Venmotifs

Tom Knowles
tknowles

Cynthia Liang
liangcyn

Jenny Zhi
jzhi21

1 Introduction

Motif analysis is a technique used in network analysis that has classical origins [11], but was rediscovered and pioneered in the field of systems biology in a milestone 2002 paper by Milo et al [6]. The technique consists of counting all possible n-node subgraphs of a graph (with some constraints for tractability), usually comparing the frequencies of these subgraphs to corresponding frequencies in some null model. The technique allows one to decompose a large graph into smaller structural units that recur significantly within the overall graph, which (particularly within systems biology) often serve some function or give some insight as to the structure of the larger graph [1].

Partially due to motif analysis's origins within systems biology, temporal graphs with edge attributes have been largely untreated by this technique, despite such graphs being very common in modern social sciences—for example, graphs of messages between users are of this form. We present a novel algorithmic generalization of motifs ("Venmotifs") which extends the concept of motifs to such graphs. We apply these algorithms to enumerate certain motifs among a large sample of Venmo transactions, providing new insight into motifs in this dataset, and proving the usefulness of the novel algorithms.

2 Problem Definition

We consider the problem of motif analysis on (directed) temporal graphs with edge attributes. These are (directed) graphs where every edge has some associated attribute, and an associated time at which that edge was created. Observe that temporal graphs can be multi-graphs, since multiple edges (with different times, and possibly different attributes) can be made between the same two nodes. Some of the largest graphs consisting of interactions between actors are therefore canonically modelled using such graphs—important examples would include messages between users in social networks (where the edge attributes are the message), payments between financial entities (where the edge attributes are a value), and internet requests (where the edge attributes are types of request and payload).

From here, this paper will use Venmo transactions as a recurring example of temporal graphs, although the approaches and motivations described easily apply to many other important graphs. Venmo is an online payment network, where users are able to send each other money, along with a message—the nodes in the associated graph are users, the edges are payments (with times paid), and the attributes are the messages associated with payments, which usually describe the payment.[1] We hypothesise that the Venmo dataset is likely amenable to motif discovery and analysis—that it consists of discoverable, smaller functional units, such as users splitting shared bills, resolving debts, or paying bills to each other at regular intervals. However, it appears that fine-grained functional units would likely not be discoverable using standard network motifs that ignore either temporal aspects or edge attributes. For example, there is a qualitative functional difference between the two star motifs in Figure 1, but this cannot be discovered without comparing the attributes, since the subgraphs are identical without attributes.

[1] In general, it would also be useful to have the amount paid, but this is generally private.

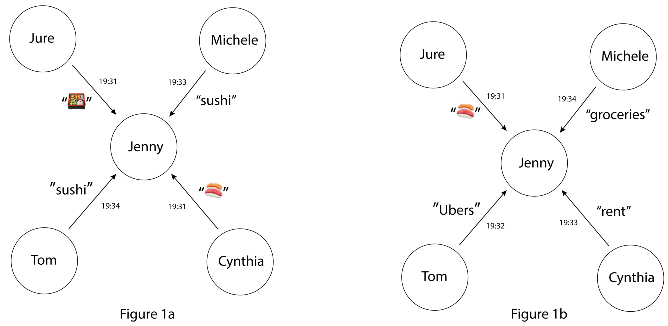

Figure 1: Figure 1a likely corresponds to a sushi dinner, whereas Figure 1b likely corresponds to Jenny collecting a diverse array of debts she is owed. These correspond to different user behaviours, but both are star motifs with 5 nodes and 4 edges without considering the edge attributes

Another example would be a motif corresponding to "recurring monthly bills", as shown in Figure 2:

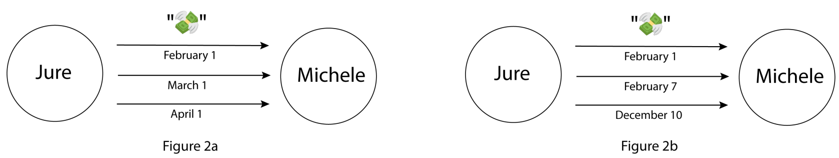

Figure 2: Figure 2a likely corresponds monthly bills, while Figure 2b corresponds to more random payments.

Differentiation between any of these functional units clearly requires consideration of both the temporal and edge attribute features of the network, which most methods in motif analysis fail to perform. We provide new algorithmic definitions, which are able to take into account both time and edge features in motif analysis. We prove the usefulness of these algorithms on the Venmo dataset.

