Moving with the Times: Investigating the Alt-Right Network Gab with Temporal Interaction Graphs

NAOMI A. ARNOLD, Queen Mary University of London, UK

BENJAMIN STEER, Queen Mary University of London, UK

IMANE HAFNAOUI, Queen Mary University of London, UK

HUGO ALEXER PARADA GELVEZ, Universidad Politécnica de Madrid, Spain

RAUL J MONDRAGÓN, Queen Mary University of London, UK

FÉLIX CUADRADO, Universidad Politécnica de Madrid, Spain

RICHARD G. CLEGG, Queen Mary University of London, UK

Gab is an online social network often associated with the alt-right political movement and users barred from other networks. It presents an interesting opportunity for research because near-complete data is available from day one of the network's creation. In this paper, we investigate the evolution of the user interaction graph, that is the graph where a link represents a user interacting with another user at a given time. We view this graph both at different times and at different timescales. The latter is achieved by using sliding windows on the graph which gives a novel perspective on social network data. The Gab network is relatively slowly growing over the period of months but subject to large bursts of arrivals over hours and days. We identify plausible events that are of interest to the Gab community associated with the most obvious such bursts. The network is characterised by interactions between 'strangers' rather than by reinforcing links between 'friends'. Gab usage follows the diurnal cycle of the predominantly US and Europe based users. At off-peak hours the Gab interaction network fragments into sub-networks with absolutely no interaction between them. A small group of users are highly influential across larger timescales, but a substantial number of users gain influence for short periods of time. Temporal analysis at different timescales gives new insights above and beyond what could be found on static graphs.

CCS Concepts: • **Mathematics of computing** → **Time series analysis**; **Graph algorithms**.

Additional Key Words and Phrases: online social networks, temporal networks

Authors' addresses: Naomi A. Arnold, Queen Mary University of London, London, UK, n.a.arnold@qmul.ac.uk; Benjamin Steer, Queen Mary University of London, London, UK, b.a.steer@qmul.ac.uk; Imane Hafnaoui, Queen Mary University of London, London, UK, i.hafnaoui@qmul.ac.uk; Hugo Alexer Parada Gelvez, Universidad Politécnica de Madrid, Madrid, Spain, hugoalexer.parada@upm.esra; Raul J Mondragón, Queen Mary University of London, London, UK, r.j.mondragon@qmul.ac.uk; Félix Cuadrado, Universidad Politécnica de Madrid, Madrid, Spain, felix.cuadrado@qmul.ac.uk; Richard G. Clegg, Queen Mary University of London, London, UK, r.clegg@qmul.ac.uk.

1 INTRODUCTION

Modern online social networks (OSN) foster different social bubbles, with some platforms hosting opinions that would not be acceptable on mainstream venues. Gab is one such platform, with prominent usage from alt-right supporters. Previous studies found that extremism and radical opinions were a cornerstone of the platform [16, 36]. There are concerns that such online extremism has real world consequences [10, 19]. To explore these communities, researchers frequently model a graph from the explicit relationships declared by users, namely followers and group memberships. Since follower relationships can't usually be given an exact starting time [20], this approach usually involves analysing a static graph which conceals any temporal evolution, e.g. the changing intensity of these relationships over time.

In this paper, we focus instead on the communication happening between users; We study the live interactions captured by the replies of users to posts and messages. This crucial information about individual and collective user behaviour is naturally embedded in the temporal dimension and cannot be observed in static graphs. A key advantage of studying the interactions is that we can observe the active, instant usage of the platform. Hence, we model Gab as a temporal interaction graph recording every interaction taking place within the platform and limit the study to that of the structure of the graph formed by the messages, dismissing the content of posts.

An important property of live interactions is that they might be relevant only during a limited time span across the network's existence. We explore this feature by windowing the Gab temporal interaction graph over multiple temporal resolutions (ranging from one hour to one year at a time). Looking at the graph at multiple temporal resolutions allows us to observe different types of dynamics around user interactions [26].

Our approach provides a new way of observing OSN communities. Some researchers characterise OSN as either social (friends interacting) or broadcast (influencers writing to strangers), and the latter may be true of Gab [39]. We investigate this more deeply by looking at the proportion of interactions between user pairs who had not previously communicated. We also investigate influence on Gab by examining whether a dominant group of most important users (measured by number of interactions) persists or whether this influence is short lived.

The complete nature of the Gab data makes this the first paper, to our knowledge, that analyses the interaction graph of a complete social network. We believe that viewing temporal networks at a number of timescales is a tool that can provide considerable insight into Gab and other social networks. Consistently, we found that the insights gained from viewing the interaction graph at one timescale were enhanced by the insights from viewing it at other timescales. Hence, we hope that the techniques used in this paper will prove valuable for other OSN researchers.

This paper makes the following contributions; 1) We present a methodology to analyse OSNs through the lenses of different temporal windows applied to an interaction graph; 2) We explore the Gab interaction network observing trends in its growth over the first eighteen months. We show that the slow growth over long time windows decomposes to rapid rises and falls at the daily and hour timescale driven mainly by events of interest to the user base; 3) We look at the cohesiveness of the Gab userbase, asking whether it forms a single connected group. We find that at a timescale of an hour the network oscillates between being mostly a single connected component at peak hours and being largely fragmented at off-peak hours. We believe this behaviour has never previously been observed; 4) we explore the relative influence of the top Gab users over different time resolutions, observing that a small number of users gather an extremely high level of attention – several users interact with more than 5% of the entire user base every month.

2 BACKGROUND AND RELATED WORK

2.1 Gab Analysis

Gab [1] is a fairly new social media platform, in a lot of ways similar to Twitter, that claims to champion "free speech, individual liberty, and the free flow of information online". Efforts to open accessibility to datasets showing the different characteristics of this platform are continuously provided for the research community [6, 36]. The first venture into studying this network was performed by [36]. They found that Gab attracted predominantly alt-right views and conspiracy theorists and both hate speech and discriminary language were highly common. [16] delves deeper into the characterisation of the network by looking into the nature of Gab users and the type of content being shared. In doing so, they link the unmoderated nature of posts to the emergence of echo chambers around alt-right-leaning content. [11] compared the spreading of the Soros Myth on Twitter and Gab and concluded that, in addition to anti-Semitic content being more prominent on Gab, its users display unabashed willingness to post and share such content. The work of [19] was motivated by the Pittsburgh synagogue shooting in which they analysed the use of language and the topics discussed and their evolution since the creation of the network until the violent attack. They note a shift from neutral terms and site-setup topics to an increase in more extreme offensive language and racist subjects. They also observe a period in time where the popular topics include non-English terms, an indication to the network expanding to a more European presence. A recent analysis [39] further explored this finding by comparing the behaviour of English and German speaking users. They observe that, although small in proportion, the German accounts tend to build tighter connections amongst themselves. A comparison with a Twitter dataset highlighted a more elitist structure to the Gab network with mostly homogeneous content. Most of the research presented so far focused on investigating the content from a language perspective, and little has been done to analyse the structure of the network; more interestingly, the type of interactions that govern this particular platform. One study [2] examined the nature of conversations on Gab by analysing cascades in the corresponding graph. They notice that cascade conversations tend to emerge from linear interactions and evolve towards more complex structures as topics become viral. A recent work by [18] explores the temporal behaviour of the userbase on Gab. By examining the nature of speech, they categorise the users into hateful and non-hateful. They report a fast increase in the number of hateful users and that new users seem to embrace hateful speech quicker. Looking at the structure of the follower graph, they notice that a large number of hateful users are at the core of the network and seem to attract a lot of attention which was steering the conversations towards hateful speech at a faster rate. In here, we aim at exploring the temporal aspect of the Gab network by focusing on the interactions among the users without looking at the content or the follower relationships but rather at the structure of the interaction graphs.

