
Structural analysis of cascade dynamics for measuring inhibition of information diffusion

Soumajyoti Sarkar · Ruocheng Guo ·
Paulo Shakarian

Received: date / Accepted: date

Abstract Most social network sites allow users to reshare a piece of information posted by a user. As time progresses, the cascade of reshares grows, eventually saturating after a certain time period. While previous studies have mostly focused on predicting when the cascade would go viral or when there would be an epidemic outbreak in the lifecycle of the cascade, in this paper, we focus on causes which inhibit the growth of a cascade after a certain time period. We focus on this problem of cascade diffusion inhibition based on the concept of network motifs and node-centric structural measures such as Pagerank, clustering coefficients among other widely used measures, which help us conclude how the network structure changes over the cascade lifecycle to impact the growth after a certain time period, eventually leading to inhibition. We make a comparative analysis of these structural measures using the cascade network structure during the time intervals preceding two important phases of the cascade lifecycle - the period of maximum growth and the start of the inhibition period which help us define causal relationships between the network structure and the decaying nature of the cascades after a certain time period. We observe how small network patterns such as motifs can be used to explain the group dynamics of the cascade network as it transitions towards a decayed resharing phase leading to a complete stop in reshares. We also study the effect of the node-centric network features on the dynamics of the reshare

Soumajyoti Sarkar
CIDSE, Arizona State University
E-mail: ssarka18@asu.edu

Ruocheng Guo
CIDSE, Arizona State University
E-mail: rguosni@asu.edu

Paulo Shakarian
CIDSE, Arizona State University
E-mail: shak@asu.edu

times within the lifecycle using Granger causality which helps us quantify the significance of each node-centric feature.

Keywords social network analysis · information cascades · motifs · causality

1 Introduction

Sharing information in online social networks has become a widespread phenomenon where multimedia information can be in the form of text, photos or links to other information. When such a piece of information is shared among multiple people over a prolonged period of time, we obtain cascades of resharing events for that information. There has been a growing interest in information cascades as they have wide range of applications in viral marketing [2] [3], cascade prediction [12], and traffic management. The increasing availability of data identifying diffusion traces that lead to such cascades has allowed researchers to obtain empirical evidences of mechanisms through which information diffusion occurs in social networks. However, majority of these studies have focused on predicting the growth of cascades over a period of time.

One of the commonly overlooked phenomenon in the cascade lifecycle is identifying the time phase when the cascade would start decaying and the causes which inhibit the growth of cascades in that interval. A study on what limits the size of cascades has been performed in [10] where the authors focus on spreading mechanisms to determine whether the cascades would be limited in size or would continue to cross an epidemic threshold. The main motivation for that paper was to study how repeated exposures impact the final size of cascades. Contrary to the size attribute studied there, in this paper we focus on social network analysis that identifies causes in network structure which inhibit cascades to grow after a certain time period. Unlike the social contagion mechanism involving all cascades studied in that paper, we adopt the independent cascade model to observe time periods within the lifecycle of each cascade and the causes which enhance or impact the growth in those time intervals.

The main contributions of this paper are as follows:

- We identify the period of maximum growth and the start of inhibition time phase in the course of the cascade lifecycle using the concept of Hawkes self exciting point process model [4][28].
- We study the dynamics of the cascade network structure as it grows over time and analyze the time intervals leading to the region of maximum growth and the region of inhibition to explain the causes in the network that augments or creates a bottleneck in the information diffusion process.
- We use recurring graph patterns called network motifs and use such patterns to study how the formation of such small interaction groups leverage the diffusion process during the cascade lifecycle. We conclude from our observations that the intervals preceding the inhibition period witness accelerating formation of such small patterns before just after which it starts decelerating to finally reach the inhibition period.

- In contrast to motifs which explains group dynamics within time-varying networks, we use node-centric social network measures to explain the phenomena of growth and decay of the resharing process by focusing on individuals as the factor for information diffusion. Our results show that while the node degree may be important as far as the connectivity of individuals are concerned, the power of the neighbors' degree is more influential as an indicator of the approaching inhibition region.
- We quantitatively measure the effect of the node-centric network features on the reshare time difference in the lifecycle by using Granger causality [42] which is identified using VAR models due to their simplicity.

The key component in our approach is studying the network structure at different time intervals in a cascade and observing how group dynamics and node connectivity can help us infer the change in the structure leading to inhibition. To the best of our knowledge, this is a first comprehensive study on causes of cascade inhibition and using social network analysis to measure them.

2 Related Work

There has been an increasing line of work surrounding the dynamics of cascade growth[8], [12]. Numerous methods have studied in social network diffusion [38] surrounding such cascades. Social network diffusion has been an important component in predicting cascade growth. To this end, [4][37] used Hawkes model to measure time-varying social influence and to model viral network diffusion. Applications of Hawkes process dates back to studies describing self-exciting processes of earthquakes [30]. Our approach to identify the period of maximum growth and start of the inhibition region in a cascade life based on Hawkes process is performed along the line of work introduced in [28] where the authors use Hawkes point process model to predict the final number of reshares of a post.

Using social network features to identify superspreaders has been comprehensively studied in [33][24][23] where the authors compare various features like PageRank, degree centrality, core number in a k -core decomposition of network, to identify influential users. Similar to such network features, graph motifs have been used in studying collaboration networks of authors [14], biological genes [16], human communication networks [15] and thereby such studies give a hint as to why motifs form core components in studying time varying networks. Extending this concept, temporal motifs, which forms the basis of one of our measures used in this paper, have been studied in the context of communication networks [17][19] in which the authors use the concept of motif patterns distinguished by times of events to study homophily in networks.

Table 1: Table of Symbols

Symbol	Description
C	Information cascade
T_C	Total span of cascade C in minutes (time difference between the first and the last reposting)
S_C	Total size of cascade C equal to the total number of reshares for C .
V_C^τ	Nodes which participated in the cascade C in the time interval τ in the cascade lifecycle.
E_C^τ	The social interactions between pairs of individuals denoted by $e=(i,j)$ in the time interval τ in the cascade lifecycle.
rt_v^C	Time at which node $v \in V_C$ (set of cascade nodes) reshared C (such that $rt_v^C \leq T_C$)
w_ϕ	Time interval (<i>window</i>) of ϕ region where $\phi = \{\text{steep}, \text{inhib}\}$
t_ϕ	The mid-point of τ_ϕ where $\phi = \{\text{steep}, \text{inhib}\}$
M_{ms}	Set of motif patterns of size ms .
m_i	A motif pattern belonging to the set M_{ms} for some ms .

3 Technical Preliminaries

We represent a social network as a directed graph $G = (V, E)$ where V is a population of individuals and the edge $(i, j) \in E$ refers to individual i having the ability to influence individual j . We will assume that these influence relationships are known a-priori. As a practical matter, we determine these relationships from previously observed microblog relationships that occur prior to the microblogs analyzed in this paper¹. This is similar to our previous work [1] - this will be described in detail in our dataset section. Following conventions established in previous work [1][9][12], we will use the symbol C to denote an arbitrary information cascade (i.e. a microblog that spreads in the social network). In Figure 1(a) we show an example plot of time (t) versus cumulative cascade size, size denoting the number of reshares for C .

We denote the set of reshare times for C as $\tau_C = \{0, \dots, t, \dots, N_C\}$ ordered by time, where all the reshares range between the timepoints $[0, N_C]$. So by our definition of a social network, we denote $G_C = (V_C, E_C)$ as the graph for C , induced by all the nodes V_C who participated in the cascading process and where an interaction between a pair of individuals maps to an edge in the set E_C . We will drop the subscript C from the notations when they are applicable for all cascades.

Given the resharing time sequence τ , a subsequence $\tau' \subset \tau$ is a sampling of length $|\tau'| < N_C$. In this context, we denote $G^{\tau'} = (V^{\tau'}, E^{\tau'})$ as the subgraph induced by the nodes $V^{\tau'}$. In our work, we sample the subsequences in a way such that the following conditions hold: We denote an instance of a subsequence by adding the subscript

In our work, we keep $|V^\tau|$, the number of distinct nodes which participated in the cascading process during subsequence τ , fixed for every window.

¹ Note that even in our machine learning experiments, these relationships are based on microblog influence behaviors that occurred *prior* to both the training and test data.

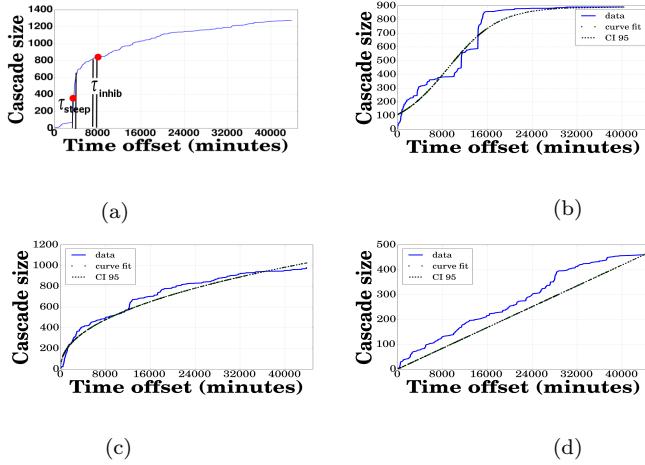


Fig. 1: (a) *Growth curve* depicting the progress of a Type I cascade (b) Type II curve fitted to logistic function (c) Type III curve fitted to a concave increasing function (d) Type III curve fitted to a straight line curve.

As would be described in the network evolution model in later sections, the advantage of selecting $|V^w|$ a-priori is that we can be agnostic about the cascade size and analyze the network generated by the windows as the cascade progresses over time until we reach the end of the cascade. So in this context our approach avoids retrospective analysis of the social network generated by the cascade.

A *temporal network* within a cascade denoted by N_i is defined as the cumulative subgraph induced by the nodes and edges in two consecutive subsequences τ_{i-1} and τ_i . In this context, a temporal representation of a cascade is a sequence of temporal networks $\langle \tau_1, \dots, \tau_W \rangle$.

3.1 Steep and Inhibition intervals

Figures 1(a) to (d) show the kinds of *growth curves* which depict the increase in cumulative cascade size ($|V^{rt}|$) versus the reshare time rt . In Figure 1(a) we show an example plot of *growth curve*.

We define three such *Growth curves* (see examples in Figure 1):

1. *Type I cascades*: Cascades which follow the logistic function as shown in Figure 1(a).
2. *Type II cascades*: These cascades exhibit a step-like pattern of growth shown in Figure 1(b).
3. *Type III cascades*: These cascades do not follow the logistic function as shown in Figures 1(c) and (d) and one of the many reasons for these cascade curves is that they probably do not complete their lifecycle within the period of 3 months that we have considered.

The motivation behind defining these three types of *Growth Curves* lies in the way we define the *inhibition* interval as will be described below. It arises from the intuitive explanation of an inhibition period in the cascade lifecycle being a period whereafter the cascade fails to regain any surge in the growth rate and hence finally dies. Since we aim to tag only a single period in the cascade lifecycle as the *inhibition* interval, we only consider Type I cascades for our study where the cascade lifecycle follows such growth dynamics as would allow us to tag a single interval as an inhibition interval. As can be seen in Figures 1(b), (c) and (d), there are multiple regions of slackness in growth which makes it difficult to tag one particular region as an inhibition phase as per our definition and hence for this study, we do not consider Type II and Type III cascades. Also, since we observe the dynamics of growth of each cascade over a period of one month, any anomaly arising due to time scaling issues would be captured by Type II and Type III cascades and therefore we prevent taking up such cases in our inhibition study. In our diffusion data, most of the information cascades we observed were of Type I. Hence, we generally assume a logistic fit to the growth curve.

The major challenge in our study is the lack of ground truth data of the inhibition time interval of a cascade, to study the dynamics of the network structure surrounding that interval. To overcome the issue cause by the absence of ground truth, we use retrospective analysis on selected cascades to infer the approximate inhibition time phases. We use the concept of point processes and maximum likelihood estimation to develop a parametric model to find the approximate inhibition time points of the available cascades. Given scaling parameter α , we divide the *Growth curve* of C into sequence of time intervals of uniform size $\alpha \log(T_C)$ (refer Table 1 for symbols). Note that these sequence of intervals are different from the subsequences τ introduced in Technical preliminaries notations above, as here it is necessary to use retrospective analysis to assume that we know T_C a-priori for the cascades we select for parameter estimation. Amongst these intervals, two are of interest in this work: the *steep interval* and the *inhibition interval*, an example of which is shown in Figure 1(a). Intuitively, the *steep interval* is the interval where maximum, “spiky” diffusion activity occurs characterized by a sharp increase in the adoption over the previous intervals that is also maintained in subsequent intervals. Likewise, the *inhibition interval* represents a significant, “spike” decline in adoption relative to previous intervals - which is followed by subsequent intervals where adoption continues to decay or saturate.

To formalize these ideas, the simplest notion would have been to find the slope or the first derivative of the reshare times rt - but we found this method to have some significant drawbacks:

1. It is difficult to define a threshold for the slope values at the start of the inhibition phase as the rate of adoption in that region for each cascade varies significantly.