3 Related work

The standard modern definition of motifs was originally proposed and defined by Milo et al in 2002 [6]. This is the definition and model most commonly used, which has had great success in certain fields, but fails to resolve the complexities discussed above with temporal edges and edge features. It also fails to apply in multi-graphs, due to its focus on nodes rather than edges. However, a number of more recent papers since then have showed that motifs are both theoretically and empirically important to the higher-order structure of networks [2].

More recent papers have attempted to deal with motif analysis in multigraphs and specifically temporal networks. A fairly influential and recent paper is from Paranjape et al [7]. Its motivation is that that motifs within a network of messages where the edges are separated by significant periods of time should be interpreted to motifs happening within a short interval. They therefore define a δ-temporal motif as some (time-ordered) graph motif "that appears within a δ-period of time". The paper then presents a number of algorithms for counting δ-temporal motifs with dynamic programming. This is somewhat useful, but the time-ordered nature may be useful to some datasets and not others, so we will are more agnostic about this in our definition of motif. However, our work will build directly off of this work by Paranjape et al.

Other definitions of motifs in temporal graphs [5] often place very strong requirements on the edges in the motif, such as edges having to be consecutive in time. This tends to be much less useful in large graphs, like the Venmo dataset.

An approach that does not feature time-ordered motifs is by Zhao et al [12]. This approach is fairly interesting, but makes a number of assumptions about networks consisting of "communication about events" between users, which may be less useful to more general datasets. It also generally fails to usefully consider edge attributes. An important related paper on social communication networks is from Kossinets et al [4], but this fails to consider motifs, being more focused on information flow through networks.

Lastly, we note that number of recent papers combine edge features with local network structure (as we hope to do) to obtain useful results. One example is from Kahanda et al [3]. This paper finds that network-transactional features considering not only edge features between two users, but also edge attributes in a neighborhood, are very useful to link strength prediction. This suggests a model of motifs including the features of a local neighborhood may be very useful to characterizing such graphs using motifs. We view our algorithm as a first step to using a neighborhood's transactional features within the context of motif analysis.

4 Algorithm

4.1 Prerequisites

We assume we have a temporal network $G = (V, E, T, F)$, which has both temporal annotations and attributes on the edges. We let T be the edge-indexed set of times (so T_e is the time of edge e), while F is an edge-indexed set of attributes (so F_e is the attribute on edge e.) We'll then define some similarity metric d between edges, that will always use the attributes on the edges—so $d(e, e')$ is the similarity between edges e and e'. An example of this might be a metric embedding of the attributes F_e and $F_{e'}$, followed by taking the cosine similarity of that metric embedding clipped to 0, to get a similarity value $r \in [0, 1]$.

4.2 Temporal Motifs with Edge Attributes

We now define a **temporal motif with edge attributes.** This is just another graph $M = (V^M, E^M, T^M, F^M)$ that has its own times and attributes. Given some M, we can then use some (possibly temporal) motif enumeration approach to enumerate all isomorphic subgraphs $H \cong M$ in G (ignoring attributes) and then take the average similarity between each edge in H and the corresponding edge in M:

$$S(H, M) = \mathbb{E}_{e_h \in H, e_m \in M} \left[d(e_h, e_m) \right].$$

This is a score corresponding to how close H is to M, when taking into account the attributes on the edges. This is shown diagrammatically in Figure 3, and allows us to quantify how often M appears in such a network in a much more fine-grained manner—[possibly by taking the sum of this score over all isomorphic subgraphs H, so $\sum_{H \cong M} S(H, M)$.]

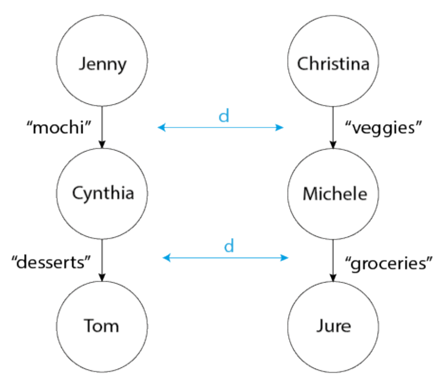

Figure 3

Figure 3: Diagrammatic view of the similarity score between two temporal motifs with attributes

4.3 Venmotifs

One difficulty with the above definition is it requires one to specify both the attributes and times on any motif M that one wishes to enumerate. While this makes the definition useful for looking for and comparing the frequency of certain user behaviours, it makes it less useful for motif discovery. To allow motif discovery without specifying times and features one could therefore also define **Venmotifs**, which are motifs scored by the self-similarity of their edges.

