



OPEN Insights and caveats from mining local and global temporal motifs in cryptocurrency transaction networks

Naomi A. Arnold^{1,2}✉, Peijie Zhong², Cheick Tidiane Ba^{2,4}, Ben Steer³, Raul Mondragon², Felix Cuadrado⁵, Renaud Lambiotte^{3,6} & Richard G. Clegg²

Distributed ledger technologies have opened up a wealth of fine-grained transaction data from cryptocurrencies like Bitcoin and Ethereum. This allows research into problems like anomaly detection, anti-money laundering, pattern mining and activity clustering (where data from traditional currencies is rarely available). The formalism of temporal networks offers a natural way of representing this data and offers access to a wealth of metrics and models. However, the large scale of the data presents a challenge using standard graph analysis techniques. We use temporal motifs to analyse two Bitcoin datasets and one NFT dataset, using sequences of three transactions and up to three users. We show that the commonly used technique of simply counting temporal motifs over all users and all time can give misleading conclusions. Here we also study the motifs contributed by each user and discover that the motif distribution is heavy-tailed and that the key players have diverse motif signatures. We study the motifs that occur in different time periods and find events and anomalous activity that cannot be seen just by a count on the whole dataset. Studying motif completion time reveals dynamics driven by human behaviour as well as algorithmic behaviour.

Digital assets have a growing importance in the modern world with cryptocurrencies and Non-Fungible Tokens (NFTs) being important examples: cryptocurrencies are analogous to digital wealth and NFTs are analogous to digital ownership. An advantage for researchers is that the transfer is recorded, almost always publicly, on a blockchain giving valuable data-sets and opportunities to bring sophisticated analysis tools to bear. This enables us to gain insight into the nature of these ecosystems. An additional challenge in this kind of analysis is that these systems are relatively young; Bitcoin (BTC) was created in 2008¹, NFTs in 2014 and hence they change rapidly over time. The systems can be viewed as a set of exchanges between a pair of users with each exchange having a timestamp and an amount. It is natural to view and analyse such systems as weighted temporal graphs, where nodes are users and a directed edge represents a transaction from the source to the destination node.

The paper shows that temporal motifs are a powerful tool for analysis of transaction networks. We extend the work of² by providing methods that allow temporal motifs to be examined locally (that is, counting the number and type of temporal motifs centred around each node) and temporally (that is showing how the number and type of motifs evolve over time). These extensions rely on integrating our algorithms with our Raphtory software³ that allows efficient parallel computation of temporal structures in networks. This analysis would not be possible using the software developed in². By taking a deep dive into datasets centred on cryptocurrency we show how these tools can be used in practice. The paper also presents a number of caveats: (i) Traditional assumptions about temporal ordering within a motif do not hold precisely with transactions within a single block (we show that the effects of this on our results are minor but this might not be the case for all possible analysis). (ii) Motif counts of the transaction graph as a whole might give a distribution of motifs that places undue attention on events limited to a single node or a small number of nodes and (iii) limited to a short span of time. Using motif analysis on these networks requires researchers to investigate the distribution both across time and across individual nodes. We show how the tools and techniques developed for this research can achieve this.

¹Network Science Institute, Northeastern University London, London E1W 1LP, UK. ²School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK. ³Pometry Ltd, 69 Wilson St, London EC2A 2BB, UK. ⁴Department of Computer Science, University of Milan, Via Giovanni Celoria, 18, 20133 Milano, MI, Italy. ⁵School of Telecommunications Engineering, Universidad Politecnica de Madrid, Av. Complutense, 30, Moncloa - Aravaca, 28040 Madrid, Spain. ⁶Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK. ✉email: naomialicearnold@gmail.com

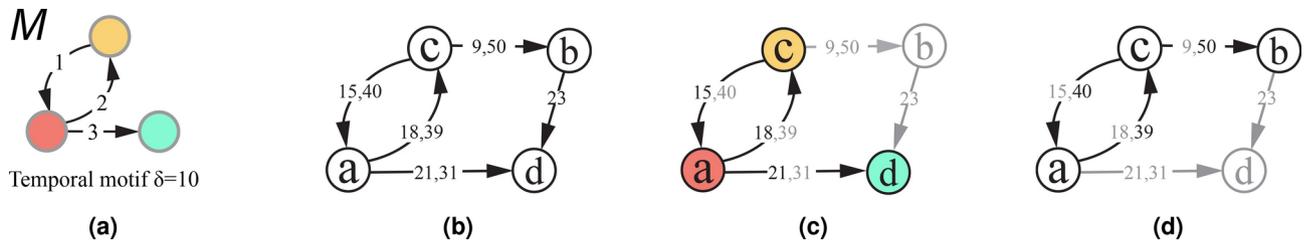


Fig. 1. An example of extracting a particular temporal motif from a temporal graph. (a) is an example of δ -temporal motif M with a given $\delta = 10$; (b) is a temporal graph with edges appearing at the times shown on each edge; (c) shows an instance of δ -temporal motifs in the temporal graph; (d) is not a δ -temporal motif because the difference between the timestamp of the first temporal edge and the timestamp of the last temporal edge exceeds the given δ .

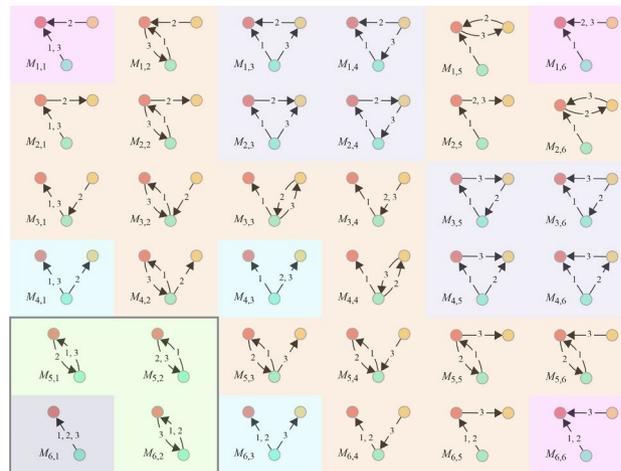


Fig. 2. Possible three-edge up-to-three node motifs considered in this work, using the same enumeration as in Paranjape et al.² with added colouring to indicate sub-types. For local motif counts, we consider an additional four two-node motifs which are exactly those pictured but with directions reversed (the direction becomes important when counting from the perspective of a node).

Background
Temporal motif definition

Motifs^{4,5} are small commonly occurring subgraphs in a network, which can be thought of as measuring the organisation of that network at a small scale. The study of motifs has been of great importance in the fields of biology⁶, anatomy⁷ and social science⁸. Temporal motifs^{2,9} add a requirement that the motif takes place in a set order and within a certain timeframe. We begin by defining what we mean by a temporal motif following Paranjape et al.². Let the tuple (u, v, t) define a transaction from user u to user v at time t , and let $\delta > 0$ represent a time interval. Define a k -node l -edge temporal motif as a sequence of edges $M = (u_1, v_1, t_1), \dots, (u_l, v_l, t_l)$ such that $t_1 < t_2 < \dots < t_l$ and $t_l - t_1 \leq \delta$, and the graph induced by these edges is connected with k nodes. $M' = (u'_1, v'_1, t'_1), \dots, (u'_l, v'_l, t'_l)$ is an instance of motif M if there is a mapping $u'_i \mapsto u_i, v'_i \mapsto v_i$ for $1 \leq i \leq l$ and the same constraints as M are obeyed. Figure 1 shows a temporal motif M , a timestamped graph, an example instance and non-instance of M for a given δ . In this paper we consider three-edge up-to-three node temporal motifs.

