Temporal Patterns of Motifs

Sameera Horawalavithana

May 1, 2017

Abstract

A social network evolves over time through the creation or deletion of ties among a set of actors [22].The volatile nature of social ties provides a strong platform to identify the dynamic community structure. This change of structural patterns can be well represented by the existence of motifs (or graphlets). Further, they can be enriched with the temporal information of social ties to define the recurrent subgraphs of interest. This would yield important insights about the correlation between patterns of ties in a social network [12]. Major contribution of our study is the analysis of dynamicity over the social ties via motifs.

Contents

1 Introduction

Modeling social interactions by temporal properties has been attracted as a rich characterization of social ties. Temporal patterns provide key insights to identify inherent properties of social ties, since, they are usually packaged with the tightly coupled duo: network structure and dynamic. Hence, social networks can be viewed as highly complex dynamic entities which facilitate the instantaneous nature of human activity patterns [12].

When characterizing temporal properties of a social network, most of the studies consider persistent models by aggregating temporal information into an aggregated snapshot (LSN 1 - Figure 1) [19, 1]. Hence, these techniques fail to capture most of state changes in a dynamic network. Thus they lose rich corpus of information in the go, which leads to bring insufficient insights about the network [14]. Thus, it's required to analyze the network behavior of nodes in their structural positions and participation to explore dynamicity of an evolving network [22].

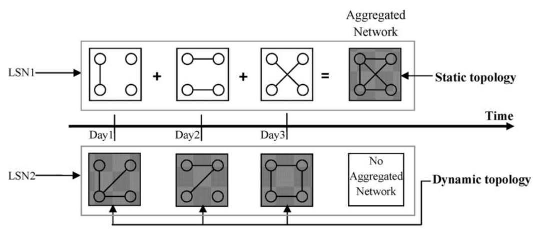


Figure 1: Static and dynamic topologies of an evolving longitudinal social network (LSN) [22]

A network which spans over multiple short time intervals can be well studied by modeling it as a dynamic topology (LSN 2 - Figure 1). Many dynamic structural properties emerge over the evolving network, which can be understood by studying position and participation dynamicity measures [22]. In our study, we focus on the position dynamicity, which models the change of node's structural positions in the network, specifically via motifs (or graphlets).

The notion of motif is well established in static networks, which is defined as the small induced subgraph patterns of the original structure of the underlying network [2]. Further, it has been extended to capture brokerage positions (i.e. orbits) in subgraph patterns, which gives a detailed view of the network structure local to a node specification [13].

Motivation: "For static networks, a triangle is a triangle, and such subgraphs can readily be counted. For temporal networks, we first have to define what a triangle is; here, we have defined temporal subgraphs based on the time adjacency of events sharing nodes." [8]

As an example, without the notion of temporal motifs, we would fail to describe a triangle relation of three individuals in the context of temporal causality, where we can not describe the relationship of Alice with Carl, is due to the common friend Bob or by chance with no former communications. Also, the mesoscale structure of temporal network can be well represented by motifs which describe the topological change over time.

2 Temporal Motifs

In our study, we are based on a recent specification of temporal motifs, that defines the subgraph patterns which consists of all edges observed within a δ time-units in a dynamic network [14].

Definition 2.1 (Temporal graph)**.** Let V be a set of nodes, such that $\forall v \in V$, and E be a set of temporal edges, such that $\forall(u_i, v_i, t_i) \in E$, where $\forall v_i$ are instances of $v \in V$, and $t_i \in \mathbb{R}$ depicts the time-stamp of edge formation. Also t_i could be extended to $d_i = t_i^l - t_i$ to define the duration of an active edge, where t_i^l define the dissolve time of an edge u_i, v_i. Thus, $(V, E) \in G_d$ defines a temporal graph within a observation time window $d \geq \forall d_i, i = 1, ...m$, such that m is the number of edges.

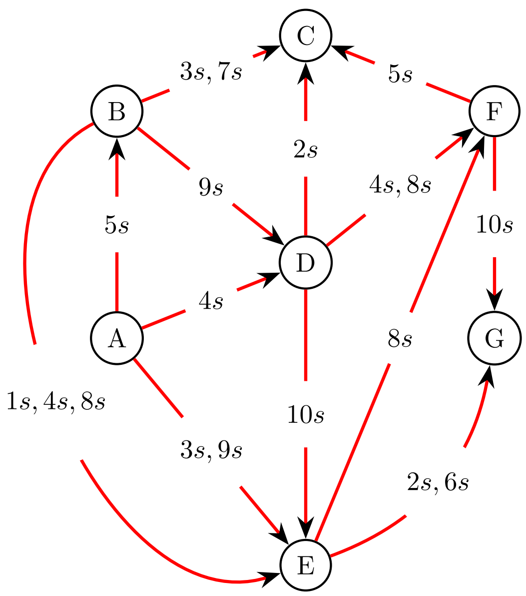


Figure 2: Sample Temporal Graph (G)

Based on the definition 2.1, temporal graph would be a multi-graph which consists of many temporal edges between any two nodes $u, v \in V$, and such edges are strictly ordered based on the timestamp t_i attached with an edge (Figure 2). Note that we can induce the underlying static graph, by avoiding the temporal property.

Definition 2.2 ($\delta-$temporal motifs [14])**.** A k-node, q-edge, temporal motif is a sequence of q edges, $M = (u_1, v_1, t_1), ..., (u_q, v_q, t_q)$ that are time-ordered within a duration, i.e. $t_1 < t_2 ... < t_q$ and $\delta \geq t_q - t_1$, such that the induced static graph from the edges is connected and has k nodes.

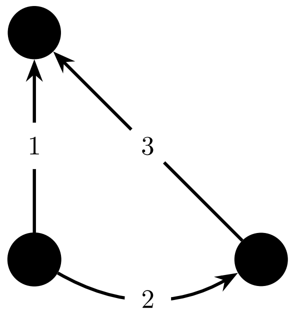


Figure 3: $M_{4,6}$

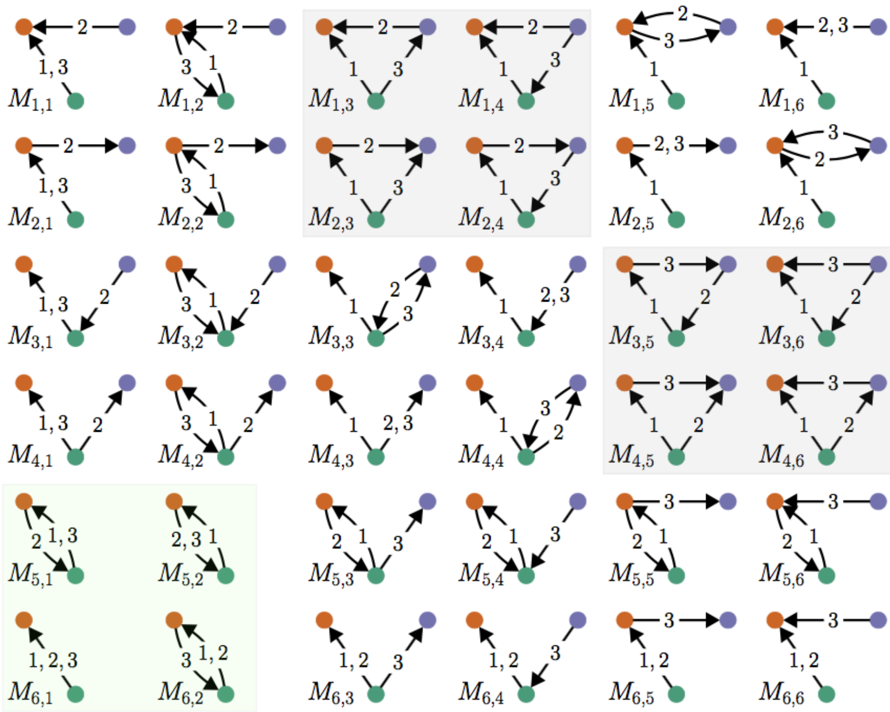


Figure 4: Temporal motif structures $k = 2, 3;\ q = 3;$

Based on the definition 2.2, temporal motif defines a pattern of subgraph (i.e. multi-graph), which consists of ordered edges on timestamps (Figure 4). In the sense, any temporal graph in a observation window could be a instance of a given temporal motif structure, as it follows the $k - node, q - edge$ in the given order of edge pattern. While the definition 2.2 is defined for directed temporal graph with timestamped edges, it can be specialized for other variations too (e.g. undirected, signed network).

In general, Paranjape et al. [14] explain that any time-ordered sequence $S = (x_1, y_1, t'_1), ..., (x_q, y_q, t'_q)$ of q unique edges is an instance of the motif $M = (u_1, v_1, t_1), ..., (u_q, v_q, t_q)$ if,

- $\exists f$ bijective function on the vertices such that $f(x_i) = u_i$ and $f(y_i) = v_i, i = 1, ..., q$, and
- the edges all occur within the time, i.e., $\delta \geq t'_q - t'_1$

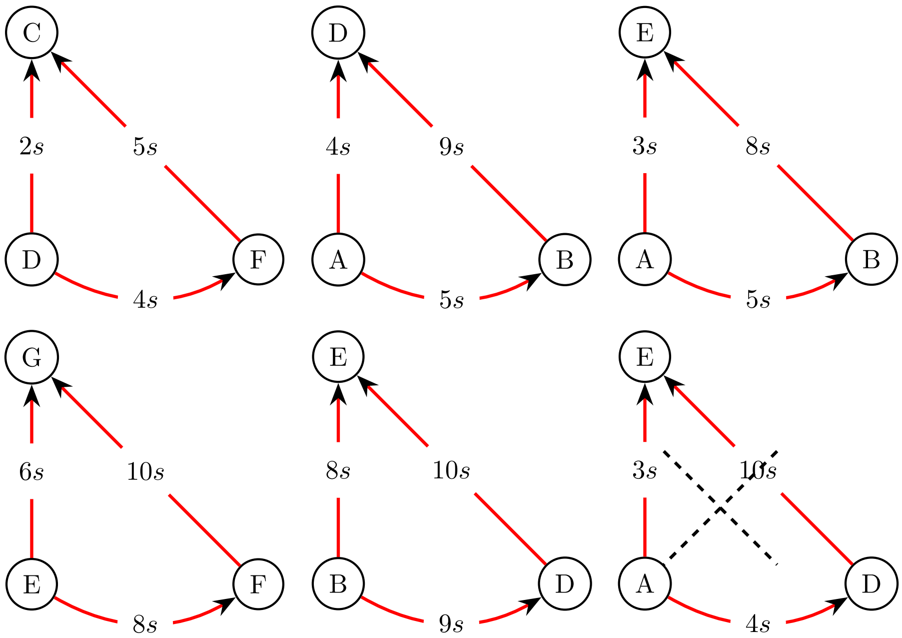


Figure 5: Sample Instances of Motif Structure ($M \subset G$), $\delta = 5s$

Figure 5 demonstrates the instances of $\delta-$temporal motifs ($k = 3, q = 3, \delta = 5$) for the sample graph G as shown in Figure 2.

