



Analysis of Witnesses in the Steem Blockchain

Barbara Guidi¹ · Andrea Michienzi¹ · Laura Ricci¹

Accepted: 28 March 2021 / Published online: 14 April 2021
© The Author(s) 2021

Abstract

Online Social Networking platforms (OSNs) have become part of people's everyday life and their usage covers the deep-rooted need for communication among humans. During recent years, as people are questioning more and more OSN service providers, a new generation of proposals, based on blockchain became very popular thanks to the ethics adopted by these platforms. Steemit is the most important blockchain-based social networking site, which integrates, as main novelty an economic layer to the social media service. Steemit is implemented on top of Steem which, as in other blockchains, awards miners of the blocks with cryptocurrency. Steem miners, called witnesses, are not chosen based on the solution of a mathematical problem, as in Proof of Work based systems, but must be voted by other users. In this work, we decide to study the witnesses on Steem and their contribution to the social platform Steemit, and their social impact. We performed a set of analyses to shed light concerning their behaviour and to understand how they are socially perceived by other users. Analyses show an important social impact but, at the same time, some of them have a negative social impact. Their discussion is polarized towards content concerning Steem, Steemit, witnesses, and other platforms hosted on Steem.

Keywords Blockchain social media · Online social networks · Steem · Steemit · Blockchain

1 Introduction

Online Social Networks (OSNs) are extremely popular internet services which, and their study was proven useful for solving many problems [3, 8, 9, 26, 28]. More importantly, they have changed the way of how people interact. Indeed, people can build a public profile where share public and private information with others. Current OSNs are free to use, people do not need to pay to share their information, because the owner of these social applications can gain money through advertising, etc. Understandably, this sharing activity has been extensively studied [4, 19, 20, 23] focusing not only on the consequences but also on the underlying network properties amplifying such behaviours.

Several recent scandals, such as the well known Cambridge Analytica scandal [2], have highlighted the weaknesses of OSNs, principally concerning privacy. Researchers and not only, have tried to face this issue by proposing alternative platforms. One of the proposed alternatives consists of the decentralization of social services. Decentralized Online Social Networks (DOSNs) [7] have been proposed and several, public and research, proposals have been produced [1, 6, 13, 14]. However, they did not have a big impact on the daily life of people. Thanks to the blockchain technology and the idea of rewarding valuable content, Blockchain Online Social Media (BOSM) [12] have been proposed by exploiting the most well-known blockchain [11], such as Ethereum, or EOSIO. The most well-known is Steemit,¹ which is based on the blockchain Steem. It combines blockchain with blogging along with a monetary system, allowing participants in the platform to receive rewards for their activities.

While Steemit has been often compared to Reddit, the former financially rewards people for creating and identifying popular content. Altogether, there are three types of rewards for users in Steemit: producer, author, and curator rewards. The producer reward is given to specific users, called witnesses (which are 21 on the platform), that generate the blocks of the blockchain Steem. With the

✉ Barbara Guidi
guidi@di.unipi.it

Andrea Michienzi
andrea.michienzi@di.unipi.it

Laura Ricci
laura.ricci@unipi.it

¹ Department of Computer Science, University of Pisa,
Pisa, Italy

¹<https://steemit.com/>

producer award, users are further incentivised to produce content to become a witness.

Even if blockchain is widely studied, BOSMs are still unclear, at least in terms of how rewards affect the nature of these platforms, which should provide an alternative to the current centralized Social Media applications [12].

Current studies are trying to study more in detail the role of rewards in BOSMs and the social nature of these platforms [5, 15, 16, 18, 22] by exploiting the Steem blockchain, which is publicly available and it has more than 1 million users. In the Steem blockchain, one of the main roles is the witness, which is responsible for creating a block and part of the governance body of the platform. Their activity has been partially studied in [22], however, considering their importance in the platform as a governance body, it is not clear which kind of activity they provide in the platform and if they really are the most trustworthy users.

In this paper, we provide analyses to understand the role of the witness users in the Steemit platform and their social impact on other users. We study their posts and find out that they are mostly covering topics such as Steem, witnesses, Steemit, and similar platforms, and are highly evaluated by the users in Steemit. Moreover, we discover that they tend to mention many users to increase their audience. The analysis of their accounts confirm their social impact and uncover details concerning their economic rewards and reputation.

This paper proposes an extended study of users' behaviour and the content created by users initially proposed in [18]. We extend our previous work by focusing with a much higher level of detail in a specific set of users, the witnesses, which fulfil a crucial set of tasks for the blockchain Steem.

The rest of the paper is structured as follows. We propose in Section 2 the relevant notions concerning blockchain and the most important BOSM proposal available. We describe in Section 3 the main features of Steem, the hosting blockchain, and Steem, the social networking platform. We explain in Section 4 the figure of witnesses in the blockchain Steem. We propose in Section 5 our experimental analyses performed on the set of most important witnesses. We conclude the paper in Section 6 drawing our conclusions and pointing possible future works.

2 Background

In this section, we provide an overview of the state of the art concerning the topic of the paper. Principally, we provided an overview of the blockchain technology and the application of the blockchain technology to Social Media.

2.1 The blockchain technology

A blockchain is a public ledger of all transactions ever executed [25]. All of the transactions and data are stored in

a distributed database. Each time the database is updated, all updates are done together in a batch called a 'block'. Each time a new block is produced/added, it is appended on to all of the previous blocks - hence the name "blockchain". The blockchain is built in such a way that is cryptographically hard to tamper [10, 27].

The Steem blockchain is the publicly accessible distributed database, which records all posts and votes, and distributes the rewards across the network [21]. It is where all of the text content and voting data is stored, and it is where all of the reward calculations and payouts are performed. Blockchains like Steem and Bitcoin produce new tokens each time a block is produced. Unlike Bitcoin, where all of the new coins go to the block producers (called miners) [24], the Steem blockchain allocates a majority of the new tokens to a reward fund called the "rewards pool" [15]. The rewards pool gives users tokens for participating in the platform based on the value they add.

2.2 Blockchain technology and social media

Thanks to the revamp of the technology caused by the massive success of Bitcoin, blockchain technology has been considered in several research fields, including social media. Social media services implemented exploiting the blockchain technology are usually called Blockchain Online Social Media (BOSM) [12]. The most impactful BOSM proposals rely on the blockchain and use it for several aspects, such as the implementation of many functionalities, and the storage of data. A blockchain, which can be easily understood as a chain of blocks cryptographically hard to tamper, is one of the possible implementations of the distributed ledger technology (a collection of records, distributed among a set of nodes in the network). The blocks that make a blockchain contain the transactions between users of the peer-to-peer network that manage it through a P2P protocol. The main innovation is that it is not stored in a single, centralized, site, but each node participating in the P2P network 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 [29].

