IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 28, 2020, accepted June 5, 2020, date of publication June 18, 2020, date of current version June 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003370

Charting the Landscape of Online Cryptocurrency Manipulation

LEONARDO NIZZOLI[1,2], SERENA TARDELLI[1,2], MARCO AVVENUTI[2], (Member, IEEE), STEFANO CRESCI[1], (Member, IEEE), MAURIZIO TESCONI[1], AND EMILIO FERRARA[3], (Senior Member, IEEE)

[1]Institute of Informatics and Telematics, National Research Council (IIT-CNR), 56124 Pisa, Italy
[2]Department of Information Engineering, University of Pisa, 56122 Pisa, Italy
[3]Information Sciences Institute, University of Southern California, Marina Del Rey, CA 90292, USA

Corresponding author: Serena Tardelli (serena.tardelli@iit.cnr.it)

This work was supported in part by the Air Force Office of Scientific Research (AFOSR) under Award FA9550-17-1-0327, and in part by the Defense Advanced Research Projects Agency (DARPA) under Contract W911NF-17-C-0094 and Grant D16AP00115.

ABSTRACT Cryptocurrencies represent one of the most attractive markets for financial speculation. As a consequence, they have attracted unprecedented attention on social media. Besides genuine discussions and legitimate investment initiatives, several deceptive activities have flourished. In this work, we chart the online cryptocurrency landscape across multiple platforms. To reach our goal, we collected a large dataset, composed of more than 50M messages published by almost 7M users on Twitter, Telegram and Discord, over three months. We performed bot detection on Twitter accounts sharing invite links to Telegram and Discord channels, and we discovered that more than 56% of them were bots or suspended accounts. Then, we applied topic modeling techniques to Telegram and Discord messages, unveiling two different deception schemes – ''pump-and-dump'' and ''Ponzi'' – and identifying the channels involved in these frauds. Whereas on Discord we found a negligible level of deception, on Telegram we retrieved 296 channels involved in pump-and-dump and 432 involved in Ponzi schemes, accounting for a striking 20% of the total. Moreover, we observed that 93% of the invite links shared by Twitter bots point to Telegram pump-and-dump channels, shedding light on a little-known social bot activity. Charting the landscape of online cryptocurrency manipulation can inform actionable policies to fight such abuse.

INDEX TERMS Cryptocurrency manipulation, social bots, Twitter, Telegram, Discord.

I. INTRODUCTION

Cryptocurrencies have attracted considerable public attention over the years, carving out a significant social presence in online environments. The vast conversation space offered by social media is the perfect venue to promote the cryptocurrency world, supporting its rise in popularity [1], [2]. As more and more online users adopt cryptocurrencies as a practical means of investment and payment, scientists investigate their interaction with social media [3] to predict prices fluctuation [4], to monitor the users' trust in cryptocurrencies [2], and to pave the way to new disruptive applications [1].

However, social media have already proven to be a suitable habitat for deception in many domains, such as politics and healthcare, and the cryptocurrency market is no exception [5]. In other words, social ecosystems enable manipulation to run wild in a domain that already thrives on anonymity, decentralization, high volatility, and self-regulation [6], [7]. A growing strand of research has been focusing on cryptocurrency manipulation. Indeed, many studies examined and characterized specific online frauds to understand their influence on the online markets. Some works focused on cryptocurrency thefts [8]. Others dissected the mechanisms behind pump-and-dump schemes, in which willing participants collectively aim to artificially inflate a currency price through coordinated, simultaneous buying (''pump''). Once outside unaware investors notice the surge in price and start investing in the asset, the participants sell it to them, thus making a profit and causing a price collapse (''dump''). Generally, there are orchestrators behind the curtains, who profit even at the expense of the witting participants themselves, let alone the other unaware investors [9]. Other studies focused on analyzing Ponzi schemes, financial scams that rely on acquiring investors by promising high returns in exchange of a minimum amount of currency. Those funds are used to generate profits for old investors and organizers. When the rate

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

of new investors is not large enough to sustain the process, the chain breaks, and last comers lose their investment [10].

However, such works addressed each fraud in a targeted way. There's still little understanding of the interactions between different types of scams and wherein the online landscape they fall. This is crucial to spot the key players involved in scams and to support actionable policies to fight such abuses.

This work aims to map and assess the extent of cryptocurrency manipulations within and across the online ecosystems of Twitter, Telegram, and Discord. Such wide, horizontal exploration is of the utmost importance to have a broad overview of the crypto environment and the actors involved. In fact, many successful scams depend on attracting a large mass of users. Therefore, it is necessary to address the problem by evaluating a large user base. Indeed, Twitter provides the perfect showcase to attract potential investors and to deceive users, as demonstrated in many scenarios before [11]. Likewise, Telegram and Discord are the ideal habitats for crypto scams to proliferate, as they offer anonymity and low levels of moderation. Moreover, they feature channels as a way to broadcast public messages to a large audience and to invite investors to join.

This is why we pose our focus on the diffusion of invite links – that is, unique URLs allowing users to join channels. In fact, our controlling idea for this study is based on the intuition that fraudsters can exploit invite links to scam channels as an effective way of recruiting participants to the scam. If proven right, detecting and monitoring significant hubs spreading links can help mitigate the effects of scams in the real world, defend the markets from manipulation, and protect people's investments. In addition, diffusion patterns of invite links already provided valuable information for detecting homophily and common interests in online communities [12], which can be useful in order to spot malicious actors orchestrating scam campaigns.

A. CONTRIBUTIONS OF THIS WORK

To shed light on cryptocurrency manipulation, we follow invite links in a snowball strategy and collect a rich dataset of 16M tweets, 10M Discord, and 23M Telegram messages discussing cryptocurrencies. As we add layers of knowledge about the authenticity of Twitter accounts and the content of Telegram and Discord channels, we uncover and trace the path and evolution of manipulation attempts from platform to platform.

By cross-checking these different types of manipulation, we show that deceptive channels are the ones receiving the most part of the invite links, confirming the founding intuition of our controlling idea. We report that the strategy of following invite links is ideal for tracking online crypto scams. We highlight the vast presence of Twitter bots and their key part in promoting pump-and-dump channels, thus contributing in mapping the role of automated accounts in spreading misinformation through social media. Based on our findings, researchers and stakeholders can enforce novel actionable policies to mitigate online cryptocurrency scams, by tracking and cutting off the main hubs of manipulation, thus severely impairing the efficacy of cryptocurrency frauds.

Our main contributions are summarized as follows:

- We collect and share a large dataset for studying online cryptocurrency manipulations, comprising more than 50M messages and describing the online cryptocurrency ecosystem across three major platforms: Twitter, Telegram and Discord.
- We uncover the pivotal role of Twitter bots in broadcasting invite links to deceptive Telegram and Discord channels, exposing a little-known social bot activity.
- Instead of focusing on specific frauds, we let manipulation patterns naturally emerge from data, highlighting the existence of two different manipulations – namely, *pump-and-dump* and *Ponzi* schemes.
- Our results describe Discord as a reasonably healthy online cryptocurrency ecosystem. In contrast, more than 56% of crypto-related Telegram channels are involved in manipulations. Moreover, these deceptive activities are massively broadcast with the help of Twitter bots.

Reproducibility. To ensure reproducibility, we released an anonymized, privacy-preserving version of the dataset.[1]

B. ROADMAP

The remainder of this paper is organized as follows. In Section II, we survey the recent literature on online manipulations, focusing on those involving cryptocurrencies. In Section III, we depict and motivate our data collection strategy, and we provide the first descriptive statistics about the resulting dataset. Section IV provides the invite link network as an effective way to represent the diffusion of invite links, and to highlight its underlying patterns. In Section V, we enrich the network with a layer of information about the Twitter account genuineness. In Section VI, we add network layers related to the social media content traits. In Section VII, we uncover the most relevant cryptocurrency manipulation patterns, and we characterize their underlying promotion strategies. Finally, Section VIII summarizes the main results, contextualize them in the general research effort against online manipulations, and presents the limitations and future work.

II. RELATED WORKS

In this Section, we survey the recent literature about manipulations perpetrated in the online ecosystems. We focus in particular on cryptocurrency manipulations, since our contribution also falls in this category.

A. CRYPTOCURRENCY MANIPULATIONS

This paper fits into the literature of cryptocurrency-related social media content analysis. Several works addressed the cryptocurrency topic with a focus on manipulation patterns. In [22], authors provided a first look of cryptocurrency pump-and-dump schemes and found that such frauds target specific

[1] http://doi.org/10.5281/zenodo.3895021

TABLE 1. Comparative analysis with related works on cryptocurrency-related social media content analysis.

Study	Platform	Analysis		Dataset		Focus	
		predictive	*descriptive*	*data-informed*	*data-driven*	*coin*	*scheme*
Xu et al. [9]	Telegram	✓		✓		comprehensive	pnd
Li et al. [13]	Telegram		✓	✓		comprehensive	pnd
Mirtaheri et al. [14]	Telegram	✓		✓		comprehensive	pnd
Hamrick et al. [15]	Telegram, Discord		✓	✓		comprehensive	pnd
Glenski et al. [16]	Reddit		✓	✓		Bitcoin, Ethereum, Monero	–
Kim et al. [17]	Forum	✓		✓		Bitcoin, Ethereum, Ripple	–
Kim et al. [18]	Forum	✓		✓		Bitcoin	–
Vasek et al. [19]	Forum	✓		✓		Bitcoin	Ponzi
Bartoletti et al. [20]	Reddit, Forum	✓		✓		Bitcoin	Ponzi
Linton et al. [21]	Forum		✓	✓		comprehensive	–
Our contribution	Twitter, Telegram, Discord		✓		✓	comprehensive	comprehensive (pnd, Ponzi emerged)

cryptocurrency coins and exchanges. However, their work was preliminary and did not yet analyze the coordination of pump-and-dumps in online chat groups, nor the means by which misinformation about specific coins is spread on, like on social media.

Focusing on studies that also take into account social media, authors in [9] provided a detailed description of pump-and-dump schemes that take place on known deceptive Telegram channels and developed a model to predict the likelihood of a coin being the target for manipulation. While they offered important key insights into a specific fraudulent scheme and platform, our goal is to offer a more general vision into the bounds of cryptocurrency manipulation activities that take place across multiple social media platforms. Similarly, authors in [13] focused exclusively on pump-and-dump schemes that occur on known deceptive Telegram channels, demonstrating the harmful effects of pump-and-dump events on the liquidity and price of cryptocurrencies.