To formalize this notion algorithmically, given a motif *without time or features* $M = (V, E)$, and similarity metrics over both time and edge attributes (d_T and d_F), we let $d(e, e') = d_T(T_e, T_{e'}) \cdot d_F(F_e, F_{e'})$, and we score all isomorphic subgraphs H using the following self-similarity score:

$$S(H) = \mathbb{E}_{e_h \neq e'_h \in H} \left[d(e_h, e'_h) \right].$$

This is shown diagramatically in Figure 4:

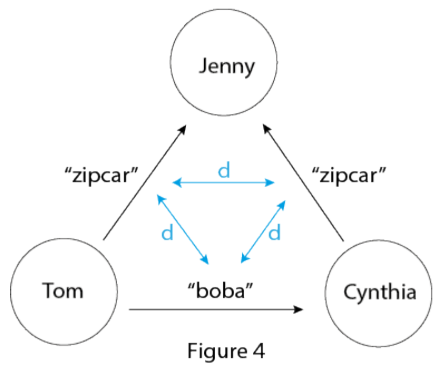

Figure 4: Diagrammatic view of the self-similarity score for a Venmotif

We expect motifs with particularly high or low self-similarity scores to correspond to different behaviours, and these can be discovered without specifying time annotations or attributes on the motif M. These will be our main focus in this paper.

4.4 Instantiations

There are still several ways that these "venmotifs" and "temporal motif with attributes" can be instantiated, depending particularly on two main factors that are so as of yet undefined:

1. **The motif enumeration method (or definition of motif) used.**
   As an example, one could use either the standard definition of motifs [6], or δ-temporal motifs [7], or communication motifs [12]. Motifs could be time-ordered or not time-ordered. This will also affect the tractability of the approach.
   In our experiments, we worked with both with standard and δ-temporal motifs.

2. **The similarity metric used.**
   One approach would be to simply enumerate motifs ignoring the temporal aspect, and to treat the time itself as another edge attribute. We can then define an (approximately normalized) similarity metric over time d_T, and an (approximately normalized) similarity metric over features d_F, and let the total similarity metric between the edges be the product of these two similarity metrics:

   $$d(e, e') = d_T(T_e, T_{e'}) \cdot d_F(F_e, F_{e'}).$$

   This could be used to find similar motifs that occurred around the same time, or with the same time differences, as the input motif M.
   Another approach would be to enumerate δ-temporal motifs as described by Paranjape et al [7], and then use just a similarity metric on edge features alone (so $d = d_F$), ignoring

the time in the similarity metric. This would result in motifs that are comparable to H, that occurred within a δ-window of time.

These are the two methods we work with in our experiments.

5 Experimental Methods

5.1 Dataset

Over the last several years, the Venmo API has been left fully accessible to the public. Some researchers have consequently scraped and publicised reasonably sized samples of the set of Venmo transactions. Notably, these transactions include a **message**. The most common in 2017 was: 🍕.

The original dataset is a MongoDB-exported dataset from GitHub [10] containing 7,076,585 transactions that occurred between July 26, 2018 and February 21, 2019. Each is labeled with elements such as payer and recipient user identification and profile photos, payment transaction message, timestamp, etc. [2]

While this is an interesting and novel dataset, we mostly choose to use it because we have some idea of behaviours that should exist in this graph, but are aware that other approaches to motif analysis (which don't take into account either time or edge features) cannot find them. We therefore mainly view it as an excellent dataset to validate our algorithms.

5.2 Network representation

The samples of Venmo transactions form a (sparsely connected) temporal multi-graph with edge attributes. To represent this data in a graph, we have a node for each user and edge for each transaction between two users. We unified payments and charges such that all edges are directed from payer to recipient. Each edge has the time and message as an attribute. We then defined both similarity metrics over time and the messages as follows.

5.3 Similarity Metric

In order to convert the messages on the edges into features with a similarity metric d_F, we use a metric embedding. We use a bag of words model to embed sentences, using a modified GLoVE [8] embedding. This includes fixing common or purposeful misspellings of words, which we defined using regex (such as '^u+b+e*r+(z|s)*$' $\rightarrow$ 'uber'), as well as converting emojis into descriptive plain text. Both of these steps are important for the Venmo dataset because the vast majority of the messages include either a misspelling or an emoji. We use word embeddings rather than a more complex sentence embedding (such as BERT), due to the short length of the messages and high frequency of emojis.

