

RESEARCH

Open Access



# Do you trade with your friends or become friends with your trading partners? A case study in the $\tilde{G}1$ cryptocurrency

Nicolas Gensollen\*  and Matthieu Latapy

\*Correspondence:  
nicolas.gensollen@lip6.fr  
Laboratoire d'informatique de Paris  
6 (LIP6) 4 Place Jussieu (25-26 302),  
75005 Paris, France

## Abstract

We study the interplay between social ties and financial transactions made through a recent cryptocurrency called  $\tilde{G}1$ . It has the particularity of combining the usual transaction record with a reliable network of identified users. This gives the opportunity to observe exactly who sent money to whom over a social network. This social network is a key piece of this cryptocurrency, which therefore puts much effort in ensuring that nodes correspond to unique, well identified, real living human users, linked together only if they met at least once in real world. Using this data, we study how social ties impact the structure of transactions and conversely. We show that users make transactions almost exclusively with people they are connected with in the social network. Instead, they tend to build social connections with people they will never make transactions with.

## Introduction

When seeing two friends inviting each other, who could tell whether they are friends because they invite each other regularly, or if they invite each other regularly because they are friends? Such questions are common in systems where interactions occur over a social network. Obviously, interactions happen preferentially between socially connected individuals, but new ties are also created or reinforced through interactions. Understanding the mechanisms driving the evolution of these systems is an active field of research.

Most of the time, the social network is unknown and one can only record the interactions over time using various kinds of sensors. In such settings, a classical problem consists in inferring the network from traces of interactions. With the development of GPS and Bluetooth equipped devices such as smart phones, interaction data get more and more available, and there is a growing need to design efficient algorithms to infer the underlying relationships between entities. Alternatively, one may also study the opposite problem where the network is known and one wishes to infer associated interactions. Solving this kind of problems have various applications from traffic simulation to anomaly detection. Nonetheless, these problems are difficult to solve, and even difficult to study because of the lack of suitable data sources. We believe that better understanding the

underlying mechanisms at play in such complex systems is a useful step within this line of work.

In this paper, we thus investigate the interplay between social ties and financial transactions using real data from a specific cryptocurrency. Similarly to friends inviting each other, we wish to better understand whether transactions occur between individuals who were already socially connected, or if individuals build a new tie because they are involved in regular trades. Studying these questions is often challenging because financial transactions are often considered as sensitive data, and rarely made public; even when they are, interactions are anonymized.

In 2008, the blockchain technology (Nakamoto 2009) opened the doors to new virtual currencies which do not rely on a central authority. Transactions are written in a public distributed ledger, such that anybody can obtain the full list of transactions. Since then, the number and diversity of applications relying on the blockchain has been continuously growing (Al-Saqaf and Seidler 2017; Hileman and Rauchs 2017). Although Bitcoin is still a benchmark cryptocurrency (Hileman and Rauchs 2017; Gohwong 2018), many new currencies relying on different kinds of blockchains have been introduced since then (Gohwong 2018). Most provide (to some extent) anonymity to the entities making transactions. Indeed, users of these systems are often encouraged to create a new address when they want to make a new transaction, making the association of users and addresses a problem on its own (Meiklejohn et al. 2013; Remy et al. 2018). Some heuristics have been proposed to tackle this challenge but they mostly work for big users and they are difficult to assess. In addition, even if users were identified properly, underlying social ties would remain unknown.

The  $\checkmark$ 1 (BL et al. 2017) cryptocurrency under study in this paper relies on explicit social ties to strengthen the robustness of the system. It maintains an accurate network of identified users with reliable social ties, and uses it for monetary growth. This offers a unique opportunity to study the interplay between financial transactions and social ties between human beings of a specific community.

## Related Work

### *Inferring relations from interactions*

Scientific works tackling the general problem of inferring a network of social relationships from a sequence of interactions span several domains from sociology to computer science. With the increase of geotagged data availability due to the popularization of smart phones and other GPS equipped devices, a large portion of these studies focuses on the inference of social ties from mobility traces. Indeed, social networks are embedded in geography such that it is commonly assumed that interacting probability increases with physical proximity. For instance, the authors of Toole et al. (2015) use a call detail records (CDR) to explore the interplay between mobility and social ties, while the authors of Xu et al. (2019) rely on transactions made through student ID cards to build the students' social network.

It is well-known that all social relationships are not equivalent (Bapna et al. 2017; Xiang et al. 2010). Indeed, they can, among other things, be of different nature and have different strength. Being able to not only infer social ties from interactions, but also quantify their nature and strength is a key challenge within this line of work. In this direction, the authors of Gelardi et al. (2019) study the interactions in a group of baboons in which

they observe proximity and grooming as bounding activities. Similarly, the authors of Kobayashi et al. (2019) present a method to filter strong ties from temporal networks of interactions. Their method computes for each pair of nodes the distribution of their number of interactions in a null model based on node activities. Significant pairs of nodes are thus defined as those with a number of interactions that cannot be explained by the null model.

Slightly different studies aim at predicting missing links of a network from known links or other external features. For example, the authors of Crandall et al. (2010) use the proximity, in space and time, of geo-tagged photographs over the Flickr social network to infer the likelihood of a social tie between users. This paper shows that this probability increases by orders of magnitudes as the number of co-locations increases. Focusing on topological features, the authors of Hristova et al. (2016) explore the combined effect of multiple social networks for link prediction. More precisely, they represent social ties as a multiplex network where each layer represents a specific social platform, and they show how this additional information can be used to improve link prediction.

In Khosravi et al. (2013), the authors investigate link strength prediction in a social network based on social transactions (likes, comments, etc). They propose a new type of multiple-matrix factorization model for incorporating a transaction matrix between users, and test their method on *Cloob* (Cloob), a popular Iranian social network where users can rate their friendship relationships.

#### ***Inferring interactions from relations***

If inferring the network of relationships (or missing links of this network) from traces of interactions is often studied, the inverse problem of simulating traces of interactions from the network is also an interesting area of research. In Barrat et al. (2013) for example, the authors propose a procedure to generate dynamical networks from any weighted directed graph. This graph is considered as the accumulation of paths between its nodes, and the proposed procedure unfolds these paths using random walks of variable lengths. The authors show that their approach is able to generate dynamical networks with bursty, repetitive, or correlated behaviors.

#### ***The case of financial transactions***

Pioneering work studying both the nature and strength of social ties as well as the way people make transactions can be found in social sciences. In Zelizer (1996) for example, the authors propose to split payments into three categories: gifts, entitlements, and compensations, and show that each category corresponds to a specific set of social relationships and systems of meanings.

More recently, the authors of Martens and Provost (2011) use real but anonymized transaction records to infer a pseudo-social network of users in which two users are connected if they transferred money to the same entity. Then, they use this pseudo-social network for social targeting and obtain better performances than traditional models.

To the best of our knowledge, there is no previous work studying financial transactions and social interactions simultaneously from a reliable data source, even in specific settings. The recent development of cryptocurrencies is creating new opportunities for this kind of studies. Contrary to transactions relying on usual payment methods, blockchain based transactions are public and can be analyzed freely as long as the blockchain itself

is public. In Kondor et al. (2014), for instance, the authors extract the transactions from the Bitcoin blockchain and build the network of transactions. They provide a graph-based analysis of this network and show that linear preferential attachment drives its growth. In Popuri and Gunes (2016) the authors study the network of transactions of both Bitcoin and Litecoin, while the authors of Maesa et al. (2019) recently studied the structure of the Bitcoin users graph, exhibiting a bow tie like structure between its components. In Kim et al. (2016), the authors analyzed user comments in online communities of Bitcoin, Ethereum, and Ripple to predict the price and number of transactions in these cryptocurrencies.

The main limitation is often that, in most of these systems, most public keys are used only once such that there is no obvious way to link real users to the set of keys they used to make transactions (Meiklejohn et al. 2013; Remy et al. 2018). In addition, even if users were identified properly, underlying social ties would remain unknown.