2.2 Temporal graphs and windowing

Graphs are a popular means for modelling relationships between members of online social networks, and are useful for understanding collective behaviours and the spread of information and content on these platforms. There are different models for constructing such a graph depending on the purpose of the study and available data. Much early work in this area focused on static graphs, with edges representing a binary friendship or directed 'following' relationship [21]. Later work coined the phrase 'interaction graph' [34], focusing rather on actual communications between users, with the idea that these might bring more insight into the relationship between users than previously used friendship graphs. Contact graphs are another type where the interaction is associated with a duration [22]. A similar concept of an 'activity

---

[1]http://gab.ai

Table 1. Details on the userbase and interactions within Gab that span the period between August 2016 – May 2018.

Property	Quantity
Number of users	169,745
% of users who only post and never reply	48.5 %
Number of posts	19,091,476
% as original posts	62.5 %
% as replies	37.5 %
% as self-replies	1.53 %

network' was explored in a temporal setting, examining the persistence of interactions between users on Facebook [33] finding that only a minority of declared friendships were maintained with a monthly 'wall post' interaction.

An important question when looking at temporal graphs is that of the appropriate time scale to study. Analysing the evolution of the aggregate graph (every link that occurs before some time t) has an important problem: the aggregate graph grows continually and even relatively large events would be eventually eclipsed. An obvious approach is to study the graph of contacts within a particular 'window'. A small number of authors have investigated how the study of a network depends critically on the time scale at which you study it. In [27, 30] the authors look at bluetooth contact data for 1,000 individuals. They show that studying the data in five minute windows produced identifiable social groups but aggregating this to a day long network blurred these groups into a single connected component. The authors of [5] look at a small contact network (67 individuals) and similarly find considerable variation in network measures using time scales, varying from five minutes to a full day. In [24] the authors consider three graph datasets – two citation networks and one social network – and look at the frequency of pairwise interactions over a variety of window lengths. They observed that these datasets make up sparse, disconnected networks for shorter window lengths, with giant connected components forming as the window lengths became longer. Unfortunately, they were only able to do a limited amount of exploration on three different window lengths. In [13] the authors investigate mobile telephone network data within various window sizes, from one hour to several days, with a particular interest in diurnal behaviour. They used the same null model of shuffled data as this paper and showed that the large connected component could be much larger than in the shuffled model for window sizes of one to two hours. Choice of window size to analyse can affect how systems are built. In [35] the authors build delay tolerant networks modelled using contact graph datasets and note that "too large a window size cannot capture the time-varying features [..] too small a window size may [..] lead to inaccurate calculation of social metrics." This emphasises the need to explore temporal networks at different window sizes since diverse patterns of behaviour would be distinct in one window and obscured in others. The work by [14] offers an upper-limit on the range of possible periods that could be analysed. They looked at whether graph snapshots of a given window size were preserving the propagation properties of the original temporal network. They defined an automatic way to determine a saturation scale by tracking the window that brings the occupancy rate distribution closest to a uniform density distribution. They show that this describes a threshold on the window size where the aggregated graph would lose its temporal dynamics as it relates to properties of propagation if the window was larger than this scale. The work presented here doesn't aim at determining a single window to perform the analysis but rather explores different window sizes to uncover dynamic network patterns that have different rate of change and necessitate looking at the network changes at multiple scales.

3 DATASET COLLECTION AND ANALYSIS

This paper investigates the structural characteristics of the Gab social network as they evolve through time by focusing on user interactions. Our starting dataset is obtained from crawling the Gab network by relying on the Rest API provided and consists of 95 GB of raw data. The relevant details of the dataset are summarised in Table 1. It contains every post between 10th August 2016 (the start of the Gab network) and 5th May 2018. Initially Gab numbered posts consecutively and our crawler tried to access those with numbers from 1 to 24,999,999 (at the time of our crawl this was almost all of the content). Shortly after this an API update made it more difficult to identify new posts as post numbers were then hashes. The methodology in [36] could be used to access these. Some messages were deleted either by the users who posted them or by moderators; the initial crawl accessed posts authored by 169,745 users and 5,908,523 of the 24,999,999 posts were deleted or had empty content. The study [36] found 336,752 users, but includes not only users who post or reply but also users who simply follow others without ever posting. This implies that at least 49% of those users never made a post or interacted via replies. Similar to Twitter, the platform allows users to follow and be followed by other users. However, since the time at which follow events occur is not available the follower graph could only be subject to static graph analysis. Instead of looking at follower/friendship relationships, we rather explore the interactions between the users in the form of posts. Users of Gab share posts, which may consist of text, images or urls, much like Twitter. They can interact by making a *post* or interact with other posts by *reply post*, as well as liking posts or reposting (in the same manner as retweeting). This work focuses on interactions via replies as it is impossible to obtain liking/reposting timestamps from the data. We formally define an interaction as a tuple (u, v, T) where u is a user replying at time T to a post authored by v.

4 METHODOLOGY

4.1 Windowing and graph construction

With temporal graphs, snapshots of graphs are generally built at time t which includes every activity that occurred at time t. When windowing is concerned, graph snapshots within a window of length τ are aggregated together to form a single graph. Formally, a graph $G(t, \tau)$ at time t and window size of τ is constructed by adding a single directed edge e_{uv} between users u and v if an interaction (u, v, T) occurs for $T \in (t - \frac{\tau}{2}, t + \frac{\tau}{2}]$. This means necessarily that for $0 < \tau_1 < \tau_2$, the graph $G(t, \tau_1)$ is a subgraph of $G(t, \tau_2)$. Figure 1 illustrates an example of such windowed aggregation. Notice that, although users A and B interact more than once during this interval, only one edge is added between them. Also, only users participating in an interaction during a window are included as a node in the graph. To observe temporal patterns throughout the history of the network, we construct graphs by sliding windows of size τ and an offset Δ. Unless stated otherwise, for the rest of the work, we use the window sizes of $\tau \in \{1 \text{ hour}, 1 \text{ day}, 7 \text{ days}, 30 \text{ days}\}$ and offset $\Delta = 1$ day.