3 Measures

3.1 Motif Count

Motif degree count defines the frequency of appearance of the given motif in a network. The interesting observation would be to measure the frequency of temporal motifs in different time scales. Many insights could be demonstrated about the formation of motifs, which are discretized along the lifespan of the observed pattern.

As an example, the motif M as given in Figure 3, can be seen 5 times within a $\delta = 5$ observation window in the sample temporal graph (Figure 5). Further, we used snap[1] package to count the number of several temporal motifs in our sample network. Results are demonstrated at Table 1 and 2 for $\delta = 5$ & 10 respectively for the sample network.

[1]http://snap.stanford.edu/temporal-motifs/index.html

$M_{1,1} = 2$	$M_{1,2} = 0$	$M_{1,3} = 2$	$M_{1,4} = 1$	$M_{1,5} = 0$	$M_{1,6} = 2$
$M_{2,1} = 4$	$M_{2,2} = 0$	$M_{2,3} = 1$	$M_{2,4} = 0$	$M_{2,5} = 1$	$M_{2,6} = 0$
$M_{3,1} = 0$	$M_{3,2} = 0$	$M_{3,3} = 0$	$M_{3,4} = 0$	$M_{3,5} = 0$	$M_{3,6} = 1$
$M_{4,1} = 3$	$M_{4,2} = 0$	$M_{4,3} = 1$	$M_{4,4} = 0$	$M_{4,5} = 0$	$M_{4,6} = 5$
$M_{5,1} = 0$	$M_{5,2} = 0$	$M_{5,3} = 0$	$M_{5,4} = 0$	$M_{5,5} = 0$	$M_{5,6} = 0$
$M_{6,1} = 0$	$M_{6,2} = 0$	$M_{6,3} = 2$	$M_{6,4} = 2$	$M_{6,5} = 1$	$M_{6,6} = 2$

Table 1: Counts of instances of temporal motifs with $k = 2, 3; \delta = 5s$. Counts in the i^{th} row and j^{th} column is the number of instances of motif $M_{i,j}$ (see Figure. 4);

$M_{1,1} = 5$	$M_{1,2} = 0$	$M_{1,3} = 4$	$M_{1,4} = 3$	$M_{1,5} = 0$	$M_{1,6} = 3$
$M_{2,1} = 9$	$M_{2,2} = 0$	$M_{2,3} = 1$	$M_{2,4} = 0$	$M_{2,5} = 1$	$M_{2,6} = 0$
$M_{3,1} = 0$	$M_{3,2} = 0$	$M_{3,3} = 0$	$M_{3,4} = 2$	$M_{3,5} = 0$	$M_{3,6} = 2$
$M_{4,1} = 7$	$M_{4,2} = 0$	$M_{4,3} = 3$	$M_{4,4} = 0$	$M_{4,5} = 1$	$M_{4,6} = 9$
$M_{5,1} = 0$	$M_{5,2} = 0$	$M_{5,3} = 0$	$M_{5,4} = 0$	$M_{5,5} = 0$	$M_{5,6} = 0$
$M_{6,1} = 1$	$M_{6,2} = 0$	$M_{6,3} = 8$	$M_{6,4} = 5$	$M_{6,5} = 3$	$M_{6,6} = 9$

Table 2: Counts of instances of temporal motifs with $k = 2, 3; \delta = 10s$. Counts in the i^{th} row and j^{th} column is the number of instances of motif $M_{i,j}$ (see Figure. 4);

3.2 Higher order organization

Motifs provide rich extensions for analyzing higher order organization of a network. Instead of edges, we adopt a motif based clustering to present higher order connectivity patterns [2].

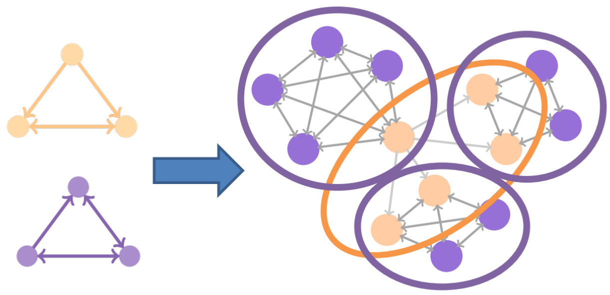


Figure 6: Analyzing modules from motifs [2]

We discuss about the vast space of possible motif instances that could be generated from our definition 2.2. As higher order organization modules are dependent on the small subgraph patterns, such space of motifs could lead us to have more fine grained optimal modules or community structures in the network (Figure 6).

Benson et. al [2] introduce a generalized framework for clustering based on a new metric of motif conductance.

Definition 3.1 (Motif conductance metric [2])**.** Given a motif instance M, find a set of nodes S that minimizes

$$\phi_M(S) = \frac{cut_M(S, \overline{S})}{min[vol_M(S), vol_M(\overline{S})]} \tag{1}$$

where $\overline{S} = V/\ S$, $cut_M(S, \overline{S})$ is the number of instances of motif M with at least one node in S, and one in $\overline{S}$, and $vol_M(S)$ is the number of nodes in instances of M that reside in S.

Hence, this grouping needs to minimize the given motif cuts, while having most of the edge points in the clusters (Figure 7).

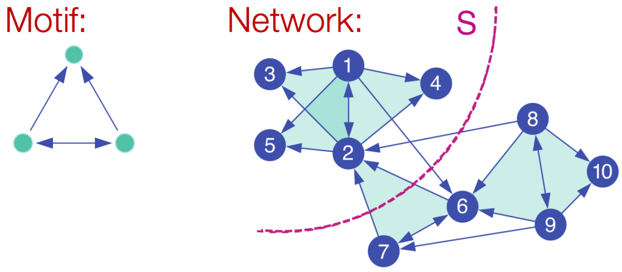


Figure 7: Clustering based on motif conductance metric [2]

We use snap-higher-order[2] python package to find clusters of low motif conductance, which is implemented based on motif spectral clustering methods. Snap is only capable on static graphs currently, therefore we induce static graphs from all the temporal versions to study the near optimal higher order organization.

Figure 8 lists the motif types and naming conventions supported by snap[3] currently. We use them for our analysis on higher order structures.

[2]http://snap.stanford.edu/higher-order
[3]http://snap.stanford.edu/higher-order/code.html

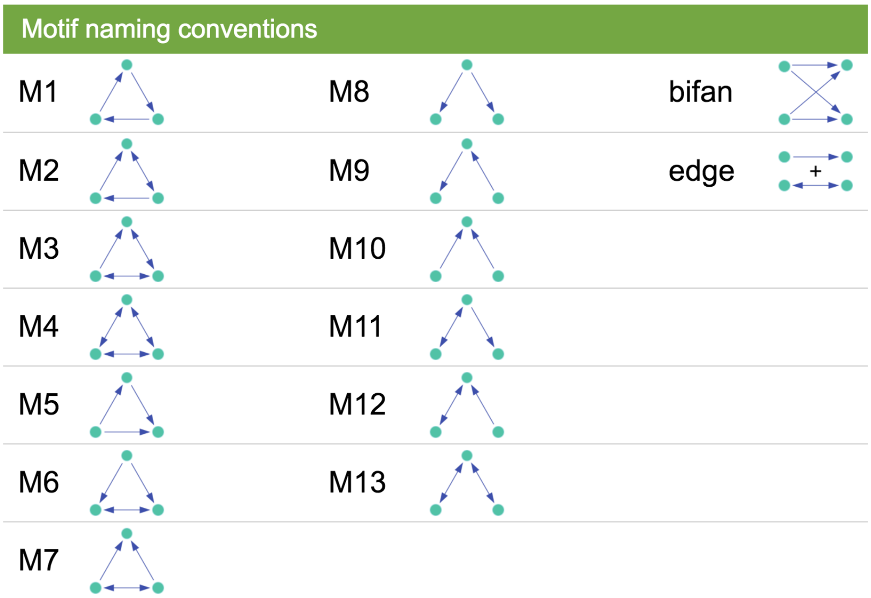


Figure 8: (Static) Motif naming conventions for higher order organization

As an example, given the motif M7, the network shown at Figure 7 can be partition into two sets, $S = \{1, 2, 3, 4, 5\}$ and $\overline{S} = \{6, 7, 8, 9, 10\}$ which has the lowest motif conductance value 0.125 (Table 3).

Motif	Largest CC size	Cluster size	Motif conductance in largest CC	Eigenvalue
M7	10	5	0.125000	0.115140
M8	10	5	0.285714	0.279742
M9	8	4	0.750000	0.855662
M10	6	3	0.250000	0.316987
M11	4	2	0.666667	1.000000
M12	5	2	0.600000	0.903144

Table 3: Higher-order organization

3.3 Brokerage positions

We would like to extend our focus to analyze brokerage positions in temporal motifs. Such positions were established in the context of static motifs, as known as orbits [13].