### Our contribution

In this paper, we study a specific cryptocurrency which offers both a recording of transactions and of social bounds between identified human beings. This means that we know exactly who sent money to whom and when. Our main objective is to understand the interplay between these transactions and social ties between users.

More precisely, we first explore whether users start making transactions before creating a tie, or if they tend to make transactions with people they are already friends with. Going further, we study the different neighborhood structures and their evolution over time. We tackle questions such as: Are my transaction partners the same as my friends? How do my friends exchange between them compared to my transaction partners? Are my friends and transaction partners more and more homogeneous over time?

As we will see in “[Dataset and link stream modeling](#)” section, although the data is rather simple at first glance, the proper modeling of interactions is challenging and no unique, commonly accepted approach exists. We leverage here the recently introduced link stream model, which captures both the temporal and structural nature of data (Latapy et al. 2017; Latapy et al. 2019). We start our analysis with basic metrics targeting the questions above and we define link stream concepts as we need them in the analysis.

Therefore, our contribution is two-fold: it gives a modeling of the data that incorporates time and structure and it sheds lights on fundamental questions on the interplay between transactions and social ties, in the  $\check{G}1$  system. These insights are important for progress in several areas, like in particular the inference of social networks from interaction traces.

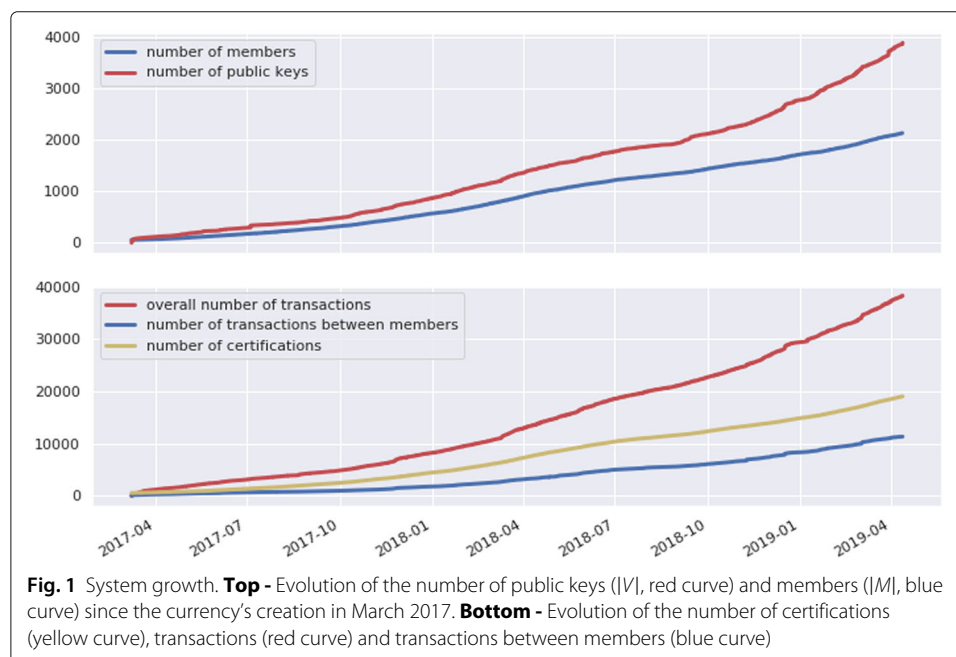
This paper is organized as follows. “[The  \$\check{G}1\$  cryptocurrency](#)” section introduces the  $\check{G}1$  cryptocurrency and explains the main ideas and mechanisms behind it. In “[Dataset and link stream modeling](#)” section, we present the dataset under study and show how link streams model interactions from this dataset. In “[Overview of certification and transaction streams](#)” section, we use time series and static graph concepts to give a first insight on the global structure and dynamics of the system. In “[How do new certifications and transactions appear between members?](#)” section we consider time and structure together but stay at a basic link level in order to understand the interplay between social ties and transactions. Finally, in “[Certifications and transactions neighborhoods](#)” section we use more complex stream concepts mixing time and structure in order to investigate this interplay further.

### The $\check{G}1$ cryptocurrency

$\check{G}1$  (BL et al. 2017) was introduced in France in March 2017 and relies on two kind of accounts: *member* and *anonymous* accounts. Both types are linked to a pair of cryptographic keys enabling them to make transactions. In the following, we denote by  $V$  the set of all accounts involved in  $\check{G}1$ , and by  $M \subseteq V$  the subset of member accounts. Therefore, the set  $A = V \setminus M$  contains all anonymous accounts. The top plot of Fig. 1 shows how the total number of anonymous and member accounts evolved since the currency’s creation in March 2017. At the time we downloaded the blockchain, in April 2019, there were 3872 accounts, among which 2128 member accounts.

While there is no control over the ownership of anonymous accounts, identification of entities behind member accounts is a key concern. Indeed, it is essential that a member account belongs to only one real and living human being, meaning that institutions such as companies or services have to rely on regular anonymous accounts. This constraint is linked to the monetary growth mechanism implemented within  $\check{G}1$ : each day the monetary mass increases and new units of the currency are injected in the system. These new units are distributed evenly between members such that each member receives exactly one share of the monetary growth.

This amount of money that members receive every day is called the *universal dividend* and is denoted by UD in the following. The growth rate of the monetary mass depends on the number of members  $|M|$  in the system and is updated every six months in the current implementation (BL et al. 2017). The purpose of this inflation mechanism is two-fold: first, it ensures that all members of a given generation are equal in terms of currency creation. Second, it ensures that the relative value of a dividend is constant over time, and that no generation is privileged over another by the currency creation itself. These two concepts were developed in Laborde () as the *spatial* and *temporal symmetries* of a currency, and  $\check{G}1$  is, to the best of our knowledge, the first cryptocurrency implementing them.



**Fig. 1** System growth. **Top** - Evolution of the number of public keys ( $|V|$ , red curve) and members ( $|M|$ , blue curve) since the currency’s creation in March 2017. **Bottom** - Evolution of the number of certifications (yellow curve), transactions (red curve) and transactions between members (blue curve)

It is easy to see that most currencies, independently of their use of the blockchain, do not implement these rules: some individuals benefit from the monetary creation at the expense of others (Kondor et al. 2014; Rose 2015). With fiat currencies, creation of new units is often a privilege given to states and banks and remains obscure to most citizens. One of the main objectives of Laborde () is to show that a currency can fulfill its purpose of enabling transactions of goods, while preserving fairness and equity in terms of monetary creation.

In order to identify its members,  $\check{G}1$  maintains a *web of trust* (WoT) between them. A link from person  $a \in M$  to person  $b \in M$  in the WoT is called a *certification* and means that  $a$  certifies that  $b$  is a real living human being, with no other member account:  $\nexists b' \in M, identity(b') = identity(b)$ . Members of the system wishing to participate in the currency creation process must hold enough certifications at all time and sign a license in which they commit to only give certifications to persons they know and trust.

When there is a pre-existing social tie between two persons, a certification can be seen as a projection of this bound, meaning that all possible relationships (family, work, friends...) get encoded as certification links because they imply enough trust. For two persons who don't already know each other, these links need to be built, which involves, most of the time, real life meetings. The  $\check{G}1$  community organizes regular local social events to facilitate the construction of these links. A certification graph, i.e. a web of trust, is thus a simplification of the real social network in which different types of co-existing relationships are encoded in the same type of bound. Despite this simplification, a certification link is supposed to exist only between human beings who have met at least once in real life, which might be more than one can expect in other online social networks. Therefore, we assume in this paper that certification links accurately reflect social ties strong enough to imply trust between members. In addition, certifications expire and have to be renewed which means that members have an incentive to build new connections in order to ensure their status. Moreover, building new connections between members is a way to reinforce the web of trust and protect the system against sybil attacks. The bottom plot of Fig. 1 shows the evolution of the number of certifications in the WoT, which increases up to 19049 certifications for 2128 members in April 2019.