To enable such high-granularity analysis, we employ an efficient graph processing system called Raphtory [29]. Raphtory models ingested data as a temporal property graph, where the full structural and property history of entities within a network are maintained in-memory. This allows the user to view the exact state of a graph at any point throughout its evolution. A user may additionally apply a 'windowed view' which only returns entities that have been updated within the window size, looking back from the chosen view time. In this fashion, the scraped Gab data was ingested and built into the interaction graph described above. Windowed views were then produced at each day throughout its lifetime, applying the chosen window sizes. For several of the results below, this translates into over 17,000 windowed views of the graph being materialised. This is handled automatically by the tool making it possible
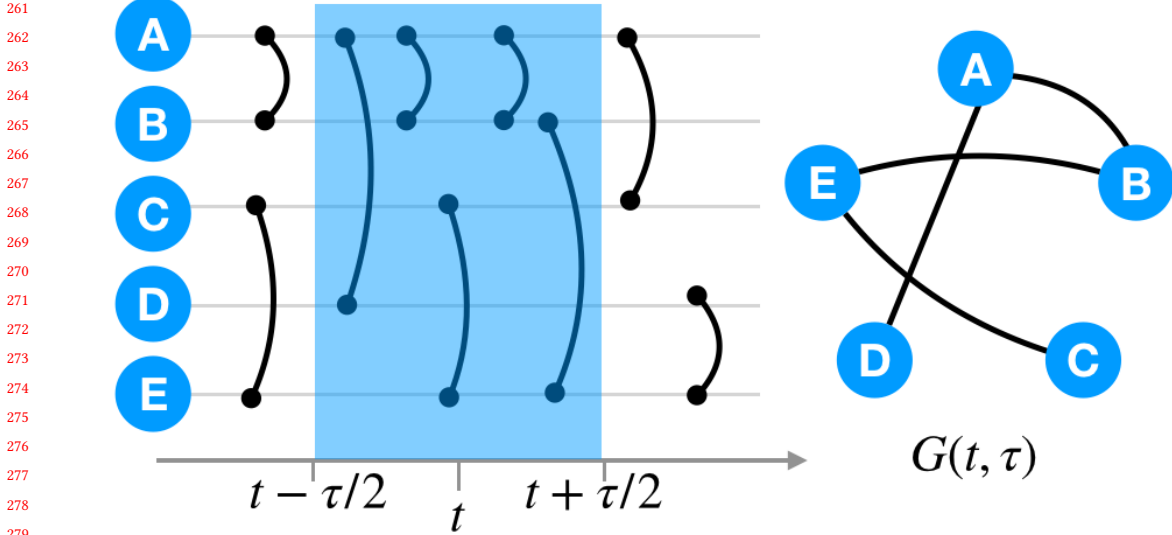
Fig. 1. Windowing procedure to construct the graph $G(t, \tau)$.

to run different analysis algorithms on varied temporal scales and observe patterns that may be hidden with larger window sizes.

4.2 Randomised reference models

We augment our analysis where appropriate with the use of two randomised reference models to investigate potential structural and temporal factors determining the phenomena we observe. Our first reference model is a "shuffled timestamps" model [7, 9]. This model randomly shuffles the timestamps T of all interactions (u, v, T) while keeping (u, v) fixed. In this way, the graph obtained by aggregating all interactions within the dataset is conserved within this null model, as well as the set of interaction times, but removed are the temporal correlations between interactions from adjacent pairs of users. These correlations are common to interaction networks, arising, for example, as burst trains [12] (sequences of repeated interactions between node pairs) or as temporal subgraph motifs [37]. The second reference model we use, commonly known as the restrained randomisation model [17], generates a random graph snapshot with a given degree sequence, usually chosen to match that of the dataset. After creating a graph for a specific window we use this model to create a graph with the same degree sequence. This may be used to test whether observed graph features can be explained just by the degree sequence of the network or whether structural correlations such as assortativity or community structure play a role.

5 RESULTS

Our investigation of Gab considers three important questions about the network. *(i) What is the nature of growth in the Gab network and what drives that growth?* We address this question by looking at how the number of nodes and edges in the interaction network changes over time and the proportion of those nodes and edges that have never been seen before. *(ii) Is the Gab network best characterised a single group of users connected directly or indirectly through*
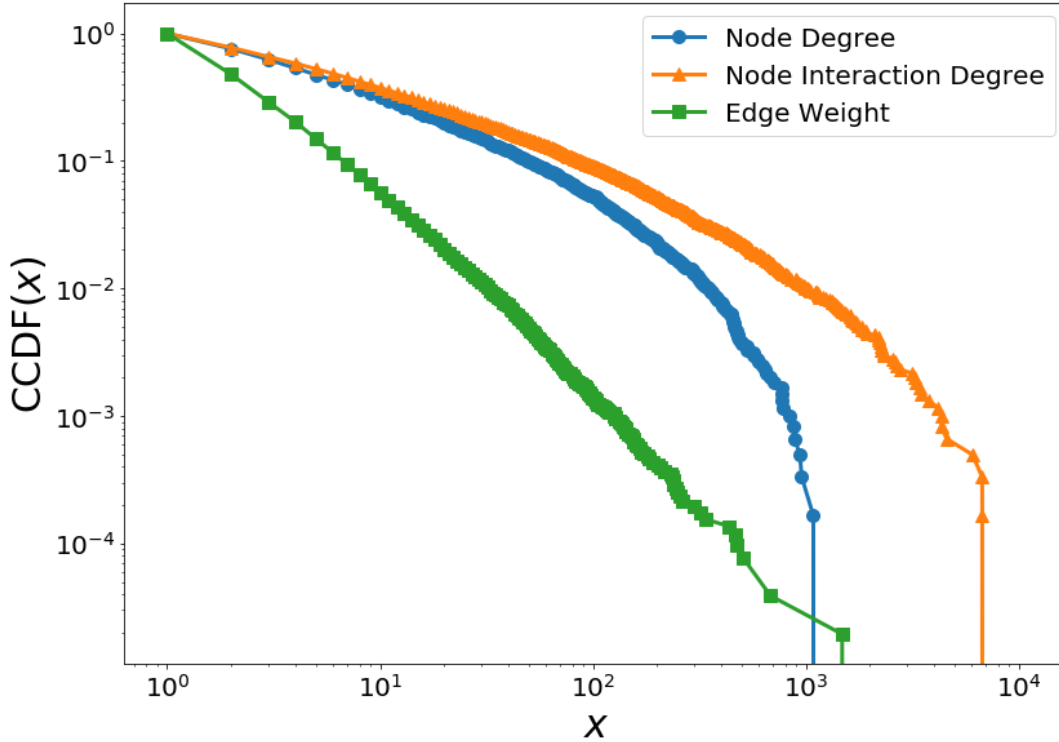
Fig. 2. Distributions of node degree (the number of unique users a user interacts with), node interaction degree (the number of interactions a user is involved with) and edge weight (the number of times a given pair of users interacts).

their interactions or is it a number of disconnected groups? We consider this by investigating the behaviour of the large connected component for the graph. Finally *(iii) we ask whether Gab is dominated by a consistent group of influential users or whether the most influential users change over time.* We address this by looking at the most connected users in the network.

As this paper focuses on interactions rather than declared friendship/follower relations, we first examine how these interactions are distributed across the network, with Figure 2 showing some of these quantities pertaining to interactions. Node degree refers to the number of distinct users that node has interacted with, whereas interaction degree refers to the total number of interactions that node has had. Edge weight refers to the number of interactions between each pair of users, excluding pairs which never interact. We notice that 52% of the user pairs that interact do so once and never again, with a long tail of few edges over which many interactions occur.

In all of these investigations, we analyse how the network evolves as time goes on but also how the network quantities change as we consider different sizes of time window. The latter parameter is important; the larger the time window to investigate the network, the more interactions we will be looking at. Figure 3 shows the change of some basic network properties when varying the size of the window. As would be expected, larger windows include a greater number of nodes and edges. A less intuitive but well-known observation for growing networks is that the longer the time period

considered, the higher the average degree (or edge/node ratio); a process sometimes called densification [15]. A similar pattern was observed when using window based analysis to study proximity networks [5].

5.1 Drivers of Growth in Interactions

An aggregate graph viewed over time can only ever add nodes and edges and may not reveal clear patterns in the temporal behaviour of these entities. In this section, we use different window sizes to explore whether we can identify factors affecting growth or decline in these interactions. We use the definition of nodes and edges within a window τ which make up the graph $G(t, \tau)$ defined in the previous section. We distinguish a "new" node or edge within a window as entities that have never been seen at any time before this window. In other words, these represent users and pairwise interactions that occur in the window $T \in (t - \frac{\tau}{2}, t + \frac{\tau}{2}]$ and were never active before $(t - \frac{\tau}{2})$. Adjusting the window size τ provides different definitions of who is an "active" user. We investigate whether the system is driven by returning users who continue to contribute over time or new users who have never been seen before the time window under analysis.