Definition 3.2 (Orbits [23])**.** Two nodes within a network are said to belong to the same automorphism orbit (or automorphic equivalence class) if there exists a relabeling of nodes in the graph that exchanges the two nodes while preserving the graphs adjacency structure.

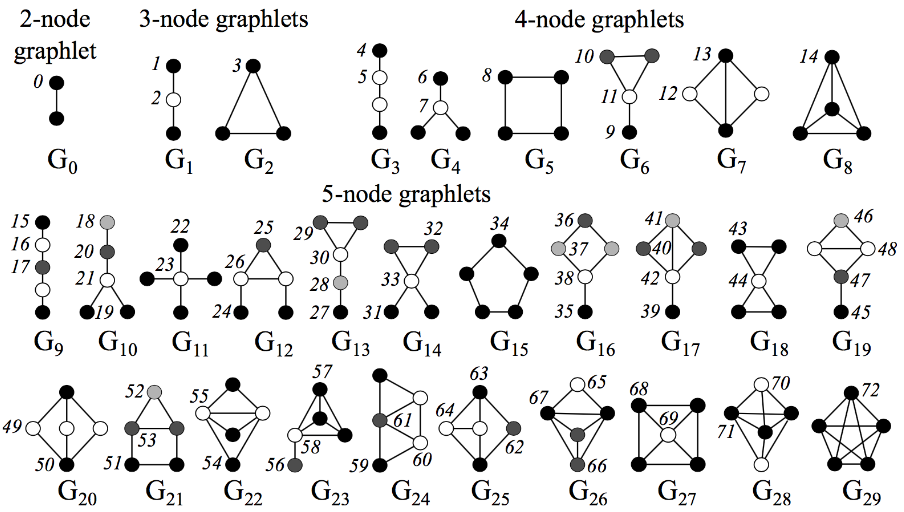


Figure 9: All undirected motifs (or graphlets) $G_0, ..G_{29}; k = 2, ..5$, and respective orbits $0, ..72$ [15]

Figure 9 visualizes all the possibilities of motifs in the given space of $k = 1, ..5$ nodes, and the respective orbits in a static context.

Temporal Orbits When defining orbits in the context of temporal motifs, we need to study the problem of multi-graph automorphism. In the meanwhile, [4] shed light on finding the automorphism groups of multigraphs to the case of irreducible multigraphs, that is to multigraphs having no twin vertices. We would like to follow the automorphism groups as proposed by the definition 3.3 in our study.

Definition 3.3 (Multigraph automorphism [4])**.** Let G be a multigraph. For any $u, v \in V$, the multiplicity $\mu(u, v)$ is the number of edges (possibly 0) having u and v as endpoints. An automorphism of G is a one-to-one mapping $g : V \to V$ such that $\mu(g(u), g(v)) = \mu(u, v)$ for any $u, v \in V$. The automorphisms of G define the automorphism group $Aut(G)$ of G.

Since, temporal motifs are strictly defined by the order of edge patterns, we guess that a vast space of temporal orbits could be explored, given the possibility of generating $q!\binom{m}{q}$ ordered length of q sequences of edge patterns in a temporal graph.

3.4 Motif Degree

In the static context, motif degree of a node reflects the number of motifs that it participates at a given orbit. Thus, we could have a vector of motif degree which is local to a node in the given range of k (e.g. $k > 1$). Hence, **motif degree vector** is a generalization of degree (i.e. when $k = 2$), which details about the local view of the network structure.

Temporal Motif Degree The definition of temporal motif degree can be parameterized by δ, and would provide a dynamic insight about the change in temporal orbits in a evolving network.

4 Analysis

We analyze several temporal and static datasets (Table 4) by observing the time it takes for motifs to form. In general, we group the frequency of motifs into time bins of $[60(i-1), 60i]$ seconds, which represent the difference of motif counts with $\delta = 60i$ seconds minus the count with $\delta = 60(i-1)$ seconds.

Dataset	No. of nodes	No. of static edges	No. of temporal edges	Timespan (Days)
StackOverflow (Answers to Questions)	16,266,395	2,464,606	17,823,525	2774
StackOverflow (Comments to Answers)	1,646,338	11,370,342	25,405,374	2774
Wiki-Talk	1,140,149	3,309,592	7,833,140	2320
Bitcoin (Active users)	1288	-	7255	600
CollegeMsg	1,899	59,835	20,296	193
Amazon	334,863	925,872	-	-
DBLP	317,080	1,049,866	-	-

Table 4: Basic statistics on datasets, Note that datasets without temporal edges are static, otherwise we induce static version as required.

4.1 Stack-Overflow

Stack-overflow is an online portal where registered users can interact with others, by posting questions, answering them and commenting on both questions and answers. We consider several datasets from the recent stack-overflow dump for our analysis [9].

In the temporal network, users depicts nodes while any possible interaction between users (e.g. answer or comment on others' questions etc.) represent ties. They are directed and timestamped, span over 2773 days.

4.1.1 Type of interaction: answers to questions

Stack-overflow questions are long lived, and span over long periods due to the burst nature of interest in the community. However, we limit our analysis the temporal motifs with $\delta = [0, 3600]$ *secs.* to demonstrate the behavior of human conversation patterns.

Temporal Dyadic Behavior Figure 12 demonstrates the number of instances observed on the given dyadic instance of temporal motif against the time they spent to complete. It includes all the possibilities of dyadic behaviors (M(5,1),M(5,2),M(6,1),M(6,2)) from our temporal motif instances (see Figure 4). They reflect the behavior of answering patterns for any question between two users.

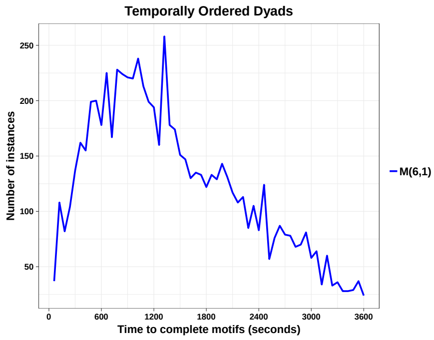


Figure 10: Dyadic behavior $M_{6,1}$

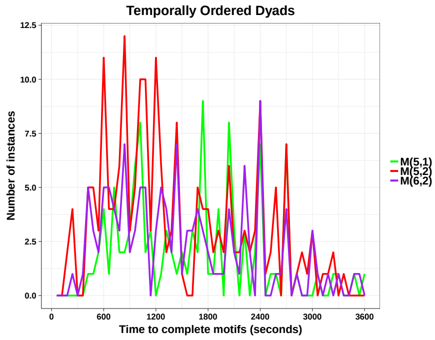


Figure 11: Dyadic behavior $M_{5,1}, M_{5,2}, M_{6,2}$

Figure 12: Temporal Ordered Dyads

Figure 10 represents the continuous feedback by an user answering to a question from another user. It's clearly noted that such any user tend to answer the same question until 20-25 minutes frequently, where we see a significant decline of the time to complete M(6,1) motif in the middle. In fact, this is an observation by Ubaldi et al. [21] which also conclude that burstiness in human communication is mostly a link property between two people.

By comparing with the continuous feedback pattern from a user, it's to hard to see a reciprocity nature in the order of communication, since the frequency of motif instances $M_{5,1}, M_{5,2}, M_{6,2}$ does not dominate over. Nevertheless, it's good to see the continuous responses by the party who's beneficial from getting an answer to her question, since M(5,2) has the average upper bound (Figure 11). This might reveal the generosity nature of humans.

Temporal Triadic Behavior We partition the possibilities of temporal triangles into two classes by the participation of the originator to close the ordered triad. As an example, M(1,3) reflects that the originator participates on closing the triad with a direct edge, and M(3,5) reflects the vice-versa. Figure 13 and 14 demonstrates the behavior of these classes in the given order, which demonstrate the number of instances observed on the given triadic instance of temporal motif against the time they spent to complete.

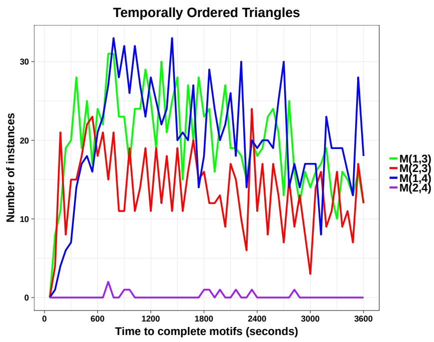


Figure 13: Originator participation = yes

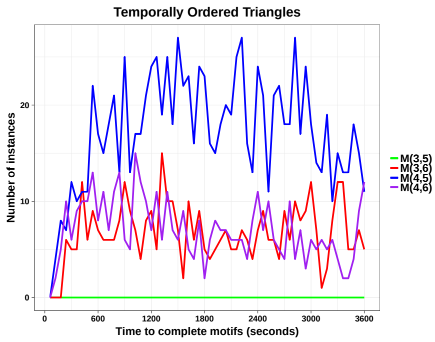


Figure 14: Originator participation = no

Figure 15: Temporal Ordered Triads

In both classes (Figure 15), we could see that the cyclic triangles (M(2,4) and M(3,5) have a significant deviation from other patterns. This is a usual breed to happen in the context of online forums, hence it's rare to find such conversations.

The growth over other patterns are identical, when the originator is participating on closing the ordered triad (Figure 13), and they reach the same frequency at the end of the period. The higher growth of M(4,5) in Figure 14, represents two users tend to participate on an unknown user's question. This could be the fact that recent sessions of questions motivate users to involve on answering in the same context. Because, it's highly unlikely two users share another question by answering if they don't have a previous experience recently.

Non-blocking interaction behavior In an online forum like stack-overflow, it's important to have close attention from all the users to have a clear resolvent for an answer. Such that, one would not wait for another to proceed on the answers or comments, but try to give the feedback as frequent as possible. We model this behavior as non-blocking motifs, which is represented by the temporal motif instances of M(4,1), M(4,3) and M(6,3).