Transactions can be done between any kind of accounts, members or anonymous. There is no transaction rate, and the only ways to earn currency units are either to receive them from someone already having coins (by selling an item or a service for example), or to become a member and earn the daily universal dividend. The bottom plot of Fig. 1 shows the evolution the overall number of transactions in the system and of transactions between members. In April 2019, there were 38296 transactions in the blockchain, among which 11345 occurred between members. Transactions are written in the blockchain by miners against retribution. A large difference with other cryptocurrencies like Bitcoin is that miner retributions come from donations and not from the process of mining itself. Indeed, the money creation is done by the members themselves through the universal dividends and has nothing to do with mining. In order to compensate the miners for their voluntary work, there is a special account in  $\check{G}1$ , called *Remuniter*, which receives donations and uses them to retribute miners.

### Dataset and link stream modeling

All the data we consider here is extracted from the G1 blockchain, which is publicly available (G1 blockchain explorer). It contains three key pieces of information: *identities*, *certifications*, and *transactions*. An *identity* associates a public key to a user name.

**Definition 1** (Certification) *A certification is a directed link between two members  $u \in M$  and  $v \in M$  and can be represented as a  $(t, u, v)$  triplet, meaning that entity  $u$  certifies entity  $v$  at time  $t$ .*

**Definition 2** (Transaction) *A transaction is a directed weighted link between two public keys (member or not)  $u \in V$  and  $v \in V$ , and can be represented as a  $(t, u, v, a)$  quadruplet, meaning that entity  $u$  sends an amount  $a$  to entity  $v$  at time  $t$ .*

This data has both structural and temporal components which make its analysis far from trivial. Classical methods such as static graphs or time series simplify data and information is lost in the process. In this paper, we use *link streams*, recently introduced in Latapy et al. (2017) and Latapy et al. (2019). Below, we give basic link stream notations and we refer the interested reader to these papers for more details.

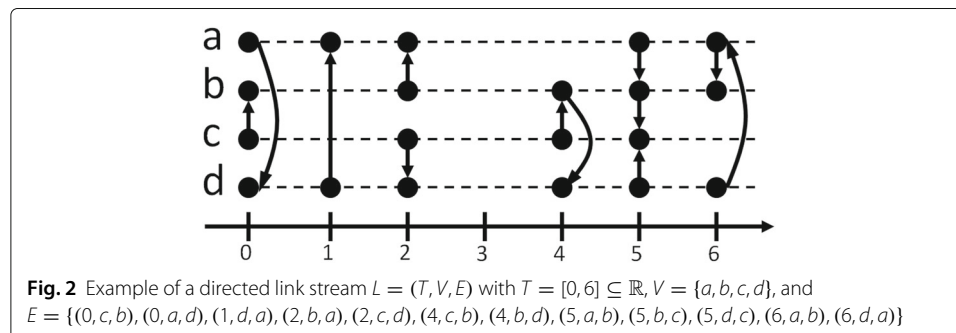
### Link streams

**Definition 3** (Directed Link Stream) *A directed link stream  $L$  is a triplet  $(T, V, E)$  where  $T$  is a set of time instants,  $V$  is a finite set of nodes, and  $E \subseteq T \times V \times V$  is a set of directed links. Having  $(t, u, v) \in E$  means that  $u$  is linked to  $v$  at time  $t$ .*

Figure 2 displays an example with four nodes  $a, b, c$ , and  $d$  (node labels are represented on the y-axis) over a time period going from  $t = 0$  to  $t = 6$  (time is represented on the x-axis). Each interaction  $(t, u, v)$  is represented as an arrow going from the horizontal line corresponding to node  $u$  to the one of node  $v$  at time  $t$ . For example, interaction  $(2, b, a)$  appears as an arrow between  $a$  and  $b$  at time  $t = 2$ .

Since link streams encode both time and structure, it is natural to define the *graph induced* by a stream and the *activity* of a stream:

**Definition 4** (Graph Induced by a Stream) *The graph induced by a directed link stream  $L = (T, V, E)$  is  $G(L) = (V, \bar{E})$  in which two nodes are linked if they interacted at least once in  $L$ , i.e.  $\bar{E} = \{(u, v), \exists (t, u, v) \in E\}$ .*



**Definition 5** (Activity of a Stream) *The activity of a link stream  $L = (T, V, E)$  is defined as the number of active links as a function of time, ie  $a_L(t) = |\{(u, v), (t, u, v) \in E\}|$ .*

#### Modeling certifications and transactions using link streams

**Definition 6** (Certification Stream  $\mathcal{C}$ ) *We denote by  $\mathcal{C} = (T_{\mathcal{C}}, M, E_{\mathcal{C}})$  the stream of certifications between members:*

- $T_{\mathcal{C}} = [1488987127, 1555054577] \subseteq \mathbb{R}$  is the **time interval** going from the first certification (on 2017-03-08 15:32:07) to the latest one considered in this study (on 2019-04-12 07:36:17)
- $E_{\mathcal{C}}$  is the set of **certification links**:  $(t, u, v) \in E_{\mathcal{C}}$  if  $u$  certifies  $v$  at time  $t$  (see Definition 1).

**Definition 7** (Transaction Stream  $\mathcal{T}$ ) *We denote by  $\mathcal{T} = (T_{\mathcal{T}}, V, E_{\mathcal{T}})$  the stream of transactions:*

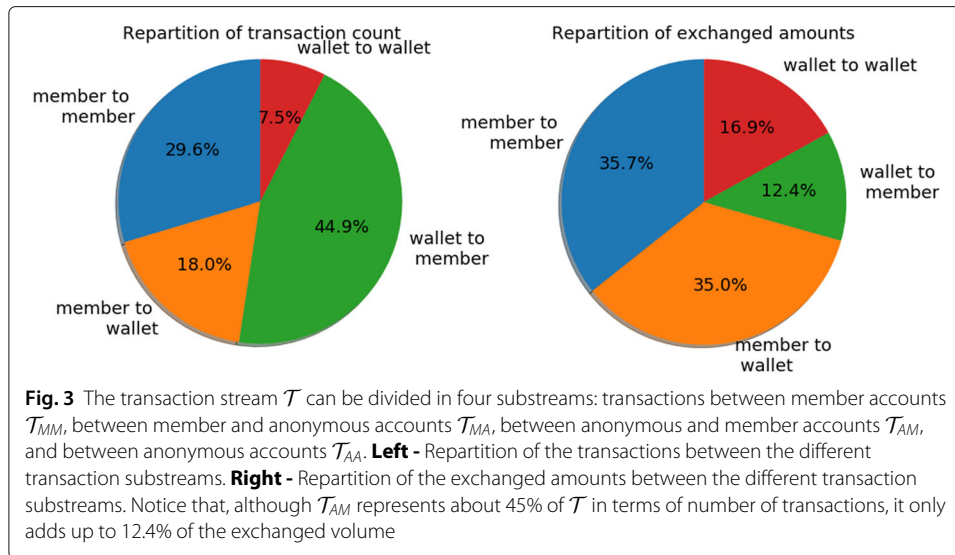
- $T_{\mathcal{T}} = [1488990898, 1555052722] \subseteq \mathbb{R}$  is the **time interval** going from the first transaction (on 2017-03-08 16:34:58) to the latest transaction considered in this study (on 2019-04-12 07:05:22)
- $E_{\mathcal{T}}$  is the set of **transaction links**:  $(t, u, v) \in E_{\mathcal{T}}$  if  $u$  sent a payment to  $v$  at time  $t$  (see Definition 2).

Notice that  $\mathcal{T}$  is defined over the node set  $V$  and contains transactions between members and anonymous accounts. We can define the *substream* induced by a set of nodes as the stream of links between these nodes (see Latapy et al. (2017) for details). Transactions can thus occur between member accounts (elements of  $M \subseteq V$ ), between anonymous accounts (elements of  $A \subseteq V$ ), or between member and anonymous accounts. Therefore, we split  $\mathcal{T}$  into four disjoint substreams:  $\mathcal{T}_{MM}$  which holds all transactions between member accounts,  $\mathcal{T}_{MA}$  which holds all transactions from a member account to an anonymous account,  $\mathcal{T}_{AM}$ , which holds all transactions from an anonymous account to a member account, and  $\mathcal{T}_{AA}$  which holds the transactions between anonymous accounts, such that  $\mathcal{T} = \mathcal{T}_{MM} \cup \mathcal{T}_{MA} \cup \mathcal{T}_{AM} \cup \mathcal{T}_{AA}$ .

