Alakorkko, T.; Saramaki, J.

Circadian rhythms in temporal-network connectivity

Published in:
CHAOS

DOI:
10.1063/5.0004856

Published: 01/09/2020

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Circadian rhythms in temporal-network connectivity

Cite as: Chaos 30, 093115 (2020); https://doi.org/10.1063/5.0004856
Submitted: 14 February 2020 . Accepted: 11 August 2020 . Published Online: 04 September 2020

T. Alakörrkö and J. Saramäki

COLLECTIONS

Paper published as part of the special topic on Dynamics of Social Systems

ARTICLES YOU MAY BE INTERESTED IN

Power-law distribution in the number of confirmed COVID-19 cases
Chaos: An Interdisciplinary Journal of Nonlinear Science 30, 093123 (2020); https://doi.org/10.1063/5.0013031

Rare events in complex systems: Understanding and prediction
Chaos: An Interdisciplinary Journal of Nonlinear Science 30, 090401 (2020); https://doi.org/10.1063/5.0024145

Neural network representation of the probability density function of diffusion processes
Chaos: An Interdisciplinary Journal of Nonlinear Science 30, 093118 (2020); https://doi.org/10.1063/5.0010482
Circadian rhythms in temporal-network connectivity

Cite as: Chaos 30, 093115 (2020); doi: 10.1063/5.0004856
Submitted: 14 February 2020 · Accepted: 11 August 2020 · Published Online: 4 September 2020

T. Alakörkkö and J. Saramäki

AFFILIATIONS
Department of Computer Science, Aalto University, 02150 Espoo, Finland

Note: This article is part of the Focus Issue, Dynamics of Social Systems.

Affiliated Authors. T. Alakörkkö to whom correspondence should be addressed: tuomas.alakorkko@aalto.fi
J. Saramäki: jari.saramaki@aalto.fi

ABSTRACT
Human activity follows a circadian rhythm. In online activity, this rhythm is visible both at the level of individuals as well as at the population level from Wikipedia edits to mobile telephone calls. However, much less is known about circadian patterns at the level of network structure, that is, beyond the day–night cycle of the frequency of activity. Here, we study how the temporal connectivity of communication networks changes through the day, focusing on sequences of communication events that follow one another within a limited time. Such sequences can be thought to be characteristic of information transfer in the network. We find that temporal connectivity also follows a circadian rhythm, where at night a larger fraction of contacts is associated with such sequences and where contacts appear more independent at daytime. This result points out that temporal networks show richer variation in time than what has been known thus far.

The circadian rhythm is an omnipresent feature in human behavior—our activity levels and sleep patterns are locked to the diurnal cycle. This daily variation is generally visible in human activity, including our online behavior. Time-stamped data on human online activity have revealed circadian patterns at the population level (e.g., in the overall frequency of Wikipedia edits or phone calls) as well as at the level of individuals (see, e.g., Refs. 3 and 4). Here, we ask what happens between these microscopic and macroscopic levels of individuals and populations—do the daily activity patterns of individuals translate into daily rhythms in the mesoscopic connectivity of temporal networks?

I. INTRODUCTION
The dynamics of social systems and population-level behavioral patterns emerge from the behavior of individuals. While the emergence of certain social-network features, such as the abundance of triangles, is well understood, the picture is far less clear when the temporal dimension is included in the picture. Here, we focus on one prominent temporal feature of social activity: circadian patterns in communication.

Circadian patterns in mobile phone use have been shown to be strong enough that data recorded with tracking apps allow one to determine the so-called chronotypes of individual people (their morningness or eveningness). Mobile telephone data also hint that our social behavior might be different at day and at night: often, people’s communication narrows down to selected, specific others at the later hours. However, beyond this picture of individual behavior, not much is known beyond the obvious circadian rhythms in the overall frequency of online activity (see, e.g., Refs. 1, 2, 5, and 6).

In this article, we study what happens between the microscopic level of individuals and the macroscopic level of entire (temporal) networks. We ask how the daily activity patterns of individuals translate into the daily rhythms of temporal networks at the mesoscopic level, between individual behavior and the daily population-level cycle in the frequency of communication events. We focus on the circadian variation of temporally correlated multi-node and multi-link patterns in human communication. More specifically, we look at the daily variation in so-called $\Delta t$-triggered communication events. We label an event $\Delta t$-triggered if its initiator node has participated in other events within a time window of $\Delta t$ time units before the focal event. We also monitor variation in the properties of larger temporal-network structures, $\Delta t$-connected subgraphs.
We show that in mobile telephone communication, the fraction of \( \Delta t \)-triggered events shows a clear circadian pattern, where they are more frequent at night than expected. \( \Delta t \)-connected subgraphs show a similar variation with respect to their expected size. Part of this behavioral shift can be attributed to a specific, more night-active population with idiosyncratic communication habits.