If any message has the vector 0 (meaning all words were not found), we instead hash that message to a random vector, so it will at least be similar to any other identical message. The final result is a 300-dimensional metric embedding for each of the approximately 7,000,000 edges, since GLoVE itself can be viewed as a metric embedding.

We define the similarity metric over vectors as the zero-clipped cosine similarity:

$$d_F(F_e, F_{e'}) = \max\left[0, \frac{F_e \cdot F_{e'}}{\|F_e\|\|F_{e'}\|}\right]$$

(noting that the cosine similarity is generally positive anyway due to the use of GLoVe vectors.)

Let ΔT be the difference between T_e and $T_{e'}$ in days. We define the similarity metric over time as:

$$d_T(T_e, T_{e'}) = \begin{cases} 1 & \Delta T < 1 \\ \frac{1}{\sqrt{\Delta T}} & \text{otherwise} \end{cases}$$

[2] In an attempt to respect the privacy of Venmo users, we've anonymized participants as far as we can without removing message contents. We fully parsed and cleaned this dataset, removing less interesting cases such as isolated pairs of users with only one transaction.

5.4 Finding three-edge temporal motifs

Experimentally, we primarily examine 3-edge motifs. We define a *3-edge* motif as a recurring subgraph containing three distinct edges, ignoring time ordering. This definition differs from traditional methods of classifying motifs by number of nodes [9], but is the same definition used in defining δ-temporal motifs. [3]

We note that 3-edge motifs can be fairly interesting and complex in multi-graphs (compared to regular graphs), and can include a number of different behaviours, due to the possibility for multiple edges between nodes:

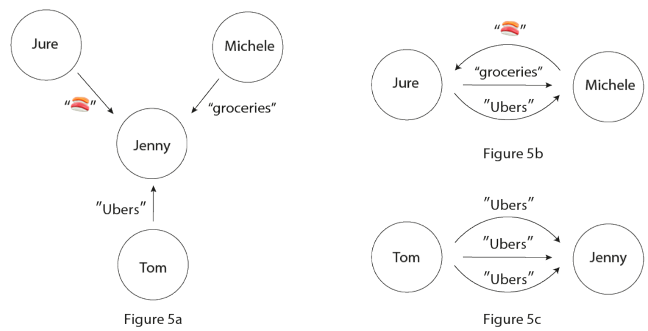

Figure 5: Examples of highly different 3-edge motifs in a multi-graph with messages

Due to our non-standard definition, we wrote our own code to enumerate all 3-edge motifs and 3-edge δ-temporal motifs—this code is in itself novel, as Paranjape et al only deal with counting motifs [7]. However (as in some similar papers on motif analysis) we chose to ignore “line" motifs (consisting of three nodes in a line, with edges in any direction) in this process, which are generally of less interest and also highly intractable to enumerate due to their high frequency in null models. We also choose to ignore ordering in our definition of motif.

We compare to the switching model as our null model, with 300,000,000 edge switches.

6 Results

Our initial experiments are with **Venmotifs**. Our first experiment is using standard motif enumeration, and a similarity metric over both time and messages. Enumeration of all 3-edge motifs turns out to be highly intractable. Regardless, we can uniformly sample a large number of motifs (by uniformly sampling a set of nodes, then taking all motifs including that node). The results are plotted in Figure 6.

Triangle motifs are near negligible in this sample (and thus excluded). Certain star motifs are over-represented—particularly inward star motifs, and the motif consisting of two users paying a single person, but one paying twice. Notably, all motifs except for the four star motifs are negligible in the null model. However, the similarity score distributions are generally an unremarkable, negative result that is comparable to our null model. The only significant difference is outward star-motifs, which are somewhat less self-similar in the null model. By sampling the scores, one finds that the vast majority of motifs have edges that are very far apart in time, so are very uninteresting. This indicates that the majority of motifs are meaningless as user behaviours (due to the large time differences), so constraining to temporal motifs may be useful.

[3]This is a necessary decision, because in multi-graphs like the Venmo graph, there are theoretically an infinite number of possibilities for n-node motifs, because these n nodes may have any of $k \geq n-1 \in \mathbb{N}$ number of edges between them.