Figure 3 shows the repartition of transactions between these four substreams as well as their repartition in terms of the total exchanged volume. As can be seen, transactions between member accounts (i.e. belonging to  $\mathcal{T}_{MM}$ ) represent only 30% of  $\mathcal{T}$  while transactions between an anonymous account and a member account (i.e. belonging to  $\mathcal{T}_{AM}$ ) represent almost 45% of  $\mathcal{T}$  and only 12.4% in terms of exchanged money, meaning that there are many small transactions from anonymous to member accounts. A reason for this asymmetry is linked to the way miners are retributed in  $\check{G}1$ . Recall that money creation is done by the members themselves and that it has nothing to do with mining, such that miners are doing the work for free by design. In order to compensate the miners for this voluntary work, there is a special anonymous account in  $\check{G}1$ , called *Remuniter*, which receives donations and uses them to retribute miners. This specific account is therefore a large hub in  $\mathcal{T}$  and stands for almost 40% of all transactions.

**Observation 1** *Almost 40% of all transactions are related to miner retribution.*

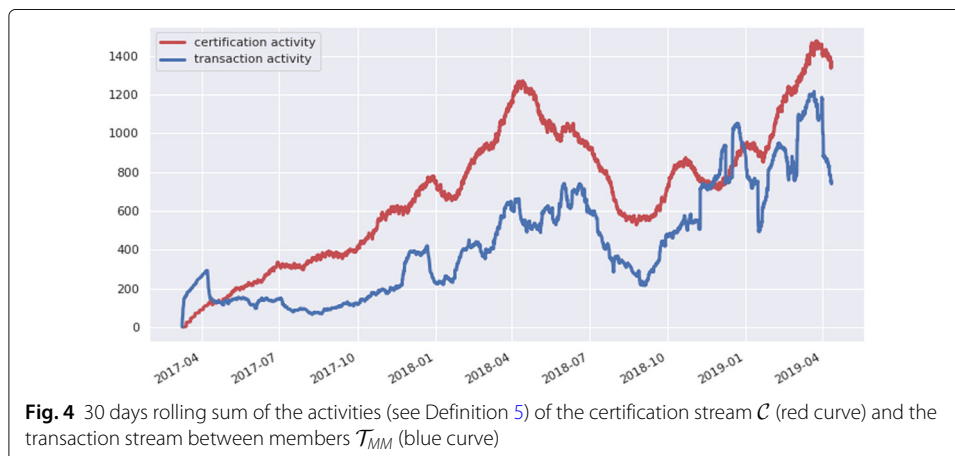




Removing all transactions involving *Remuniter* changes the transaction count repartition to: 49.2% for  $\mathcal{T}_{MM}$ , 28.4% for  $\mathcal{T}_{MA}$ , 10.3% for  $\mathcal{T}_{AM}$ , and 12.2% for  $\mathcal{T}_{AA}$ . Note that, thanks to this specific wallet, we can easily identify the miners of  $\check{G}1$  as the outgoing neighbors of *Remuniter* in  $\mathcal{T}_{AM}$  (i.e. all members who received money from *Remuniter*). At the time we downloaded the blockchain, in March 2019, there were 158 miners.

### Overview of certification and transaction streams

In this section, we present general properties of  $\mathcal{C}$  and  $\mathcal{T}$  in order to gain a better understanding of their structure and dynamics. We start by studying the activities (see Definition 5) of the certification stream  $\mathcal{C}$ , and of the transaction stream between members  $\mathcal{T}_{MM}$ . Figure 4 shows the 30 days rolling sum of these activities from March 2017 to March 2019. It appears clearly that they follow very similar trends over this period. A first growth period goes from March 2017 to April 2018 before a strong decrease in activity until early October 2018. Since then, both transaction and certification activities increase, with more volatility for transactions.



**Observation 2** *The certification dynamics is strongly correlated to the number of transactions among members.*

By aggregating interactions into a static graph, we obtain two induced directed graphs  $G(\mathcal{C})$  and  $G(\mathcal{T}_{MM})$  (see Definition 4). Figure 5 shows the in and out-degree distributions of each graph. The out-degree of a node in  $G(\mathcal{C})$  is the number of certifications given by this member, while its in-degree is the number of certifications it received. In  $G(\mathcal{T}_{MM})$ , the out-degree of a node is the number of transactions initiated by this member, while its in-degree is the number of transactions it received. First, note that there is no node with an in-degree smaller than 5 in  $G(\mathcal{C})$  because the minimum number of certifications required to become a member is set to 5 in the current implementation of  $\check{G}1$  (BL et al. 2017). Both graphs clearly display a heavy tailed degree distribution meaning that some members are involved in many more certifications or transactions than the majority.

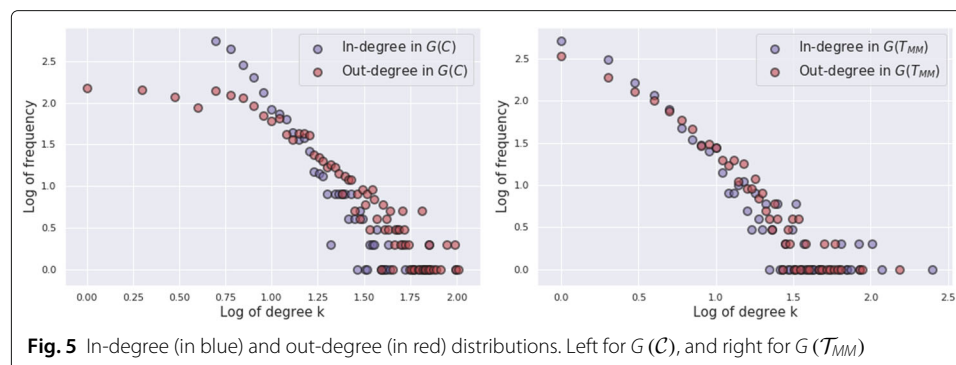
**Observation 3** *The certification graph and the member transaction graph have heterogeneous in and out degree distributions.*

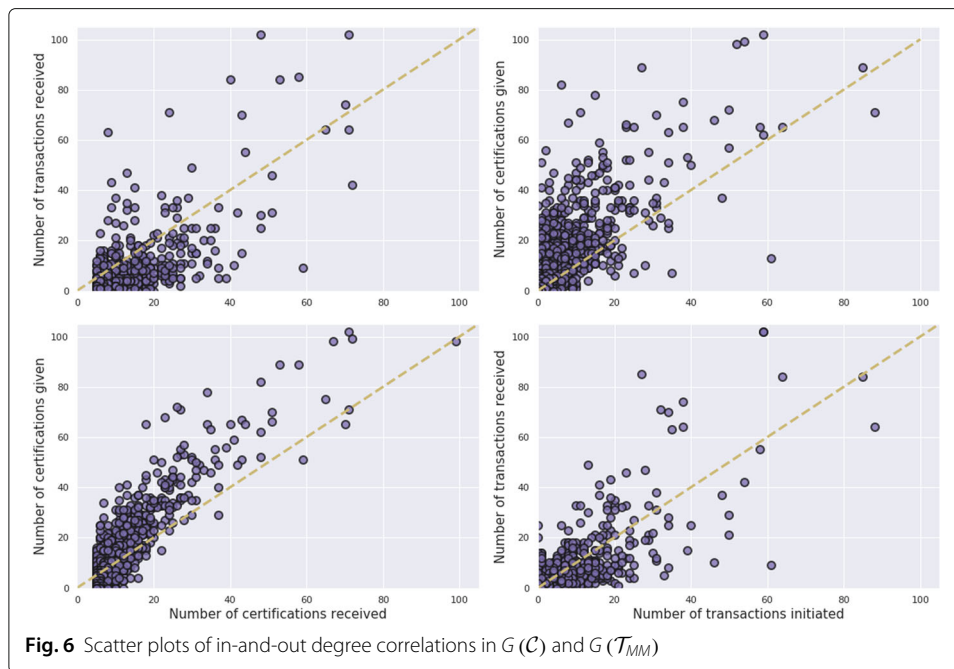
Figure 6 displays correlations between node degrees in  $G(\mathcal{C})$  and  $G(\mathcal{T}_{MM})$ . The top right subplot shows that members tend to give more certifications than they tend to initiate transactions, especially for high degree values. The bottom left subplot shows that nodes in-degree and out-degree values in  $G(\mathcal{C})$  are strongly correlated, meaning that people who receives a lot of certifications tend to also give a lot of them, and tend to actually give more certifications than they received.