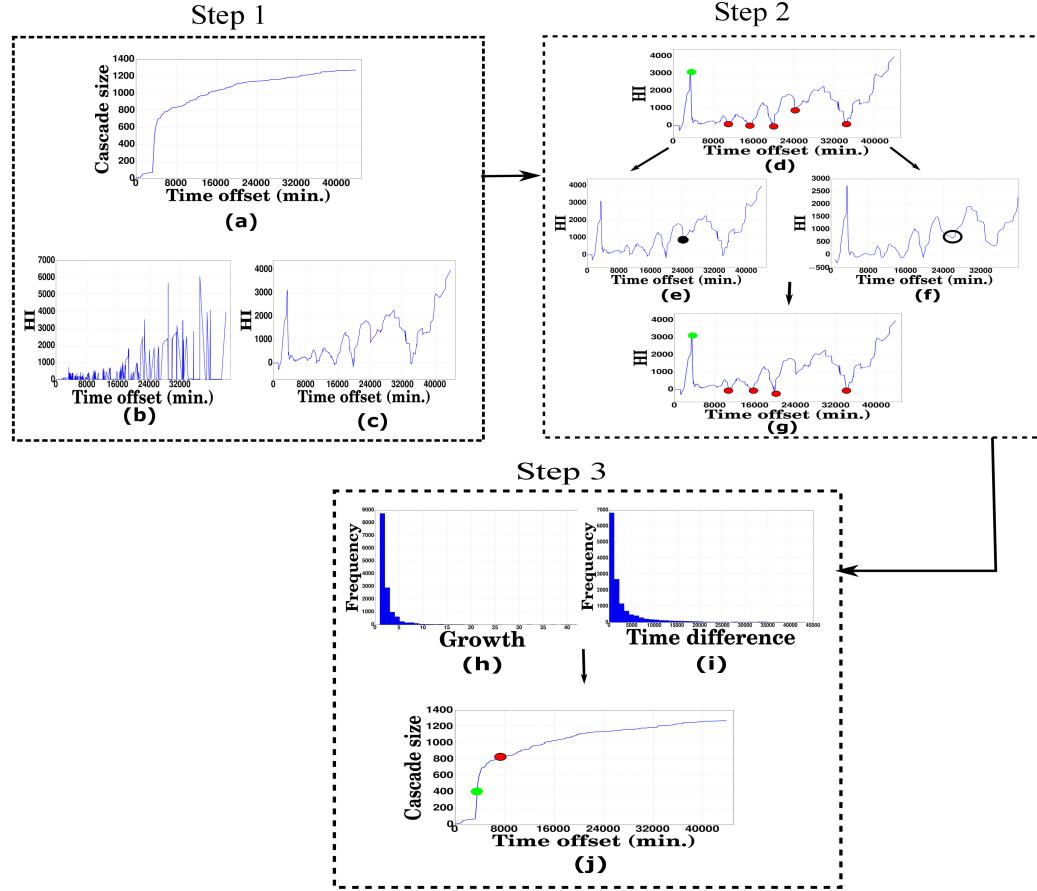


Fig. 2: This flowchart lists the steps used in the identification of the *steep* and the *inhibition* intervals for C . Refer to **Appendix A¹** section for algorithm details. Refer to Table 1 for symbols. **Step 1:** We obtain the Hawkes interval curve H (fig. (c)) from the *Growth curve* C (fig. (a)) by splitting the Hawkes curve (fig. (b)) into intervals of size K_C . **HI** on the Y-axis refers to Hawkes Intensities. **Step 2:** We set t_{steep} (marked in green point) by selecting the first local maxima among neighborhood, of Hawkes values and obtain a set of t_{inhib} (marked in red points) by selecting the local minima among neighborhood shown in fig. (d). Using the moving mean curve M (fig. (f)) we filter out outliers. One such outlier is marked by the black dot in fig. (e) where $H[t] > M[t]$, marked in circle in fig. (f). At the end of this step we obtain t_{steep} and a set I_C of t_{inhib} points shown in fig. (g). **Step 3:** We consider the corpus of cascades C_{all} and obtain the *Growth ratio* histogram (fig. (h)), where $Growth[t_{inhib}] = \frac{|V_{C_{inhib}}|}{|V_{C_{steep}}|}$ and the *Time difference* histogram (fig. (i)), where $TG[t_{inhib}] = t_{inhib} - t_{steep}, \forall t_{inhib} \in I_C, \forall C \in C_{all}$. We use maximum likelihood estimation on the probability distribution of these two attributes to obtain the estimated thresholds of *Growth* and *Time difference*. We set the first time point t in the lifecycle of cascade C , whose $Growth[t]$ and $TG[t]$ cross the estimated thresholds, as the final t_{inhib} for C shown in fig. (j). Note that t_{inhib} in each cascade depends on when t_{steep} occurs.

2. There will be multiple regions in the same cascade with nearly equal slopes
- though most intervals do not fall into our described category of *inhibition interval*.

Put together, the first order derivative approach does not incorporate sufficient information about the time taken by users to adopt the cascade and only takes into account the cumulative size of cascade at each time point which is insufficient to identify the *steep* and *inhibition* intervals. We identify these intervals in a three-step process, which we provide technical details for in the Appendix . This process, illustrated in Figure 2 is described intuitively below:

1. Based on recent findings that relate point processes to network diffusion (i.e. [28][4], we calculate the Hawkes intensity at each reshare time point rt , which is a function of the cumulative number of adopters till rt as well as a function of the number of adopters at previous time points and the time taken by the users to adopt the cascade C . We convert this curve into *Hawkes interval curve* shown in Figure 2(c) by summing the intensities of time points in each interval in K_C .
2. We then identify intervals with local maxima (which are candidates for the *steep* interval) and local minima (which are candidates for the *inhibition* interval).
3. Based on ideas from [4], we then use a maximum-likelihood approach to filter the points in step 2 shown in *Step 3* of Figure 2 to obtain the parameters that would be used to infer the *steep* and *inhibition* time points of the new cascades.

We follow the above three steps for identifying the *steep* and *inhibition* time points of the rest of the cascades in the corpus except that we directly use the parameters obtained in Step 3 instead performing a maximum likelihood approach. At the end of this procedure, we identify the time points rt_{steep} and rt_{inhib} , which we use for social network analysis.

3.2 Network Evolution Model

Previous works on cascade dynamics have mostly focused on taking static networks or subgraphs representing the cascade and then analyzing them over several features [43]. One of our main contributions has been to study how the cascade network evolves over the time windows specially in the regions preceding the *steep* and the *inhibition* regions and what characterizes those changes. We consider all the temporal networks as the cascade grows until we reach t_{steep} and t_{inhib} . We denote the subsequences containing those two points as τ_{steep} , τ_{inhib} and the networks N_{steep} and N_{inhib} respectively. We note that τ_{inhib} is not necessarily the last subsequence in the cascade, as there may be few more reshares before the cascade finally dies down but since we are interested in the network dynamics in the time windows leading to the *inhibition* region, we focus on intervals preceding τ_{inhib} .

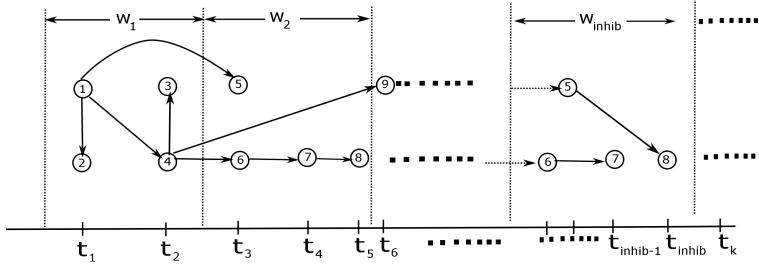


Fig. 3: Example showing how the network is partitioned over windows for feature analysis. Observe that w_l denoting the last window considered in the analysis contains t_{inhib} . Each w_i contains the same number of nodes equal to 4 in this example. Also observe that same timepoints can have multiple reshares for example, t_2 and t_3 both have 2 reshares.

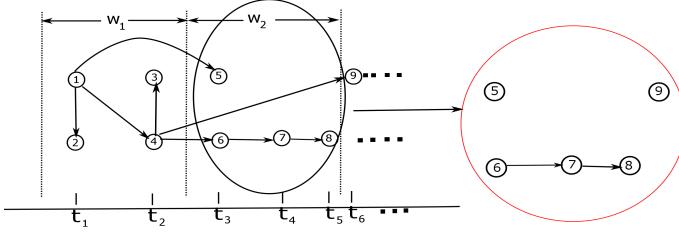


Fig. 4: Example to show the forest structure of $N_i \forall i > 1$. Here we show an example w_2 whose structure would be 3 components with 2 single nodes as individual components.

Figure 3 depicts the way we consider the evolution of the network for studying the cascade dynamics leading to inhibition.

The diffusion network $G_D = (V_D, E_D)$ is created by linking any two users who are involved in a microblog reposting action in time periods preceding the initial posting time of all cascades in our analysis, details of which will be provided in the experiment section.

Since the cascade topology resembles a tree structure, the graphs mapping $N_i \forall i \in [2, W]$ to the topological space would all depict a forest or a disjoint set of trees except N_1 which contains the source node of the cascade. An example of this is shown in Figure 4.

To handle the issue of too many disconnected components which might impact our network analysis, we add the edges from E_D to N_i . Considering each N_i , for any pair of individuals $e = (u, v) \forall u, v \in V^{N_i}$, such that $e \notin E^{N_i}$, we perform the following operation: $E^{N_i} = E^{N_i} \cup (u, v)$ if $(u, v) \in E_D$.

An example of this procedure is shown in Figure 5 where the left figure denotes the original window network structure and the red edges in the right figure denote the diffusion edges. The right figure is the final window network that we consider for network analysis.

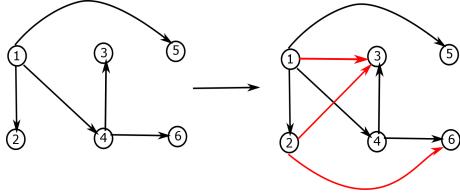


Fig. 5: Example to show the addition of diffusion edges to the cascade network structure.

Table 2: Motif patterns of size 4

Motif ID	Pattern	Edges	Density	Motif ID	Pattern	Edges	Density
m_1		3	0.5	m_4		4	0.67
m_2		3	0.5	m_5		5	0.83
m_3		4	0.67	m_6		6	1

4 Measurements for inhibition

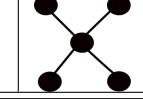
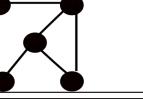
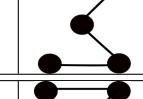
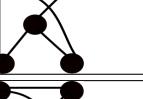
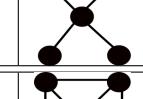
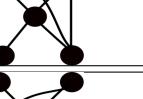
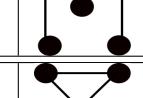
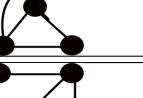
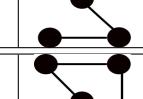
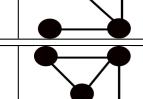
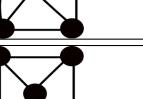
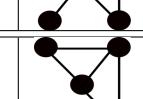
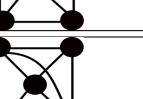
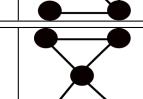
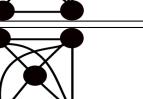
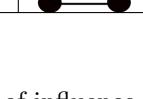
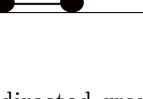
We consider each temporal network N separately and compute the features described in the Section 4 over these networks and observe how the values change as the cascade progresses.

4.1 Motif based network features

“Network motifs” are patterns of interconnections occurring in complex networks at numbers that are significantly higher than those in randomized networks (significance here being determined by z -scores with the null model being motifs extracted from a random network with the same degree sequence as the original network, as described in [20]).

We use network motifs to study how social interactions are driven by individuals forming small patterns of graphs within themselves and how these patterns occur over the windows of the cascade. Studying such group dynamics using cliques and communities has been a subject of research as done in [45]. Similar to the study in the paper where authors use clique percolation algorithm to study communities, we use several motif features to investigate how groups of nodes direct the network change over time. Since we are more interested in finding the patterns of subgraphs rather than the

Table 3: Motif patterns of size 5

Motif ID	Pattern	Edges	Density	Motif ID	Pattern	Edges	Density
m_1		4	0.4	m_{11}		6	0.6
m_2		4	0.4	m_{12}		7	0.7
m_3		4	0.4	m_{13}		7	0.7
m_4		5	0.5	m_{14}		7	0.7
m_5		5	0.5	m_{15}		7	0.7
m_6		5	0.5	m_{16}		7	0.7
m_7		5	0.5	m_{17}		8	0.8
m_8		6	0.6	m_{18}		8	0.8
m_9		6	0.6	m_{19}		9	0.9
m_{10}		6	0.6	m_{20}		10	1

direction of influence in patterns as used in [12], we use undirected graph motifs for our features by converting the directed network N_i into an undirected graph for all $i \in [1, W]$.

For each of the networks N_i , we extract undirected graph motifs of size 5. We choose the motif size based on the observation that with 3-sized and 4-sized motifs, the number of possible patterns occurring from the permutations on edges, are too low to conclude anything significant about the contribution of such motifs to the cascade growth.

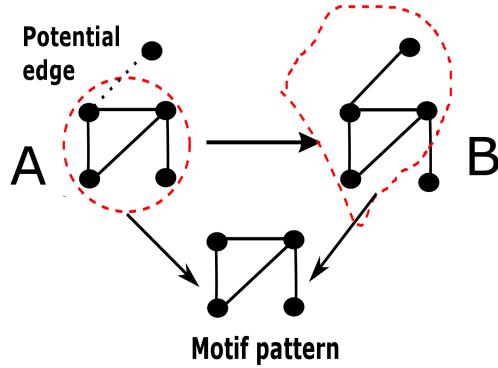


Fig. 6: Role of motifs in a diffusion process

As mentioned before, here we are more concerned about motifs as patterns of subgraphs. As shown in Figure 6, notice that in the transition from A to B, there is a high probability that the new user would adopt the cascade diffusion as the new edge would lead to the motif pattern. Motifs closely replicate small patterns important for diffusion in a network, as illustrated in Figure 6, which are otherwise not captured by topological patterns like nodes with large shell number or high degrees.

The list of all motif patterns of size 5 is shown in Table 3 and all patterns of size 4 are shown in Table 2.

4.1.1 Motif counts

Let M_{ms} be the global list of all motif patterns considering all cascades where ms denotes the number of nodes in the motif instances belonging to the set of patterns (which we call motif size hereon). Going along the notation of windows, $M_{ms}^{N_i}$ represents the list of motifs of size ms found in the network N_i of a particular cascade. So $M_{ms}^{N_i} \in M_{ms}$. Similarly we use the shorthand notation of the letters, m_i , $i \in [0, |M_{ms}|]$ to denote a particular pattern within the list of motifs.

To examine the occurrence of particular motifs in specific time regions preceding w_{inhib} in the cascade, we observe the count of motifs of each pattern m_j of size ms which we denote by $MC_{ms}^{N_i}[m_j]$, that occurs in network $N_i \forall m_j \in M_{ms}^{N_i}, \forall i \in [0, W]$ where the notation W holds its original meaning in a cascade.

4.1.2 Motif weights

Weighted networks in which we assume a weight on each edge has been used in social network analysis for sometime now [44]. It has been used to mine heavier subgraphs in networks [46] where the 'heaviness' parameter is based

on an edge weighting scheme and also for quantifying the influence exerted by one individual on another in the network. Formally we define an edge weight as a function $w : e \rightarrow \mathbb{R}$ where $e \in E$ is an edge in the graph $G = (V, E)$. The idea of using weights to determine motif weights as will be described below is to measure the importance of particular motif patterns within the temporal networks in the cascade lifecycle. This stems from the observation that while motif count denotes the importance of motifs from the frequency perspective, it is not easy to say whether all motifs carry the same importance in individuals forming groups to drive the cascade growth since a single node might be more influential in forming multiple motifs of the same or different patterns in a network, in which case we would assign greater weight to those motifs formed by that individual. To this end, we model edge weights in the temporal networks below using a generative edge weights and then using them to measure the importance of weighted motifs.

4.1.2.1 Edge weight model

To incorporate edge weights in the network for each N_i , we adopt the likelihood model for network evolution used in [47]. The model proposes two hypotheses for edge creation.

