The structured backbone of temporal social ties

Teruyoshi Kobayashi, 1 Taro Takaguchi, 2 and Alain Barrat∗3, 4
1Department of Economics, Kobe University, Kobe, Japan
2Toda, Saitama 335-0021, Japan
3Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France
4Data Science Laboratory, ISI Foundation, Torino, Italy
∗Corresponding author : Alain.Barrat@cpt.univ-mrs.fr

In many data sets, crucial information on the structure and temporality of a system coexists with noise and non-essential elements. In networked systems for instance, some edges might be non-essential or exist only by chance. Filtering them out and extracting a set of relevant connections, the "network backbone", is a non-trivial task, and methods put forward until now do not address time-resolved networks, whose availability has strongly increased in recent years. We develop here such a method, by defining an adequate temporal network null model, which calculates the random chance of nodes to be connected at any time after controlling for their activity. This allows us to identify, at any level of statistical significance, pairs of nodes that have more interactions than expected given their activities: These form a backbone of significant ties. We apply our method to empirical temporal networks of socio-economic interest and find that (i) at given level of statistical significance, our method identifies more significant ties than previous methods considering temporally aggregated networks, and (ii) when a community structure is present, most significant ties are intra-community edges, suggesting that inter-community edges are random. Most importantly, our filtering method can assign a significance to more complex structures such as triads of simultaneous interactions, while methods based on static representations are by construction unable to do so. Strikingly, we uncover that significant triads are not equivalent to triangles composed by three significant edges. Our results hint at new ways to represent temporal networks for use in data-driven models and in anonymity-preserving ways.

I. INTRODUCTION

The analysis of large-scale empirical data sets, and in particular of complex networked data, is often made difficult by the nature of the data itself: data may be noisy [1–3], and contain both robust, generalizable properties and details specific to the collected data set under investigation, which would change if the data had been collected at a different moment or for a different sample of the same system. For instance, data describing interactions between individuals (face-to-face [4, 5], phone calls [6, 7], or online interactions [8, 9]) might show robust group structures in different days or weeks but with a different set and timing of interactions each day, so that the exact timing of an interaction in a specific day might not be relevant to the understanding of the population’s characteristics. Another issue might arise if the network under scrutiny is dense but with very heterogeneous weights on the edges. The importance of edges might then not be easily reducible only to their own weight, nor to the local properties of the nodes they link, such as their degree (number of neighbours in the network) or their strength (sum of the weights of their edges).

In order to extract the most relevant information from the data, several approaches have been put forward for static networks. For instance, the k-core decomposition focuses on more and more connected parts of a network and has been established as an important tool to analyze and visualize complex networks and to determine influential spreaders in networks [10–12]. Another approach consists in determining a “backbone” of significant edges in the network, and to filter out the remaining non-essential edges. Several methods have been proposed to this purpose in the case of static weighted networks. The simplest way of filtering edges is through thresholding: all the edges with weight below a given threshold value are removed. Such a method however imposes an arbitrary cutoff scale, while many systems of interest display broad distributions of weights and complex patterns at multiple scales. Other methods have thus been put forward to filter out edges simultaneously at different scales, using statistical tests based on null models [13–18]: in all cases, the fundamental idea is to test whether the weight of an edge is distinguishable from the hypothetical one that would be generated at random by a certain null model. Filtering is performed by fixing a desired significance level and selecting only those edges whose weight cannot be explained by the null model at the chosen significance level. These significant edges form a backbone of the network.

Various null models have been proposed in the literature to deal with static weighted networks [13–18]. The recent surge in the availability of temporally-resolved high-resolution data on social and economic networks highlights however the need for methods specifically designed to extract backbones from temporal networks or temporally aggregated networks [19, 20]. Obviously, each method defined on static weighted methods can be applied to a temporally aggregated network: for instance, a simple threshold could be applied on the number of contacts between two nodes. However, a highly active node could in principle have a large number of (non-essential) ties, so that one needs to control for the difference in
intrinsic activity levels across nodes to extract statistically significant ties that cannot be explained by random chance.

Here, we develop a method to extract an irreducible backbone from a sequence of temporal contacts between nodes, by defining an adequate temporal null model. This null model can be interpreted as a (temporal) configuration or fitness model, whose parameters are estimated by using global information, namely the numbers of contacts for all node pairs, similarly to the enhanced configuration model (ECM) filter defined for static networks [17]. Thanks to this null model, we determine the set of significant ties, at any significance level, among all the pairs of nodes with at least one interaction in the data. These ties form an irreducible backbone in the sense defined in [17], as their significance cannot be reduced to the activity of the involved nodes. Most importantly, the temporal nature of the null model allows us to attribute a significance also to higher order structures such as e.g., simultaneously occurring triplets of interactions or other temporal motifs [21], a task that would be impossible when defining significant ties and backbones directly from a temporally aggregated network.

