchanger services. The lower number of edges impacted also on the density which is much lower. In particular, in this graph we consider the *transfer* transactions, which are the most common monetary transactions, and we count approximately 23M of them, but only 4M edges in the graph. The fact is a clear sign that most of the transactions happen between the same nodes. A very interesting result is that sink nodes have increased, despite the lower number of nodes, again as consequence of the lower number of edges. This means that approximately 80% of the nodes were recipient of at least one monetary transaction, but never issued any monetary transaction towards other nodes. This shows that users try not to spend their savings, but they prefer to accrue it. A surprising result is that the maximum degree is only slightly lower than the one found in the Interactions Graph, uncovering that a user performed monetary transactions only. This is a direct consequence that some users tend to send small amount of STEEM for advertisement purposes. Together with the transfer, users can send a direct messages to other users and advertise a service, such as a resteemer service or an upvote service. There are more than 900k strongly connected components, most of which must be made of a single node, indeed the largest component is made of 160,949 nodes. The fact that the largest strongly connected component comprehends only a small fraction of the nodes, may have a big impact on the bow-tie structure of the graph.

TABLE 5. Distribution of the components size.

Size	Number
700,842	1
13	1
10	1
4	4
3	7
2	89
1	543,809

TOP NODES BY DEGREE

Figure 5 shows the in-degree and out-degree distributions of the Money Interactions Graph, which both follow a Zipf's law, highlighting the scale-freeness of the graph. However the two distributions have a different tail, indeed the maximum in-degree is 41,542 and the maximum out-degree is 999,576. We decided to investigate more in detail the most central nodes to have a clearer idea of who are the actors around which the graph is structured.

To do so, we inspect the top nodes per in-degree and out-degree, as shown in Figure 6, where for each node we show its in-degree and out-degree. The top node for out-degree

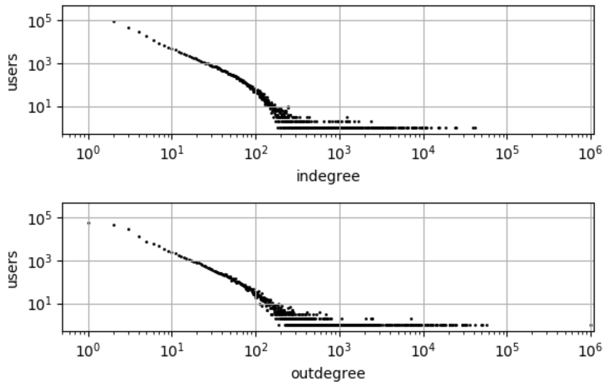

FIGURE 5. In-degree and out-degree distributions of the Monetary Interactions Graph.

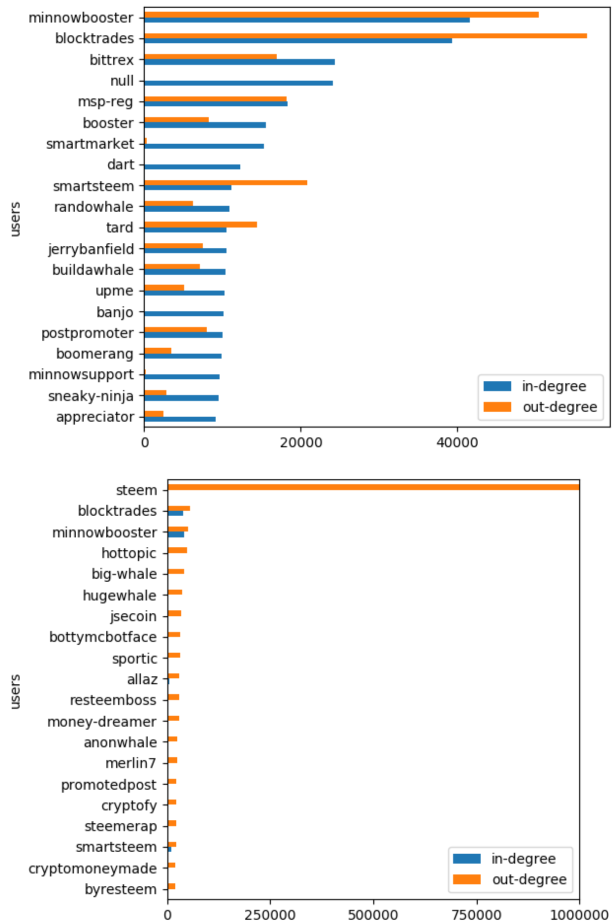

FIGURE 6. Top nodes per in-degree (top) and out-degree (bottom) in the Monetary Interaction Graph.

is the account `steem`, which is a system account. The reason behind its very high out degree lies in the fact that it usually lends few VEST to newly registered users. This operation is also referred to as VEST delegation, and it consist in a temporary loan of VEST which must be returned in full when the lender request them back. VESTS are perhaps the most impactful currency from the social point of view, because they give more power in the case of votes and resteems. While VEST can be delegated to another account at any time, the system account `steem` usually delegates few VEST to new users to make them more visible at the very first stages of an account lifetime. Other accounts that appear in the list, such as `minnowbooster` and `bottymcbotface` offer similar service, but sometimes they require payment, which is the reason why they were not as active as `steem`. The account of the node with second highest out-degree, `blocktrades`, is owned by the homonym platform which is a digital currency exchanger, where people can buy and sell STEEM. Most of the other accounts, namely `hottopic`, `cryptomoneymade`, `big-whale`, `hugewhale`, `sportic`, `resteemboss`, `anonwhale`, `merlin7`, `steemerap`, `smartsteem`, `byresteem`, `msp-reg`, `allaz`, and `promotedpost`, are bot accounts which promote by upvoting and resteeming a content upon payment. While this seems counter-intuitive, these are accounts that usually send small fraction of STEEM for advertising their service to other users, indeed in their out-degree we mostly find transfer transactions. Other profiles include `jsecoin` and `steemfuzzy` which send small amounts of STEEM to thank other users for upvoting their posts. Instead, users like `money-dreamer` and `cryptofy` appear in the list because they advertise mini-games or contests they create.