Edge attachment by target node in-degree: The preferential attachment model [48] which posits that the probability that an edge chooses a target node is proportional to the node's degree. Using this idea, we compute the final in-degree of each target node in the diffusion network. Let d_{tgt} denote the final in-degree of tgt in the diffusion network. Figure 7 shows the distribution of the final in-degree of the target nodes.

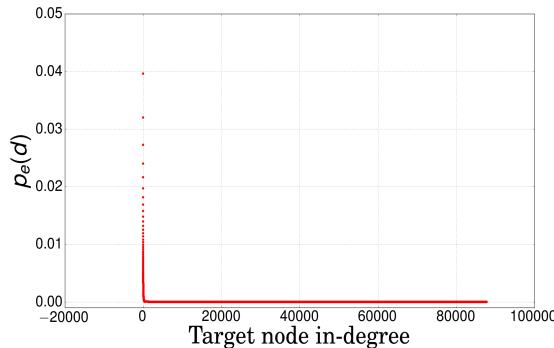


Fig. 7: Probability $p_e(d)$ of a new edge e choosing a target node of in-degree d .

Edge attachment by source node age : The hypothesis used in [47] which proposes that an edge is more biased towards an older node as its source

since more experienced users are supposed to be more influential. Let $rt_u[i]$ denote the time when a user u reshares or posts for the i^{th} time in the diffusion network. We define the node age of u as $na_u = rt_u[t_l] - rt_u[0]$ where t_l denotes the last time a user was active in the diffusion network. Figure 8 shows the distribution of the source node's age.

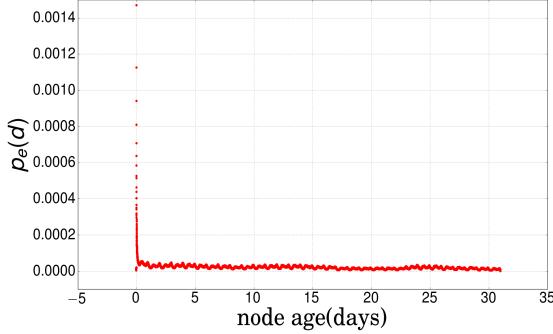


Fig. 8: Probability $p_e(a)$ of a new edge e choosing a source node of age a .

Maximum Likelihood principle: Using the maximum likelihood principle, we devise a way to compute our weight likelihood function. We propose that for each stage of reshare in a cascade, the likelihood of an edge given the network is directly proportional to the age of source node src , na_{src} raised to the power λ and the in-degree of the target node tgt , d_{tgt} . So $p(e|G_D) \propto (na_{src})^\lambda \cdot d_{tgt}$. We assign the proportionality constant beforehand. For finding the value of λ from the diffusion network data, we take a vector of values for λ and calculate the log-likelihood of that function for those selected values of λ . We take the maximum among these values for our final λ to be used for the cascades.

We note that edges with higher weights would be a result of high influential sources nodes or preferential target nodes with high in-degree or both. For calculating the edge weights for each edge $e = (src, tgt)$ in a network N_i for a cascade C , we proceed with the following steps:

1. We assign $w(e) = 1 \forall e \in E_D$. That is in the scenario of an independent cascade model, we assume that the diffusion edges should not play any role in the evolution of the cascade over time except for adding more influence from the source node.
2. We calculate $na_{src} = \tau_C^{[crt]} - rt_{src}[0]$ where crt denotes current reshare time. We access d_{tgt} from the diffusion network data. We note that we do not increase d_{tgt} after this reshare as it would not affect future edge weights since a target node wouldn't reshare the same microblog more than once in the cascade lifecycle and we do not consider the degree of source nodes even if an individual reshares from t in future.

3. We calculate the edge weight value as $w(e) = (na_{src})^\lambda \cdot d_{tgt}$.

4.1.2.2 Computing the motif weights

Once the edges have been assigned the weights in $N_i \forall i \in [1, W_C]$, we generate the motifs and compute the weight of a motif in the following way: we multiply the edge weights for each edge $e \in E(m_i)$, $\forall m_i \in M_5^{N_i}$. So for each motif instance of a particular motif pattern m_i among M_{ms} , we obtain a weight mapping $MW : m \rightarrow \mathbb{R}$. To aggregate the weights for all motif instances belonging to a pattern m_i , $i \in [1, W]$, to find the final weight for it in a particular network N_i , we compute the following equation (assuming a fixed ms):

$$MW[m_i] = \frac{1}{MC^{N_i}[m_i]} \sum_{k=1}^{MC^{N_i}[m_i]} \left(\prod_{e \in m_i} w[e] \right)^{\frac{1}{|m_i|}} \quad (1)$$

Here $|m_i|$ denotes the number of edges of m_i .

4.1.3 Motif transitions

One of the primary objectives of this study of motifs is to observe how the patterns change over time and one important feature with respect to this purpose is to inspect whether the transition of one group of motif patterns to another is indicative of an approaching *steep* or *inhibition* region.

An example of motif transition is shown in Figure . We extract all 4-sized motifs for N_i and for each motif $m_i^4 \in M^{N_i}$, we extract all 5 sized motifs M^{N_i} . We aim to find the number of 4 sized-motifs in N_i that change into 5-sized motifs in N_i . Since the combinations for this can be very large, we use some pruning algorithm to find transitions for specific set of patterns.

For modeling the transitions, we first extract motifs M_{ms-1} of size $ms - 1$ for each network N_i in a cascade C . We then extract motifs M_{ms} of size ms for each of the networks N_i in C . We then compute the number of motif instances of $m_j \in M_{ms-1}$ that have transitioned into motif instances of $m_k \in M_{ms}$. We define motif pattern transition as a transformation of an instance m_{ji} into an instance m_{kl} such that $V(m_{ji}) \cap V(m_{ki}) \geq 2 \forall i \in MC^{N_{i-1}}[m_j], \forall l \in MC^{N_i}[m_k]$, where $V(m)$ denotes the nodes of motif m . An example of a motif transition is shown in Figure 9.

The major issue in computing this motifs transition procedure is the time complexity of the algorithm defining it. Observe that the maximum number of combinations of the pairs of motifs for transition among all network intervals in a cascade is $|M_{ms}| \times |M_{ms-1}|$, where $||$ symbol denotes the count. To prune only the significant and the relevant transition pairs among all of them, we use two filtering techniques to prune them as described in Algorithm 1.

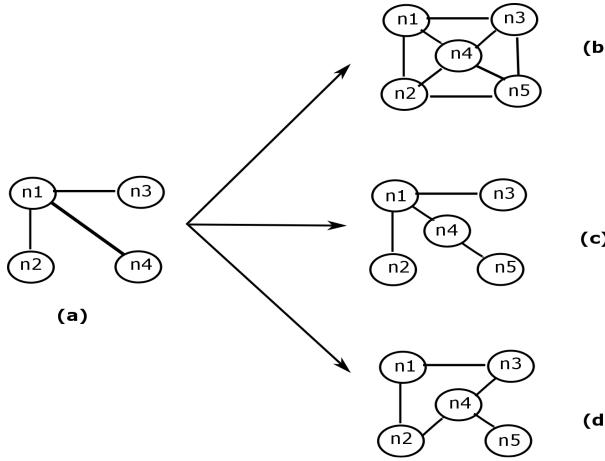


Fig. 9: An example of motif transition. Figure (a) denotes a 4-size motif in some network N_i for a cascade. Figures (b), (c) and (d) show 3 possible transition patterns into 5-sized motifs in the network N_{i+1} . Observe that in this ideal transition, all the nodes in motif in (a) are present in (b), (c) and (d) which may not always be the case.

4.1.4 Temporal motifs

We use temporal motifs to examine the role of out-neighbors of V^{N_i} , which are not yet part of V^{N_i} , but which have adopted C in the next window succeeding N_i . To this end, Let us denote the set of out-neighbor nodes of a node v in the diffusion network G_D as $ON_D(v)$, where the subscript D denotes the network in consideration. Formally the frontiers of a node $v \in N_i$ which we call an adopter node in N_i , is defined as $F_{N_i}(v) = \{v_k \in V^{N_{i+1}} \setminus V^{N_i} \mid v_k \in ON_D(v) \text{ and } (v, v_k) \in E^C\}$. Using the same notation, we denote λ_{lt} -frontiers [1] of a node $v \in N_i$, as $F_{N_i, \lambda_{lt}}(v) = \{v_k \in V^{N_{i+1}} \setminus V^{N_i} \mid v_k \in ON_D(v), (v, v_k) \in E^C, 0 < (rt_{v_k}^C - rt_v^C) \leq \lambda\}$ (refer to Table 1 for the symbols). We extend this notation to define λ_{gt} -frontiers as $F_{N_i, \lambda_{gt}}(v) = \{v_k \in V^{N_{i+1}} \setminus V^{N_i} \mid v_k \in ON_D(v), (v, v_k) \in E^C, (rt_{v_k}^C - rt_v^C) > \lambda\}$.

So the above definitions imply that the set of nodes in N_i who have $|F_{N_i, \lambda_{lt}}(v)| \geq 1$ or $|F_{N_i, \lambda_{gt}}(v)| \geq 1$ are the *frontier-carriers* for N_i and the *frontier-scarce* nodes in N_i are defined as the set of nodes whose $F_{N_i, \lambda_{lt}}(v) \cup F_{N_i, \lambda_{gt}}(v) = \emptyset$, that is they do not have any frontiers.

Using these notations, we define *frontier motifs* as those motifs which have at least one *frontier-carrier* node.

We study the appearance of two motif types for N_i in this regard:

1. λ_{gt} -frontier motifs: These motifs in Figure 10(a) contain at least one *frontier-carrier* node v such that $|F_{N_i, \lambda_{gt}}(v)| \geq 1$ but these motifs do not contain any v such that $|F_{N_i, \lambda_{lt}}(v)| \geq 1$.

Algorithm 1 Motifs transition algorithm

```

1: procedure MOTIFS_TRANSITION
2:   for each  $N_i$  do
3:     motifs_4  $\leftarrow$  motifs_function( $M_4^{N_i}$ )            $\triangleright$  Extract 4-sized motifs for  $N_i$ 
4:     motifs_5  $\leftarrow$  motifs_function( $M_5^{N_i}$ )            $\triangleright$  Extract 5-sized motifs for  $N_i$ 
5:     for each  $m_4 \in$  motifs_4 do
6:       if  $MC^{N_i}[m_4] <$  threshold_count then
7:         Skip this motif and continue
8:       end if
9:       for each  $m_5 \in$  motifs_5 do
10:        if  $m_4$  is not a subgraph of  $m_5$ 
11:          Skip  $m_5$  and continue with other motifs       $\triangleright$  Filtering by subgraph
12:        patterns
13:        end if
14:        if  $MC^{N_i}[m_5] <$  threshold_count then
15:          Skip this motif and continue
16:        end if
17:        for each node_list_4  $\in m_4$  do
18:          for each node_list_5  $\in m_5$  do
19:            nodes_intersect  $\leftarrow$  node_list_4  $\cup$  node_list_5
20:            if length(nodes_intersect)  $\geq 2$  then
21:              transition_array[m_4][m_5]  $\leftarrow$  transition_array[m_4][m_5] + 1
22:            end if
23:          end for
24:        end for
25:      end for
26:    end for
27: end procedure

```

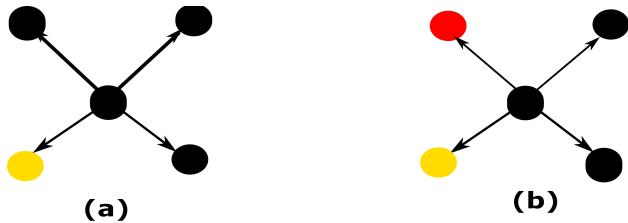


Fig. 10: Two kinds of temporal motifs. Colored nodes indicate a λ -frontier node while black nodes indicate an adopter. Yellow node v indicate $|F_{N_i, \lambda_{gt}}(v)| \geq 1$ while red node v indicate $|F_{N_i, \lambda_{lt}}(v)| \geq 1$.

2. λ_{lt} -frontier motifs: These motifs shown in Figure 10(b) contain at least one *frontier-carrier* node v such that $|F_{N_i, \lambda_{lt}}(v)| \geq 1$.

Using temporal motifs, we observe how the distribution of the *frontier-carrier* nodes and *frontier-scarce* nodes affect the dynamics of the cascade growth.

4.2 Node based features

Centrality of nodes or identification of ‘*key*’ nodes in spread of information has been an important area of research in social network analysis. But in majority of these analyses, measurement of importance through some network statistics are performed on static networks evolving on a cumulative basis. Unlike motifs, where we consider the emergence of groups of nodes as structural measures, here we consider individual node features over time and observe how they change over the lifecycle of the cascade that is to say, whether emergence of crucial nodes with high or low feature values in the middle of the cascade minimizes inhibition or the absence of such nodes augments the rate of cascade decay.

We describe the node-centric features in the following sections.

4.2.1 Degree entropy

We used the idea proposed in [49] to see how the degree of the out-neighbors of a node v affects its influence power when $d(v)$ is in the lower percentile of the set of node degrees of the network, where $d(v)$ denotes the out-degree of v . Traditionally, as proposed in [49], when $d(v)$ is low, influence is a function of the degree of out-neighbors but as $d(v)$ increases, its own influence power dominates that of its neighbors in that more users reshare from v itself.

Similar to their approach, we define the degree entropy as follows:

$$En(v) = - \sum_{j=1}^{|d(v)|} \bar{d}_j \log(\bar{d}_j) \quad (2)$$

where $\bar{d}_k = \frac{d_k}{\sum_j d_j}$, $\forall k \in d(v)$

4.3 Betweenness centrality

Betweenness centrality measures the extent to which a node lies on paths between other nodes. Nodes with high *betweenness* may have considerable influence within a network by virtue of their control over information passing between others. They are also the ones whose removal from the network will most disrupt flow of information between other vertices because they lie on the largest number of paths taken by messages.

4.4 Clustering Coefficient

The clustering around a node u is quantified by the clustering coefficient C_u , defined as the number of triangles in which node u participates normalized by the maximum possible number of such triangles, formally

$$C_u = \frac{2t_u}{k_u(k_u - 1)}$$

where t_u denotes the number of triangle around u and k_u denotes its degree in the network. The main crux in connectivity through forming quick and small clusters has been studied before and pose two important explanations: first while higher clustering suggests groups forming circles to an extent the message does not circulate beyond certain nodes and secondly more small clusters suggest that users who exhibit higher clustering coefficients are more eager to form such loops easily thereby increasing the diffusion rate. To test these two theoretical observations, we use clustering coefficients of nodes as a measure of information diffusion extent.

4.5 Nodal Degree

We observe the nodal degree of nodes as a measure of how connectivity to immediate neighbors can affect the extent of diffusion spread.

