electronics

MDPI

Article

Applying Social Network Analysis to Model and Handle a Cross-Blockchain Ecosystem

Gianluca Bonifazi, Francesco Cauteruccio, Enrico Corradini, Michele Marchetti, Domenico Ursino * and Luca Virgili

Department of Information Engineering, Polytechnic University of Marche, 60121 Ancona, Italy
* Correspondence: d.ursino@univpm.it

Abstract: In recent years, the huge growth in the number and variety of blockchains has prompted researchers to investigate the cross-blockchain scenario. In this setting, multiple blockchains coexist, and wallets can exchange data and money from one blockchain to another. The effective and efficient management of a cross-blockchain ecosystem is an open problem. This paper aims to address it by exploiting the potential of Social Network Analysis. This general objective is declined into a set of activities. First, a social network-based model is proposed to represent such a scenario. Then, a multi-dimensional and multi-view framework is presented, which uses such a model to handle a cross-blockchain scenario. Such a framework allows all the results found in the past research on Social Network Analysis to be applied to the cross-blockchain ecosystem. Afterwards, this framework is used to extract insights and knowledge patterns concerning the behavior of several categories of wallets in a cross-blockchain scenario. To verify the goodness of the proposed framework, it is applied on a real dataset derived from Multichain, in order to identify various user categories and their "modus operandi". Finally, a new centrality measure is proposed, which identifies the most significant wallets in the ecosytem. This measure considers several viewpoints, each of which addresses a specific aspect that may make a wallet more or less central in the cross-blockchain scenario.

Keywords: cross-blockchain ecosystem; Social Network Analysis; user behavior; wallet centrality; cross-blockchain money swaps; cross-blockchain data transfer

Citation: Bonifazi, G.; Cauteruccio, F.; Corradini, E.; Marchetti, M.; Ursino, D.; Virgili, L. Applying Social Network Analysis to Model and Handle a Cross-Blockchain Ecosystem. *Electronics* **2023**, *12*, 1086. https://doi.org/10.3390/electronics12051086

Academic Editors: Rashid Mehmood, Juan M. Corchado, Byung-Gyu Kim, Carlos A. Iglesias, In Lee and Fuji Ren

Received: 30 January 2023
Revised: 20 February 2023
Accepted: 20 February 2023
Published: 22 February 2023

1. Introduction

Since the introduction of the blockchain concept and the Bitcoin cryptocurrency by Satoshi Nakamoto in 2008 [1], a huge number of ideas, proposals, approaches and systems having blockchains as their fundamental core have been presented [2–4]. The main context in which blockchains initially developed was cryptocurrencies. This led to the emergence of a large number of cryptocurrencies, each managed by a blockchain, and a thriving ecosystem of blockchains supporting the creation and the exchange of amounts of money expressed in cryptocurrencies or tokens [5–7]. As evidence of this, Decentralized Finance (hereafter, DeFi) [8,9] is one of the most intriguing among the emerging technological evolutions in the context of global finance.

DeFi enables the decentralized delivery of financial services through a set of infrastructures, markets, technologies, systems and applications having blockchains as their core component. A challenging issue in the context of DeFi is cross-blockchain interoperability [10–14]. It aims to allow wallets to make transactions between different blockchains. The blockchains involved may have different architectures, consensus mechanisms, security models and privacy requirements. The problem of cross-chain interoperability is becoming increasingly relevant since the number and variety of blockchains involved in the DeFi context are growing enormously. This makes the solutions thought in the past to address this issue may now be ineffective and inefficient; hence, the definition of new approaches is compulsory. Regarding this scenario, there are several aspects that are attracting the interest of both academia and industry [15]. Examples of them concern the use and role of cross-blockchain tokens (i.e., tokens that can be exchanged across multiple blockchains [16]) and cross-blockchain data storage (i.e., data transfer across different blockchains [17]). All of these aspects require complex interactions and methodologies, which are the subject of current investigations by researchers.

A key component to facilitate cross-blockchain interoperability is represented by the cross-blockchain router protocols. Such a protocol, also known as cross-blockchain bridge, is a mechanism that allows different blockchains to communicate and share data with each other [18]. It provides the bridging infrastructure for enabling data to flow across different blockchains and, therefore, allows an asset to be transferred from one blockchain to another. The former is called "source blockchain", the latter is called "target blockchain", and the transfer is called swap or transaction. The transferred asset is typically an amount of money expressed in a cryptocurrency or a token, and the transfer requires the payment of a fee. Currently, there are already a few services that adopt cross-blockchain router protocols [19]. An example is Multichain (www.multichain.org (accessed on 29 January 2023)), formerly known as Anyswap (www.anyswap.org (accessed on 29 January 2023)), which supports the transfer of tokens, NFTs and general data across multiple blockchains.

The existing cross-blockchain interoperability solutions certainly implement some of the desiderata for this kind of system. As the cross-blockchain scenario is becoming increasingly complex and heterogeneous, the need for new solutions arises. One attempt in this direction is represented by the Token Atomic Swap Technology research project [16,20]. It aims to build an advanced platform for cross-blockchain interoperability and represents a generational leap over existing solutions. However, there are still many challenges to address in this context. These include: (i) the representation of a cross-blockchain scenario that involves multiple blockchains coexisting and interacting with each other; (ii) the efficient and effective management and tracking of money and data exchange activities between wallets in a cross-blockchain context; (iii) the development of new metrics and methods to capture the centrality and importance of different wallets in a cross-blockchain scenario; (iv) the possibility of identifying some interesting categories of users, along with their corresponding "modus operandi"; (v) the discovery of frauds and money washing activities in such a scenario.

Following such a reasoning, this paper wants to provide a contribution in this setting; in fact, it aims to define a set of tools that can be used by those researchers that strive to define and build next-generation cross-blockchain interoperability platforms. In particular, we first propose a model to represent the scenario of interest, i.e., a cross-blockchain ecosystem consisting of several blockchains in which wallets perform swaps to transfer money from one blockchain to another. Then, we propose a multi-dimensional and multi-view framework to manage a scenario represented through our model. Our framework starts from the idea that a cross-blockchain ecosystem can be effectively represented through particular multi-arc social networks. This allows the usage, in this scenario, of all the results found in the past Social Network Analysis research.

In order to demonstrate the feasibility of this idea, we focus on three typical concepts of Social Network Analysis, namely walk, path and cycle, and propose an approach for their detection in a cross-blockchain ecosystem handled by our framework. Then, we apply our approach to a real dataset derived from Multichain and show how it allows the detection of insights, through the in-depth investigation of the characteristics of the paths and cycles found. In turn, these insights allow us to discover and define real "modi operandi" of wallets operating in a cross-blockchain scenario.

As a final contribution, we propose a new centrality measure specifically tailored to detect the most important wallets in this context. This centrality takes into account several factors, each providing a viewpoint regarding the more or less central role played by a wallet in this ecosystem.

We believe that each of the four contributions mentioned above is already interesting in itself. For instance, in a DeFi scenario, no wallets should be able to accumulate sufficient power to monopolize the scenario and exclude others from participating [21]. An attempt by a wallet to monopolize the scenario can be easily monitored and highlighted through the centrality measure we are proposing in this paper. At the same time, collectively our contributions lay the foundations for a next-generation cross-blockchain interoperability platform, based on the use of Social Network Analysis, and capable of modeling and handling a cross-blockchain ecosystem.

In Summary, the main contributions of this paper are:

- A network-based model for representing and managing a cross-blockchain scenario;
- A multi-dimensional and multi-view framework for managing a cross-blockchain ecosystem;
- An approach that uses the concepts of Social Network Analysis to identify some particular categories of users and to define the "modus operandi" of each of these categories;
- A new centrality measure defined specifically to identify the most important wallets in a cross-blockchain context.

The outline of this paper is as follows. In Section 2, we provide an overview of related literature. In Section 3, we illustrate our model to represent a cross-blockchain ecosystem. In Section 4, we describe our social network-based framework and highlight that, thanks to it, several approaches, solutions and results found in the past Social Network Analysis research can be exploited in this context. In Section 5, we illustrate our approach for detecting walks, paths and cycles that ultimately allow the detection of "modi operandi" of several categories of users. In Section 6, we propose our centrality measure. In Section 7, we illustrate our experimental campaign and the insights derived from it. In Section 8, we present a discussion regarding our approach, its contributions, implications, applications, limitations and possible future developments. Finally, in Section 9, we draw our conclusions.

2. Related Works

Blockchains have been widely investigated and have stimulated a huge interest from both academia and industry researchers. Indeed, they have found applications both in classical contexts, such as cryptocurrencies [22,23], and in new ones, such as security [24,25] and IoT [26–29]. Due to their versatility, blockchains have given rise to several technologies and paradigms in recent years [30].

The massive diffusion of blockchains and the corresponding approaches and systems has led to a fragmentation of this scenario, on the one hand, and has made evident the need to define approaches for blockchain interoperability, on the other hand. This last problem has started to be investigated very recently, so there are only few approaches addressing it currently. Our paper falls into this context; specifically, it considers cross-blockchain interoperability in the case where transactions are performed to exchange tokens (and, thus, money) from one blockchain to another. This issue is part of the broader concept of Decentralized Finance (DeFi), which, in the last couple of years, has had, and is having, enormous growth.

Due to the small number of approaches to blockchain interoperability already proposed in the past literature, only very few papers can be directly compared with ours. However, if we look at the general blockchain context, we can find some aspects and features presenting correlations with our approach also in papers that do not specifically deal with blockchain interoperability; think, for instance, of the idea of modeling blockchain transactions using graphs.

For this reason, we thought of splitting this section into two subsections. The first is dedicated to the general topic of blockchains. In it, we will consider surveys and approaches that, while not dealing with cross-blockchain interoperability, share some similarities with ours. The second is specifically devoted to cross-blockchain interoperability.

Before going into detail in our discussion, we consider it necessary to highlight that, to the best of our knowledge, our approach is the first to use Social Network Analysis and, more specifically, a multi-arc model and a multi-view and multi-dimensional framework for representing and handling a cross-blockchain scenario. As we will see below, while this representation is more complex than those generally used in cross-blockchain scenarios, it allows us to extract some knowledge that would be much more difficult to derive through simpler models.

2.1. Investigations on Blockchains

According to Semantic Scholar (www.semanticscholar.org (accessed on 29 January 2023)), more than 90,000 blockchain-related studies have been published only in the last five years. To give an overview, albeit not exhaustive, on this topic and allow a systematization of these works, several surveys have been proposed. A survey that analyzes blockchains, along with the corresponding challenges and opportunities, is presented in [31]. In it, the authors first introduce a taxonomy on blockchains and, then, examine applications, technical challenges and advancements in tackling such challenges. This survey covers all the main perspectives that need to be considered when discussing the blockchain scenario. An interesting context in which blockchains are used extensively is the Internet of Things (IoT). In [32], the authors propose a survey investigating the role of blockchains for IoT. They propose what they call "Blockchain of Things" (BCoT), a sort of synthesis between blockchain and IoT. In [33], the authors propose a survey on blockchain security. In particular, they present a systematic examination of the security risks existing in some popular blockchains. They also illustrate an overview of real attacks, along with an in-depth description of the vulnerabilities exploited by attackers to conduct them. An aspect orthogonal to that of security in blockchains regards privacy in these systems. A survey on this issue can be found in [34].

In [35], the authors present a formal model of the transaction semantics that is guaranteed by a blockchain-based system. This model exploits on an acyclic directed graph that represents transactions in the blockchain. The latter are subject to several validation rules. The authors also define and study several properties of such a graph, with the goal of defining the valid states in which a blockchain could be found. Although the authors of [35] also use a network-based model, this study and ours have few similarities. Indeed, the study of [35] focuses on single-blockchain activities, while we focus on cross-blockchain transactions. Moreover, the approach of [35] concentrates on a very specific topic, namely the semantics of the blockchain scenario. Instead, we focus on several aspects and consider real scenarios, such as those arising from the framework Multichain.

In [36], the authors propose an approach that studies the transaction graph of a blockchain to learn and predict Bitcoin price dynamics. To this end, they focus on graph-based learning and topological feature extraction. Specifically, their approach is based on a combination of persistent homology and machine learning techniques. The approach described in [36] and ours are very different from each other; however, they can be considered orthogonal. In fact, the approach of [36] can be extended to transactions in a cross-blockchain scenario, where several features could be extracted to predict the dynamics of blockchain prices and token values.

In [37], the authors propose an approach to characterize the behavior of specific malicious activities on a blockchain scenario using time cycles. They first use a Depth First search algorithm to find such time cycles from the information associated with arcs. Next, they use such cycles to characterize specific malicious activities, such as gambling, phishing and money laundering. Both the approach of [37] and ours are graph-based and both of them search for cycles. However, the cycles detected in [37] are temporal, while the ones found by our approach have a completely different semantics. Finally, the approach of [37] works on a single blockchain, while our approach is cross-blockchain.

2.2. Investigations on Cross-Blockchain Interoperability

Along with numerous surveys that address general issues related to blockchains, there are also specific ones focusing on blockchain interoperability. In [38], the authors provide an overview on the current landscape concerning blockchain interoperability, from both an academic and an industry perspective. They propose a methodological framework assessing the criteria needed to define a blockchain interoperability solution. Then, based on it, they present a systematic literature review. Finally, they discuss the challenges and obstacles that characterize the development of blockchain interoperability solutions and present several use cases that could benefit from blockchain interoperability. A different point of view is presented in [39]. Here, the authors investigate the role of smart contracts in blockchain interoperability solutions. In particular, they organize their review based on the functionalities offered by smart contracts. Consequently, they address the functional, rather than the architectural, aspect of the interoperability between blockchains. They also present a taxonomy of interoperability solutions based on smart contracts and highlight three different combinations between smart contracts and blockchains.

In [40], the authors focus on the state of the art of cross-blockchain technologies. Here, they consider some interesting issues and challenges. They also present several open issues regarding normal and cross-blockchain smart contracts, cross-blockchain data exchange, and so on. In [41], the authors illustrate leading cross-blockchain technologies and highlight the differences between them. They also provide several details on the components of mainstream cross-blockchain technologies, such as hash lock and notary mechanisms.

A seminal work on blockchain interoperability is described in [10]. Here, the authors present the needs underpinning such an activity. Specifically, they consider two main aspects, namely cross-blockchain token transfer, and cross-blockchain smart contract invocation and interaction. Finally, they examine potential solutions and approaches for both levels. Both the work of [10] and ours focus on a cross-blockchain scenario. However, they have an essential difference in that the former presents the general problem of blockchain interoperability at a high abstraction level, while the latter discusses some aspects of this problem in detail.

In [42], the authors propose an approach to handle situations of mispricing in cryptocurrency market networks. It uses a graph whose nodes represent a combination of currency and blockchain market. Arcs indicate that an exchange between the corresponding pairs of currency and blockchain market is possible. Each arc is associated with a weight that indicates the rate to perform the corresponding exchange. The approach operates by searching for cycles in this particular graph. The approach of [42] and ours are both graph-based and both make use of cycles. However, the structure of graphs and the goals for which they extract cycles from graphs are completely different.

In [43], the authors present a method to support cross-blockchain interoperability. This method aims to address cross-communication between blockchains without an intermediary. The lack of an intermediary implies that transactions are verified through the consensus system of the blockchains involved. This method aims to be light, in the sense that it aims not to alter the main properties of the blockchains involved. The approach proposed in [43] and ours share the goal of handling cross-blockchain communication. However, the concepts and techniques used to achieve this goal are very different in the two cases.

In [44], the authors propose an approach for modeling and managing a cross-blockchain scenario. They represent such a scenario through a directed graph, whose nodes correspond to parties and whose arcs denote asset transfers. The approach presented in [44] and ours share the general goal, namely the management of a cross-blockchain scenario, and the use of graphs as a general model to perform such a management. However, there are several differences between the two approaches. First, the model used in [44] is a single-arc directed graph. Therefore, there can be at most one arc between two nodes. By contrast, our model allows the presence of multiple arcs between two nodes because it must represent a scenario in which a wallet can perform multiple transactions. Moreover, the model of [44] focuses on transactions where several parties cooperate. Instead, our approach considers swaps between the same wallet in different blockchains. Finally, the authors of [44] define a concept called "sequence", which indicates whether a party transferred the same asset. This concept has some similarities with our concept of cross-blockchain walk. However, there are also significant differences between the concept of sequence in [44] and our concept of walk. Indeed, sequence was thought for assessing whether a party has the considered asset as a resource; hence, it is only defined on transactions occurring in the same blockchain. By contrast, our concept of walk involves two or more blockchains.