We illustrate the application of our filtering method on temporal networks of social and economic relevance and compare its results with several static filtering methods. Interestingly, at a given level of significance, our method identifies more significant edges than other filters and is less biased towards edges with large weights. Moreover, in cases where the aggregated network has a clearcut community structure, corresponding e.g. to classes in a school, the significant ties turn out to be mostly intra-community ones: at high significance levels, the network of significant ties (i.e., backbone) breaks into several connected components, each corresponding to one community, and inter-community edges turn out to be non-significant. This suggests that inter-community edges, while playing a crucial role in reducing the diameter of the network [22, 23], are here indistinguishable from randomly created edges between groups, once nodes’ activities are fixed. We also investigate significant triads, defined as sets of three nodes that interact simultaneously with each other more than expected given their activities, and show that these significant triads are not necessarily composed of three significant edges. This shows the crucial importance of taking into account temporality when defining a null model to detect the significance of structures in temporal networks, as such information could not be obtained from a purely static null model.

II. DATA

We consider five data sets of social and economic interest described by temporal networks (Table I). Four data sets correspond to face-to-face contacts among individuals in different contexts, recorded using wearable sensors by the SocioPatterns collaboration with a temporal resolution of 20 seconds and publicly available [24]. We consider data sets collected in contexts with very different activity levels, constraints on the schedule of individuals, duration and group structures, namely: i) a high school ("Highschool") [24], ii) a primary school ("Primaryschool") [5], iii) an office building ("Workplace") [26] and iv) a hospital ward ("Hospital") [27]. In addition, we also use the data on temporal financial networks in which nodes and edges represent banks and overnight lending-borrowing relationships, respectively. Since the overnight loan contracts last only for one day, we can construct a sequence of daily snapshot networks (i.e., time resolution is one day). We consider here the data on the online interbank market in Italy, called e-MID, between June 12, 2007 and July 9, 2007 (i.e., 20 business days). The data is commercially available from e-MID SIM S.p.A. based in Milan, Italy [28].

III. RESULTS

A. Temporal fitness model

We consider a set of \( N \) nodes and a sequence of interactions that occur at arbitrary points in time between these nodes [20, 29]. We fix a temporal resolution \( \Delta \) by dividing the whole data temporal window of length \( T \) into \( T/\Delta \) time intervals, and we build on each interval a binary adjacency matrix \( A_t \) with elements \( A_{i,j,t} \) equal to 1 if there is at least one interaction between \( i \) and \( j \) during \( (t-\Delta, t] \) and zero otherwise.

In the temporal fitness model, each node \( i \) is assigned an activity level \( a_i \in (0,1] \) and the probability \( u \) that nodes \( i \) and \( j \) interact (e.g., through a face-to-face contact, a bilateral financial transaction, etc) during any given time interval is simply given by the product of their activity levels [30]:

\[
u(a_i, a_j) = a_i a_j,
\]

In a static network context, this class of network model is called the fitness model and has been used to model
network generative processes [31][33]. The null model we obtain is thus a sequence of successive independent realizations of the (static) fitness model. It settles a baseline of how much two nodes are expected to interact, given their activities, if interaction partners are selected at random at each time step. In the simplest version of this null model, we consider constant activity values for each node, which will allow us to explicitly compute the probability distribution of the total number of interactions between each pair of nodes. Note that this number is at most \( \tau = T/\Delta \), i.e., the number of time intervals given the resolution \( \Delta \).

Note also that this temporal null model does not contain any a priori knowledge of the group structure of the nodes. An interesting generalization could be to superimpose group labels or node properties (e.g., gender or age for nodes representing individuals) and interaction probabilities depending on the nodes’ properties.

In addition, we present in section SI of Supporting Information (SI) a refined version that takes into account temporal variations of the overall interaction activity in the system, through the introduction of a time-varying parameter \( \xi(t) \). In that case, the null model is defined by the fact that the probability of nodes \( i \) and \( j \) establishing a connection at time \( t \) is \( w(a_i, a_j, t) = a_i a_j \xi(t) \). We present here the case of constant \( \xi(t) = 1 \) as we can then obtain an analytical form for the probability distribution of the number of interactions of a pair of nodes, while only an approximate formula is available in the more general case. We show however in SI that both methods yield almost identical results in the case studied here.

B. Significant ties

To uncover significant ties with respect to the null model described above, we proceed in two steps (Fig. 1). First, we perform a maximum likelihood estimation of all the node activities \( \mathbf{a} \equiv (a_1, \ldots, a_N) \), as described in the Materials and Methods section, obtaining the estimates \( \mathbf{a}^* \equiv (a_1^*, \ldots, a_N^*) \). We then compute for each pair of nodes \( i \) and \( j \) the probability distribution of their total number of interactions \( m_{ij} \) in the null model, which is given by the following binomial distribution:

\[
g(m_{ij}|a_i^*, a_j^*) = \binom{\tau}{m_{ij}} u(a_i^*, a_j^*)^{m_{ij}} (1 - u(a_i^*, a_j^*))^{\tau - m_{ij}}. \tag{2}
\]