Along this line, authors in [14] collected pump-and-dump Telegram channels starting from known seeds, and continuing in a snowball fashion. They leveraged such data to predict the success of pump operations in terms of meeting the anticipated price targets. They also collected tweets in close proximity to the attacks and discovered a prevalence of bots involved in the discussion. This study is more akin to our work, as it looks for deception on multiple social platforms. However, they still focused on a specific cryptocurrency manipulation domain. Instead, we are interested in a more general overview of the cryptocurrency ecosystem, in order to understand the extent of deception in the network. Indeed, other studies observed that pump-and-dump phenomena are also widespread on Discord, as well as on Telegram [15], emphasizing the importance of taking into consideration more platforms.

In [16], authors measured the spread of cryptocurrency discussion on Reddit. They focused on conversations around *Bitcoin*, *Ethereum* and *Monero* cryptocurrencies, noticing that users discussing coins used for shady activities engage in a deeper and longer debate. However, more work needs to be done to address cryptocurrency-related discussion across a variety of coin types or across multiple social media platforms. In [17], [18], authors focused on known pump-and-dump events involving specific cryptocurrencies on dedicated forums and developed a model to predict fluctuation in cryptocurrency prices.

Pump-and-dump is not the only financial fraud under scrutiny. In [19], authors investigated online Ponzi schemes advertised on threads of the Bitcointalk forum. They discovered that shill interactions lengthen the life of a scam, whereas constant daily interactions between scammers and victims shorten it. However, more work is needed to measure the spread of scams between social media with different interaction paradigms. In [20], authors looked for Ponzi schemes involving the *Bitcoin* cryptocurrency on both Reddit and the Bitcointalk forum, proposing a model to predict Ponzi schemes involving cryptocurrencies.

Authors in [21] focused on discussion forums to identify and characterize communities by applying topic modeling in order to model trends in the community and to see how real-life events affect the topics discussed and *vice versa*. Interestingly, pump-and-dump activities spontaneously emerged. Similarly, we let manipulation emerge in this way, although considering multiple platforms.

In Table 1, we summarize the related works according to four dimensions. First, we consider the platform on which each work looks for deceptive activities. Second, we characterize the type of analysis of the paper as predictive or descriptive. Third, we specify the dataset collection strategy adopted as data-informed – when the starting seeds of deceptive messages collected is known for being deceptive, or data-driven – when generic messages are collected and then deception is found. Finally, we define the focus of each work under both the coin and the deceptive scheme perspectives. As shown, our work opted for a comprehensive, data-driven strategy, from which deceptive schemes naturally emerge from the data.

B. OTHER ONLINE MANIPULATIONS

The existence of manipulative, deceptive, synthetic content in online discussions has already been witnessed in a wide variety of societal topics. For instance, it has been demonstrated that bots are exploited to promote online financial

content [23], as well as health content [24], [25]. Other studies showed that bots tampered with US [26], [27], Japanese [28], South Korean [29], French [30], Italian [31], and German [32] political elections.

In other recent work [11], it is reported the emergence of new waves of social bots, capable of mimicking human behavior in social media better than ever before. As social bots evolve, online content manipulation goes undetected even by platform administrators [33], with consequent profound impact on content popularity and activity in social media [34], [35]. Scholars and platform administrators reacted by proposing more advanced detection techniques based on the analysis of both individual [36] and collective [33], [37], [38] behaviors. The current research trend with regards to online manipulation is shifting from a focus on individual malicious accounts (e.g., bots, trolls) to a broader and more sophisticated model that embraces the interplay between both automated and human-driven behaviors [39]–[41]. However, to the best of our knowledge, the latter model is yet to be exploited and operationalized.

III. DATASET

In this Section, we first introduce preliminary considerations motivating the design of our snowball data collection strategy. Then, we focus on the procedure details, and we provide some descriptive statistics about the obtained dataset.

A. PRELIMINARIES

Previous works about cryptocurrency manipulation [9], [19] focused on a specific scheme (e.g., pump-and-dump or Ponzi), aiming to outline its anatomy, assess its efficacy, or predict its occurrence. Accordingly, they relied on datasets specifically designed to include only data pertinent to the cryptocurrency manipulation scheme under exam and on a specific platform. Conversely, here we are interested in performing a wide, horizontal exploration of the online cryptocurrency ecosystem across multiple platforms. The goal is to find manipulation patterns by looking at intra- and cross-platform community interactions. As mentioned before, many works [9], [19] collected cryptocurrency manipulation data by using a snowball approach strategy, starting from a known seed of deceptive channels. Conversely, here we opted for a crawling snowball approach starting from generic invite links occurring in cryptocurrency-related tweets. In this way, we avoid any bias towards legitimate or deceptive communities. We also have the chance to (i) observe deceptive schemes naturally emerging from the data, (ii) assess their spread within the online multi-platform cryptocurrency ecosystem, (iii) identify legitimate and deceptive agents (e.g., accounts, channels), and (iv) study the interplay between them. In order to obtain a dataset with the desired features, we designed and implemented a crawling strategy based on a snowball approach. We focused on the Twitter microblogging platform, and on the two instant messaging platforms Telegram and Discord. We decided to focus on Twitter because literature already provided evidence of the presence of financial and cryptocurrency-related content on such platform [23], [42]. Regarding Telegram and Discord, literature proved that their encryption, programmability and anonymity encourage the presence of cryptocurrency communities [9], [15]. Telegram features two types of group chats: (i) *groups* – where all members are allowed to share content by default, and (ii) *channels* – where usually only administrators broadcast content to their audience. They can be joined by means of specific *invite links* (URLs), which can contain the required password in case the group or channel is private. Discord features *servers* (also referred to as *guilds*), in which admins can create several channels – each one usually devoted to a specific topic – and handle the writing privileges. Authorized users can generate invite links (URLs) for the server, which are specific for the user who created them. Hereafter, we use the generic term ''channel'' for Telegram groups and channels as well as for Discord servers, and the term ''invite link'' for every type of URL allowing users to join a channel.

TABLE 2. Counts of distinct channels, users and messages for each considered platform.

	channels	users	messages
Discord	1,755	211,409	10,331,720
Telegram	3,813	920,925	23,812,537
Twitter	–	5,745,944	16,840,312
total	5,568	6,878,278	50,984,569

B. DATA COLLECTION

Firstly, we leveraged the Twitter's Streaming API[2] to collect all tweets mentioning at least one of the 3,822 cryptocurrency cashtags[3] provided by the *CryptoCompare*[4] public API. This data collection covered a three months-long time window spanning from March to May 2019, resulting in the acquisition of more than 16M tweets. Then, we retrieved all the invite links contained in these tweets pointing to Telegram or Discord channels, and we used them as seeds for an iterative snowball crawling strategy. In particular, this first set of channels, pointed by those invite links, represents the *hop 0* of our crawl. By leveraging Telegram[5] and Discord[6] APIs, we collected the message histories of hop 0 channels. Then, we parsed such messages looking for more invite links; we retrieved the message histories of the related (hop 1) channels, and we continued iterating this data collection pipeline. At hop 3, we retained only invite links pointing to channels already found at hops 0-2, and we concluded our crawling. In Table 2, we provide some aggregates of the obtained dataset. As shown, our dataset includes more

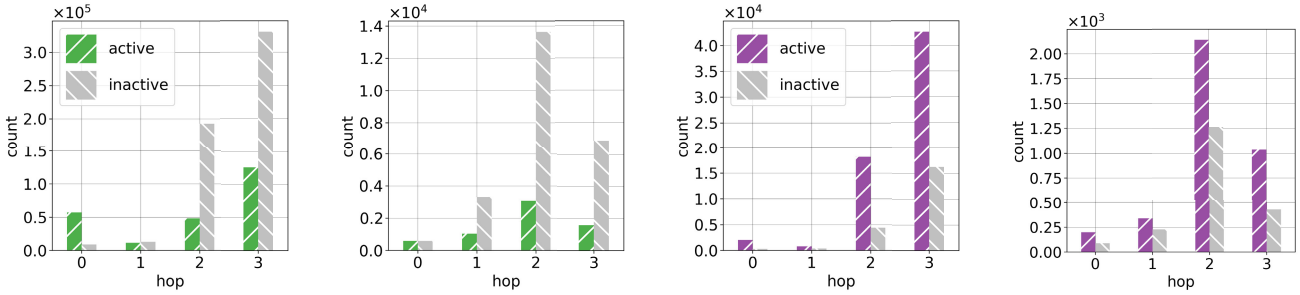
[2]https://developer.twitter.com/en/docs

[3]The cashtag of a cryptocurrency is composed of a dollar sign followed by the ticker symbol of the cryptocurrency (e.g., $BTC for Bitcoin). Similarly to hashtags, they can be used to efficiently tag and filter tweets.

[4]https://min-api.cryptocompare.com/

[5]https://core.telegram.org/

[6]https://discord.com/developers/docs/reference



(a) Total invite links (Telegram). (b) Distinct invite links (Telegram). (c) Total invite links (Discord). (d) Distinct invite links (Discord).

FIGURE 1. Counts of active and inactive invite links to Telegram (1a, 1b) and Discord (1c, 1d) channels, retrieved at each hop of our snowball crawling strategy. Telegram attracts much more invite links than Discord (79.2%). The large number of inactive invite links may reflect the practice of publishing "expiring" invite links.

than 50M messages, published by almost 7M distinct users across the three platforms. In particular, we highlight the unprecedentedly large number of Telegram (3,813) and Discord (1,755) channels, that guarantees a sound coverage of the cryptocurrency ecosystem on such platforms. Focusing on the two instant messaging platforms, we notice that 68.5% of the retrieved channels, 81.3% of distinct users and 69.7% of messages belong to Telegram. In Figure 1, we depict the count of active and inactive invite links, retrieved at each hop of our snowball crawling strategy. Considering the combined amount of invite links for both platforms, Telegram accounts for 79.2% of active and 96.2% of inactive links. We highlight that our data collection strategy is impartial with respect to the two instant messaging platforms. Hence, our sample suggests that Telegram is much more used than Discord within the online cryptocurrency ecosystem, reflecting its larger share of users in general. Finally, figures 1a, 1b show that at hop > 1 inactive Telegram invite links largely exceed active ones, as opposed to Discord (figures 1c, 1d). The large number of inactive invites may reflect the practice of publishing "expiring" links, to promote more elitist, limited access channels.