Figure 4a shows the number of users according to daily, weekly, monthly, quarterly and 6-monthly window sizes (bottom), and the proportion of those users that are new (top). In addition, the running total (aggregate) number of users is plotted as a reference. The number of users that are active monthly or more regularly makes up a small proportion of
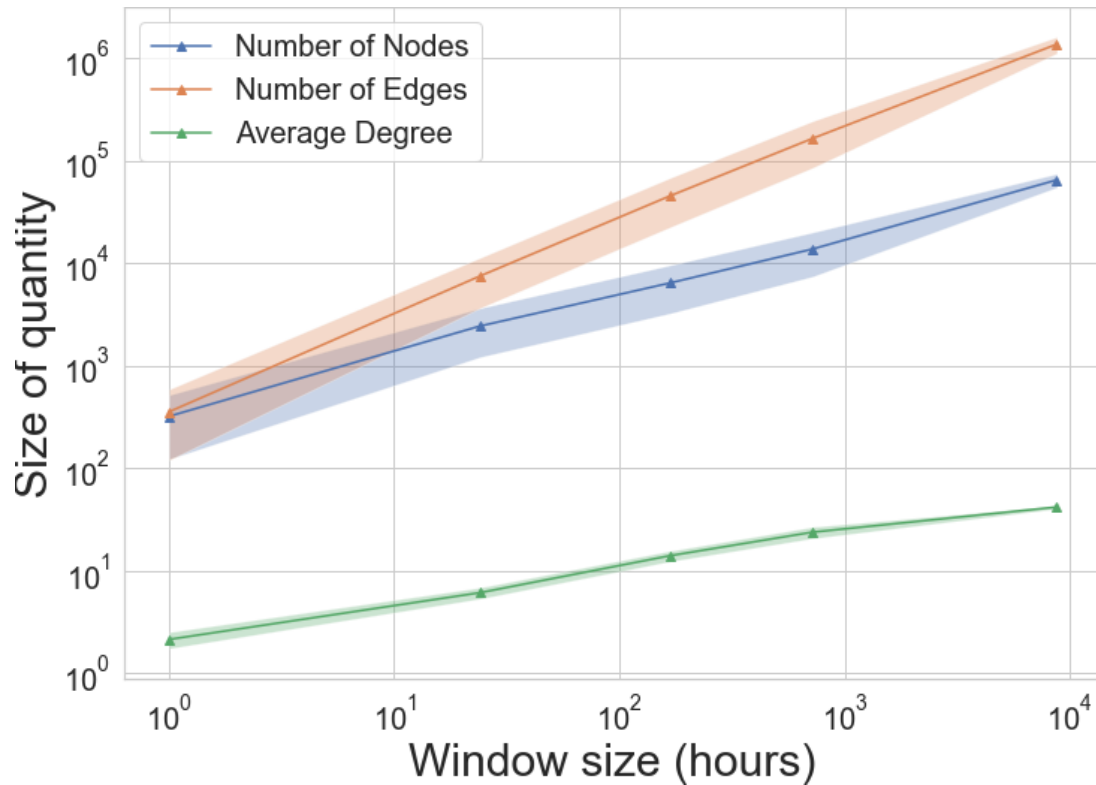

Fig. 3. Average value of the number of nodes, edges and average degree as the window size grows; the shaded region represents one standard deviation above and below the mean.
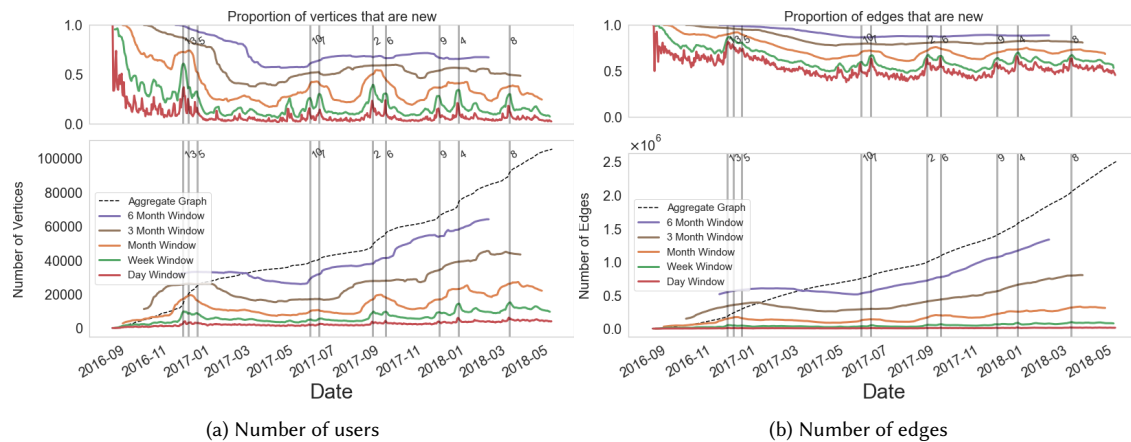
Fig. 4. Number of (a) users and (b) edges between unique user pairs across different window sizes (bottom) and the proportion of these that are new (top). 'New' here means that this is the first window in which they are present. Shown also as a reference is the number of nodes and edges in the aggregate graph, i.e. from aggregating the interactions from time 0 to time t.

the total userbase, which is not surprising given the prevalence of bots and duplicate accounts on social networks [28]. On average 19% of monthly users are active on a given day, compared to 44% for Twitter [31].

Similarly, Figure 4b shows the number of edges and the proportion of new interactions in different windows. Surprisingly, the majority of communications in most windows of a day or greater are between users who have never previously interacted. The ratio holds even in the later months of our dataset where the mass of users was higher. This is in contrast to work [8] which has found that many edges in temporal social networks tend to be persistent, with the probability of an edge returning depending positively on whether it exists already. They observe that 12% of weekly edges are new interactions, a figure which stays low throughout the dataset (after the first week). This is in contrast to an approximate proportion of 60% we see in the Gab data. It does, however, add weight to studies [39] that have characterised interactions on Gab as broadcast behaviour rather than engaging in conversations; where interactions exist not to build up social relationships but to reinforce hierarchies.

Multiple window sizes help in identifying events that are either hidden with larger window scales or lost within the noise of smaller windows. With this particular dataset, the weekly ratio seems to provide the most clear indicator of bursty spikes of new users. We identify 10 peaks by finding the highest local maxima which are at least 7 days apart from each other, excluding the first 100 days of data which contain a high proportion of new users due to their proximity to the site's creation. The times of these peaks in new user proportions coincided more or less with the peaks found in the proportions of edges that are new, so we refer to the same list of peaks for both. We have compared these peaks with key events from the alt-right movement. The ten highest peaks labelled in decreasing order of height have been marked in Figures 4a and 4b; for the top five of these we also identify relevant events which closely precede and may explain them which we discuss in the following. We consolidate these beliefs by measuring the frequency of posts containing a given keyword on the event day, compared with a corpus of 250,000 randomly selected posts. As a benchmark comparison, 'Trump' is mentioned 3.5 times more on the day after the 2016 US election than on a randomly chosen day.