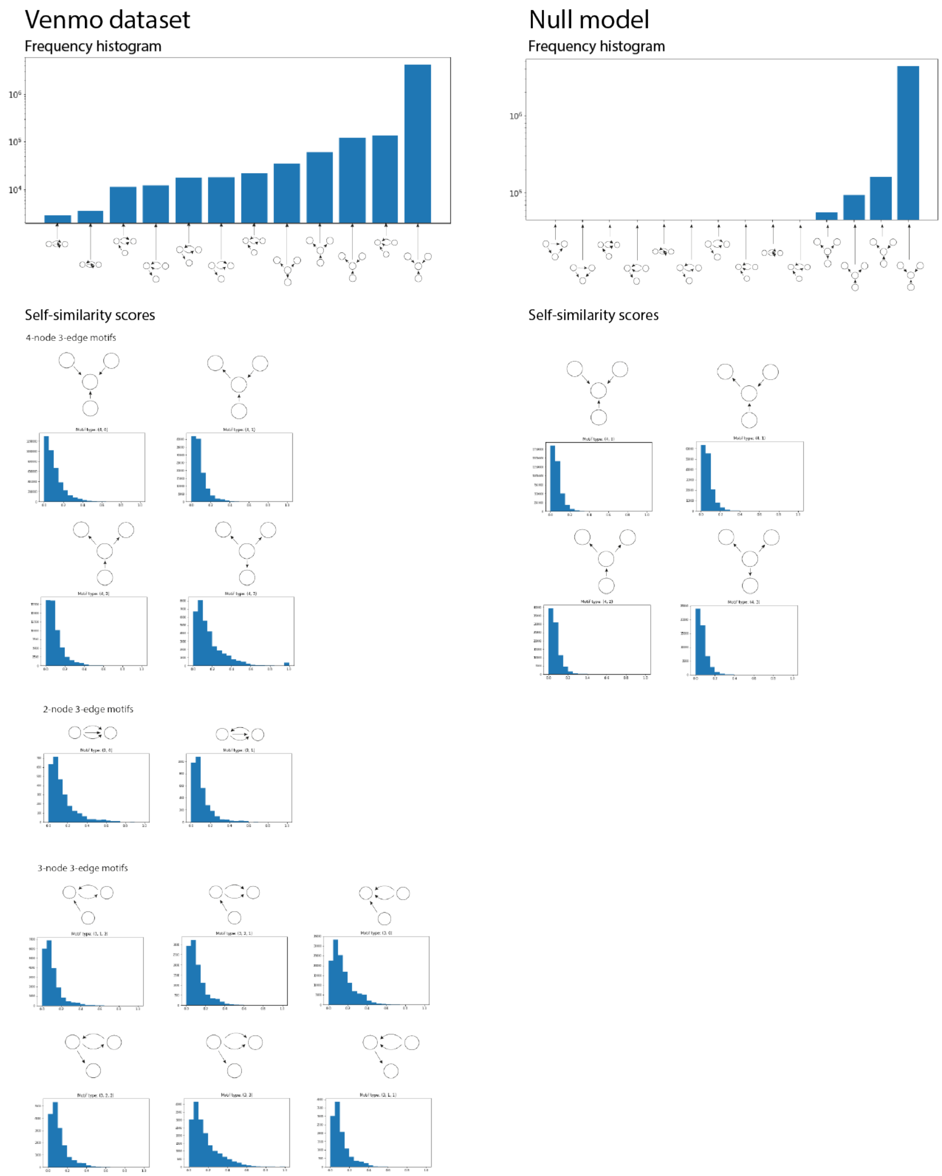

Figure 6: Results of motif frequency and self-similarity, using standard (non-temporal) motif enumeration, and a similarity metric over both messages and time.

Our algorithm definition allows for some flexibility in method, so we now experiment with δ-temporal motif enumeration (with $\delta = 1$ day), and using a similarity metric over just the message vectors. This is more tractable, and the result (plotted in Figure 7) is much more interesting—this indicates that the level of flexibility in our algorithm definition is in fact useful.

The same motifs are over-represented as before. However, using temporal motifs, we now see that a few motifs have significantly different score distributions from the null model. Most notably, **fully inward and outward star motifs** (red dotted outlines) are very differently distributed from the null model, indicating that these motifs likely serve a different purpose in the graph. We also observe that most of the score distributions in the Venmo dataset are shifted to be somewhat somewhat more "self-similar" than in the null model—particularly the motifs with two nodes, as well as the (notably over-represented) motif consisting of two users paying one other user (orange dotted outlines). Several motifs have interesting spikes of highly self-similar scores.