Let \( m_{ij}^c \) denote the \( c \)-th percentile (0 \( \leq c \leq 100 \)) of \( g(m_{ij}|a_i^*, a_j^*) \), i.e., \( c/100 = G(m_{ij}^c|a_i^*, a_j^*) \), where \( G \) is the cumulative distribution function (CDF) of \( g(m_{ij}|a_i^*, a_j^*) \). If the actual empirical number of interactions \( m_{ij}^a \) between \( i \) and \( j \) is larger than \( m_{ij}^c \), it means that this number cannot be explained by the null model at significance level \( \alpha \equiv 1 - c/100 \), indicating that \( i \) and \( j \) are connected by a significant tie. For a given significance level \( \alpha \), we can test the significance of the set of interactions of each pair of nodes independently from the others [17]. Note that, even if the significance of a tie is determined from an aggregated number of interactions, a significant tie does not correspond here to a static edge but to an interacting pair of nodes with their set of temporally resolved interactions, and the backbone given by all interactions in the significant ties remains a temporal network.

Thanks to the use of the null model, a pair of interacting nodes can be significant even if their number of interactions is small, as long as their individual activity levels are sufficiently low. Reciprocally, ties with a large number of interactions might not be significant if the two involved nodes are very active. The ST filter controls indeed for the difference between nodes in terms of intrinsic activity levels. As a consequence, the significant ties identified by our method are “irreducible” in the sense that their significance cannot be attributed to local node-specific properties [17], such as the node degree and strength in the aggregated network: the probability of interaction between two nodes under the null hypothesis is determined by an interplay of global and local information through the maximum likelihood estimation (S4) in Materials and Methods). The resulting network of interactions between the significant pairs of nodes may thus be regarded as an irreducible backbone of the temporal network under study [17].
C. Significant temporal structures

Using a temporal fitness model as null model allows us to go beyond the usual tests concerning the significance of ties, and to assign a significance to higher order structures such as temporal motifs. To illustrate this point, let us consider the simple case of a triadic interaction between three nodes $i, j, k$; the empirical number of time intervals in which the three pairs $(i,j)$, $(j,k)$ and $(i,k)$ are simultaneously interacting, denoted by $r^o_{ijk}$, can be compared to the probability distribution of the number $r_{ijk}$ of occurrences of such triangles in the null model. For each time interval, the probability that $i, j, k$ are forming a triangle of interactions in the temporal fitness model is

$$v(i, j, k) = u(a_i, a_j) \cdot u(a_j, a_k) \cdot u(a_k, a_i),$$

so that $r_{ijk}$ obeys the following probability distribution in the null model:

$$h(r_{ijk} | a_i^*, a_j^*, a_k^*) = \left( \frac{\tau}{r_{ijk}} \right) v(i, j, k)^{r_{ijk}} (1 - v(i, j, k))^{\tau - r_{ijk}}. \tag{4}$$

Similarly to the case of dyads, we define for each significance level $\alpha = 1 - c/100$ the significant triads as those such that $r^o_{ijk}$ is larger than the $c$-th percentile of

$$h(r_{ijk} | a_i^*, a_j^*, a_k^*).$$

Note that this method can easily be generalized to any set of temporally constrained interactions (e.g., occurring in a sequence of successive snapshots) or motifs. On the contrary, any filtering method based directly on the aggregated network, and not taking into account the temporality of the data, is by construction unable to define a null hypothesis for simultaneous interactions (or interactions with a given temporal sequence) and thus to assign a significance to such patterns.

D. Case studies

We now use the ST filter on the data sets presented above, varying the significance level $\alpha$ and comparing the results, when possible, with two other filtering methods that use directly the static, temporally aggregated network, namely the disparity filter (DP filter) \cite{jaccard} and the enhanced configuration model (ECM filter) \cite{ECM}, whose computation is recalled in section \textbf{S2} of the SI. We present in the main text the main results corresponding to the Primarschool and Highschool data sets, with temporal resolution $\Delta = 15\text{min}$ (except when mentioned otherwise) and refer to the SI for the results concerning other temporal resolutions and other data sets.

1. Number of significant ties

Figure 2 first displays the number of significant ties as a function of the significance level $\alpha$, for the three methods. As $\alpha$ decreases, this number decreases sharply for all methods. Interestingly however, the number of significant ties remains much larger for our method than for the DP and ECM filters, as soon as $\alpha$ enters a regime of high statistical significance (e.g., $\alpha < 10^{-2}$). As $\alpha$ becomes very small, i.e., at very high statistical significance, DP and ECM filters retain only a very small number of edges, while the ST filter still uncovers a relatively large number of significant node pairs. The resulting backbone of interactions between significant pairs at very low $\alpha$ (e.g., $\alpha$ between $10^{-10}$ and $10^{-5}$) might be regarded as a fundamental backbone of the data set. See Fig. \textbf{S4} for the similar results obtained with other data sets and different parameters.