The data collection process was carried out by using Python 3.7 as the programming language to collect public data from Telegram and Discord through official APIs. For collecting public tweets, we leveraged the Twitter Monitor tool described in [43], a framework that implements and handles requests to Twitter APIs. Finally, we used the distributed search engine Elasticsearch to index and store the dataset collected. The following analyses and code implementations were performed by using Python 3.7 and its libraries.

As an additional contribution of our work, we publish an anonymized, privacy-preserving version of this dataset,[7] for reproducibility purposes and to foster further research on this important topic.

IV. THE INVITE LINK NETWORK

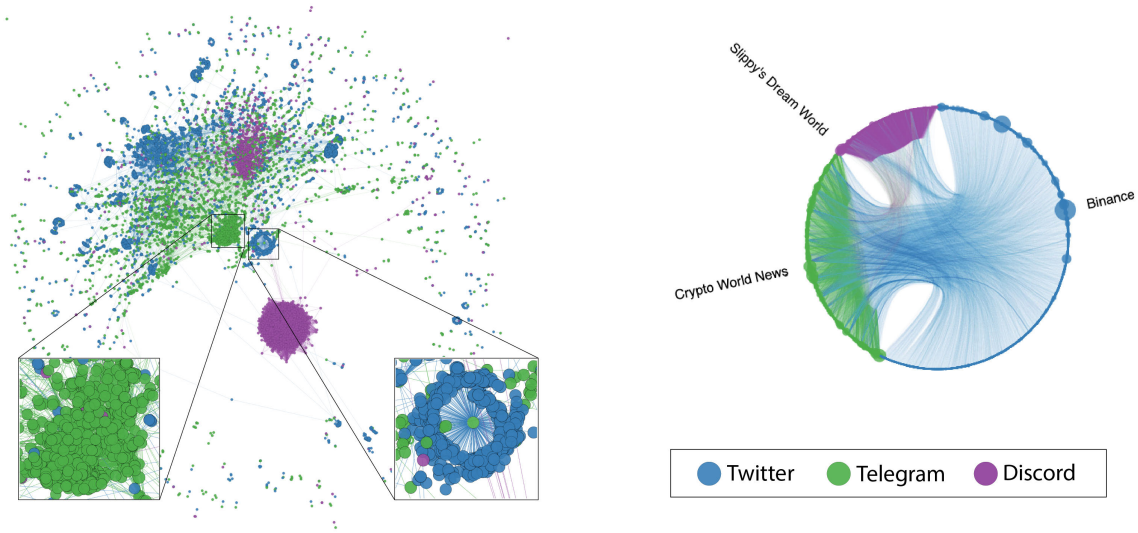
The diffusion of invite links plays a major role in the growth of online platforms and communities. As an effective means for recruiting people to channels, we expect the invite link diffusion to be particularly relevant in cryptocurrency manipulation schemes since their effectiveness strongly depends on the number of participants involved. Moreover, there is a strong interplay between the structural properties underlying the diffusion of invites and the characteristic features of the source and target agents involved in those processes [12]. In particular, the exchange of invites can be an excellent proxy for homophily or common goals. Hence, we hypothesize deceptive agents to give a major contribution to the diffusion of invite links. Consequently, characterizing the invite link diffusion can allow to spot malicious agents, track their activities, and reveal possible interplay patterns.

We study the invite link diffusion process by building the invite link network, which is shown in Figure 2a. It is a directed, weighted network composed of 13,009 nodes and 62,278 edges. Nodes represent agents sharing or receiving at least one invite link. In detail, 7,441 (57.2%) nodes are Twitter accounts, 3,813 (29.3%) are Telegram channels and 1,755 (13.5%) are Discord channels. Edges are directed from a source node – representing an agent who broadcasts an invite link, to a target node – representing the channel pointed by the invite link. Their weights account for the number of existing invite links between the two. It is worth noticing that Twitter nodes can only have outgoing edges since Twitter accounts cannot receive invite links.

Figure 2a shows a node-link diagram of the network, realized with the *ForceAtlas* algorithm. Node size is proportional to the number of channel members or the number of followers of a Twitter account – that is, to the size of the potential audience of the agent. Nodes are colored according to the corresponding platform. Edge thickness is proportional to the weight, and their color is the same as that of the source node. ForceAtlas determines the layout of nodes so that nodes connected by strong links appear close to each other in the diagram. Figure 2a shows the presence of a giant component, including 91% of nodes and having a diameter of 19. Since the giant component includes most of the nodes and links, we focus the rest of our analyses on it.

The giant component includes a strongly clustered community of Discord channels, weakly connected to the rest of the nodes and represented as an isolated (violet) "hairball",

[7] http://doi.org/10.5281/zenodo.3895021



(a) ForceAtlas node-link diagram of the invite link network.

(b) Radial node-link diagram of the invite link network.

FIGURE 2. **Invite link network where nodes represent Twitter accounts (blue), Telegram (green) and Discord channels (violet) sharing or receiving at least one invite link. Edges are colored by their source node color. Figure 2a shows the presence of peculiar network structures, such as a dense cluster of Telegram channels (bottom-left inset) and star structures (bottom-right inset). Figure 2b highlights the role of Twitter as a bridge, while Telegram and Discord channels exchange invite links within their own platform.**

located near the center of Figure 2a. Within the rest of the giant component, there is still a clear separation between Discord and Telegram nodes. A very dense cluster of Telegram channels (green-colored) is magnified in the bottom-left corner of the plot. This kind of structure reflects channels engaged in mutual promotion within the same platform. Twitter nodes (blue-colored) appear as frequently arranged in a ring surrounding a single channel – usually a Telegram one – thus forming a star structure. An example of this feature is magnified in the bottom-right corner of the plot. We counted 135 of these structures having a size of at least 10 accounts, 107 of which (79.2%) are centered around a Telegram channel. Those structures reveal multiple Twitter accounts promoting a single channel (usually belonging to Telegram). Interestingly, these preliminary observations highlight the presence of peculiar network structures (e.g., dense clusters, stars), likely representative of interesting real-world phenomena.

In order to provide a better visualization of the flow of invite links within and across the different platforms, in Figure 2b we represent the network according to a radial node-link diagram, applying a bundle between edges originating from the same platform. Also in this case, node size is proportional to the number of members of a channel (in case of Telegram and Discord) or the number of followers of an account (in case of Twitter). Out of curiosity, we show the names of the biggest public channels in terms of members and of the Twitter account with the largest number of followers. Notably, Telegram and Discord nodes exchange invite links almost exclusively with nodes belonging to the same platform. Focusing on the edge counts, Discord-Discord edges (40,040) account for 64.3% of the total, followed by Telegram-Telegram (11,098, 17.8%) and Twitter-Telegram (8,371, 13.4%) edges. Twitter-Discord edges (1,686) are 2.7% of the total, while Discord-Telegram (527) and Telegram-Discord (556) edges are less than 1%. According to this result, Discord emerges as a highly interconnected environment, where each channel exchanges invite links with many others. In particular, the skewness in the edge count distribution is mainly due to the aforementioned community of Discord channels, that is isolated from the rest of the network, but so strongly clustered to approximate a clique. When also taking into account edge weights – that is when accounting for the actual number of invite links – the results are very different. Telegram-Telegram (186,903) links are 60.1% of the total, whereas Discord-Discord (60,972) ones end up in second place (19.6%). Discord-Telegram and Telegram-Discord links are still very few (less than 1%). Twitter exhibits a strong relationship with Telegram (57,377, 18.5%), but the amount of invites toward Discord is now very little (less than 1%). As opposite to Discord, Telegram channels are connected with fewer other channels but with much stronger links. In particular, the main contribution to the distribution skewness is due to the aforementioned cluster of Telegram channels, actively engaged in a mutual promotion.

A. NETWORK ANALYSIS

To better understand and measure the properties of this network, we applied standard network analysis approaches. The weighted degree distribution of the network is well approximated by a power-law with exponent ~ -2.1, in the typical range of scale-free social networks [44]. Interestingly, the weighted degree distribution exhibits an anomalous peak around degrees $\gtrsim 10^2$. This peak is related to the

hairball-shaped cluster of Discord channels, further confirming the high amount of connections between its nodes.

In order to measure randomic [45], small world [46] and preferential attachment [47] properties of the network, we built Erdős-Rényi, Watts-Strogatz and Barabási-Albert synthetic networks. For all the synthetic networks, we chose the same number of nodes and edges/degrees of the giant component of the invite link network. In the comparison, we disregarded direction and weight of the giant component edges. We measured an average clustering coefficient around 0.13, much larger with respect to the Erdős-Rényi one (0.0007) but smaller than the Watts-Strogatz (0.55). The average shortest path length resulted equal to 6.05, longer than Erdős-Rényi (4.45) but shorter than Watts-Strogatz (8.89). From these results, we conclude that the invite link network exhibits small world properties. As a consequence, the presence of the already mentioned clustered communities and cliques is expected. Notably, this kind of networks exhibits a certain robustness to random perturbations, meaning that the demise of random nodes should not dramatically affect the overall network properties. At the same time, in this kind of network most average shortest paths pass through nodes with high degree (hubs). As a consequence, removing hubs can severely compromise the network properties. Finally, we obtained a Pearson degree correlation of -0.15, significantly lower than that of the Barabási-Albert model (-0.02). The overall slightly disassortative behaviour of our invite link network confirms the relevance of the previously mentioned star structures of Twitter accounts. In fact, when disregarding edge weights, those structures correspond to many nodes with very low degree (typically equal to one) connected to a node with very high degree, determining the overall disassortative behaviour of the network.

Until now, we built and analyzed the network by only considering its structural properties. The only attribute distinguishing nodes and edges was the corresponding social media platform. In the next sections, we deepen our analysis of the invite network by enriching the nodes with semantic features, with the goal of providing explanations for its peculiar structures.

V. INVITE LINK NETWORK ENRICHMENT: ASSESSING THE NATURE OF TWITTER ACCOUNTS

Tracking cryptocurrency manipulation schemes within the multiform ecosystem enclosed in our dataset requires to overlap layers of knowledge over the map sketched until now. In particular, we want to characterize the Twitter nodes of our network according to their genuineness, in order to drive further analyses.