When considering the global motif count there are thirty six possible motifs of this type (pictured in Fig. 2). We divide them here into six classes according to the number of nodes, the subgraph formed by aggregating the motif's edges, and the direction of the transactions involved, along with their colour in Fig. 2: (i) Two-node motifs, all transactions in the same direction (dark grey); (ii) two-node motifs with transactions in mixed directions (green); (iii) a three node "star" (a centre node with two nodes connected to it) with all transactions incoming to the centre (pink); (iv) a three node "star" with all transactions outgoing from the centre (blue); (v) a three node star with mixed transaction directions (orange) and finally (vi) triangular motifs (light grey) which can have transactions in a cyclic direction (e.g. $M_{2,4}$ and $M_{3,5}$) or mixed directions. For a global motif count these are the only possible motifs with three edges and up to three nodes. Figure 2 is shaded to distinguish these motifs. With a local motif count then we can distinguish some extra motif types. For example, three transactions in the same direction between the same pair of nodes (the lower left box in Fig. 2) would be seen as "all outgoing"

from the point of view of one node but “all incoming” from the point of view of the other. The two node motifs that can be distinguished from a local node (but not a global perspective) are surrounded by a box in Fig. 2.

Related work

Temporal motifs have been used for studying financial transactions¹⁰, call networks¹¹, email interactions¹² and patent networks¹³. In the context of a *transaction networks*, temporal motifs can give us an insight into the pattern of trade. For example, if we often see the motif: A sends to B then B sends to C, it implies node B is often a “middleman” in the trade but if we often see the motif: A sends to B then A sends to C it suggests that node A may be a buyer (sending virtual currency for goods or services) or an on-ramp (trading virtual currency for real world currency). Motif counting has been used within cryptocurrency networks to detect addresses associated with mixing services¹⁴ (services designed to obfuscate the intended sender and recipient of a transaction) or for detecting phishing scams in Ethereum¹⁵ where authors posed this as a classification problem. These authors found that the inclusion of temporal motif features improved their classification results beyond using basic transaction features. Motifs have also been used to understand the longer-term formation of the socio-economic network of traders, using *graph evolution rules* to highlight how different structures like closed triangles emerge between traders^{16,17}. More generally, static motifs have been used as a descriptive graph metric for comparing cryptocurrency networks to other social networks¹⁸, as features for a token price prediction pipeline¹⁹ and to investigate the rise of smart contracts in Ethereum, typically involving groups of pre-programmed transactions²⁰. Motifs are also related to *cycles* which have been used to study wash trading and market manipulation in NFT markets by our team²¹ and others²².

Most of the literature mentioned uses temporal motifs for a specific task such as detecting a type of entity or anomalous activity. In this paper, we step back and take a more explorative approach, highlighting the insights that can be gained using temporal motifs into cryptocurrency systems as a whole as well as individuals within them, making use of three different datasets. Two of the datasets are BTC transactions centred around different “dark web marketplaces”^{23,24} (services on the dark web that exist to facilitate the exchange of goods and services, often illegal) and a third is a collection of NFT sales. We use temporal motifs with three edges and a maximum of three nodes² but with two further steps: (i) we disaggregate the motifs by user, that is to obtain a *local* motif count for every vertex in the graph and (ii) we study the whole graph motif counts over different times and different timescales. We find that counting local motifs produces a very different view of the network from simply looking at node degree or weighted node degree. It gives prominence in the analysis to a different set of nodes (the highest degree node does not have the largest motif count) and the motif count over time shows a very different trend to the transaction count over time. It also allows insight into the nature of and importance of nodes in the network that might be overlooked by more traditional temporal graph methods.

Datasets

Three cryptocurrency datasets are used in this paper. The Alphasay and Hydra datasets are bitcoin transactions centred around the Alphasay and Hydra dark web marketplaces respectively. These marketplaces act as an escrow service, such that a typical (successful) purchase on these platforms involves a transaction from the buyer to the market node followed by a transaction from the market node to the seller after the goods have been confirmed as received. They include transactions up to two hops away from the marketplace nodes (i.e. the transactions of all users who have bought or sold on these markets); this includes transactions to and from the marketplace itself. The NFT dataset is formed of transactions representing the purchase of an NFT, usually but not always in Ethereum. Unlike the dark web marketplace purchases, NFT purchases on a marketplace like OpenSea are executed using a smart contract initiating a direct transaction from buyer to seller and a small platform fee is taken by OpenSea. We do not observe transactions representing a platform fee, just the buyer to seller transactions.

In Bitcoin, one person tends to have multiple different wallets (for example, a basic transaction involving user u sending an amount to user v will involve an additional change wallet belonging to user u). The Alphasay and Hydra datasets are therefore pre-processed by Chainalysis Inc. such that wallets who are believed to be associated with the same person are merged into one entity in the graph. The NFT purchases network is not preprocessed in this way. In the NFT network the direction of the transfer is the direction in which money is sent, not the direction in which the NFT is sent, this makes it consistent with the other two datasets. Part of the nature of blockchain systems is that transactions are grouped in “blocks”. If two transactions are within the same block it is not certain which order they came in, although it might be argued that a transaction at the end of the block is highly likely to occur after one at the beginning. This issue is discussed in depth in section “[Transactions where time order is uncertain](#)”.

In all datasets, nodes represent wallets (potentially a group of wallets as discussed above for Alphasay and Hydra); a directed edge from node u to node v at time t represents a transaction from the user represented by node u to the user represented by node v . Table 1 gives some statistics on the number of nodes, the number of edges (the number of ordered pairs (u, v) such that there is at least one transaction from u to v) and the total number of transactions.

Throughout the paper we will refer to the data sets as Alphasay, Hydra and NFT. It’s important to remember that Alphasay and Hydra can be viewed as different subsets of the same underlying dataset (the global bitcoin transfer network).

Results

Motifs carry different information to (weighted) node degree

It is important to establish that motif counts are giving different information than simply looking at, for example, node degree. In Fig. 3 for each node we plot the number of motifs (with δ of one hour for Alphasay/Hydra

Dataset	Time range(mm/yy)	Number of nodes	Number of directed edges	Number of transactions
Alphabay	09/2010–12/2020	11,021,692	15,890,235	33,588,967
Hydra	07/2010–12/2020	20,137,164	35,765,997	69,811,848
NFT	11/2017–04/2021	532,945	2,991,602	6,071,027

Table 1. Dataset summary. Time range (earliest to latest recorded transaction), number of nodes, number of directed edges and number of transactions. Note the distinction between *edges*, the number of unique pairs (u, v) such that there is at least one transaction from u to v , and transactions.

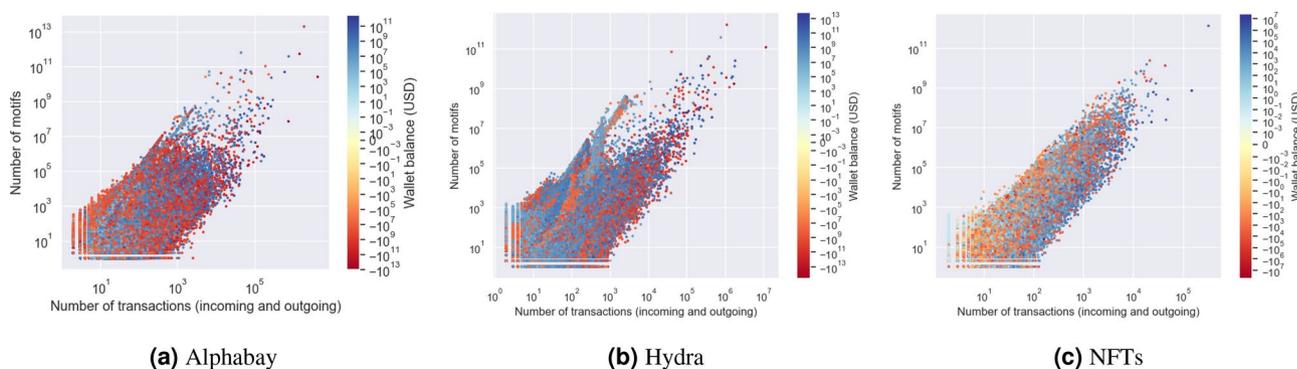


Fig. 3. Local motif counts for nodes in the Alphabay network with δ set to one hour. The plot shows total motif count versus total transactions (both incoming and outgoing). The nodes are coloured by their wallet balance (total incoming transaction value minus total outgoing transaction value). Note the scale is logarithmic in both axes and in the colouring.