4.6 PageRank

PageRank centrality [31] has been used for ranking the spreading capability of users in diffusion networks and till date, most of the research done on PageRank has been on simulation of spreading dynamics to validate it as a predictor. In [32], the authors study the PageRank centrality in relation to stochastic processes and conclude that in general, the PageRank measure does not perform well. Similarly, in [33], Pei et al. study different indicators of influential users, but find that PageRank performs poorly as compared to k-core. PageRank describes a random walk process on the links between the nodes and a parameter p is assigned to each link that decides its traversal.

5 Dynamics of centrality changes

We extend the study of the above node centrality measures over the interval to quantify how the node centrality changes over time.

Definition 1 Node centrality function A node centrality function nc_v of a node v in the lifecycle of the cascade is defined as a mapping $nc_v: \mathcal{T} \rightarrow \mathcal{V}$ such that $\mathcal{T} \in \mathbb{R}$ stores reshare times rt such that $rt[s] = v$ and \mathcal{V} stores the corresponding node centrality values of v at the time instant rt .

Definition 2 Centrality change Given a node centrality measure \mathcal{M} , centrality change between two consecutive temporal networks N_i and N_{i-1} is defined as follows:

$$d(N_i, N_{i-1}) = \frac{1}{|V^{N_i} \cap V^{N_{i-1}}|} \sum_{v \in V^{N_i} \cap V^{N_{i-1}}} nc_v[rt_1] - nc_v[rt_2]$$

where rt_1 denotes the last time a node reshared microblog corresponding to cascade C from v in the time span of N_i and similarly rt_2 denotes the last time in N_{i-1} .

Thus the centrality change measures the average magnitude by which the node centralities change when the network evolves from N_{i-1} to N_i while the size of the network remains the same. Here we focus on 1-norm as the distance measure between the values. The intention of measuring the change in the values as well as the direction of change stems from the first order dynamics principles of N_i where the speed of centrality change can be modeled as:

$$\frac{\partial nc_v(t)}{\partial t} = -C_1 nc_v(t) + g(t) \quad (3)$$

where $m(t)$ denotes the node centrality value of v at time t and $g(t)$ is a forcing function. However in paper, we only calculate $\delta nc_v(t) = nc_v[t] - nc_v[t-1]$ instead of calculating the velocity as the rate of change due to the following two reasons:

1. *Reaction-time bias*: The reaction time for a particular reshare in the lifecycle of the cascade is defined as the time difference between the current reshare and the previous time ordered closest reshare. As we observe in Figure , the reaction time increases near-linearly that is to say, the general trend is that as the cascade approaches inhibition, the reaction time increases. Therefore we do not use $\Delta t = rt_i - rt_{i-1}$ to measure the dynamics as then the node centrality velocities described by such dynamics would also become biased.
2. *Network Size*: Since $|V^{N_i}|$ is fixed for any N_i in the lifecycle of the cascade, so the time period spanning each N_i would vary and hence because of the reaction time bias as mentioned above the velocity would also be biased.

6 Statistical testing

Traditional approaches to quantify the importance of the networked structure of social networks like the one used in Bass model [50] involve using various network features comprising the user friendship network characteristics and the cascade network characteristics to form regression models and then estimating the parameters of the regression model to reject or accept hypotheses using statistical significance measures regarding the network structure as has been done for Youtube diffusion model [51]. The aim of such models is to mainly infer the linear monotonicity relation between the response and the predictor variables without implicitly incorporating time as an element. Such models do not implicitly characterize whether such network features would temporally

be good early indicators of some phenomena like virality or inhibition in the cascading process of the diffusion models. To resolve this issue, causality in time series data [52] has been recently put to practice to quantify the cause and effect. However such non-parametric methods of logic based causality involves computational complexities of high order. Granger causality [42] has been widely used as a parametric model to measure cause and effect in time series social network data as has been studied in [53]. We first introduce the concept of Granger causality as a tool for measuring the contribution of each of the network features toward inhibition. We measure causality as a means of quantifying the impact of network features on the reshare time increase attribute in a cascade for only the node-centric measures. This is because motif features incorporate multiple nodes at a time to study the network structure so it is not conventional to obtain nodal importance out of such measures,

6.1 Granger Causality

Assume two jointly distributed vector valued stochastic variables \mathbf{X} and \mathbf{Y} , we say that \mathbf{Y} does not Granger-cause \mathbf{X} if and only if \mathbf{X} conditional on its own past is independent of the past of \mathbf{Y} . Formally a p^{th} order vector autoregressive model (VAR) takes the form:

$$\mathbf{U}_t = \sum_{k=1}^p \mathbf{A}_k \mathbf{U}_{t-k} + \boldsymbol{\epsilon}_t \quad (4)$$

where $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ constitutes a multi-variate time series where for each time t , \mathbf{u}_t is a real valued n -dimensional(column) vector with components $u_{1t}, u_{2t}, \dots, u_{nt}$, the $n \times n$ real-valued matrices A_k are the regression coefficients and the n -dimensional stochastic process $\boldsymbol{\epsilon}_t$ the residuals, which are independently and identically distributed (i.i.d.) and serially uncorrelated. Using this notation, in the time-domain, Granger causality is motivated by the following: suppose that \mathbf{U}_t is split into two interdependent processes:

$$\mathbf{U}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix} \quad (5)$$

Under a predictive interpretation, Granger Causality from \mathbf{Y} to \mathbf{X} quantifies the “degree” to which the past of \mathbf{Y} helps predict \mathbf{X} over and above by which \mathbf{X} is already predicted by its own past. These comprise two regression models to test for significance of causality of \mathbf{Y} on \mathbf{X} .

$$\mathbf{X}_t = \sum_{k=1}^p \mathbf{A}_{xx,k} \mathbf{X}_{t-k} + \sum_{k=1}^p \mathbf{A}_{xy,k} \mathbf{Y}_{t-k} + \boldsymbol{\epsilon}_{x,t} \quad (6)$$

and

$$\mathbf{X}_t = \sum_{k=1}^p \mathbf{A}'_{xx,k} \mathbf{X}_{t-k} + \boldsymbol{\epsilon}'_{x,t} \quad (7)$$

Then we apply some F-test to obtain a p -value for whether or not Equation (6) results in a better regression model than with Equation (7) with statistically significant better results.

6.2 Obtaining the feature time series

A univariate time series is modeled as a function $\mathcal{F}:t \rightarrow v_t$ where $v_t \in \mathbb{R}$ is a random variable denoting the feature value at time t and in our case t is represented in the discrete time domain.

In this context, we model a social network feature time series for each cascade C as follows: For each reshare, we have the resharer tgt that reshares from the source node src at time rt_i . So for each network feature f among the set of node centralities we have considered in Section 4.2, we have a function $NC_f:rt \rightarrow v_{rt[src]}$, that is to say, for a new reshare at time rt , we record the new centrality value of src at time rt by considering the cumulative cascade network till time rt . But since there can be lot of reshares in a cascade and it would be infeasible to construct the cascade network for every new reshare and then recompute the centralities at every reshare time instant, we resort to computing them over time intervals as has been done in the network evolution phase described in Section 3.3.

Taking these into consideration, we compute the centralities taking each network $N_i = G^{w_i} \cup G^{w_{i-1}}$ into consideration and considering only $NC[v] \forall v \in V^{w_{i-1}}$. This ensures that we take all the windows into consideration and hence all the reshares but this way we avoid doubly considering the overlapped window reshares between two successive temporal networks N_i and N_{i+1} . For reshares which occur at the same time rt we take the mean of $NC[v] \forall v \in V^{[rt]}$ and V being set of source nodes for those reshares.

Using the above procedure, we form the time series NC_f for feature f for each cascade among our corpus which form the set of causal variables we would be using for testing causality. The response variable \mathcal{L} whose effect we would studying for the causal variables is the function $\mathcal{E}:rt_i \rightarrow \Delta t_i$ where $\Delta t_i = rt_i - rt_{i-1}$. So the response variable time series depicts the aging effect of the cascade. We clip our cascade time series data till t_{inhib} that is to say we study the time series cause and effect relationships building till the inhibition region reaches for the cascade.

6.3 Measuring causality

We note that we measure the granger causality effect between a network feature modeled as a time series and the reshare time differences modeled over time as the response variable, over individual cascades unlike in a regression model where we generally assume that the model incorporates the corpus but over individual time points. To measure Granger causality between the series NC_f and E , we first fit the data to a Vector Autoregression model(VAR)

Table 4: Properties of Reposting Network and Cascades

Properties	Reposting Network
Vertices	6,470,135
Edges	58,308,645
Average Degree	18.02
Number of cascades	7,479,088
Number of cascades over 300	7407

with p maximum lags. We then choose the order of the VAR model based on the *AIC* (Akaike Information criterion) to select the best order p_b^C among all lags in the range $[0, p]$ for the cascade C . The following equation gives the two hypothesis for measuring the causality between feature NC_f and \mathcal{E} over a cascade C :

$$\mathcal{E}_t = \sum_{k=1}^{p_b^C} \mathbf{a}_k \mathcal{E}_{t-k} + \epsilon_t \quad (8)$$

and

$$\mathcal{E}_t = \sum_{k=1}^{p_b^C} \mathbf{a}_k \mathcal{E}_{t-k} + \sum_{k=1}^{p_b^C} \mathbf{b}_k \mathcal{N}\mathcal{C}_{f,t-k} + \epsilon_t \quad (9)$$

where Equation (8) represents the null hypothesis and Equation (9) represents the alternate full hypothesis.

Then we use the Wald F-test to test the hypothesis that the coefficients of the first p_b^C lagged values of \mathcal{E} are zero as shown in Equation (8). We note that we do not perform the causality test of \mathcal{E} on NC_f as we are not concerned about that direction of causality. The rejection of the null hypothesis implies a rejection of Granger *non-causality* that is to say, it supports the presence of Granger causality.

7 Dataset description

For building the diffusion network, we use the dataset provided by WISE 2012 Challenge² as has been previously used in [1]. The dataset provides us with user data and the reposting information of each microblog along with the reposting times which enables us to form the cascades for each microblog separately. The directed diffusion graph $G_N = (V_N, E_N)$ is created by linking any two users who are involved in a microblog reposting action within the period May 1, 2011 and August 31, 2011, where a directed edge between two users denotes a reposting action for some cascade C . Similar to most social networks, this network also exhibits a power law distribution of degree [1]. Table 4 shows the statistics of the diffusion network and the corpus of cascades used in our

² <http://www.wise2012.cs.ucy.ac.cy/challenge.html>

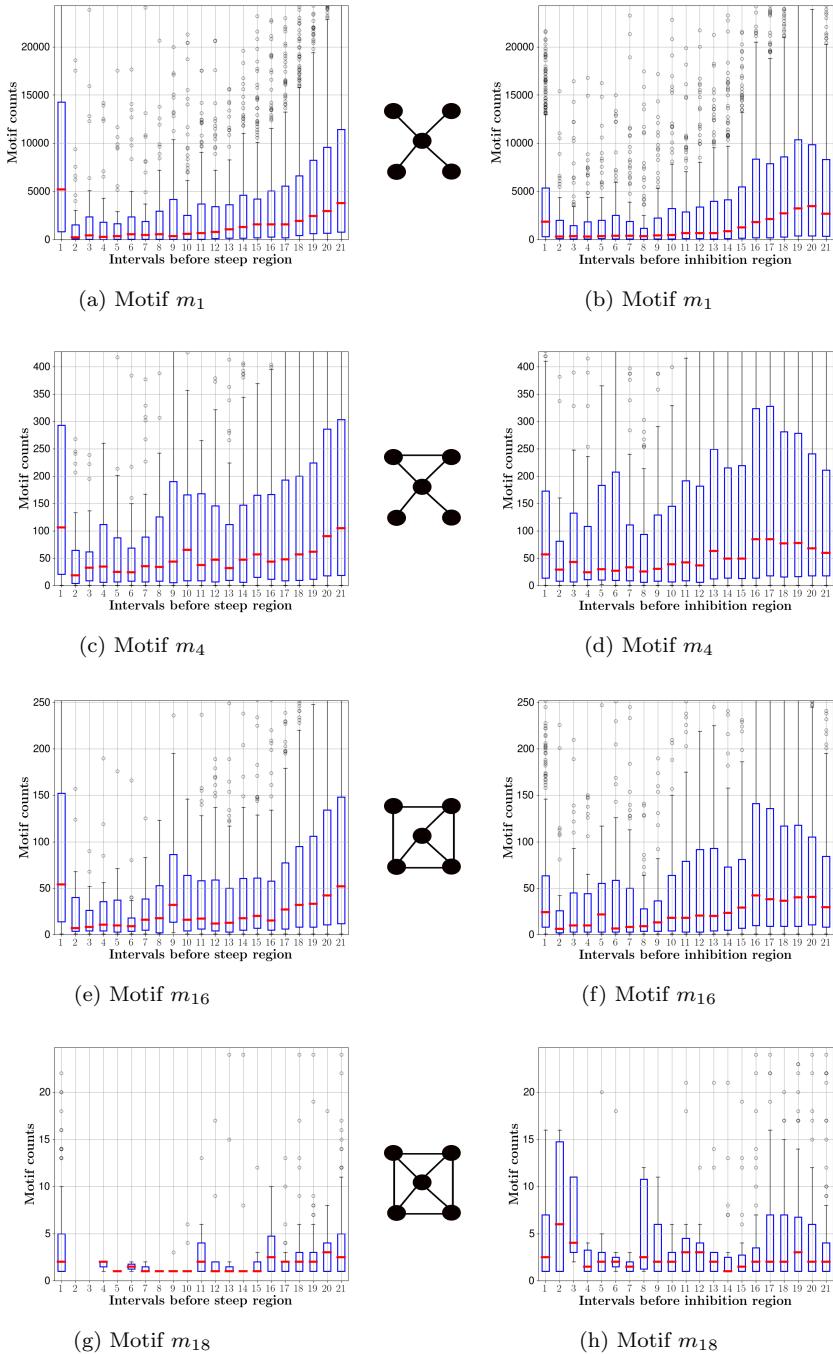


Fig. 11: Plots of motif counts. The left column denotes the plots of the network intervals before w_{steep} . The right column denotes the plots of the network intervals between w_{steep} and w_{inhib} .

experimental study. From the corpus of cascades which spanned between June 1, 2011 and August 31, 2011, we only work with cascades with more than 300 nodes. Since we are considering 50 nodes for forming the cascade subgraph for both *steep* and the *inhibition* intervals as mentioned before, it is more meaningful to have enough nodes for measuring the network features of the cascade.