2. Comparison with other filtering methods

Given the differences in the definition of the various filters, it is important to understand to what extent these filters select distinct or similar sets of ties. We quantify the similarity between the results of the filtering methods in two different ways. First, the Jaccard index $J(J_{\mu}, J_{\mu'}) = \frac{|J_{\mu} \cap J_{\mu'}|}{|J_{\mu} \cup J_{\mu'}|}$ gives the fraction of common edges between the backbone obtained by the ST filter at significance level $\alpha$ and another backbone $\mathbf{x} \in \{\text{DP, ECM}\}$ at significance level $\alpha'$. A Jaccard equal to 1 means that both methods yield the same exact set of edges, while $J = 0$ means that the backbones are disjoints. As the different methods yield very different backbone sizes for a fixed significance level, we show in Fig. \textbf{S5} a color plot of the Jaccard index as a function of the number of node pairs retained by each filtering method. In both DP and ECM cases, the largest Jaccard indices are obtained when the number of edges are similar: they reach at most of $\sim 80\%$ and decrease as the backbone size decreases. The ST filter thus defines a backbone significantly different from the one obtained from either the DP or the ECM filters, even at fixed number of edges.

The Jaccard index however does not take into account the fact that different ties can correspond to very different number of interactions. We therefore also consider a
A visualization of the backbones obtained by the different weighted measure of the similarity between the backbones various backbones seem thus to all retain similar sets of the network into connected components corresponding to the backbone obtained by the ST filter seems to separate a filtering methods, shown in Fig. 4 confirmed by the values of the weighted modularity of the are more mixed in the other data sets [27]. This is con-

mstruct [5, 25, 26], while nodes from different groups place cases, these groups define a clear-cut community hospital. In the Highschool, Primaryschool and Work-

3. Backbones and community structure

In the data sets we consider, nodes can be classified into different groups, corresponding e.g. to classes in schools, in the workplace and roles in the hospital. In the Highschool, Primaryschool and Workplace cases, these groups define a clear-cut community structure [25, 26], while nodes from different groups are more mixed in the other data sets [27]. This is con-

rning of the weighted modularity of the partition corresponding to these groups shown in Fig. [4]. A visualization of the backbones obtained by the different filtering methods, shown in Fig. [4] and b, indicates that the backbone obtained by the ST filter seems to separate the network into connected components corresponding to these communities more efficiently than the other filters, at fixed number of edges. We show that this is indeed the case, through two more quantitative indicators. First, we measure, as a function of the backbone size, the fraction of intra-group edges (Fig. S7). It is larger than the random baseline (in which edges are kept completely at random) for all filters, approaching one as the number of edges decreases, and maximal for the ST filter. Second, we consider each filtering method as a prediction task for finding intra-group edges: we use \( \alpha \) as parameter and measure, for each \( \alpha \), the true and false positives (edges in the backbone that are/are not intra-group edges) and true and false negatives (edges not in the backbone that are not/are intra-group edges), building thus a ROC curve (see section S3 of the SI for details). The area under the curve (AUC) of the ROC curve of the ST filter is higher than for the other filters for the three data sets for which the group structure corresponds to a community structure (large value of the modularity) while, for the other data sets in which the modularity is small, all filters lead to an AUC close to 0.5, i.e., close to a random baseline (Fig. [4]).

The fact that the ST backbone selects mostly intra-

community node pairs suggests that inter-community edges are largely due to random interactions, given the activity levels of the nodes. Inter-community edges, which act as bridges, play an important role in propagating information, spreading ideas, and diffusion of influence [22, 56]. Nevertheless, our analysis shows that the actual inter-community edges are statistically indistinguishable from randomly connected ties, hence not "significant" with respect to a null model of random interactions between nodes of fixed activities. This hints at a way to represent the original temporal network as a superposition of (i) a backbone of significant ties and (ii) connections extracted at random between nodes of different groups, in a way that would refine the contact matrix of distributions put forward in [37, 38].

4. Triadic relationships

As described above, we can extract triads with a significant number of simultaneous interactions only by using the filtering method presented here, as the other filters do not use temporally resolved null models. We show in Fig. [4] the number of such significant triadic relationships as a function of the significance level \( \alpha \), for the Highschool data set and temporal resolution \( \Delta = 1 \) min. Similar results are shown in Fig. S8 for the other data sets and different values of \( \Delta \). For filtering levels \( \alpha > 10^{-4} \), almost all the triangles present in the aggregated network are considered as significant. However, the number of significant triadic relationships strongly decreases as \( \alpha \) becomes lower, determining a set of triads such that their number of simultaneous interactions cannot be explained by the temporal null model and the individual nodes’ activity levels.