3. Modeling the Scenario of Interest

The scenario of interest consists of m blockchains in which p multi-blockchain wallets operate. These are capable of operating on multiple blockchains with the same address. A wallet performs a series of transactions called swaps. A swap is a transaction performed by a wallet for exchanging tokens from a source blockchain to a target one. We assume that in our ecosystem there are q different tokens. Several information can be associated with each swap; they can be inferred through bridge services.

We model this scenario starting with a set $\mathcal{S} = \{s_1, \ldots, s_n\}$ of swaps. Each swap s_i can be represented as:

$$s_i = \langle id_i, wallet_i, bchain_i^{src}, bchain_i^{tgt}, token_i, val_i, qnt_i, fee_i, time_i, tval_i, conf_i \rangle$$

Here:

- id_i is the identifier of s_i;
- $wallet_i$ is the identifier of the wallet that executed s_i;
- $bchain_i^{src}$, $bchain_i^{tgt}$ ($bchain_i^{src} \neq bchain_i^{tgt}$) are the identifiers of the source and target blockchains, respectively; this implies that s_i is carried out from $bchain_i^{src}$ to $bchain_i^{tgt}$;
- $token_i$ is the identifier of the token involved in s_i;
- val_i is the average daily price of $token_i$;
- qnt_i is the quantity of $token_i$ involved in s_i;
- fee_i is the fee paid to perform s_i;
- $time_i$ is the timestamp of s_i;
- $tval_i$ is the time (in seconds) needed to validate s_i;
- $conf_i$ is the number of confirmations needed to validate s_i.

A further parameter that can be defined starting from the previous ones is the amount of money $money_i$ exchanged during the swap $s_i \in \mathcal{S}$. It can be defined as:

$$money_i = qnt_i \cdot val_i$$

Once we have defined $\mathcal{S}$, we can introduce some of its subsets. In particular, let b_k be a blockchain and w_j be a wallet. We define the subset $\mathcal{BS}_{kj}^{src}$ of swaps performed by w_j having b_k as the source blockchain as:

$$\mathcal{BS}_{kj}^{src} = \{s_i | s_i \in \mathcal{S}, b_k = bchain_i^{src}, w_j = wallet_i\}$$

Analogously, we define the subset $\mathcal{BS}_{kj}^{tgt}$ of swaps performed by w_j having b_k as the target blockchain as:

$$\mathcal{BS}_{kj}^{tgt} = \{s_i | s_i \in \mathcal{S}, b_k = bchain_i^{tgt}, w_j = wallet_i\}$$

Afterwards, we define the subset $\mathcal{BS}_{kj}$ of swaps performed by w_j and involving b_k (as source or target blockchain) as:

$$\mathcal{BS}_{kj} = \mathcal{BS}_{kj}^{src} \cup \mathcal{BS}_{kj}^{tgt}$$

Then, we define the subsets $\mathcal{BS}_j^{src}$ (resp., $\mathcal{BS}_j^{tgt}$) of swaps performed by w_j from (resp., to) any blockchain of the scenario:

$$\mathcal{BS}_j^{src} = \bigcup_{k=1}^{m} \mathcal{BS}_{kj}^{src} \qquad \mathcal{BS}_j^{tgt} = \bigcup_{k=1}^{m} \mathcal{BS}_{kj}^{tgt}$$

Clearly, in our cross-blockchain ecosystem, $\mathcal{BS}_j^{src} = \mathcal{BS}_j^{tgt} = \mathcal{S}_j$, i.e., the subset of swaps performed by w_j in the cross-blockchain ecosystem.

Continuing with the definition of the subsets of $\mathcal{S}$, let t_l be a token and w_j be a wallet. We define the subset $\mathcal{TS}_{lj}$ of the swaps performed by w_j to exchange money through the token t_l as:

$$\mathcal{TS}_{lj} = \{s_i | s_i \in \mathcal{S}, t_l = token_i, w_j = wallet_i\}$$

Next, we define the subset $\mathcal{TS}_j$ of swaps performed by w_j (independently of the tokens adopted) as:

$$\mathcal{TS}_j = \bigcup_{l=1}^{q} \mathcal{TS}_{lj}$$

Clearly, $\mathcal{TS}_j = \mathcal{S}_j$.

Now, we can define some other sets. The first of them is the set $\mathcal{B} = \{b_1, \ldots, b_m\}$ of the blockchains involved in the swaps of $\mathcal{S}$. A blockchain b_k belongs to $\mathcal{B}$ if there is at least one swap whose source or target blockchain is b_k. Formally speaking, we first define the set $\mathcal{B}^{src}$ of blockchains being the source of at least one swap of $\mathcal{S}$:

$$\mathcal{B}^{src} = \{b_k | s_i \in \mathcal{S}, b_k = bchain_i^{src}\}$$

Then, we define the set $\mathcal{B}^{tgt}$ of blockchains being the target of at least one swap of $\mathcal{S}$:

$$\mathcal{B}^{tgt} = \{b_k | s_i \in \mathcal{S}, b_k = bchain_i^{tgt}\}$$

Clearly:

$$\mathcal{B} = \mathcal{B}^{src} \cup \mathcal{B}^{tgt}$$

We can now introduce the set $\mathcal{W} = \{w_1, \ldots, w_p\}$ of the wallets involved in the swaps of $\mathcal{S}$. A wallet w_j belongs to $\mathcal{W}$ if it executed at least one swap in $\mathcal{S}$. Formally speaking:

$$\mathcal{W} = \{w_j | s_i \in \mathcal{S}, w_j = wallet_i\}$$

Let b_k be a blockchain of $\mathcal{B}^{src}$. We define $\mathcal{W}_k^{src}$ as the subset of wallets that performed at least one swap of $\mathcal{S}$ having b_k as source blockchain. $\mathcal{W}_k^{src}$ can be formalized as follows:

$$\mathcal{W}_k^{src} = \{w_j \in \mathcal{W} | s_i \in \mathcal{S}, b_k = bchain_i^{src}, w_j = wallet_i\}$$

Analogously, let b_k be a blockchain of $\mathcal{B}^{tgt}$. We define $\mathcal{W}_k^{tgt}$ as the subset of wallets that performed at least one swap of $\mathcal{S}$ having b_k as target blockchain. $\mathcal{W}_k^{tgt}$ can be formalized as follows:

$$\mathcal{W}_k^{tgt} = \{w_j \in \mathcal{W} | s_i \in \mathcal{S}, b_k = bchain_i^{tgt}, w_j = wallet_i\}$$

Finally, we define the subset $\mathcal{W}_k$ of wallets that have performed at least one swap of $\mathcal{S}$ involving a blockchain b_k of $\mathcal{B}$ as source and/or as target:

$$\mathcal{W}_k = \mathcal{W}_k^{src} \cup \mathcal{W}_k^{tgt}$$

A further set we introduce is the set $\mathcal{T} = \{t_1, \ldots, t_q\}$ of the tokens involved in the swaps of $\mathcal{S}$. A token t_l belongs to $\mathcal{T}$ if it is the reference token for at least one swap of $\mathcal{S}$. Formally speaking:

$$\mathcal{T} = \{t_l | s_i \in \mathcal{S}, t_l = token_i\}$$

Let w_j be a wallet of $\mathcal{W}$. We define the subset $\mathcal{T}_j$ of $\mathcal{T}$ as the set of tokens that have been used at least once by w_j:

$$\mathcal{T}_j = \{t_l | s_i \in \mathcal{S}, w_j = wallet_i, t_l = token_i\}$$

We conclude this section by introducing some parameters that model the amounts of money exchanged in the reference context. Specifically, we define:

- The quantity $\mathcal{M}_j$ of money exchanged by w_j as:

$$\mathcal{M}_j = \sum_{s_i \in \mathcal{S}_j} money_i$$

- The quantity $\mathcal{TM}_{lj}$ of money exchanged by w_j using the token t_l as:

$$\mathcal{TM}_{lj} = \sum_{s_i \in \mathcal{TS}_{lj}} money_i$$

- The quantity $\mathcal{BM}_{kj}$ of money exchanged by w_j involving b_k (as source or target blockchain) as:

$$\mathcal{BM}_{kj} = \sum_{s_i \in \mathcal{BS}_{kj}} money_i$$

4. A Framework to Represent and Handle Cross-Blockchain Scenarios

After modeling the scenario of interest, we are able to propose a multi-dimensional and multi-view framework for representing and analyzing the swaps performed. Our framework adopts several network-based representations of the swaps involved. It is multi-dimensional because it represents and analyzes the phenomenon of multi-blockchain swaps under different dimensions or perspectives (e.g., blockchains, tokens, wallets, etc.). It is multi-view because it represents the scenario of interest from different points of view (i.e., abstraction levels). It offers a self-contained way to handle all the data that can be associated with a set $\mathcal{S}$ of swaps represented by means of the model described in Section 3. Thanks to it, we can perform extensive data analytics activities on these data for extracting insights and knowledge patterns from them.

Our cross-blockchain framework can be represented as a triplet:

$$\mathcal{F} = \langle \mathcal{S}, \mathcal{H}, \mathcal{N} \rangle$$

Here, $\mathcal{S}$ is the set of swaps involved, while $\mathcal{H}$ and $\mathcal{N}$ are two network-based views of the scenario of interest. $\mathcal{H}$ is a hypergraph representing wallets and their involvement in each blockchain. $\mathcal{N}$ is a multi-arc n-partite social network representing the cross-blockchain transactions performed by wallets.

4.1. The View $\mathcal{H}$

The view $\mathcal{H}$ aims to represent wallets and their involvement in each blockchain. It can be defined as:

$$\mathcal{H} = \langle N, E \rangle$$

Here, $N = \{n_1, \ldots, n_p\}$ is the set of nodes of $\mathcal{H}$. There is a node $n_j \in N$ for each wallet $w_j \in \mathcal{W}$. $E = \{E_1, \ldots, E_m\}$ is the set of hyperedges of $\mathcal{H}$. There is a hyperedge $E_k \in E$ for each blockchain $b_k \in \mathcal{B}$. A hyperedge E_k links the nodes corresponding to the wallets that made at least one swap in b_k, regardless of whether the latter was the source or target blockchain. Based on graph theory, E_k also represents the subset of nodes connected by the hyperedge. Therefore, $E_k = \mathcal{W}_k$. The view $\mathcal{H}$ allows us to easily represent the meso-structure of the wallets operating within the cross-blockchain scenario. Each hyperedge is associated with a blockchain and connects the set of wallets operating in it. The intersection between two or more hyperedges tells us which wallets operate in all the blockchains corresponding to them.

4.2. The View $\mathcal{N}$

The view $\mathcal{N}$ is an n-partite multi-arc social network representing the swaps occurring in our reference context, along with the wallets performing them. In particular, $\mathcal{N}$ is constructed starting from $\mathcal{H}$ and allows the representation of the transactions performed by wallets when they swap tokens from a blockchain to another. $\mathcal{N}$ has a set of nodes for each blockchain and, more precisely, for each hyperedge of $\mathcal{H}$, and an arc for each swap. More formally:

$$\mathcal{N} = \langle E_1, \cdots, E_m, A \rangle$$

$\mathcal{N}$ has m sets of nodes. Each set E_k of nodes corresponds to a hyperedge of $\mathcal{H}$ or, equivalently, a blockchain of $\mathcal{B}$. As a consequence, each node $n_{j_k} \in E_k$ corresponds to a wallet $w_j \in \mathcal{W}_k$. A is the set of arcs of $\mathcal{N}$. There exists an arc $a_i \in A$ for each swap $s_i \in \mathcal{S}$. Recall that, in our reference context, a swap involves the same wallet in two distinct blockchains. A wallet can make one or more swaps between the same pair of blockchains; therefore, $\mathcal{N}$ is a multi-arc network. Since a swap involves the same wallet transferring money from a blockchain b_k to a blockchain b_h, we have that an arc $a_i \in A$ connects two nodes n_{j_k} and n_{j_h} corresponding to the same wallet w_j in b_k and b_h.

Observe that there exists a biunivocal correspondence between an arc $a_i \in A$ and the corresponding swap $s_i \in \mathcal{S}$. Therefore, in the rest of the paper, we use these two terms interchangeably. Furthermore, we associate with a_i all the information that we can derive for s_i and that we described in Section 3.

5. Modelling and Detecting Cross-Blockchain Walks, Paths and Cycles

Thanks to the view $\mathcal{N}$ introduced above, we were able to define a biunivocal correspondence between arcs and swaps. Having done such modeling, we are now able to perform more complex studies by applying concepts typical of graph theory on $\mathcal{N}$.

A first concept that can be extremely interesting to investigate is the one of walk, i.e., a sequence of arcs such that the target of an arc represents the source of the next one. In our reference context, the study of walks allows us to investigate contiguous swaps, i.e., a series of swaps by which a wallet moves a token through a succession of blockchains such that: (i) the transfer from one blockchain to another occurs through a swap; (ii) two successive swaps are performed at contiguous times (or, more formally, such that the time interval between the corresponding timestamps is less than a value ts^*). There might be various motivations for a user to perform a series of contiguous swaps. We will not deal with this issue here since it goes beyond the objectives of this paper. The interested reader is referred to [37,45,46] for an in-depth discussion of this topic. In the following, we will only study contiguous swaps (and, therefore, cross-blockchain walks) from a technical point of view.

Let r be a positive integer and let ts^* be a time interval. A cross-blockchain walk $W = (a_1, a_2, \cdots, a_r)$ in $\mathcal{N}$ is a succession of arcs such that:

- $a_i \in A, 1 \leq i \leq r$; this is equivalent to saying that all the arcs of W must belong to $\mathcal{N}$;
- $token_i = t_l, 1 \leq i \leq r$; this is equivalent to saying that the swaps corresponding to all the arcs of W must refer to the same token t_l;
- $bchain_i^{tgt} = bchain_{i+1}^{src}, 1 \leq i < r$; this is equivalent to saying that the target blockchain of one arc of W coincides with the source blockchain of the next arc;
- $time_{i+1} - time_i \leq ts^*, 1 \leq i < r$; this is equivalent to saying that the time interval between the timestamps of the swaps corresponding to two consecutive arcs of W is less than or equal to a value ts^*.

After formalizing the concept of cross-blockchain walk, we can define an algorithm that searches for all cross-blockchain walks of length r made by a wallet w_j to exchange a token t_l among different blockchains. It is formalized in Algorithm 1.

Algorithm 1 Algorithm for detecting cross-blockchain walks of length r in a cross-blockchain ecosystem.

Input

- $\mathcal{N}$: the view representing the cross-blockchain ecosystem
- w_j: a wallet
- t_l: a token
- r: a positive integer
- ts^*: a time interval

Output

- $WS^r_{j_l}$: a set of cross-blockchain walks of length r performed by w_j to exchange a token t_l through the blockchains of $\mathcal{N}$

Require: A_{j_l}: a set of arcs; Q: a queue of arcs; *Visited*: a set of arcs; W: an ordered list of arcs representing a walk

```
WS^r_{jl} = ∅
A_{jl} = selectArcs()
for a_i ∈ A_{jl} do
    enqueue(Q, a_i)
    Visited = {a_i}
    while not isEmpty(Q) do
        a_u = front(Q)
        if length(a_i, a_u) = r then
            W = reconstructWalk(a_i, a_u)
            WS^r_{jl} = WS^r_{jl} ∪ {W}
        else
            for a_v ∈ selectNextArcs(a_u) do
                if a_v ∉ Visited and bchain_u^tgt = bchain_v^src and (time_v − time_u) ≤ ts* then
                    Visited = Visited ∪ {a_v}
                    label(a_v, a_u)
                    enqueue(Q, a_v)
                end if
            end for
        end if
    end while
end for
return WS^r_{jl}
```

Our algorithm receives: (i) a view $\mathcal{N}$ modeling the reference scenario; (ii) a wallet w_j; (iii) a token t_l; (iv) a positive integer r denoting the length of the desired cross-blockchain walks; (v) a time interval ts^* indicating the maximum time interval between two swaps for considering them contiguous. It returns the set $WS^r_{j_l}$ of cross-blockchain walks of length r performed by w_j in the blockchains of $\mathcal{N}$ to exchange a token t_l.