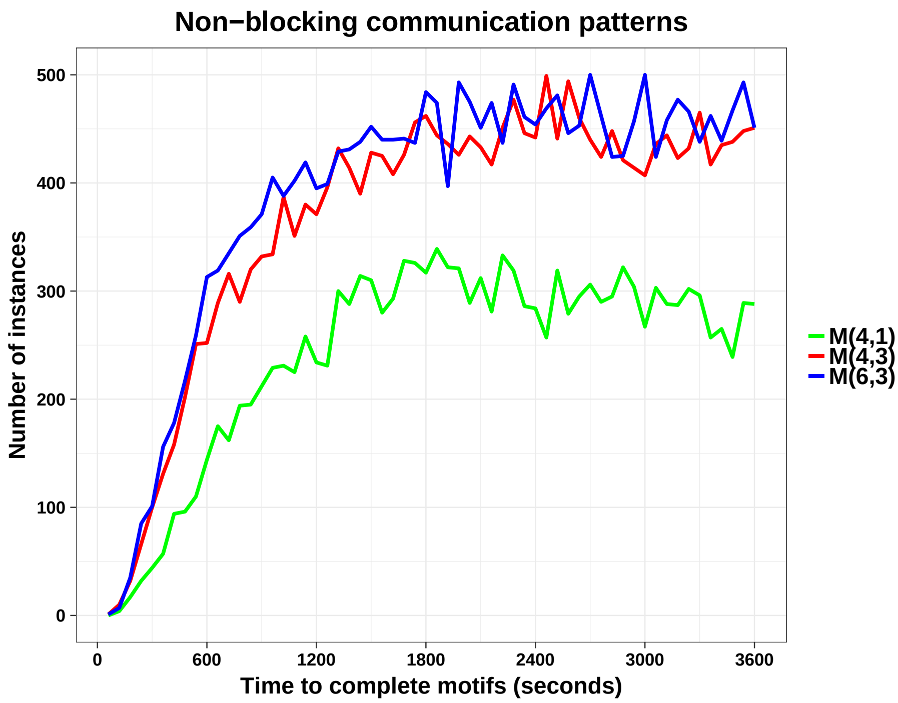


Figure 16: Non-blocking behavior

M(4,3) and M(6,3) are dominant on reaching an identical higher growth on completing motif instances, while M(4,1) has shown a relative low growth, and also a decline at the end. When answering questions in the forum, it's not common to go back and forth to the same question due to the fact of novelty of other answers in the thread.

4.1.2 Type of interaction: comments to answers

In stack-overflow forum, any user could clarify, modify or suggest an answer by proceeding with a thread of comments. We observe that comments are actively made by users for an answer within a very short interval, which is 5 minutes approximately. Hence, we limit our analysis to the period δ of [0,300] to analyze the behavior of temporal motifs.

Temporal Dyadic Behavior Figure 19 demonstrates the number of instances observed on the given dyadic instance of temporal motif against the time they spent to complete. It includes all the possibilities of dyadic behavior (M(5,1),M(5,2),M(6,1),M(6,2)) from our motif representation (Figure 4). They reflect the behavior of commenting patterns for any answer between two users.

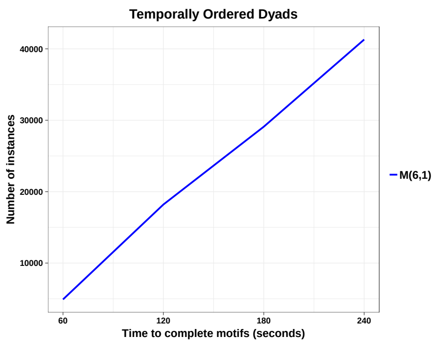


Figure 17: Dyadic behavior $M_{6,1}$

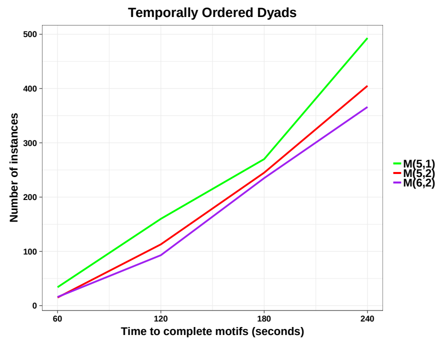


Figure 18: Dyadic behavior $M_{5,1}, M_{5,2}, M_{6,2}$

Figure 19: Temporal Ordered Dyads

M(6,1) represents a continuous feedback by an user for other user's answer, which is a significant large growth over the life span. This pattern is clearly dominated over other dyadic patterns (M(5,1),M(5,2),M(6,2)) which require bi-directional feedback by both users to continue.

It's interesting to see M(5,1) has relatively higher growth by comparing with M(5,2),M(6,2), which reflects the normal conversation behavior, where any comment is blocked until a response by other party.

Temporal Triadic Behavior Figure 22 demonstrates the number of instances observed on the given triadic instance of temporal motif against the time they spent to complete.

Figure 20: Originator participation = yes

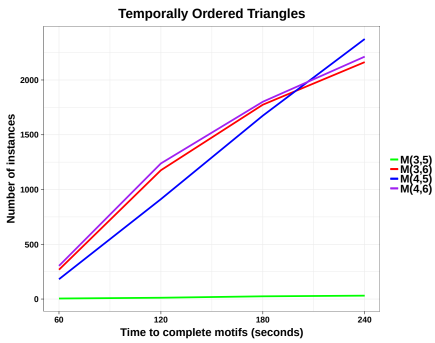


Figure 21: Originator participation = no

Figure 22: Temporal Ordered Triads

However, when the originator is participating on closing the triad conversation, it's interesting to see the sudden gap between the growth of M(1,3) and M(2,3). While M(1,3) reflects the appearance of causal relationship with a neighbor due to a common neighbor, which might be due to the similar interested event, e.g. assume there is an answer that both parties are interested on, hence they are attracted to each other since they comment on same. M(2,3) is missing that attractive pulse. Also, M(4,5) reflect such attraction in the case where the originator is not participating on closing the triad conversation.

Non-blocking interaction behavior As we model non-blocking conversation behavior by the temporal motif instances of M(4,1), M(4,3) and M(6,3), the latter group is identical in their growth, but M(4,1) has shown a sudden incline. It can be thought of a diverse nature of normal commenter who try to cover a range of clarifications.

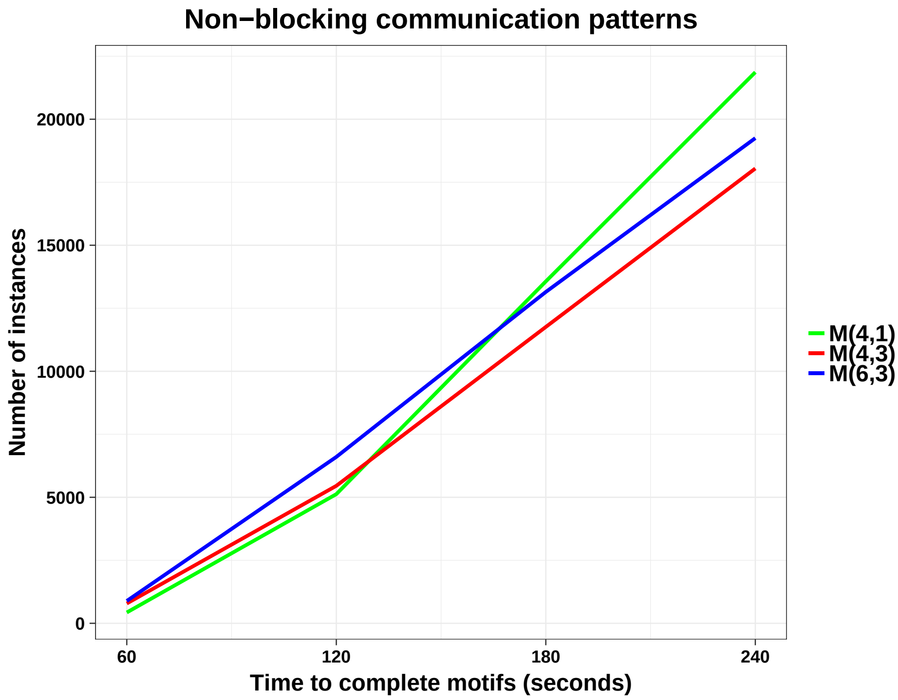


Figure 23: Non-blocking behavior

However, there is a significant change over the patterns of growth between answers-to-questions and comments-to-answers. While the former shows relatively low growth of the distribution, the latter shows the opposite. We believe that comments are more frequent to happen in the context and they are not blocked, while answers are more solid and not frequent.

Also, we could see a significant difference when considering the frequency distribution of non-blocking motif instance M(4,1). As time to complete such motif increases, the relative growth is inclined on comments, but not in the answers (ref M(4,1) at Figures 16 and 23).

4.2 Wiki-Talk

Wikipedia has user talk pages for all editors, where they can communicate with other editors via posts or messages. This dataset represents edits on user talk pages on Wikipedia [14]. An edge represents that an user edited other users talk page at a particular time.

Temporal Dyadic Behavior Figure 26 demonstrates the number of instances observed on the given dyadic instance of temporal motif against the time they spent to complete. It includes all the possibilities of dyadic behavior (M(5,1),M(5,2),M(6,1),M(6,2)) from our motif representation (Figure 4). They reflect the behavior of talking patterns between two users via talk pages.