Besides human users, the Twitter platform is populated by bots. These are accounts controlled by computer algorithms, able to automatically produce content and interact with other accounts, emulating human behaviour. Some bots perform neutral or even useful tasks, but some others instead attempt to manipulate and deceive genuine social media users, pursuing malevolent purposes [11]. Hence, in this section we aim at measuring the contribution of Twitter social bots to the diffusion of invite links, in order to evaluate their role in cryptocurrency manipulation schemes.

We performed bot detection on the 7,441 Twitter accounts broadcasting invite links. We used Botometer [36], a well-known Twitter bot detection service, publicly-available via REST API.[8] Botometer is a supervised machine learning classification model, combining more than a thousand features extracted from profile metadata, friends, social network structure, temporal activity patterns, language and sentiment. The service takes an account ID as input and returns two scores: one is called ''universal'', because it disregards language and sentiment features; the other is specific for English accounts. Since our dataset includes several non-English accounts, we used the universal score, labeling as bots those accounts having a score ≥ 0.5.

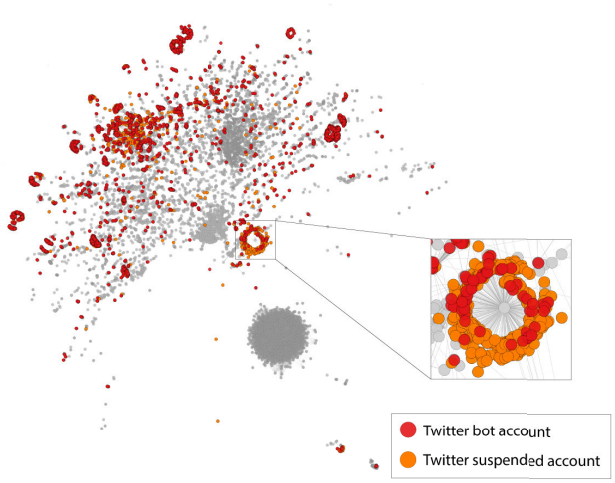


FIGURE 3. Invite link network highlighting deceptive Twitter accounts. A large portion of Twitter accounts has a deceptive nature (56.3%). The typical star structures frequently correspond to botnets promoting a single channel. We found 69 botnets with a size of at least 10 elements.

Botometer classified 2,710 accounts as bots, resulting in a remarkable fraction of 36.4% of the total. In addition, other 1,483 (19.9%) accounts were already suspended by Twitter, again testifying malicious behaviors. By grouping together bots and suspended accounts, we discover that more than a half (56.3%) of the Twitter accounts involved in invite link broadcasting have a deceptive nature. This remarkable fraction of deceptive accounts largely exceeds previous estimations of overall Twitter bot population, ranging from 9% to 15% [48]. Instead, it approaches the fraction of 71% of bots, recently observed when considering most active accounts broadcasting stock-related messages [23]. To this regard, our finding reinforces the knowledge that social bots proliferate in those scenarios involving strong economical incentives. In the remainder, we address both bots and suspended

[8] https://rapidapi.com/OSoMe/api/botometer

accounts as ''deceptive'' or bots, whereas we define groups of such accounts as ''botnets''.

In Figure 3, we highlight Twitter deceptive accounts in the invite link network by coloring the corresponding nodes. Clusters of deceptive accounts clearly emerge, frequently assuming the star shape mentioned in the previous Section. At that time, we counted 135 of them in the network, having a size of at least 10 accounts. Now, we find 69 botnets with the same minimum size, 56 of which (81.1%) are promoting a single Telegram channel. Hence, those star structures can be confidently interpreted as Twitter botnets.

VI. INVITE LINK NETWORK ENRICHMENT: CHARACTERIZING ONLINE CRYPTOCURRENCY DISCUSSIONS

In the previous Section, we characterized Twitter nodes according to their genuine or deceptive nature. Now, we focus on the content of the messages posted within Telegram and Discord channels, and by Twitter accounts. In detail, we highlight the main topics of discussion within each platform by applying topic modeling. Then, we refine the granularity of our analysis by labeling each channel/account according to its dominant topic. Notably, a similar approach has already been applied in [21] to online forums, with interesting results.

To perform topic modeling, we adopted a recent, state-of-the-art algorithm known as Anchored Correlation Explanation (CorEx) [49]. As opposed to generative models – such as Latent Dirichlet Allocation (LDA) – CorEx learns latent topics over a collection of documents without assuming any particular data generating model. Instead, it leverages the dependencies of words in documents through latent topics, by maximizing the total correlation between groups of words and the respective topic. This approach ensures greater flexibility, enabling hierarchical and semi-supervised variants [49]. In particular, it features word anchoring, a semi-supervised technique improving topic separability with minimum human intervention. In fact, by providing some sets of anchor words relevant for specific topics, it is possible to push the model to better identify and separate them.

This emerging topic modelling technique has already proven useful for extracting relevant topics in social media data. In particular, authors in [49] applied CorEx on tweets related to the shooting of a Black teenager in 2014. They managed to disambiguate the tweets in favour of the protest and those against it. Such work demonstrated the capability of this novel technique to extract topics able to separate multiple facets within each analyzed discussion. Along this line, authors in [50] applied CorEx to tweets concerning eating disorders. They extracted novel topics and provided insights into factors that may foster the perpetuation of such behaviors, thus contributing to better understand – and possibly deal – with this serious issue.

A. UNSUPERVISED TOPIC EXTRACTION

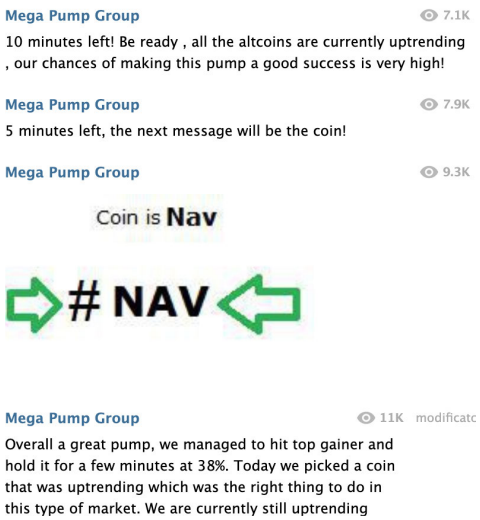
We first applied CorEx without anchoring (i.e., in a completely unsupervised fashion), in order to discover topics spontaneously emerging from our data. Input documents consist in the textual content of each Discord/Telegram channel, and in a concatenation of all the available tweets for each Twitter account. To increase the accuracy of our results, we learned separate models for each platform, in order to account for possible differences in topics and forms of speech. In addition, we also filtered channels/accounts based on the prevalent language of their messages. In particular, we used the Python library *polyglot* [51] to estimate the prevalent language, and we neglected non-English instances. As a result, we retained 64.6% of all Telegram channels, 89.5% of all Discord channels, and 88.8% of Twitter accounts. In this way, we obtained much more accurate results in terms of detected topics, at the cost of discarding just a minority of all the available data. Following the recommendation in [49], we experimented with several configurations, by increasing the number of expected topics as long as new topics add a significant contribution to the total correlation of the model. We found that 12 is a suitable number of topics for all the three platforms, since new topics would add negligible contributions. Hence, we chose to keep it uniform across the three models, for better comparability. Finally, we ranked the obtained topics according to the fraction of the total correlation that they explain.

For Discord, a topic related to ''gaming and entertainment'' (characterized by words like `anime`, `memes`, `gamers`) explains most of the total correlation of the model, reflecting the user community originally targeted by the platform. In fourth position, we find a topic related to cryptocurrencies, characterized by generic words like `wallet`, `coin`, `exchange`, `btc`. Regarding Telegram, the most important topic learned by CorEx is characterized by words such as `referral`, `withdraw` and `bonus`. By leveraging results of previous studies [10], we are able to connect this topic to the well-known financial scam called *Ponzi scheme*, previously described in the Introduction. As a further confirmation of our labeling, the Telegram messages belonging to this topic share all the features outlined by the U.S. Securities and Exchange Commission [52] as red flags for recognizing Ponzi schemes: (i) promises of high investment returns with little or no risk, (ii) overly consistent returns, (iii) unregistered investments, (iv) unlicensed seller, (v) secretive and/or complex investment strategies, and (vi) no minimum investor qualifications. Another interesting finding is that channels associated to this topic are characterized by many similar messages repeatedly posted by Telegram bots, as shown in Figure 4a. The fourth topic is characterized by words like `pump`, `buy`, `sell` and `resistance`, that can be easily related to *pump-and-dump schemes* [9], previously discussed in the Introduction. Figure 4b provides a typical example of a chat in which organizers mobilize participants for the upcoming pump signal. They provide the target coin (e.g., $NAV) at the scheduled time, and they subsequently comment the results of the operation. In sixth, seventh and ninth positions, we find topics related to legitimate cryptocurrency discussions. One topic includes

TABLE 3. Topic modeling results, obtained by applying Anchored Correlation Explanation (CorEx) to Telegram and Discord channels, and Twitter users. Two online cryptocurrency manipulation schemes emerge: Ponzi and pump-and-dump.

rank	words	label
Discord		
1	anime, chill, roblox, nsfw, memes, hangout, gamers, giveaways, chats, fortnite	gaming and entertainment
2	**wallet**, **coin**, **exchange**, **crypto**, **blockchain**, **token**, **cryptocurrency**, btc, coins, address	legitimate crypto
9	**pump**, **signal**, **target**, people, money, high, big, pretty, better, buy	pump-and-dump
11	**referral**, **bonus**, **ref**, hosting, **withdraw**, services, service, network, website, opportunity	Ponzi scheme
Telegram		
1	**ref**, **referral**, **withdraw**, **bonus**, paying, doubler, instant, legit, doge, automatic	Ponzi scheme
3	**pump**, **target**, **signal**, **stoploss**, market, dump, chart, price, resistance, sell	pump-and-dump
4	**exchange**, **token**, **coin**, **crypto**, **blockchain**, **wallet**, **cryptocurrency**, tokens, exchanges, listed	legitimate crypto
11	game, games, fun, players, play, items, item, multiverse, edition, dragons	gaming and entertainment
Twitter		
1	**bullish**, **bull**, **currency**, **blockchains**, work, years, market, really, lot, buying	legitimate crypto
3	**pump**, **target**, **signal**, chart, looks, alts, term, break, bought, resistance	pump-and-dump
4	**bonus**, **withdraw**, ll, coins, **ref**, end, usd, let, worth, tweet	Ponzi scheme

(a) Example of Telegram "Ponzi scheme" chat.