Figure [5] highlights moreover a striking feature of the
FIG. 4. Backbones and community structure. Visualization of the backbones obtained by different filtering methods, at similar numbers of edges, for the (a) Highschool and (b) Primaryschool data sets. The nodes are shown in the same position for all the backbones of a given data set. “High contact” correspond to a simple thresholding procedure on the aggregated network. Different colors denote different classes. For the Primaryschool data, black circles represent the teachers. The ST filter detects a larger fraction of intra-class edges than the other filtering methods. (c) AUC of the ROC curve for the identification of intra-community edges. The red solid curve denotes the value of the weighted modularity $Q$, calculated by regarding the actual groups (classes for Primaryschool and Highschool, departments for Workplace, roles for Hospital) as communities on the original aggregate networks. For the Primaryschool data set, the teachers are classified into classes with which they have the largest number of edges. For the Interbank data set, banks are classified into Italian banks and foreign banks.

FIG. 5. Significant triadic relationships in the Highschool data set, for a temporal resolution $\Delta = 1$ min. (a) Number of significant triads as a function of $\alpha$. (b) Fraction of significant triadic relationships with a given number of significant dyads, as a function of $\alpha$.

significant triads, namely, that they do not necessarily correspond to three significant ties. In fact, the number of significant ties in a significant triad can take any value between 0 and 3 (see also Fig. S9). Reciprocally, not all triangles made by three significant ties turn out to be significant triads (Fig. S10). This clearly shows how the temporal null model allows us to go beyond the definition of significant ties and find significant higher order structures that could not be unveiled by a static approach. Indeed, considering triangles made by significant ties does not guarantee that the corresponding triads have a significant number of simultaneous interactions, while on the other hand ties $(i,j)$ with non-significant number of interactions when considered as dyads can turn out to interact a significant fraction of times simultaneously with two other dyads $(j,k)$ and $(i,k)$ for a certain $k$, forming thus a significant triad.

IV. DISCUSSION

In this paper, we have presented a new method to find significant ties and structures in temporal network data sets. To this aim, we have defined a null model of interactions that takes into account the heterogeneity in the activity of individual nodes and the temporal dimension of the system, and can potentially be extended to include temporal variations in the overall network activity, due for instance to circadian or weekly rhythms, imposed schedule constraints, etc. We compute for each pair of nodes the distribution of their number of interactions in the null model, and compare the empirical value to this distribution. For any chosen significance level, we thus define as significant pairs of nodes those with a number of interactions that cannot be explained by the null model. As the null model includes the heterogeneous activity of nodes, the temporal network backbone composed by the ties with significant numbers of interactions is not reducible to the nodes local properties and contains ties with a broad distribution of numbers of interactions. Varying the significance level allows us to tune the number of node pairs in this backbone.

The comparison with other backboning methods built for weighted static network, and hence applied here on the temporally aggregated network, reveals interesting similarities and differences. Our method yields, at a
given significance level, more pairs of nodes than the other benchmarks. A more detailed comparison shows that the difference does not come from ties with large numbers of interactions, which are similarly selected by all methods, but rather by the fact that our filter uncovers more significant pairs of nodes with small number of interactions. The resulting distribution of weights of the significant ties is less biased towards large weights. Moreover, it turns out that, for networks with a strong community structure, our method tends to uncover mostly intra-community ties, showing the random nature of the inter-community ones, once the node activities are given.

Thanks to the temporal nature of the null model considered, our method can also attribute a significance to more complex temporal structures, such as sets of simultaneous interactions, which have a clear importance in social terms (it is clearly not the same to have three interactions \((i, j)\) \((j, k)\), \((i, k)\) between three individuals at the same moment or at different times), but also for processes such as epidemic spread occurring on top of a temporal network \cite{39}. We have in particular shown that significant triads of simultaneous interactions are not equivalent to triangles of three significant dyads, illustrating the need to take into account the temporality of these structures, which could not be uncovered by static filtering methods.

Our work hints at several perspectives and future research directions. First, it would be interesting to refine data representations such as the ones put forward in \cite{37, 38}, by combining a backbone at a certain significance level and a contact matrix representing in a summarized fashion the non-significant ties. Another possibility would be to represent the data as a backbone plus the set of activities \(\{a^*_i, \xi(l)\}\). In both cases, these representations should be validated by numerical simulations of various types of processes on top of the data. The relevance of such representation is two-fold: on the one hand, they allow to summarize and generalize complex data sets in a way that can be fed into data-driven models of dynamical processes such as epidemic or information spreading; on the other hand, they keep the minimum amount of detailed information on the precise interactions, summarizing less relevant details as distributions or averages, and thus in a way that might be easier to render data anonymous. Another direction of research would be to define a backbone at a finer resolution, namely that would be composed of (sets of) significant interactions instead of ties or set of ties.