This paper is structured as follows: first, we will introduce the necessary concepts (\( \Delta t \)-triggering, \( \Delta t \)-connectivity, and temporal subgraphs) and the data that we use. Then, we will first show how \( \Delta t \)-triggering varies throughout the day in mobile telephone data and compare this to a randomized reference. We then “zoom out” and look at broader patterns spanned by \( \Delta t \)-connected events in the form of temporal subgraphs and their daily cycles. We then also address the issue of whether this population-averaged behavior is representative, or whether the observations are caused by shifts in the relative activity levels of some specific subpopulations. Finally, we discuss our conclusions.

II. METHODS

A. Data

Call detail records used in this study consists of \( 1.35 \times 10^8 \) calls and \( 465 \times 10^5 \) text messages from a 29 week period. Calls were made by \( 32 \times 10^6 \) individuals and messages were sent by \( 16.6 \times 10^6 \) individuals in a single European country. The aggregated call network consists of \( 100 \times 10^6 \) directed edges, and the message network consists of \( 39.6 \times 10^6 \) directed edges.

B. \( \Delta t \)-triggering in temporal networks

We study temporal networks \( G = (V_G, E_G, T) \), where \( V_G \) is the set of nodes and \( E_G \subset V_G \times V_G \times [0, T] \) is the set of time-stamped communication events between the nodes. The events happen within some limited period of observation \([0, T]\). We denote a communication event initiated by \( i \) and received by \( j \) at time \( t \) with \( e(i,j,t,\tau) \), where \( \tau \geq 0 \) is the duration of the event; the duration is set to 0 for text messages. We require that one node is only allowed to participate in one event at any given point in time. The time difference between two consecutive events \( e_1(i,j,t_1,\tau_1) \) and \( e_2(j,k,t_2,\tau_2) \) with \( t_1 > t_2 + \tau_1 \) is defined as \( \delta(t_1, e_2) = t_1 - (t_2 + \tau_1) \).

Because we want to detect circadian rhythms of temporal patterns beyond the activity levels of individual users, we first focus on the temporal–topological proximity of events—our aim is to detect events that are related through common node(s) and that rapidly follow one another. The motivation here is that such events may have a causal relationship: Alice calls Bob who then calls Carol in response.

To this end, we go through every event (call, text message) of every user. For each event \( e(i,j,t,\tau) \), we check whether the user has been involved in another event in the time window \([t - \Delta t, t]\), where \( \Delta t \) is a parameter. In other words, we check whether there is an event \( e'(k,l,t',\tau') \) so that \( t > t', \delta(t,e') \leq \Delta t \) and \( \{i,j\} \cap \{k,l\} > 0 \). If an event \( e' \) that satisfies these conditions exists, we label the event \( e \) as \( \Delta t \)-triggered. See Fig. 1 for a visual example.

Because it is well known that the time series of events in communication networks contain many return events, visible as bursty trains of repeated communication exchange between two users,\(^{11}\) we also use a “non-backtracking” definition of \( \Delta t \)-triggering where two events may share only one node and return events are not allowed. That is, the conditions for event \( e \) to be labeled as \( \Delta t \)-triggered are that there is at least one event \( e' \) so that \( e'(k,l,t',\tau') \) so that \( t > t' \), \( \delta(t,e') \leq \Delta t \) and \( \{i,j\} \cap \{k,l\} > 0 \). This is to separately count events where it is possible that information is being forwarded from one user to another through an intermediary. We will clearly indicate when this additional constraint is being imposed.

Now, given a temporal network and a choice of the parameter \( \Delta t \), we for each hour of the day compute the ratio \( f_{\Delta t}(t) \) between the number of \( \Delta t \)-triggered events that take place during that hour and the number of all events that take place during that hour.