The main aim of all BOSMs is to overcome the problems of current OSNs, in particular Facebook. In [12], five key characteristics are identified: No Single-Point of Failure, Censorship Resistance, Economic Rewards for Valuable Contributions, Content Authenticity, and Truthfulness.

There are already numerous active BOSMs platforms available online [12, 16], we will list the most relevant ones in the remainder of this section. The most famous is **Steemit**, which has reached almost 1.5 million users, growing by the day, and represents an important alternative

to centralized OSNs. The description of the relevant system features is detailed in Section 3.

Peepeth² is a Twitter-like system hosted on Ethereum, and it includes Peepeth, an open-source smart contract running on the Ethereum blockchain (data storage); and Peepeth.com, the front-end for the smart contract. Data is saved to the Ethereum blockchain, and anyone can monitor Peepeth's data. Instead of a canonical "Like" button, Peepeth users have a once-per-day like, called Ensō. Since it is very hard to obtain it is more special, and it should encourage the creation of "dignified, beautiful, and timeless content".³ 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 content only with truthful and important information. Additionally, an Ensō is forever: there is no way to revoke it. Peepeth is moderated, and the process is transparent because of the open nature of the blockchain.

Sapien provides a platform for users to publish, and view content. Sapien provides a common platform for many different media types: articles, videos, images, and much more. Content can be made public or private, which means that the platform guarantees a level of visibility of social data. All the main social services are offered: add friends, form groups and tribes, build public profiles, share and comment on published media. Sapien cuts out the intermediary by rewarding content creators directly through peer-to-peer (P2P) transactions. Other important projects are: **Minds**, **FORESTING**,⁴ and **SocialX**. For what concerns academic proposals, **BCOSN** [17] represents a very important example of BOSM focused on privacy issues, where the blockchain is employed to provide decentralised access control services.

To the best of our knowledge, BOSMs have not been studied with sufficient details yet, in particular concerning the social properties of their users and the impact of the rewarding system. They are rather new platforms, with a limited amount of users, making them not suitable for a general evaluation. In some cases like Minds, data cannot be easily collectable due to the lack of APIs and documentation concerning how to interpret the data stored in the blockchain. A platform with a relevant amount of users and with APIs to retrieve data is Steemit, which represents a perfect case study given its constant evolution through the last years. In [22] 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 or the witness users. In [16] the authors provide a study of the transactions graph, but their analyses are limited to the comparison of the structure of

the graph with relevant other structures. Lastly, in [15] the authors provide an analysis of the Steemit follower-following graph. They also consider a restricted number of posts, but the analyses are limited.

3 The Steem Blockchain and Steemit

Steem is a social blockchain that grows communities and makes immediate revenue streams possible for users by rewarding them for sharing content. It is currently the only blockchain that can power real applications via social apps, like Steemit or DTube. The Steem blockchain adopts the Delegated Proof of Stake consensus (DPoS) protocol, meaning that the creation of blocks is delegated to a set of accounts, called *witnesses*. All the relevant details concerning witnesses will be given in Section 4. Block creation, as in Bitcoin, produces now cryptocurrency which is not entirely awarded to the block producer, but is instead allocated as follows:

- 65% go to the reward pool, which is split between authors and curators.
- 15% of the new tokens are awarded to holders of Steem Power.
- 10% of the new tokens are awarded to the Steem Proposal System.
- the remaining 10% pays for the witnesses.

The peculiarity of Steem is that it uses the blockchain as support for the storage and every action, being it a social interaction between two users or exchange of cryptocurrency, is modelled as a transaction. This happens thanks to a complex transaction system, which includes 38 different transaction types, which covers all the expected functionalities of a social network (such as writing a post), and all the functionalities of a traditional blockchain (such as money transfers).

Steemit is the main service implemented on top Steem and lets users build their social network through an asymmetric relationship called *follow*. Users have their content organised in a structure called *blog* by the platform. A piece of content can be of two types: a *post* or a *repost*. A post in Steemit is very similar to the ones of other social media platforms and can include text, images, videos, links, and so on. Other users can be mentioned in the body of the post to directly point to them or to increase the visibility and awareness of the post, and posts can be tagged with arbitrary strings to facilitate the search for posts labelled with the same tags. A post can be reposted and it will appear in the reposter's blog alongside the information that the content was reposted and the name of the original poster.

A piece of content can be voted by other users in Steemit, thus providing feedback to the content creator and the other

²<https://peepeth.com/welcome>

³<https://peepeth.com/about>

⁴<https://foresting.io/>

readers on the quality of the content. Users can provide two kinds of feedback, or votes: an up-vote expresses positive feedback; a down-vote expresses negative feedback. Votes are extremely important in Steemit, because they directly influence the reward assigned to a content. The impact of a vote on a content potential reward is called the *reward shares* of the vote. The computation is rather complex, but it is mainly based on two parameters:

- The SP held by the voter added to the SP received through delegation, at the time of voting;
- The weight power of the vote, which will determine whether the vote is an up-vote or down-vote, and how much the vote will influence the reward assigned to a content.

When a content is created it will be editable for 7 days, after which the content is frozen, and the reward is computed. The rewards assigned to a content are split into two parts. One part, called **Posting Reward** goes to the creator of the content, is always at least 50% of the total rewards pool. The other part, called **Curation Reward** goes to the curators. The curators of a content are those users that up-voted the content. The curation reward is not evenly split among curators, but it's based on the voting power of the vote, and the voting position: being the first curator (casting the first up-vote) pays better. Since down-votes decrease the reward assigned to a content, down-voters will not receive any curation reward.

In Steemit, every user has a reputation score used in the platform to measure the amount of value a user has brought to the community. Furthermore, this represents a mechanism designed to help reduce abuse of the Steemit platform. Every new user starts with a reputation score of 25. Then, the reputation of a user can increase or decrease. The reputation of users goes up when his/her content is up-voted by others. A down-vote instead can decrease the reputation. However, only users with a lower reputation score are unable to affect your reputation.