Some of these users, such as `minnowbooster`, `blocktrades`, `minnowsupport`, `postpromoter`, `smartsteem`, and `msp-reg` are also amongst the top users sorted by in-degree. Since these users have very high values of both in-degree and out-degree, we expect them to make the core part of the economic sphere of the platform. New accounts that offer upvote or resteem services appear: `booster`, `randowhale`, `tard`, `dart`, `bittrex`, `appreciator`, `buldawhale`, `upme`, `minnowhelper`, and `boomerang`. Another interesting case is the one of `streemian` which used to offer a partial automation account service to other users. Users were able to subscribe to its service, upon a SBD payment, and the subscribing account would *automatically* cast upvotes to specific contents to receive high curation rewards. In practice this is done through the complex key system of Steemit, in particular using the posting key, which lets a user cast a vote as by impersonating another user. Surprisingly enough, we also find a non bot user, namely `jerrybanfield`, which is a real user producing high quality contents, who sometimes receive donations from his followers. Also user `null` offers a promotion service, but it has its own peculiarities. Indeed it is a system account which, in return of SBD, makes a payer's post appear among the trending posts of the platform on the official website. However the currency received by this account, since the 17^{th} hard fork, is not collected by the account but it instead destroyed such that the balance of the account have no impact on the inflation.

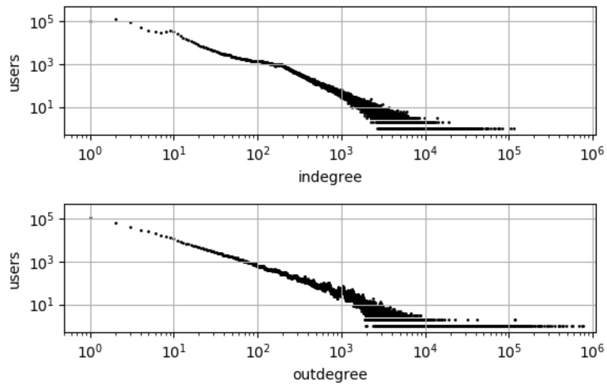

FIGURE 7. In-degree and out-degree distributions of the Social Interactions Graph.

B. SOCIAL INTERACTIONS GRAPH

In this section we present a more detailed study about the Social Interactions Graph. The transactions considered for building this graph are listed in Table 3.

The Social Interactions Graph properties, shown in Table 4, are similar to the one of the Interactions Graph. The graph contains 1.23M nodes and 187.9M edges, which makes up almost the total graph, thus suggesting that the economic incentives to the social actions of the platform greatly helps its usage. The number of sink nodes of the graph is 524,000, or 42% of the nodes; this number corresponds to the number of users that never performed a social transaction. We recall that in this graph the social transactions that result in an edge incident on two different nodes are *vote*, *comment*, and *custom_json*. This result shows that there is a high number of users that manage to attract the social attention of other users, but never performed any social interaction, such as a comment or an upvote. This is the case, for instance, of users that showcase a product on their platform, such as a curation bot. Self-loops, which in this graph are generated in case a user deletes her/his own comment or modifies a post settings, are almost 470,000, meaning that most users that performed a social transaction, also performed one of the two mentioned. The number of strongly connected components and the size of the largest SCC is also similar to the one of the Interactions Graph. This study of the Social Interactions Graph is suggesting that, at a macroscopic level, the graph has a very similar structure to the Interactions Graph.

TOP NODES BY DEGREE

Max values of the in-degree and out-degree are 115.825 and 765.879 respectively. Figure 7 shows the distribution of the out-degree (upper) and in-degree (lower) of the nodes of the Social Interactions Graph. As expected, the two distributions follow a Zipf's law, highlighting the scale freeness of the graph.

We investigate which are the top nodes per in-degree and out-degree, shown in Figure 8, to have a more deep comprehension of the key nodes of the graph. We see that a nice

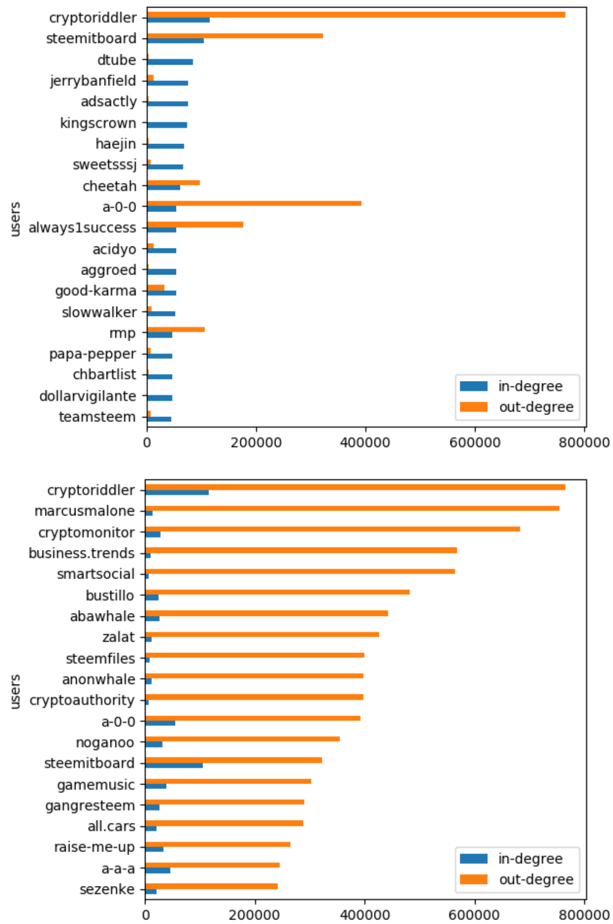

FIGURE 8. Top nodes per in-degree (top) and out-degree (bottom) in the Social Interaction Graph.