The core of our algorithm is a Breadth-First traversal [47] of $\mathcal{N}$ through which all the walks of length r starting from a certain arc a_i of $\mathcal{N}$ are derived. Our algorithm uses the following data structures:

- A_{j_l}: it stores the set of arcs of $\mathcal{N}$ representing swaps made by the wallet w_j to transfer the token t_l across the blockchains of $\mathcal{N}$.
- Q: it is a queue of arcs of $\mathcal{N}$; the following primitives are provided for its management:
  - `enqueue()`, which inserts an arc at the end of Q;
  - `front()`, which removes and returns the first element of Q;
  - `isEmpty()`, which returns `true` if Q is empty, `false` otherwise.
- *Visited*: this is a set containing the arcs of $\mathcal{N}$ that have already been traversed.

- $WS^r_{j_l}$: it stores an ordered list of arcs representing a walk of length r performed by w_j to transfer a token t_l through the blockchains of $\mathcal{N}$.

Our algorithm also uses some support functions, namely:

- `selectArcs()`: it returns the set A_{j_l} of the arcs of $\mathcal{N}$ associated with the swaps performed by w_j to exchange the token t_l across the blockchains of $\mathcal{N}$.
- `selectNextArcs()`: it receives an arc a_u corresponding to a swap s_u and returns all the arcs of A_{j_l} corresponding to swaps performed after s_u.
- `label()`: it receives two arcs a_v and a_u, such that a_v is contiguous with a_u and $time_v > time_u$, and assigns a_u as the label of a_v.
- `reconstructWalk()`: it receives two arcs a_i and a_u from A_{j_l} and reconstructs the minimum walk from a_i to a_u. To perform its task, it uses the labels assigned to the arcs by the function `label()`.
- `length()`: it receives two arcs a_i and a_u and returns the length of the minimum walk between a_i and a_u; if $a_u = a_i$ it returns 1. To perform its task, it uses the labels assigned to the arcs by the function `label()`.

Our algorithm starts by setting $WS^r_{j_l}$ to empty and calling the function `selectArcs()` to construct the set A_{j_l} of arcs corresponding to the swaps made by w_j to transfer t_l.

For each arc a_i of A_{j_l}, it performs an iteration that aims to compute all cross-blockchain walks of length r starting from a_i. To achieve this goal, it first inserts a_i into the queue Q and the set *Visited*.

Then, until Q is empty, it begins an iteration consisting of the following steps. Initially, it picks up the first element a_u of Q and computes the distance between a_i and a_u by applying the function `length()`.

If this distance is equal to r, it calls the function `reconstructWalk()` to reconstruct the walk W from a_i to a_u. Once performed this task, it adds W to the set $WS^r_{j_l}$.

On the other hand, if the distance between a_i and a_u is less than r, it calls the function `selectNext- Arcs()`, which returns all the arcs of A_{j_l} whose timestamps follow that of a_u. For each arc a_v with such characteristics, if it has not already been visited, is contiguous to a_u (which implies that $bchain^{tgt}_u = bchain^{src}_v$) and the time interval between its timestamp and that of a_u is less than or equal to ts^*, then it is inserted into Q and *Visited*. Furthermore, a_u is assigned as the label of a_v.

When all iterations are complete, our algorithm returns the set $WS^r_{j_l}$. This set contains all the walks of length r present in $\mathcal{N}$, which were performed by w_j to exchange t_l. From graph theory, recall that a walk is open if the first and last nodes are different; otherwise, it is closed. A walk is simple if no node is crossed twice. A path is an open and simple walk. A cycle is a closed and simple walk [47]. Algorithm 1, as it is structured, searches only for simple walks. Consequently, the set $WS^r_{j_l}$ returned by it can be partitioned into two subsets, namely the set $\mathcal{P}^r_{j_l}$ of paths of length r performed by w_j to exchange t_l in the blockchains of $\mathcal{N}$, and the set $\mathcal{C}^r_{j_l}$ of cycles of length r performed by w_j to exchange t_l in the blockchains of $\mathcal{N}$. Clearly, $\mathcal{P}^r_{j_l} \cup \mathcal{C}^r_{j_l} = WS^r_{j_l}$ and $\mathcal{P}^r_{j_l} \cap \mathcal{C}^r_{j_l} = \emptyset$.

As specified above, we do not analyze the motivations that may lead a user to perform a succession of swaps that transfer a token across the blockchains of $\mathcal{N}$. In fact, this is beyond the scope of this paper. Here, we only want to emphasize that there are several motivations that can drive users to perform such an activity [42,45,48].

6. A Centrality Measure for Wallets

In a cross-blockchain ecosystem, it can be extremely useful to determine the so-called "power wallets", which are the most important wallets playing a leading role in the scenario. To reach this objective, a new centrality measure, tailored to the peculiarities of this ecosystem, could be helpful.

In particular, it should consider the swaps in which a wallet is involved and the tokens it has swapped. Furthermore, it should reflect the meso-structure of wallets, and thus the information stored in the view $\mathcal{H}$. For example, it should take into account the blockchains in which a certain wallet operates and the importance of that wallet in them. Having such guidelines in mind, we can define the wallet centrality wc_j of a wallet w_j. It can be represented by the following formula:

$$wc_j = \alpha \cdot sc(w_j) + \alpha' \cdot sdc(w_j) + \beta \cdot tc(w_j) + \beta' \cdot tdc(w_j) + \gamma \cdot mc(w_j) + \gamma' \cdot bmdc(w_j) + \gamma'' \cdot tmdc(w_j)$$

Here, the function $sc(\cdot)$ receives a wallet w_j and returns a value in the real interval $[0,1]$ that takes into account the set of swaps made by w_j compared to the ones carried out by the other wallets. A possible implementation of this function is the following:

$$sc(w_j) = \frac{|\mathcal{S}_j|}{\mathcal{S}_{max}}$$

where $\mathcal{S}_j$ has been introduced in Section 3 and $\mathcal{S}_{max}$ is the maximum number of swaps performed by a single wallet of $\mathcal{W}$.

The function $sdc(\cdot)$ receives a wallet w_j and returns a value in the real interval $[0,1]$ that takes into account the distribution of the swaps of w_j in the various blockchains. The rationale underlying $sdc(\cdot)$ is related to the fact that a wallet is more central the more it is able to uniformly distribute its activity across all the blockchains of the scenario. A possible implementation of $sdc(\cdot)$ can be obtained by computing the Herfindahl–Hirschman Index (HHI) [49] of the fractions of swaps that w_j made in the blockchains of $\mathcal{B}$. This index has been widely used in various fields of economics research for several decades. For example, it has been adopted to evaluate the concentration ratio in a certain market. In this case, it is defined as $HHI = \sum_{i=1}^{N} s_i^2$, where N is the number of firms operating in the market and s_i is the market share of the i^{th} firm. HHI ranges in the real interval $[\frac{1}{N}, 1]$; the higher HHI, the higher the concentration rate in that market. Following this reasoning, $sdc(\cdot)$ can be defined as follows:

$$sdc(w_j) = 1 - \sum_{k=1}^{m} \left[\frac{|\mathcal{BS}_{kj}^{tgt}|}{|\mathcal{S}_j|} \right]^2 = 1 - \frac{1}{|\mathcal{S}_j|^2} \cdot \sum_{k=1}^{m} |\mathcal{BS}_{kj}^{tgt}|^2$$

Here, $\mathcal{BS}_{kj}^{tgt}$ and $\mathcal{S}_j$ have been defined in Section 3.

The function $tc(\cdot)$ (resp., $mc(\cdot)$) receives a wallet w_j and returns a value in the real interval $[0,1]$ that takes into consideration the set of tokens (resp., the amount of money) exchanged by w_j compared to the ones carried out by the other wallets. Possible implementations of $tc(\cdot)$ and $mc(\cdot)$ are the following:

$$tc(w_j) = \frac{|\mathcal{T}_j|}{\mathcal{T}_{max}} \qquad mc(w_j) = \frac{\mathcal{M}_j}{\mathcal{M}_{max}}$$

where $\mathcal{T}_j$ and $\mathcal{M}_j$ have been introduced in Section 3 and $\mathcal{T}_{max}$ (resp., $\mathcal{M}_{max}$) is the maximum number of tokens (resp., the maximum amount of money) exchanged by a single wallet of $\mathcal{W}$.

The function $tdc(\cdot)$ receives a wallet w_j and returns a value in the real interval $\left[0, \frac{q-1}{q}\right]$ that takes into consideration the distribution of the swaps of w_j against the tokens of $\mathcal{T}$. A possible implementation of $tdc(\cdot)$ can be obtained by computing the Herfindahl–Hirschman Index of the fraction of swaps that w_j made for exchanging the tokens of $\mathcal{T}$.

$$tdc(w_j) = 1 - \sum_{l=1}^{q} \left[\frac{|\mathcal{TS}_{lj}|}{|\mathcal{S}_j|} \right]^2 = 1 - \frac{1}{|\mathcal{S}_j|^2} \cdot \sum_{l=1}^{q} |\mathcal{TS}_{lj}|^2$$

Here, $\mathcal{TS}_{lj}$ and $\mathcal{S}_j$ have been defined in Section 3.

The function $bmdc(\cdot)$ (resp., $tmdc(\cdot)$) receives a wallet w_j and returns a value in the real interval $\left[0, \frac{m-1}{m}\right]$ (resp., $\left[0, \frac{q-1}{q}\right]$) that takes into account the distribution of the amount of money $\mathcal{M}_j$ exchanged by w_j against the blockchains of $\mathcal{B}$ (resp., the tokens of $\mathcal{T}$). Again, a possible implementation of $bmdc(\cdot)$ (resp., $tmdc(\cdot)$) can be obtained by computing the HHI of the amounts of money that w_j exchanged through the blockchains of $\mathcal{B}$ (resp., the tokens of $\mathcal{T}$):

$$bmdc(w_j) = 1 - \sum_{k=1}^{m} \left[\frac{\mathcal{BM}_{kj}}{\mathcal{M}_j}\right]^2 = 1 - \frac{1}{\mathcal{M}_j^2} \cdot \sum_{k=1}^{m} \mathcal{BM}_{kj}^2$$

$$tmdc(w_j) = 1 - \sum_{l=1}^{q} \left[\frac{\mathcal{TM}_{lj}}{\mathcal{M}_j}\right]^2 = 1 - \frac{1}{\mathcal{M}_j^2} \cdot \sum_{l=1}^{q} \mathcal{TM}_{lj}^2$$

Here, $\mathcal{BM}_{kj}$, $\mathcal{TM}_{lj}$ and $\mathcal{M}_j$ have been defined in Section 3.

The reasoning behind the formula of $wc(\cdot)$ is as follows. A wallet is more central: (i) the greater the number of swaps it makes; (ii) the more such swaps are equally distributed across the blockchains; (iii) the greater the number of tokens it is able to use; (iv) the more its swaps are equally distributed across the various tokens; (v) the greater the amount of money it is able to transfer; (vi) the more such amount of money is distributed across the various blockchains, and (vii) across the various tokens.

In the definition of $wc(\cdot)$, the parameters α, α', β, β', γ, γ' and γ'' specify the weights assumed by the various functions that contribute to the wallet centrality. Their sum must be equal to 1. We obtained the optimal values of these parameters using an empirical approach. Specifically, we performed a grid-search [50] considering that each parameter ranges between 0 and 1. Starting from a combination of the values of parameters, in the grid-search, we computed the value of wc_j for all the wallets in our dataset and obtained a distribution of wallet centrality. The objective of the grid-search we performed was to maximize the distance between the minimum and the maximum values of the wallet centrality distribution. As a consequence, at the end of this task, we were provided with the optimal values of the parameters capable of maximizing the distance above. Specifically, these optimal values are the following: $\alpha = 0.30$, $\alpha' = 0.20$, $\beta = 0.15$, $\beta' = 0.10$, $\gamma = 0.15$, $\gamma' = 0.05$, $\gamma'' = 0.05$.

7. Experimental Campaign

In this section, we illustrate the experimental campaign we conducted to evaluate the approach and centrality measure proposed in this paper and to verify how they can support the extraction of insights and knowledge patterns related to the cross-blockchain scenario of our interest. To carry out these experiments, we used the model described in Section 3 and the framework introduced in Section 4. This section is organized as follows. First, we illustrate the dataset at the base of our experiments (Section 7.1). Then, we propose an Exploratory Data Analysis of it (Section 7.2). Afterwards, we describe our experiments aimed at evaluating our approach for the extraction of cross-blockchain paths and cycles (Section 7.3). Finally, we present our tests performed to evaluate our wallet centrality proposal (Section 7.4).

7.1. Dataset

To support our experiments, we built a dataset by extracting data from *Multichain*, an infrastructure (formerly known as *AnySwap*) developed for arbitrary cross-blockchain interactions. In particular, we implemented a web scraper in Python, which downloads all transaction data recorded by the Multichain explorer. The latter keeps track of all the transactions made on Multichain. Each transaction represents a swap of tokens between different blockchains. Since each transaction records data of a swap and each swap corresponds to a transaction, in the following we will use these two terms interchangeably.

We downloaded data on all transactions made between 13 February 2022 and 21 February 2022. We obtained data regarding 88,087 transactions between 26 different blockchains involving 442 different tokens. The Multichain explorer does not provide any token pricing data. Since this was essential to us, we had to compute it autonomously. For this purpose, we used historical price data provided by *CoinGecko* (www.coingecko.com (accessed on 29 January 2023)), a service that supplies financial information about cryptocurrencies.

We performed an ETL activity on the raw data scraped from the Multichain explorer. Specifically, we first removed all transactions with null or inconsistent values, such as missing value transferred, missing token and missing blockchain identifier. At the end of these activities we kept 79,523 transactions. After that, we decided to limit our analysis to the five blockchains with the highest number of transactions. This choice was motivated by the fact that many blockchains had a very low number of transactions, which would have made the analyses on them meaningless. As evidence of this, consider Table 1 showing the number of transactions for the top six blockchains in our dataset. As we can see from this table, the sixth blockchain (Moonbeam) has a number of transactions that is 19.92% of the fifth one (Ethereum) and 5.65% of the first one (Fantom). As further evidence in support of our decision, we highlight that, with this choice, the transactions kept in the dataset represent 91.31% of the original ones (Recall that a transaction involves exactly two blockchains in our ecosystem).

Table 1. First six blockchains ordered by the number of transactions involving them.

Blockchain	Number of Transactions
Fantom	51,955
Smart Chain	46,612
Avalanche	29,977
Polygon	17,576
Ethereum	14,736
Moonbeam	2936

At the end of this selection, our dataset consisted of 72,612 transactions related to 313 tokens and 5 blockchains, i.e., Fantom, Smart Chain, Avalanche, Polygon and Ethereum.

The reference dataset, the corresponding algorithms and other information about the experimental campaign are available at the following address https://github.com/daisy-univpm/Multichain-Analysis (accessed on 20 February 2023).

7.2. Exploratory Data Analysis

In this section, we present an Exploratory Data Analysis on the dataset described in Section 7.1. Specifically, we organize our analysis into two macro-areas, the former focused on the analysis of transactions (Section 7.2.1) and the latter concerning the analysis of blockchains (Section 7.2.2).

7.2.1. Analysis of Transactions

We start our Exploratory Data Analysis by providing an initial overview of our dataset. It is shown in Table 2. To quantify the real economic value of the swaps, we report the total volume of money (in USD) exchanged through them and the total fees (in USD) paid to make them. For each swap, the volume of money exchanged is obtained by multiplying the amount of tokens exchanged by its average daily value (expressed in US Dollars) and adding the fee required for the exchange. Note that even though the time span of our dataset is one week, the total volume of money exchanged by the wallets in the five blockchains is considerable.

Table 2. A general overview of our dataset.

Property	Description	Value
$\|\mathcal{S}\| = n$	Total number of swaps	72,612
$\|\mathcal{B}\| = m$	Total number of blockchains	5
$\|\mathcal{W}\| = p$	Total number of wallets	32,775
$\|\mathcal{T}\| = q$	Total number of tokens	313
$\sum_{s_i \in \mathcal{S}} val_i \cdot qnt_i + fee_i$	Total volume of swaps (in USD)	1,778,842,387.63
$\sum_{s_i \in \mathcal{S}} fee_i$	Total volume of fees (in USD)	860,150.52

In Figure 1, we report the variation of the number of swaps in our dataset over time, with a granularity of one hour. Note that the average number of swaps does not show significant changes over time, as highlighted by the orange line indicating the value of a 24 h rolling average. This suggests a regular usage of cross-blockchain services by wallets. We also observe the presence of some positive and negative peaks.