8 Experiments

Amongst the corpus of cascades, the number of Type I cascades are 5924 while the total number of Type II and Type III cascades are 1483. The total number of cascades of Type I are roughly around 80 % of the total number of cascades that are more than size 300. For dividing the cascade curve C into intervals of size $K^C = \alpha \log(T^C)$ as described in Technical Preliminaries section, we set the scaling factor α to 5 which we found suitable through experimental evaluations, in order to obtain window sizes which are optimal. For finding the suitable α and the parameters from the maximum likelihood method, we carefully choose 1000 cascades from among the entire corpus on which the evaluation is done. The sensitivity of α as a parameter for the optimal window sizes has been studied in detail in Appendix Section 9. We separate Type I cascades from Type II and Type III cascades by setting a threshold of time (t_{th}). We mark those cascades with $t_{steep} \leq t_{th}$ as Type I. We set $t_{th} = 5000$ minutes, as we found that majority of cascades following the Type I pattern in Figure 1(a) experience the steep phase before our selected threshold time. The reason for using this threshold to label Type I cascades instead of using more complex curve fitting methods is two fold: firstly, since the shape of curves even within Type I cascades vary based on when t_{steep} occurs, it is difficult to manually select a set of Type I cascades to obtain the parameters based on logistic fit and secondly, like mentioned before, we observed that the Type I cascades are mainly characterized by situations where t_{steep} occurs within a very short time after the cascade starts and cascades where t_{steep} occurs after a certain amount of time do not exhibit the Type I pattern.

For our social network analysis using the network evolution model, as mentioned in Definition ??, we fix the number of nodes in each window denoted by ws to 40. So $|V^{N_i}| = 80 \forall i \in [0, W]$. However we note that E^{N_i} would vary for each N_i . For evaluation of the intervals leading to w_{steep} and w_{inhib} , we consider the last 20 networks preceding N_{steep} and also for N_{inhib} . This is to ensure that we do not miss out on any time regions that may be early signs of an approaching *steep* or an *inhibition* region.

We make a comparative analysis of the feature values by plotting the measure values in the regions preceding w_{steep} and w_{inhib} . We separate the regions before w_{steep} and between w_{steep} and w_{inhib} to obtain two sets of plots for each of those measures.

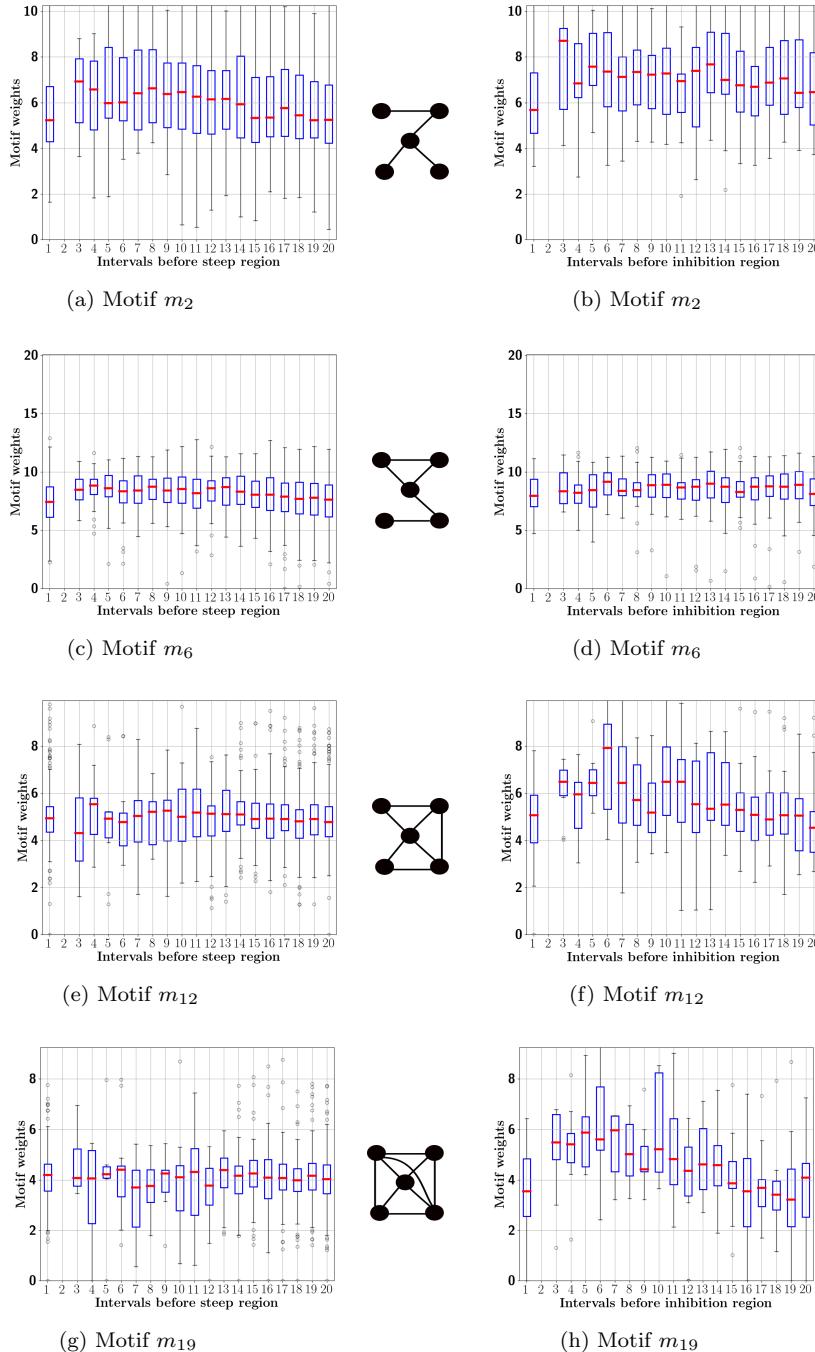


Fig. 12: Plots of motif weights. The left column denotes the plots of the network intervals before w_{steep} . The right column denotes the plots of the network intervals between w_{steep} and w_{inhib} .

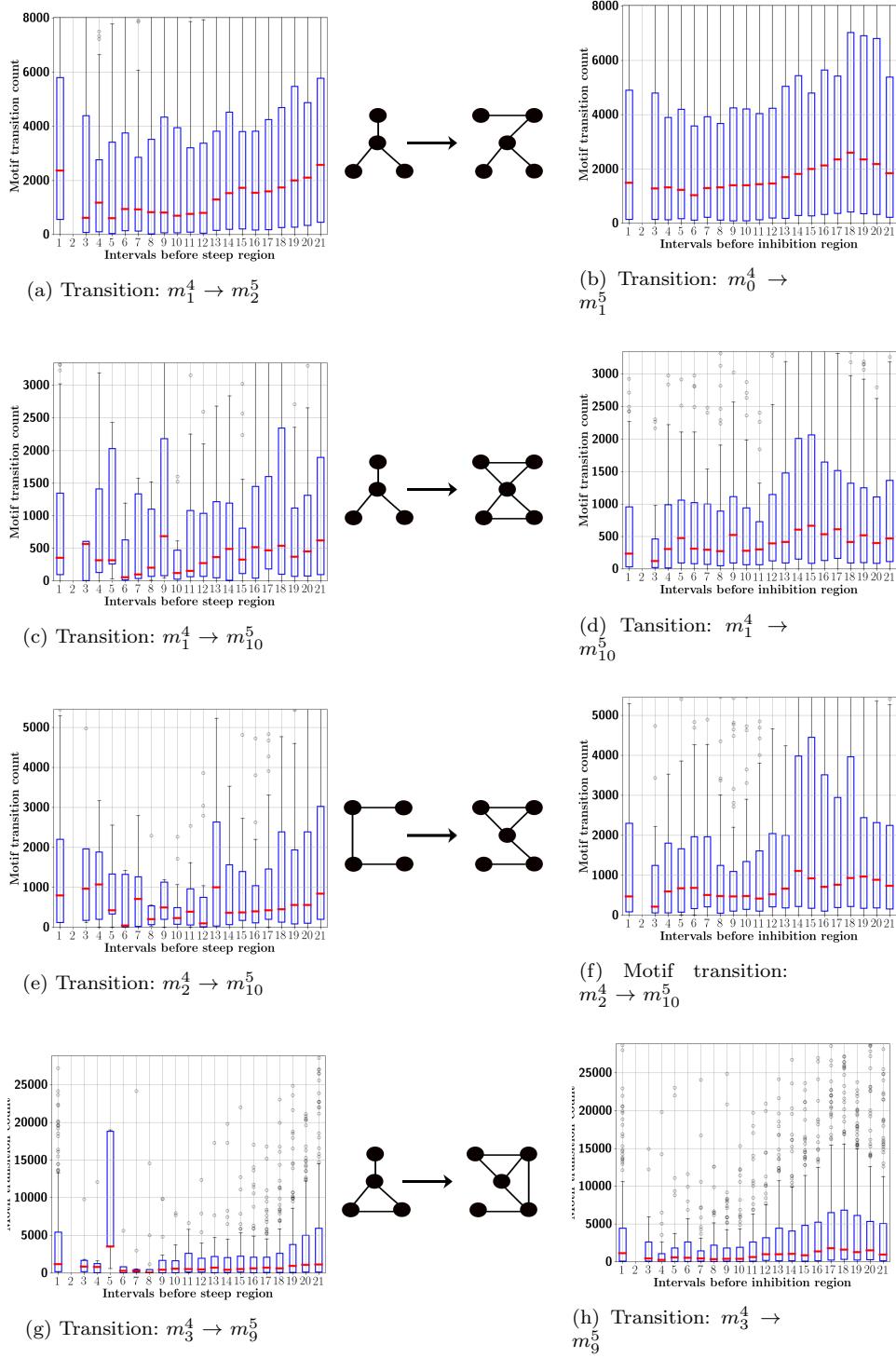


Fig. 13: Plots of motif transitions. The left column denotes the plots of the network intervals before w_{steep} . The right column denotes the plots of the network intervals between w_{steep} and w_{inhib} .

8.1 Motifs

The interesting thing about studying motifs as a tool for measuring structural dynamics is that while we can observe and study how the various motif features change over time for each motif, we can additionally also study how the feature values change within the set of motifs from less dense motifs to higher density motifs for each feature. For each of the features below, we present and compare the results for only those motifs which carry information about the onset of the *steep* and *inhibition* intervals while trying to compare the dynamics of the values between these intervals for each motif. For each of the motif experiments, since it is not possible to display the results of all the motifs, all of which are not necessarily as well, we display the best results from both the less dense and higher dense motifs to explain the network structure and growth.

8.1.1 Motifs count

] We obtained a total of 20 patterns of M_5 which are shown in Table 3 which are listed in order of decreasing significance. We plot the motif counts for the last 20 networks by each motif pattern as shown in Figure ??.

We come across two important observations with respect to motif counts: for motifs m_2 and m_5 which do not have any squares, the network intervals leading to w_{steep} show sharp increases in motif counts with higher quantities whereas although the network intervals leading to w_{inhib} show an increase towards the end of the inhibition region, the motif counts decay after hitting a bottleneck peak towards the end compared to those around steep region. This can be attributed to the fact that in the intervals around the steep region, the adoption mechanism is more spread out in the sense individuals tend to share with new individuals who have not interacted before in the past history in the diffusion network period whereas such a trend is not as dominant in the region around w_{inhib} .

On the contrary, for motifs m_{16} and m_{18} which are densely clustered with more triangles and squares, we observe that although the regions around both the *steep* windows show a steady increase in the motif counts, the motif counts increase are more prominent in the intervals preceding w_{inhib} . Such interaction mechanisms in the dense motifs indicate an impending inhibition region.

8.1.2 Motif Weights

On computing the log-likelihood of the weight function for the vector of parameters, we obtain λ as 0.015 as has been shown in Figure 20.

There are two interesting observations with respective to motif weights: first the intervals preceding the steep region does not show any sharp changes over the intervals for any of the motifs which does not equip us to conclude any such edge preferences towards source node age or target node in-degree for the regions around w_{steep} . However when we look at the motif weights for the

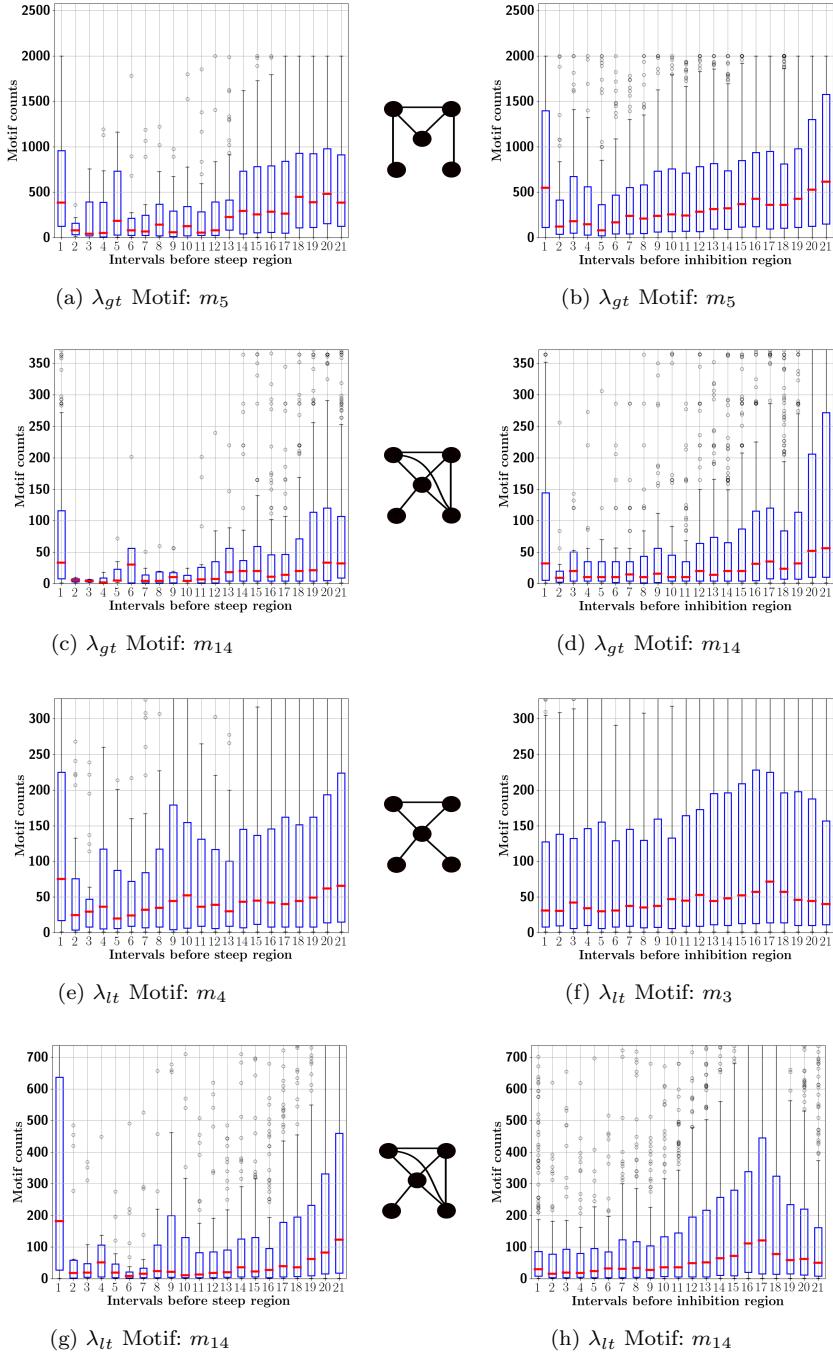


Fig. 14: Plots of temporal motif counts. The left column denotes the plots of the network intervals before w_{steep} . The right column denotes the plots of the network intervals between w_{steep} and w_{inhib} .