4 Steem witness

The Steem blockchain requires a set of people to create blocks and uses a consensus mechanism called Delegated Proof of Stake (DPoS), as explained before. For this reason, the witnesses are one of the most important social roles in the Steem environment. The election of witnesses is provided by the Steem community, and they are elected through a voting system that is based on the stake of the voters. Every user on Steem can cast up to 30 witness votes, and through these votes, weighted by the stake of the users, it is possible to form a ranking of the witness nodes. This ranking will be used to determine which witness will be able to add new blocks to the blockchain. Witness votes cannot be cast on the same witness, and cast votes can be withdrawn

if a user no longer wants to support a witness. A withdrawn witness vote can be cast again, freely, to any witness node.

The top 20 witnesses will work as full-time block miners, producing a block every 63-second round. A 21st position is shared by backup witnesses, who are scheduled proportionally to the amount of stake-weighted community approval they have. Witnesses are compensated with STEEM Power for each block they create, but, unlike in Bitcoin, where all the newly minted cryptocurrency goes to the block creator, in Steem the rewards are split between some actors. Indeed, only the 10% of the new coins are paid to block producers (witnesses) [21], as we anticipated in Section 3. Formally, a Steem Witness is a person who operates a witness server (which produces blocks), and publishes a price feed of STEEM/USD to the network.

If a witness does not produce a block in his/her time slot, then that time slot is skipped, and the next witness produces the next block.

A witness is paid proportionally to how high they are in the witness ranks, excluding the top 19 witnesses (who get 1 block every 63 seconds). A witness at rank 30 can produce as many as 4 blocks/hour, compared to a witness at rank 50, which may produce less than 1 block/hour.

5 Witnesses analysis

The main goal of this paper is to focus the attention on the witnesses activities to better understand their role in the Steem community and how they have collected popularity in the platform. Our analysis consists of an analysis of the post and resteem in terms of popularity and topic. Furthermore, we analyse their accounts to discover more information concerning their popularity.

We collected the blockchain of Steem, as explained in [18], and additionally, we collect the information concerning the current witnesses accounts updated on the 16th of November 2020, which are 100.⁵ Indeed, Steem shows a top 100 Witness live list and explains that the very round of block production begins with the shuffling of 21 witnesses: the top 20 witnesses (by vote), plus one randomly-selected standby witness. By exploiting the 100 witnesses, we collect their history in terms of posts and resteems, and the account information. We collect 34,944 posts and 28,485 resteems. All the analyses have been provided by taking into account the time. Indeed, we analyze the activity of the witnesses year per year, to follow the guideline provided in our previous work [18].

5.1 Witnesses Blogs analysis

We show in Fig. 1 the bivariate distribution of the number of posts and resteems made by the witnesses over the

⁵<https://steemitwallet.com/witnesses>

considered years. The plots show that in the early days of Steemit (see Figure A), witnesses tend to have a similar number of posts and resteems, meaning that approximately 50% of the pieces of content that appear in their blog is not original but created by someone else and then resteemed (shared) on their blog. It is also interesting that there was a shift through the years (see Figure D) towards a higher number of posts created by the witnesses. Additionally, we notice that witnesses tend to be overall very active users as the combined number of posts and number of resteems often exceeds 365, meaning that they add content to their blogs at least once every day. However, we notice that through the years their activity decreased considerably, probably due to the hard fork that started the Hive community in March 2020.

We proceed by analysing the average number of votes received by the posts and resteems by each witness, which will help us to understand how much they are socially attractive to other users. In this case, as shown in Fig. 2 we notice an even stronger similarity, where the average number of votes received by posts and resteems often exceeds 100. It is also interesting to notice that, contrarily to the number of posts and resteems which decreases over the years, the number of votes received increases over the years, especially in the case of resteems, exceeding several hundreds of votes in 2020 (see Figure D). Interestingly enough, on average, resteems receive more votes as the majority of the points are over the diagonal line. This shows

the huge impact of resteems on the platform, up to the point where it's more profitable and engaging to become a content sharing hub, rather than trying to create original content. This is somewhat an expected behaviour concerning our specific study concerning witnesses, as they are perceived as the most trustworthy users of the network.

We decide to study how many common users and how many witnesses are mentioned in posts (see Fig. 3) and resteems (see Fig. 4) appearing in witnesses' blogs. When analysing these graphs, we must be aware that the number of witnesses that can be mentioned in a piece of content is capped at 100, indeed, as explained at the beginning of Section 5, there are 100 witnesses in total in our dataset. On the other hand, there are almost 1.5 million common users. Concerning the posts, we see that in the vast majority of them only up to 10 witnesses and up to 100 common users are mentioned. Additionally, the number of mentions lowers year after year. This is a sign that, in the early days of the platform, witnesses tried to mention many other users so that their posts would gain additional visibility, and possibly would encourage other users to vote for them as witnesses. However, over the years we see a general decrease in the number of mentions in posts. Similar considerations can be made regarding the number of mentions in resteems, in which we also observe an average lower number of mentions if compared to posts.