and one day for NFTs) against the number of transactions. We colour the node by the wallet balance, that is the amount in dollars coming in to that node minus the amount going out within the time period. An obvious feature is that while there is a clear correlation between the number of transactions and the number of motifs, this varies by many orders of magnitude; for example a node that is involved in 1000 transactions may be part of fewer than 10 motifs but may be involved in as many as 100 billion. The node in Fig. 3a with the most transactions (corresponding to Alphabay itself) is not in the top ten for motif count. Besides the large fluctuations in the relations between number of transactions and number of motifs, one also observes non-trivial patterns, such as in Fig. 3b, which may be associated to the existence of different types of users in the system.

Global motif counts show fanout patterns of trade

In Fig. 4a–c we look at the global number of motifs for each dataset. A first thing to note is the huge variation in motif counts. The highest motif count for Alphabay is 9.3 trillion for an outgoing star motif whereas the smallest motif count is 164 thousand for a mixed triangular motif count. It is clear that some types of motif have extremely high prevalence in this data compared to others. Still, one may wonder if these high counts of motifs occur by chance due (say) to an entity having a high number of outgoing transactions and the motifs are simply a product of a large number of edges between certain node pairs. To investigate this, we use a randomised reference model (see section “Null model”) which preserves the source and destination of all transactions but permutes the time at random, so that the overall activity rate and aggregate graph structure is preserved but the order of events is lost. Figure 4d–f show the ratio between the motif counts in the original and shuffled network, showing that for Alphabay the four all-outgoing motifs are thousands of times overrepresented. All motifs are represented more than would be expected in the null reference model. Similar results hold for the Hydra dataset. An outbound pattern corresponds to an entity which is largely sending currency, perhaps buying physical goods, performing automated services or exchanging to some other (perhaps fiat) currency.

For the NFT data we found δ of one day was a more appropriate time scale (again see section “Choosing and examining the effects of time period δ ”). This may be because the NFT data set has a slower rate of arrival of edges. The NFT data set shows more diversity of motif types. The extreme difference in count magnitudes is still present but is reduced (the largest motif has a count of 1.6 billion for an outgoing star, the smallest, 471 thousand for a mixed direction triangle). While outbound star patterns are still the most common motifs, mixed star patterns with one node always inbound and one node always outbound are also common, perhaps a user selling one NFT to fund the purchase of another or buying an NFT and selling it shortly after. It is also worth noting that triangular motifs are rare in the NFT marketplace (again this is not unexpected, there is no particular reason three traders would wish to exchange NFTs between themselves) and in some cases the number of triangle motifs is lower than the chance results from the time-shuffled null model ($M_{1,4}$ and $M_{4,5}$).

A note on interpreting motif count plots In Fig. 4 the cell i, j corresponds to the count $M_{i,j}$ in Fig. 2. However, the axes can be used to interpret which motif is being used without referring to that figure. The x and y axes show the directions and participants for the three transactions. The row (y-axis) corresponds to the first two

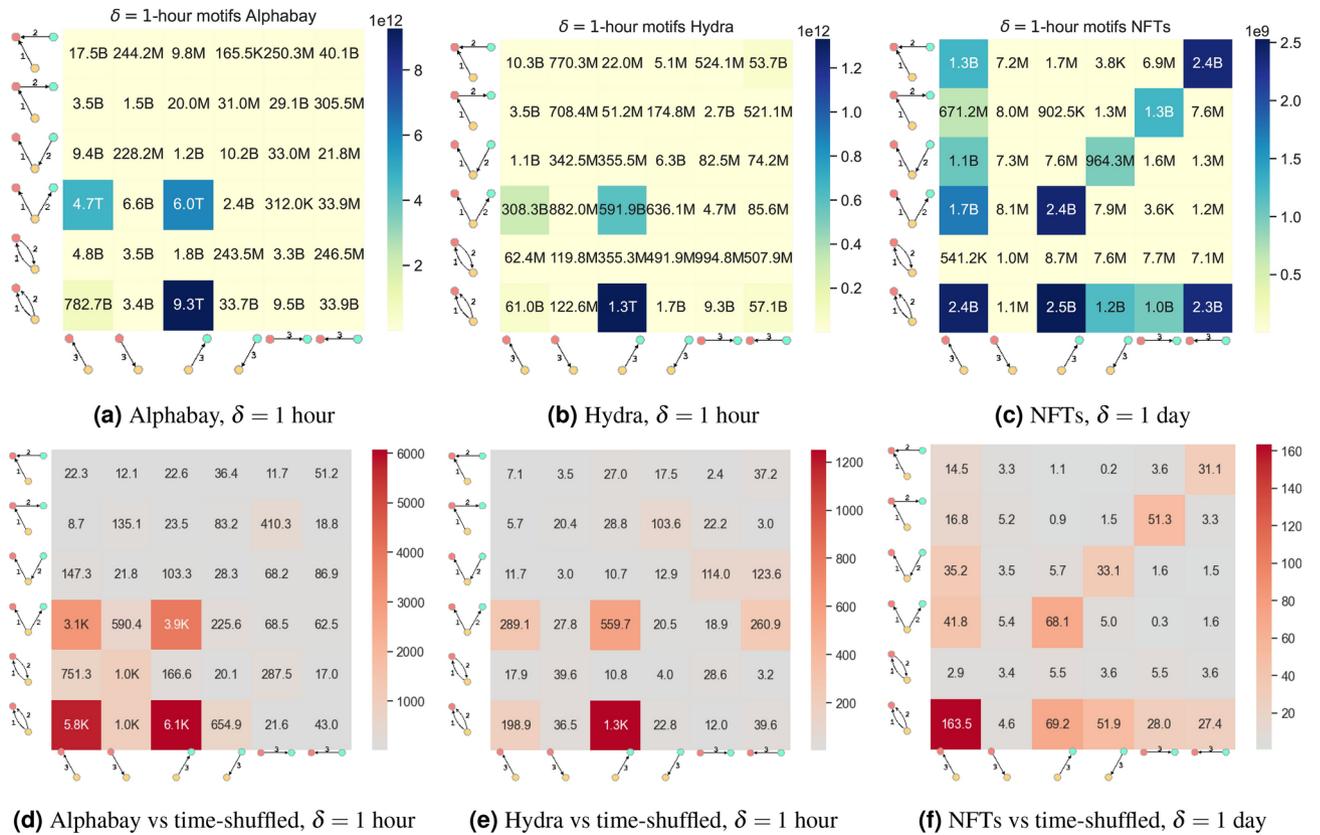


Fig. 4. Global motif counts for each dataset (top) compared with a null time shuffled model (bottom) where the lower graphs show the ratio of motif count in the unshuffled versus shuffled data. (See text for axes interpretation).