(b) Example of Telegram "pump-and-dump" chat.

FIGURE 4. Chats of the cryptocurrency manipulation channels, showing the typical deception patterns outlined as red flags for recognizing Ponzi and pump-and-dump schemes.

words related to technological aspects (blockchain, technology, platform), the other two are oriented to finance (trading, investment). Similarly to Discord, also Telegram has a "gaming and entertainment" topic, occurring in the twelfth position. Moving to Twitter accounts, the first topic includes words like bullish, market, buying, cap, and it concerns legitimate cryptocurrency discussions focusing on financial speculations.

Also the other topics extracted appear as legitimate cryptocurrency discussions, focusing on different issues related to cryptocurrencies. For example, in the third position we find a topic addressing cryptocurrency reliability (`privacy`, `consensus`), whereas in sixth position the focus is on technology (`app`, `adoption`, `technology`). However, no traces of cryptocurrency manipulation schemes spontaneously emerge from Twitter data.

B. SEMI-SUPERVISED TOPIC EXTRACTION

Since we are interested in studying manipulations within the cryptocurrency ecosystem, we also leveraged the word anchoring feature of CorEx to improve topic separability, focusing on legitimate cryptocurrency, Ponzi scheme and pump-and-dump topics. We leveraged previous findings, obtained with the unsupervised approach, and domain knowledge derived from existing literature to choose appropriate anchor words. Despite Ponzi scheme and pump-and-dump do not emerge spontaneously on Discord and Twitter, we leveraged the capability of anchored topic modeling to find underrepresented topics, by forcing the same anchor words as Telegram. The results of this analysis are resumed in Table 3, where topics are ranked according to the amount of total correlation explained. For each topic, words are ordered according to mutual information with the topic, and anchors are highlighted in bold. Discord is still dominated by the ''gaming and entertainment'' topic. Thanks to anchoring, the legitimate cryptocurrency topic jumped to the second position and improved its quality, as confirmed by the coherence of non-anchored words. Despite anchoring, pump-and-dump and Ponzi schemes confirmed low contribution to correlation and poor internal coherence, showing marginal diffusion among Discord channels. For Telegram, anchoring increased the contribution of our topics of interest to the model correlation. Excellent topic quality was confirmed by the occurrence of non-anchored words with high coherence within each topic, like `dump` and `resistance` for pump-and-dump, or `doubler` and `instant` for Ponzi schemes. Finally, when applied to Twitter accounts, word anchoring induces the appearance of a dominant legitimate cryptocurrency topic. Moreover, a pump-and-dump topic emerges in third position, showing a decent coherence of the non-anchored words (`alts`, `resistance`). Instead, the Ponzi scheme topic appears spurious, as a result of a low significance in Twitter discussions.

C. NODE LABELING

We used the semi-supervised models to label Discord and Telegram channels, and Twitter accounts, according to their prevalent topic. In particular, we are interested in topics related to legitimate or deceptive cryptocurrency discussions. We leveraged CorEx to compute the correlation of an instance with each possible topic. Then, we labeled each instance with the most correlated topic. Notably, the incidental mentioning of just a few words of a topic is not sufficient to assign that topic as the instance label. Conversely, prevalent topics are determined by the systematic co-occurrence of the related words. Despite its accuracy cannot reach the one of a supervised classification model, our technique prevents possible biases towards specific cryptocurrency deception schemes that may be introduced by human annotators. In particular, Ponzi and pump-and-dump schemes spontaneously emerged from the data, whereas other well-known schemes (e.g., cryptocurrency thefts) did not.

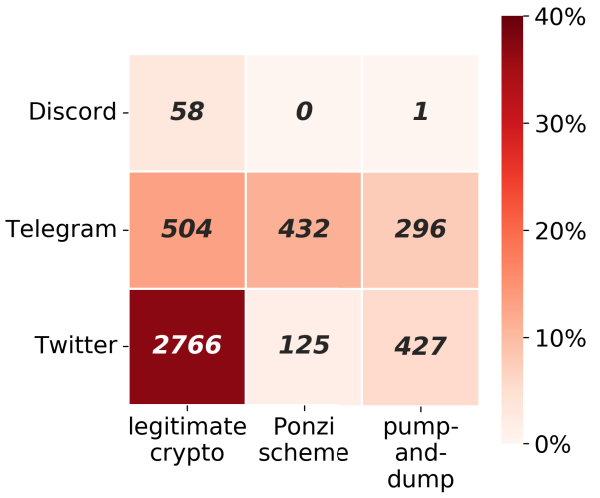


FIGURE 5. Heatmap of the percentage of channels/accounts per topic by platform. Each cell is annotated with the respective count. As opposite to Discord, Telegram shows high correlation with cryptocurrency-related topics, together with a remarkable presence of deceptive channels. Twitter accounts mostly debate about legitimate cryptocurrency topics, with a minor presence of pump-and-dump.

In Figure 5, a heatmap shows the channel/account percentage per topic per platform, focusing on cryptocurrency-related topics. Moreover, each cell is annotated with the respective instance count. Consistently with topic modeling results, Discord has low correlation with cryptocurrencies, with 58 channels labeled as legitimate cryptocurrency (3.3%), only one pump-and-dump and zero Ponzi scheme. On the contrary, Telegram hosts 504 legitimate cryptocurrency (13.2%), 432 Ponzi scheme (11.3%), and 296 pump-and-dump (7.8%) channels. Hence, the high correlation with cryptocurrency-related topics is confirmed, together with a remarkable presence of cryptocurrency manipulation channels. Instead, Twitter exhibits 2,766 (37.2%) accounts correlating with legitimate cryptocurrency, 427 (5.7%) with pump-and-dump, and 125 (1.7%) with the spurious Ponzi scheme topic.

In Figure 6, we color the invite link network nodes according to the assigned topic. Uncolored nodes correspond to non-English instances, or instances with a non-labeled (i.e., generic, uninteresting) topic. The isolated community of Discord channels, already mentioned in the previous Section, is clearly dominated by the gaming and entertainment topic (86.2% of nodes). A cluster of Ponzi scheme Telegram channels clearly emerges, which approximately corresponds to the cluster of Telegram channels magnified in Figure 2a.

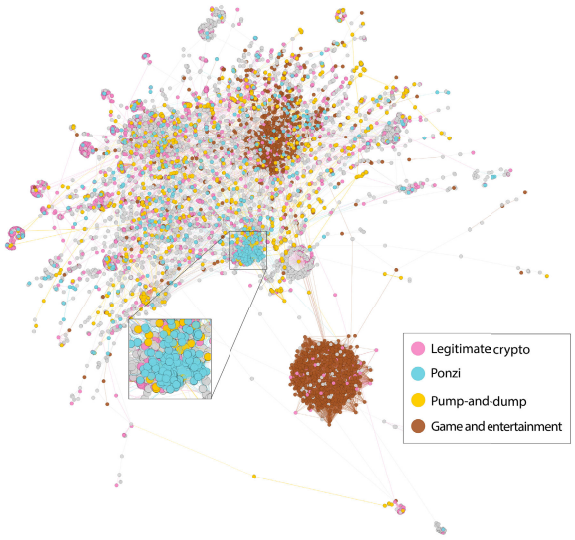


FIGURE 6. Invite link network with nodes colored according to their prevalent topic. It shows a dense cluster of Ponzi scheme channels engaged in mutual promotion. Instead, pump-and-dump channels are scattered across the network. Most of the labeled Twitter accounts are engaged in legitimate cryptocurrency discussions. The weakly connected Discord community is mainly engaged in game and entertainment.

Conversely, pump-and-dump and legitimate cryptocurrency Telegram channels are scattered across the network.

Finally, we observe the presence of star structures of legitimate cryptocurrency or unlabeled Twitter accounts, promoting Telegram channels. Conversely, we do not observe significant patterns involving the small percentage of Twitter pump-and-dump and Ponzi scheme accounts. We explain this behaviour by considering the different level of exposure of contents published on Twitter with respect to the more secluded Telegram or Discord channels. In fact, Twitter accounts promoting deceptive channels may prefer to dissimulate their purposes by posting seemingly-benign messages, to elude unwanted attention and possible interventions by platform administrators. Hence, no clear patterns of deception emerge from Twitter account characterization by topic, as opposed to the impressive star structures of bots emerged from the analysis of Section V. Therefore, for the sake of simplicity, in the remainder of our analysis we focus on the genuineness of Twitter accounts, leaving aside their topic characterization.

VII. UNCOVERING CRYPTOCURRENCY MANIPULATIONS

In previous sections, we sketched a map of the online cryptocurrency landscape by building the invite link network. Then, we added two semantic layers. The first one allowed us to label Twitter nodes according to their genuine or deceptive nature. The second one characterized Telegram and Discord channels according to their prevalent topic of discussion. In this way, two schemes of deception naturally emerged: pump-and-dump and Ponzi scheme. Now that we have charted the online landscape of cryptocurrency manipulations, we leverage our map for investigating the tracks of manipulation. Firstly, we ''zoom in'' to focus on the portions of the original network in direct contact with the deceptive channels. Then, we ''zoom out'' to interpret our results within the general framework of online manipulation.