Some growth bursts seem to centre on events around the Trump administration, with the site's largest peak in proportion of new users after the site's birth coming shortly after Trump's election (peak 1) and the 3rd and 5th highest
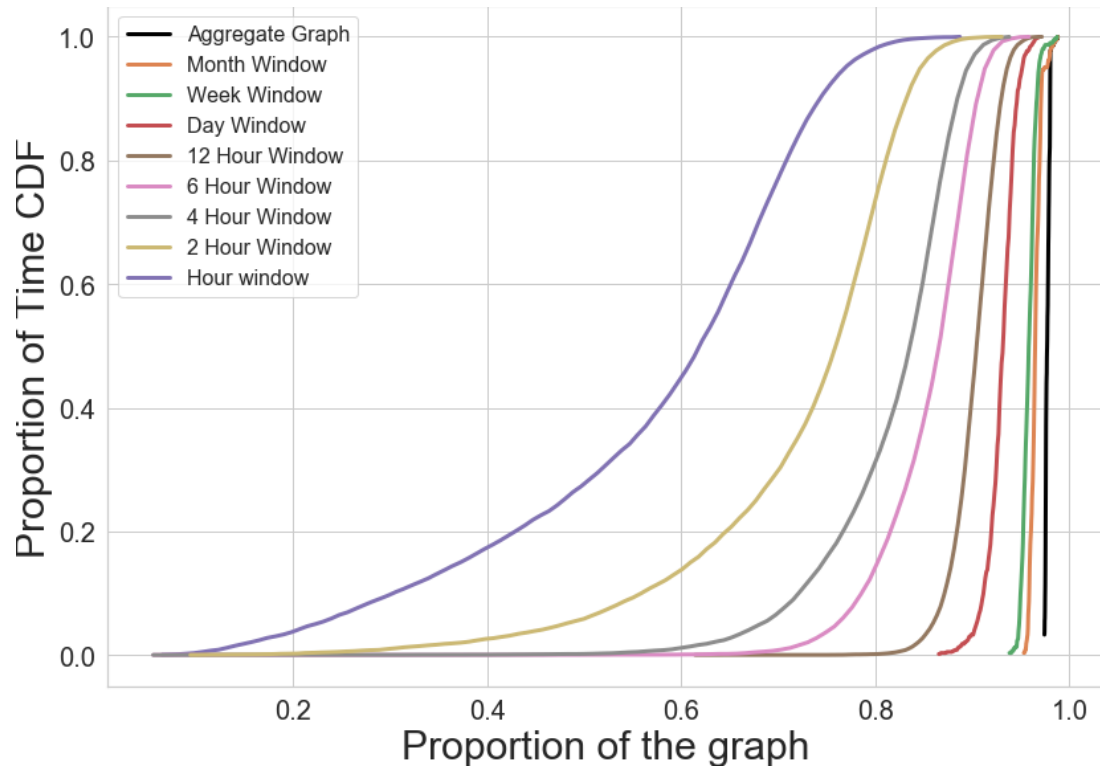
Fig. 5. CDF of the size of the largest connected component (LCC) as a proportion of the total graph for each window size. A point on this graph shows the proportion of the time (y-axis) the proportion of graph within the largest connected component was smaller than the value on the x-axis.

peaks shortly following this. Some peaks in new users seem to be driven by migration from Twitter in response to clampdowns on the alt-right, the site founded shortly after the high profile banning of Milo Yiannopoulos from Twitter in July 2016. One such clampdown is on 18 Dec 2017 (event 4) when Twitter removed numerous accounts associated with the far right, (including for example, Paul Golding and Jayda Fransen of Britain First who joined Gab the same day) prompting a 2.2 times increase in occurrences of 'Twitter' among Gab posts. This peak precedes a sustained (if small) increase in total number of users across all time windows. What seems to be the site's largest and most sustained increase in number of users after the birth of the site is around the Charlottesville Unite the Right rally (peak 2) and its aftermath. The number of active users in this month roughly doubles from a month previous, with over half of these users never seen before on the site at its peak.

5.2   Considering the Gab Userbase as a Connected Community

The Gab userbase is characterised by homogeneity, especially driven by political topics. It is useful thence to investigate the extent to which they form a single community[2]. We might consider the ties among the Gab community by looking again at the data in Figure 4b and in particular the proportion of new edges. The day trace for new edges shows that

---

[2]Here we refer to a community in the informal sense and not to the mathematics of community detection.

40% to 70% of daily interactions are between users who have never interacted before that day. This is consistent with the idea that Gab is not well characterised as a community of friends who consistently interact with each other.

We address this question by observing the size of the largest connected component (LCC). Aggregate graphs of OSN almost always show that the majority of users in a network are part of a large connected component. Considering time windows of data gives a much more detailed insight. Smaller window sizes imply fewer nodes and fewer links (analogous to edge removal for graph robustness [1]).

Figure 5 shows the proportion of users connected by the LCC for different window sizes τ. For example, at a window size of one day (and all longer windows) the LCC connects 90% of users for the vast majority of windows (the LCC proportion is above ninety percent nearly one hundred percent of the time). However, the smaller the window size we observe, the higher variability we find in the size of the LCC. As might be expected the number of users (and hence the absolute size of the LCC) varies considerably with time of day when studied at the hour level. This is consistent with [13] looking at the same effect in a telecommunications network. What is less predictable is what will happen when we look at the *proportion* of users who are within the LCC.

It is not obvious whether an OSN will form a single LCC when viewed at the finest time scale (one hour). In fact, at the hour time-scale, the Gab network moves between two different regimes. This can be seen in Figure 6 where we focus further on the proportion of users in the LCC using the hour long window. When we zoom in on a short period of time the reason for the variability becomes clear. The proportion of users who are part of the LCC varies hugely from the peak to off-peak. Gab usage is highly diurnal, driven by a userbase that is largely US- and Europe-based [39]. At peak hours, 70 – 80% of active users are part of a single LCC. At off-peak hours, the LCC contains no more than 30% of users and Gab becomes several smaller completely disconnected networks that can be thought of as isolated groups (sometimes just user pairs) talking among themselves. We believe that this daily shattering and reforming of the LCC has never previously been observed in OSN data.

We look a little closer at the nature of activity arising from the diurnal habits of the userbase by applying a low-pass filter to the data in Figure 6. The red line represent the data with a low pass filter keeping only the lowest frequency components. We notice twin peaks, a distinct peak with a slightly smaller peak lagging behind. This pattern could arise from two superimposed groups of slightly different sizes that are six hours apart. This is consistent with the work [39] which found a German and French presence alongside English language content on the platform.