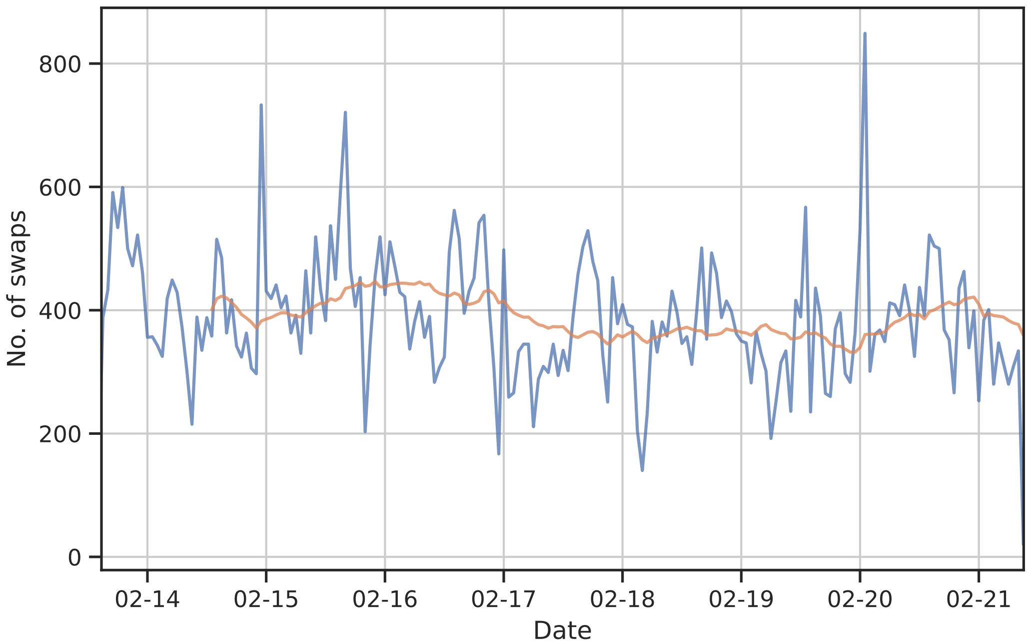

Figure 1. Variation of the number of swaps in our dataset over time.

In Figure 2, we show the variation of the volume of swaps in our dataset over time. Again, the granularity is one hour. We observe that, as in the case of the number of swaps reported in Figure 1, also for volumes there are no substantial variations over the whole observation period, as shown by the orange line reporting the 24 h rolling average. It is also possible to notice the presence of some peaks, three of which are very evident. In particular, the highest peak corresponds to a very consistent single swap from Smart Chain to Polygon involving the token BANANA. This peak suggests that high volumes are not always obtained by the sum of the (possibly low) volumes of a considerable number of swaps.

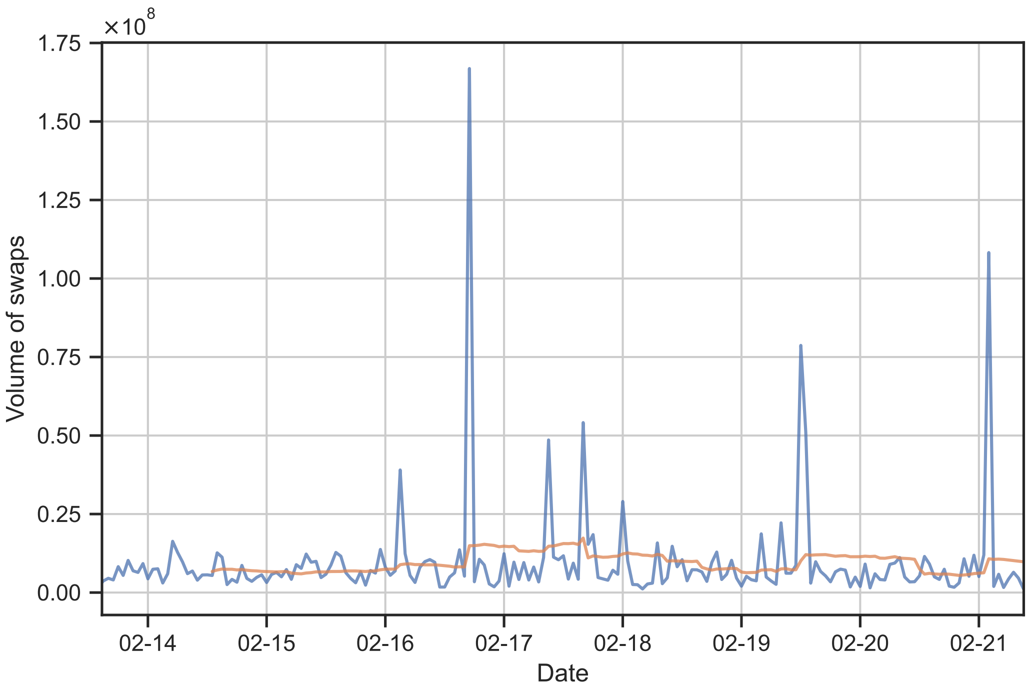

Figure 2. Variation of the volume (in USD) of swaps in our dataset over time.

Figure 3 shows the distribution of swaps against the number of wallets in log–log scale. From the analysis of this figure, we can conclude that, for the performed swaps, the wallets of our dataset can be partitioned into two categories. In fact, we have a few wallets (in particular, 47) that performed at least 50 swaps. These are present in the left part of the figure. On the other hand, in the right part of it, we have most of the wallets, each of which made a small number of swaps. This behavior highlights an interesting insight about wallets, namely their heterogeneity. In particular, it is presumable to think that some of the wallets that made many swaps are associated with smart contracts, which might encode a systematic behavior, rather than human users.

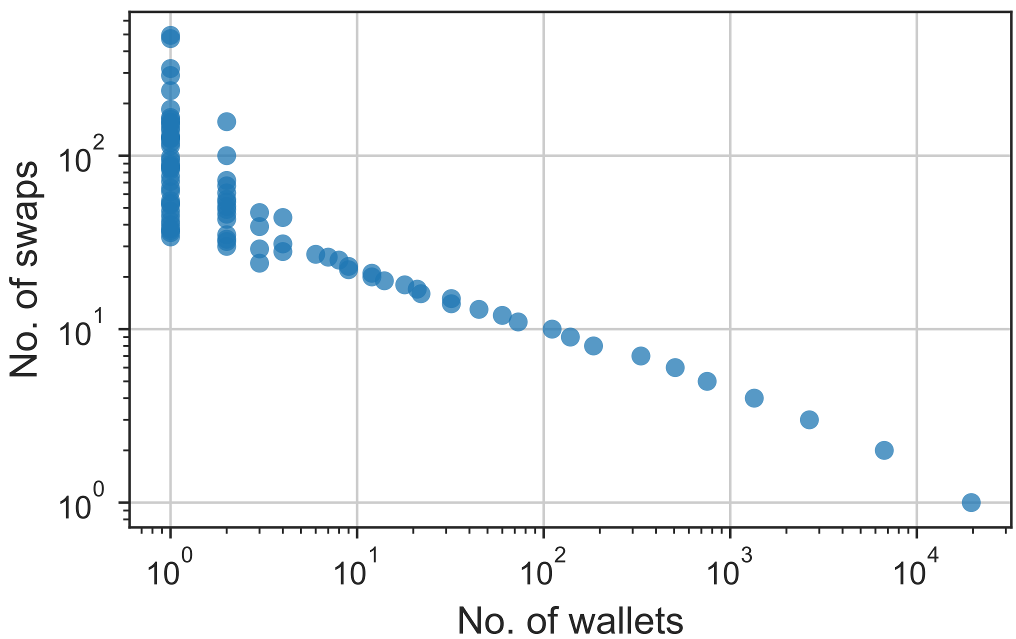

Figure 3. Distribution of swaps against the number of wallets.

Finally, Figure 4 shows the distribution of swaps against the number of tokens. Furthermore, this figure gives us some interesting insights. Indeed, we observe that there are many tokens (in particular, 54) involved in just one swap. Then, there is a significant number of tokens (in particular, 129) involved in less than 10 swaps. Furthermore, there are 130 tokens involved in at least 10 swaps. Finally, there is one token, namely USDC (USD Coin) occurring in 24,202 swaps. The high success of USDC could be motivated by the fact that USDC is a stable coin pegged to the United States Dollar, and it is claimed that each USDC is backed by fully reserved assets (https://www.centre.io/usdc-transparency (accessed on 20 February 2023)). This offers a certain degree of trust in the coin, which makes it widely used.

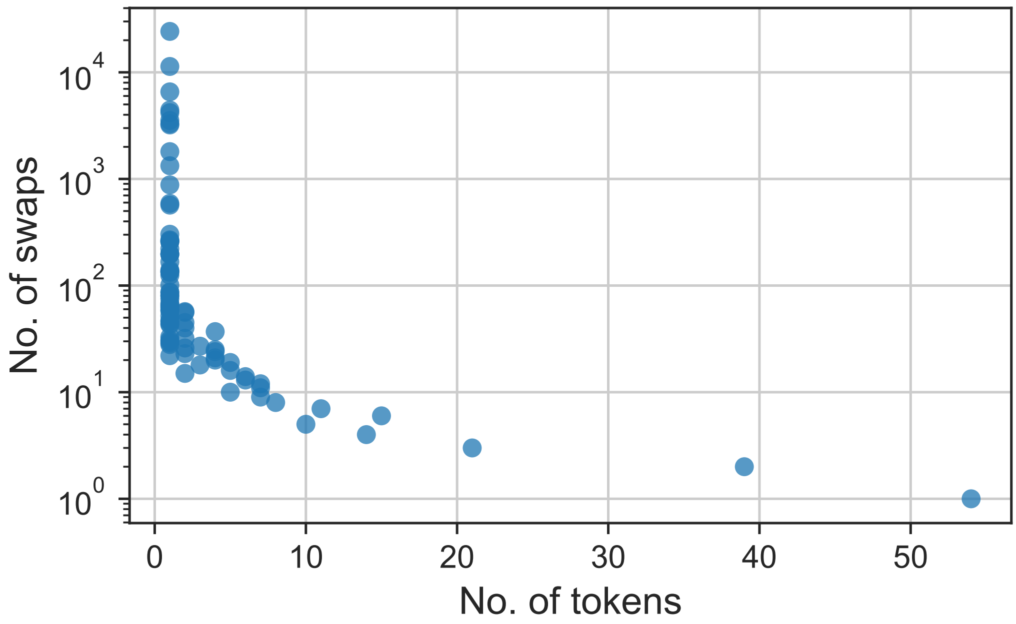

Figure 4. Distribution of swaps against the number of tokens.

7.2.2. Analysis of Blockchains

We start considering the number of swaps made to and from each of the five blockchains we are interested in; it is reported in Figure 5. This figure reveals that, as for this aspect, the blockchains show heterogeneous behaviors.

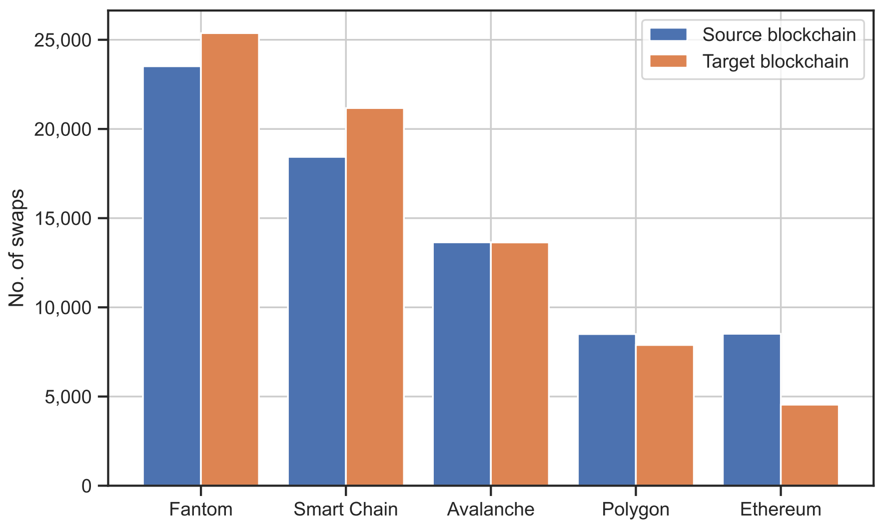

Figure 5. Number of incoming and outgoing swaps for each blockchain.

We hypothesized that the reason for this heterogeneity was primarily due to the fee charged for each swap. Recall that the fee for a swap does not depend on the money exchanged and is tied to the target blockchain. In order to verify our hypothesis, we computed the distribution of the values of the fees in USD for each blockchain. They are shown in Figure 6. The comparison between this figure and Figure 5 shows that our assumption was correct. In fact, Fantom is the blockchain with the lowest fees and, at the same time, one of the blockchains with the number of input swaps higher than the number of output ones. In contrast, Ethereum is the blockchain with the highest fees and, at the same time, the greatest imbalance between outgoing and incoming swaps.

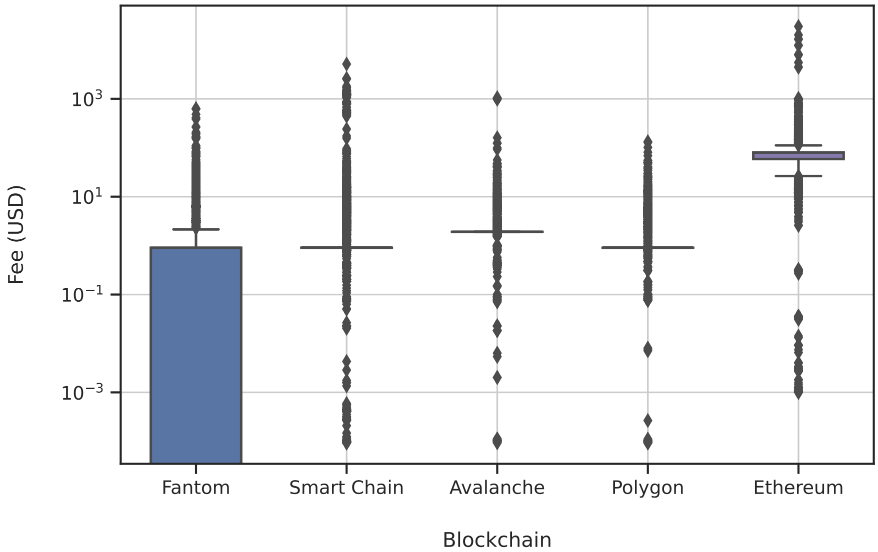

Figure 6. Distribution of fees (in USD) for each target blockchain.

Figure 7 shows the percentage of swaps occurring between each pair of blockchains. Figure 8 reports the total and average volume (expressed in USD) between each pair of blockchains. The comparison between this figure and Figure 7 is very interesting. In fact, the swaps having Fantom as target (i.e., 34.95% of swaps) exchanged a total volume of USD 462 million (i.e., 26.03% of total volumes) and an average volume per swap of USD 18 thousand. The swaps having Ethereum as target (i.e., 6.25% of swaps) exchanged a total volume of USD 414 million (i.e., 23.32% of total volumes) and an average volume per swap of USD 91 thousand. In the same wake as Ethereum we find Polygon. In fact, it represents the target for 10.87% of the total number of swaps but it exchanged a total volume of USD 355 million (i.e., 20.00% of total volume) and an average volume per swap of USD 45 thousand.

Figure 9 illustrates the total and average fees (expressed in USD) paid for swaps between each pair of blockchains. This figure shows that Ethereum is by far the target blockchain that charges the highest fees as both total values and average ones. The other target blockchains have very low values and are essentially comparable to each other. The only exceptions are Smart Chain, for transactions starting from Polygon, and partially Avalanche, for transactions starting from Fantom.

Finally, in Figure 10, for each pair of blockchains, we report the average time, expressed in seconds, needed to perform a swap (on the left) and the average number of confirmations needed to validate it (on the right). From the analysis of this figure, we can deduce some interesting insights. In particular, we first observe that the average time required to execute a swap having Avalanche as a target blockchain is extremely high and not comparable with that of the other blockchains. If, instead of considering average times, we take into account the number of confirmations, the scenario changes drastically. In fact, in this case, the target blockchain that requires more confirmations by far is Polygon, while those that require fewer confirmations are Smart Chain and Ethereum.