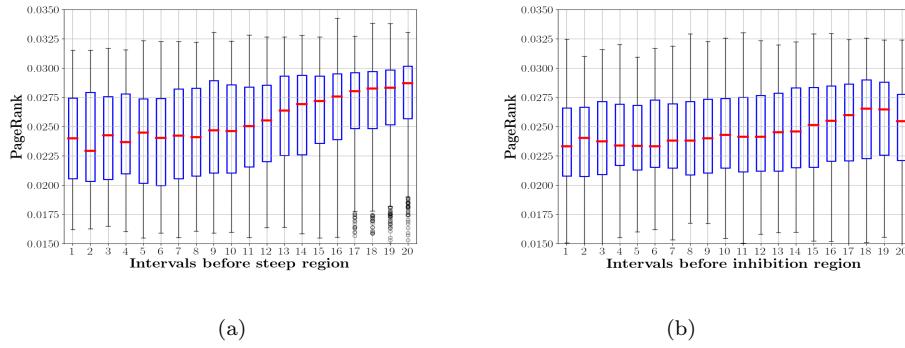


Fig. 15: PageRank centrality values

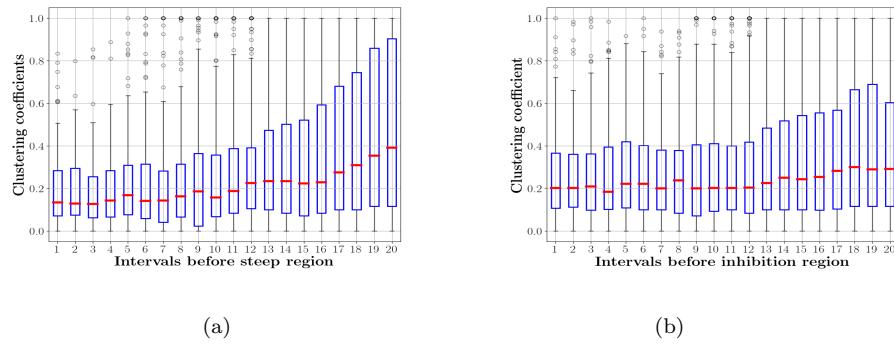


Fig. 16: Clustering coefficient values

intervals preceding w_{inhib} , we observe two different pattern of dynamics, one for motifs with lower density and for motifs with higher densities. As shown in Figure 12, we observe that for motifs with lower density like m_2 and m_6 , there is a general decreasing trend, although not sharp, as the cascades start approaching w_{inhib} . This stems from the fact that the source node age and the target node in-degree are not very influential factors for the transitions between intervals when we compare lesser dense motifs. However when we look at the motifs with higher density like m_{12} and m_{19} we see a very sharp downfall in the weights as the cascades approach w_{inhib} . This points to the fact that for the reshares around the inhibition region, the main infleuntial nodes are the nodes with lesser age or target nodes with lesser in-degree or both, all of which point to the adoption by newer nodes in the network ecosystem during such time intervals. So we see that the denser motifs turn out to be a better indicator about the age and the kind of user population that adapt the cascade towards its end.

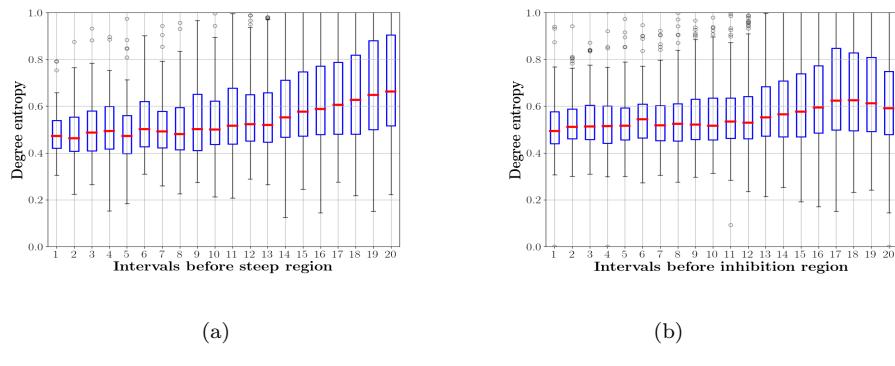


Fig. 17: Degree entropy values

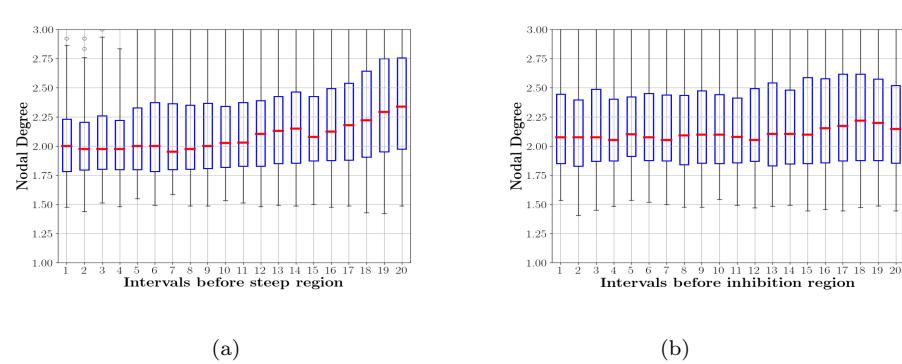


Fig. 18: Nodal degree values

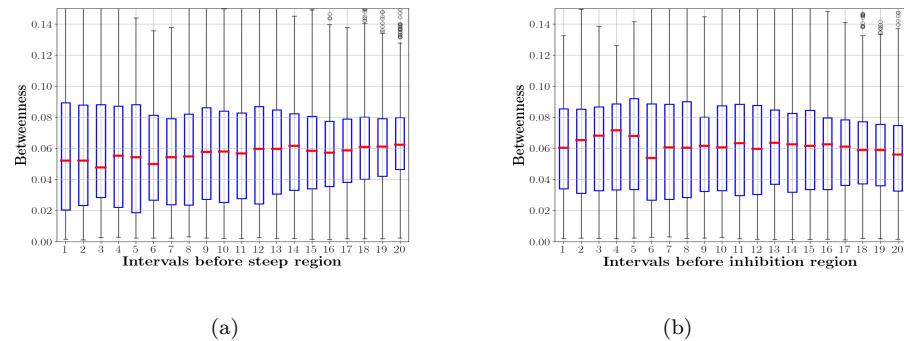


Fig. 19: Betweenness values

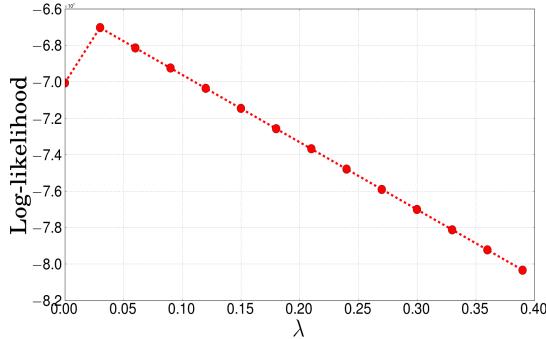


Fig. 20: Log-likelihood for various λ values. As can be seen the maximum reaches at $\lambda = 0.03$ which serves as the final parameter value for the edge weight model.

8.1.3 Motif transitions

To study how the evolution of motifs over the intervals by forming more edges within themselves impacts the diffusion process, we observe two kinds of situations: one where a low density 4 sized motif transitions into a low density 5 sized motif and one where a lesser dense 4 sized motif transitions into a denser 5 sized motif. Since it is not easy for 4-sized motifs of lower density to easily transition into highly dense 5-sized motifs within one interval, primarily because it is not easy to add many edges within a group of nodes within a short period of time, we avoid such extreme transitions in our discussion.

For transitions where a sparse 4 sized motif transitions into sparse 5 sized motifs, for example in Figures 8(a) to (d), we find that while the transition count in the intervals leading to both w_{steep} and w_{inhib} is high, the intervals preceding w_{inhib} show signs of rapid increase before the count starts decaying shown by similar dynamics of the mean values in the boxes. For the other case where transitions from a sparse 4-sized motif to a denser 5-sized motif happen, for example in Figures 8(e) to (f), we find this pattern of dynamics more revealing where there is spiky increase in the transitions in the intervals preceding w_{inhib} before the transitions die down which hint at the possibility of individuals forming small groups quickly before the bottleneck happens. In contrast we see that in the intervals preceding w_{steep} there is generally an increasing monotonicity towards the transition count which concur with the observation of motif appearances as described in Section 8.1.1 above.

8.1.4 Temporal motifs

We first describe the way we select λ time units for computing temporal motifs. The *Reaction-time* bias in cascade reshares described in Section 8 makes it difficult to fix λ in the optimal way. So we resort to increasing λ units similar to

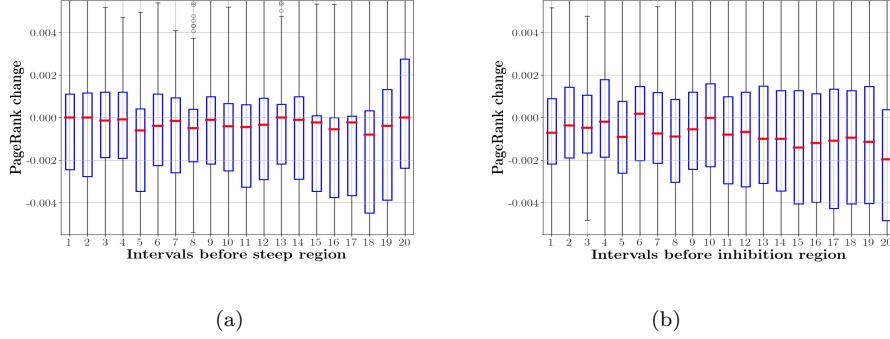


Fig. 21: PageRank centrality first-order changes

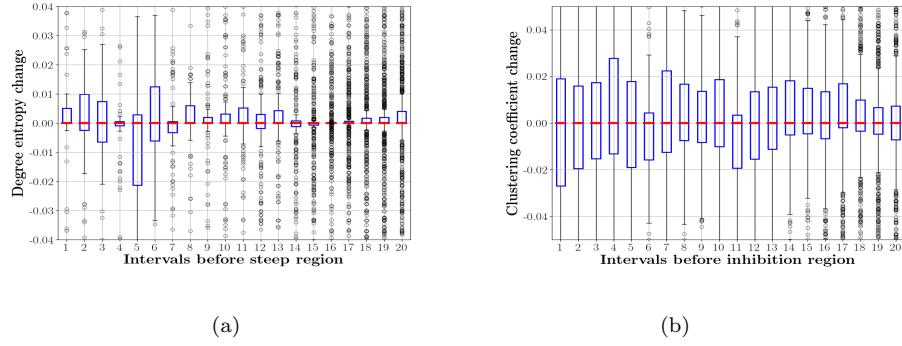


Fig. 22: Clustering coefficient first-order changes

the way we select α while calculating Hawkes intensity described in Section 5. We put $\lambda = \alpha = 5$, as the way to select the nodes' most influential time window. For temporal we observe two the results for the two kind of motifs described in Section 4.1.4. The two motifs explain two different kind of dynamics of motif appearances.

- For λ_{gt} frontier motifs in Figures 14 (a) to (d), which do not contain any λ_{lt} nodes, we find that for both dense and sparse motifs, the count significantly increases which indicates that nodes interact more with their past friends but with not with them within λ time units. This is explained through the rapid increase in the λ_{gt} frontier motifs preceding w_{inhib} .
- In contrast for λ_{lt} frontier motifs as shown in Figures 14 (e) to (f), the trend of interactions follow the more general motif count dynamics as explained in Section 8.1.1.

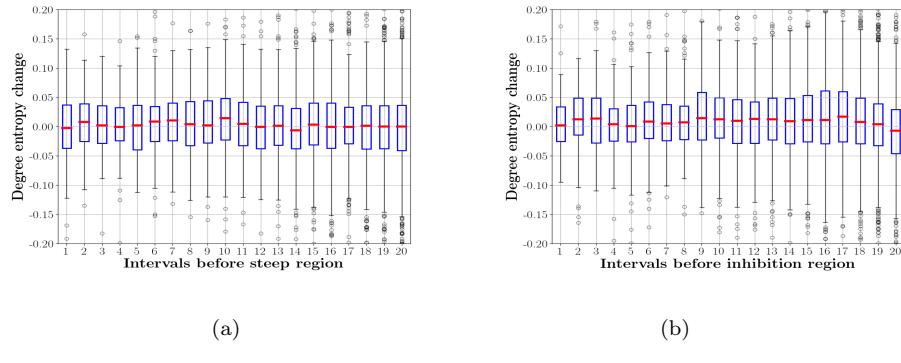


Fig. 23: Degree entropy first-order changes

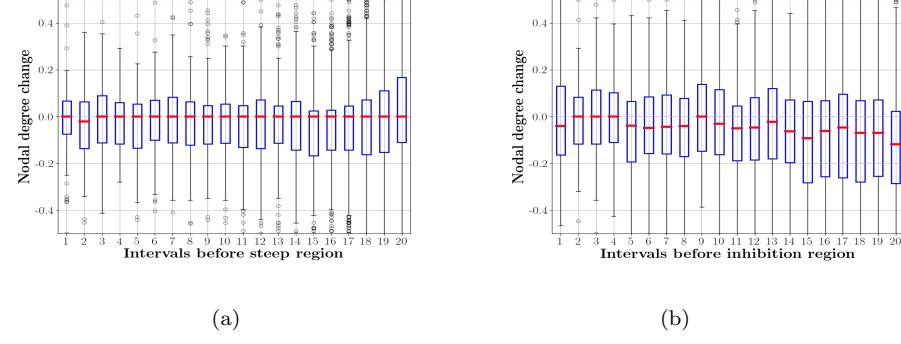


Fig. 24: Nodal degree first-order changes

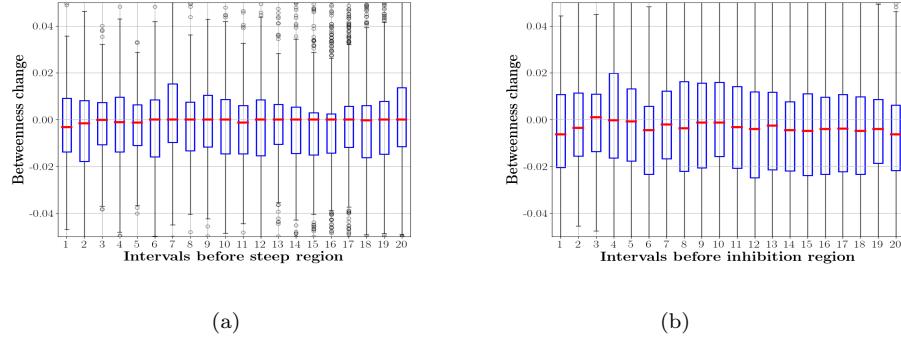


Fig. 25: Betweenness first-order changes

8.2 Node measures

In this section, we first present the basic observations following the box plots similar to the presentation of results for our motifs section. We extend these results to present the statistical results of the causality framework presented in Section 5.