We can expect that much of the intraday variation of \( f_{\Delta t}(t) \) comes trivially from the circadian pattern of the overall frequency of events: the more events, the more likely it is that some of them will be \( \Delta t \)-triggered. To account for this, we also employ the standard time-shuffled randomized reference model.\(^{2,14}\) In this reference model, the time stamps of all events in the data are randomly shuffled. The outcome is a new event timeline that has the same circadian variation as the original in terms of overall event frequency. The number of events between each pair of nodes is conserved. All other temporal correlations are, however, destroyed. Therefore, measuring \( f_{\Delta t}(t) \) in the reference model yields variation that is purely caused by the daily pattern in event frequency. Consequently, if we normalize the original \( f_{\Delta t}(t) \) by that of the reference model, we arrive at the normalized \( f(t) \) that measures the excess

---

**FIG. 1.** A schematic diagram of the concept of \( \Delta t \)-triggered events. Both panels show the timelines of three nodes, \( i, j, \) and \( k \), with arrow indicating events, pointing from the originator of the communication event to its recipient (e.g., caller to callee). (a) The red arrows indicate two events whose initiator node has participated in other events within a preceding time span of \( \Delta t \); we call these events \( \Delta t \)-triggered. The first triggered event is a return event \( j \) returning a call to \( i \) who earlier called \( j \). The fraction of \( \Delta t \)-triggered events is 2/5. (b) The same, but with an additional restriction that return events are not counted as \( \Delta t \)-triggered (“non-backtracking”), leaving out the \( j - i \) event. Therefore, the fraction of \( \Delta t \)-connected events is smaller, \( f_{\Delta t} = 1/5 \).
C. $\Delta t$-connectivity and temporal subgraphs

In order to investigate larger temporal-network patterns, we extend the notion of $\Delta t$-triggering to $\Delta t$-adjacency and $\Delta t$-connectivity and finally temporal subgraphs. We follow the approach and terminology of Saramäki et al.\textsuperscript{14}

To begin, we define a pair of events $e_i = (i, j, t_i)$ and $e_j = (k, l, t_j)$ to be temporally adjacent, denoted $e_i \rightarrow e_j$, if they share at least one node, $|(i, j) \cap (k, l)| > 0$, and they are consecutive in time but not simultaneous, that is, $\delta t(e_i, e_j) > 0$.

If, similarly to $\Delta t$-triggering, we now impose a constraint on the time difference between the two events, we arrive at the definition of $\Delta t$-adjacency: two events $e_i$ and $e_j$ are $\Delta t$-adjacent, denoted $e_i \xrightarrow{\Delta t} e_j$, if they are temporally adjacent and the time difference $\delta t(e_i, e_j)$ between them is less than $\Delta t$, i.e., $\delta t(e_i, e_j) \leq \Delta t$. This means that every $\Delta t$-triggered event and its triggering event are also $\Delta t$ adjacent.

To move beyond pairs of events, we first define weak temporal connectivity as follows: Two events, $e_i$ and $e_j$, are temporally weakly connected if there is a sequence of temporally adjacent events between them.

Similarly to $\Delta t$-adjacency, we can then impose an additional constraint on the temporal relationship between adjacent event pairs: two events, $e_i$ and $e_j$, are weakly $\Delta t$-connected if there is a sequence of $\Delta t$-adjacent events between them, when taking the directions of the adjacency relationships into account.

Using the above definition, we arrive at our key object of interest, the temporal subgraph. A $\Delta t$-connected temporal subgraph consists of a set of events where all pairs of events are weakly $\Delta t$-connected. The subgraph is called valid if no events are skipped.
Our focus will be on the variation of the average size of temporal subgraphs by the hour of the day. There are several complementary ways of measuring temporal subgraph size. First, as temporal subgraphs are collections of events, their size can be measured as the number of constituent events, \( S_E \). Second, as the events take place along a timeline, one can measure the lifetime or duration \( S_T \) of a temporal subgraph as the length of the time interval from its first to its last event. Finally, as all the subgraph’s events are associated with nodes, one can measure the number of nodes involved in the subgraph \( S_N \). Note that it is possible that \( S_N \ll S_E \) because events can recurrently take place within some small subset of nodes—at one extreme, a temporal subgraph may consist of a large number of events between two nodes. Also note that when the hourly averages of the above-mentioned measures are calculated, the subgraph is taken into account only in the average of the hour when the subgraph was started, i.e., the time of the first event in the subgraph.