**Observation 4** *In-coming and out-going activities are strongly correlated both from certification and transaction points of views.*

### How do new certifications and transactions appear between members?

A key goal of this paper is to gain insight on how certifications impact transactions and vice-versa. We therefore focus on the certification stream  $\mathcal{C}$  and the transaction stream restricted to identified members  $\mathcal{T}_{MM}$ . In this section we use a direct link-based approach to understand how new links appear in these two streams. More specifically, we wish to understand if a relationship between two members tends to exist in both streams and if it rather starts with a certification (seen as a social tie) or through transactions.



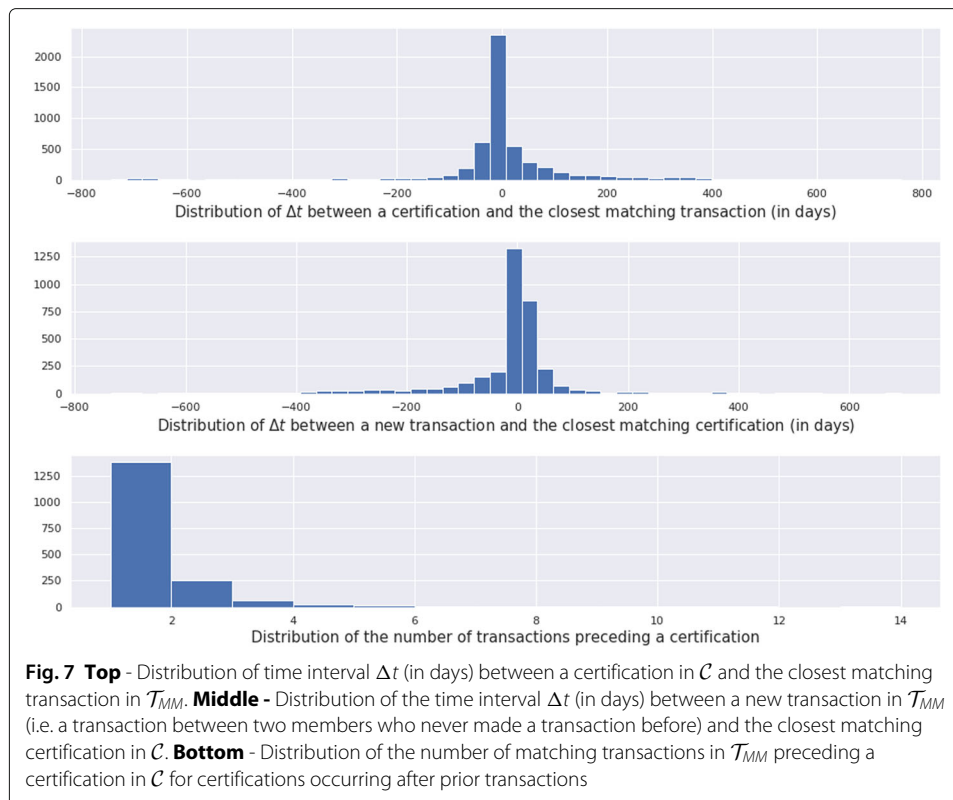


### Do social ties come before transactions?

When a certification occurs between two previously disconnected nodes in  $\mathcal{C}$  we search for the first transaction between these two nodes. This transaction can happen before the certification, after it, or never. For 73% of certifications, the two involved nodes never made a transaction. Instead, 16% of certifications have a pre-existing transaction, and 11% will make a transaction in the future. The top plot of Fig. 7 shows the distribution of the delay between a certification (between two nodes in  $\mathcal{C}$ ) and the closest matching transaction (between these two nodes in  $\mathcal{T}_{MM}$ ). The delay tends to be small in most cases. One possible explanation is that new members are often involved in small transactions (welcome gifts or acknowledgments) shortly after (or before) being certified. The bottom plot of Fig. 7 shows the distribution of the number of transactions in  $\mathcal{T}_{MM}$  preceding a certification in  $\mathcal{C}$  for these certifications that occur after transactions. Almost all such certifications occur after only one or two transactions, but in a very few cases, a certification can happen after as many as 14 prior transactions.

**Observation 5** *Most certification relationships in  $\mathcal{C}$  do not have a matching transaction in  $\mathcal{T}_{MM}$ , but when they do, most are close in time.*

Let us now investigate the other way round: When a transaction occurs in  $\mathcal{T}_{MM}$ , are the two involved nodes already linked by a certification or will they certify each other in the future? It turns out that 64% of all transactions from  $\mathcal{T}_{MM}$  occur between two members linked by a certification relationship in  $\mathcal{C}$ . More precisely, 42% occur between two already certified members, while only 22% occur between members who will certify themselves in the future. These numbers also mean that about 36% of all transactions from  $\mathcal{T}_{MM}$  occurs between members who never certify themselves directly in  $\mathcal{C}$ . We can also consider only new transactions in  $\mathcal{T}_{MM}$ , that is transactions between two members who never made a transaction before. The middle plot of Fig. 7 shows the distribution of time intervals  $\Delta t$



(in days) between a new transaction in  $\mathcal{T}_{MM}$  and the closest matching certification in  $\mathcal{C}$ . As can be seen, the majority of new transactions in  $\mathcal{T}_{MM}$  occurs shortly after a matching certification in  $\mathcal{C}$ .

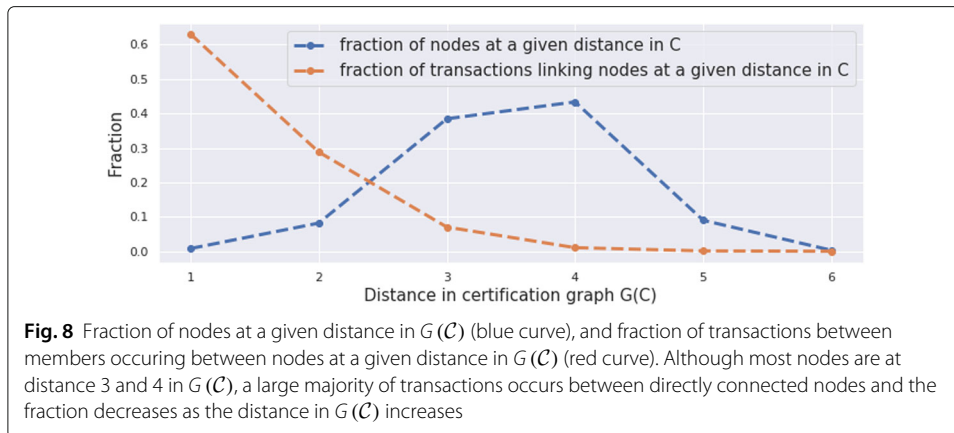
**Observation 6** *Almost two thirds of transactions between members occur between members linked by a certification.*

Because of the constraints imposed on the WoT, the undirected version of  $G(\mathcal{C})$  is connected and we can compute a certification distance between any two members. Because the WoT results from social interactions, we expect that nodes lying far apart in  $G(\mathcal{C})$  have a small probability of knowing each other, and therefore, of making transactions. Figure 8 shows the fraction of nodes at different distances in  $G(\mathcal{C})$  as well as the fraction of transactions between members occurring between nodes at different distances in  $G(\mathcal{C})$ . Although most nodes are at distance 3 and 4 (38% and 43% respectively) in  $G(\mathcal{C})$ , a large majority of transactions occurs between directly connected nodes and the fraction decreases strongly as the distance in  $G(\mathcal{C})$  increases.