\footnote{1 We solve the equation by using Matlab function \texttt{fsolve} which is based on a modified Newton method, called the trust-region-

\section{METHODS}

\textit{Estimation of nodal activity}

We perform a maximum likelihood (ML) estimation of \(a \equiv (a_1, \ldots, a_N)\), taking the \(\tau\) temporal snapshots as input, where \(\tau = |T/\Delta|\). If two individuals are independently matched in each time interval according to probability \(u(a, a')\), then the number of times temporal edges are formed between nodes \(i\) and \(j\) over \(\tau\) time intervals is a random variable \(m_{ij}\) that follows a binomial distribution with parameters \(\tau\) and \(u(a_i, a_j)\). Therefore, the joint probability function leads to

\[
p\{m_{ij}\}|a) = \prod_{i,j:i\neq j} \left( \frac{\tau}{m_{ij}} \right) u(a_i, a_j)^{m_{ij}} (1 - u(a_i, a_j))^{\tau - m_{ij}},
\]

where \(m_{ij} \leq \tau\) denotes the count of temporal edges between \(i\) and \(j\) observed over \(\tau\) periods in the null model. The log-likelihood function for the empirical data \(\{m_{ij}^o\}\) is thus given by

\[
\mathcal{L}(a) = \log p(\{m_{ij}^o\}|a) = \sum_{i,j:i \neq j} \left[ m_{ij}^o \log (u(a_i, a_j)) + (\tau - m_{ij}^o) \log (1 - u(a_i, a_j)) \right] + \text{const.,}
\]

where “const.” denotes the terms that are independent of \(a\). The ML estimate of \(a\) is the solution for the following \(N\) equations:

\[
H_i(a^*) \equiv \sum_{j:j \neq i} \frac{m_{ij}^o - \tau a_i^* a_j^*}{1 - a_i^* a_j^*} = 0, \forall i = 1, \ldots, N, \quad (7)
\]

The first-order condition \cite{54} is obtained by differentiating the log-likelihood function \cite{6} with respect to \(a_i\). The system of \(N\) nonlinear equations, \(H(a) = 0\), can be solved by using a standard numerical algorithm.\footnote{The obtained ML estimates of \(a\) is denoted by \(a^* \equiv (a_1^*, \ldots, a_N^*)\). The numbers of contacts obtained from the model and the empirical data are compared in section \cite{54}.}

The extension of the method to include time-varying probabilities of creating interactions is shown in section \cite{51}.

\section*{ACKNOWLEDGMENTS}

TK acknowledges financial support from the Japan Society for the Promotion of Science Grants no. 15H05729 and 16K03551.
The initial values of $a_i$ are given by the configuration model: $a_i = \sum_{j \neq i} (m_{ij}/\tau)/\sqrt{2 \sum_{i<j} m_{ij}/\tau}$, where the numerator and the denominator represent the means of $i$’s temporal degree and the doubled number of total temporal edges, respectively.
S1. TIME-VARYING MATCHING PROBABILITY

In constructing the temporal fitness model, it is possible to take into account the possibility that the probability of an interaction between two nodes can vary over time even when individuals’ intrinsic activities \( \{ a_i \} \) are constant. This can happen when, for example, a school schedule has a certain rhythm (e.g., lunch time, class schedule, etc), or due to circadian or weekly rhythms. The probability \( u \) for the existence of an interaction between two nodes \( i \) and \( j \) at time \( t \) is then given as

\[
u(a_i, a_j, t) \equiv a_i a_j \xi(t), \quad t = 1, \ldots, \tau,
\]

where \( \xi(t) \) denotes a time-varying parameter. We assume that there is no correlation between the values of \( \xi \) at different times, and the interaction probabilities are independent across time intervals.

The joint probability function for a certain temporal network \( \{ A_t \} \) is obtained as

\[
p(\{ A_t \}| \{ a, \xi \}) = \prod_{t=1}^{\tau} \prod_{i,j \neq j} u(a_i, a_j, t)^{A_{ij,t}} (1 - u(a_i, a_j, t))^{1 - A_{ij,t}},
\]

where \( A_{ij,t} \) is the \((i,j)\) element of the adjacency matrix in time interval \( t \), denoted by \( A_t \), and \( \xi \equiv (\xi(1), \ldots, \xi(\tau))^\top \). The log-likelihood function is thus given by

\[
\mathcal{L}(a, \xi) = \log p(\{ A_t \}| \{ a, \xi \}) = \sum_{t=1}^{\tau} \sum_{i,j \neq j} [A_{ij,t} \log (a_i a_j \xi(t)) + (1 - A_{ij,t}) \log (1 - a_i a_j \xi(t))],
\]

The maximum-likelihood estimate of \((a, \xi)\) is the solution for the following \( N + \tau - 1 \) equations:

\[
\begin{align*}
H^\text{act}_{i}(a^*, \xi^*) &\equiv \sum_{t=1}^{\tau} \sum_{j \neq i} A_{ij,t} - a^*_i a^*_j \xi^*(t) \left( 1 - a^*_i a^*_j \xi^*(t) \right) = 0, \quad i = 1, \ldots, N, \\
H^\text{time}_{t}(a^*, \xi^*) &\equiv \sum_{i \neq j \neq i} A_{ij,t} - a^*_i a^*_j \xi^*(t) \left( 1 - a^*_i a^*_j \xi^*(t) \right) = 0, \quad t = 2, \ldots, \tau,
\end{align*}
\]

The first-order conditions \([\text{S4}]\) and \([\text{S5}]\) are obtained by differentiating the log-likelihood function Eq. \([\text{S3}]\) with respect to \( a_i \) for \( i = 1, \ldots, N \) and \( \xi(t) \) for \( t = 2, \ldots, \tau \). For \( t = 1, \xi(1) \) is normalized as one since otherwise there would arise a linear dependency between the optimality conditions and therefore the solution would be indeterminate. This reflects the fact that any combination of \( a_i, \ a_j \) and \( \xi(t) \) would satisfy the optimality conditions if \( a^*_i a^*_j = c \cdot a_i a_j \) and \( \xi^*(t) = \xi(t)/c \).