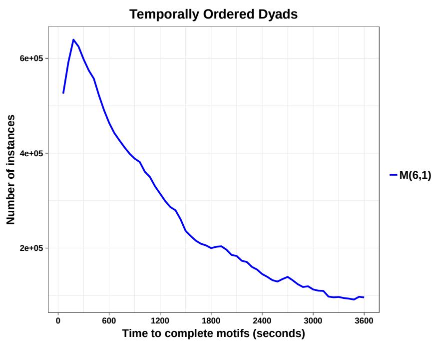


Figure 24: Dyadic behavior $M_{6,1}$

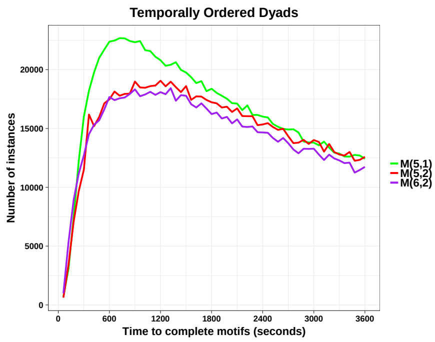


Figure 25: Dyadic behavior $M_{5,1}, M_{5,2}, M_{6,2}$

Figure 26: Temporal Ordered Dyads

More frequent edits between two users can be seen initially, but decreases heavily over time as shown in Figure 26. M(6,1) represent several edits by an user for other user's talk page, and it dominates other dyadic patterns by the frequency levels (Figure 24 and 25). However, we could see a decline growth of the motif frequency distribution, specially M(6,1). This is due to the nature of short discussion of between two users that relate with a shared Wikipedia article edit. The sessions are shortly lived, hence we would not see motifs that take a long time to complete.

Temporal Triadic Behavior Figure 29 demonstrates the number of instances observed on the given triadic instance of temporal motif against the time they spent to complete. We analyze how Wikipedia editors form triangles in the order of edge occurrences. We partition the instances into two classes by the participation of originator on closing the triad.

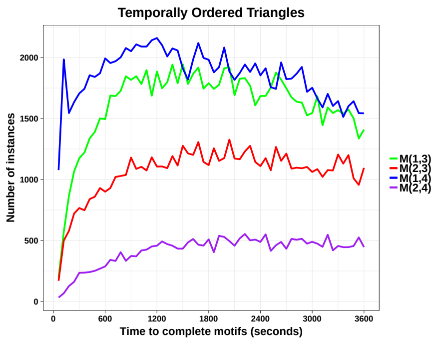


Figure 27: Originator participation = yes

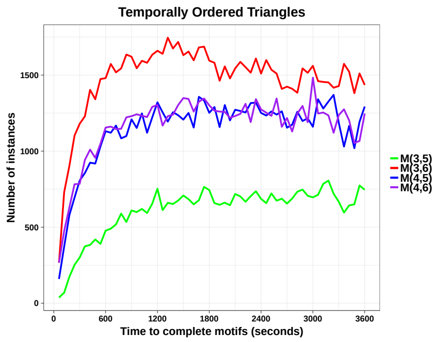


Figure 28: Originator participation = no

Figure 29: Temporal Ordered Triads

As shown by the Figure 29, the distribution of motif frequencies are identical, and shows an uniform nature after a initial growth. It's interesting to see when two users edit a shared talk page, they tend to edit each others' pages to have a triangle relationship, since M(1,4) and M(1,3) has higher growth over others (Figure 27). Further, the originator is attracted more to have direct closing edge in the triangle formation. Also, we observe a less number of cyclic triangles in the order of edge occurrences in both classes.

Non-blocking interaction behavior As we model non-blocking conversation behavior by the temporal motif instances of M(4,1), M(4,3) and M(6,3), we group them into two classes based on the number of switches that the originator has to deal with. Such that, M(4,1) has 2-switches while others only have 1 switch to proceed with the non-blocking interaction.

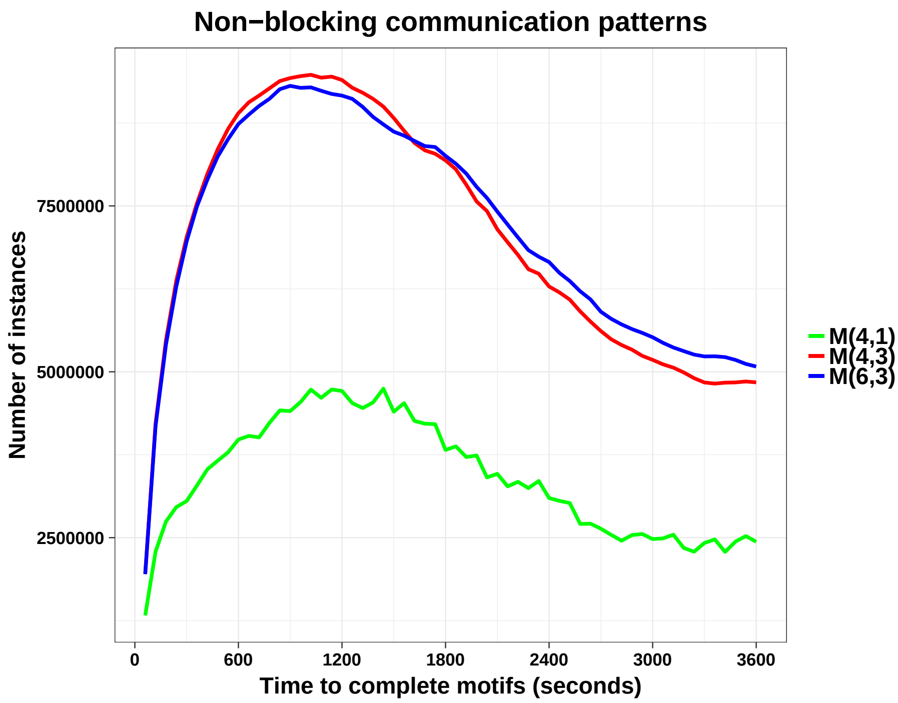


Figure 30: Non-blocking behavior

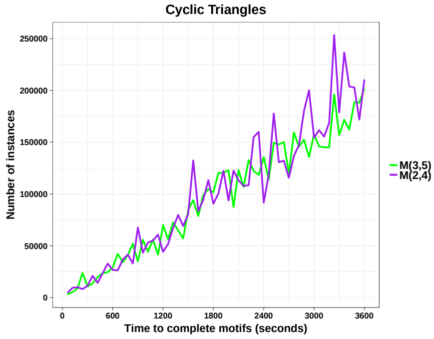
We observe a clear dominance of 1 switch non-blocking interactions over 2-switch, where an user doesn't like to switch the session in the middle. It's shown to have a persistent nature of user engagement and associated discussion in Wikpedia talk page edits. Hence, M(4,1) frequency distribution is shown to have significant low growth comparing with others.

4.3 Bitcoin

Apart from conversation networks, we try to shed some light on analyzing online transactions of the digital currency system - Bitcoins. They act in the same way of monetary transactions in daily lives, but it's digital and decentralized[4]. To exchange Bitcoins in a transaction, any user would be required to have a Bitcoin address, which is one-to-one relationship with <user,transaction>. A temporal edge signifies the exchange of Bitcoin from an user to another in a particular time. Specifically, we analyze 11,971,481 Bitcoin transactions by most active users, which is extracted from the list of all transactions up to 2013.12.28 [6].

Cyclic Triangles We observe a significant growth of cyclic triangles in the Bitcoin network as shown in Figure 31 over other networks. Such motif instances M(3,5) and M(2,4) depict the sending and receiving order of digital money which form a triangle between three users. Since, transactions have to be balanced globally, we would anticipate cyclic triangles to be present more in the Bitcoin network.

[4]http://www.vo.elte.hu/bitcoin/default.htm



Dataset	Mean	Std. Dev.
Bitcoin	96170	56026.48
Stack-overflow (c2a)	18.25	11.87083
Wiki-Talk	609.2	167.5881

Figure 31: Cyclic triangles

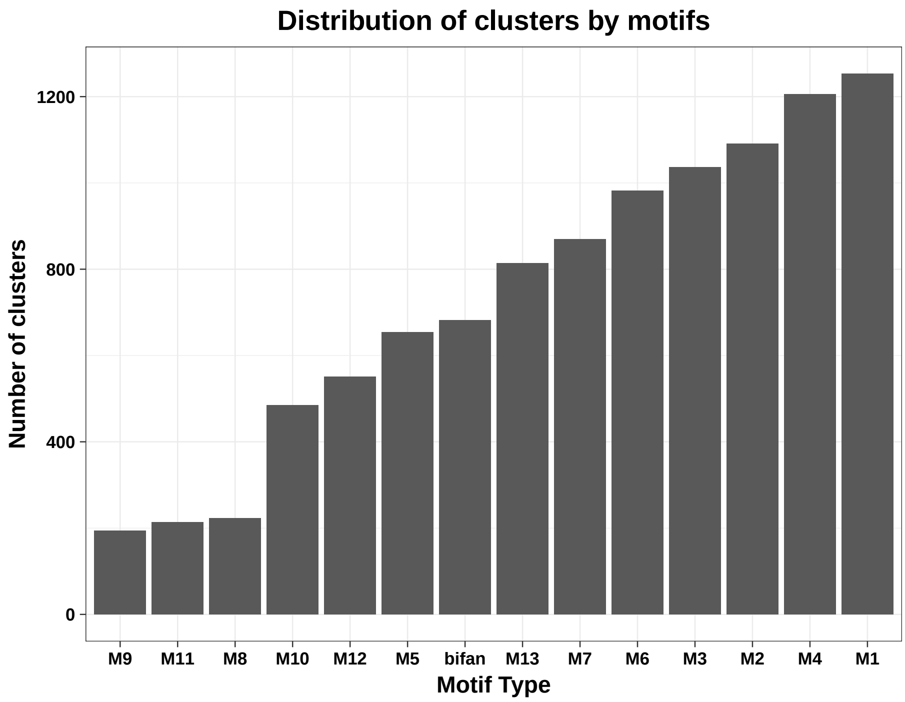
Higher-order organization Table 5 demonstrates the basic results of motif spectral clustering method introduced at section 3.2, separately for each of the 14 motif instances (see Figure 8). This summarizes the size of largest connected component which is available for spectral clustering, the size of cluster with the lowest motif conductance and eigenvalue associated with. As an example, given the motif M1, there is a size 115 of connected component available to partition, and we found the best cluster (S) with the lowest motif conductance 0.135417. Please note that we induce the associated static graph from the last temporal version for the analysis of higher order organization.