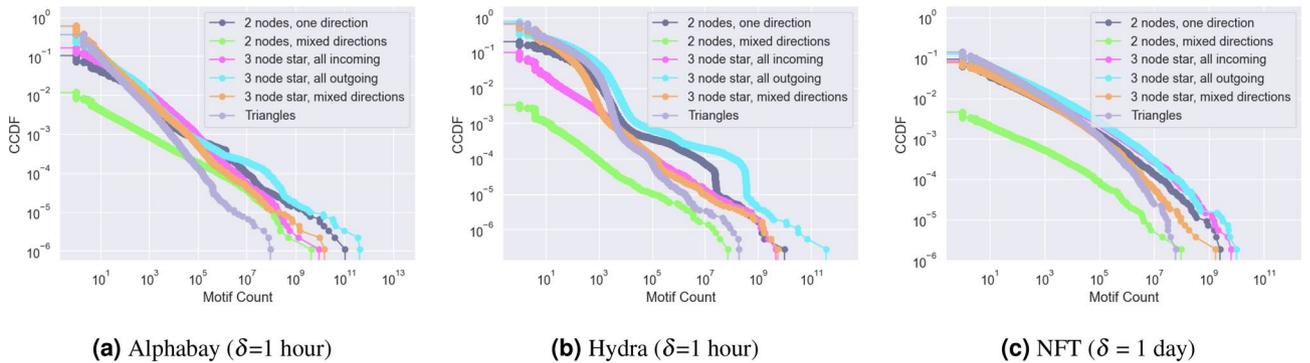


Fig. 5. Complementary cumulative distribution function of the motifs each node participates in, grouped into categories of similar motifs for Alphabay and Hydra and NFT. Both axes are log-scaled.

transactions and the column (x-axis) corresponds to the third. For example, the top right box is the motif count for transactions from yellow to red, then blue to red then blue to red a second time (a type of inbound star motif).

Motif counts are dominated by a very small number of nodes

Given the magnitude of the motif counts found in the graph-wide counts, we next ask how these motifs are dispersed across the nodes in the network. Figure 5 shows the distribution of motifs across the nodes in the network. For Alphabay we largely see straight lines in the log-log CCDF which is characteristic of power law behaviour and shows huge dominance by the nodes with the most motifs. In Hydra only the three-node star with all incoming edges and two-node same direction motifs are power-law, the others display sudden downward “bumps” (e.g. around 2×10^7 motifs for 3 node star all outgoing). These “bumps” indicate a group of nodes with extremely similar motif counts, one possible explanation for this is the presence of groups of nodes implementing similar automated behaviour.

For the NFT data Fig. 5c the situation is similar but less extreme. The curves are smoother than in the Alphabay and Hydra and fall off towards higher motif counts. However, they still indicate a situation where the bulk of the count of motifs is with a small number of nodes. The two-node mixed direction motif has a much smaller count than any other in this data set. This is unsurprising as it would indicate two people buying NFTs from each other in quick succession.

Motifs are varied among the top 10 nodes

Among the top 10 nodes in each dataset (ordered by total motif counts), Fig. 6 shows that there are a number of distinct ‘signatures’ which may signify different functional roles of the nodes in this network. One of the patterns (e.g. shared by nodes 3, 6, 8, 9 and 10 in Alphabay and 8 in Hydra) is mostly having just pairwise incoming motifs, another (nodes 1, 4, 5 in Alphabay and 1, 2, 4 in Hydra) being star motifs with all outgoing edges. The Alphabay node itself, despite its central role in the network as an escrow, placed 17th in total number of 1 h motifs. The top node in Alphabay has 19 trillion motifs in total, a slightly negative transaction balance (about 2% of its incoming transaction value) and a large out-degree (1831183) compared to in-degree (668). By checking with a blockchain explorer the addresses involved in the transaction and referencing with the GraphSense wallet/address labelling service²⁵, these addresses seem to be associated with an illicit service that was very popular for a short amount of time.

Among the highest motif nodes in the NFT dataset, some (e.g. nodes 4, 5, 7, 9 and 10) show similarities to top nodes in Alphabay and Hydra with all-outgoing patterns being common, but we also see more incoming-based star motifs (nodes 6, 8) and the top node has more of a variety of motifs.

Motifs reveal events not observable by studying transaction volume

Next we study how the motif counts are spread out through the time periods of the different datasets. To do this, we group the transactions by month and count the temporal motifs within that month with δ set to one hour for Alphabay and Hydra and δ set to one day for the NFT sales following previous experiments. Figure 7 shows these results, in terms of total motif counts and relative counts. In the Alphabay data, we see a spike in motif counts some months after the market opens in 2014, which at first seems to follow the trend in transaction volume. However, these counts drop off long before the transactions do, and long before the market is shut down. Given that the entity with the most motifs is possibly an illicit service or mixer (discussed in section “Motifs are varied among the top 10 nodes”), the drop-off could be due to users moving away from using it. In the Hydra market, we observe shadows of Alphabay’s active period between 2014 and 2017 (some of Hydra’s userbase were Alphabay users who migrated when it was shut down²³). This is followed by a similar spike and premature drop-off of motifs to Alphabay that is not mirrored in the transaction volume, which is still increasing at the end of our observation period. The vast majority of the NFT motifs occur within early 2020 when the technology was taking off followed by another spike at the end of the study period.

Following this, we investigate the composition of motifs over time, that is the counts of each category of motif as a proportion of all motif counts. In the Alphabay dataset, most of the motifs between 2013 and 2016 are two-node motifs representing transactions between peers (it is unlikely that these two-node motifs would involve the marketplace at this point as this would additionally invoke star motifs). At some point coinciding with the spike observed in all motif counts, this behaviour transitions to being all-outgoing stars which dominate, which could be the result of entities like the marketplace itself or an exchange facilitating transactions, or switching behaviour of users making purchases among different vendors. However, instead of the two-node motifs returning after Alphabay is shut down in 2017, the star motifs are still the most prevalent, fluctuating between all-incoming and all-outgoing. In the NFT dataset, apart from early 2020 where there is a spike in motif volume, most of the motifs are all-incoming stars. This could be due to the practice of sellers auctioning a number of their NFTs with the same closing period (and hence the transactions from winning bidders come in at the same time).

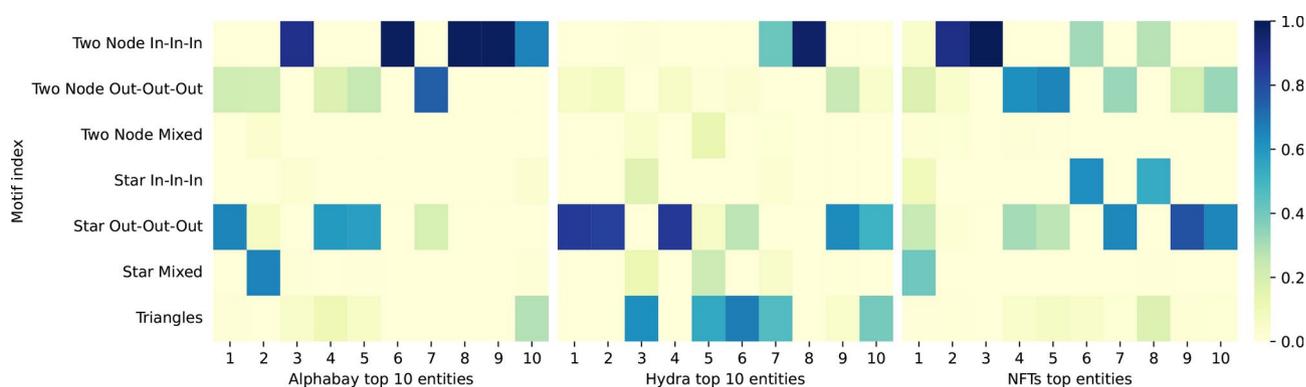


Fig. 6. Motif ‘signatures’ among the top 10 entities in Alphabay, Hydra and NFTs. Each column is obtained by taking an entity’s motif vector and dividing by the total number of motifs for that entity, relative to the motif split in the timestamp-shuffled model. The entities are displayed left to right in descending order of total motif counts.