**Observation 7** *Members who make transactions without being certified are linked by very short chains of certifications.*

### Connection probability

What makes two members more likely to be socially connected from a transaction perspective? Are two members more likely to know each other if they made several transactions with the same orientation or if they have made at least one reciprocal transaction? To answer such questions, we introduce:



$$\overline{\mathcal{C}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_C \text{ or } (t, j, i) \in E_C\} \tag{1}$$

$$\overrightarrow{\mathcal{C}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_C \text{ and } \nexists t : (t, j, i) \in E_C\} \tag{2}$$

$$\overleftrightarrow{\mathcal{C}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_C \text{ and } \exists t : (t, j, i) \in E_C\} \tag{3}$$

In other words,  $(ij) \in \overline{\mathcal{C}}$  if there exists at least one certification between  $i$  and  $j$  at some point in time,  $(ij) \in \overrightarrow{\mathcal{C}}$  if there exists only unidirectional links between  $i$  and  $j$  (either  $i \rightarrow j$  or  $j \rightarrow i$ ), and  $(ij) \in \overleftrightarrow{\mathcal{C}}$  if there exists at least one bi-directional certification link between  $i$  and  $j$ . Notice that  $\overleftrightarrow{\mathcal{C}} \subseteq \overline{\mathcal{C}}$  and  $\overrightarrow{\mathcal{C}} \subseteq \overline{\mathcal{C}}$ . We use the same notations for transactions between members:

$$\overline{\mathcal{T}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_{\mathcal{T}_{MM}} \text{ or } (t, j, i) \in E_{\mathcal{T}_{MM}}\} \tag{4}$$

$$\overrightarrow{\mathcal{T}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_{\mathcal{T}_{MM}} \text{ and } \nexists t : (t, j, i) \in E_{\mathcal{T}_{MM}}\} \tag{5}$$

$$\overleftrightarrow{\mathcal{T}} = \{(i, j) \in M \otimes M, \exists t : (t, i, j) \in E_{\mathcal{T}_{MM}} \text{ and } \exists t : (t, j, i) \in E_{\mathcal{T}_{MM}}\} \tag{6}$$

where  $E_{\mathcal{T}_{MM}}$  is the set of links of stream  $\mathcal{T}_{MM}$ . Table 1 gives the repartition of interactions among these sets. Although the number of certification links is higher than the one of transactions between members, i.e.  $|\overline{\mathcal{C}}|/|\overline{\mathcal{T}}| = 2$ , the fraction of bidirectional certifications is much higher than the one of bidirectional transactions:  $|\overleftrightarrow{\mathcal{C}}|/|\overleftrightarrow{\mathcal{T}}| = 3.75$ .

**Table 1** Repartition of interactions among the sets defined by Eqs. 1 - 6

$ \overline{\mathcal{C}} / M \otimes M $	$ \overrightarrow{\mathcal{C}} / M \otimes M $	$ \overleftrightarrow{\mathcal{C}} / M \otimes M $
0.008	0.005	0.003
$ \overline{\mathcal{T}} / M \otimes M $	$ \overrightarrow{\mathcal{T}} / M \otimes M $	$ \overleftrightarrow{\mathcal{T}} / M \otimes M $
0.004	0.0039	0.0008
$ \overline{\mathcal{T}} \cap \overline{\mathcal{C}} / \overline{\mathcal{C}} $	$ \overline{\mathcal{T}} \cap \overrightarrow{\mathcal{C}} / \overrightarrow{\mathcal{C}} $	$ \overline{\mathcal{T}} \cap \overleftrightarrow{\mathcal{C}} / \overleftrightarrow{\mathcal{C}} $
0.318	0.271	0.389
$ \overline{\mathcal{C}} \cap \overline{\mathcal{T}} / \overline{\mathcal{T}} $	$ \overline{\mathcal{C}} \cap \overrightarrow{\mathcal{T}} / \overrightarrow{\mathcal{T}} $	$ \overline{\mathcal{C}} \cap \overleftrightarrow{\mathcal{T}} / \overleftrightarrow{\mathcal{T}} $
0.535	0.497	0.715

**Observation 8** *Reciprocity of relationships in the social network is stronger than reciprocity in the transaction network.*

Table 1 also shows that the probability for two members to be linked in one stream increases a lot if they are connected in the other stream. More precisely, when two members are connected by a certification, they are almost eighty times more likely to make a transaction than two random members. Similarly, when two members made at least one transaction, they are 67 times more likely to be linked by a certification.

**Observation 9** *The probability two nodes to be linked in one stream is orders of magnitudes larger if they are linked in the other stream.*

In addition, reciprocity of transactions seems to play an important role in the probability of having a social tie. Indeed, two members having done at least one reciprocal transaction are almost ninety times more likely to be connected in the WoT than two random members.

**Observation 10** *Two members involved in reciprocal transactions are very likely to be socially connected.*

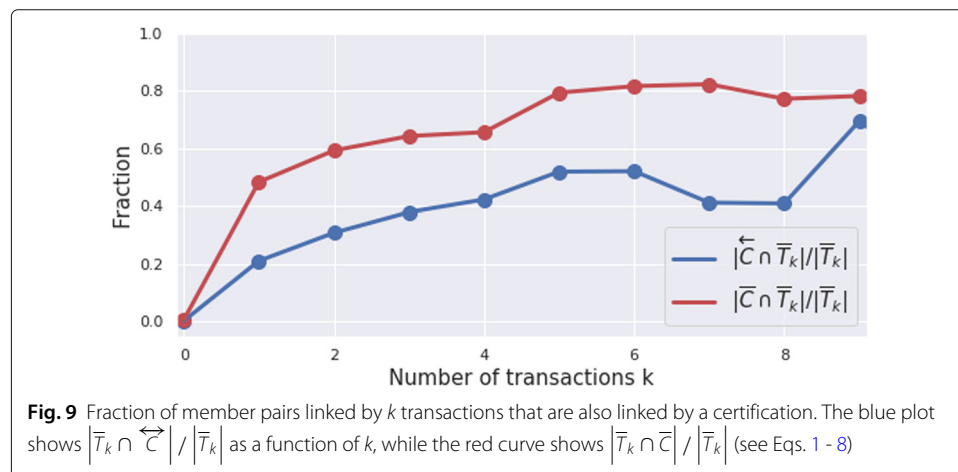
We now study how the number of transactions between two members influence their probability of having a social tie. We denote the number of transactions involving members  $i$  and  $j$  by:

$$|\tau_{ij}| = |\{(t, i, j) \in E_{\mathcal{T}_{MM}}\}| + |\{(t, j, i) \in E_{\mathcal{T}_{MM}}\}| \tag{7}$$

And we denote the set of member pairs linked by  $k$  transactions by

$$\bar{T}_k = \{(i, j) \in M \otimes M, |\tau_{ij}| = k\} \tag{8}$$

Figure 9 shows the fraction of member pairs linked by  $k$  transactions that are also linked by a certification. The blue plot shows  $|\bar{T}_k \cap \overleftarrow{C}| / |\bar{T}_k|$  as a function of  $k$ , while the red curve shows  $|\bar{T}_k \cap \bar{C}| / |\bar{T}_k|$ . These fractions globally increase with the number of transactions  $k$ . More than 80% of members linked by more than five transactions are connected in the WoT.



**Observation 11** *The probability for two members to be socially connected increases with the number of transactions they make.*

In conclusion, this section shows that certifications are often the first kind of link appearing between two  $\mathcal{G}1$  members, and that transactions occur afterwards. Moreover, these transactions do not happen randomly but constrained by the social network: they preferably take place between friends or between friends of friends. In addition, we showed that the existence of a social relationship between two members increases the probability of making a transaction. Conversely, the probability of being connected with a certification increases when members make transactions, especially when they are bidirectional. In the next section, we will investigate these questions further using more advanced concepts mixing time and structure.