variety of accounts appear in the list. `cryptoriddler`, the account with the highest in-degree, belongs to a user that claim to be loving riddles, however the majority of its last contents created on Steemit are about Exode and NextColony, two games implemented over the Steem blockchain. The account `rmp` have similar contents. In this list we also find a number of accounts with different interests and purposes, such as `jerrybanfield`, a youtuber, `a-0-0` and `papa-pepper`, interested in flowers and gardening, `slowwalker`, an historian, `adsactly` is a user that creates a large number of posts concerning literature, arts and blockchains, `acidyo`, which is a person that manually curates posts of Steemit, `always1success` is an account that, on teemit, daily publishes a short list of accounts that produce high quality contents for Steemit, and lastly `cheetah` is a bot account that, when invoked by other users, checks for potential plagiarism on posts published on Steemit. Special mentions go to accounts that are related to the field of economy: `kingscrown` shares a lot of news about cryptocurrencies, and `haejin` evaluates indexes of the stock market and makes predictions about their near-future values. The accounts `good-karma`, `aggroed`, and `teamsteem` are three witnesses, Steem and cryptocurrencies enthusiasts. Steemitboard was a very well known account that, by periodically analysing the transactions that appear in the blockchain, implements an achievement system for Steemit. Lastly, dtube is the official account of DTube, a platform to upload videos, built on the Steem blockchain.

The list of the top 20 nodes per out-degree shows a number of differences if compared to the one found for the in-degree. Three nodes have already appeared in the top nodes by in-dgree, namely critoriddler, steemitboard, and a-0-0. We find several services which resteems and/or upvotes contents behind payment: abawhale, anonwhale, gangresteem, all.cars, and raise-me-up. Other accounts which probably offer similar services because they only resteem posts from other users are zalat and a-a-a. The users marcusmalone, business.trends, and smartsocial are very special because they mainly resteem contents from a specific user called socky. We also find some other canonical accounts, such as cryptomonitor for food and cryptocurrencies, bustillo and noganoo which cover many topics, such as blockchain, economy, programming, ecology, agronomy and gardening, cryptoauthority which makes posts about cryptotrading and cryptocurrencies, gamemusic which is interested in videogames' music, and the aforementioned a-0-0. A special notion goes to the user steemfiles because it lets access premium content, outside the platform Steem, behind payment of SBDs. The other major change is that a new account sezenke appears in the list: a resteem and upvote service like abawhale and anonwhale.

C. SOCIAL NO BOT INTERACTIONS GRAPH

In the top degree rankings we showed in Figure 8, it emerges that the vast majority of nodes are either system and/or bot accounts. The platform explicitly allows and, at a certain degree, it supports the presence of such accounts, however we also wanted a graph in which we consider only the organic social relationship among the users, made by real people. to do so, we decided to produce a subgraph of the Social Interactions Graph, in which we try to remove the most important nodes which we know they refer to system or bot accounts. This subgraph, called Social No Bot Interactions Graph, is obtained by removing from the Social Interactions Graph the most central nodes, by degree centrality, that were also identified as bots. The basic properties of this graph are shown in Table 4.

If we compare the properties of this graph with the ones of the properties of the Social Interactions Graph (see Table 1) we see that removing the identified bot accounts did not have impact also on the total amount of nodes in the graph, which decreased by 14. The number of self-loops decreased accordingly, but only by 13 units, meaning that one of the removed nodes did not have a self-loop. On the other hand, the number of edges decreased by approximately 5 millions, as result of the removal of all the edges connected to the 14 nodes removed (more than 350.000 edges per node

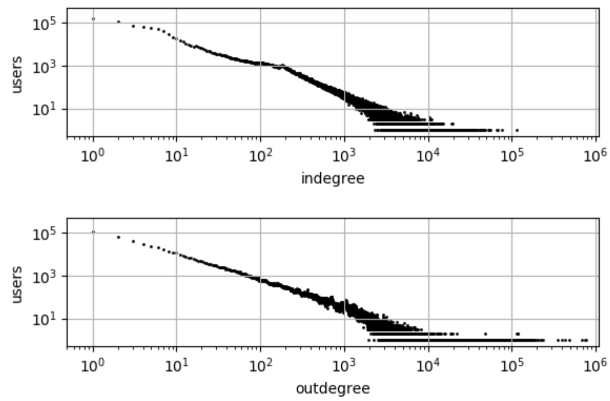

FIGURE 9. In-degree and out-degree distributions of the Social No Bot Interactions Graph.

removed). The density is slightly decreased accordingly to the less number of edges; the number of sink nodes increased by approximately 500 units. The removal of these nodes also caused the average out-degree to drop by 4.1, showing how impactful were the 14 removed nodes over the distribution of the out-degree. The maximum out-degree decreased by few units, indeed the top influential node by out-degree in the Social Interactions Graph was not removed because identified as human. Finally, the number of Strongly Connected Components is increased by 4688 demonstrating that those nodes were very important to the macro-structure of the graph. As direct consequence, the largest SCC lost 5000 nodes.