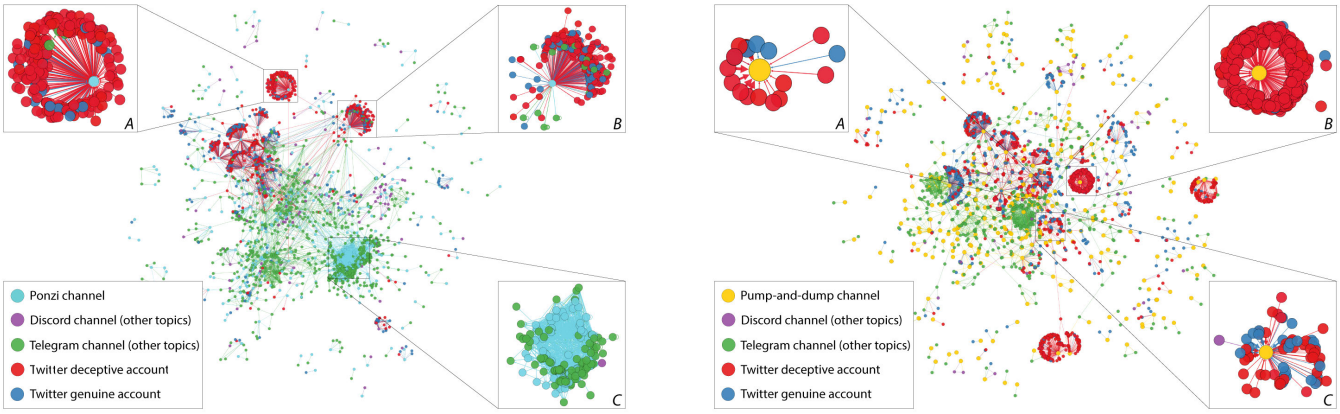
A. PONZI SCHEMES

In Figure 7a, we isolate Ponzi scheme nodes, their first neighbors, and the related edges. We color the 432 Ponzi scheme channels in pale blue, whereas we color the other channels according to the platform they belong to. We distinguish Twitter accounts according to their genuine (blue-colored) or deceptive (red-colored) nature. The scene is dominated by Telegram and Twitter platforms (1,124 and 1,696 nodes, respectively), with only 106 Discord nodes. We count 600 genuine and 1,096 deceptive Twitter accounts, resulting in a fraction of deceptive accounts of 64.6%, significantly higher than the one measured for the whole network (56.3%). There are 11 Twitter botnets with a size of at least 10 nodes, promoting a Ponzi scheme channel. They account for 15.9% of the total, while the Ponzi scheme channels are only 7.8% of the total amount of Discord and Telegram channels.

Two examples of those botnets are magnified in Figure 7a, in panels *A* and *B*. As shown, they feature the typical star structure that we previously highlighted. In panel *C*, we also highlight a dense cluster of Ponzi scheme Telegram channels, roughly corresponding to the one shown in Figure 6. This cluster is the largest cryptocurrency manipulation hub found in our study. It is composed of 166 Telegram nodes, 63 (39.8%) of which are Ponzi scheme channels. To understand its role within the Ponzi scheme ecosystem, in Figure 8a, we represent the heatmap of the number of invites per source node platform and target node topic. For source nodes, we also separate genuine Twitter accounts from deceptive ones. Results show that Ponzi scheme channels collect 71.4% of invite links shared by Telegram source nodes in the whole network. Moreover, in 92.3% of cases, invites targeting Ponzi scheme channels originated from other Ponzi scheme channels. Hence, most of the diffusion of invites to Ponzi scheme channels was carried out, within the examined cluster, by other Ponzi scheme channels. The engagement on mutual promotion within the Ponzi scheme cluster is further confirmed in Figure 8b, showing that the top-10 channels with the highest weighted out-degree perform Ponzi schemes.

B. PUMP-AND-DUMP

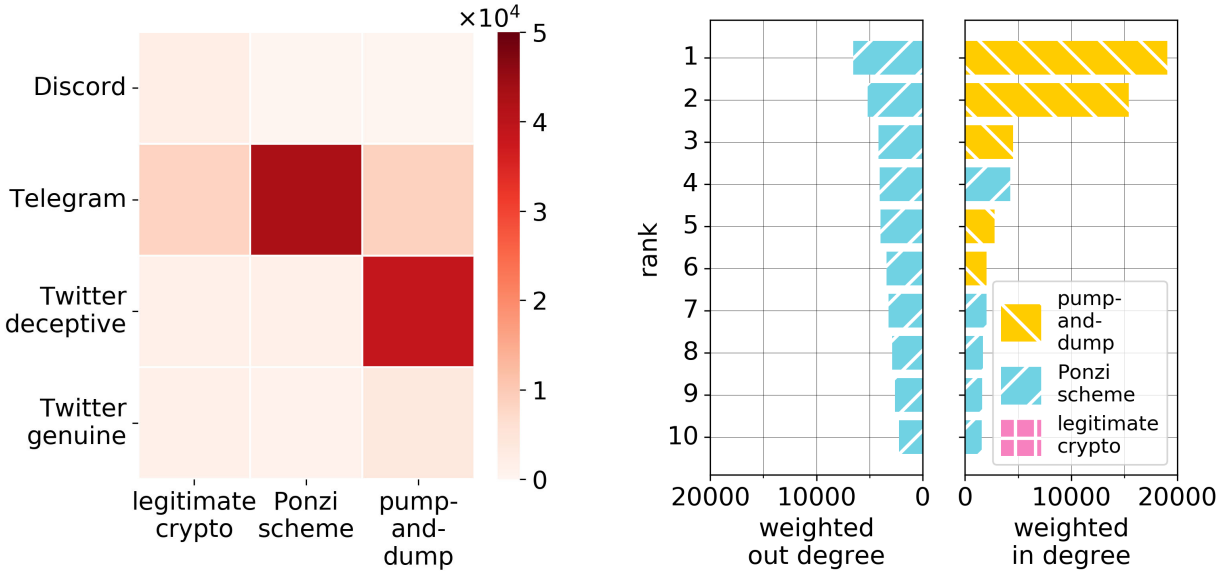
In Figure 7b, we depict the neighbour network of pump-and-dump channels. Pump-and-dump channels are colored in yellow, whereas for other nodes, we apply the same convention as before. Besides the 297 pump-and-dump nodes, the network is composed of 1,917 Twitter, 504 Telegram, and 52 Discord nodes. Pump-and-dump nodes are scattered across the network, and it is not possible to identify any cluster of them. As a result, they do not exhibit



(a) Neighbour network of Ponzi scheme channels.

(b) Neighbour network of pump-and-dump channels.

FIGURE 7. **Portions of the invite link network in direct contact with deceptive Ponzi scheme (7a) and pump-and-dump channels (7b). While Ponzi scheme channels are strongly engaged in mutual promotion, pump-and-dump channels are mainly endorsed by star structured Twitter botnets.**



(a) Heatmap of the number of invites per source node platform and target node topic.

(b) Top-10 cryptocurrency-related channels, sorted by weighted out (left) and in (right) degrees. Bars are colored according to the channel prevalent topic.

FIGURE 8. **Interplay between invite link diffusion and channel behaviour. Figure 8a shows that deceptive channels attract the most part of invites (87.8%). Ponzi scheme channels are mainly promoted within Telegram, while pump-and-dump channels receive most of the invites from Twitter deceptive accounts. Figure 8b confirms deceptive channels as the major hubs of invite diffusion. In fact, no legitimate cryptocurrency channel ranks in the top-10 in/out degree nodes.**

a strong engagement in mutual promotion, as opposed to Ponzi scheme channels described in Section VII-A. Also, in this sample, the fraction of Twitter accounts having a deceptive nature (65.4%) significantly exceeds the one measured on the whole network. They are frequently organized in botnets. In detail, we spot 15 botnets with a size of at least 10 accounts, promoting a pump-and-dump channel. They account for 21.7% of the observed botnets, resulting overrepresented if we consider that pump-and-dump is only 5.3% of the total channels. To estimate the contribution of those botnets in promoting pump-and-dump channels, we again resort to the heatmap of Figure 8a. We find out that Twitter deceptive accounts contribute to the 75.4% of all the invite links to pump-and-dump channels. Conversely, 92.9% of invite links, diffused by Twitter deceptive accounts, point to pump-and-dump channels. The effectiveness of Twitter deceptive accounts in promoting pump-and-dump channels is further proved by Figure 8b, showing that five of the top-6 channels with the highest weighted in-degree are labeled as pump-and-dump. The first three of

them are magnified in Figure 7b. They appear surrounded by their respective botnets, responsible for the high weighted in-degree of their target channels. The botnet in panel *B* promotes the MET∆.Symetra Telegram channel, resulting in the star structure that was magnified in figures 2a and 3. This channel and its botnet represent the largest invite diffusion hub in our study. Yet, to the best of our knowledge, the existence of this pump-and-dump channel was unknown prior to our analysis since it was never mentioned in existing studies, nor is it reported in authoritative lists of known pump-and-dump channels [9]. This result further supports the soundness of our method and the impact of our findings.

C. ZOOMING OUT TO THE GENERAL FRAMEWORK

Our exploration of the online cryptocurrency ecosystem confirms the concerns about the susceptibility of cryptocurrency markets to online manipulation, raised by authoritative agencies [52], [53]. While Discord appears as an overall healthy environment in our data sample, Twitter and Telegram reveal a strong interplay between numerous deceptive agents engaged in promoting scams. The choice of the invite link diffusion as the compass orienting our route proved to be particularly suitable for tracking online cryptocurrency manipulations. It was motivated by two hypotheses: (i) cryptocurrency manipulation stimulates the invite link diffusion because the efficacy of deceptive schemes strongly depends on recruiting a large number of participants, and (ii) the exchange of invite links implies homophily and common goals between the involved agents. The first assumption is supported by results shown in Figure 8b, proving that legitimate channels collect a negligible fraction of the overall invite links (12.2%). In contrast, cryptocurrency manipulation emerges as the main trigger to invite link diffusion in the online cryptocurrency ecosystem. The second assumption is supported by the existence of the dense cluster of Telegram Ponzi scheme channels, strongly committed in a mutual promotion. Further confirmation comes from the finding of several Twitter botnets, specially created to promote pump-and-dump channels. In both cases, agents sharing similar features, behaviors, and goals are strongly connected by the invite link diffusion.

Our study allows estimating the extent of deceptive content in the online cryptocurrency ecosystem that we explored. In fact, 56.5% of the cryptocurrency-related channels in our dataset appear to be involved in the deception. This result is even more significant if we consider that we retrieved channels starting from generic invite links occurring in cryptocurrency-related tweets. Despite avoiding to bias our data crawling by starting from a known seed of deceptive channels, as done in other works [9], [14], we still end up with a remarkable number of deceitful channels. Moreover, Twitter botnets emerge as the main vehicle for spreading pump-and-dump invites. This result enriches our knowledge on Twitter bot activities with a new element, relating our work with the flourishing line of research that aims to estimate how social bots manage to condition human activities in various ways, from contaminating the social debate [26], [54], [55] to adulterating the economic processes [23], [56].