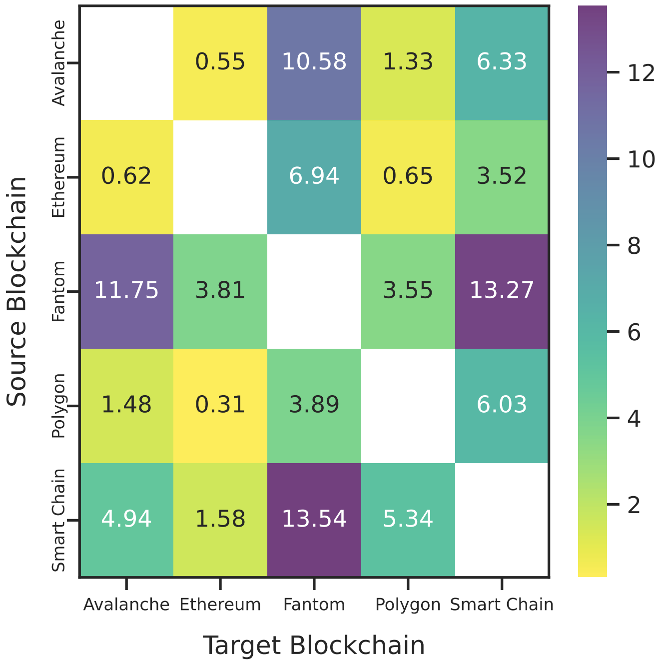

Figure 7. Percentage of swaps between blockchains.

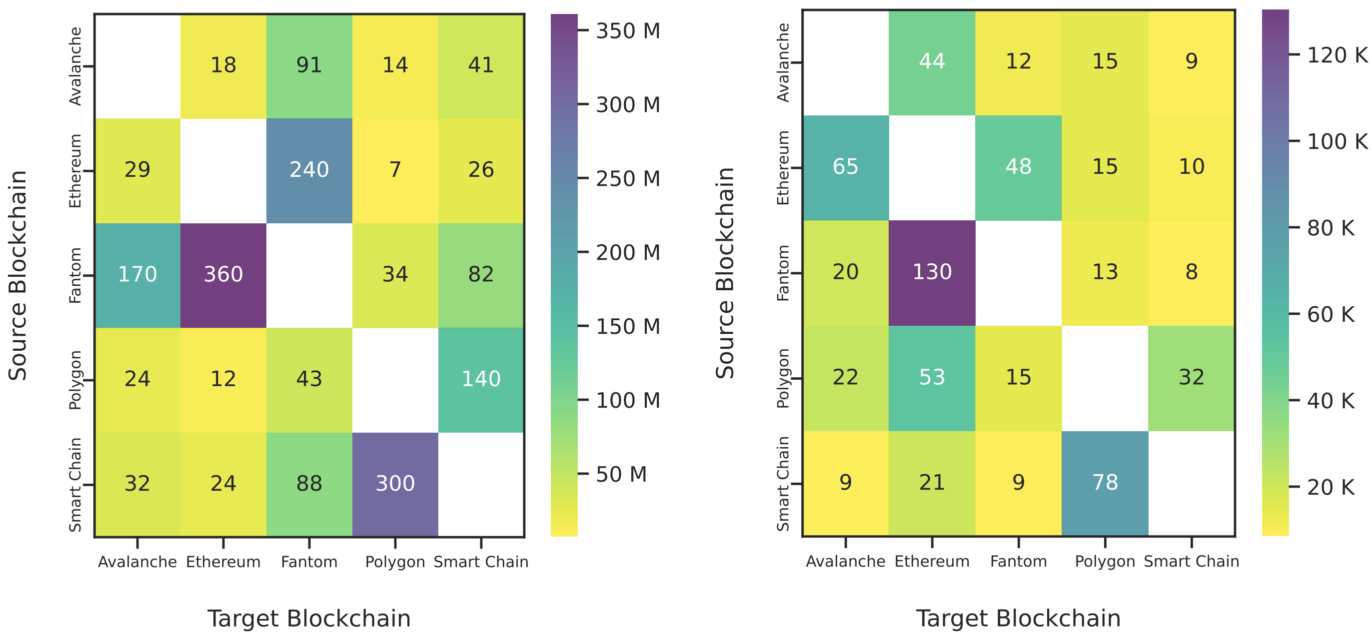

Figure 8. Total (on the left) and average (on the right) volumes (in USD) swapped between blockchains.

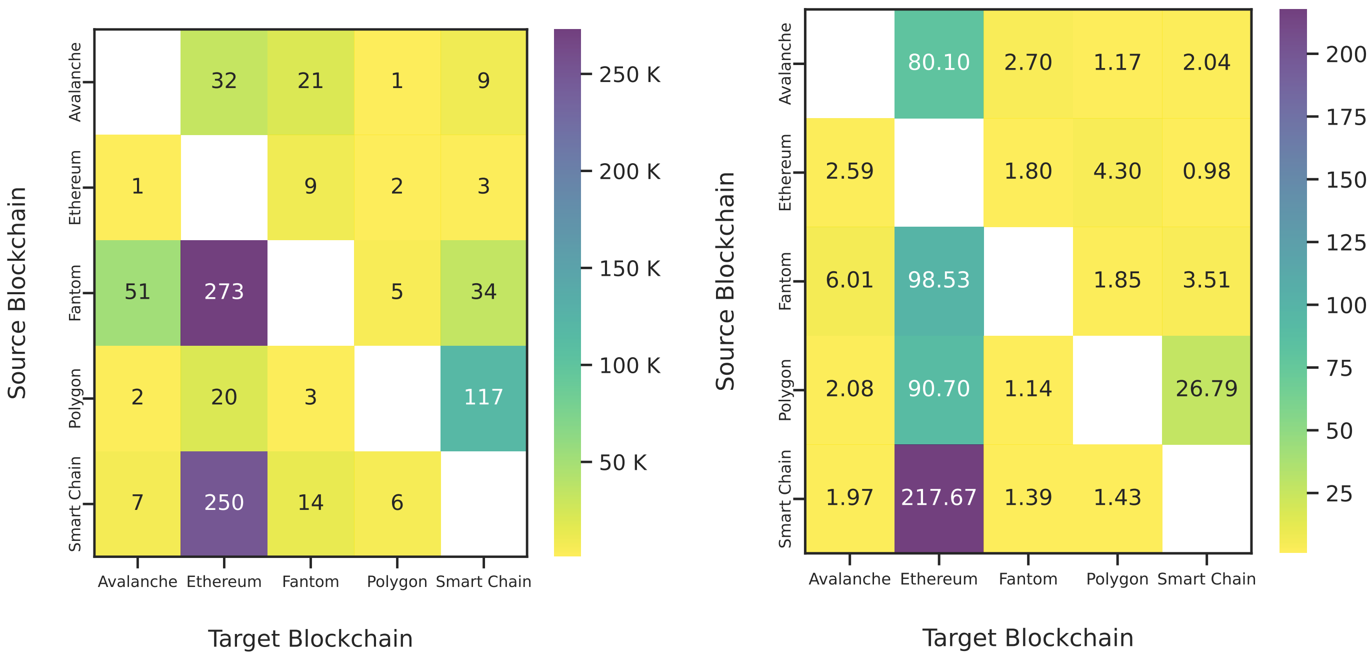

Figure 9. Total (on the left) and average (on the right) fees (in USD) paid for swaps between blockchains.

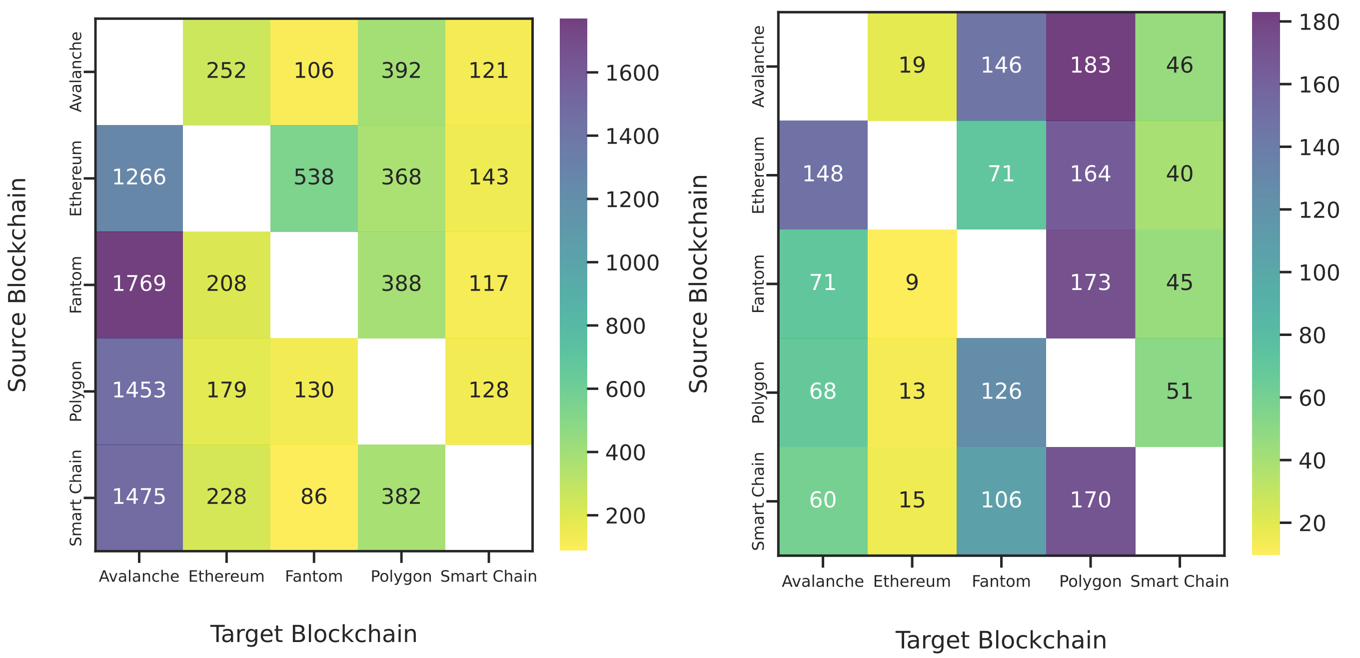

Figure 10. Average time, expressed in seconds, needed to create a swap (on the left) and average number of confirmations needed to validate it (on the right).

7.3. Mining Cross-Blockchain Paths and Cycles

In this section, we want to test our approach for path and cycle detection and extract knowledge about the behavior of wallets by analyzing the paths and cycles performed by them. This knowledge could concern the presence of paths or cycles typical of many wallets, which define some "modi operandi" of wallets, or, on the contrary, the presence of particular paths or cycles that can be seen as anomalies and allow us to discover unexpected behaviors by some wallets.

To detect cross-blockchain paths and cycles, we need to define the maximum time interval ts^* that can elapse between one swap and the next one in the path or cycle. In our tests, we have considered several values of ts^*. Specifically, we set ts^* to 0.1, 0.5, 1, 2, 6, 12 and 24 h. We selected this variety of values for ts^* to detect paths and cycles of different temporal lengths and to test the extent to which a high value of ts^* allows for the detection of more paths and cycles. In our analysis, besides considering different values of ts^*, we took paths and cycles of different lengths into account (By the length of a path or a cycle we mean the number of arcs composing it). Specifically, we set the desired length of paths and cycles equal to 2, 3, 4, 5, 6 and 7.

In Table 3 (resp., Table 4), we report the number of paths (resp., cycles) detected by our approach as the values of ts^* and the length of paths increase. A first consideration that emerges from the comparison of the two tables is that the number of cycles is much lower than the number of paths for the same parameter values. Moreover, we detected only cycles of length 2 and 3. The growth in the number of paths and cycles detected as ts^* increases is considerable when their length is small, while it is limited as we search for paths and cycles of greater length. This trend is evident in Figure 11, where we show the variation of the number of paths detected against the increase of ts^* for paths consisting of 2, 3 and 4 arcs.

Table 3. Number of paths detected by our approach against the variation of their length and the value of ts^*.

No. of Arcs | Path Length	0.1 h	0.5 h	1.0 h	2.0 h	6.0 h	12.0 h	18.0 h	24.0 h
2	222	482	622	758	1087	1366	1565	1802
3	18	53	81	107	202	322	437	545
4	1	11	15	24	47	97	140	174
5	0	1	2	4	5	32	48	66
6	0	0	0	0	0	6	8	27
7	0	0	0	0	0	5	7	12

Table 4. Number of cycles detected by our approach against the variation of their length and the value of ts^*.

No. of Arcs | Cycle Length	0.1 h	0.5 h	1.0 h	2.0 h	6.0 h	12.0 h	18.0 h	24.0 h
2	88	239	337	435	693	900	1050	1226
3	0	3	3	4	7	10	15	17

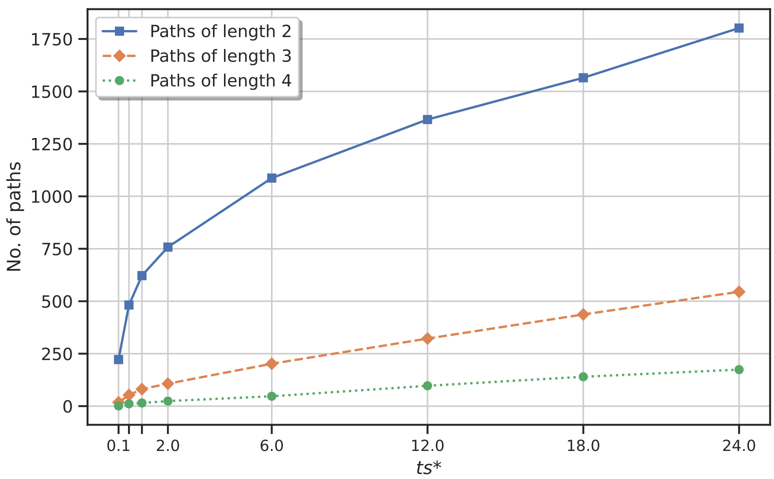

Figure 11. Variation of the number of paths against the increase of ts^*.

It is worth noting that there is a high number of cycles, and especially of paths, consisting of 2 swaps already with $ts^* = 0.1$ h. Given the very short maximum time interval between the two swaps of the path or cycle (i.e., 6 min), we could be led to think that there are cases in which the creation of a cycle or a path consisting of two swaps is a "modus operandi", perhaps encoded in a suitable smart contract. To better understand this phenomenon, we first considered the distribution of the paths of length 2 against ts^* when the values of the latter range from 0 to 360 s (i.e., 0.1 h). The results obtained are shown in Figure 12. From the analysis of this figure, we can observe the presence of a peak at 90 s. This distribution reinforces the idea that: (i) these paths are generally performed by smart contracts and not by humans; (ii) they encode very precise "modi operandi".

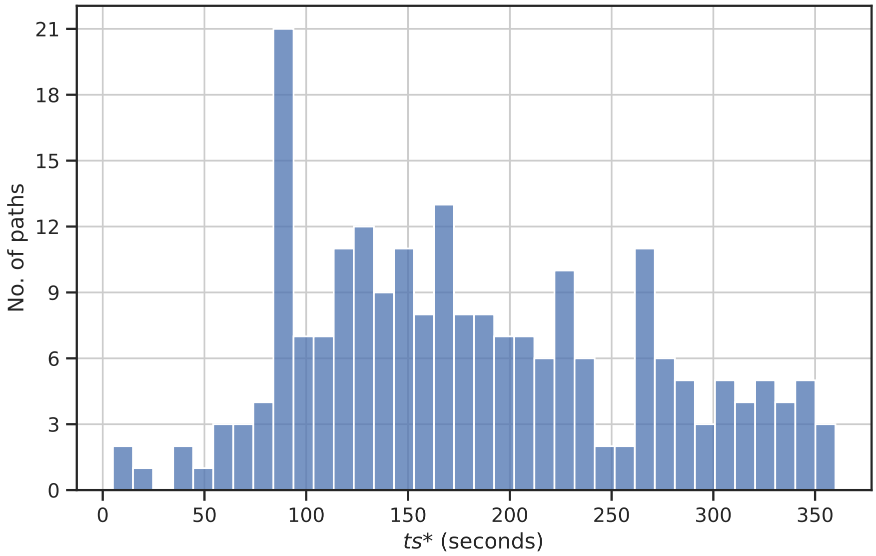

Figure 12. Distribution of the paths of length 2 against ts^* for values ranging from 0 to 0.1 h.