NODES' DEGREE

The nodes' in-degree and out-degree distributions are shown in Figure 9. Similarly to the case of the Social Interactions Graph, also these two distributions follow a Zipf's law.

VII. STEEM: BOW-TIE ANALYSIS

The core of this study is to verify whether or not the graph induced by the transactions performed by users on Steemit has a botwtie structure similar to the one observed in previous works. to search for this structure we used the same procedure exploited in [28].

Given the presence of two highly intertwined layers in the platform (i.e. social and economic), we decided to study the presence of the bow-tie structure considering them. Indeed, firstly, we performed the analysis considering all the transactions of Steem, thus the induced graph, which was named Interactions Graph in Section IV, which describes a general view of the interactions. Then, we study the structure of the two layers separately: the Social Interactions Graph for the social layer, and the Monetary Interactions Graph for the economic layer. As emerged from Section V, the two graphs show very different characteristics, making it worth to study them and understand how the structural differences impact on the macroscopic topology. Lastly, since we identified a non negligible number of bot accounts, and since bots are capable

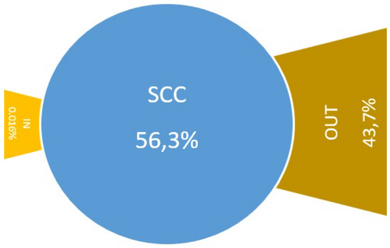

FIGURE 10. Visual representation of the of the bow-tie structure in the Interactions Graph.

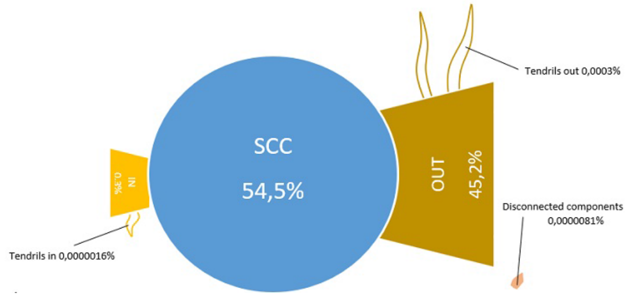

FIGURE 11. Visual representation of the bow-tie structure in the Social Interactions Graph.

of performing a huge amount of transactions compared to people, we also decided to investigate what happens to the structure if key bot accounts are removed, as in the Social No Bot Interactions Graph, introduced in Section IV.

The study of the bow-tie structure of the graphs considered in the paper can be helpful in casting light concerning the user activity on the platform. For instance, in the Monetary Interactions Graph it shows how the cryptocurrency flows in the system. In this case, users in the IN component are users that are able to start the currency flow because they managed to acquire the newly created cryptocurrency. Users in the SCC component are the ones keeping the economy alive, while users in OUT are those users who tend to accrue cryptocurrency with the aim of getting richer. On the other hand, if we consider the Social Interactions Graph, the users in the IN component can identify spammers, stalkers, or other unpleasant behaviours which are detrimental for the platform. Users in the OUT component are instead those that manage to attract social interaction even if they do not belong to the active core of the platform, which is instead represented by SCC.

A. THE BOW-TIE STRUCTURE IN THE INTERACTIONS GRAPH

The size of each graph, in terms of nodes and edges, can be inspected in Table 4, while we show in Table 6 the relative and absolute sizes of the components of the graph (first column).

There are no other nodes outside the three main components, namely IN, SCC, and OUT, meaning that all nodes contribute to the macrostructure and they are in one of the three main components (namely, SCC, OUT or IN). The first fact that stands out is that in the graph we consider almost all nodes belong to just two components: SCC and OUT. The two components have a size of the same order of magnitude, but SCC is bigger by approximately 157,000 nodes. Only a negligible number belong to IN and, surprisingly, other minor components such as TENDRILs, TUBEs, and FRINGEs are absent. There is a high disparity compared to the structure found in the graph of the Internet [16], where nodes where approximately evenly split among the SCC, IN, OUT, and TENDRIL components. The main reason of this effect lies in the fact that in Steemit not all accounts are equally powerful, indeed some specific accounts are needed to perform delicate operations which requires a ''trusted'' account, such as account creation or account recovery. These accounts may belong either to the Steemit team, such as `steemit` and `steem`, but they can also belong to private users, as the case of `anonsteem` which offers a (paid) service to register to the platform. These accounts tend to have many transactions originated from them and that they are connected with a huge number of other users, which will heavily impact on the overall connectivity of the graph. The same effect is emphasised also by other bot accounts we identified in Section V, such as the gambling services, or the account promotion services. Indeed they tend to reach out with transactions a considerable amount of other account in order for them to be considered in a number of ways, such as sending small sums of cryptocurrency or through automatically generated comments. This is the mainly reason to investigate separately the two layers of the platform, with a special focus on the social layer.

B. THE BOW-TIE STRUCTURE IN THE SOCIAL INTERACTIONS GRAPH

The first subgraph we analyse is the Social Interactions Graph which, we recall, it is the graph induced by considering only the social transactions. We show in Table 6 the relative and absolute sizes of the components of the graph (second column).