The change of proportion and size of the LCC is not simply an effect of the number of active users at a given time as can be seen by comparing against randomised models. Figure 7 shows the proportion of the graph taken up by the LCC (blue), other components with a size strictly greater than two (orange) and pairs of users only interacting with each other (green). This view shows us not only that the proportion of users in the largest connected component enlarges and shrinks again over the course of a 24 hour period but that at off-peak hours this shrinkage comes mainly with an increased presence of isolated interactions in the network as opposed to more intermediate sized components. First, ignoring any temporal aspect, we test whether the behaviour of the connected component is owing to any structural correlations within the network, using a restrained randomisation model (see section 4.2) as a reference (middle panel of Figure 7). The restrained randomisation model produces qualitatively similar results to the real data. To check for statistical significance we compared with a 95% confidence interval (not shown) obtained by rerunning the randomisation 5 times. The proportion of users in the LCC remained significantly lower in the real data, particularly in the troughs where, for example, in the real data the LCC was only 13% of the graph on 24th February in one of the off-peak hours but in the restrained randomisation data for the lower bound of the confidence interval the LCC was 56% of the graph. Similarly the proportion of users making pairwise interactions is slightly higher in the real data,
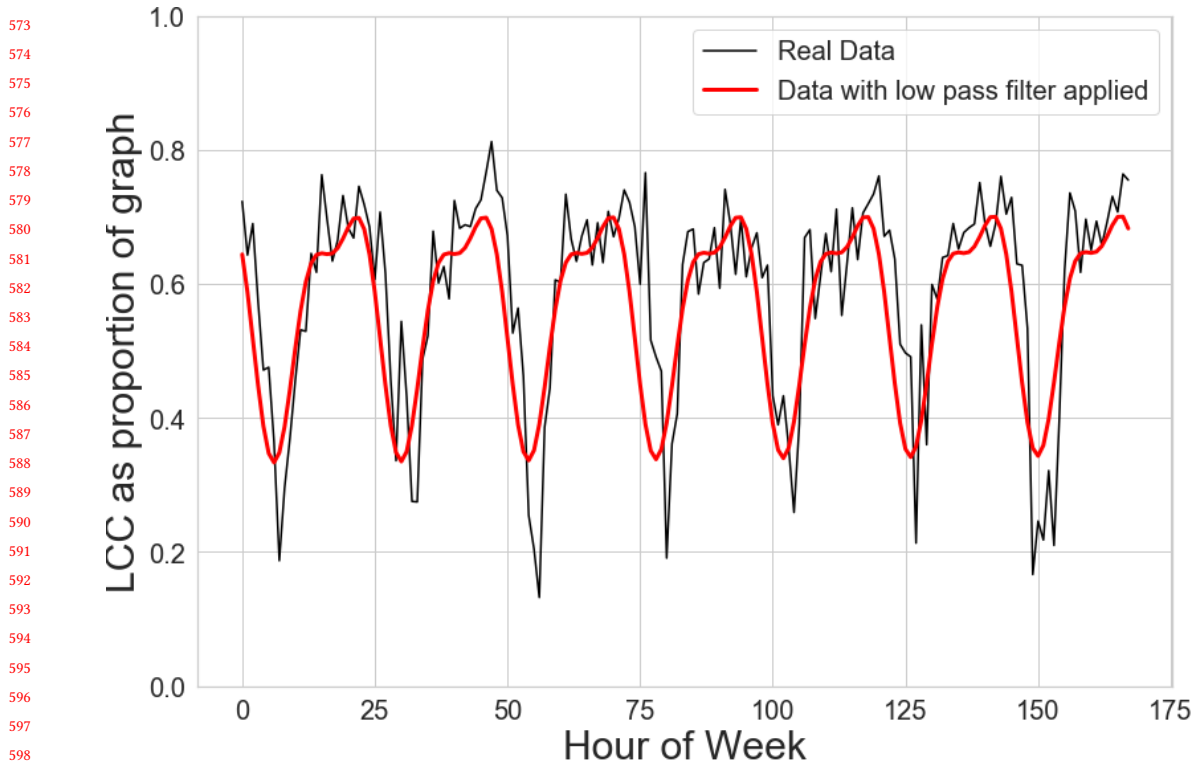
Fig. 6. A week of data showing the proportion of the graph that is within the largest connected component when the graph is viewed in hourly windows. The red line smooths this with a low pass filter. With the hourly window most of the graph is in a single connected component at peak hours but at off peak hours the graph "shatters" into smaller components.
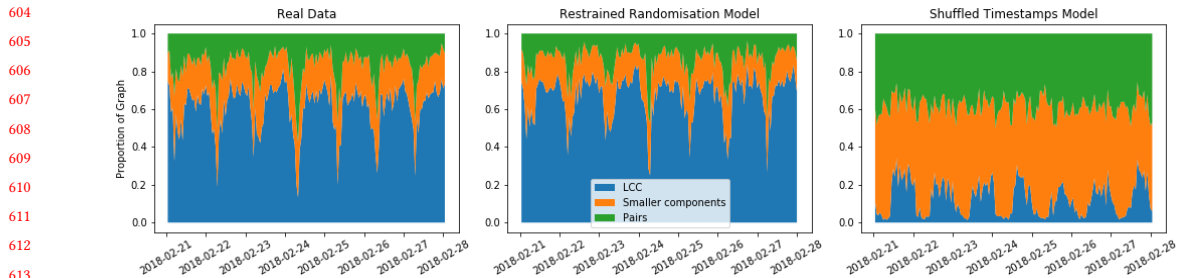
Fig. 7. Comparison of the proportion of the graph that is located within the largest connected component, components of size 2 (pairs) and intermediate sized components when the graph is viewed with an hour sized window. Values are given for the same week-long time period, in the real data (left), the restrained randomisation model (centre), and the shuffled timestamps model (right). Graphs are obtained over the hourly window for a week period.

particularly in the off peak. In this way, the majority of the behaviour is a result of the number of interactions each present user is engaged with but some arises from Gab's network structure which favours pairwise interactions over group discussions.
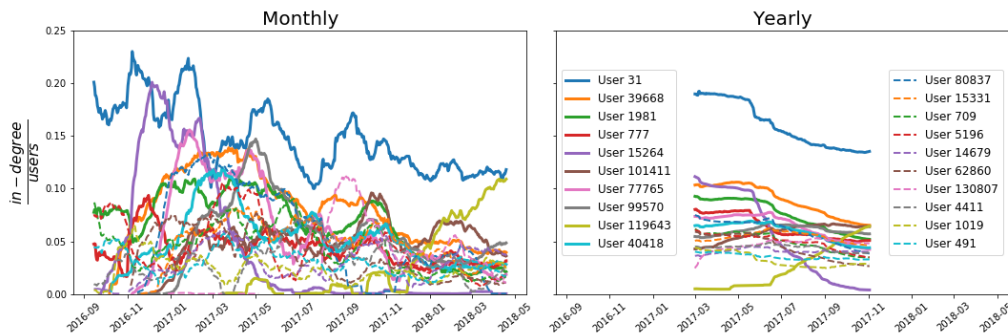
Fig. 8. In-degree for a set of top-twenty users for the month (left) and year (right) window sizes normalised by the number of nodes within that window. The top-twenty are chosen by their in-degree in the aggregate graph.
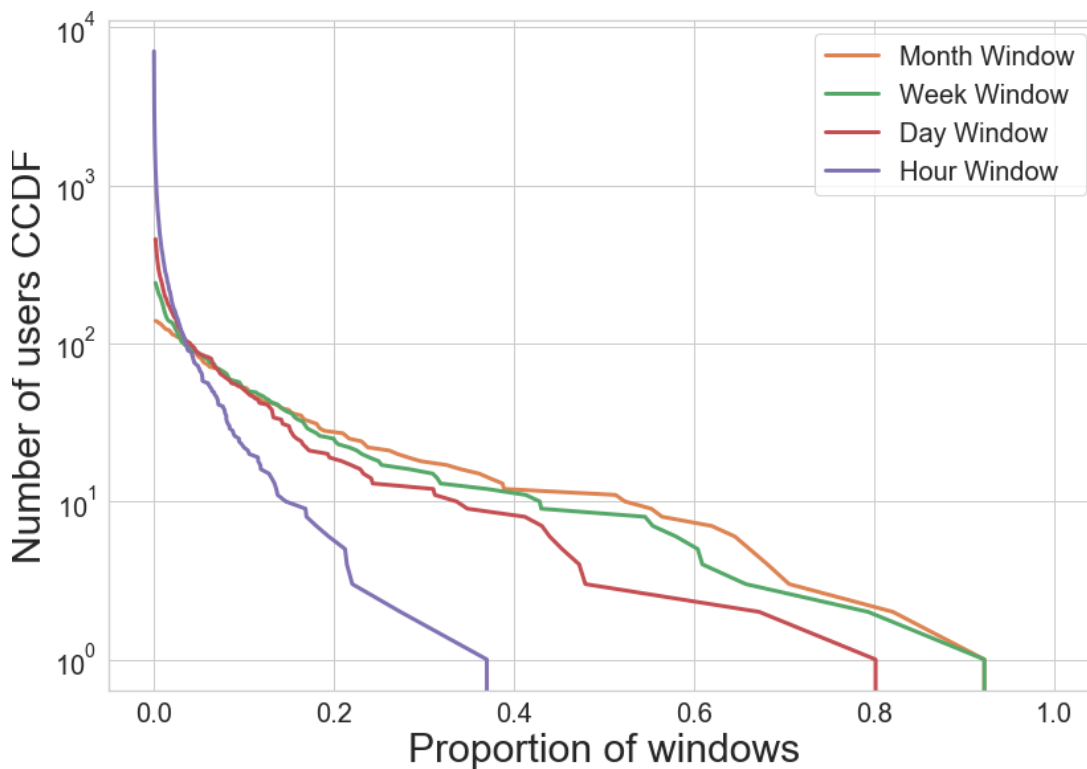