Fig. 1 Post-Resteem number bivariate distributions

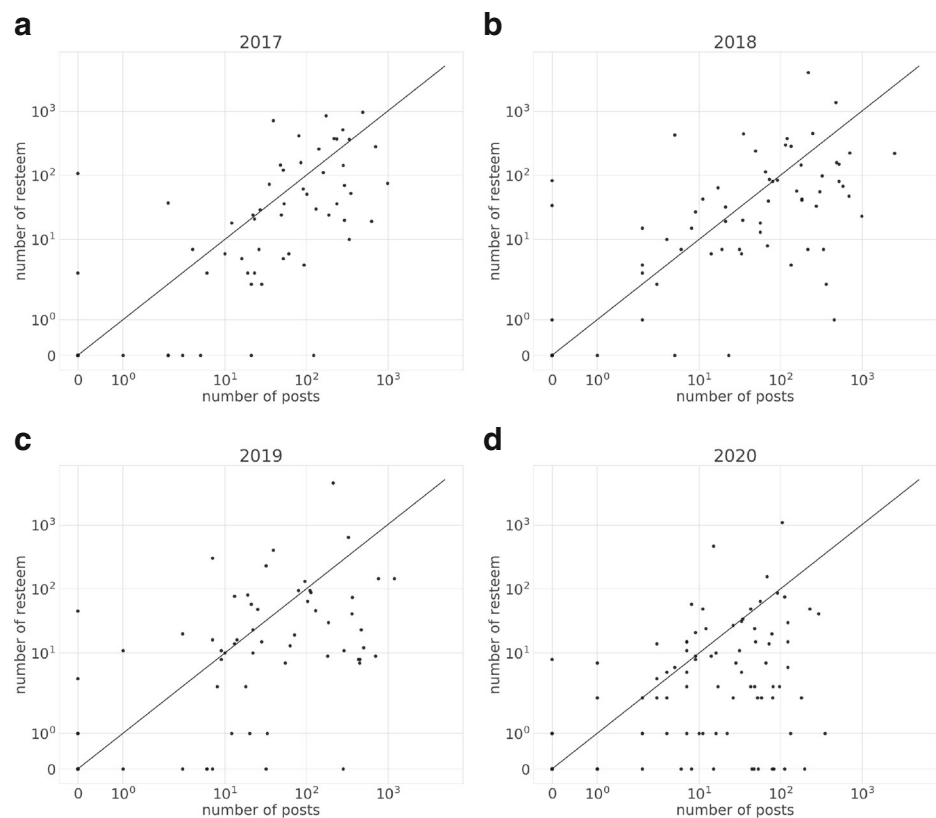


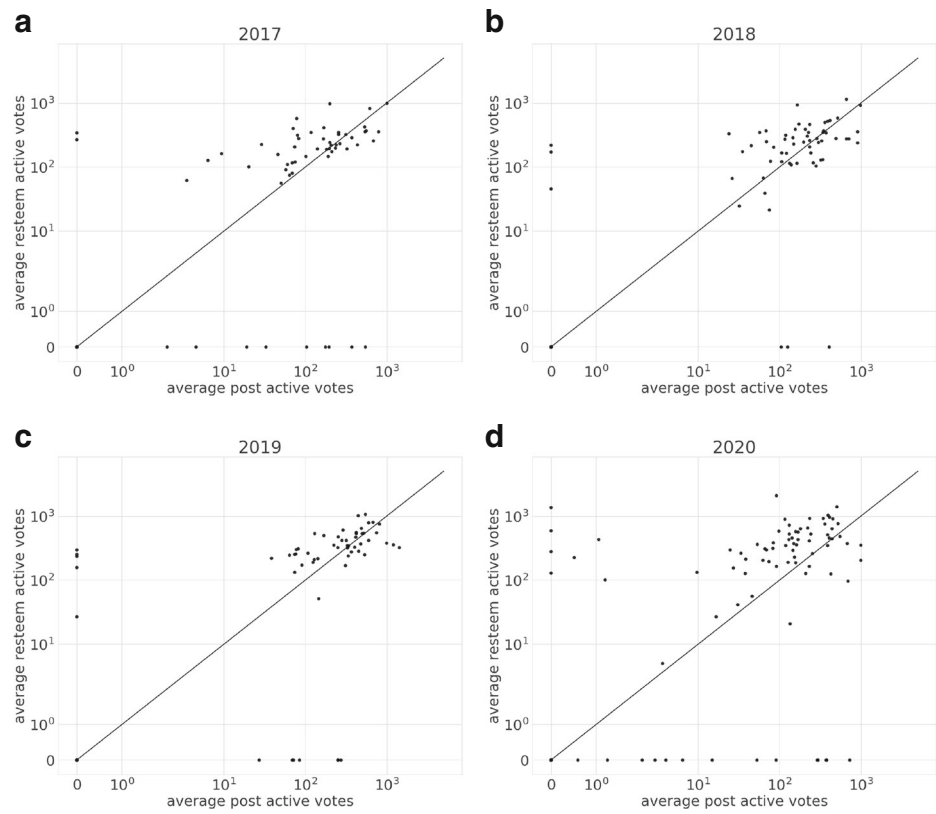
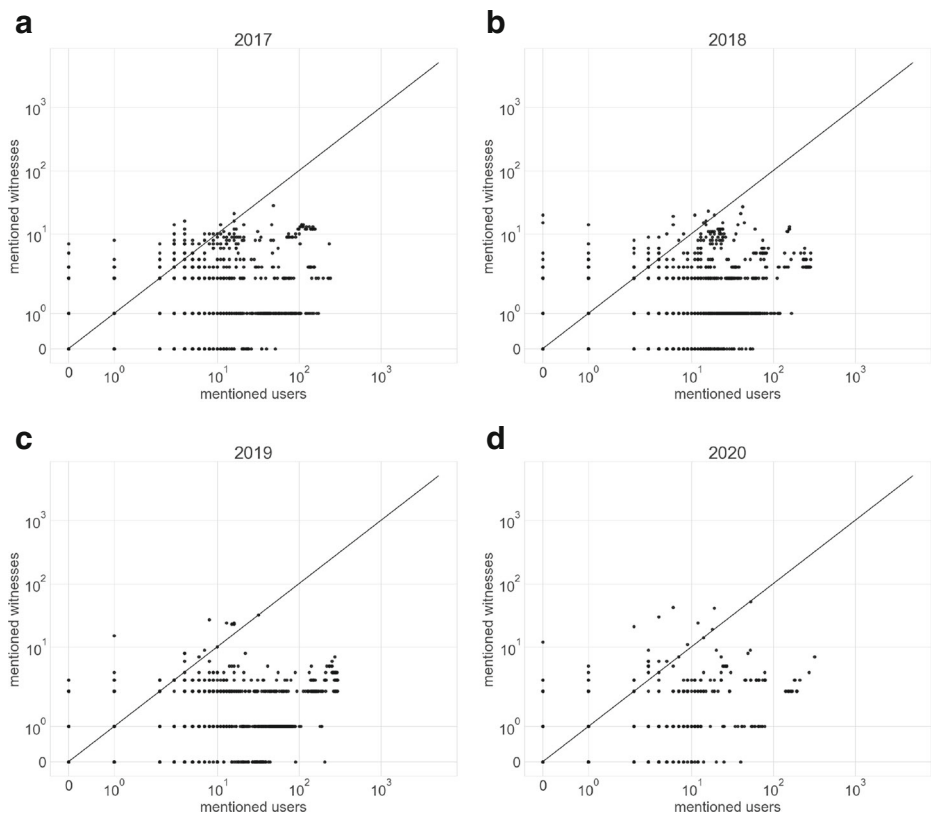
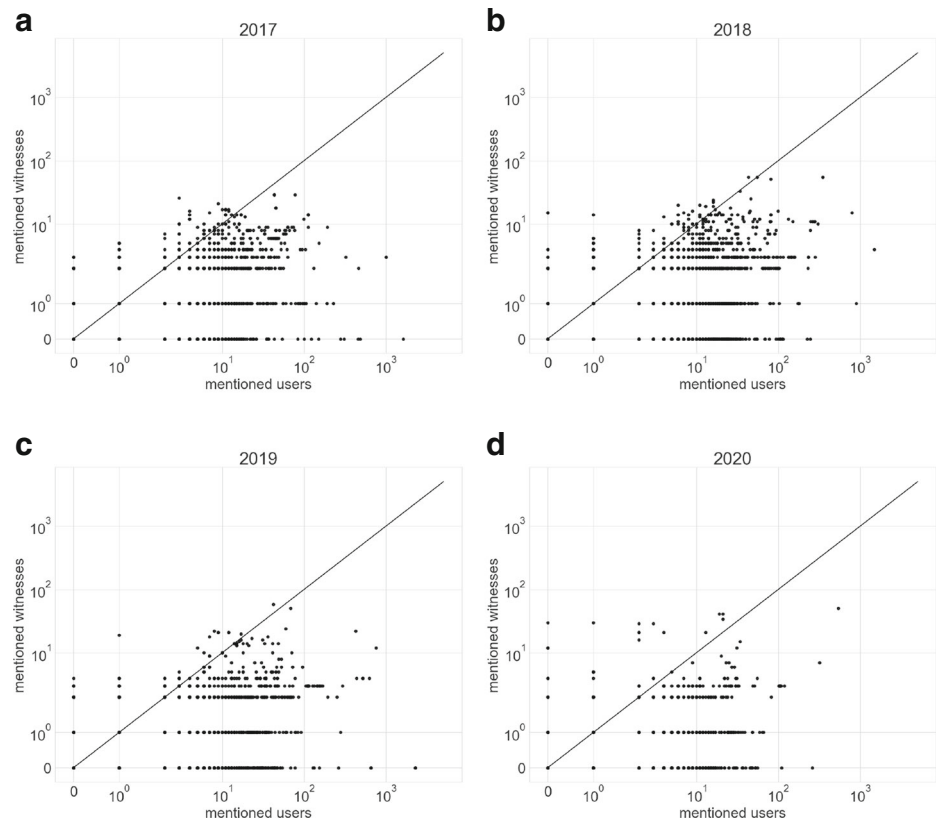
Fig. 2 Post-Resteem active votes bivariate distributions**Fig. 3** Mentions in posts