From the Table 5, we could observe that an near optimal clustering with the lowest motif conductance is shown at the motif M12 when partitioning the largest connected component. Due to the induction of static relations from the temporal Bitcoin transactions for this analysis, we can not derive the exact pattern in the order, but we could sat that's a representation of an unbalance triad with regards to money exchange.

Motif	Largest CC size	Cluster size	Motif conductance in largest CC	Eigenvalue
M1	115	32	0.135417	0.103735
M2	264	4	0.200000	0.096393
M3	317	4	0.125000	0.105369
M4	123	39	0.114650	0.076408
M5	717	3	0.250000	0.190097
M6	389	6	0.200000	0.188391
M7	491	5	0.142857	0.077313
M8	1146	5	0.142857	0.120632
M9	1177	6	0.173913	0.114873
M10	884	5	0.090909	0.061481
M11	1159	5	0.200000	0.185293
M12	819	7	0.058824	0.050965
M13	548	105	0.113835	0.083970
bifan	691	3	0.333333	0.333140

Table 5: Higher-order organization

It's interesting to note that cycle triangles (M1) would generate more weak clusters by size in the static induction graph (Figure 4.3). On average M12 generates 551 clusters across Bitcoin users. However, M11, M8 and M9 make large communities by size, which represents different broker types in a triad money exchange.



4.4 CollegeMsg - UCIrvine

This temporal dataset includes the interactions performed by UCIrvine college students in a local social network via private messages. An edge (u, v, t) means that student u sent a private message to user v at time t [9].

Motif	1)	2)	3)	4)	5)	6)
M(1,	67,043	39,518	1,460	1,214	64,883	85,581
M(2,	49,155	34,481	1,319	1,062	50,309	59,903
M(3,	39,830	43,391	65,090	72,227	1,086	1,400
M(4,	81,302	39,830	118,003	63,778	1,446	1,423
M(5,	98,400	82,667	52,491	62,484	55,027	64,921
M(6,	147,417	85,000	104,530	56,718	59,127	86,686

Table 6: Motif counts: $\delta = 30\ minutes$

Motif	1)	2)	3)	4)	5)	6)
M(1,	126,687	75,313	2,666	2,050	132,211	184,134
M(2,	92,042	64,313	2,309	1,655	109,695	125,020
M(3,	79,491	85,000	134,890	157,501	1,938	2,503
M(4,	161,035	79,491	276,981	136,792	2,595	2,437
M(5,	170,112	149,984	111,087	132,017	113,096	133,772
M(6,	278,779	156,065	244,525	129,352	131,504	188,249

Table 7: Motif counts: $\delta = 1\ hour$

Table 6 and 7 present the counts of temporal motifs, which complete within the duration of 30 and 60 minutes consecutively. M(4,1) depicts a significant growth over 30 - 60 minutes to complete the given structure, which highlights the preference for recent conversation to proceed with between two students.

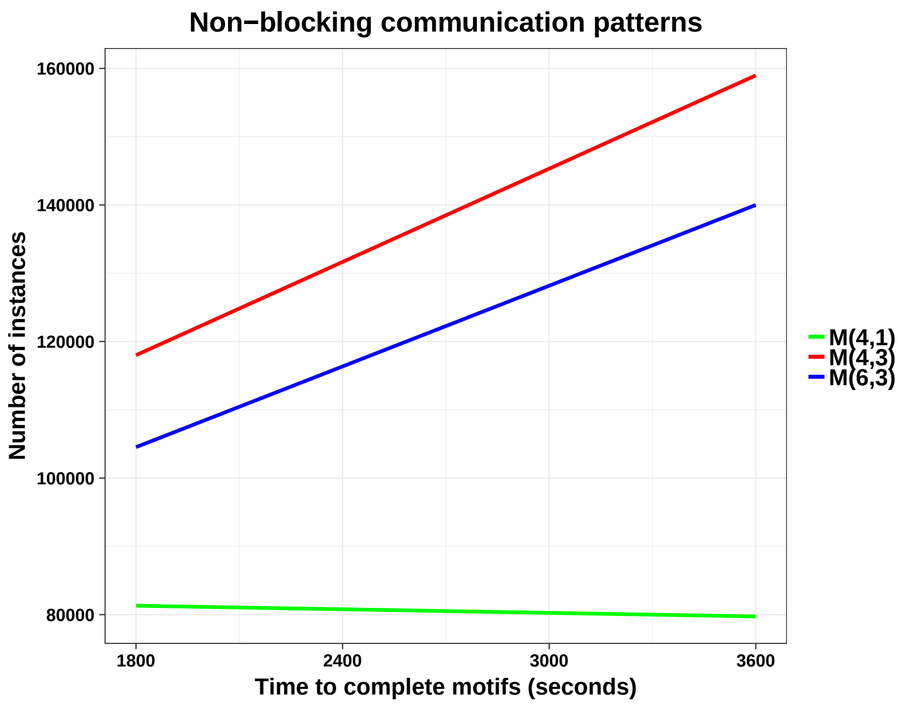


Figure 32: Non-blocking behavior

The non-blocking behavior is identical to WikiTalk as it represents a general conversation behavior, but it's interesting to see the continuous growth by M(4,3) and M(6,3), where the originator has only 1-switch between the conversation with two parties. It's more likely students like to maintain a conversation without switching to other, and usually it's the most recent conversation.

Motif	Largest CC size	Cluster size	Motif conductance in largest CC	Eigenvalue
M1	73	26	0.039216	0.020367
M2	564	8	0.285714	0.190130
M3	872	3	0.250000	0.168959
M4	679	6	0.090909	0.067224
M5	716	5	0.142857	0.103800
M6	702	3	0.200000	0.158960
M7	676	4	0.166667	0.134672
M8	1704	8	0.306306	0.210825
M9	1702	808	0.342365	0.427103
M10	1319	5	0.285714	0.210808
M11	1837	858	0.365982	0.489463
M12	1325	594	0.429393	0.498692
M13	1264	626	0.312469	0.475830
bifan	1030	281	0.328400	0.371212

Table 8: Higher-order organization

As you could observe, the triangle (M1) gives the optimal spectral clustering out of other motifs in Table 8. But it gives the maximum number of clusters, making more weak communities by the size (M1 - 1800 Figure 33). This represents the transitivity nature of forming communities across college students. Hence, they would not lead to form larger communities, but more likely to be remained as small groups.

Figure 33 shows that M11 generates only 64 clusters by having around 30 $(\frac{1899}{64})$ students forming communities in average. In general, M11, M8 and M9 represent a shared broker who intermediate the communication in a triad, such that making less number of communities, but they are more likely to be strong by size.