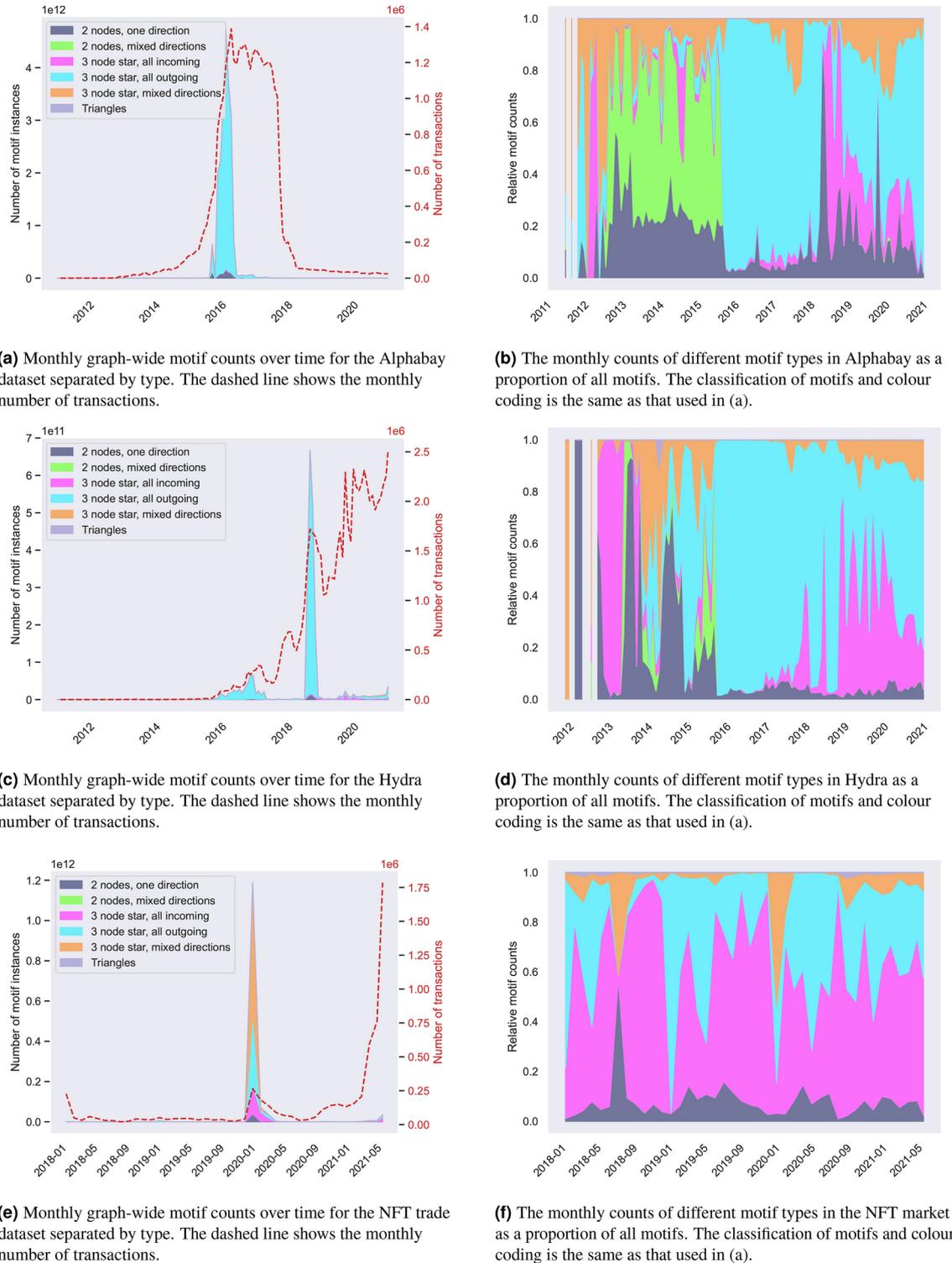
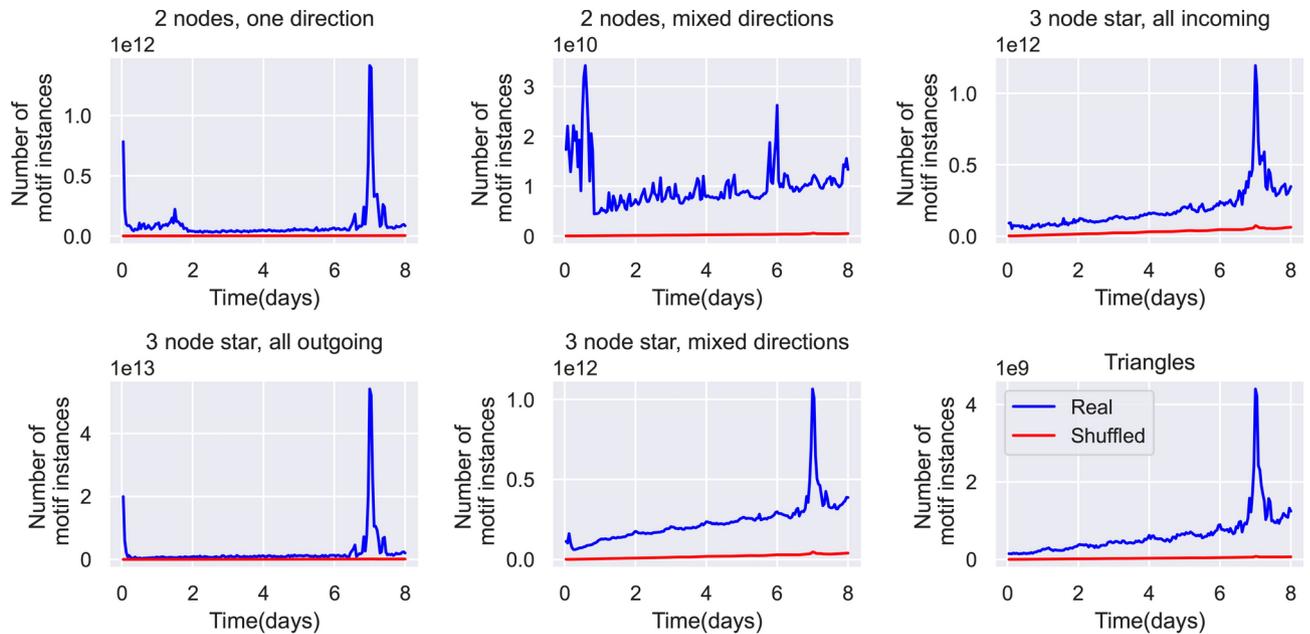


Fig. 7. Results on 1 h motif counts for both datasets on a month-by-month basis. The left hand plots show the total monthly motif counts for each dataset grouped by motif types; the right hand shows these counts as a proportion of all motifs present.