Fig. 9. The plot shows for each time window how many users are in the top twenty for a given proportion of the windows of that size. For example, in the month window 10 users are in the top 20 for 50% of the month long windows.

We investigate the role of temporal correlations by considering a randomised timestamps model (see section 4.2) which takes the original interaction dataset and permutes the timestamps. In this case Figure 7 (right) we observe that a periodic behaviour persists but that the proportion of users in the LCC is very small for all time periods. The periodicity can be explained by the fact that the higher rate of interactions during peak hours means increased densification and

hence a larger LCC. However, the activity rate alone does not explain the proportion of users present in the LCC in the real data during peak hours (the LCC is a much smaller proportion of the graph in the shuffled model), meaning that temporal correlations between interactions are indeed the vital driver in explaining this behaviour.

5.3 Churn of Influential Users on Gab

In this section, we ask how many users *ever* become influential on the platform, and for how long that status lasts. We use the in-degree (number of interactions from distinct users) a user receives as a coarse measure of their influence in that window, and we extract the top 20 users for each window. Figure 9 reports the proportion of windows a user spends in the top 20 at different timescales. The same nine users that are found in the top 20 in more than half of the month windows are found in nearly half of the week windows, and 30% of the day windows. For MathOverflow this may partly be because of the longer time period that is covered in the dataset. When looking through a timescale of a day, we see a large pool of different users that enter the top 20 at this timescale, around 800 for Gab and 2500 for MathOverflow. For the most part, however, this influence is very short lived, with nearly half of these entering the ranks for a single day and never again.

We also examine trajectories of the top ranked users by in-degree. Figure 8 shows the rise (and sometimes fall) of a subset of these users. The users shown are the 20 with the highest in-degree in the aggregate graph. Some of these users hold and maintain a huge level of attention, with the top ranked user [3] in the year window receiving replies from 20% of the active users within that year and the rest of the users presented in the year window each achieving attention from over 4% of the network. The year window shows on the most part a stable hierarchy of top users who only occasionally overtake each other, with a downward trend which is explained by the growing nature of the network. The monthly plot shows in more detail when and in what manner each user rises to prominence. For example, user 31 who tops the year window is dominant for nearly the whole time period for the smaller window as well. On the other hand, some users who obtain prominence in the year window achieve this through a short burst of fame seen only through the lens of smaller timescales (seen with Users 15264 and 130807). Not included are the plots for the day and week windows size which are mostly too noisy to draw out any salient user trajectories. We did note, however, a small number of users who gained a high amount of influence for just a single day each and never again.

5.4 Discussion

In a time where mainstream social media offer only limited access to their platform data, full records of a social media site such as Gab has become a goldmine for researchers to study the full evolution of such platforms, especially in these turbulent political times. Looking at data through multiple time windows provides insight that could not be gained from analysing a graph using just one time scale. Even a basic question like "is the Gab network growing or shrinking?" depends crucially on the time window we look at. Although over longer periods of time the number of users in the network seems to be steadily growing, when considered in small time windows of hours or days, the userbase for Gab is driven by spikes of arrivals, which may be associated with events of importance for potential users of this network. Some of these events coincide with bans or suspensions of certain individuals from other mainstream platforms, potentially challenging the question of whether deplatforming is effective in curbing extremist ideologies. A previous study on a subreddit ban [3] showed that despite seeing a slight decline in their userbase, no increase in hate speech was observed on the communities that took on the migrating users; which might be seen as a success

---

[3]Andrew Torba, founder of Gab

story for deplatforming. This however overlooks the effect of these kind of policies on the social media eco-system. In a way, deplatforming policies are seen by some as passing the problem to someone else and potentially pushing these users to a darker part of the web [3]. As a matter of fact, looking at the growth of Gab here, we see an increase both in activity and the userbase on the website after two such dismantling events. They seem to give a fresh breath to the platform considering the weakening of user activity that characterises Gab. Despite Twitter suspending multiple influential accounts such as that of Milo Yiannopoulos, we see a substantial rise in users joining Gab after two mass ban/suspension events during the period of study. This is akin to the Streisand effect [23], where the result of a particular intervention can be accompanied by results opposite to what was intended. This can be explained by the simultaneous migration of followers from one platform to another, which enables banned influencers to swiftly establish an audience on another platform, mirroring social bootstrapping approaches [38]. This could be an incentive for future policy makers to disperse these suspensions over time to avoid mass migrations.

That being said, despite this offering new life for platforms like Gab to grow, we observe a thinning in audience that has been reported by multiple studies on similar platforms [25, 32]. This could be explained in part by the broadcast behaviour that is prevalent in the Gab userbase, especially amongst "superusers". In a given month, typically more than one quarter of the users interacting with the network have never previously done so. At every time scale we consider, 40% of interactions take place between pairs of users who have never previously interacted. The most interacted with users in Gab form a small core of users who dominate most of the time, receiving extremely high levels of attention, and much wider pool who gain influence fleetingly. This supports the study by Cinellie et al. [4] who reveal that, although Gab users are more likely to receive right-leaning views, they are not more or less likely to spread the information. Typically, like the saying "fifteen minutes of fame", users in this wider pool may enter the top twenty for a single time period and then rarely, if ever, reappear. The network appears to be characterised by churn, both in new users appearing and in different users dominating the interactions which is detrimental to cultivate a cohesive network.

Our study of activity on Gab as a temporal interaction graph was focused on studying the structural aspect of the network and for such, the observations reported here are content-free and do not address other aspects that can divulge interesting information about the evolution of such platforms. For instance, future work would look at the effect of different types of moderation, or lack therefore, on radical platforms in promoting or curbing hate speech. Another point is that, despite the multitude of studies on social media sites, it is becoming more and more apparent that looking at a platform behaviour as a single entity isn't enough. Further study is required to understand how these platforms evolve and the ways that extremist behaviours emerge by addressing these platforms as one complex eco-system.

6 CONCLUSIONS

In future work we plan to extend our approach to other OSNs where the full set of interactions is available (unlike the sampled data available for Twitter). We will also assess the discrimination potential of different timescales depending on the dataset characteristics.

REFERENCES

[1] Réka Albert, Hawoong Jeong, and Albert-László Barabási. 2000. Error and attack tolerance of complex networks. *Nature* 406, 6794 (2000), 378–382.

[2] Arunkumar Bagavathi, Pedram Bashiri, Shannon Reid, Matthew Phillips, and Siddharth Krishnan. 2019. Examining untempered social media: analyzing cascades of polarized conversations. In *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 625–632.