Notably, our findings are not merely descriptive, but they provide actionable knowledge to counteract cryptocurrency manipulations. Tracking the major hubs of invite diffusion is a simple, effective way to spot malicious agents and manipulation schemes, as proven by the discovery of the previously unknown deceptive channels. Moreover, the success of these manipulations depends on the possibility to exploit invite link and Twitter bots. Hence, limiting the diffusion of invites and reducing the activity of bots would severely impair the efficacy of these frauds. In detail, in Section IV-A, we determined the small-world nature of the invite link network, underlying its potential vulnerability to selected hub suppression. Therefore, identifying and removing deceptive hubs can be an effective way to fragment the deceptive portions of the network, thus impairing the recruitment of participants. However, in Section VII-A, we observed that a significant fraction of the Ponzi scheme channels is clustered in a near-clique structure. Such type of very inter-connected structure is resilient to the shut down of a few selected hubs since other surviving edges can easily cover for the removed ones. As a consequence, more extensive actions are required to make it harmless. Hence, automatically detecting suspected Ponzi scheme channels is essential to cordon the compromised portion(s) of the network and correctly target the interventions. In Section VII-B, we reported that pump-and-dump channels are scattered throughout the network and poorly connected. At the same time, we found that they receive most of the invites from the star-shaped Twitter botnets. In this case, a promising countermeasure can be leveraging state-of-the-art bot detection tools to suspend Twitter deceptive accounts promoting such scams. Enforcing such actions could be particularly relevant for authorities responsible for the safety of the online financial markets, like the U.S. Securities and Exchange Commission and the U.S. Commodity Futures Trading Commission.

Cryptocurrencies were born to empower the dream of an accountable, decentralized, democratic payment method, preserving user privacy, and subtracting the consumer habits to the undesired scrutiny of governments and corporations [57]. In the same way, the Web was meant to realize the promise of an ecosystem granting free speech and equal access to information, goods, and opportunities to every human being [58]. Unfortunately, the conjunction between the potentialities of cryptocurrencies and the Web has opened the Pandora's box of criminal darknet markets, wild financial speculation, money laundering, criminal, and terrorist organization financing and deceptive manipulations [59]–[61]. This work addresses a peculiar example of those threats. Despite its specificity, typical patterns of online deception emerged, confirming the pervasiveness of these nasty phenomena across multiple aspects of online human activities. This work thus contributes towards raising collective awareness about the risks and the opportunities offered by cryptocurrencies to

our society. It also stimulates further research for designing countermeasures to the related threats.

VIII. CONCLUSIONS

Motivated by the increasing alarm raised by institutions about cryptocurrency manipulation, we mapped the online cryptocurrency ecosystem to identify, assess, and characterize possible threats. We analyzed the diffusion of invite links to cryptocurrency-related channels by cross-checking over 50M messages across Twitter, Telegram, and Discord platforms. Results confirmed our controlling idea, based on the hypothesis that the invite link exchange is a characteristic pattern related to deceptive schemes, as well as a proxy for homophily and common goals between the involved agents. First, we observed that two cryptocurrency manipulation schemes emerged – "pump-and-dump" and "Ponzi" – both affecting Telegram much more than Discord. Then, we identified a dense cluster of Ponzi scheme channels, so engaged in mutual promotion as to contribute to the 71.4% of the overall invite link diffusion measured on Telegram. Finally, we reported on 15 Twitter botnets that are responsible for the 75.4% of invite links to pump-and-dump channels, thus adding a new piece of knowledge about social bot activities.

Since institutions are evaluating the eligibility of cryptocurrencies as a legal payment method, our research community must raise awareness and design countermeasures to possible threats related to this emerging scenario. This work provides actionable knowledge, suitable to enforce more effective responses.

A. LIMITATIONS AND FUTURE WORK

Due to its explorative approach, this work presents some limitations.

First, although our study takes a step forward in tracking cryptocurrency manipulations across a wide and manifold online ecosystem, we still rule out relevant social media – especially forums – known to host fraudulent schemes [21]. Future work may include new platforms in order to further extend the analysis.

Second, despite Botometer offers a simple, effective way to detect Twitter bots by relying on individual account features, recent studies highlighted that social bots acting in coordination are harder to identify when inspected individually [33], [62]. Future work may extend such analysis by employing additional bot detection techniques, that also take into account coordinated behaviours [31], [63].

Finally, since we did not know in advance the deception schemes affecting our dataset, attempting a manual annotation according to predefined labels would introduce biases. Consequently, we let manipulation schemes spontaneously emerge from the data by leveraging a state-of-the-art, unsupervised approach. However, a supervised model would be more accurate in identifying the channels promoting deceptive actions. Hence, to overcome the accuracy limitations imposed by an unsupervised approach, we plan to perform a manual annotation, based on the deception schemes spontaneously emerged in this exploratory analysis. In this way, we can both further improve the accuracy of our results and provide a foundation for developing supervised models, capable of automatically detecting malicious actors and deceptive actions. Notably, the manual annotation can also enable more challenging, predictive tasks, targeting the occurrence of deception schemes (e.g., a pump and dump events), as well as the cryptocurrencies affected and their price fluctuations on the market.

REFERENCES

[1] M. Thelwall, "Can social news websites pay for content and curation? The SteemIt cryptocurrency model," *J. Inf. Sci.*, vol. 44, no. 6, pp. 736–751, Dec. 2018.

[2] J. C. Mendoza-Tello, H. Mora, F. A. Pujol-López, and M. D. Lytras, "Social commerce as a driver to enhance trust and intention to use cryptocurrencies for electronic payments," *IEEE Access*, vol. 6, pp. 50737–50751, 2018.

[3] R. C. Phillips and D. Gorse, "Mutual-excitation of cryptocurrency market returns and social media topics," in *Proc. 4th Int. Conf. Frontiers Edu. Technol.*, 2018, pp. 80–86.

[4] H. Jang and J. Lee, "An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information," *IEEE Access*, vol. 6, pp. 5427–5437, 2018.

[5] M. Tsikerdekis and S. Zeadally, "Online deception in social media," *Commun. ACM*, vol. 57, no. 9, pp. 72–80, Sep. 2014.

[6] P. M. Krafft, N. D. Penna, and A. S. Pentland, "An experimental study of cryptocurrency market dynamics," in *Proc. Conf. Hum. Factors Comput. Syst. (CHI)*, 2018, pp. 605:1–605:13.

[7] M. Rahouti, K. Xiong, and N. Ghani, "Bitcoin concepts, threats, and machine-learning security solutions," *IEEE Access*, vol. 6, pp. 67189–67205, 2018.

[8] U. W. Chohan, *The Problems of Cryptocurrency Thefts and Exchange Shutdowns*, 2018, doi: 10.2139/ssrn.3131702.

[9] J. Xu and B. Livshits, "The anatomy of a cryptocurrency pump-and-dump scheme," in *Proc. 28th USENIX Conf. Secur. Symp. (SEC)*, 2019, pp. 1609–1625.

[10] W. Chen, Z. Zheng, J. Cui, E. Ngai, P. Zheng, and Y. Zhou, "Detecting Ponzi schemes on Ethereum: Towards healthier blockchain technology," in *Proc. Web Conf. (WWW)*, 2018, pp. 1409–1418.

[11] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social Bots," *Commun. ACM*, vol. 59, no. 7, pp. 96–104, Jun. 2016.

[12] A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec, and M. Tiwari, "Global diffusion via cascading invitations: Structure, growth, and homophily," in *Proc. 24th Int. Conf. World Wide Web (WWW)*, 2015, pp. 66–76.

[13] T. Li, D. Shin, and B. Wang, *Cryptocurrency Pump-and-Dump Schemes*. 2018, doi: 10.2139/ssrn.3267041.

[14] M. Mirtaheri, S. Abu-El-Haija, F. Morstatter, G. V. Steeg, and A. Galstyan, "Identifying and analyzing cryptocurrency manipulations in social media," pp. 1–10, 2019, *arXiv:1902.03110*. [Online]. Available: https://arxiv.org/abs/1902.03110

[15] A. Feder, N. Gandal, J. Hamrick, T. Moore, A. Mukherjee, F. Rouhi, and M. Vasek, "The economics of cryptocurrency pump and dump schemes," C.E.P.R. Discussion Papers 13404, 2018.

[16] M. Glenski, E. Saldanha, and S. Volkova, "Characterizing speed and scale of cryptocurrency discussion spread on reddit," in *Proc. 28th Int. Conf. World Wide Web (WWW)*, 2019, pp. 560–570.

[17] Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, T. H. Kim, S. J. Kang, and C. H. Kim, "Predicting fluctuations in cryptocurrency transactions based on user comments and replies," *PLoS ONE*, vol. 11, no. 8, pp. 1–17, 2016.

[18] Y. B. Kim, J. Lee, N. Park, J. Choo, J.-H. Kim, and C. H. Kim, "When Bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation," *PLoS ONE*, vol. 12, no. 5, pp. 1–14, 2017.

[19] M. Vasek and T. Moore, "Analyzing the Bitcoin Ponzi scheme ecosystem," in *Proc. 23rd Int. Conf. Financial Cryptogr. Data Secur. (FC)*. Berlin, Germany: Springer, 2019, pp. 101–112.