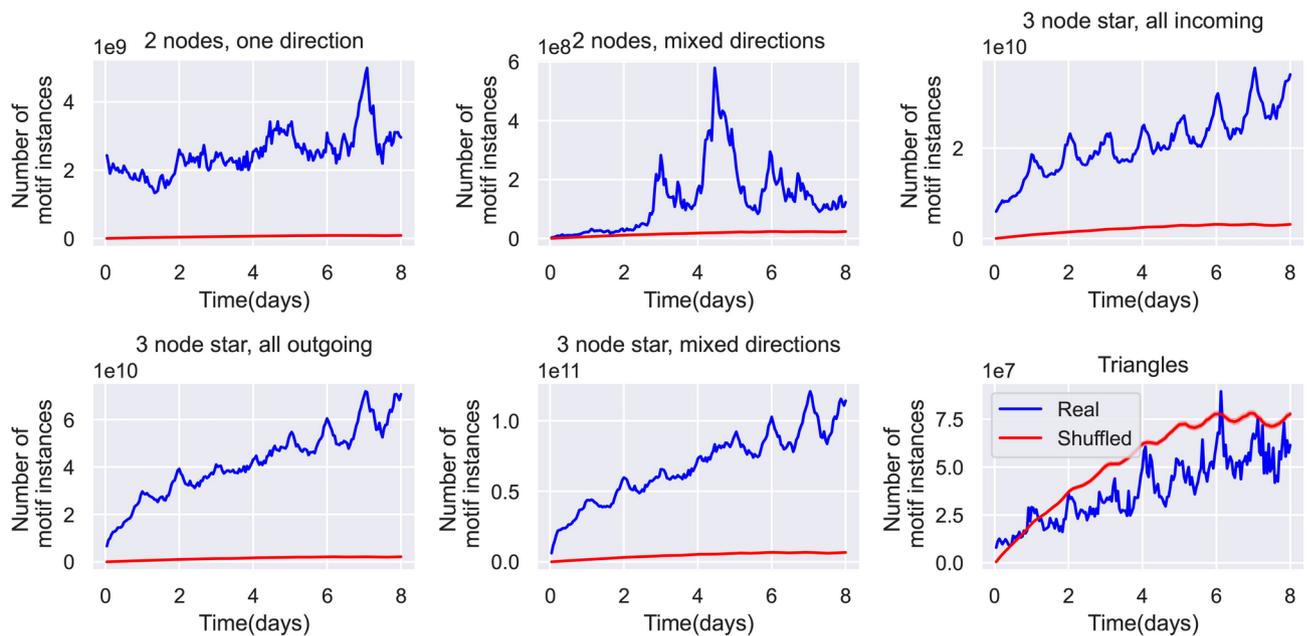
Choosing and examining the effects of time period δ

Obviously, the choice of δ the time period is important. In this section we justify the choice of one hour for Alphabay and Hydra and one day for NFT. We do this by asking how many extra motifs are introduced by increasing the time window by a single hour. This highlights time periods of particular importance. By subtracting the motifs obtained with time window δ plus one hour from those with time window δ we can investigate the new motifs introduced as the time window used is increased. Moreover we can assess the robustness of the δ

parameter as a measure of the number of motifs. Figure 8 shows the effect of increasing δ from one hour up to just over one week (168 h) by one hour increments. First, we see that for Alfabay, our chosen size for most of the analysis $\delta = 1$ h captures the vast majority of short-range motifs, seen by the dip from one to two hours across the most common motifs (those which are three nodes outgoing and three nodes mixed directions). Apart from the two nodes mixed directions motif, the increase in motifs is then stable up to seven days. The graph for



(a) Alfabay transactions



(b) NFT purchases

Fig. 8. The number of each type of motif as δ the timescale is increased a single hour at a time for (a) Alfabay (b) NFT purchases. A peak in the graph at a timescale of t hours shows that a large number of motifs are introduced by moving the timescale from $t - 1$ to t hours. In each graph, the blue line shows the count for the real dataset and the red line shows the mean value for the timestamp shuffled data calculated over 10 realisations, with standard deviation error bars.

two nodes mixed directions shows a variety of peaks at between one hour and twelve hours; the reason for this is not completely certain but these could be human-moderated trades between two individuals done on the same day. Finally, all motifs show a peak at one week, this is due to an autofinalisation escrow mechanism in Alfabay where trades are cleared if not confirmed by the buyer within seven days. This leads to a peak at precisely seven days where a transaction directly causes another exactly seven days later. This is not a particularly useful timescale for analysis though, firstly because it is very long and would lead to extremely large motif counts but mainly because the motifs are a result of a “refund” mechanism rather than genuine trades.

For the NFT trading the three node motif counts all show large peaks at intervals of one day. The peaks increase in size as the number of days increases. A possible explanation is that the motif counts are related to the human diurnal cycle and people tend to trade with people in the same geographic area. Whatever the mechanism, this makes δ of 1 day the most obvious time period to study. In the case of motifs involving only two nodes there is no clear peak in the motif counts but these motifs are relatively rare compared with three node motifs (less than 1% of the number of three node motifs). It is also interesting to note that the time-shuffled null model produces as many or more triangles than the real data showing that triangular trade patterns between three individuals are rare in the NFT marketplace within the one week time scale.

Discussion and conclusion

This paper investigated the use of temporal network motifs to explain trade in cryptocurrency networks. We showed how to optimise the time period δ (the maximum time for a temporal motif) and how varying it gave insight into the behaviour of the studied systems. Efficient temporal graph computation in the Raphtory software allowed us to investigate multiple values of δ and multiple time windows. This was impossible using existing bespoke temporal motifs software. It allowed us to identify behaviours we that we would not have been able to see from a single or small number of δ values. We could see both human (diurnal) and mechanism (e.g. escrow timeouts) induced behaviours in the graphs.

In all three datasets a naive analysis of graph-wide temporal motif counts would be misleading. It is vital to consider the distribution of motifs over the time dimension and individual nodes (with particular attention to “heavy-hitters”). All the networks studied had extreme nodes that contributed a high proportion of the motifs. Simply counting motifs for the Alfabay network, for example, would not tell you that a high proportion of all motifs were from a single node in a short (six month) period. For the networks with observations centered around dark markets (Alfabay and Hydra) it was unexpected that the nodes representing the markets themselves did not top the list of nodes by motif count.

In the Hydra network we could clearly see evidence of synchronisation between multiple nodes likely indicating coordinated, automated behaviour. The top ten nodes by motif count could be grouped into distinct classes: mainly buying behaviour (multiple transfers to one or more nodes); mainly selling behaviour (multiple transfers from one or more other nodes) and a small number exhibiting a mixture of sell/buy behaviours. Triangular motifs were rare in all cases and lower than the null model (chance) rate for the NFT data. Done with care this form of analysis of networks leads to insights into trade behaviours not obtainable from other analysis tools.

Given the diverse motif profiles of heavy-hitting users and the ability to uncover events not observable by simply counting transactions, we believe local temporal motifs may prove useful as features for classification or clustering tasks, such as in the domains of anti-money laundering or fraud detection.

Methods

In this paper, we study three-edge, up-to-three node, temporal motifs, defined in section “[Temporal motif definition](#)”. Three is chosen for the following reasons. Firstly, it lies in the middle of a tradeoff between descriptive power and computational complexity; there are only six possible two-edge temporal motifs but an algorithm for counting four-edge, up-to-four node, temporal motifs is not feasible to run on this scale of data because of the combinatorial explosion. In addition, with so many possible motifs interpretability becomes a problem. Secondly, a number of datasets have been studied using this technique^{2,10,13}, hence our work is more comparable with that of other researchers.