In this graph, the SCC is approximately 30k nodes smaller, but at the same time IN and OUT grow in size, while other lesser components emerge. Nevertheless, SCC still makes up more than half of the nodes of the graph, followed by OUT with more than 45% of the nodes. The SCC contains the socially active nodes of the platform. In a BOSMs, being socially active does not necessarily mean that the users write a lot of contents on the platform, because also comments, upvotes, and follows are considered social actions that end up in creating an edge in this graph. Thanks to the fact that these users socially interacted with other users and were able to attract social interactions from other users, they become an important part of the community.

We find in the IN component users which did not manage to attract social interactions from the active community of Steemit. This may be caused by the fact that these users are new to the platform or because they are immediately identified by other users as malicious, or spamming and therefore

TABLE 6. Structure of the Interactions Graphs according to the bow-tie components. For each component the relative size is reported, the corresponding number of nodes is reported in parentheses.

Feature	Interactions Graph	Social Interactions Graph	Social no bot Interactions Graph	Monetary Interactions Graph
IN	0.016% (198)	0.3% (3,284)	0.55% (6,773)	1.73% (18,676)
SCC	56.3% (700,842)	54.5% (670,580)	53.79% (661,876)	14.9% (160,949)
OUT	43.7% (543,849)	45.2% (556,303)	44.18% (543,677)	83.2% (894,746)
DISCONNECTED	-	0.000016% (2)	1.44% (17,751)	0.003% (31)
TENDRILS-IN	-	0.000008% (1)	0.00106% (13)	0.002% (27)
TENDRILS-OUT	-	0.0003% (393)	0.0034% (457)	0.03% (293)
TUBES	-	-	0.0006% (7)	-

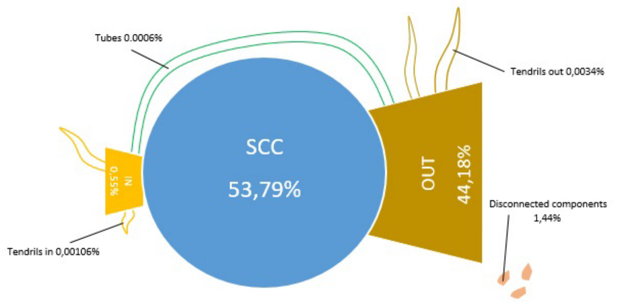

FIGURE 12. Visual representation of the bow-tie structure in the Social No Bot Interactions Graph.

they do not want to waste resources on them. On the other hand, OUT contains users which are not prone to interact with other people, but they somehow attracted attention. Other users in this component include the so-called *early adopters*: users that joined the platform back in the day when it first came out, but then they stopped using it before it attracted the attention of the Internet users.

C. THE BOW-TIE STRUCTURE IN THE SOCIAL NO BOT INTERACTIONS GRAPH

We pose the attention to understand what is the impact of the bots in the social layer of Steemit concerning the bow-tie structure, so we once again categorized the users according to the bow-tie components, but this time we used the Social No Bot Interactions Graph. We show in Table 6 the relative and absolute sizes of the components of the graph (third column).

With respect to the components identified in the Social Interactions Graph, presented in Section VII-B, there are some changes concerning the SCC, OUT and DISCONNECTED components. Indeed, SCC lost approximately 9k nodes and OUT lost approximately 13k nodes, most of which ended up in DISCONNECTED and a small shard ended up in IN. This is an important results concerning the role of a bot in a BOSM. Indeed, we removed only 14 nodes from the graph, which are the identified bots. This is a clear sign of how bots are well-established in the platform, how they are capable of attracting social interactions with other users, and how they actively contribute to the connectivity of the graph. Overall, the macrostructure of the graph has not suffered major changes, indeed nodes still distribute mainly among the SCC (53%) and OUT (44%). While SCC is still made of the active users of the platform, OUT mostly contains nodes with a low to zero out-degree. This is typical behaviour of accounts made for showcasing products and services, which have a lot of followers that are daily active users which comment and upvote their contents. They make content for other users for advertisement reasons, but they hardly interact with their followers. By a closer inspection, we also noticed that there are also users in OUT whose edges originating from them are only self-loops. This is an important result that showcases how much commonly users tend to upvote their contents with the aim to receive higher production reward and a portion of the curation reward for their contents. It also emerges that there are users who understand how their social action can be exploited to increase their personal economic gain on the platform and unscrupulously decide to act with the only aim of maximising their gain, instead of creating a better community.

To have a more clear view of the nature of the users that are now part of the IN component, we decide to take a closer these users and the transactions that they usually make or receive. It turns out that they are mostly inactive users that did not actively participate in the social activity of Steemit in a relevant way, but are more interested in gaining cryptocurrency using the rewards coming from the curation activity of the platform. Indeed, a portion of these users only have upvotes among the transactions issued by these accounts. Sometimes they also correspond to users not able to attract other users' attention, for instance when a user creates contents and comments, but they rarely (by other users in IN) or never get commented or upvoted. Another case is one of the users that are now appearing in the component called DISCONNECTED. Indeed these users have a lot of transactions with users that were formerly in the SCC component, either the 14 nodes that were removed from the graph or other nodes that end up in the SCC as a side-effect of the removed nodes.

D. THE BOW-TIE STRUCTURE IN THE MONETARY INTERACTIONS GRAPH

We tackle the Monetary Interactions Graph for which we foresee there will be major differences because the economic side of the platform is much more tied to specific accounts such as bots and Internet relevant people present on the

TABLE 7. Comparison of the relative sizes of the components.