Fig. 4 Mentions in resteems

Next, we analyse the tags of posts and resteems, which are represented with word clouds in Figs. 5 and 6 respectively. The size of the tags is determined by the number of posts labelled with that tag. Tags in Steemit are the equivalent of hashtags in Instagram and Twitter, and serve the purpose of a creator-driven labelling mechanism. These tags can be arbitrary strings and should help the readers of a piece of content to quickly understand what its topic.

We find a very nice variety of tags in witnesses' posts. For instance, some of them are connected with mundane or personal information, such as *life*, *blog*, *news*. Other ones are more related to the platform itself, such as *steem*, *steemdev*, *steemit*, *community*, and *curation*, or to blockchain in general, such as *money* and *cryptocurrency*. We see also some tags connected to witnesses in general, such as *witness-category*, *witness-update*, and *witness*, or to witness accounts in particular, such as *firepower*, and *cervantes*. Some interesting tags are *sct* (short for SteemCoinpan), *dblog*, *marlians*, *zzan* and *palnet*, which are platforms very similar to Steemit, built on SteemEngine,⁶ or *steemhunt*, which is a platform built on top of Steem. Lastly, there are also some common tags used to label posts in a specific language, such as *cn* for the Chinese language or *kr* for the Korean language.

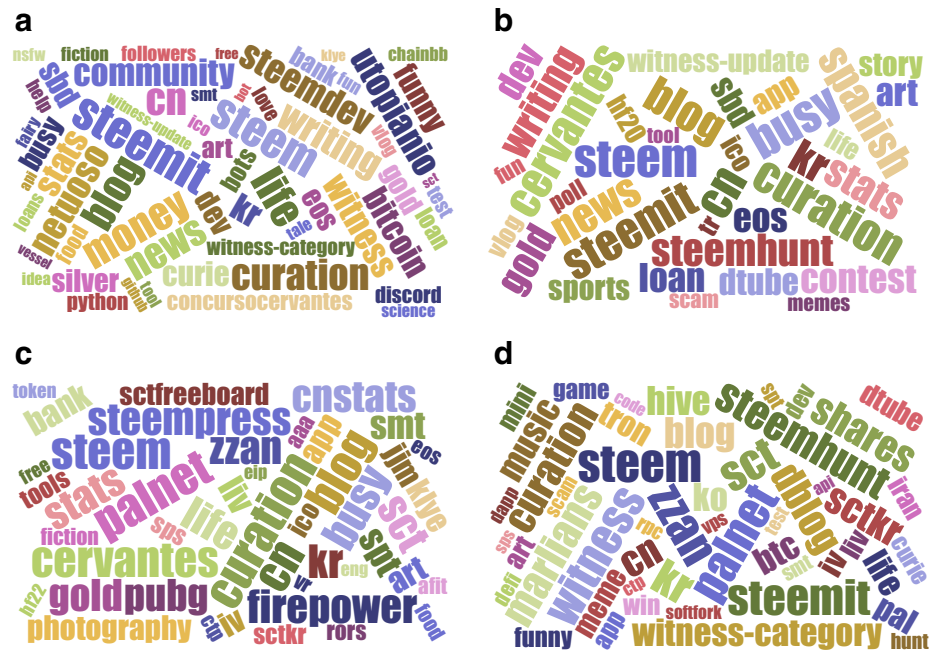
We find similarities concerning the most common tags appearing in resteems. Tags like *steem* and *steemit* are obviously recurring here, but tags like *art*, *photography*, *story*, *writing*, and *music* become fairly more common. A special mention goes to the tag *thealliance*, which is a group of self-regulating Steemit users who try to create interesting content for the platform and use this tag to help each other gaining increased visibility. Another important tag is *neoxian*, which is connected to an exchanger service, a platform built on SteemEngine, a Steem witness, and a Steemit user with high interest especially in the economic side of the platform.