At this point, we carried out a manual investigation on each of the 222 paths involved. The first result we obtained was that 210 out of 222 paths were determined by smart contracts. The second result concerned the possibility of partitioning these paths into four distinct clusters, each of which actually encodes a "modus operandi". These clusters are the following:

- Cluster 1: The wallets belonging to this cluster tend to move much more money in the first swap of the path than in the second one. The overall amount of money moved in the swaps of these paths is much higher than that moved in the paths of the other clusters. The times between two swaps are very low (less than one minute), so we can think that the behavior of these wallets is driven by a very solid strategy. The execution time of the swaps is very low, which means that the wallets performing them are willing to pay very high fees. A total of 32 of the 222 wallets under examination belong to this cluster. Using a technical term typical for investors, we can call these wallets "whales".
- Cluster 2: The wallets belonging to this cluster tend to move the same amount of money in the two swaps. The overall amount of money moved in the swaps of these paths is much smaller than that moved in the paths of Cluster 1 but higher than that moved in the paths of Cluster 3. The times between the two swaps are medium (between 2 and 4 min). The execution time of swaps is average, which leads us to think that these wallets tried to obtain fees allowing the swap execution in a reasonably short time without paying an excessive cost. A total of 111 of the 222 wallets under examination belong to this cluster. The profile described for the wallets performing them corresponds to that of "experienced users".
- Cluster 3: The wallets belonging to this cluster tend to move significantly more money in the second swap than in the first one. The overall amount of money moved in the swaps of these paths is slightly less than that moved in the paths of Cluster 2

and much less than that moved in the paths of Cluster 1. The times between the two swaps are long (between 4 and 6 min) (Clearly, the term "long" is to be intended with reference to the fact that we are examining paths with very low values of ts^*). This makes us think that the first swap is only functional for the second one. The execution time of the swaps is a bit higher than the previous two clusters, leading us to think that these wallets are trying to pay low fees. A total of 67 out of the 222 wallets under examination belong to this cluster. The profile described for these wallets corresponds to that of "normal users".

- Cluster 4: The wallets belonging to this cluster move much less money in a path than the wallets in the previous three clusters. The times between the two swaps are the longest ones (between 5 and 6 min) among the paths under consideration. The execution time of the swaps is much higher than in the previous three clusters, which leads us to think that these wallets are willing to pay very low fees. 12 of the 222 wallets into examination belong to this cluster. The swaps of these paths are executed by humans and not by smart contracts. All these features lead us to think that these wallets correspond to "inexperienced users".

We repeated this process in the case of cycles and obtained that 84 out of the 88 cycles were determined by smart contracts. Moreover, by partitioning these cycles we obtained the same distinct clusters we had obtained for paths.

Another interesting aspect to analyze is the relative position of the blockchains in the paths. In particular, we want to analyze which blockchains the paths originate from or terminate to the most. In Figure 13 (resp., Figure 14), we show the percentage of paths of length 2 (on the left) and 3 (on the right) originating from (resp., terminating to) each blockchain. In these figures, for each blockchain, we report three percentages relative to a value of ts^* equal to 0.1, 0.5 and 1.0 h, respectively.

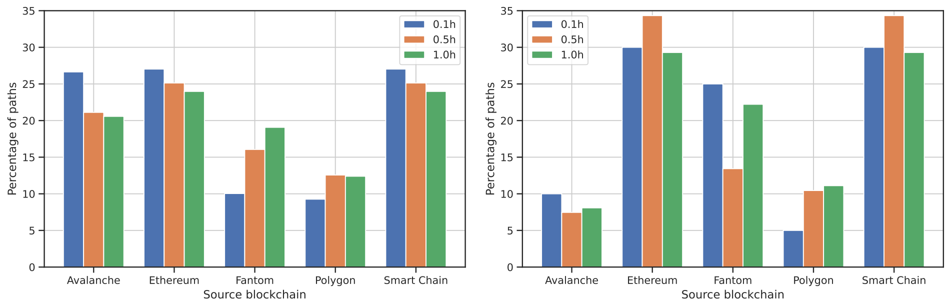

Figure 13. Percentage of paths of length 2 (on the left) and 3 (on the right) originating from a specific blockchain.

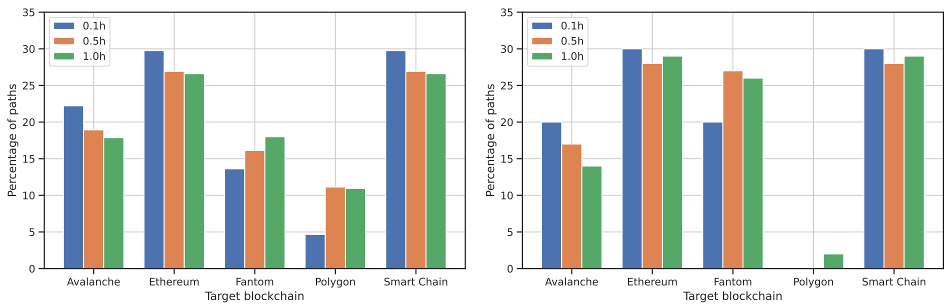

Figure 14. Percentage of paths of length 2 (on the left) and 3 (on the right) terminating to a specific blockchain.

Figure 13 allows us to derive several interesting insights. First, we can observe that in both cases Polygon is under-represented, suggesting that there is a negligible number of paths starting from it. One particular behavior is that of Avalanche. Indeed, it is the starting blockchain for many paths of length 2 but for few paths of length 3. Figure 13 also shows that, as the value of ts^* increases, the percentage of paths of length 2 decreases for Ethereum, Avalanche and Smart Chain. This does not happen when we consider the paths of length 3 neither for these blockchains nor for any other. Let us now consider Figure 14. We can see that the trends characterizing the various blockchains are similar to those in Figure 13. A noticeable insight is the very small or null representativeness of Polygon as the blockchain to which a path terminates. In general, the discovered paths have a tendency to start or end mainly in two blockchains, namely Ethereum and Smart Chain.

After examining the distribution of paths against blockchains, we performed the same task for cycles of length 2. In Figure 15, we report the results obtained. From the analysis of this figure, we can see that Ethereum and Smart Chain are the blockchains where cycles are most involved. Comparing the results with those of Figures 13 and 14 on the left, regarding paths of length 2, we can observe a lower presence of Avalanche and a higher presence of Fantom for cycles than for paths.

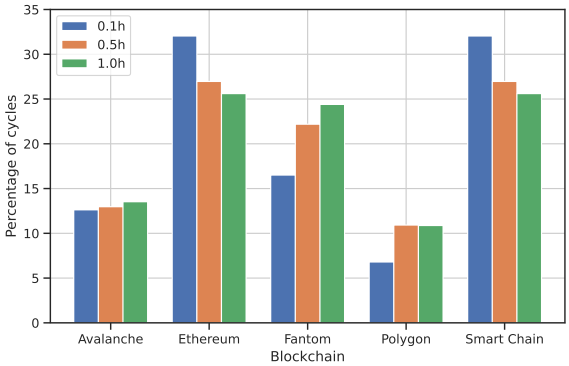

Figure 15. Percentage of cycles of length 2.

To date, we have analyzed the main paths and cycles with respect to blockchains. However, there is another viewpoint, i.e., the token viewpoint, orthogonal to the previous one, which is worth investigating. Regarding this, in Table 5 we present the top 10 most frequent tokens in the paths of length 2, along with the corresponding number of occurrences, when ts^* = 0.1 h. We do not consider other paths and cycles because the number of their occurrences is too low and the corresponding results would be insignificant. Instead, in Table 6 (resp., Table 7), we present the top 10 most frequent tokens in paths of length 2, 3 and 4 and in cycles of length 2, along with the corresponding number of occurrences, when ts^* = 6 h (resp., 24 h).

Table 5. Top 10 most frequent tokens in paths of length 2 (ts^* = 0.1 h).

Paths of Length 2	
USDC	79
USDT	29
MIM	17
DAI	12
BNB	10
FABRIC	9
ETH	6
AVAX	5
BIFI	5
ALPACA	3

Table 6. Top 10 most frequent tokens in paths of length 2, 3 and 4 and in cycles of length 2 (ts^* = 6 h).

Paths of Length 2		Paths of Length 3		Paths of Length 4		Cycles of Length 2	
USDC	432	USDC	75	USDC	21	USDC	220
USDT	176	USDT	25	USDT	5	USDT	97
MIM	88	MIM	14	BNB	4	MIM	62
BNB	60	BNB	13	DAI	3	BNB	51
DAI	56	DAI	10	MIM	3	AVAX	49
AVAX	49	ALPACA	8	ALPACA	2	DAI	30
ETH	29	ETH	8	DERC	2	ALPACA	16
ALPACA	16	AVAX	6	ETH	2	ETH	15
BIFI	11	BIFI	4	BANANA	1	DERC	9
DERC	10	DERC	4	MATIC	1	FABRIC	9

Table 7. Top 10 most frequent tokens in paths of length 2, 3 and 4 and in cycles of length 2 (ts^* = 24 h).

Paths of Length 2		Paths of Length 3		Paths of Length 4		Cycles of Length 2	
USDC	715	USDC	194	USDC	67	USDC	392
USDT	298	USDT	87	USDT	29	USDT	195
MIM	147	MIM	46	MIM	18	MIM	110
BNB	111	BNB	34	DAI	14	BNB	97
AVAX	88	AVAX	25	BNB	7	AVAX	87
DAI	82	DAI	25	FABRIC	7	DAI	45
ETH	42	ETH	12	DERC	5	ETH	23
ALPACA	20	ALPACA	10	YOSHI	4	ALPACA	20
BIFI	16	FABRIC	10	ETH	3	FABRIC	14
DERC	16	DERC	8	AVAX	3	DERC	14

From the analysis of these tables, we can see that the distribution of paths and cycles against tokens follows a power law. In particular, the most frequent token is always USDC (USD Coin). The second token is always USDT (Tether) and has a number of occurrences that is generally less than half of USDC. The third token is almost always MIM (Magic Internet Money), whose number of occurrences is generally a little more than half that of USDT. It is worth pointing out that all of these tokens are pegged to the United States Dollar, and this is probably one of the reasons for their success. Another interesting insight that can be gained by examining Tables 5–7 is the strong overlap between the top 10 most frequent tokens in the various cases.

7.4. Wallet Centrality

In Section 6, we proposed a new centrality measure specifically suited for evaluating the importance of a wallet in a cross-blockchain scenario. In this section, we want to evaluate the goodness of this centrality measure, the role played by the various components, and focus on the wallets with the highest centrality value.

We begin our study by considering the distribution of wallets against wallet centrality. It is shown in Figure 16. In this distribution, we computed the centrality of each wallet using the optimal values of the weights obtained experimentally and reported in Section 6. From the analysis of this figure we can observe that the wallet centrality is really able to assign different scores to the various wallets. In other words, it is able to partition the wallets into many subsets characterized by different values of centrality. In particular, looking at Figure 16, we can see the presence of two macro-sets of wallets. Both of them have approximately a Gaussian distribution. However, the distribution on the left is much wider and lower than the one on the right. Moreover, we can observe the presence of some outliers associated with values of wallet centrality close to 0.5.

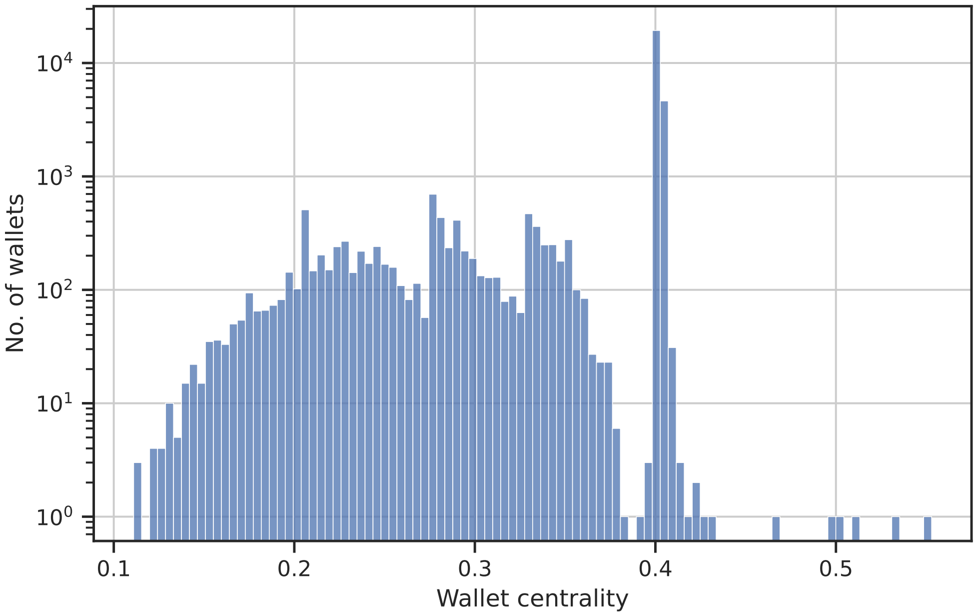

Figure 16. Distribution of wallets against wallet centrality.

It should be noted that, overall, the wallet centrality assumes medium-low values since there are no wallets with centrality values close to 1. This leads us to hypothesize that the various parameters that constitute the wallet centrality are actually able to capture different aspects of the importance of a wallet in a cross-blockchain context.

To better understand the role of the various parameters involved in the wallet centrality, we decided to compute the distribution of wallets against them. Specifically, in Figure 17, we illustrate the distribution of wallets against the components of the formula of wc_j taking into account only the swaps made and their equidistribution with respect to blockchains. It can be obtained by considering the general formula of wc_j and setting $\alpha = 0.60$, $\alpha' = 0.40$, and all the other weights equal to 0.

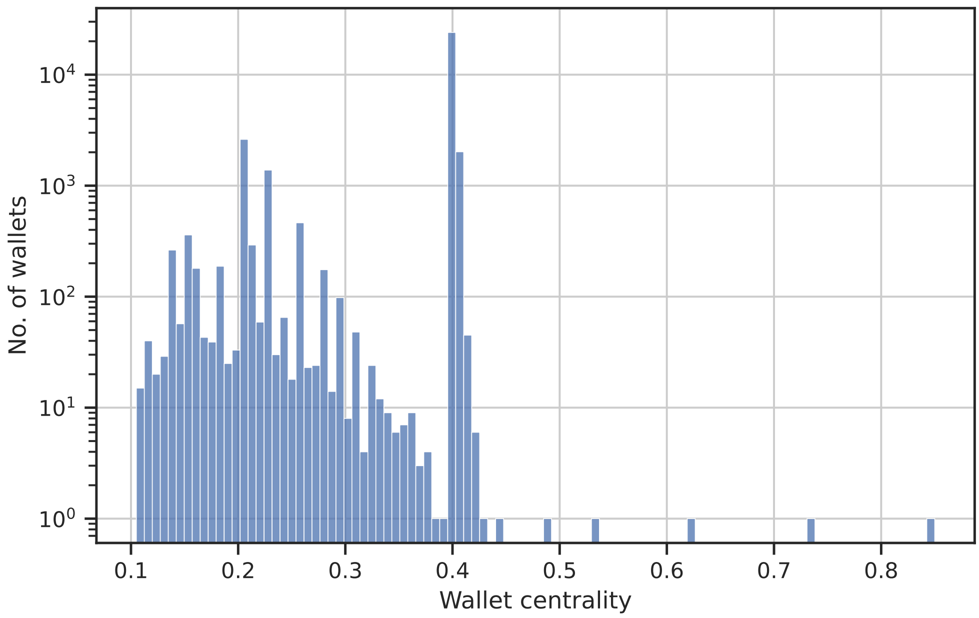

Figure 17. Distribution of wallets against the swaps made and their equidistribution with respect to blockchains.

Similarly, in Figure 18, we report the distribution of wallets against the tokens used and the equidistribution of the corresponding swaps with respect to tokens. It can be obtained by considering the general formula of wc_j and setting $\beta = 0.55$, $\beta' = 0.45$, and all the other weights equal to 0.