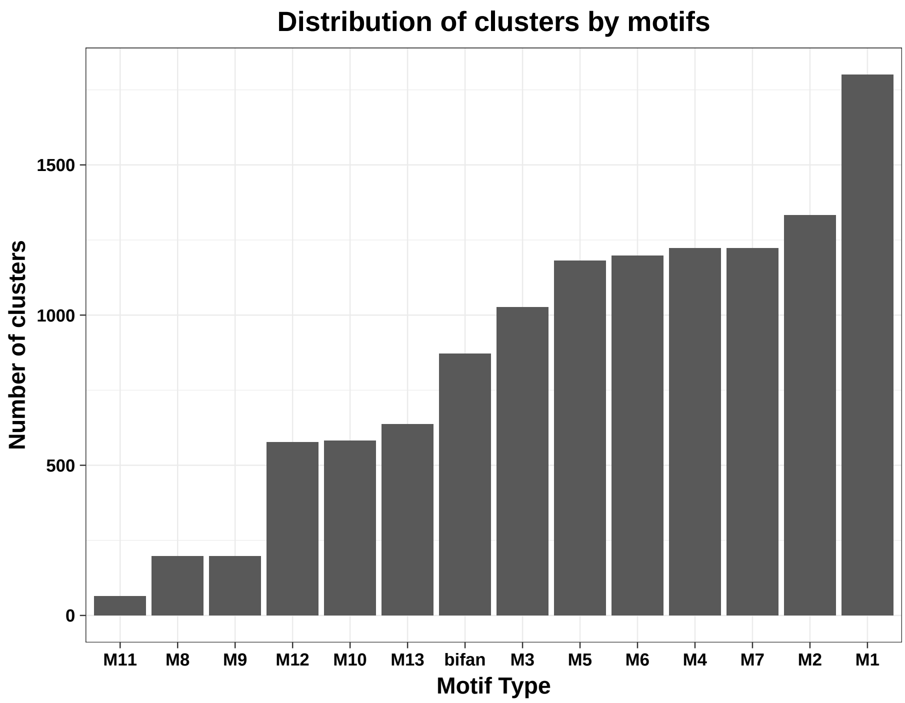


Figure 33: Distribution of clusters by motifs

4.5 Ground-Truth communities

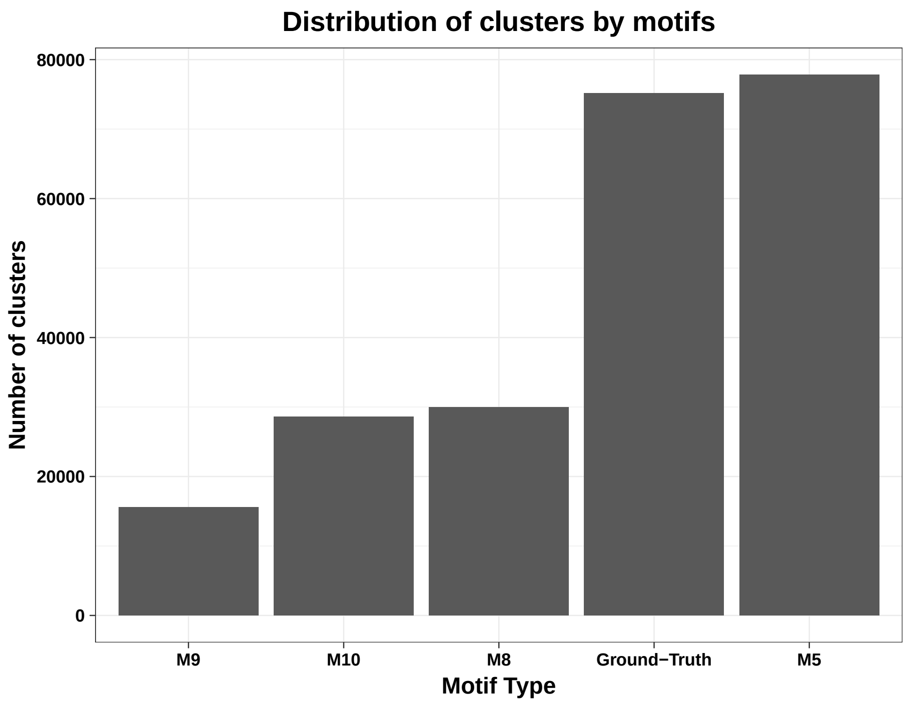
We try to evaluate the spectral clustering performed using motif conductance with ground-truth communities, Yang et al. [25] detect that community definitions related with the classes of edge conductance and triad-participation-ratio, consistently give the best performance in identifying ground-truth communities. Here, we extend their experiments to proceed with motif conductance at several static networks.

4.5.1 Amazon product co-purchase

Nodes in the network represent Amazon products that were available for shopping until 02.03.2003, an edge is defined as following: "If a product i is frequently co-purchased with product j, the graph contains a directed edge from i to j. Each product category provided by Amazon defines each ground-truth community." [9].

Motif	Largest CC size	Cluster size	Motif conductance in largest CC
M5	189145	504	0.000342
M8	300717	170	0.000529
M9	316029	115	0.000272
M10	301727	44	0.000830

Table 9: Higher-order organization



We could see that the largest connected component to partition lies at motif M9 (Table 9), which also depicts the lowest motif conductance. However, M5 based spectral clustering is identical with the ground-truth community partitions by the number of clusters.

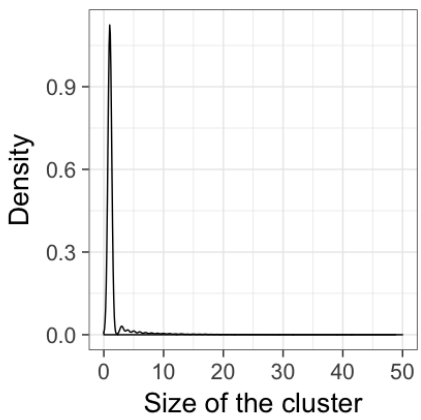


Figure 34: M5

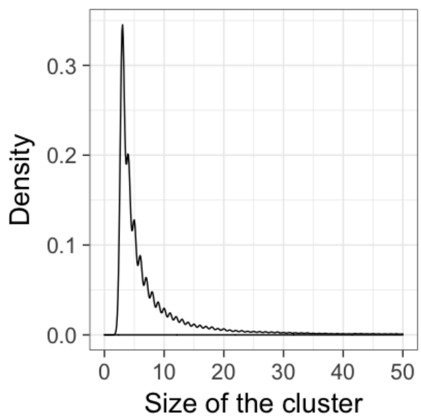


Figure 35: Ground-Truth

Figure 36: Density distribution of spectral clustering and ground truth

Figure 36 shows how the density of clusters varies with the size of the cluster, generally it measures the spread of clusters by size. Density distributions are used to compare M5 motif spectral clustering with ground truth. We can observe an identical shape, but the density varies a lot with the size of cluster.

4.5.2 DBLP co-auhorship network

DBLP is a repository that maintain research papers listed in computer science. "We construct a co-authorship network where two authors are connected if they publish at least one paper together. Publication venue, e.g, journal or conference, defines an individual ground-truth community; authors who published to a certain journal or conference form a community" [9].

Motif	Largest CC size	Cluster size	Motif conductance in largest CC
M5	259533	142	0.000997
M8	297216	20	0.003021
M9	306243	25	0.008850
M10	133090	28	0.019608

Table 10: Higher-order organization

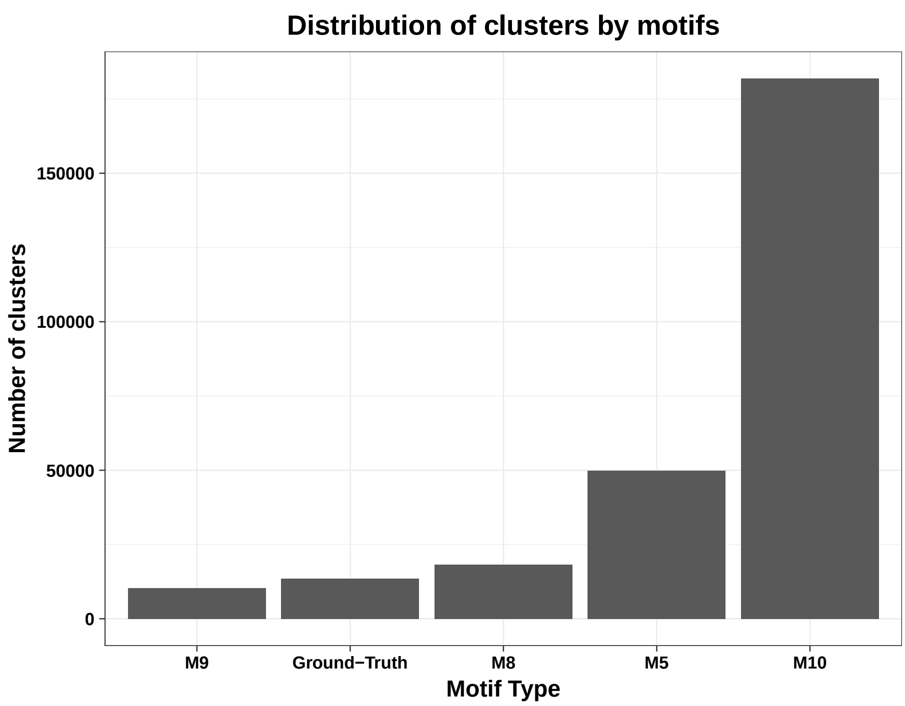


Figure 37: Distribution of clusters by motifs

Figure 37 shows that the optimal number of ground-truth communities is within the range of clusters produced by M9 and M8. However, we couldn't see any identical distribution on cluster frequencies by size across M8, M9, and ground-truth communities (Figure 41).

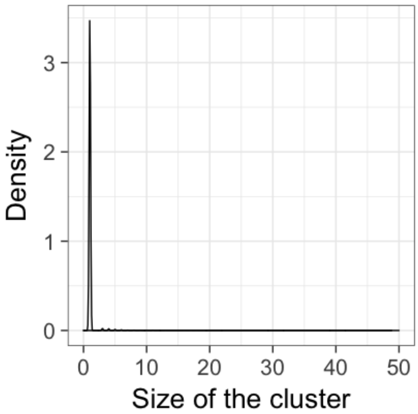


Figure 38: M8

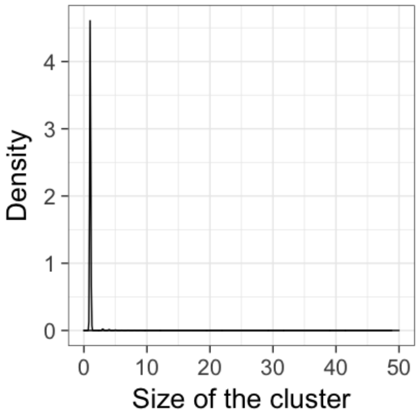


Figure 39: M9

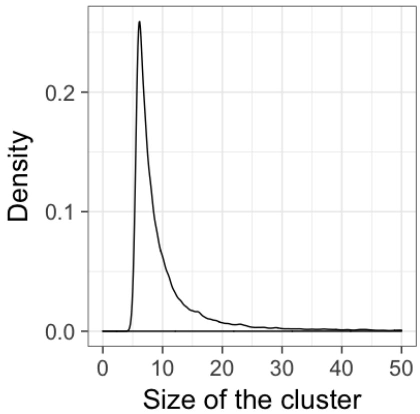


Figure 40: Ground-Truth

Figure 41: Density distribution of spectral clustering and ground truth

5 Related Work

5.1 Information Diffusion

Zhao et al. [26] discuss the possibility of inferencing information propagation and behavior patterns of social networks (specifically communication networks.) through the lenses of temporal motifs. It's assumed that the amount of information is proportional to the duration of the interaction, and adjacent interactions between a common user tend to propagate the same set of information. By annotating temporal attributes, both the lifespan and the frequency of patterns of interactions were measured.

Event driven nature of communication networks is modeled by analyzing the frequency of temporal motifs in different contexts, range from call logs, to facebook wall-posts. It's been observed that the communication behavior is common in different environments, by having a shared temporal motif instances (e.g star and chain motifs). But however, the distribution of temporal motifs' frequencies is not identical, by generating different patterns of information propagation which are dependent on the (a)synchronous behavior of interactions. Also, the speed and amount of information is correlated, and dependent on the individual's ego.

Liu et al. [10], define the notion of stochastic temporal network motif (STNM) based on first-order markov chain to study the patterns of communication in a mobile network. Further, it's been extended to characterize the hidden communication patterns by learning a discrete markov chain. The major contribution is the proposal of probabilistic approach to infer the evolution of temporal motifs.