In Fig. 7 we show the boxplots concerning the number of votes received by posts and resteems appearing in witnesses' blogs. The boxes cover from the 25th percentile to the 75th percentile with the 50th percentile highlighted with an orange line, and the whiskers extend to the minimum and the maximum. The plots show that resteems tend to receive a higher number of votes meaning that these pieces of content are usually more socially engaging. Moreover, the number of votes does not show a relevant evolution over time. Interestingly, some of the posts and resteems have a low social impact as they didn't manage to attract a single up-vote, which questions the impact that witness should have to become block producers.

The votes we studied in Fig. 7 does not take into account the fact that there are two types of votes in Steemit

⁶<https://steem-engine.com/>

Fig. 5 Tags in posts



(up-vote and down-vote), and that votes have certain parameters (voting power and SP of the voter) as explained in Section 3. In Fig. 8 we study the distributions of the reward shares of the votes cast to witnesses' posts and resteems. The plots show in both cases that the vast majority of votes are up-votes (positive feedback) and that their weight towards the computation of the reward is similar and quite stable over time. However, it must be noted that, although they are a small fraction, some pieces of content received negative feedback (down-votes), as shown

by the whiskers reaching negative values. This is rather counter-intuitive as the witnesses are supposed to be highly trusted nodes of the network, sharing truthful and relevant information. On the other hand, the reasons why a user, in general, would receive a down-vote are multiple and each vote is usually given for its own unique reason. Lastly, we notice that the order of magnitude of the highest reward shares is similar to the one of the lowest reward shares, probably meaning that very influential nodes cast these down-votes.

Fig. 6 Tags in resteems



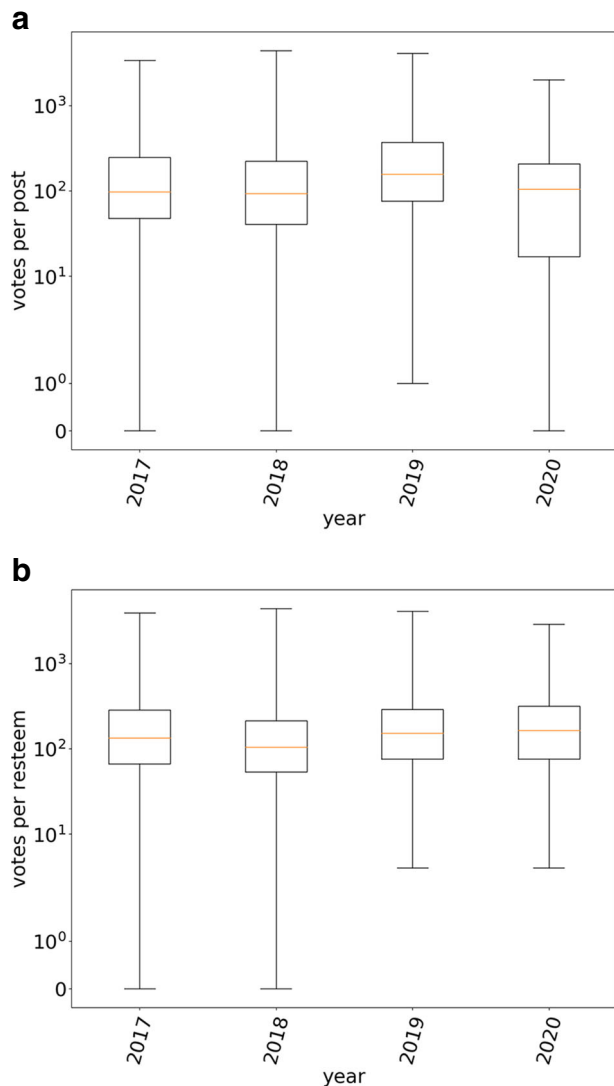


Fig. 7 Votes per posts boxplots

5.2 Witnesses account analysis

We collect all the information concerning the account of each witness to evaluate their characteristics.

Considering the importance of witnesses we decide to analyse their sociality in the platform by exploiting the followers-following information.

Figure 9 shows the relation between followers and followings of each witness account. Witnesses do not follow several other users in Steemit. Indeed they usually have more followers than followings, even if the number of both is small. Only 10% of witnesses have more than 10,000 followers, instead, only four witnesses are involved in the activity of the other users and have more than 3,000 followings.

Another step of our analysis is to compare their activity and the rewards obtained. Figure 10 shows the number

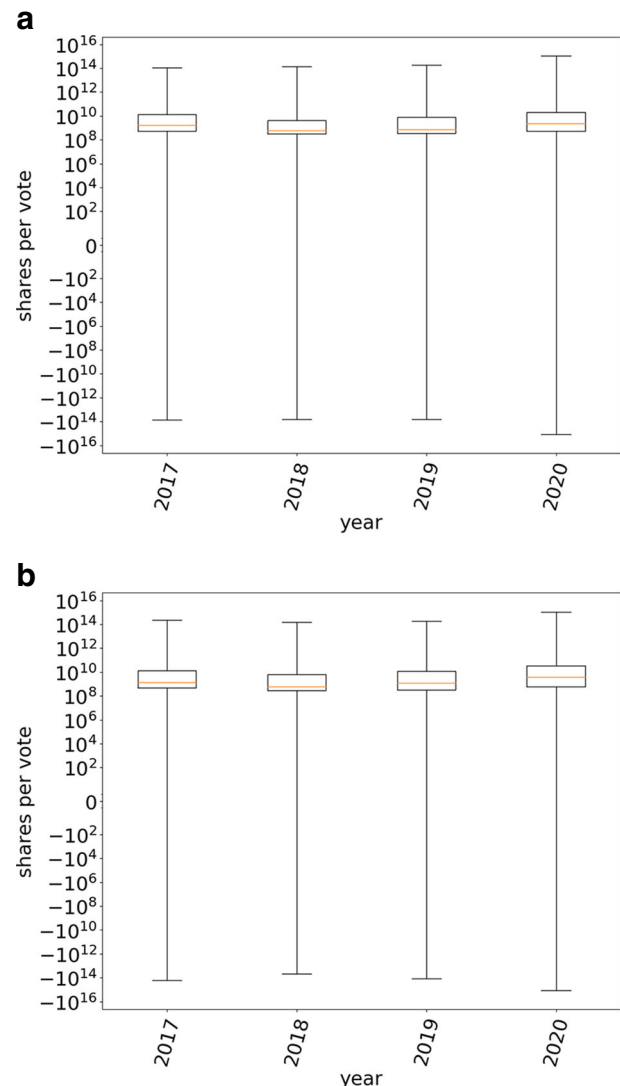


Fig. 8 Shares per votes boxplots

of Steem power gained through the curation and posting rewards considering all the history. Witnesses are involved not only in the curation of content but considering that they