In solving the nonlinear equations \([\text{S4}]\) and \([\text{S5}]\), the initial values for \( a_i \) and \( \xi(t) \) are set as \( a_i = \sum_{j \neq i} (m_{ij}/\tau)/\sqrt{2 \sum_{j < j} m_{ij}/\tau} \) and 0.999, respectively.

Under the null model with a time-varying term, the average number of contacts between \( i \) and \( j \) over \( \tau \) periods is given by

\[
\lambda_{ij} \equiv \sum_{t=1}^{\tau} u(a_i, a_j, t), \quad \forall \ i, j.
\]

Thus, the number of contacts obeys a Poisson binomial distribution with mean \( \lambda_{ij} \) and variance \( \sigma_{ij} \equiv \sum_{t=1}^{\tau} (1 - u(a_i, a_j, t)) u(a_i, a_j, t) \). Since an exact functional form for a Poisson binomial distribution is intractable, we approximate the distribution of \( \{ m_{ij} \} \) with a Poisson distribution \([\text{H}0][\text{H}2]\):

\[
f(m_{ij}|a, \xi) \approx \frac{\lambda_{ij}^{m_{ij}} e^{-\lambda_{ij}}}{m_{ij}!} \equiv \tilde{f}(m_{ij}|a, \xi),
\]

\([\text{S7}]\)
where the error bound is given by Le Cam’s theorem [40–42]:

$$\sum_{m_{ij}=0}^{\infty} \left| f(m_{ij} a, \xi) - \frac{\lambda_{ij}^{m_{ij}} e^{-\lambda_{ij}}}{m_{ij}!} \right| < 2(1 - e^{-\lambda_{ij}}) \sum_{t=0}^{\tau} u(a_i, a_j, t)^2, \forall i, j.$$  \hspace{1cm} (S8)

We use Eq. (S7) in testing the significance of edge \((i, j)\) for a given observation \(m_{ij}^0\).

A comparison between the models with and without time-varying parameter is shown in Fig. S1. It shows that the test results are almost identical between the two null models; the numbers and the degree of overlap of identified significant ties suggests that the introduction of a time-varying parameter for capturing an activity rhythm does not affect the results shown in the main text.

**S2. BACKBONING METHODS FOR STATIC NETWORKS**

We recall here two well-known ways to assign a significance to edges and build backbones for static weighted networks. In our context, it can be done by first aggregating the temporal network on the available time-window, obtaining a network where the degree of a node is given by its number of distinct neighbors, the weight of an edge is total interaction time between two nodes, and the strength of an edge is given by

$$s_i = \sum_{j,t} A_{ij,t}.$$  \hspace{1cm} (a)

**A. Disparity filter**

The Disparity (DP) filter [13] is a filtering algorithm to classify the edges of a static weighted network into significant and insignificant ones. The DP filter uses only local information: the weight of an edge, \(\omega_{ij}\), the nodal degree, \(k_i\), and the strength, \(s_i\). The idea is that if node \(i\) has no specific relationship with its neighbors, then its strength (i.e., sum of weights) is distributed uniformly at random on the \(k_i\) edges incident to it. The authors of [13] show that the link between \(i\) and \(j\) is regarded as significant at filtering level \(\alpha\), if it satisfies the following condition:

$$1 - (k_i - 1) \int_0^{p_{ij}} (1 - x)^{k_i-2} dx < \alpha,$$  \hspace{1cm} (S9)

where \(p_{ij} = \omega_{ij}/s_i\). The LHS of Eq. (S9) represents the \(p\)-value for the null hypothesis that the edge weights are distributed uniformly at random. In fact, as argued in Gemmetto et al. [17], the significance of edge between \(i\) and \(j\) is not necessarily identical to that between \(j\) and \(i\) even for an undirected network. Therefore, one needs to test the significance of “two edges” \((i, j)\) and \((j, i)\) independently, and then the (undirected) edge is regarded as significant if at least one of the two “edges” satisfies the criterion (S9).
B. ECM filter