Counting local and global motifs

In this paper, as well as studying global motif counts over the whole graph, we study motif counts from the perspective of the nodes involved in them. We refer to these as the *local motif counts* and we also study them as they change in time. We first define a temporal graph $G = (V, T)$ where V is a set of *vertices* and T is a sequence of tuples (u_i, v_i, t_i) , $1 \leq i \leq m$ with $u_i, v_i \in V$ and $t_1 \leq t_2 \leq \dots, \leq t_{m-1} \leq t_m$ which we will refer to as *temporal edges*, representing an event (e.g. a trade of currency) from node u_i to node v_i at time t_i . In our data the event often also has a weight w_i representing the amount of currency involved in the transaction. The local motif count of vertex u for a given motif M (motif definition in section “[Temporal motif definition](#)”) refers to the number of times vertex u has participated in motif M . To count the local motifs, we implement a modified version of the algorithm by Paranjape et al.² in our temporal graph library Raphtory^{3,26} which performs the counts in parallel across the vertices. Raphtory was built from the ground up to study large-scale temporal networks efficiently, including the possibility to use ‘windowed’ approaches which enabled the study of temporal motifs over time (section “[Motifs reveal events not observable by studying transaction volume](#)”). Global temporal motifs counting code from SNAP²⁷ was tested on the Alfabay dataset but didn’t complete due to out-of-memory issues, whereas the implementation in Raphtory took 3 min on the same hardware (2021 MacBook Pro with 16GB RAM). Note that our motif counting algorithm is fundamentally the same; in both implementations, motifs are counted for each vertex using the Paranjape et al. counting algorithm². In their work, these local counts are then summed to

obtain graph-wide motif counts whereas we also study the local counts. Differences in runtime/memory usage are therefore likely due to differences in the underlying graph storage/low level functions of SNAP and Raptory. With this in mind, we point the readers to their work² for pseudocode of the algorithm. We instead focus here on the details of what is meant by a local motif in terms of which nodes in a motif count that motif. A speed comparison with² is in section “[Speed of execution compared to state-of-the-art](#)”.

Two-node motifs

There are eight (2^3) possible two-node, three-edge, motifs when measured from the perspective of an individual node and this corresponds to three links in temporal order with two possible directions. When viewed globally, this collapses to the four possibilities shown in the bottom left of Fig. 2 when not viewed from the perspective of a single node. Consider the following example: from the point of view of a node pair (u, v) it makes sense for the node u to count a three edge temporal motif with v that as “outgoing, incoming, outgoing” and from the perspective of v this same motif would be counted as “incoming, outgoing, incoming”. From a global perspective this is a single motif with transactions in alternating directions. For plotting, we often group these into all-outgoing, all-incoming and mixed directions. The time-complexity of counting two-node motifs is $O(E)$ where E is the number of temporal edges.

Three-node stars

The star motifs (all motifs in Fig. 2 excluding the two-node motifs and triangle motifs) involve one central node having transactions with two other leaf nodes. In our work, we include an instance of this motif in a node u 's local motif count only if u is the central node of this motif. One might argue that the two leaf nodes should also count this motif instance. However, we argue that (i) the central node is more of interest in such a motif, (ii) having all three nodes count the motif would result in more noise, and (iii) would be more computationally intensive to run, as it may instead involve running the global motifs count across the two-hop neighbourhood of each node. As with two-node motifs, the time-complexity of counting star motifs is $O(E)$.

It is useful to aggregate these star motifs into different sub-categories of all-outgoing, all-incoming and mixed-direction stars since they represent different behavioural patterns: it is more likely to see three outgoing transactions happen at a similar time than three incoming transactions since they originate from the same source, and mixed direction stars may involve some dependent sequence of events. This classification is used in some of the figures for clarity instead of showing all 36 global/40 local motifs.

Triangles

Triangle motifs are counted locally for each of the nodes involved in a triangle since, for each of those motifs, it is hard to make an argument for one of its constituent nodes being the most important. Additionally, there is little extra difficulty counting triangle motifs for each node (compared to leaf nodes of the star motifs for example), they are simply over-counted once for each node, with respect to the global counting algorithm which counts each triangle once. The time-complexity of counting triangle motifs in Raptory is $O(VE + \tau E)$ where τ is the number of static triangles in the graph. The first term relates to the enumeration of static triangles in the graph, the second to counting the number of triadic temporal motifs given the edges involved in each static triangle. We argue that in cryptocurrency transaction networks, the first term tends to dominate since the number of triangles in such a network is fairly small; this may not be the case in social networks.

Choice of timescale

Naturally, the choice of timescale δ plays a role in what motifs are discovered. Human communication patterns have been known to have characteristic timescales²⁸ based on memory/attention²⁹ and diurnal behaviour³⁰. In transaction networks, we may also expect to see timescales resulting from pre-programmed rules such as preset durations for NFT auctions, transaction clearing times or payroll. While we perform a deeper study of which motifs arise at which timescale in section “[Choosing and examining the effects of time period \$\delta\$](#) ”, for most of the results we choose a timescale of one hour for the Alphabay/Hydra datasets and one day for the NFT datasets, corresponding to the first peak for a number of the motifs in Fig. 8. One observation which may be useful to practitioners is that one can study motifs which have a duration between values δ_1 and $\delta_2 > 1$ by subtracting the motifs obtained with $\delta = \delta_1$ from those obtained with $\delta = \delta_2$. In some scenarios for example, motifs from transactions that all occur within the same block may be indicative to an algorithmic mechanism. If it is preferable to exclude such motifs in an effort to exclude pre-programmed behaviour, this can therefore be done by subtracting motifs with δ set to some time period much less than the block processing time e.g. 1 s.

To encourage researchers to explore this important effect of timescale choice on results using experiments like Fig. 8, our implementation of the motif mining algorithm in Raptory includes the possibility to provide a range of delta values as input for the algorithm. This makes use of the observation that enumerating the static triangles in the network is typically the bottleneck of the algorithm's runtime. Having identified the edges which participate in these static motifs, running the motif counting procedure on these edges for each delta value becomes significantly faster than re-running the full algorithm for each delta value.

Speed of execution compared to state-of-the-art

Figure 9 shows the time to count motifs for a range of δ values in the Raptory implementation compared with what is arguably the most competitive alternative, the SNAP temporal motifs implementation. For each δ value on the x -axis, the global motifs algorithm is run for an array of timescales: 1 h, 2 h, 3 h, ..., δ hours with the corresponding value on the y axis showing the total runtime in seconds. Therefore, here δ represents both the number of timescales processed and the size of the largest timescale in hours. For fewer than 6 windows, SNAP is faster, which is likely because Raptory as a more general-purpose library incurs a slightly higher

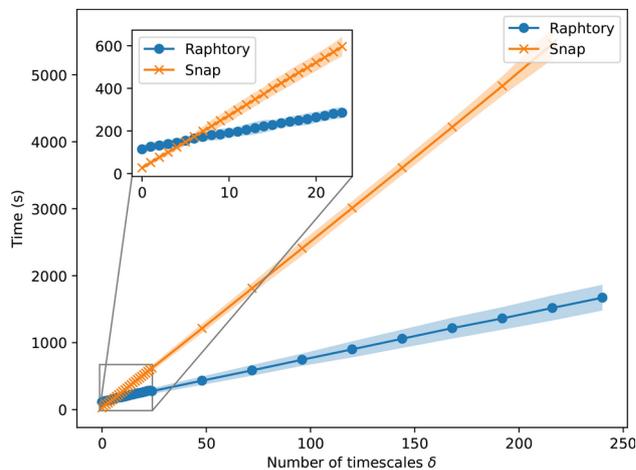


Fig. 9. The time taken to count motifs for a range of δ values. For each value δ on the x -axis, the corresponding y value is the time taken to count motifs for the input array $[1h, 2h, \dots, \delta h]$. The result is averaged over 10 experiments with shaded 95% confidence intervals.

initial overhead for building the graph. However, the time increment for adding more windows in Raphorty is much smaller than the alternative in SNAP which can be seen by Raphorty's smaller slope. Note that in both implementations, the time is linear in the number of δ values processed and that the effect of the size of timescale δ itself on the time taken to process is negligible.

Null model

To investigate whether motifs arise from temporally correlated behaviour or are simply a result of a large number of transactions we use a null model. In this case we use a randomised reference model where for our set of transactions (u_i, v_i, t_i) the nodes u_i and v_i are kept the same and the t_i are shuffled randomly among the transactions. This has the effect that between any node pair the same absolute number of transactions occurs over the whole study period and for any chosen time period the same number of transactions occurs in the network. We will refer to this model as the *timestamp-shuffled* model³¹. For this paper, we produce 10 shuffled versions of each dataset and use the term “relative to the null model” to mean that we divide the value as found in the real dataset by the mean of the values found on the 10 shuffled versions for that dataset.