III. RESULTS

A. The fraction of \( \Delta t \)-triggered events peaks at night for mobile telephone calls and text messages

We begin by counting, for each hour of the day, the fraction of calls as well as the fraction of \( \Delta t \)-triggered calls, averaged over week-days (\( N = 203 \) in total). In this dataset, the overall daily pattern of call frequency shows a shape with two peaks, one before noon and one in the late afternoon/early evening; call frequency temporarily goes down in the early afternoon [Figs. 2(a) and 2(b)]. Note that this population-level activity rhythm can reflect cultural conventions, and it may differ between countries.

The fraction of \( \Delta t \)-triggered events, \( f_{\Delta t}(t) \), follows the overall daily pattern of frequency for the day but remains higher at night. In contrast, as expected, the same fraction calculated for the time-shuffled null model [Figs. 2(a) and 2(b)] closely mirrors the variation of call frequency. This suggests that at night, there is an excess in the fraction of \( \Delta t \)-triggered calls that is not explained by the variation in call frequency. Indeed, the normalized fraction of \( \Delta t \)-triggered events, \( \tilde{f}_{\Delta t}(t) \), reveals this clearly: \( \Delta t \) triggering is surprisingly frequent at night [Figs. 2(c) and 2(d)]. This means that a smaller share of calls is isolated than at daytime: nighttime...
calls tend to trigger further calls within a short period of time. These triggered calls are not only return calls, as visible in the plots in panels (b) and (d) that are computed using forwarded calls only.

For text messages, we see a similar pattern (Fig. 3). The overall daily variation differs slightly from phone calls in that the early-afternoon drop is less pronounced. There is also a larger difference between $\Delta t$-triggering of forward-only text messages and all text messages [Fig. 3(a) and Fig. 3(b), respectively]. This is natural, given that a typical text message is part of a conversation involving several messages between the two parties. However, when looking at the normalized fraction of $\Delta t$-triggered events, $f_{\Delta t}(t)$, we see that
varies similarly to mobile phone calls: there is a clear peak at night and in the very early morning hours.

B. The size and duration of $\Delta t$-connected temporal subgraphs also increases at night

Next, we expand our analysis to larger temporal–topological patterns in the form of $\Delta t$-connected temporal subgraphs (see Sec. II C). To this end, we detect all $\Delta t$-connected temporal subgraphs, using the weighted event graph approach; each temporal subgraph is a component of the thresholded weighted event graph. For each component, we compute the three different measures of component size: the size in events $S_E$, the size in nodes $S_V$, and the lifetime $S_T$. These are then averaged over each hour of each day of the week and then over all the 29 weeks. We also compute reference values for these measures under the time-shuffled null model, similarly as in the analysis of the fraction of $\Delta t$-connected events.

The results for mobile phone calls are shown in Figs. 4(a), 4(c), and 4(e). For all measures of component size—its duration, its number of calls, and the number of users involved in the component—there is a clearly daily pattern, which is very clearly seen in the curves normalized by the time-shuffled null model numbers. This indicates that the pattern in $\Delta t$ connectivity shown above is not limited to pairs of call events alone; these pairs form larger components. Therefore, not only the density of the network shows a circadian pattern, but there is a topological circadian pattern as well. For text messages, we see a very similar picture [Figs. 4(b), 4(d), and 4(f)].

The timeline of Fig. 4 begins at Thu 00:00, and it is interesting to see the variation in the peak heights of temporal subgraph size between the weekends (peaks 2–4) and weekdays (the rest). This is in line with Krings et al., who observed that the structural features of call networks differ between weekends and weekdays.

C. The behavioral shift reflected in $\Delta t$-connectivity is partially attributable to a specific subpopulation

Next, we investigated the question of whether the increased $\Delta t$-triggering at night reflects a behavioral shift in the whole population of users, or whether it is caused by a night-active subpopulation whose call behavior differs from the rest of the population.

To this end, we computed the hourly fraction of calls for the top 5% of the population in terms of $\Delta t$-triggering and the rest of the population [Fig. 5(a)]. It is seen that the most $\Delta t$-triggered users display slightly higher call activity in the evening and at night, indicating that part of the increase at night might be due to this subpopulation. As a complementary approach, we also computed the fraction of $\Delta t$-triggered events between 0 and 6 AM for each user. The histogram of these values is shown in Fig. 5(b). The histogram shows that while most of the users do not have any $\Delta t$-connected events at night, there is a tail of users for whom this fraction is non-vanishing. This, again, indicates that at least part of the observed nighttime increase in $\Delta t$-triggering might be due to a specific subpopulation.