[3] Eshwar Chandrasekharan, Umashanthi Pavalanathan, Anirudh Srinivasan, Adam Glynn, Jacob Eisenstein, and Eric Gilbert. 2017. You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 31 (Dec. 2017),

22 pages.

[4] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021).

[5] Aaron Clauset and Nathan Eagle. 2012. Persistence and periodicity in a dynamic proximity network. *arXiv preprint arXiv:1211.7343* (2012).

[6] Gabriel Fair and Ryan Wesslen. 2019. Shouting into the void: A database of the alternative social media platform Gab. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13(01). 608–610.

[7] Laetitia Gauvin, Mathieu Génois, Márton Karsai, Mikko Kivelä, Taro Takaguchi, Eugenio Valdano, and Christian L Vestergaard. 2018. Randomized reference models for temporal networks. *arXiv preprint arXiv:1806.04032* (2018).

[8] Peter Grindrod and Mark Parsons. 2011. Social networks: Evolving graphs with memory dependent edges. *Physica A: Statistical Mechanics and its Applications* 390, 21-22 (2011), 3970–3981.

[9] Petter Holme and Jari Saramäki. 2012. Temporal networks. *Physics reports* 519, 3 (2012), 97–125.

[10] N. F. Johnson, R. Leahy, N. Johnson Restrepo, N. Velasquez, M. Zheng, and P. Manrique. 2018. Social media cluster dynamics create resilient global hate highways. *arXiv preprint arXiv:1811.03590* (2018).

[11] Ivan Kalmar, Christopher Stevens, and Nicholas Worby. 2018. Twitter, Gab, and racism: The case of the Soros myth. In *Proceedings of the International Conference on Social Media and Society*. 330–334.

[12] Márton Karsai, Kimmo Kaski, and János Kertész. 2012. Correlated dynamics in egocentric communication networks. *PLOS One* 7, 7 (2012), e40612.

[13] Gautier Krings, Márton Karsai, Sebastian Bernhardsson, Vincent D Blondel, and Jari Saramäki. 2012. Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science* 1, 1 (2012), 4.

[14] Yannick Léo, Christophe Crespelle, and Eric Fleury. 2015. Non-Altering Time Scales for Aggregation of Dynamic Networks into Series of Graphs. In *Proceedings of the ACM Conference on Emerging Networking Experiments and Technologies (CoNEXT '15)*. Association for Computing Machinery, New York, NY, USA, Article 29, 7 pages.

[15] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. 2005. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of ACM International Conference on Knowledge Discovery in Data Mining*. 177–187.

[16] Lucas Lima, Julio CS Reis, Philipe Melo, Fabricio Murai, Leandro Araujo, Pantelis Vikatos, and Fabricio Benevenuto. 2018. Inside the right-leaning echo chambers: Characterizing Gab, an unmoderated social system. In *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 515–522.

[17] Sergei Maslov and Kim Sneppen. 2002. Specificity and stability in topology of protein networks. *Science* 296, 5569 (2002), 910–913.

[18] Binny Mathew, Anurag Illendula, Punyajoy Saha, Soumya Sarkar, Pawan Goyal, and Animesh Mukherjee. 2019. Hate begets Hate: A Temporal Study of Hate Speech. *arXiv preprint arXiv:1909.10966* (2019).

[19] Reid McIlroy-Young and Ashton Anderson. 2019. From "Welcome New Gabbers" to the Pittsburgh Synagogue Shooting: The Evolution of Gab. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13(01). 651–654.

[20] Brendan Meeder, Brian Karrer, Amin Sayedi, R Ravi, Christian Borgs, and Jennifer Chayes. 2011. We know who you followed last summer: inferring social link creation times in twitter. In *Proceedings of the International conference on World Wide Web*. 517–526.

[21] Alan Mislove, Massimiliano Marcon, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. 2007. Measurement and analysis of online social networks. In *Proceedings of the ACM Internet Measurement Conference*. 29–42.

[22] Vincenzo Nicosia, John Tang, Cecilia Mascolo, Mirco Musolesi, Giovanni Russo, and Vito Latora. 2013. Graph metrics for temporal networks. In *Temporal Networks*. Springer, 15–40.

[23] Abby Ohlheiser. 2020. Twitter's ban almost doubled attention for Biden story. *MIT Technology Review* (2020). https://www.technologyreview.com/2020/10/16/1010644/twitter-ban-hunter-biden-emails-backfires Accessed: 2021-04-15.

[24] Nicola Perra, Bruno Gonçalves, Romualdo Pastor-Satorras, and Alessandro Vespignani. 2012. Activity driven modeling of time varying networks. *Scientific Reports* 2 (2012), 469.

[25] Richard Rogers. 2020. Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication* 35, 3 (2020), 213–229.

[26] Jari Saramäki and Esteban Moro. 2015. From seconds to months: an overview of multi-scale dynamics of mobile telephone calls. *The European Physical Journal B* 88, 6 (2015), 164.

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

[28] Sérgio SC Silva, Rodrigo MP Silva, Raquel CG Pinto, and Ronaldo M Salles. 2013. Botnets: A survey. *Computer Networks* 57, 2 (2013), 378–403.

[29] Benjamin Steer, Felix Cuadrado, and Richard Clegg. 2020. Raphtory: Streaming analysis of distributed temporal graphs. *Future Generation Computer Systems* 102 (2020), 453–464.

[30] Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttone, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. 2014. Measuring large-scale social networks with high resolution. *PLOS One* 9, 4 (2014).

[31] Twitter. 2019. Q1 2019 Earnings Report. https://s22.q4cdn.com/826641620/files/doc_financials/2019/q1/Q1-2019-Slide-Presentation.pdf.

[32] Aleksandra Urman and Stefan Katz. 2020. What they do in the shadows: examining the far-right networks on Telegram. *Information, Communication & Society* 0, 0 (2020), 1–20.

[33] Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P Gummadi. 2009. On the evolution of user interaction in Facebook. In *Proceedings of the ACM workshop on Online social networks*. 37–42.

[34] Christo Wilson, Alessandra Sala, Krishna PN Puttaswamy, and Ben Y Zhao. 2012. Beyond social graphs: User interactions in online social networks and their implications. *ACM Transactions on the Web* 6, 4 (2012), 1–31.

[35] Dongliang Xie, Xin Wang, Lanchao Liu, and Linhui Ma. 2016. Exploiting time-varying graphs for data forwarding in mobile social Delay-Tolerant Networks. In *Proceedings of IEEE/ACM International Symposium on Quality of Service*. IEEE, 1–10.

[36] Savvas Zannettou, Barry Bradlyn, Emiliano De Cristofaro, Haewoon Kwak, Michael Sirivianos, Gianluca Stringini, and Jeremy Blackburn. 2018. What is Gab? A bastion of free speech or an alt-right echo chamber?. In *Proceedings of the The Web Conference*. 1007–1014.

[37] Qiankun Zhao, Yuan Tian, Qi He, Nuria Oliver, Ruoming Jin, and Wang-Chien Lee. 2010. Communication motifs: a tool to characterize social communications. In *Proceedings of the International conference on Information and Knowledge management*. 1645–1648.

[38] Changtao Zhong, Mostafa Salehi, Sunil Shah, Marius Cobzarenco, Nishanth Sastry, and Meeyoung Cha. 2014. Social Bootstrapping: How Pinterest and Last.Fm Social Communities Benefit by Borrowing Links from Facebook. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14)*. Association for Computing Machinery, New York, NY, USA, 305–314.

[39] Yuchen Zhou, Mark Dredze, David A Broniatowski, and William D Adler. 2019. Elites and foreign actors among the alt-right: The Gab social media platform. *First Monday* 24, 9 (2019).