5.2 Community

Creusefond et al. [3] study the structure of communities via temporal motifs, specially in networks that have short-lived interactions. They are motivated by the assumption that the nature of communities shared by any two users could be depicted by the nature of their interaction. From the experiments, they observe star and chain motif instances are likely to be inside explicit communities, while spams, ping-pong and triangles could occur in cross communities. Hence, the diffusion of information over global network structure could be supported by former instances that interact over cross communities.

Xuan et al. [24] use temporal motifs to reveal collaboration patterns of a software development environment. Specifically, they study the composition of two people and an artifact in task-oriented social networks. Hence, they try to reveal how two people interact to achieve a shared task by monitoring temporal collaborations and communications. They observe the number of temporal motifs is larger than the chance model most of time in task-oriented social network. Further, they infer about the team structure, and conclude that central individuals and the more cohesive teams are more productive.

6 Discussion

Motifs are initially used to define the patterns of interest in static complex networks (e.g biological network), thus to study the functional behavior of local structures [11]. With the introduction of equivalence set of isomorphic instances, they have been utilized to reason about the universality classes of network.

Definitions: When defining motif instances in temporal networks, we have to concern about the generalized concept of induced subgraphs and isomorphism with the time dimension. We have seen the constraints introduced at temporal graphs in the definition 2.1, to have a configuration model of $\delta-$temporal motifs. It's assumed to have connected subgraphs, which is active within a δ time window. The analysis of unconnected subgraph patterns is out of the study, where we consider any two events are causally related within a short time interval δ to be represented in a valid motif instance. We adopt Goldilocks approach [24] to have a representative δ, which is not too small, since responding to an event takes time; and not too large, to avoid the expired associations. Thus, we vary δ, as a small sized sliding window (usually 60 seconds), and produce fine-grained results.

By the definition 2.2, it would give us an exponential number of subgraphs, which also lead to have constraints on computation. Also, this vast space of motif structures, might be quite misleading since the choice of the configuration model is dependent on the context of analyzed graph, and hard to interpret results [8]. Such that, we control the motif types by k; number of nodes, and q; number of edges, but allow to have a version of multi-graph. It's assumed that the multiplexity of ties has been achieved by associated timestamps of edges, which could be considered as discrete set of types.

Also, we try to conceptualize the notion of brokerage positions with respect to temporal motifs. Multigraph automorphism (definition 3.3) is studied to define the equivalence classes of such positions, which could be extended into the study of temporal motif degree (section 3.4). Further work including solid definitions, development of new tools etc. is required in this section.

Higher order organization of a network is only considered over static snapshots that unifies over motif analysis and network partitioning. As Benson et al. [2] suggest, motif-based clustering could be used to find out the motif that organize the network structurally. This study is an extended version of hypergraph partitioning, where they interpret motifs as hyperedges in a graph.

Impact: Also, we present our results mostly derived from the frequency of motif appearances over different activation window. Overall, they could be used to find the impact on human communication behavior which is attributed by two universal properties - burstiness and causality [7]. However, we observe that motifs which interpret causal events are more common. Such patterns of causality is existed over similar distributions of motifs over similar networks (e.g. different causality patterns wrt communication vs. transaction).

In our study, we try to evaluate motif-based clustering with some networks which have predefined ground-theory community structures. Our benchmark parameters include the number of obtained communities, and the density distribution of community sizes. However, the results are not up to the expectations, where density distributions in both of our examples (section 4.5) don't align with ground-truth. We might argue that the structural position of a node is highly effected by some network forces of homophily, reciprocity etc. In that sense, motif-based clustering does not account any of node's attributes to define the force. This pay-off might be represented by the distance between obtained distributions of communities over ground-truth and spectral clustering.

Computational complexity: We use snap for our analysis including temporal motif counts and higher order organization due it's scalability. Along with the optimized algorithms, we were capable to process more than 16 million nodes (e.g. StackOverflow), in it's temporal version of edges. Snap is capable to model the temporal graphs in a data structure, modeled with multi-graphs.

Extensions: As an extension, we would like to consolidate our study using the concept of null models, to better claim results using null hypothesis. It would enable us to carefully examine the hidden causality in the network.

Also, the definition of temporal motifs could be generalized to represent the simultaneous events happening around in the network, in fact, such definition would degenerate to the definition 2.2, but edges are attributed by the number of simultaneous events. Hence, several modifications are required for equivalence classes. We have done a basic survey on the context of synchronous behaviors, which is presented optionally in the Appendix A.

However we observe that granularity levels of temporal and structural dimensions would vary over the evolution of the network. Evolution of a population was detected using traits, and is dependent on the strength of such evolation force (i.e. growth rate, fitness) [5]. An interesting question would be to claim about the fitness of the population and the contribution of different traits to the evolution of the network through the study of temporal motifs.

References

[1] Miguel Araujo, Spiros Papadimitriou, Stephan Günnemann, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos E Papalexakis, and Danai Koutra. Com2: fast automatic discovery of temporal (comet) communities. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 271–283. Springer, 2014.

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

[3] Jean Creusefond and Remy Cazabet. *Characterising Inter and Intra-Community Interactions in Link Streams Using Temporal Motifs*, pages 81–92. Springer International Publishing, Cham, 2017.

[4] Hubert De Fraysseix and Patrice Ossona De Mendez. Regular embeddings of multigraphs. In *Topics in Discrete Mathematics*, pages 553–563. Springer, 2006.

[5] Martin Hilbert, Poong Oh, and Peter Monge. Evolution of what? a network approach for the detection of evolutionary forces. *Social Networks*, 47:38 – 46, 2016.

[6] Dniel Kondor, Istvn Csabai, Jnos Szle, Mrton Psfai, and Gbor Vattay. Inferring the interplay between network structure and market effects in bitcoin. *New Journal of Physics*, 16(12):125003, 2014.

[7] 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):P11005, 2011.

[8] Lauri Kovanen, Márton Karsai, Kimmo Kaski, János Kertész, and Jari Saramäki. *Temporal Motifs*, pages 119–133. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[9] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`, June 2014.

[10] Kai Liu, William K. Cheung, and Jiming Liu. Detecting stochastic temporal network motifs for human communication patterns analysis. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '13, pages 533–540, New York, NY, USA, 2013. ACM.

[11] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[12] Giovanna Miritello. *Temporal patterns of communication in social networks*. Springer Science & Business Media, 2013.

[13] Yavero Nebil, Sean M Fitzhugh, Maciej Kurant, Athina Markopoulou, Carter T Butts, Nataa Prulj, et al. ergm. graphlets: A package for erg modeling based on graphlet statistics. *Journal of Statistical Software*, 65(i12), 2015.

[14] Ashwin Paranjape, Austin R. Benson, and Jure Leskovec. Motifs in temporal networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, WSDM '17, pages 601–610, New York, NY, USA, 2017. ACM.

[15] Nataša Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):e177–e183, 2007.

[16] Daniel M. Romero, Brian Uzzi, and Jon Kleinberg. Social networks under stress. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 9–20, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee.

[17] Serguei Saavedra, Kathleen Hagerty, and Brian Uzzi. Synchronicity, instant messaging, and performance among financial traders. *Proceedings of the National Academy of Sciences*, 108(13):5296–5301, 2011.

[18] Vedran Sekara, Arkadiusz Stopczynski, and Sune Lehmann. Fundamental structures of dynamic social networks. *Proceedings of the National Academy of Sciences*, 113(36):9977–9982, 2016.

[19] Chayant Tantipathananandh, Tanya Berger-Wolf, and David Kempe. A framework for community identification in dynamic social networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, pages 717–726, New York, NY, USA, 2007. ACM.

[20] Colin J. Torney, Tommaso Lorenzi, Iain D. Couzin, and Simon A. Levin. Social information use and the evolution of unresponsiveness in collective systems. *Journal of The Royal Society Interface*, 12(103), 2014.

[21] Enrico Ubaldi, Alessandro Vezzani, Márton Karsai, Nicola Perra, and Raffaella Burioni. Burstiness and tie activation strategies in time-varying social networks. *Scientific Reports*, 7, 2017.

[22] Shahadat Uddin, Arif Khan, and Mahendra Piraveenan. A set of measures to quantify the dynamicity of longitudinal social networks. *Complexity*, 2015.

[23] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

[24] Qi Xuan, Huiting Fang, Chenbo Fu, and Vladimir Filkov. Temporal motifs reveal collaboration patterns in online task-oriented networks. *Phys. Rev. E*, 91:052813, May 2015.

[25] Jaewon Yang and Jure Leskovec. Defining and evaluating network communities based on ground-truth. *Knowledge and Information Systems*, 42(1):181–213, 2015.

[26] Qiankun Zhao, Yuan Tian, Qi He, Nuria Oliver, Ruoming Jin, and Wang-Chien Lee. Communication motifs: A tool to characterize social communications. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 1645–1648, New York, NY, USA, 2010. ACM.

Appendix A - Related work on Synchronicity

Synchronous trading has been evaluated in the context of financial stock market by taking stock selection of 66 day traders [17]. The results suggest the timing is a key factor that drives the decision of traders in highly critical financial systems. Also, the level of uncertainty would let the synchronous behavior to be emerged among the traders. Finally, they conclude synchronous trading does not appear due to co-ordination, but it helps individuals to perform better in the context.

Another way of thinking about dynamic behavior is the collective reaction of a community for an external event, such that explains about the individuals' cognitive reactive process aggregated in a social network. Romero et al. [16] examine a social network under stress to observe that a structure of a network is diagnostic of the sudden execution of new events, and it can be used to predict behavioral patterns in a glance.

It's been shown that collective intelligence of a community could eradicate human errors of individual decisions by making use of social information. As a downside, this might lead to have a lack of responsiveness for rapidly changing information streams, while mostly relying on synchronous behavior of others [20].

From a longitudinal study over 1000 university students, Sekara et al. [18] observe that individuals participate in subsequent gatherings may have seen repeatedly as a core subset. Thus, they represent a dynamic social network with strong temporal and spatial regularity.