**Certifications and transactions neighborhoods**

In previous section, we studied how new links occur in  $\mathcal{C}$  and  $\mathcal{T}_{MM}$  using a direct link-based approach. In this section, we go further using the topology of both streams to gain more insights.

A key concept in graph theory is the one of neighborhoods. For a directed link stream  $L = (T, V, E)$  the definition of the neighborhood of node  $v \in V$ , as given in Latapy et al. (2019), is the following cluster:

$$N_S(v) = \{(t, u), (t, u, v) \in E \text{ or } (t, v, u) \in E\} \tag{9}$$

A neighborhood is thus composed of temporal nodes rather than simple nodes for static graphs. We also consider the neighborhood of the undirected induced graph  $G(S)$ :

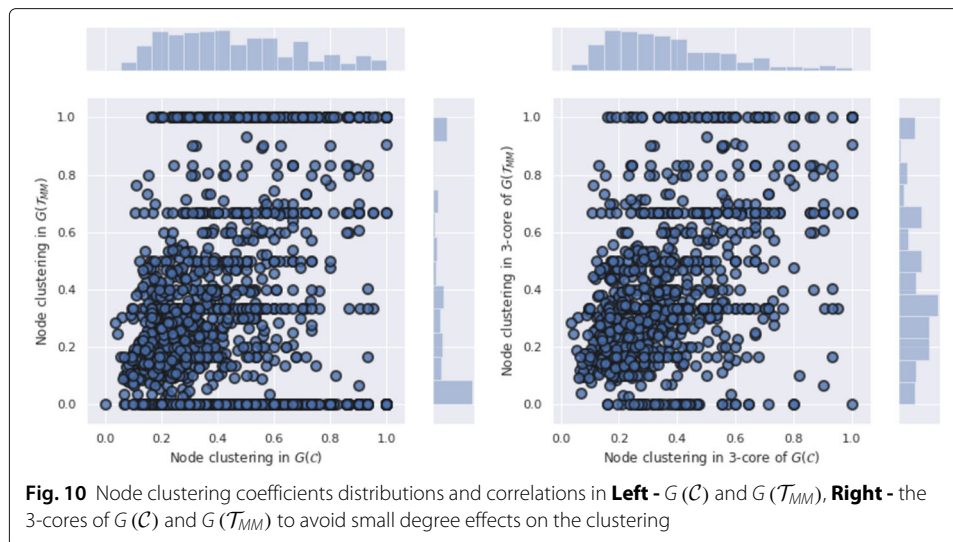
$$\bar{N}_S(v) = \{u \in V, \exists (t, u, v) \in E \text{ or } \exists (t, v, u) \in E\} \tag{10}$$

In other words,  $\bar{N}_S(v)$  contains all nodes which interacted with  $v \in V$  at least once, while  $N_S(v)$  keeps track of interaction times.

In addition to neighborhoods, we propose to study triangles in  $\mathcal{C}$  and  $\mathcal{T}_{MM}$ , which are a famous and very important concept in network theory. They intuitively convey the idea that *people I interact with tend to also interact between themselves*. A commonly used metric mixing triangles and neighborhoods in graphs is the clustering coefficient. The clustering coefficient of a node is the density of its neighborhood, and the clustering coefficient of the whole graph is defined as the average value over its nodes.

The average clustering coefficient for  $G(\mathcal{C})$  (resp.  $G(\mathcal{T}_{MM})$ ) is 0.49 (resp. 0.31), while the average for  $G(\mathcal{T}_{AA})$  is 0.13. Note that, by construction, the clustering coefficients for  $G(\mathcal{T}_{AM})$  and  $G(\mathcal{T}_{MA})$  are both equal to 0. In terms of triangles,  $G(\mathcal{C})$  contains 6589 triangles,  $G(\mathcal{T}_{MM})$  1990 triangles, and  $G(\mathcal{T}_{AA})$  only 393. These values are much higher than for random graphs with the same degree distributions (about 23 times higher for  $G(\mathcal{C})$ , and about 4 times higher for  $G(\mathcal{T}_{MM})$ ), meaning that both  $G(\mathcal{C})$  and  $G(\mathcal{T}_{MM})$  are very clustered networks. Figure 10 shows the distributions of the node clustering coefficients in both networks. Note that, although their average clustering values are similar, the distributions over nodes have different shapes.

The distribution for  $G(\mathcal{T}_{MM})$  exhibits two peaks around 0 and 1 which are due, to some extent, to nodes of degree 1 and 2 in the transaction graph, while the certification graph does not have nodes of degree less than 5 because of the minimum number of



certifications required to become a member. The right plot of Fig. 10 shows the same distributions for the 3-cores of these graphs, in which nodes of degree 1 and 2 were removed. We can see that, although there are less nodes with clustering 0, there are still nodes with very high clustering values in  $G(\mathcal{T}_{MM})$ , meaning that some members tend to make transactions in a very clustered way despite a less clustered certification neighborhood.

**Observation 12** Both  $G(\mathcal{C})$  and  $G(\mathcal{T}_{MM})$  are strongly clustered. However, the clustering coefficient distributions have different shapes. Some groups of members tend to make transactions in a very clustered way.

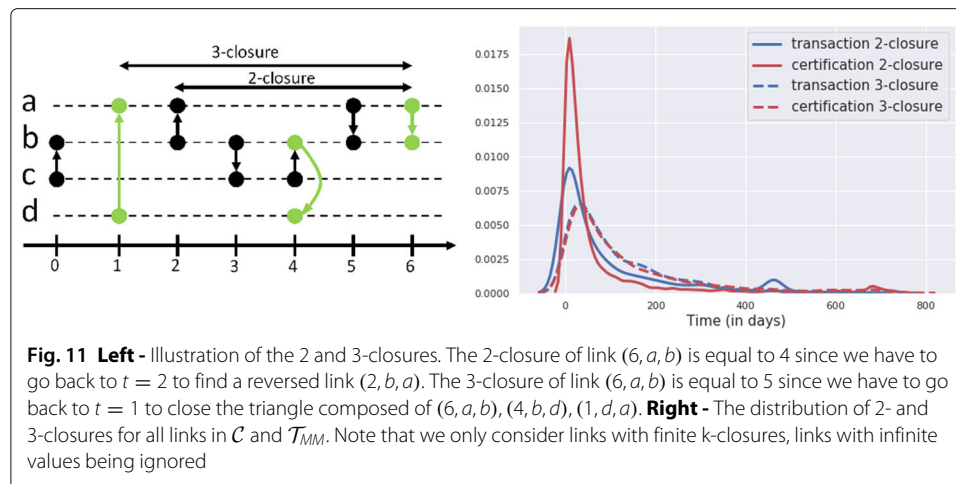
Previous results are built on induced graphs and so they do not take time into account. Clustering coefficient has been extended to link streams (Latapy et al. 2017), but provide little insight here. Indeed, the slow dynamics of  $\mathcal{C}$  make results very similar to the ones obtained for the induced graph  $G(\mathcal{C})$ , and  $\mathcal{T}_{MM}$  is a very sparse stream such that there is almost no time instant where transactions form triangles.

We propose another stream concept to study triangles with time: directed  $k$ -closure of links, where  $k$  is an integer larger than 1: the 2-closure, for example, is the time one has to go back to find a link between the same nodes but in the opposite direction. The 3-closure is defined as the amount of time required to find a triangle containing the considered link.

The left plot of Fig. 11 illustrates 2 and 3-closures in a simple case. The 2-closure of link  $(6, a, b)$  for example is equal to 4 since we have to go back to  $t = 2$  to find a link in the opposite direction, namely link  $(2, b, a)$ . The 3-closure of the same link  $(6, a, b)$  is equal to 5 since, this time, we have to go back to  $t = 1$  to find the triangle  $(6, a, b), (4, b, d), (1, d, a)$ .