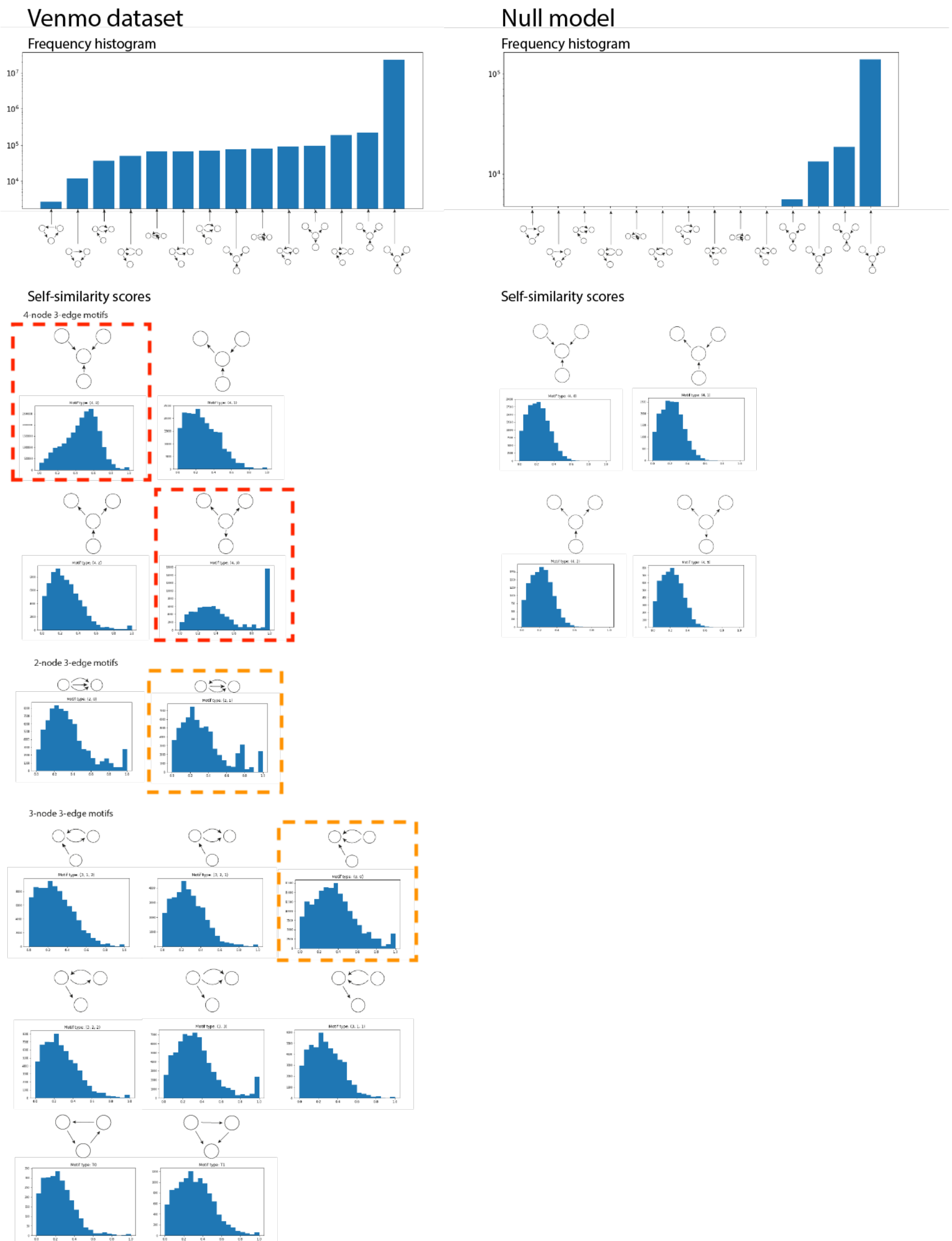

Figure 7: Results of motif frequency and self-similarity, using $\delta-$temporal motif enumeration ($\delta = 1$ day), and a similarity metric over both messages. Some particularly significant score distributions are highlighted. The most significant differences from the null model are outlined in red, while slightly less significant differences are outlined in orange.