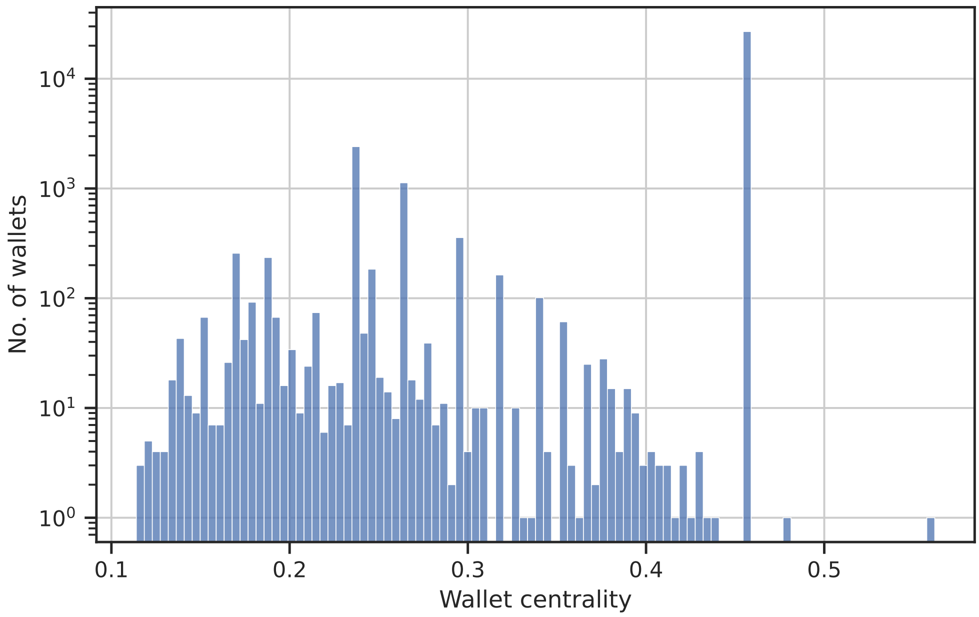

Figure 18. Distribution of wallets against the tokens used and the equidistribution of the corresponding swaps with respect to them.

Finally, in Figure 19, we show the distribution of wallets against the exchanged money and its equidistribution with respect to blockchains and tokens. It can be obtained by considering the general formula of wc_j and setting $\gamma = 0.55$, $\gamma' = \gamma'' = 0.225$, and all the other weights equal to 0.

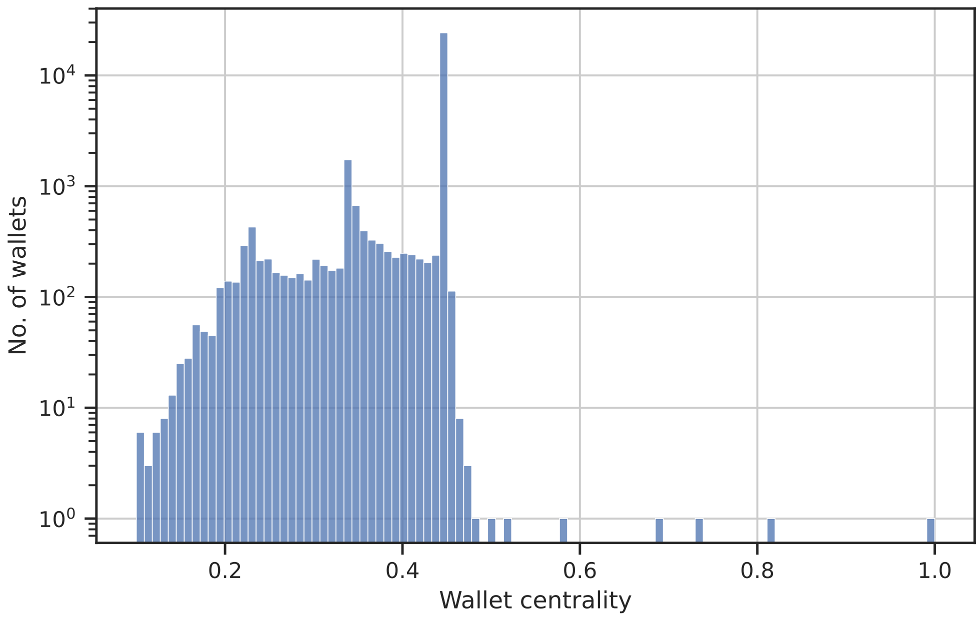

Figure 19. Distribution of wallets against the exchanged money and its equidistribution with respect to blockchains and tokens.

The distribution in Figure 17 is the most similar to the overall distribution in Figure 16. The distribution in Figure 18 is quite similar to that of Figure 16, but with some differences. Finally, the distribution of Figure 19 differs significantly from the others. It is much more irregular and, at a certain point, decays rapidly. It has outliers too, some of which have particularly high values. Considering Figures 17–19 on the one hand, and Figure 16 on the other hand, it is quite evident that the values of the overall wallet centrality actually represent an average of the three components shown in Figures 17–19. In particular, the overall distribution does not distort the individual ones but, in a sense, harmonizes them.

After having studied the role played by the three components related to swaps, tokens and money in the determination of the overall wallet centrality, now we want to investigate other components, orthogonal to the previous ones. In fact, if we consider the formula of wc_j from another point of view, we can see that it consists of a component related to quantity (taking the number of swaps, the number of tokens and the amount of money into consideration) and another one related to equidistribution. In the following, we aim at investigating how these two components contribute to the overall wallet centrality.

In Figure 20, we show the distribution of wallets against the component of the wallet centrality related to quantity. It can be obtained by considering the general formula of wc_j and setting $\alpha = 0.34$, $\beta = 0.33$, $\gamma = 0.33$, $\alpha' = \beta' = \gamma' = \gamma'' = 0$. Instead, in Figure 21, we show the distribution of wallets against the component of the wallet centrality related to equidistribution. It can be obtained considering the general formula of wc_j and setting $\alpha' = 0.33$, $\beta' = 0.33$, $\gamma' = \gamma'' = 0.17$, $\alpha = \beta = \gamma = 0$.

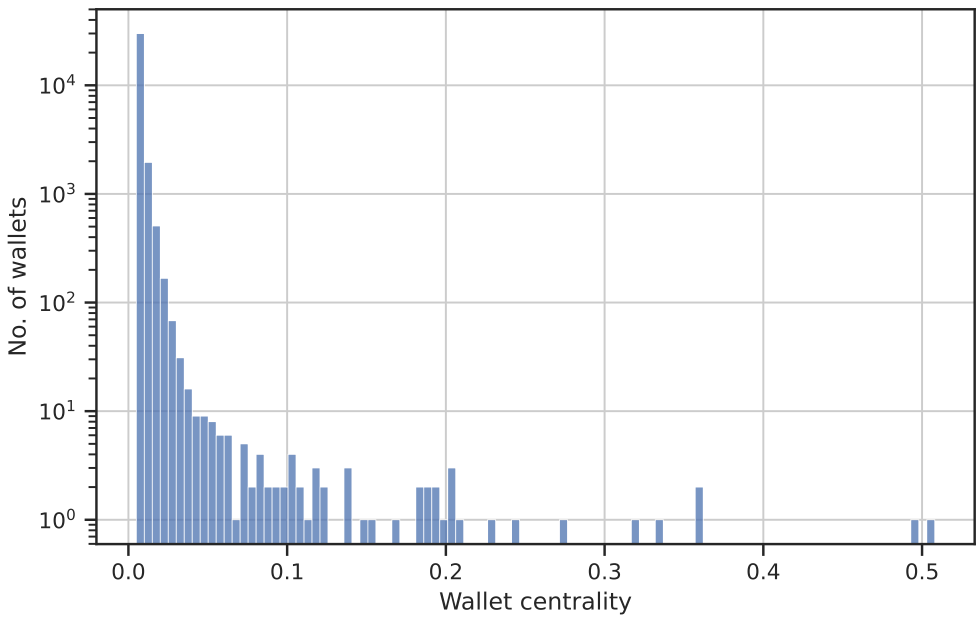

Figure 20. Distribution of wallets against the component of the wallet centrality related to quantity.

The distribution in Figure 20 is very different from the overall distribution in Figure 16. In fact, it is very similar to a power law distribution. Moreover, the distribution of Figure 21 is different from that of Figure 16 and totally different from that of Figure 20. Comparing Figures 20 and 21 on the one hand, and Figure 16 on the other hand, it can be seen that, also in this case, the values of the overall wallet centrality represent an average of the values of the component related to the quantity and that related to the equidistribution. However, in this case, since the two partial distributions are very different from each other, the overall one "distorts" them and acts, in a sense, as a smoother. In fact, the two partial distributions are very extreme, although for opposite reasons. Instead, the overall distribution managed to dampen these extremes by making them smooth each other.

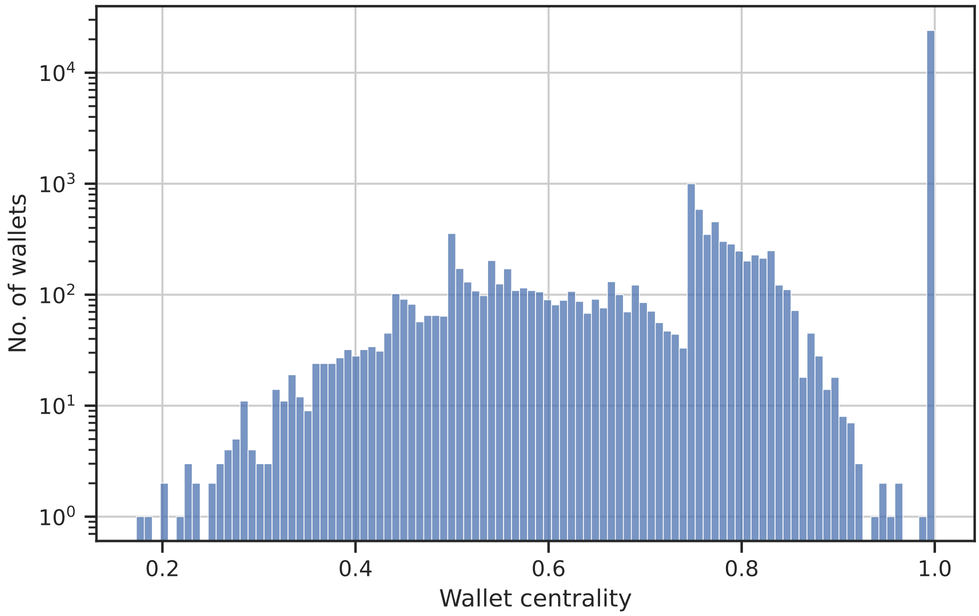

Figure 21. Distribution of wallets against the component of the wallet centrality related to equidistribution.

After analyzing the role played by the various components of the wallet centrality, we want to focus on the wallets with the highest values of this measure. More specifically, we want to analyze the top wallets in terms of both overall centrality and each of its components identified above. Our goal is to test whether these sets of top wallets are disjoint or overlapping. In the following, we use the following notations to denote the top wallets: (i) τ_ω indicates the top wallets obtained considering the overall wallet centrality; (ii) τ_α denotes the top wallets obtained considering only the swaps performed: (iii) τ_β indicates the top wallets obtained considering only the token adopted, i.e., the component associated with the distribution in Figure 18; (iv) τ_γ denotes the top wallets obtained considering only the money exchanged, i.e., the component associated with the distribution in Figure 19; (v) τ_δ indicates the top wallets obtained considering only the component related to quantity; (vi) τ_ϵ denotes the top wallets obtained considering only the component related to equidistribution.

In our analysis, we considered the top 100, top 1000, top 5000 and top 10,000 wallets for each set defined above and performed a series of intersections between them. The values obtained are shown in Table 8.

Table 8. Intersections between different sets of top wallets.

	$\tau_\alpha \cap \tau_\beta$	$\tau_\alpha \cap \tau_\gamma$	$\tau_\alpha \cap \tau_\omega$	$\tau_\beta \cap \tau_\gamma$	$\tau_\beta \cap \tau_\omega$	$\tau_\gamma \cap \tau_\omega$	$\tau_\delta \cap \tau_\epsilon$	$\tau_\delta \cap \tau_\omega$	$\tau_\epsilon \cap \tau_\omega$
Top 100	2	4	64	0	1	21	1	10	1
Top 1000	14	70	556	40	25	159	1	19	24
Top 5000	652	1127	3372	997	870	1950	8	84	978
Top 10,000	3823	3970	5461	3793	3672	8149	700	2173	3981

From the analysis of this table, we can derive some interesting insights. In fact, if we consider the sets of the top 100 wallets, we can see that the pairs (τ_α, τ_β), (τ_α, τ_γ) and (τ_β, τ_γ) have negligible intersections. This confirms that the various components of the wallet centrality are able to capture different aspects of wallet behavior. A similar reasoning is valid for the pair (τ_δ, τ_ϵ). This confirms what we had already observed from examining the corresponding distributions. Instead, looking at the intersections of each of these sets with τ_ω, associated with the overall wallet centrality, we note that there is an overlap between τ_α and τ_ω. This trend had already been observed by comparing the distributions in Figures 16 and 17. In contrast, the overlap, albeit partial, between τ_γ and τ_ω is surprising. In fact, the distributions in Figures 16 and 19 are significantly different. Considering the top 1000, we can see that most of the trends observed for the top 100 are confirmed.

Instead, these trends begin to change when we consider the top 5000 (observe, for instance, the much greater overlap of the sets τ_ϵ and τ_ω, compared to what happened for the top 100 and top 1000) and consolidate towards an equilibrium completely different from the previous one when we consider the top 10,000. In the latter scenario, all pairs of sets, except for (τ_δ, τ_ϵ), have much more significant overlaps than what happened for the top 100 and top 1000. By far the highest overlap is observed between τ_γ and τ_ω, which overlap by more than 80%, with an enormous growth in this phenomenon compared to the cases relative to the top 100, top 1000 and top 5000. τ_α and τ_ω continue to overlap significantly (more than 54%) but the trend of this overlap has a linear growth compared to the previous tops. The only exception is represented by the overlaps involving the set τ_δ. These remain much smaller than the others, even if, when passing from the top 5000 to the top 10,000, we observe an enormous growth of them (equal to about 2487% in the case of (τ_δ, τ_ω) and about 8650% in the case of (τ_δ,τ_ϵ)).

8. Discussion

In this section, we present a discussion on the approach proposed and the activities we conducted to evaluate it. In particular, we outline the main contributions of this paper with respect to the existing literature, as well as the implications, applications and limitations of our approach. The section terminates with a look at possible future developments of the research proposed in this paper.

8.1. Contributions

The main contribution of our paper is the idea of applying Social Network Analysis to represent and handle a cross-blockchain ecosystem. This contribution is then declined in several sub-contributions. The first one concerns the definition of a model for representing a cross-blockchain ecosystem. This model provides a clear and concise representation of the various actors involved in the cross-blockchain scenario, including blockchains, wallets and transactions for transferring money from one blockchain to another. It represents a crucial first step in analyzing and understanding the complex interactions that can occur in a cross-blockchain ecosystem.

The second sub-contribution concerns a social network-based framework for managing a cross-blockchain scenario. This framework allows for multi-dimensional and multi-view modeling and management of the cross-blockchain ecosystem, enabling the application of a wide range of concepts and solutions derived from Social Network Analysis research. Due to this feature, it greatly enhances the ability of researchers and practitioners to analyze and understand a cross-blockchain scenario.

The third sub-contribution regards the application of our framework to extract walks, paths and cycles from a cross-blockchain ecosystem. This in turns enables the extraction of insights and knowledge patterns related to the behavior of various categories of wallets in the cross-blockchain scenario. Such information can be used to better understand and improve the cross-blockchain ecosystem, as well as to look for any potentially malicious or fraudulent behavior.

The fourth sub-contribution is a new centrality measure to identify the most important wallets in the cross-blockchain ecosystem. This measure takes into account multiple factors, including the number of transfers made, the number of blockchains involved, and the amount of money transferred. In this way, it provides a comprehensive view of the centrality of a wallet in a cross-blockchain scenario.

8.2. Implications

The approach presented in this paper has several implications for the cross-blockchain ecosystem. First, it provides a new and innovative way to analyze and understand the interactions and relationships among the various actors involved in a cross-blockchain scenario. This can lead to a deeper understanding of the ecosystem and the identification of potential areas for its improvement.