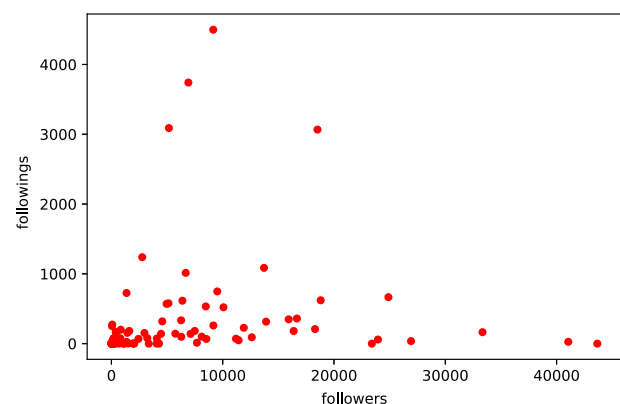


Fig. 9 Follower-following analysis

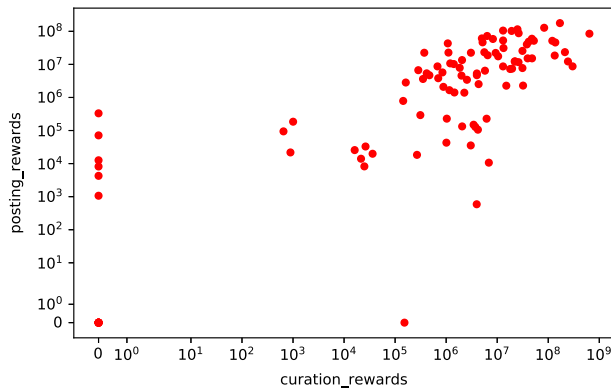


Fig. 10 Curation vs. Posting Rewards

collect high posting rewards, they also provide content to the community.

We decide to investigate more in detail on this aspect by analysing the correlation between the number of following and followers with the total amount of rewards they collected.

Figure 11 shows the bivariate distribution of the curation reward and the number of following accounts for the witnesses. The plot shows that the two distributions are not correlated, indeed the witnesses with the highest curation reward accrued follow a small set of users. This phenomenon may be caused by the fact that witnesses rarely follow other users for social purposes, but they still try to upvote to receive curation rewards. The reason why some witnesses follow a large number of other users may be linked to techniques used to increase the awareness of their profile.

Figure 12 shows the bivariate distribution of the posting rewards received and the number of followers of the witnesses. The plot shows that the witnesses that achieve the highest rewards are usually the ones with the highest number of followers. On the other hand, many witness users succeed in receiving high posting rewards even though

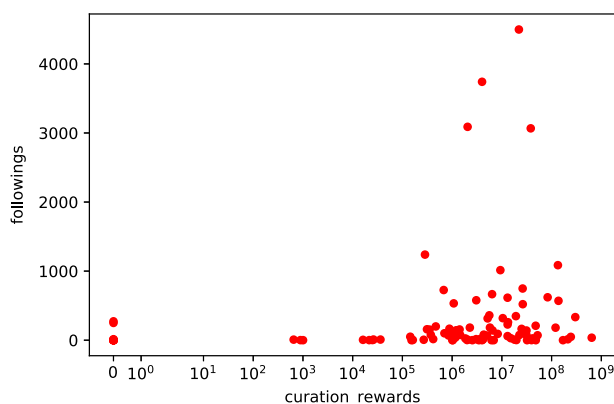


Fig. 11 Curation Rewards vs. Following

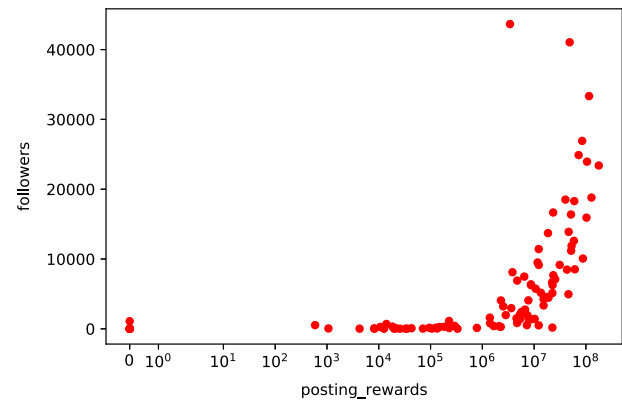


Fig. 12 Posting rewards vs Followers

they have a small number of followers. Multiple reasons can cause this phenomenon, for instance the fact that some witnesses can buy up-votes through the usage of bot accounts, or even the fact that some bots are popular among users but not followed by many.

Finally, we investigate the reputation level of witnesses. Reputation is an important point of the Steem environment. As explained in Section 3, a reputation score is an indicator of many things, including how long the user has been on the platform, how much their content is valued by the whole of the Steemit community, and how much their account has been punished by flags/down-votes. Considering the importance of witnesses in the Steem environment, we expect that the average reputation score is higher than 40. Figure 13 shows the CDF of the reputation value, and we can show that about 20% of witnesses have less than 40, as reputation score. This is an unexpected result, which probably has been affected by the scandal concerning TRON and Steem, when several important witnesses left the platform to build another one.

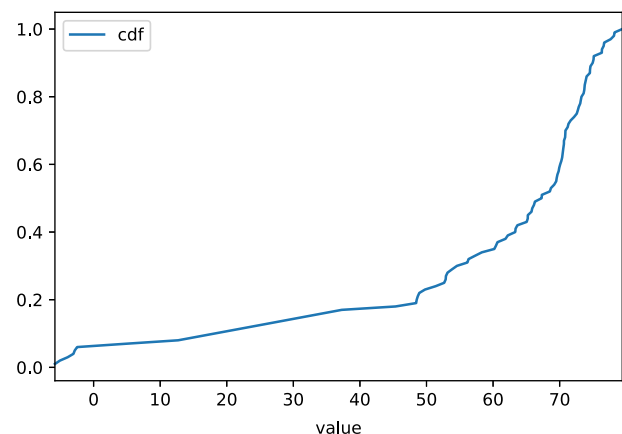


Fig. 13 Witnesses Reputation

6 Conclusion and future works

In this paper we studied the witness users of Steemit, an innovative Blockchain Online Social Media service implemented on the blockchain Steem. The main novelty introduced by the blockchain in this field is the possibility to reward users for their activity on the platform. Of particular interest is the activity of the witness users, considering the very important role they occupy in the platform. We provide two sets of analyses that cover their blogs and their accounts. The blog analyses show that witnesses are socially impactful users, they create lots of content, and the content they create are highly evaluated by the users in Steemit. Moreover, we discover that they tend to mention tens of users in each content, probably to increase the audience of their blogs. The analysis of the tags used by the witnesses shows a certain degree of polarization towards Steem, witnesses, Steemit, and similar platforms. The accounts analyses confirm the social activity of witnesses and uncover additional details concerning their rewards, in particular as content creators and content curators. Lastly, we find that not all witnesses have high reputation, despite their important role.