Among all the other structural measures that we examined, the box plots in Figures 15 and ?? reveal that the measures *PageRank* and *degree entropy* of nodes follow the same structural dynamics. While the intervals preceding w_{steep} show a steady increase in the value of the measures, the intervals preceding w_{inhib} show a gradual increase indicated by the increasing means statistics of the boxes before starting to decay.

As described in Section 4.2, clustering has been believed to be an interference in the cascade progress [39]. To test our hypothesis about nodes forming quick clusters being an augmenter of diffusion spread or being a bottleneck in the spread, we observe the clustering coefficient of nodes over the time intervals. As shown in Figure ??, we observe that the clustering coefficient rapidly increases in the intervals preceding w_{steep} , whereas although there may be a slight increase towards the clustering coefficients in the intervals preceding w_{inhib} , the change is predominant.

The features *nodal degree* and *betweenness* do not show any such signs of increasing or decreasing trend before the w_{inhib} although we test both the features for statistical significance for causality. The failure of *nodal degree* shows that the connectivity of the nodes with their immediate neighbors is not as important as the power of the neighbors' degree as would be revealed by the fact that when the degree entropy measure shows small signs of hitting the maximum value before decaying in value, the nodal degree fails to show any signs.

8.3 Node measures' first-order changes

We observe that Pagerank centrality offers the best results in terms of the first order difference. While the nodes' individual pagerank value rapidly increases in the regions preceding w_{steep} , the difference decelerates in the intervals preceding w_{inhib} .

8.4 Statistical testing for causality

The Granger causality results show that the while degree entropy and Betweenness both prove to be the better indicators of the cascade dynamics leading to inhibition, clustering coefficient proves to be the worst performer. One of the many reasons for the failure of clustering coefficients to be a good indicator is that since clustering measures the number of triangles around a node, it requires a large network for evaluation and as such temporal networks

which are constrained by the size of the network prove to be more harmful. A second reason for the poor performance of clustering coefficient as an indicator of inhibition is that unlike in social networks containing all cascades, the traces of individuals are recorded over a period long enough to measure individual tendencies towards large group formation. However in a cascade setting, the influencers keep changing very rapidly so that it becomes very difficult for an individual to form large groups within such a short span.

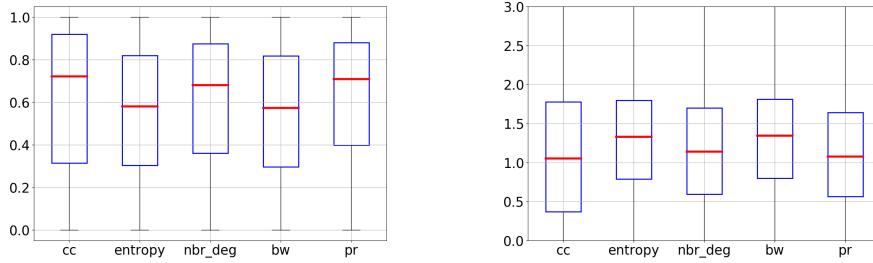


Fig. 26: VAR model causality test results.

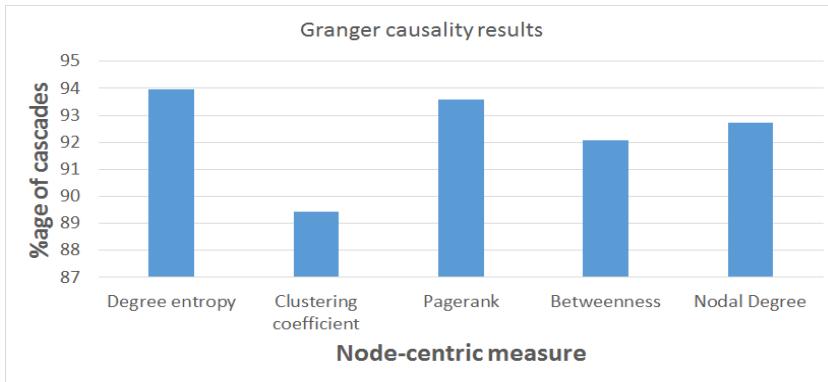


Fig. 27: Granger causality results. The bar plot shows the percentage of cascades where the network feature NC_f granger causes \mathcal{E} .

9 Conclusion

Detecting the inhibition time interval is a complicated task due to lack of ground truth for validation. In this paper we try to formalize inhibition us-

ing point processes and then define certain structural features which identify causes behind inhibition.

Acknowledgements Some of the authors are supported through the AFOSR Young Investigator Program (YIP) grant FA9550-15-1-0159, ARO grant W911NF-15-1-0282, and the DoD Minerva program grant N00014-16-1-2015.

Appendix

A1. Defining steep and inhibition intervals:

In this section we describe the 3-step algorithm used for identifying the *steep* and *inhibition* intervals and that serve as the two intervals for social network analysis for Type I cascades. The lack of ground truth in validating the two regions makes it difficult to identify the regions with accuracy, however we resort to an approach which is unsupervised in a way that we use the information from correctly identified intervals in majority cascades to rectify and remove the incorrect ones.

Now, we briefly introduce the two main concepts in point processes which have been used to model the reposting events of C .

Definition 1: Point Processes Let t_i be a set of random variables $\forall i \in [N]$. Each of these values map a certain time interval in which an event occurs (the event being reposting in our case). The stochastic process quantifying these events defined by these variables t_i gives rise to a point process. The counting process $N(t) = \sum_{i,t_i < t} 1$ is an alternative description of the point process. A point process can be defined in terms of $N(t)$ as below:

$$P[N(t + \Delta t) - N(t)] = \lambda(t)\Delta t + o(\Delta t) \quad (10)$$

Equation 10 says that the probability of an event occurring in a small time interval Δt is proportional to the time-varying intensity function $\lambda(t)$ (which is a probability density function) added to the time-invariant function $o(\Delta t)$. So if $\lambda(t)$ is constant like for example, a Poisson process with $\lambda(t) = \mu$, the process has no memory or the intra-event duration Δt does not depend on previous events and thus is i.i.d. To overcome this shortcoming we use the improvement introduced by Hawkes.

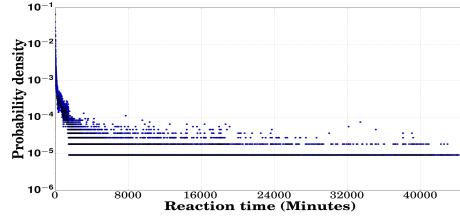


Fig. 28: Probability density function h or memory kernel - Reaction time distribution. The reaction time is plotted in logarithmic axes.

Definition 2: Single variable Hawkes process: Hawkes process is a self-exciting point process that incorporates a response function (or kernel, in fact which is a probability density function) $h(t - t_i)$ which uses the influence of history of events on the current event to define $\lambda(t)$.

$$\lambda(t) = \mu(t) + \sum_{t_i < t, i > 0} h(t - t_i) \quad (11)$$

The first term in Equation 11 ($\mu(t)$) is the base intensity of the model that determines the rate of arrival of first order events per unit of time, which in our case is assumed to be 0 as we consider the final intensity to be independent of this base intensity (or which, without loss of generality could be assumed constant). The response function $h(t)$ is the probability density function of the human reaction times shown in Figure 28. It follows a long-tailed power law with high density only at the beginning. In our work we define the reaction time of a user as the time gap between the time at which the user reposts the blog and the time at which its parent node reposted or posted the blog. This is the kernel or the response function $h(t)$ described in Hawkes equation.

The concept of Hawkes process as a self exciting process can be observed from the fact that each of the previous t_i observations in Equation 11 contribute to the intensity at time t . Our definition of Hawkes intensity at time t is based on the measure used in [28] which defines the rate of obtaining a re-share as below:

$$H[t] = p_t \sum_{(t - \Delta t) \leq t_i \leq t} n_i h(t - t_i), \quad t \geq t_0 \quad (12)$$

where p_t is the infectiousness parameter of the cascade and n_i refers to the out-degree of the node that contributed to the reposting at time t_i . The Δt in the limits of the sum is where we deviate from the normal Hawkes equation in that we do not consider the contribution of all the past events in the intensity value at time t , but only events occurring in the time interval Δt before event at t occurs. Using experimental evaluations, we set Δt to 5000 as we found that with higher values, the Hawkes intensity increases to very huge amplitudes in the later half of the cascade lifecycle thereby making it selecting the inhibition points.

The parameter p_t measures the infectiousness of the cascade at time t or in other words it is a parameter that defines the influence of infectivity of events at time t on the intensity λ_t . Learning parametric models to define p_t by maximizing a likelihood function has been studied in [4]. However, in our work we assume it to be constant over time and that this parameter does not affect the reposting event over time. So essentially we assume that it is an interplay of the reaction time of users as well as the user out-degree that contributes to the intensity. The node reaction time PDF h is a distribution of the time taken by the users to adopt that cascade as shown in Figure 28.

The advantage of using Hawkes intensity measure over the normal peak detection method of slope curve for detecting the *steep* and *inhibition* intervals is that Hawkes process is driven by spike dynamics and hence any sharp increase or decrease in the growth of the cascade is captured by the Hawkes intensity equation defined in Equation 12.

Algorithm for identifying Steep and Inhibition regions For estimating the steep and the inhibition regions, we follow a three step algorithm:

Step 1 - Obtaining the Hawkes curve: In the first step, we obtain the Hawkes curve H^C using the cascade curve C and which would be used for subsequent estimations of the steep and inhibition regions.

In our work, the interval size Δt , the number of time intervals before time t_i for calculating the Hawkes intensity at time t_i is defined as $\Delta t = \alpha * e^{\frac{t_i}{T^C}}$, where α is a scaling parameter which is constant in our case. The important observation in the way we define Δt is that it increases with increasing t_i , the time for which Hawkes intensity is calculated. We use such a kind of penalty measure for increasing t_i to account for the increase in the human reaction time as t_i increases in the cascade lifecycle. However since very large intervals also would hamper the Hawkes values, we freeze the interval size after Δt grows larger than a certain time threshold.

Step 2: First estimates of Steep and the inhibition time points: For estimating the peak points or the points denoting maximum growth or a sharp downfall in the Hawkes curve, we first divide the entire cascade curve into K intervals of fixed window size given by $K^C = \alpha * \log(T^C)$, where α is the same scaling parameter used in Step 1 of the algorithm. We then sum the Hawkes intensity of the time points within each interval so as to get a single Hawkes value for each interval. We denote this new curve obtained after the split as H_I^C . Here on, every time point in the Hawkes curve H_I^C we talk about, would refer to the start time of the corresponding interval in H^C .

The idea of identifying the *steep* interval and the *inhibition* interval can be derived from the fact that the points of local maxima in H^C map to regions in C where there was a surge in reposting compared to previous time points and similarly points of local minimum in H_I^C point to regions where there is a decrease in activity relative to its previous time intervals in C . So H_I^C closely

resembles the slope of the logistic curve of C, where each point in $H_I^C[t]$ would resemble the derivative at time point t .

After obtaining H_I^C , we perform a 2-step procedure for obtaining the estimates of the steep and the inhibition intervals. For detecting the points of local minima and local maxima, we filter out points in H_I^C which are larger than their immediate neighborhood or, $t_i > t_i - \Delta d$ and $t_i > t_i + \Delta d$ for local maximum and $t_i < t_i - \Delta d$ and $t_i < t_i + \Delta d$ for local minimum $\forall t_i \in H_I^C \forall \Delta d \in 1, 2, 3$. However this step identifies a lot of outliers as peak points. To remove the outliers, we calculate the moving mean curve M^C of the intensities for each time point t_i in H_I^C based on the previous $\Delta k = 300$ minutes interval of t_i . Based on this simple technique of finding the maximum among neighborhood, we first find all the local maxima and local minima points. However, since the *steep* interval shows the maximum rate of growth in the cascade lifecycle through our empirical observations (for Type I cascades), we set the global maximum of H_I^C to be the time point identifying the *steep* interval.

However the same cannot be concluded about the inhibition time points. Therefore using M^C we filter out all the inhibition time points which are greater than the Hawkes value at those points, that is to say, we keep only those inhibition points t in H_I^C where $H_I^C[t] < M^C[t]$. The reason behind this step of the algorithm is that the inhibition point t exhibits a very sharp change (and hence lower $H_I^C[t]$) in the reposting rate as compared to its previous time points t_i where the intensities are not lower than $H_I^C[t]$ with $t_i - \Delta k \leq t_i < t$. Therefore any point t which has lower values $H_I^C[t_i]$ will have $H_I^C[t] > M^C[t]$ and hence not an inhibition point. Let us denote all the inhibition points obtained after this step as I^C where $|I^C| \geq 1$.

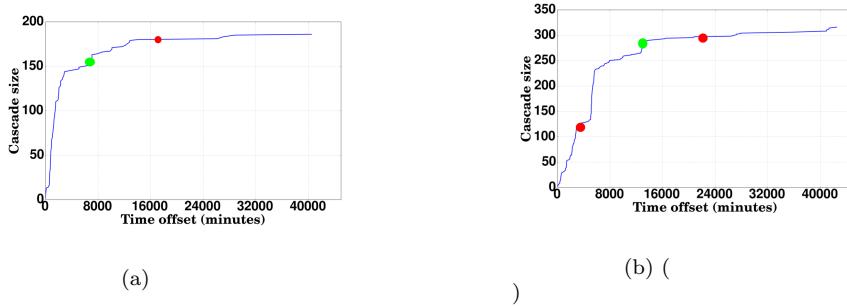


Fig. 29: Two examples of incorrect identification (a) one where the inhibition point gets identified before the actual start. (b) inhibition point gets identified after the actual start. Green points denote the correct t_{inhib} . Red points show the incorrect region which the algorithm identifies as t_{inhib} .