Transactions where time order is uncertain

In a blockchain setting, transactions are released in *blocks* according to when they are successfully verified by a miner, typically in time intervals of around 8–10 min. This means that while a transaction can be initiated by a user/service at any continuous time, it is only possible to observe the block-wide timestamp of each timestamp. The order in which they occur in the block is the order in which they were verified by a miner which comes with a topological order guarantee that if transaction τ_2 spends an output of τ_1 then τ_1 must come before τ_2 in the block¹, however there is no other guarantee on time ordering within the block. In fact, miners are incentivised to verify the transactions with the highest transaction fee first. For this work we take the block order to be the order of the transactions, but show in our [supporting information](#) that randomising intra-block orderings causes a deviation in the motif counts with a maximum at around 3%.

Data availability

Each of the algorithms (a local motif counting algorithm, a global motif counting algorithm, and a global motif algorithm optimised for multiple timescales) has been implemented in the open-source library Raphorty³ in Rust and Python. The Alfabay and Hydra datasets are proprietary but the NFT dataset is publicly available³² and scripts for reproducing the NFT results can be found at <https://github.com/narnolddd/motif-paper-reproduce>. All plots were generated using Matplotlib 3.8.2 (<https://matplotlib.org/stable/>) and Seaborn 0.13.2 (<https://seaborn.pydata.org/>).

Received: 2 February 2024; Accepted: 4 October 2024

Published online: 04 November 2024

References

1. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. *Decent. Bus. Rev.* 21260 (2008).
2. Paranjape, A., Benson, A. R. & Leskovec, J. Motifs in temporal networks. In *Proceedings of the tenth ACM International Conference on Web Search and Data Mining*, 601–610 (2017).
3. Steer, B. et al. Raphorty: The temporal graph engine for Rust and Python. *J. Open Source Softw.* 9, 5940 (2024).
4. Milo, R. et al. Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
5. Yaveroglu, Ö. N. et al. Revealing the hidden language of complex networks. *Sci. Rep.* **4**, 4547 (2014).

6. Vázquez, A. et al. The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *Proc. Natl. Acad. Sci.* **101**, 17940–17945 (2004).
7. Sporns, O. & Kötter, R. Motifs in brain networks. *PLoS Biol.* **2**, e369 (2004).
8. Holland, P. W. & Leinhardt, S. Local structure in social networks. *Sociol. Methodol.* **7**, 1–45 (1976).
9. Kovanen, L., Karsai, M., Kaski, K., Kertész, J. & Saramäki, J. Temporal motifs in time-dependent networks. *J. Stat. Mech. Theory Exp.* **2011**, P11005 (2011).
10. Liu, P. et al. Temporal motifs for financial networks: A study on Mercari, JPMC, and Venmo platforms. arXiv preprint [arXiv:2301.07791](https://arxiv.org/abs/2301.07791) (2023).
11. Kovanen, L., Kaski, K., Kertész, J. & Saramäki, J. Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences. *Proc. Natl. Acad. Sci.* **110**, 18070–18075 (2013).
12. Barnes, M. R. et al. Temporal network analysis of email communication patterns in a long standing hierarchy. *Proceedings of the International AAAI Conference on Web and Social Media* (2024).
13. Liu, P., Masuda, N., Kito, T. & Sariyüce, A. E. Temporal motifs in patent opposition and collaboration networks. *Sci. Rep.* **12**, 1917 (2022).
14. Wu, J. et al. Detecting mixing services via mining Bitcoin transaction network with hybrid motifs. *IEEE Trans. Syst., Man, Cybern. Syst.* **52**, 2237–2249 (2021).
15. Wang, Y., Wang, H., Lu, X., Zhou, L. & Liu, L. Detecting Ethereum phishing scams with temporal motif features of subgraph. In *2023 IEEE Symposium on Computers and Communications (ISCC)*, 631–636 (IEEE, 2023).
16. Galdeman, A., Zignani, M. & Gaito, S. Disentangling the growth of blockchain-based networks by graph evolution rule mining. In *2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*, 1–10 (2022).
17. Ba, C. T., Zignani, M. & Gaito, S. Characterizing growth in decentralized socio-economic networks through triadic closure-related network motifs. *Online Soc. Netw. Media* **37–38**, 100266 (2023).
18. Lee, X. T., Khan, A., Sen Gupta, S., Ong, Y. H. & Liu, X. Measurements, analyses, and insights on the entire Ethereum blockchain network. In *Proceedings of the Web Conference 2020*, 155–166 (2020).
19. Chen, Y. & Ng, H. K. T. Deep learning Ethereum token price prediction with network motif analysis. In *2019 International Conference on Data Mining Workshops (ICDMW)*, 232–237 (IEEE, 2019).
20. Wu, Z. et al. Know your transactions: Real-time and generic transaction semantic representation on blockchain & web3 ecosystem. In *Proceedings of the ACM Web Conference*, vol. 2023, 1918–1927 (2023).
21. Yousaf, H. et al. Non-Markovian paths and cycles in NFT trades. arXiv preprint [arXiv:2303.11181](https://arxiv.org/abs/2303.11181) (2023).
22. von Wachter, V., Jensen, J. R., Regner, F. & Ross, O. NFT wash trading: Quantifying suspicious behaviour in NFT markets. arXiv preprint [arXiv:2202.03866](https://arxiv.org/abs/2202.03866) (2022).
23. Nadini, M. et al. Emergence and structure of decentralised trade networks around dark web marketplaces. *Sci. Rep.* **12**, 1–9 (2022).
24. ElBahrawy, A. et al. Collective dynamics of dark web marketplaces. *Sci. Rep.* **10**, 1–8 (2020).
25. Haslhofer, B., Stütz, R., Romiti, M. & King, R. Graphsense: A general-purpose cryptoasset analytics platform. arXiv preprint [arXiv:2102.13613](https://arxiv.org/abs/2102.13613) (2021).
26. Steer, B., Cuadrado, F. & Clegg, R. Raptory: Streaming analysis of distributed temporal graphs. *Future Gener. Comput. Syst.* **102**, 453–464 (2020).
27. Leskovec, J. & Sosič, R. Snap: A general-purpose network analysis and graph-mining library. *ACM Trans. Intell. Syst. Technol. (TIIST)* **8**, 1 (2016).
28. Aoki, T., Takaguchi, T., Kobayashi, R. & Lambiotte, R. Input-output relationship in social communications characterized by spike train analysis. *Phys. Rev. E* **94**, 042313 (2016).
29. Williams, O. E., Lacasa, L., Millán, A. P. & Latora, V. The shape of memory in temporal networks. *Nat. Commun.* **13**, 499 (2022).
30. Arnold, N. A. et al. Moving with the times: Investigating the alt-right network Gab with temporal interaction graphs. *Proc. ACM Hum. Comput. Interact.* **5**, 1–17 (2021).
31. Gauvin, L. et al. Randomized reference models for temporal networks. *SIAM Rev.* **64**, 763–830 (2022).
32. Nadini, M. et al. Mapping the NFT revolution: Market trends, trade networks and visual features. *Sci. Rep.* **11**, 20902 (2021).

Acknowledgements

We thank Chainalysis Inc. for providing two of the datasets used in this research. The work of Felix Cuadrado was supported by the HE CEDAR Project under Grant 101135577.

Author contributions

N.A., R.C., R.M. and F.C. designed the research. N.A. and B.S. wrote the main algorithm implementation. N.A., P.Z., and C.B. ran the experiments. R.L. advised on experimental design and interpretation of null model results. N.A., R.C., C.B. and P.Z. wrote the manuscript. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-75348-7>.

Correspondence and requests for materials should be addressed to N.A.A.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024