Graph	#Nodes	SCC	IN	OUT	Other	Name
Interactions Graph	1.2M	56.3%	0.016%	43.7%	0%	-
Social Interactions Graph	1.2M	54.5%	0.3%	45.2%	0.0003%	-
Social No Bot Interactions Graph	1.2M	53.8%	0.5%	44.2%	1.5%	-
Monetary Interactions Graph	1M	14.9%	1.7%	83.2%	0.2%	-
AltaVista [16]	203.5M	27.7%	21.3%	21.2%	29.8%	bow-tie
Italy Web [17]	41.3M	72.3%	0.03%	27.6%	0.07%	daisy
Indochina Web [17]	7.4M	51.4%	0.66%	45.9%	2.04%	daisy
UK Web [17]	18.5M	65.3%	1.7%	31.8%	1.2%	daisy
WebBase [17]	135.7M	32.9%	10.6%	39.3%	17.2%	-
CommonCrawl [33]	3.56G	51.3%	31.9%	6.0%	10.8%	-
Chinese Web (page level) [18]	836.7M	44.1%	25.5%	14.6%	15.8%	teapot
Chinese Web (host level) [18]	16.9M	50.7%	1.4%	47.4%	0.5%	-
Chinese Web (domain level) [18]	0.78M	63.3%	0.7%	34.9%	1.1%	-
Bitcoin [28]	46.1M	86.3%	2.0%	2.5%	9.2%	bow-tie

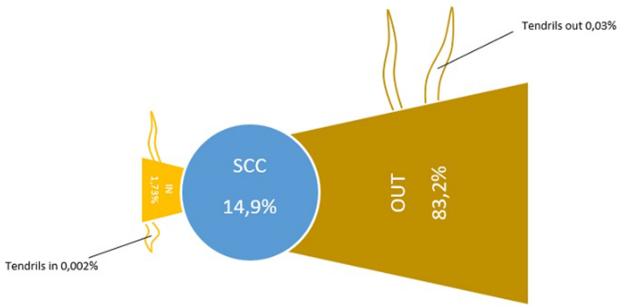

FIGURE 13. Visual representation of the bow-tie structure in the Monetary Interactions Graph.

platform (developers, steemians, and so on). We show in Table 6 the relative and absolute sizes of the components of the graph (fourth column).

As we expected, the proportion of the size of the components obtained in this version of the graph is rather different concerning the ones obtained in the social graphs, which is a clear hint that users behave much differently if we consider the two layers separately. The major difference is that, in the Monetary Interactions Graph, the OUT component is now made of 895k nodes (approximately 83% of the nodes of this graph), while the SCC is now much smaller, barely reaching 15% of the total number of nodes of the graph. Many nodes in the SCC correspond to the accounts that have the highest out-degree and to accounts that promote their services on the platform by sending small sums of cryptocurrency, as explained in Section VI-A. Since these accounts tend to advertise their services massively, it is the case that they sometimes they issue these kinds of transactions also to similar accounts, resulting in creating a strongly connected component. Also, users who then decided to use the advertised services should appear here, because they usually require the payment of cryptocurrency to be accessed, resulting in connecting these nodes to the SCC. On the other hand, users who decided not to have access to such services will be part of the OUT component. The high disparity of the sizes of SCC and OUT is an index that a small fraction of users make use of these services, and that they are not easily affordable to everyone, posing more questions about the rewarding system, its unfairness, and how it is easily cheated. Another relevant node that is part of this component is the one associated with the account `null` because, as we discussed in VI-A, it used to receive cryptocurrency but it is unable to spend it. Finally, the IN component is probably populated by accounts owned by speculators, indeed these accounts do not show any social activity and only send cryptocurrency to accounts owned by exchange services.

VIII. BOW-TIE STRUCTURE: SUMMARY AND COMPARISON

As the last contribution of this paper, we compare the structure of the Interactions Graphs we defined with other analogous structures found in the literature. The comparison is presented in Table 7. In the four graphs we analysed, we see that the relative size of the major components is not respected as in the original paper of the bow-tie structure [16]. On the other hand, we see a very strong similarity between the sizes of the components of the Social No Bot Interactions Graph and the ones of the Indochina Web [17]. The Interactions Graph and the Social Interactions Graph have similar component sizes with a large SCC comprising more than half of the nodes, a barely present IN component, and almost all the other nodes belonging to OUT. Lastly, the Monetary Interactions Graph has a very bizarre shape, with a very small SCC and an enormous OUT. As we showed in Section VI-A, the number of edges of this graph is extremely low, and the number of components is very high, up to the point where, besides the SCC, all the remaining strongly connected components are made of just one node. This is a clear sign that the monetary side of the blockchain Steem is not yet broadly used by its users, and that barely 15% of them are active.

IX. CONCLUSION

In this paper we studied the Steem blockchain and the Interactions Graph, that is a graph induced by its transactions.

Since the blockchain has a social aspect and an economic aspect merged together, we additionally decided to analyse three subgraphs of the general Interactions Graph: the Social Interactions Graph, which is made only considering the social transactions, the Social No Bot Interactions Graph, which is obtained by removing the bots from the Social Interactions Graph, and the Monetary Interactions Graph, which is made only considering the economic transactions. The degree distributions of the graphs follow a Zipf's law and the other features hint that we are dealing with scale free networks. Thanks to a more detailed study of the top nodes per degree, we see that a relevant portion of them are bot users, most of which offer paid services on the platform. Additionally, we focused on detecting whether the structure of the four graphs resembles a bow-tie [16]. Results show that the relative sizes of the components of the original bow-tie structure are not respected, and are more similar to later revisions of the structure.