In addition, our approach can improve the understanding of cross-blockchain transactions by providing a more comprehensive view of the relationships between wallets and blockchains.

Finally, it can be used to look for potentially malicious and fraudulent behaviors, helping to improve the overall security of the cross-blockchain ecosystem.

8.3. Applications

The approach proposed in this paper has several possible applications. One of them is improving the security of the cross-blockchain ecosystem. In this context, it can be used to search for potentially malicious or fraudulent behavior. In fact, the presence of loops or paths in the ecosystem can be a red flag for the presence of fraud or money washing.

Another important application is to make swapping recommendations for wallets in the cross-blockchain ecosystem. In fact, our approach can suggest the optimal transfers for a given wallet. To reach this goal, it uses the wallet us past history and current situation in the ecosystem, as well as factors like the wallet us current balance, the fees involved in the transfer, and the time required to complete the transfer. This information can help wallets make informed decisions about transfers and optimize their use of the cross-blockchain ecosystem.

Furthermore, our approach can also be extended to perform smart contract profiling for DeFi. This involves analyzing the transactions in which a smart contract is involved and the effects it produces, so as to provide a more complete understanding of its impact on the cross-blockchain ecosystem. This information can be employed to improve the design and implementation of smart contracts, ensuring that they are effective and efficient in the cross-blockchain scenario. This in turn can lead to a more efficient and effective cross-blockchain ecosystem, in which smart contracts play a key role in facilitating cross-blockchain interoperability.

8.4. Limitations

We think it is also interesting to highlight some limitations of our approach, which may provide insights for future developments of this research. First, our approach is limited by the quality and availability of data, because the accuracy of its results depends on the accuracy and completeness of the data employed to derive them.

Moreover, our approach has been defined taking into consideration the current state of the cross-blockchain ecosystem. However, the latter is rapidly evolving and changing over time.

Furthermore, our approach currently uses the structure of the cross-blockchain ecosystem and the behavior of wallets for deriving interesting knowledge patterns. However, there may be other important factors that could influence the knowledge learned and lead to additional insights and knowledge about the cross-blockchain ecosystem. Consider, for example, the contribution that could be obtained by analyzing the role of intermediaries, the impact of regulations and the influence of external events. These are all factors that should be investigated in order to have a broader and deeper understanding of the reference scenario.

8.5. Future Work

In the future, we plan to continue this research along several directions. First, we aim to identify other Social Network Analysis concepts and solutions that can be applied in a cross-blockchain ecosystem. Then, we plan to integrate all the results found in this paper, and others we will possibly find in the future, to build the platform mentioned above. Afterwards, we would like to extend our approach so that it can perform smart contract profiling for DeFi. In fact, a smart contract can be easily analyzed a priori, i.e., at the design level. However, such a static analysis does not allow us to determine its impact on the cross-blockchain ecosystem at the moment it is actually used. Indeed, the latter is a dynamic analysis that can only be carried out by examining the transactions in which the smart contract is involved and the effects produced. We believe that Social Network Analysis can be a valuable tool to do this and, in the future, we would like to demonstrate this conjecture by defining an approach to address this issue. Finally, we think of extending the results of this paper to make swapping recommendations, i.e., to suggest to a wallet the next swap to make based on its past history and current situation in the cross-blockchain ecosystem in which it is operating.

9. Conclusions

In this paper, we have seen how Social Network Analysis can be applied to model and manage a cross-blockchain scenario. First, we have highlighted how the enormous growth in the number and variety of blockchains makes it urgent to define a new generation of cross-blockchain platforms. We have mentioned that there are already some attempts to move in that direction, and our paper aims to provide a contribution in this setting. Indeed, we believe that Social Network Analysis, coupled with some sophisticated network models, can enable the efficient and effective modeling of all the actors involved in a cross-blockchain ecosystem. To demonstrate the veracity of such a conjecture, we first proposed a model to represent a cross-blockchain ecosystem and then a social network-based framework to manage it. Our model and framework make it possible to apply to this scenario the huge number of concepts and solutions that had been defined in the past by Social Network Analysis researchers. To show the feasibility of this, we proposed an approach for the extraction of walks, paths and cycles from a cross-blockchain ecosystem. We also showed that, when this approach is applied to a real case, it allows the extraction of interesting insights and knowledge patterns, such as the discovery of several "modi operandi" of some categories of wallets. Finally, we defined a new centrality measure specifically designed to identify the most important wallets in a given cross-blockchain ecosystem. The four main results we found in this paper are already important on their own. Moreover, collectively they can lay the foundation for a next-generation cross-blockchain platform.

Author Contributions: Conceptualization, G.B. and L.V.; methodology, F.C. and D.U.; software, G.B.; validation, E.C. and M.M.; formal analysis, G.B. and F.C.; investigation, E.C. and D.U.; resources, M.M. and L.V.; data curation, G.B. and E.C.; writing—original draft preparation, F.C. and L.V.; writing—review and editing, M.M. and D.U.; visualization, M.M.; supervision, G.B., E.C. and M.M.; project administration, F.C., D.U. and L.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: We used data accessible through www.multichain.org (accessed on 10 March 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. *Decentralized Bus. Rev.* **2008**, 21260.
2. Zhang, J.; Zhong, S.; Wang, T.; Chao, H.; Wang, J. Blockchain-based systems and applications: A survey. *J. Internet Technol.* **2020**, *21*, 1–14.
3. Maesa, D.; Mori, P. Blockchain 3.0 applications survey. *J. Parallel Distrib. Comput.* **2020**, *138*, 99–114. [CrossRef]
4. Surjandari, I.; Yusuf, H.; Laoh, E.; Maulida, R. Designing a Permissioned Blockchain Network for the Halal Industry using Hyperledger Fabric with multiple channels and the raft consensus mechanism. *J. Big Data* **2021**, *8*, 1–16. [CrossRef]
5. Bouri, E.; Shahzad, S.J.H.; Roubaud, D. Co-explosivity in the cryptocurrency market. *Financ. Res. Lett.* **2019**, *29*, 178–183. [CrossRef]

6. Gandal, N.; Hamrick, J.T.; Moore, T.; Vasek, M. The rise and fall of cryptocurrency coins and tokens. *Decis. Econ. Financ.* **2021**, *44*, 981–1014. [CrossRef]
7. Chuen, K.; Lee, D.; Guo, L.; Wang, Y. Cryptocurrency: A New Investment Opportunity? *J. Altern. Investig.* **2018**, *20*, 16–40. [CrossRef]
8. Zetzsche, D.; Arner, D.; Buckley, R. Decentralized Finance. *J. Financ. Regul.* **2020**, *6*, 172–203. [CrossRef]
9. Metelski, D.; Sobieraj, J. Decentralized Finance (DeFi) Projects: A Study of Key Performance Indicators in Terms of DeFi Protocols' Valuations. *Int. J. Financ. Stud.* **2022**, *10*, 108. [CrossRef]
10. Schulte, S.; Sigwart, M.; Frauenthaler, P.; Borkowski, M. Towards blockchain interoperability. In Proceedings of the International Conference on Business Process Management (BPM'19), Vienna, Austria, 3–5 September 2019; pp. 3–10.
11. Jin, H.; Dai, X.; Xiao, J. Towards a novel architecture for enabling interoperability amongst multiple blockchains. In Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS'18), Vienna, Austria, 2–5 July 2018; pp. 1203–1211.
12. Yeh, K.H.; Yang, G.Y.; Butpheng, C.; Lee, L.F.; Liu, Y.H. A Secure Interoperability Management Scheme for Cross-Blockchain Transactions. *Symmetry* **2022**, *14*, 2473. [CrossRef]
13. Mohanty, D.; Anand, D.; Aljahdali, H.M.; Villar, S.G. Blockchain interoperability: Towards a sustainable payment system. *Sustainability* **2022**, *14*, 913. [CrossRef]
14. Wang, W.; Zhang, Z.; Wang, G.; Yuan, Y. Efficient cross-chain transaction processing on blockchains. *Appl. Sci.* **2022**, *12*, 4434. [CrossRef]
15. Koohi-Var, T.; Zahedi, M. Cross-domain graph based similarity measurement of workflows. *J. Big Data* **2018**, *5*, 1–16. [CrossRef]
16. Borkowski, M.; Sigwart, M.; Frauenthaler, P.; Hukkinen, T.; Schulte, S. DeXTT: Deterministic cross-blockchain token transfers. *IEEE Access* **2019**, *7*, 111030–111042. [CrossRef]
17. Sigwart, M.; Frauenthaler, P.; Spanring, C.; Sober, M.; Schulte, S. Decentralized cross-blockchain asset transfers. In Proceedings of the International Conference on Blockchain Computing and Applications (BCCA'21), Tartu, Estonia, 15–16 November 2021; pp. 34–41.
18. Tsepeleva, R.; Korkhov, V. Implementation of the Cross-Blockchain Interacting Protocol. In Proceedings of the International Conference on Computational Science and Its Applications (ICCSA'21), Cagliari, Italy, 13–16 September 2021; pp. 42–55.
19. Zuo, Y.; Yang, M.; Qiang, Z.; Fei, D.; Shao, X.; Su, S.; Mo, Q.; Liang, Z. A Review of Cross-Blockchain Solutions. In Proceedings of the International Conference on Cloud Computing (CLOUD'22), Barcelona, Spain, 11–15 July 2022; pp. 221–233.
20. Frauenthaler, P.; Borkowski, M.; Schulte, S. A Framework for Blockchain Interoperability and Runtime Selection. In Proceedings of the IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS'20), Newark, CA, USA, 5–9 April 2019; pp. 106–113.
21. Chen, Y.; Bellavitis, C. Blockchain disruption and decentralized finance: The rise of decentralized business models. *J. Bus. Ventur. Insights* **2020**, *13*, e00151. [CrossRef]
22. Bonifazi, G.; Corradini, E.; Ursino, D.; Virgili, L. A Social Network Analysis based approach to investigate user behavior during a cryptocurrency speculative bubble. *J. Inf. Sci.* **2021**, *online first*. [CrossRef]
23. Bonifazi, G.; Corradini, E.; Ursino, D.; Virgili, L. Defining user spectra to classify Ethereum users based on their behavior. *J. Big Data* **2022**, *9*, 1–39.
24. Theodorou, S.; Sklavos, N. Blockchain-based security and privacy in smart cities. In *Smart Cities Cybersecurity and Privacy*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 21–37.
25. Leevy, J.; Hancock, J.; Zuech, R.; Khoshgoftaar, T.M. Detecting cybersecurity attacks across different network features and learners. *J. Big Data* **2021**, *8*, 1–29. [CrossRef]
26. Alladi, T.; Chamola, V.; Parizi, R.; Choo, K. Blockchain applications for industry 4.0 and industrial IoT: A review. *IEEE Access* **2019**, *7*, 176935–176951. [CrossRef]
27. Dorri, A.; Kanhere, S.S.; Jurdak, R.; Gauravaram, P. Blockchain for IoT security and privacy: The case study of a smart home. In Proceedings of the International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops'17), Kona, HI, USA, 13–17 March 2017; pp. 618–623.
28. Corradini, E.; Nicolazzo, S.; Nocera, A.; Ursino, D.; Virgili, L. A two-tier blockchain framework to increase protection and autonomy of smart objects in the IoT. *Comput. Commun.* **2022**, *181*, 338–356. [CrossRef]
29. Pajooh, H.H.; Rashid, M.; Alam, F.; Demidenko, S. IoT Big Data provenance scheme using blockchain on Hadoop ecosystem. *J. Big Data* **2021**, *8*, 1–26.
30. Habib, G.; Sharma, S.; Ibrahim, S.; Ahmad, I.; Qureshi, S.; Ishfaq, M. Blockchain Technology: Benefits, Challenges, Applications, and Integration of Blockchain Technology with Cloud Computing. *Future Internet* **2022**, *14*, 341. [CrossRef]
31. Zheng, Z.; Xie, S.; Dai, H.; Chen, X.; Wang, H. Blockchain challenges and opportunities: A survey. *Int. J. Web Grid Serv.* **2018**, *14*, 352–375. [CrossRef]
32. Dai, H.; Zheng, Z.; Zhang, Y. Blockchain for Internet of Things: A survey. *IEEE Internet Things J.* **2019**, *6*, 8076–8094. [CrossRef]
33. Li, X.; Jiang, P.; Chen, T.; Luo, X.; Wen, Q. A survey on the security of blockchain systems. *Future Gener. Comput. Syst.* **2020**, *107*, 841–853. [CrossRef]
34. Feng, Q.; He, D.; Zeadally, S.; Khan, M.K.; Kumar, N. A survey on privacy protection in blockchain system. *J. Netw. Comput. Appl.* **2019**, *126*, 45–58. [CrossRef]

35. Cachin, C.; Caro, A.D.; Moreno-Sanchez, P.; Tackmann, B.; Vukolić, M. The Transaction Graph for Modeling Blockchain Semantics. Cryptoeconomic Systems. 2017. Available online: http://cryptoeconomicsystems.pubpub.org (accessed on 21 February 2023) .
36. Abay, N.; Akcora, C.G.; Gel, Y.R.; Kantarcioglu, M.; Islambekov, U.D.; Tian, Y.; Thuraisingham, B. Chainnet: Learning on blockchain graphs with topological features. In Proceedings of the International Conference on Data Mining (ICDM'19), Beijing, China, 8–11 November 2019; pp. 946–951.
37. Lal, B.; Agarwal, R.; Shukla, S.K. Understanding Money Trails of Suspicious Activities in a cryptocurrency-based Blockchain. *arXiv* **2021**, arXiv:2108.11818.
38. Belchior, R.; Vasconcelos, A.; Guerreiro, S.; Correia, M. A Survey on Blockchain Interoperability: Past, Present, and Future Trends. *ACM Comput. Surv.* **2021**, *54*, 1–41. [CrossRef]
39. Khan, S.; Amin, M.; Azar, A.; Aslam, S. Towards Interoperable Blockchains: A Survey on the Role of Smart Contracts in Blockchain Interoperability. *IEEE Access* **2021**, *9*, 116672–116691. [CrossRef]
40. Borkowski, M.; Frauenthaler, P.; Sigwart, M.; Hukkinen, T.; Hladky, O.; Schulte, S. Cross-Blockchain Technologies: Review, State of the Art, and Outlook, 2019. Available online: https://dsg.tuwien.ac.at/tast/pub/tast-white-paper-4.pdf (accessed on 21 February 2023).
41. Li, S.; Kong, Y.; Nie, S. Overview of Blockchain Cross Chain Technology. In Proceedings of the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA'21), Beihai, China, 16–17 January 2021; pp. 357–360.
42. Bortolussi, F.; Hoogeboom, Z.; Takes, F.W. Computing Minimum Weight Cycles to Leverage Mispricings in Cryptocurrency Market Networks. *arXiv* **2018**, arXiv:1807.05715.
43. Pillai, B.; Biswas, K.; Muthukkumarasamy, V. Cross-chain interoperability among blockchain-based systems using transactions. In *The Knowledge Engineering Review*; Cambridge University Press: Cambridge, UK, 2020; Volume 35.
44. Shadab, N.; Houshmand, F.; Lesani, M. Cross-chain transactions. In Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency (ICBC'20), Toronto, ON, Canada, 2–6 May 2020; pp. 1–9.
45. Baniata, H.; Anaqreh, A.; Kertesz, A. DONS: Dynamic Optimized Neighbor Selection for smart blockchain networks. *Future Gener. Comput. Syst.* **2022**, *130*, 75–90. [CrossRef]
46. Cho, H.H.; Wu, H.T.; Lai, C.F.; Shih, T.; Tseng, F.H. Intelligent Charging Path Planning for IoT Network Over Blockchain-Based Edge Architecture. *IEEE Internet Things J.* **2021**, *8*, 2379–2394. [CrossRef]
47. Cormen, T.; Leiserson, C.; Rivest, R.; Stein, C. *Introduction to Algorithms*; MIT Press: Cambridge, MA, USA, 2009.
48. Lee, X.; Khan, A.; Gupta, S.S.; Ong, Y.; Liu, X. Measurements, Analyses, and Insights on the Entire Ethereum Blockchain Network. In Proceedings of the World Wide Web Conference (WWW'20), Taipei, Taiwan, 20–24 April 2020; pp. 155–166.
49. Hirschman, A. The paternity of an index. *Am. Econ. Rev.* **1964**, *54*, 761–762.
50. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.