As future works, we plan to deepen our understanding of the witnesses. One such direction is the study of the text of their posts and performing textual analyses such as key words detection, topic detection, understand whether they ask for up-votes or witness votes or not. Witnesses are also the only users that can have a stable cryptocurrency income through the block creation rewards, so it is interesting to understand which are the most lucrative methods they use to acquire cryptocurrency. We also plan to develop advanced studies concerning the study of other types of users, such as bot users, possibly providing bot detection tools. Lastly, we look forward to investigating the usage of blockchain as support for the development of OSNs. For instance, we plan to implement a privacy policy system or a collective bot detection system.

Acknowledgements This work is supported by the HELIOS H2020 project. HELIOS has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825585.

Funding Open access funding provided by Università di Pisa within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not

included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Buchegger S, Schiöberg D, Vu L, Datta A (2009) Implementing a P2P social network - early experiences and insights from PeerSoN. In: Second ACM workshop on social network systems
2. Cadwalladr C, Graham-Harrison E (2018) Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach. *The Guardian* 17:22
3. Cecilia JM, Cano JC, Hernández-Orallo E, Calafate CT, Manzoni P (2020) Mobile crowdsensing approaches to address the covid-19 pandemic in Spain. *IET Smart Cities* 2(2):58–63
4. Chang YT, Yu H, Lu HP (2015) Persuasive messages, popularity cohesion, and message diffusion in social media marketing. *J Bus Res* 68(4):777–782
5. Chohan UW (2018) The concept and criticisms of steemit. Available at SSRN 3129410
6. Cuttillo LA, Molva R, Strufe T (2009) Safebook: a privacy-preserving online social network leveraging on real-life trust. *Comm Mag* 47(12):94–101
7. Datta A, Buchegger S, Vu LH, Strufe T, Rządca K (2010) Decentralized online social networks. Springer US, Boston, pp 349–378
8. De Salve A, Guidi B, Michienzi A (2018) Exploiting community detection to recommend privacy policies in decentralized online social networks. In: European conference on parallel processing. Springer, pp 573–584
9. De Salve A, Guidi B, Michienzi A (2021) Exploiting homophily to characterize communities in online social networks. *Concurrency and Computation: Practice and Experience* 33(8):e5929
10. Dennis R, Owenson G, Aziz B (2016) A temporal blockchain: a formal analysis. In: 2016 international conference on collaboration technologies and systems (CTS). IEEE, pp 430–437
11. Freni P, Ferro E, Ceci G (2020) Fixing social media with the blockchain. In: Proceedings of the 6th EAI international conference on smart objects and technologies for social good, pp 175–180
12. Guidi B (2020) When blockchain meets online social networks. *Pervasive and Mobile Computing* 62:101,131
13. Guidi B, Amft T, De Salve A, Graffi K, Ricci L (2016) Didusonet: a p2p architecture for distributed dunbar-based social networks. *Peer-to-Peer Networking and Applications* 9(6):1177–1194
14. Guidi B, Conti M, Passarella A, Ricci L (2018) Managing social contents in decentralized online social networks: a survey. *Online Social Networks and Media* 7:12–29
15. Guidi B, Michienzi A, Ricci L (2020) A graph-based socio-economic analysis of steemit. *IEEE Transactions on Computational Social Systems*, pp 1–12. <https://doi.org/10.1109/TCSS.2020.3042745>
16. Guidi B, Michienzi A, Ricci L (2020) Steem blockchain: mining the inner structure of the graph. *IEEE Access* 8:210,251–210,266
17. Jiang L, Zhang X (2019) Bcosn: a blockchain-based decentralized online social network. *IEEE Transactions on Computational Social Systems* 6(6):1454–1466
18. Kapanova K, Guidi B, Michienzi A, Koidl K (2020) Evaluating posts on the steemit blockchain: analysis on topics based on textual cues. In: Proceedings of the 6th EAI international

- conference on smart objects and technologies for social good, pp 163–168
19. Khuntia J, Sun H, Yim D (2016) Sharing news through social networks. *International Journal on Media Management* 18(1): 59–74
 20. Kümpel AS, Karnowski V, Keyling T (2015) News sharing in social media: a review of current research on news sharing users, content, and networks. *Social Media+ Society* 1(2):205630511 5610,141
 21. Larimer D, Scott N (2016) Steem whitepaper. <https://steem.com/wp-content/uploads/2018/10/steem-whitepaper.pdf>
 22. Li C, Palanisamy B (2019) Incentivized blockchain-based social media platforms: a case study of steemit. In: *Proceedings of the 10th ACM conference on web science*, pp 145–154
 23. Manikonda L, Meduri VV, Kambhampati S (2016) Tweeting the mind and instagramming the heart: exploring differentiated content sharing on social media. In: *Tenth international AAAI conference on web and social media*
 24. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. <https://bitcoin.org/bitcoin.pdf>
 25. Niranjnamurthy M, Nithya B, Jagannatha S (2019) Analysis of blockchain technology: pros, cons and swot. *Cluster Computing* 22(6):14,743–14,757
 26. Rathore S, Loia V, Park JH (2018) Spamspotter: an efficient spammer detection framework based on intelligent decision support system on facebook. *Appl Soft Comput* 67:920–932
 27. Singh S, Singh N (2016) Blockchain: future of financial and cyber security. In: *2016 2nd international conference on contemporary computing and informatics (IC3I)*. IEEE, pp 463–467
 28. Terroso-Saenz F, Muñoz A, Cecilia JM (2019) Quadriven: a framework for qualitative taxi demand prediction based on time-variant online social network data analysis. *Sensors* 19(22):4882
 29. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. In: *2017 IEEE international congress on big data, BigData Congress*. IEEE, pp 557–564

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.