To further examine the meaning of these different distributions, we qualitatively examined the mode of both of the score distributions for the **fully inward and outward star motifs**, to find what behaviour they represent.

For inward stars, the mode of the score distribution has similarity scores of around 0.55-0.65, and almost always corresponds to three people paying a central user for the same thing, with slightly different messages. For example, a fairly common contributor to the mode is people donating to a central fund for rescue dogs, often mentioning dogs or including dog emojis. The sentences used aren't exactly identical, but have similarity scores of around 0.6. For example, the following three messages appear on an inward star with a score of 0.57:

```
For the pups 🙌🏼🙌🏼🙌🏼
From a fellow dog rescuer, thank you ❤️🐶
My rescue pup is asleep on my knee as I type this so I gotta pay!
```

This is a common behaviour that would have been fairly hard to find and isolate without our specific algorithm. This also really indicates the different purpose served by the inward star motifs, in a way that previous motif analysis would not have been able to: **people paying a central user for the same thing, at around the same time.** This a behaviour we'd expect to find in the Venmo dataset, but prior motif analysis methods would fail to meaningfully find.

We can apply the same approach to the mode of the outward star scores (with scores of between 0.95-1.0), to discover a slightly different behaviour. Often, these are companies or users that are using Venmo to issue commissions or refunds to other users, and using nearly exactly the same message each time.

For example, two motifs that appear are:

```
Commission payment for selling hair on Mayvenn.com
Commission payment for selling hair on Mayvenn.com
Commission payment for selling hair on Mayvenn.com
```

```
Intern lunch 💵🔙 #expensedthatshiz
Intern lunch 💵🔙 #expensedthatshiz
Intern lunch 💵🔙 #expensedthatshiz
```

In fact, the words "cashback", "refund", "forgot" and "reimbursement" are all very over-represented within this mode (appearing between 7-70x more often than on a random edge) indicating that this behaviour really does have some correspondence with people giving refunds to several people.

With this knowledge, we then apply our definition of **temporal motifs with attributes**, which can be used to discover all instances of any given temporal motif. For example, we can calculate similarity scores against the outward star motif with the message from Mayvenn.com on all edges. This gives the following distribution of similarity scores:

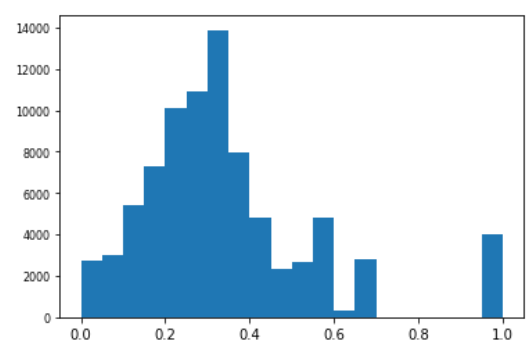

Figure 8: Distribution of similarity to the inward star motif with Mayvenn.com message on all edges, for all inward star motifs in the Venmo dataset

Examining the spikes allows us to quantify approximately how much of the inward star motif score distribution are made up of motifs with Mayvenn's messages—in fact around 4000 of the motifs feature Mayvenn's message on all three edges (the far left spike), while around 3000-7000 have Mayvenn's message on two edges (the two spikes in the middle). Interestingly, this indicates that Mayvenn also makes other transactions with different messages. We can again compare this to other motifs in the Venmo dataset, as well as in the null model, to show the statistical significance. This fairly clearly shows that this message is only gives a significant score distribution for this outward star motif, and is insignificant in the null model:

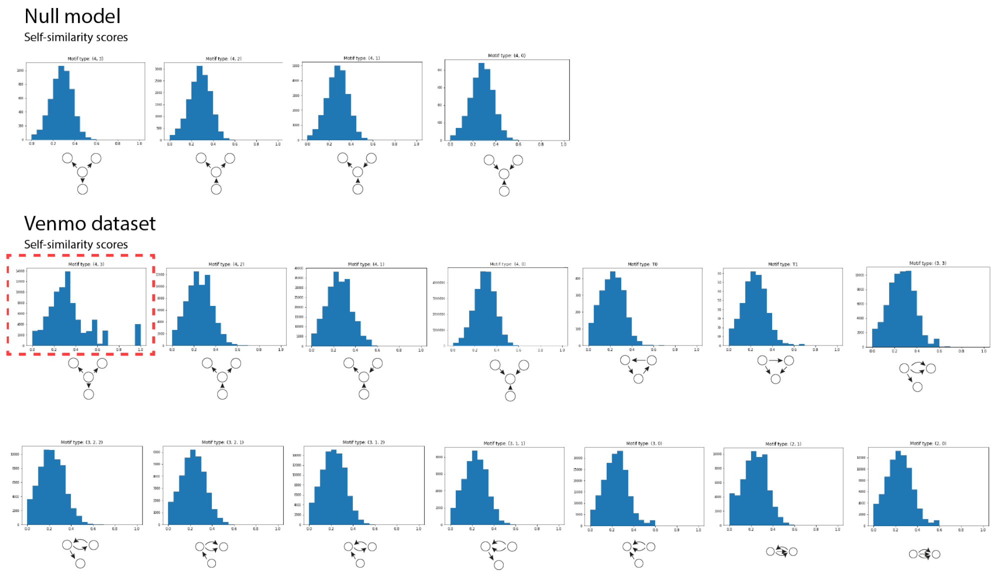

Figure 9: Distributions of similarity scores against the corresponding δ-temporal motif with the Mayvenn message on all edges

One flaw in this method is that certain behaviours we might expect to find using "temporal motifs with attributes"—(such as an inward star with the pizza or taxi emoji on all edges)—fail to give significantly different score distributions than the null model. Even though these behaviours do appear (and motifs we find with particularly high similarity scores do in fact correspond to these behaviours), the individual motifs are simply too rare and sparse to give a score distribution that is very significantly different from the null model. This is exacerbated by the fact that the Venmo dataset we have is a comparatively sparse, and likely non-uniform, sample of transactions.

Therefore, only fairly commonly repeated behaviours (such as businesses repeatedly sending the same message from the same node) are easily discovered. We view finding sparse but still significant behaviours as future work.

7 Discussion and future work

Our results provide insight into the structure of the Venmo dataset, providing significant evidence that (temporal) inward star motifs often correspond to people paying a user for the same things, while outward star motifs correspond to refunds from users, as well as commission payments by companies using Venmo. Understanding patterns in payment history may be useful to Venmo if the company plans on tweaking or building new features for its platforms or simply analyzing its customer base's needs.

More interestingly, however, our results provide new general approaches for including edge features in motif analysis, which can fairly easily be added to any algorithm performing motif enumerations, and which do not significantly affect runtime.

Additionally, note that the similarity scores give useful results using δ-temporal motifs, and less useful results using standard motifs. In our view, this provides some additional validation for the work of Paranjape et al, who suggest that δ-temporal motifs are able to provide a more useful analysis for temporal graphs, but have fairly little applied evidence for this hypothesis [7].

Regarding further work, we believe that our generalization could be expanded on to find monthly, recurring interactions in Venmo (as shown in Figure 2a) but have yet to implement this in full. The ability to find this would be further evidence for the usefulness of motif definitions which can take into account edge features. We also expect that improvements could be made to our approach to allow sparser behaviours to be found.

References

[1] Uri Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, 2007.

[2] Austin R. Benson, David F. Gleich, and Jure Leskovec. Higher-order organization of complex networks. *Science (New York, N.Y.)*, 353(6295):163–166, Jul 2016. 27387949[pmid].

[3] Indika Kahanda and Jennifer Neville. Using transactional information to predict link strength in online social networks, 2009.

[4] Gueorgi Kossinets, Jon Kleinberg, and Duncan Watts. The structure of information pathways in a social communication network. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 435–443, New York, NY, USA, 2008. ACM.

[5] Lauri Kovanen, Márton Karsai, Kimmo Kaski, János Kertész, and Jari Saramäki. Temporal motifs in time-dependent networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(11):11005, Nov 2011.

[6] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[7] Ashwin Paranjape, Austin R. Benson, and Jure Leskovec. Motifs in Temporal Networks. *arXiv e-prints*, page arXiv:1612.09259, Dec 2016.

[8] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.

[9] Natasa Przulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):e177–e183, Jan 2007.

[10] Dan Salmon. Venmo transaction dataset, 2019.

[11] Lewi Stone, Daniel Simberloff, and Yael Artzy-Randrup. Network motifs and their origins. *PLOS Computational Biology*, 15:e1006749, 04 2019.

[12] Qiankun Zhao, Yuan Tian, Qi He, Nuria Oliver, Ruoming Jin, and Wang-Chien Lee. Communication motifs: a tool to characterize social communications. In *CIKM*, 2010.