As future work, we plan to retrieve a larger portion of the blockchain, covering the latest blocks. Additionally, we will analyse more in depth the structure of the graphs to gain insights of the inner structure of each component. We also plan to study the evolution of the structure of the graph through time, to understand which are the key nodes of the graphs. Finally, we will also tackle some advanced problems, such as the study of the rich-get-richer property and the bot detection problem.

REFERENCES

[1] B. Guidi, M. Conti, A. Passarella, and L. Ricci, "Managing social contents in decentralized online social networks: A survey," *Online Social Netw. Media*, vol. 7, pp. 12–29, Sep. 2018.

[2] B. Guidi, T. Amft, A. De Salve, K. Graffi, and L. Ricci, "DiDuSoNet: A P2P architecture for distributed dunbar-based social networks," *Peer Peer Netw. Appl.*, vol. 9, no. 6, pp. 1177–1194, Nov. 2016.

[3] R. Narendula, T. G. Papaioannou, and K. Aberer, "My3: A highly-available P2P-based online social network," in *Proc. IEEE Int. Conf. Peer Peer Comput.*, Aug. 2011, pp. 166–167.

[4] R. Narendula, T. G. Papaioannou, and K. Aberer, "A decentralized online social network with efficient user-driven replication," in *Proc. Int. Conf. Privacy, Secur., Risk Trust Int. Conf. Social Comput.*, Sep. 2012, pp. 166–175.

[5] L. Cutillo, R. Molva, and T. Strufe, "Safebook: A privacy-preserving online social network leveraging on real-life trust," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 94–101, Dec. 2009.

[6] S. Buchegger, D. Schioberg, L.-H. Vu, and A. Datta, "PeerSoN: P2P social networking: Early experiences and insights," in *Proc. 2nd ACM EuroSys Workshop Social Netw. Syst. (SNS)*, Mar. 2009, pp. 46–52, doi: 10.1145/1578002.1578010.

[7] S. Nakamoto and A. Bitcoin, "A peer-to-peer electronic cash system," *Bitcoin*, vol. 4, 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[8] T. H. Davenport and J. C. Beck, *The Attention Economy: Understanding the New Currency of Business*. Brighton, MA, USA: Harvard Business Press, 2001.

[9] A. E. Kazdin, "The token economy," in *Applications of Conditioning Theory*. Evanston, IL, USA: Routledge, 2017, pp. 59–80.

[10] K. Kapanova, B. Guidi, A. Michienzi, and K. Koidl, "Evaluating posts on the steemit blockchain: Analysis on topics based on textual cues," in *Proc. 6th EAI Int. Conf. Smart Objects Technol. Social Good*, Sep. 2020, pp. 163–168.

[11] M. Ober, S. Katzenbeisser, and K. Hamacher, "Structure and anonymity of the bitcoin transaction graph," *Future Internet*, vol. 5, no. 2, pp. 237–250, May 2013.

[12] M. Fleder, M. S. Kester, and S. Pillai, "Bitcoin transaction graph analysis," 2015, *arXiv:1502.01657*. [Online]. Available: http://arxiv.org/abs/1502.01657

[13] D. Ron and A. Shamir, "Quantitative analysis of the full bitcoin transaction graph," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2013, pp. 6–24.

[14] W. Chan and A. Olmsted, "Ethereum transaction graph analysis," in *Proc. 12th Int. Conf. for Internet Technol. Secured Trans. (ICITST)*, Dec. 2017, pp. 498–500.

[15] D. Guo, J. Dong, and K. Wang, "Graph structure and statistical properties of ethereum transaction relationships," *Inf. Sci.*, vol. 492, pp. 58–71, Aug. 2019.

[16] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the Web," *Comput. Netw.*, vol. 33, nos. 1–6, pp. 309–320, Jun. 2000.

[17] D. Donato, S. Leonardi, S. Millozzi, and P. Tsaparas, "Mining the inner structure of the Web graph," *J. Phys. A, Math. Theor.*, vol. 41, no. 22, Jun. 2008, Art. no. 224017.

[18] J. J. H. Zhu, T. Meng, Z. Xie, G. Li, and X. Li, "A teapot graph and its hierarchical structure of the chinese Web," in *Proc. 17th Int. Conf. World Wide Web (WWW)*, 2008, pp. 1133–1134.

[19] A. Datta, S. Buchegger, L.-H. Vu, T. Strufe, and K. Rzadca, "Decentralized online social networks," in *Handbook of Social Network Technologies and Applications*. Boston, MA, USA: Springer, 2010, pp. 349–378.

[20] K. Graffi, C. Gross, P. Mukherjee, A. Kovacevic, and R. Steinmetz, "LifeSocial.KOM: A P2P-based platform for secure online social networks," in *Proc. IEEE 10th Int. Conf. Peer Peer Comput. (P2P)*, Aug. 2010, pp. 1–2.

[21] K. Graffi, C. Gross, D. Stingl, D. Hartung, A. Kovacevic, and R. Steinmetz, "LifeSocial. KOM: A secure and P2P-based solution for online social networks," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2011, pp. 554–558.

[22] S. Buchegger, D. Schiöberg, L.-H. Vu, and A. Datta, "PeerSoN: P2P social networking: Early experiences and insights," in *Proc. 2nd ACM EuroSys Workshop Social Netw. Syst. (SNS)*, 2009, pp. 46–52.

[23] A. Bielenberg, L. Helm, A. Gentilucci, D. Stefanescu, and H. Zhang, "The growth of diaspora–A decentralized online social network in the wild," in *Proc. IEEE INFOCOM Workshops*, Mar. 2012, pp. 13–18.