[20] M. Bartoletti, B. Pes, and S. Serusi, "Data mining for detecting bitcoin Ponzi schemes," in *Proc. Crypto Valley Conf. Blockchain Technol. (CVCBT)*, Jun. 2018, pp. 75–84.
[21] M. Linton, E. G. S. Teo, E. Bommes, C. Chen, and W. K. Härdle, "Dynamic topic modelling for cryptocurrency community forums," in *Applied Quantitative Finance*, no. 18. Berlin, Germany: Springer, 2017, ch. 18, pp. 355–372.
[22] J. Kamps and B. Kleinberg, "To the moon: Defining and detecting cryptocurrency pump-and-dumps," *Crime Sci.*, vol. 7, no. 1, p. 18, Dec. 2018.
[23] S. Cresci, F. Lillo, D. Regoli, S. Tardelli, and M. Tesconi, "Cashtag piggybacking: Uncovering spam and Bot activity in stock microblogs on Twitter," *ACM Trans. Web*, vol. 13, no. 2, pp. 11:1–11:27, 2019.
[24] J.-P. Allem and E. Ferrara, "Could social bots pose a threat to public health?" *Amer. J. Public Health*, vol. 108, no. 8, pp. 1005–1006, Aug. 2018.
[25] D. A. Broniatowski, A. M. Jamison, S. Qi, L. AlKulaib, T. Chen, A. Benton, S. C. Quinn, and M. Dredze, "Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate," *Amer. J. Public Health*, vol. 108, no. 10, pp. 1378–1384, Oct. 2018.
[26] A. Bessi and E. Ferrara, "Social bots distort the 2016 US presidential election online discussion," *1st Monday*, vol. 21, no. 11, pp. 1–14, 2016.
[27] A. Addawood, A. Badawy, K. Lerman, and E. Ferrara, "Linguistic cues to deception: Identifying political trolls on social media," in *Proc. 13th Int. AAAI Conf. Web Social Media (ICWSM)*, 2019, vol. 13, no. 1, pp. 15–25.
[28] F. Schäfer, S. Evert, and P. Heinrich, "Japan's 2014 general election: Political bots, right-wing Internet activism, and prime minister Shinzō Abe's hidden nationalist agenda," *Big Data*, vol. 5, no. 4, pp. 294–309, Dec. 2017.
[29] F. B. Keller, D. Schoch, S. Stier, and J. Yang, "How to manipulate social media: Analyzing political astroturfng using ground truth data from South Korea," in *Proc. 11th Int. AAAI Conf. Web Social Media (ICWSM)*, 2017, pp. 564–567.
[30] E. Ferrara, "Disinformation and social bot operations in the run up to the 2017 French presidential election," *1st Monday*, vol. 22, no. 8, pp. 1–33, Jul. 2017.
[31] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "Social fingerprinting: Detection of spambot groups through DNA-inspired behavioral modeling," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 4, pp. 561–576, Aug. 2018.
[32] K. Kupferschmidt, "Bot-hunters eye mischief in German election," *Science*, vol. 357, no. 6356, pp. 1081–1082, 2017.
[33] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race," in *Proc. 26th Int. Conf. World Wide Web Companion (WWW)*, 2017, pp. 963–972.
[34] L. M. Aiello, M. Deplano, R. Schifanella, and G. Ruffo, "People are strange when you're a stranger: Impact and influence of bots on social networks," in *Proc. 6th AAAI Int. Conf. Weblogs Social Media (ICWSM)*, 2012, pp. 10–17.
[35] Z. Gilani, R. Farahbakhsh, and J. Crowcroft, "Do bots impact Twitter activity?" in *Proc. 26th Int. Conf. World Wide Web Companion (WWW)*, 2017, pp. 781–782.
[36] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, "BotOrNot: A system to evaluate social bots," in *Proc. 25th Int. Conf. Companion World Wide Web (IW3C2)*, 2016, pp. 273–274.
[37] M. Jiang, P. Cui, A. Beutel, C. Faloutsos, and S. Yang, "Catching synchronized behaviors in large networks: A graph mining approach," *ACM Trans. Knowl. Discovery Data*, vol. 10, no. 4, p. 35, 2016.
[38] N. Chavoshi, H. Hamooni, and A. Mueen, "DeBot: Twitter bot detection via warped correlation," in *Proc. IEEE 16th Int. Conf. Data Mining (ICDM)*, Dec. 2016, pp. 817–822.
[39] K. Starbird, "Disinformation's spread: Bots, trolls and all of us," *Nature*, vol. 571, no. 7766, p. 449, 2019.
[40] K. Starbird, A. Arif, and T. Wilson, "Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations," in *Proc. 22nd Conf. Comput. Supported Cooperat. Work Social Comput. (CSCW)*, 2019, pp. 1–26.
[41] E. Van Der Walt and J. Eloff, "Using machine learning to detect fake identities: Bots vs humans," *IEEE Access*, vol. 6, pp. 6540–6549, 2018.
[42] J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, "Cryptocurrency price prediction using tweet volumes and sentiment analysis," *SMU Data Sci. Rev.*, vol. 1, no. 3, p. 1, 2018.
[43] S. Cresci, S. Minutoli, L. Nizzoli, S. Tardelli, and M. Tesconi, "Enriching digital libraries with crowdsensed data," in *Proc. Italian Res. Conf. Digit. Libraries*. Cham, Switzerland: Springer, 2019, pp. 144–158.
[44] A.-L. Barabási and E. Bonabeau, "Scale-free networks," *Sci. Amer.*, vol. 288, no. 5, pp. 60–69, 2003.
[45] P. Erdos and A. Rényi, "On the evolution of random graphs," *Pub. Math. Inst. Hung. Acad. Sci.*, vol. 5, no. 1, pp. 17–60, 1960.
[46] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, p. 440, 1998.
[47] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
[48] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, "Online human-bot interactions: Detection, estimation, and characterization," in *Proc. 11th AAAI Int. Conf. Web Social Media (ICWSM)*, 2017, pp. 280–289.
[49] R. J. Gallagher, K. Reing, D. Kale, and G. Ver Steeg, "Anchored correlation explanation: Topic modeling with minimal domain knowledge," *Trans. Assoc. Comput. Linguistics*, vol. 5, pp. 529–542, Dec. 2017.
[50] S. Zhou, Y. Zhao, R. Rizvi, J. Bian, A. F. Haynos, and R. Zhang, "Analysis of Twitter to identify topics related to eating disorder symptoms," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Jun. 2019, pp. 1–4.
[51] R. Al-Rfou, B. Perozzi, and S. Skiena, "Polyglot: Distributed word representations for multilingual NLP," in *Proc. 17th ACL Conf. Comput. Natural Lang. Learn. (CoNLL)*, 2013, pp. 183–192.
[52] U.S. Securities and Exchange Commission. (2013). *Ponzi Schemes Using Virtual Currencies*. [Online]. Available: https://www.sec.gov/investor/alerts/ia_virtualcurrencies.pdf
[53] U.S. Commodity Futures Trading Commission. (2018). *Customer Advisory: Beware Virtual Currency Pump-and-Dump Schemes*. [Online]. Available: https://www.cftc.gov/sites/default/files/idc/groups/public/customerpro%tection/documents/file/customeradvisory_pumpdump0218.pdf
[54] M. Stella, E. Ferrara, and M. De Domenico, "Bots increase exposure to negative and inflammatory content in online social systems," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 49, pp. 12435–12440, 2018.
[55] S. C. Woolley, "Automating power: Social bot interference in global politics," *1st Monday*, vol. 21, no. 4, pp. 1–13, 2016.
[56] S. Cresci, F. Lillo, D. Regoli, S. Tardelli, and M. Tesconi, "$FAKE: Evidence of spam and bot activity in stock microblogs on Twitter," in *Proc. 12th AAAI Int. Conf. Web Social Media (ICWSM)*, 2018, pp. 580–583.
[57] Q. K. Nguyen, "Blockchain—A financial technology for future sustainable development," in *Proc. 3rd Int. Conf. Green Technol. Sustain. Develop. (GTSD)*, Nov. 2016, pp. 51–54.
[58] T. Berners-Lee and M. Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. Darby, PA, USA: DIANE Publishing Company, 2001.
[59] S. D. Brown, "Cryptocurrency and criminality: The bitcoin opportunity," *Police J. Theory, Pract. Princ.*, vol. 89, no. 4, pp. 327–339, Dec. 2016.
[60] C. Brenig, R. Accorsi, and G. Müller, "Economic analysis of cryptocurrency backed money laundering," in *Proc. 23rd Eur. Conf. Inf. Syst. (ECIS)*, 2015, pp. 1–18.
[61] K.-K. R. Choo, "Cryptocurrency and virtual currency: Corruption and money laundering/terrorism financing risks?" in *Handbook of Digital Currency*. New York, NY, USA: Academic, 2015, pp. 283–307.
[62] K. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini, and F. Menczer, "Arming the public with artificial intelligence to counter social bots," *Hum. Behav. Emerg. Technol.*, vol. 1, no. 1, pp. 48–61, Jan. 2019.
[63] M. Mazza, S. Cresci, M. Avvenuti, W. Quattrociocchi, and M. Tesconi, "RTbust: Exploiting temporal patterns for botnet detection on Twitter," in *Proc. 10th ACM Conf. Web Sci. (WebSci)*, 2019, pp. 183–192.

LEONARDO NIZZOLI received the M.Sc. degree in physics and the M.Sc. degree in big data analytics and social mining from the University of Pisa, where he is currently pursuing the Ph.D. degree with the Department of Information Engineering. He is a Research Fellow with IIT-CNR, Italy. His interests include AI-powered applications for social media intelligence, with a focus on improving the safety, security, and health of online and offline ecosystems.

SERENA TARDELLI is currently pursuing the Ph.D. degree in information engineering with the University of Pisa. She is a Research Fellow with the Institute of Informatics and Telematics of CNR, Italy. Her research interests include web science, computational social science, and data science, within the context of social media intelligence.

MARCO AVVENUTI (Member, IEEE) received the Ph.D. degree in information engineering from the University of Pisa. He is a Full Professor of computer systems with the Department of Information Engineering, University of Pisa. His research interests include human-centric sensing and social media analysis.

STEFANO CRESCI (Member, IEEE) received the Ph.D. degree in information engineering from the University of Pisa. He is a Researcher at IIT-CNR, Italy. His research interests include the intersection of web science and data science, with a focus on information disorder, coordinated inauthentic behavior, online social networks security, and crisis informatics. In 2018, he was selected among the winners of a SAGE Ocean Concept Grant. In 2019, he won the IEEE Computer Society Italy Section Chapter 2018 Ph.D. Thesis Award and the IEEE Next-Generation Data Scientist Award.

MAURIZIO TESCONI received the Ph.D. degree. He is a Researcher in computer science and leads the Web Application for the Future Internet Laboratory, Institute of Informatics and Telematics of CNR. His research interests include big data, web mining, social network analysis, and visual analytics within the context of open source intelligence. He is a member of the Permanent Team of the European Laboratory on Big Data Analytics and Social Mining, performing advanced research and analyses on the emerging challenges posed by big data.

EMILIO FERRARA (Senior Member, IEEE) received the Ph.D. degree. He is an Associate Professor of communication and computer science with the University of Southern California. He has published over 150 articles on social networks, machine learning, and network science. His research interests include theory, methods, and applications to study socio-technical systems. He received the 2016 DARPA Young Faculty Award, the 2016 Complex Systems Society Junior Scientific Award, and the 2018 DARPA Director's Fellowship. His research was supported by DARPA, IARPA, and AFOSR.