D. Findings generalize to other temporal networks

To investigate whether the observed effects are related to our particular mobile communication dataset, we applied the $\Delta t$-triggering analysis to four temporal networks dataset that are smaller in size and that have shorter observation windows. These datasets (Messages, Facebook, Forum, and Dating) are all related to communication on online platforms; events represent messages or communication. Note that as time zone information is not available for these datasets, $t = 0$ is an arbitrary hour of the day. For details, see Table I and the references therein.

FIG. 5. (a) The fraction of $\Delta t$-triggered events for each hour of the day for 5% of phone users with the highest number of $\Delta t$ triggering events (orange circles) and the rest of the population (blue triangles). (b) Histogram of the number of phone users with a given fraction of $\Delta t$ connected events. Mean value is depicted with dashed red line.

<table>
<thead>
<tr>
<th>Name</th>
<th>$N$</th>
<th>$E$</th>
<th>$T$</th>
<th>$\delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>22 695</td>
<td>280 717</td>
<td>3 d</td>
<td>1 s</td>
</tr>
<tr>
<td>FB17</td>
<td>31 359</td>
<td>566 305</td>
<td>15 000 h</td>
<td>1 s</td>
</tr>
<tr>
<td>Dating</td>
<td>17 009</td>
<td>185 578</td>
<td>250 d</td>
<td>1 s</td>
</tr>
<tr>
<td>Forum</td>
<td>6 625</td>
<td>1 359 075</td>
<td>2 400 d</td>
<td>1 s</td>
</tr>
</tbody>
</table>

TABLE I. Properties of the temporal network datasets. $N$: number of links, $E$: number of events, $T$: time interval covered by the data, and $\delta t$: time resolution.
FIG. 6. Δ-t-triggering in different temporal-network datasets. Panels (a), (b), (e), (f): The fraction $f_{\Delta t}(t)$ of Δ-t-triggered events and the fraction of hourly events in Messages, Facebook, Dating, and Forum datasets. Blue triangles: original data, orange circles: time-shuffled reference. Panels (c), (d), (g), and (h): The ratio $\tilde{f}_{\Delta t}(t)$ between the original and time-shuffled curves of the panels above.
In all four datasets, we see similar behavior of $\Delta t$-triggering as in the mobile telephone data: the fraction of $\Delta t$-triggered events is in all cases higher in the original data than in the time-shuffled reference (Fig. 6). Similarly to calls and text messages, the ratio between the original and the reference does not remain constant throughout the day, but varies and typically peaks at around the time of the lowest level of activity in the data (which we can assume is at night). Therefore, to summarize, temporal-network connectivity displays a broadly similar circadian pattern in different types of online communication than in mobile telephone calls and text messages.

IV. DISCUSSION

We have analyzed the daily variation of connectivity in temporal networks of human communication, using anonymized mobile telephone data and the $\Delta t$-connectivity framework. We have seen that the temporal structure of the network displays circadian patterns beyond the variation in communication event frequency. At night, the calls or text messages that form the network are less isolated than expected on the basis of their number; calls and text messages appear to trigger further calls and messages more frequently than expected. This also means that if $\Delta t$-triggering is taken as indicative of possible information transmission through series of calls or texts, at night this happens more often than expected (however, the overall frequency of communication events is still very low).

This picture of a population-level behavioral shift in communication is broadly in line with other observations of different social behavior at day and at night.27,28 However, our result may also be partially explained by a subpopulation, whose communication displays more $\Delta t$-triggering and who is slightly more active at night.

In terms of the temporal network spanned by calls or text messages, our findings show that there is indeed a daily cycle, where temporal-network connectivity patterns vary throughout the day. This variation, emerging from the variation in individual user behavior, is reflected in the mesoscale properties (temporal subgraphs) of the network, which, in turn, reflect how information is (or can be) transmitted through the network.

Even though this paper has focused on circadian patterns of human communication via mobile devices, the data analysis pipeline presented here can be thought to be more general. It would be, therefore, interesting to see whether other complex systems that can be modeled as temporal networks show systematic temporal variation in connectivity and properties of temporal subgraphs.

ACKNOWLEDGMENTS

We acknowledge support from the Academy of Finland, Project “Digital Daily Rhythms,” No. 297195.

DATA AVAILABILITY

CDR data used this study are not publicly available due to third party ethical restrictions. Additional datasets used in Sec. III D are cited and available from the authors upon reasonable request and with the permission of authors of the publication in question.

REFERENCES