Here we examine both the 2 and 3-closures for  $\mathcal{C}$  and  $\mathcal{T}_{MM}$ . The right plot of Fig. 11 shows the  $k$ -closure distributions for the two streams. The 2-closures for  $\mathcal{C}$  is generally smaller than for  $\mathcal{T}_{MM}$ , meaning that people tend to certify back faster than they make a transaction in the other direction. The 3-closures distribution of  $\mathcal{C}$  and  $\mathcal{T}_{MM}$  are very similar which seems to suggest that, although there are less triangles in  $\mathcal{T}_{MM}$  than in  $\mathcal{C}$ , they tend to appear within the same time frame.





**Observation 13** *Users certify back faster than they make a backward transaction. On the other hand, triangles are built within similar time frames in  $\mathcal{C}$  and  $\mathcal{T}_{MM}$ .*

## Conclusion

We studied a recent cryptocurrency which, despite its relatively small size, provides a very interesting and publicly available dataset to study transactions within a social environment. Because the data has both a structural and temporal component, neither a pure static graph nor a time serie based model are well suited for its study. We proposed here to rely on the link stream formalism to study the streams of certifications and transactions.

We showed that certifications are often the first type of link occurring between two previously disconnected nodes, suggesting that members of  $\check{G}1$  start to meet at social events before making transactions. In this context, the social network shapes the way transactions occur: even when they occur between two members without a certification relationship, transactions have a much higher probability of occurring between two socially close nodes. Investigating further, we showed that members tend to have a transaction neighborhood included within their certification neighborhood.

We also studied clustering coefficient of certifications and transactions and discovered that some members tend to make transactions in a very clustered way. We also studied the k-closure of certifications and transactions and found that users certify back faster than they tend to make a backward transaction, meaning that certification relationships become bi-directional faster than transactions.

These results might be specific to this particular cryptocurrency and call for further investigation to explore their generality, though. Indeed,  $\check{G}1$  is still very young and used mostly by a limited number of currency-interested people. As the currency will age, we expect  $\check{G}1$  usage to grow, and that time will play a more important role in the study of the interplay between social ties and interactions. Another interesting direction would be to compare the properties of  $\check{G}1$  with other social systems with transactions, like for instance Steemit (Steem 2018).

## Abbreviations

UD: Universal dividend; WoT: Web of trust

## Acknowledgements

The authors would like to thank Hugo Trentesaux for his help.

**Authors' contributions**

All authors have contributed equally to the paper. All authors read and approved the final manuscript.

**Authors' information**

Laboratoire d'informatique de Paris 6 (LIP6), University Pierre and Marie Curie, FR. Nicolas Gensollen and Matthieu Latapy

**Funding**

Not applicable.

**Availability of data and materials**

All the data used in this study are publicly available and can be obtained from the G1 blockchain (G1 blockchain explorer).

**Competing interests**

The authors declare that they have no competing interests.

Received: 28 January 2020 Accepted: 9 May 2020

Published online: 09 June 2020

**References**

- Al-Saqaf W, Seidler N (2017) Blockchain technology for social impact: opportunities and challenges ahead. *J Cyber Policy* 1–17. <https://doi.org/10.1080/23738871.2017.1400084>
- Bapna R, Gupta A, Rice S, Sundararajan A (2017) Trust and the strength of ties in online social networks: An exploratory field experiment. *MIS Q* 41:115–130
- Barrat A, Fernandez B, Lin KK, Young L-S (2013) Modeling temporal networks using random itineraries. *Phys Rev Lett* 110:158702
- BL et al. (2017) Dunitier G1 client. <https://g1.dunitier.fr/>
- Cloob. <https://www.cloob.com>. Accessed 20 Aug 2019
- Crandall DJ, Backstrom L, Cosley D, Suri S, Huttenlocher D, Kleinberg J (2010) Inferring social ties from geographic coincidences. *Proc Natl Acad Sci* 107(52):22436–22441
- G1 blockchain explorer. <https://g1.dunitier.fr/#/app/blockchain/search>. Accessed 20 Aug 2019
- Gelardi V, Fagot J, Barrat A, Claidière N (2019) Detecting social (in)stability in primates from their temporal co-presence network. *Anim Behav* 157:239–254
- Gohwong SG (2018) The state of the art of top 20 cryptocurrencies. *Asian Adm Manag Rev* 1
- Hileman G, Rauch M (2017) Global cryptocurrency benchmarking study. <https://doi.org/10.2139/ssrn.2965436>
- Hristova D, Noulas C, Brown A, Musolesi M, Mascolo C (2016) A multilayer approach to multiplexity and link prediction in online geo-social networks. *EPJ Data Sci* 5(1):24
- Khosravi H, Bozorgkhan A, Schulte O (2013) Transaction-based link strength prediction in a social network. In: 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM). pp 191–198. <https://doi.org/10.1109/cidm.2013.6597236>
- Kim YB, Kim JF, Kim W, Im JH, Kim TH, Kang SJ, Kim CH (2016) Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PLoS ONE* 11:1–17
- Kobayashi T, Takaguchi T, Barrat A (2019) The structured backbone of temporal social ties. *Nat Commun* 10(1):220
- Kondor D, Pósfai M, Csabai I, Vattay G (2014) Do the rich get richer? an empirical analysis of the bitcoin transaction network. *PLoS ONE* 9:e86197
- Laborde S. Relative theory of money
- Latapy M, Magnien C, Viard T (2019) Weighted, bipartite, or directed stream graphs for the modeling of temporal network. *CoRR abs/1906.04840*. [https://doi.org/10.1007/978-3-030-23495-9\\_3](https://doi.org/10.1007/978-3-030-23495-9_3)
- Latapy M, Viard T, Magnien C (2017) Stream graphs and link streams for the modeling of interactions over time. *Soc Netw Anal Min* 8. <https://doi.org/10.1007/s13278-018-0537-7>
- Maesa DDF, Marino A, Ricci L (2019) The bow tie structure of the bitcoin users graph. *Appl Netw Sci* 4(1):56
- Martens D, Provost F (2011) Pseudo-social network targeting from consumer transaction data. In: Faculty of Applied Economics, University of Antwerp, Belgium; Department of Information, Operations a. <https://ssrn.com/abstract=1934670>
- Meiklejohn S, Pomarole M, Jordan G, Levchenko K, McCoy D, Voelker GM, Savage S (2013) A fistful of bitcoins: Characterizing payments among men with no names. In: Proceedings of the 2013 Conference on Internet Measurement Conference, IMC '13. ACM, New York, NY, USA. pp 127–140
- Nakamoto S (2009) Bitcoin: A peer-to-peer electronic cash system
- Popuri MK, Gunes MH (2016) Empirical analysis of crypto currencies 644. [https://doi.org/10.1007/978-3-319-30569-1\\_21](https://doi.org/10.1007/978-3-319-30569-1_21)
- Remy C, Rym B, Matthieu L (2018) Tracking bitcoin users activity using community detection on a network of weak signals. In: Cherifi C, Cherifi H, Karsai M, Musolesi M (eds). *Complex Networks & Their Applications VI*. Springer International Publishing, Cham. pp 166–177
- Rose C (2015) The evolution of digital currencies: Bitcoin, a cryptocurrency causing a monetary revolution. *Int Bus Econ Res J (IBER)* 14:617
- Steeem anincentivized, blockchain-based (2018)
- Toole JL, Herrera-Yaqüe C, Schneider CM, González MC (2015) Coupling human mobility and social ties. *J R Soc Interface* 12(105):20141128
- Xiang R, Neville J, Rogati M (2010) Modeling relationship strength in online social networks. In: Proceedings of the 19th International Conference on World Wide Web, WWW '10. ACM, New York, NY, USA. pp 981–990
- Xu J-Y, Liu T, Yang L-T, Davison M, Liu S-Y (2019) Finding college student social networks by mining the records of student id transactions. *Symmetry* 11:307
- Zelizer V (1996) *Sociological Forum* 11(3)

**Publisher's Note**

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