[24] B. Guidi, "When blockchain meets online social networks," *Pervas. Mobile Comput.*, vol. 62, Feb. 2020, Art. no. 101131.

[25] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *Proc. IEEE Int. Congr. Big Data (BigData Congress)*, Jun. 2017, pp. 557–564.

[26] L. Jiang and X. Zhang, "BCOSN: A blockchain-based decentralized online social network," *IEEE Trans. Comput. Social Syst.*, vol. 6, no. 6, pp. 1454–1466, Dec. 2019.

[27] C. Li and B. Palanisamy, "Incentivized blockchain-based social media platforms: A case study of steemit," in *Proc. 10th ACM Conf. Web Sci. (WebSci)*, 2019, pp. 145–154.

[28] D. D. F. Maesa, A. Marino, and L. Ricci, "The bow tie structure of the bitcoin users graph," *Appl. Netw. Sci.*, vol. 4, no. 1, p. 56, Dec. 2019.

[29] J. B. Glattfelder, "Backbone of complex networks of corporations: The flow of control," in *Decoding Complexity*. Berlin, Germany: Springer, 2013, pp. 67–93.

[30] S. Vitali, J. B. Glattfelder, and S. Battiston, "The network of global corporate control," *PLoS ONE*, vol. 6, no. 10, Oct. 2011, Art. no. e25995.

[31] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, and K. Ross, "Video interactions in online video social networks," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 5, no. 4, p. 30, 2009.

[32] J. B. Glattfelder, "The bow-tie centrality: A novel measure for directed and weighted networks with an intrinsic node property," 2019, *arXiv:1911.00924*. [Online]. Available: http://arxiv.org/abs/1911.00924

[33] R. Meusel, S. Vigna, O. Lehmberg, and C. Bizer, "Graph structure in the Web—Revisited: A trick of the heavy tail," in *Proc. 23rd Int. Conf. World Wide Web (WWW Companion)*, 2014, pp. 427–432.

[34] R. Baeza-Yates and B. Poblete, "Evolution of the Chilean Web structure composition," in *Proc. IEEE/LEOS 3rd Int. Conf. Numer. Simulation Semiconductor Optoelectronic Devices*, Nov. 2003, pp. 11–13.

[35] P. Boldi, B. Codenotti, M. Santini, and S. Vigna, "Structural properties of the African Web," in *Proc. 11th Int. World Wide Web Conf.*, Honolulu, HI, USA. New York, NY, USA: ACM, 2002.

[36] Steemit.io. *Steem an Incentivized, Blockchain-Based, Public Content Platform*. Accessed: Jun. 2018. [Online]. Available: https://steem.com/steem-whitepaper.pdf

[37] M. Mita, K. Ito, S. Ohsawa, and H. Tanaka, ‘‘What is stablecoin?: A survey on price stabilization mechanisms for decentralized payment systems,’’ in *Proc. 8th Int. Congr. Adv. Appl. Informat. (IIAI-AAI)*, Jul. 2019, pp. 60–66.

[38] F. Yang, W. Zhou, Q. Wu, R. Long, N. N. Xiong, and M. Zhou, ‘‘Delegated proof of stake with downgrade: A secure and efficient blockchain consensus algorithm with downgrade mechanism,’’ *IEEE Access*, vol. 7, pp. 118541–118555, 2019.

BARBARA GUIDI received the B.Sc., M.Sc., and Ph.D. degrees in computer science from the University of Pisa, Italy, in 007, 2011, and 2015, respectively. She is currently a Postdoctoral Researcher with the Department of Computer Science, University of Pisa. Her current research interests include distributed systems, P2P networks, complex networks, social network analysis, decentralized online social networks, dynamic community detection, and the blockchain technology. She has been involved in the TPC of several international conferences and workshops. She has been a reviewer of relevant journals, such as IEEE Access, and *Concurrency and Computation: Practice and Experience (CCPE)*. She received three best paper awards from the International Conference DCNET 2013, from the Workshop LSDVE 2017, and from LSDVE 2018. She was a Co-Chair of the Conference EAI GoodTechs 2017, and the several workshops.

ANDREA MICHIENZI received the bachelor’s and master’s degrees in computer science from the Università di Pisa, in ecember 2015 and October 2017, respectively, where he is currently pursuing the Ph.D. degree with the Department of Computer Science. Since April 2018, he has been an Associate Researcher with the High Performance Computing Laboratory, ISTI-CNR, Pisa. His research interests include distributed systems and complex network analysis. More in detail, he is also investigating techniques for community detection throughout time, focusing on exploiting these mechanisms for the development of Distributed Online Social Networks. Moreover, he is starting to investigate other blue-sky research topics, such as dynamics of socio-economic networks and influencer and area-of-influence detection in online social networks.

LAURA RICCI received the M.Sc. degree in computer science and the Ph.D. degree from the University of Pisa, Italy, in 1983 and 1990, respectively. She is currently an Associate Professor with the Department of Computer Science, University of Pisa. Her research interests include distributed systems, peer-to-peer networks, cryptocurrencies and blockchains, and social network analysis. In these fields, she has coauthored more than 100 articles published on international journals and conference/workshop proceedings. She has served as a Program Committee member and the Chair for several conferences. In particular, she has been a Program Chair of the 19th Edition of ‘‘DAIS, International Conference on Distributed Applications and Interoperable Systems.’’ She is the Organizer of the LSDVE, Large Scale Distributed Environments Workshop, held in conjunction with EUROPAR Conference. She has been involved in several research projects. She is also the Local Coordinator of the H2020 European Project ‘‘HELIOS: A Context-aware Distributed Networking Framework.’’