Step 3- Final Estimates of Steep and Inhibition region: As can be seen from Figure 29, even after using the moving mean filter, the output would include multiple inhibition points, some of which are clearly not the

inhibition regions that we desire to obtain, two examples of which are shown in Figures 29(a) and (b). Unlike in a prediction problem, due to lack of ground truth values to verify our inhibition points, we resort to a likelihood estimation of the probable inhibition region.

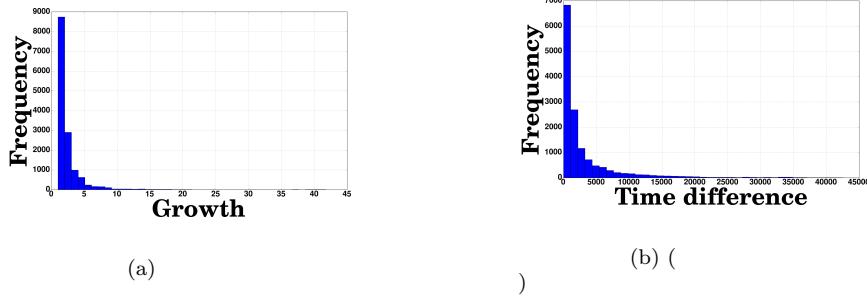


Fig. 30: (a) Growth ratio histogram (b) Time gap histogram

For this, we run Step 2 on all the cascades, and $\forall ic \in I_C \ \forall C \in C_{all}$, C_{all} denoting the set of all cascades in our data, we obtain the following two attributes:

- 1) Time gap $\Delta TG[ic] = ic - t_{steep}^C$
- 2) Growth $g[ic] = \frac{S[ic]}{S[t_{steep}^C]}$

So essentially, ΔTG and g would be a list of values for all the inhibition points for all cascades- the correctly identified as well the incorrect ones.

We minimize the negative log likelihoods of ΔTG and g by fitting a function to ΔTG and g and using an optimization algorithm to obtain the parameters of the function. We use a Poisson distribution $f(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$ with parameters λ and the corresponding log likelihood function $L(\lambda, |x|) = \ln f(x|\lambda)$, where x refers to the values in ΔTG and g . Here we run the optimization procedure twice separately for ΔTG and g to obtain two sets of parameter values. We use the Nelder-Mead optimization algorithm [41] for obtaining the parameters μ and σ . Once we obtain the parameters, we fit them to the Gaussian function used and for each cascade C we get the most probable values TG_p^C and g_p^C through $\arg \max_x f(x|\lambda) \ \forall x \in \Delta TG^C$ and g^C respectively. We select the first point t in the lifecycle of cascade C whose $\Delta TG[t]$ and $g[t]$ are both greater than TG_p^C and g_p^C respectively. We denote that t to be t_{inhib} .

A2. Parameter sensitivity evaluation for α for window scaling:

In our method for calculating Hawkes intensity at time t , we consider Δt time duration preceding t instead of starting from t_0 where $\Delta t = \alpha * e^{\frac{t_i}{T^C}}$. As mentioned before, α controls the window size and to evalaute the sensitivity of α on the final inference of t_{inhib} and t_{steep} , we evaluate the hyperparamters

λ in the Poisson distribution by considering α in the list [1, 3, 5, 7, 10, 15], that is we consider Δt very small as well as extremely large.

The procedure is briefly listed as follows: after we obtain the set of potential t_{steep} and t_{inhib} for all cascades for each α from Step 2, we fit the Growth ratio and Time Gap values each to Poisson distribution and obtain the set of hyperparameters λ using MLE. We plot the values of λ vs α for Growth ratio Figure 31(a) and Time GapFigure 31(b). We observe a monotonically non-increasing relation and as α increases λ decreases. Smaller λ denotes skewed values of Growth whereas very high values of λ denote a very uniform distribution both of which point to bias and randomness. to avoid that we pick the parameter value 5 which is an optimal value in that sense.

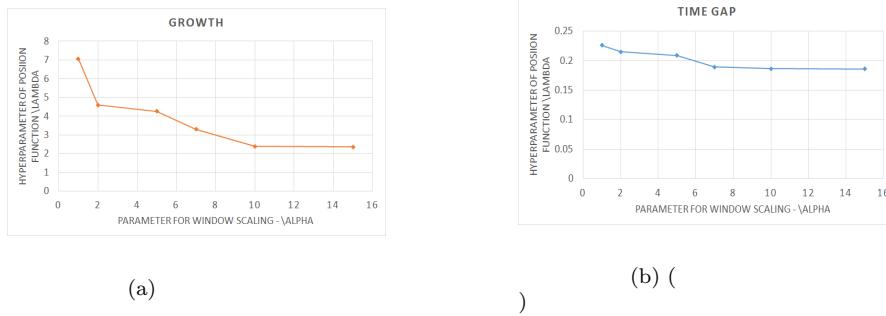


Fig. 31: λ vs α for (a) Growth ratio (b) Time gap

References

1. Ruocheng Guo, Elham Shaabani, Abhinav Bhatnagar and Paulo Shakarian, Toward Order-of-Magnitude Cascade Prediction, in Proc. ASONAM, France 2015.
2. Kempe, David and Kleinberg, Jon and Tardos, Éva, Maximizing the Spread of Influence Through a Social Network, in Proc. SIGKDD, NY, USA 2003.
3. Shakarian, Paulo and Paulo, Damon, Maximizing the Spread of Influence Through a Social Network, in Proc. ASONAM, Washington DC, USA 2012.
4. Liangda Li and Hongyuan Zha, Learning Parametric Models for Social Infectivity in Multi-Dimensional Hawkes Processes, in Proc. AAAI Québec City, Canada 2014.
5. Domingos, P, Mining Social Networks for Viral Marketing, in IEEE Intelligent Systems Journal, 2005.
6. Seo, Eunsoo and Mohapatra, Prasant and Abdelzaher, Tarek, Identifying rumors and their sources in social networks, in Proc. SPIE 2012.
7. Cui, Peng and Jin, Shifei and Yu, Linyun and Wang, Fei and Zhu, Wenwu and Yang, Shiqiang, Cascading Outbreak Prediction in Networks: A Data-driven Approach, in Proc. KDD New York, USA 2013.
8. Yu, Linyun and Cui, Peng and Wang, Fei and Song, Chaoming and Yang, Shiqiang, From Micro to Macro: Uncovering and Predicting Information Cascading Process with Behavioral Dynamics, in Proc. ICDM 2016.
9. Senzhang Wang and Zhao Yan and Xia Hu and Philip S. Yu and Zhoujun Li, Burst Time Prediction in Cascades, in Proc. AAAI, Texas, USA 2015.
10. Greg Ver Steeg and Rumi Ghosh and Kristina Lerman, What Stops Social Epidemics?, in Proc. ICWSM 2011.

11. Paltoglou, Georgios and Thelwall, Mike, Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media, in ACM Trans. Intell. Syst. Technol. Journal 2012.
12. Cheng, Justin and Adamic, Lada and Dow, P. Alex and Kleinberg, Jon Michael and Leskovec, Jure, Can Cascades Be Predicted?, in Proc. WWW Seoul, Korea 2014.
13. Jure Leskovec and Mary McGlohon and Christos Faloutsos and Natalie Glance and Matthew Hurst, Cascading behavior in large blog graphs: Patterns and a model, in Technical report 2006.
14. Chakraborty, Tanmoy and Ganguly, Niloy and Mukherjee, Animesh, An author is known by the context she keeps: significance of network motifs in scientific collaborations, in Social Network Analysis and Mining Journal, 2015.
15. Kai Liu and Cheung, W.K. and Jiming Liu, Detecting stochastic temporal network motifs for human communication patterns analysis, in Proc. ASONAM, 2013.
16. Alon, Uri, Network motifs: theory and experimental approaches, in Nat Rev Genet Journal (Nature), 2007.
17. Kovanen, Lauri and Kaski, Kimmo and Kertesz, Jnos and Saramki, Jari, Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences, in Proceedings of the National Academy of Sciences Journal, 2013.
18. Kai Liu and Cheung, W.K. and Jiming Liu, Detecting stochastic temporal network motifs for human communication patterns analysis, in Proc. ASONAM, 2013.
19. Xuan, Qi and Fang, Huiting and Fu, Chenbo and Filkov, Vladimir, Temporal motifs reveal collaboration patterns in online task-oriented networks, in Phys. Rev. E Journal, 2015.
20. Milo, R. and Shen-Orr, S. and Itzkovitz, S. and Kashtan, N. and Chklovskii, D. and Alon, U., Network Motifs: Simple Building Blocks of Complex Networks, in Science Journal, 2002.
21. Myers, Seth A. and Leskovec, Jure, The Bursty Dynamics of the Twitter Information Network, in Proc. WWW USA, 2014.
22. Myers, Seth A. and Leskovec, Jure, Menger's theorem for infinite graphs, in Inventiones mathematicae Journal, 2009.
23. Moores, Geoffrey and Shakarian, Paulo and Macdonald, Brian and Howard, Nicholas, Finding near-optimal groups of epidemic spreaders in a complex network, in PloS one Journal, 2014.
24. Gallos, L. and Havlin, S. and Kitsak, M. and Liljeros, F. and Makse, H. and Muchnik, L. and Stanley, H., Identification of influential spreaders in complex networks, in Nature Physics Journal, 2010.
25. Jenders, Maximilian and Kasneci, Gjergji and Naumann, Felix, Analyzing and Predicting Viral Tweets, in Proceedings of the 22Nd International Conference on World Wide Web, Brazil, 2013.
26. Sourav Kumar Dandapat and Swadhin Pradhan and Bivas Mitra and Romit Roy Choudhury and Niloy Ganguly, ActivPass: Your Daily Activity is Your Password, in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI, Seoul, Republic of Korea, 2015.
27. G. O. Mohler and M. B. Short and P. J. Brantingham and F. P. Schoenberg and G. E. Tita, Self-Exciting Point Process Modeling of Crime, in Journal of the American Statistical Association, 2011.
28. Qingyuan Zhao and Murat A. Erdogdu and Hera Y. He and Anand Rajaraman and Jure Leskovec, SEISMIC: A Self-Exciting Point Process Model for Predicting Tweet Popularity, in Proc. ACM KDD, Australia, 2015.
29. Masuda, Naoki and Takaguchi, Taro and Sato, Nobuo and Yano, Kazuo, Self-Exciting Point Process Modeling of Conversation Event Sequences, in Springer Berlin Heidelberg Journal, 2013.
30. Zipkin, Joseph R., Schoenberg, Frederic P. and Coronges, Kathryn and Bertozzi, Andrea L.Z, Point-process models of social network interactions: Parameter estimation and missing data recovery, in European Journal of Applied Mathematics, 2016.
31. Page, L. and Brin, S. and Motwani, R. and Winograd, T., The PageRank citation ranking: Bringing order to the Web , in Springer Berlin Heidelberg Journal, 1998 .
32. Rumi Ghosh and Kristina Lerman, Rethinking Centrality: The Role of Dynamical Processes in Social Network Analysis, in Discrete and Continuous Dynamical Systems Series B Journal, 2014.

33. Pei, Sen and Muchnik, Lev and Andrade, José S. and Zheng, Zhiming and Makse, Hernán A., Searching for superspreaders of information in real-world social media, in Scientific Reports Journal, 2014.
34. Saramäki, J. and Kivelä, M. and Onnela, J.P. and Kaski, K. and Kertesz, J., Generalizations of the clustering coefficient to weighted complex networks, in Physical Review E Journal, 2007.
35. Yu, Honglin and Xie, Lexing and Sanner, Scott, The Lifecycle of a Youtube Video: Phases, Content and Popularity, in Ninth International AAAI Conference on Web and Social Media, 2015.
36. Lagarias, Jeffrey C. and Reeds, James A. and Wright, Margaret H. and Wright, Paul E., Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions, in SIAM Journal on Optimization, 1998.
37. Shuang-Hong Yang and Hongyuan Zha, Mixture of Mutually Exciting Processes for Viral Diffusion, in Proc., ICML, Atlanta, USA, 2013.
38. Shakarian, Paulo and Bhatnagar, Abhinav and Aleali, Ashkan and Shaabani, Elham and Guo, Ruocheng, Diffusion in Social Networks, in Springer Publishing Company, 2015.
39. Miller, Joel C., Percolation and epidemics in random clustered networks, in Physical Review E Journal, 2009.
40. Leskovec, Jure and Singh, Ajit and Kleinberg, Jon, Patterns of Influence in a Recommendation Network, in Proc. of the 10th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, 2006.
41. Lagarias, Jeffrey C. and Reeds, James A. and Wright, Margaret H. and Wright, Paul E., Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions, in SIAM Journal on Optimization, 1998.
42. Granger CWJ, Investigating causal relationships by econometric models and cross-spectral methods, in Econometrica, 1969.
43. Mislove, Alan and Marcon, Massimiliano and Gummadi, Krishna P. and Druschel, Peter and Bhattacharjee, Measurement and Analysis of Online Social Networks, in Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, 2007.
44. Newman, M. E. J., Analysis of weighted networks, in Phys. Rev. E 70 Journal, 2004.
45. Palla, Gergely and Pollner, Péter and Barabási, Albert-László and Vicsek, Tamás, Social Group Dynamics in Networks, in Springer Berlin Heidelberg, 2009.
46. Bogdanov, Petko and Mongiovì, Misael and Singh, Ambuj K., Mining Heavy Subgraphs in Time-Evolving Networks, in Proceedings of IEEE 11th ICDM, 2011.
47. Leskovec, Jure and Backstrom, Lars and Kumar, Ravi and Tomkins, Andrew, Microscopic Evolution of Social Networks, in Proceedings of the 14th ACM SIGKDD, USA 2008.
48. Barabási, Albert-László and Albert, Réka, Emergence of Scaling in Random Networks, in Science Journal 1999.
49. Lawyer Glenn, Understanding the influence of all nodes in a network, in Scientific Reports 2015.
50. Frank M. Bass, Comments on “A New Product Growth for Model Consumer Durables The Bass Model”, in Management Science Journal, 2004.
51. Information Systems Research, Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube, in Information Systems Research Journal, 2012.
52. Kleinberg, Samantha, A Logic for Causal Inference in Time Series with Discrete and Continuous Variables, in Proceedings of the International Joint Conference on Artificial Intelligence, 2011.
53. Belkacem Chikhaoui and Mauricio Chiazzaro and Shengrui Wang, A New Granger Causal Model for Influence Evolution in Dynamic Social Networks: The Case of DBLP, in AAAI, 2015.