The ECM (enhanced configuration model) filter [17] is developed based on the idea that statistically significant edges are the ones whose presence cannot be explained by random chance. More specifically, the entropy-maximizing random matching probabilities are calculated with the ECM in which edge weights are distributed at random as uniformly as possible subject to two constraints: $⟨ k_i ⟩ = k^*$ and $⟨ s_i ⟩ = s^*$, where $x^*$ denotes the empirical value of variable $x$. Gemmetto et al. [17] show that the $p$-value for edge $(i, j)$ is then given by

$$\gamma^*_{ij} = p_{ij}(y_i^* y_j^*)^{u_{ij}^*-1},$$

(S10)

where

$$p^*_{ij} = \frac{x_i^* x_j^* y_i^* y_j^*}{1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*},$$

(S11)

and $x_i^*$ and $y_i^*$ represent hidden variables (or auxiliary variables) [43, 44] that solve the following conditions:

$$k_i^* = \sum_{j \neq i} \frac{x_i x_j y_i y_j}{1 - y_i y_j + x_i x_j y_i y_j} \quad \forall i,$$

(S12)

$$s_i^* = \sum_{j \neq i} \frac{x_i x_j y_i y_j}{(1 - y_i y_j)(1 - y_i y_j + x_i x_j y_i y_j)} \quad \forall i.$$

(S13)

One needs to solve a system of $2N$ nonlinear equations to obtain $\vec{x}^*$ and $\vec{y}^*$. In fact, $-\ln x_i (-\ln y_i)$ corresponds to a Lagrange multiplier associated with the constraint $(k_i) = k^*_i \ (s_i) = s^*_i)$. The backbone of a weighted network with significance level $\alpha$ is the network consisting only of edges $(i, j) \in \{(i, j) : \gamma^*_{ij} < \alpha\}$. Our implementation for the calculation of Eqs. (S12) and (S13) is based on the “Max & Sam” method proposed in [45], and the MATLAB code is available from [46].

S3. DEFINITIONS OF ROC CURVE AND AUC

The receiver operating characteristic (ROC) curve is a plot of true positive rates against false positive rates for different cutoffpoints of a test statistic. In our context, we want to know how well the significance of an edge can predict whether that edge is an intra-community edge. For this purpose, we use the $p$-value of an edge in a given filtering test as a measure of edge significance. That is, different points on an ROC curve denote different cutoff levels of $p$-values for a given filtering method (Fig. S2). For the ST filter presented in the main text, the $p$-value of an edge $i – j$ is simply $1 – G(m_{ij}^* | a_i^*, a_j^*)$.

For a given $p_0$, we consider therefore as True Positives (TP) the edges that have a $p$-value lower than $p_0$ (i.e., are considered significant) and are intra-community edges. False Positives (FP) are instead the inter-community edges with a $p$-value lower than $p_0$. Similarly, True Negatives (TM) are edges with $p$-value larger than $p_0$ and inter-community, and False Negatives (FN) the intra-community edges with $p$-value larger than $p_0$. The false positive rate is given by $FP/(TN + FP)$ and the true positive rate by $TP/(TP + FN)$.

The area under the ROC curve (AUC) quantifies the goodness of the $p$-values for the task of predicting intra-community edges. If the $p$-value of a filtering method perfectly distinguishes between intra- and inter-community edges (i.e., higher $p$-values indicate inter-community edges), then the value of AUC will be one. If the $p$-value is instead a poor indicator so that it is not different from a random prediction, then the AUC will be close to 0.5 (i.e., the ROC curve is then be a straight line joining the points $(0, 0)$ and $(1, 1)$).

We show the ROC curves and a comparison of the AUCs for different data sets in Fig. S2 and Fig. 4c in the main text.

S4. TOTAL NUMBER OF INTERACTIONS

The total number of empirical snapshot edges, denoted by $M$, is given by

$$M = \sum_{i<j} m_{ij}.$$

(S14)
FIG. S2. ROC curve for the detection of intra-community edges for the data sets with a clear community structure, for the three filters considered in the main text: Disparity filter (DP), Enhanced Configuration model (ECM) and ST filter. Here $\Delta = 15$ min.

The expected number of total snapshot edges within the null model, denoted by $M^*$, is given as

$$M^* = \tau \sum_{i<j} u(a_i^*, a_j^*). \quad (S15)$$

Figure S3 illustrates the comparison between $M$ and $M^*$, for different data sets, with temporal resolution $\Delta = 15$ min. The almost perfect fit between the estimated value of $M^*$ and the empirical $M$ indicates that the set of random matching probabilities, $\{a_i^*, a_j^*\}$, accurately explain the total number of contacts, suggesting that the maximum likelihood estimation of the activity vector works well for a wide range of temporal-network data.

FIG. S3. Total numbers of snapshot edges in the model and the real data. We set $\Delta = 15$ min for all data sets except Interbank (for which the temporal resolution is 1 day).
FIG. S4. Number of significant ties vs. $\alpha$, for different parameter values and the different filters considered.
FIG. S5. Weighted measure of the similarity between the backbones obtained by various filtering methods, vs. the number of edges retained in each case. Here we use the cosine similarity between the weights of the edges retained by two methods, defined as

\[ \sigma(x, x') = \frac{\sum_{i<j} w_{ij} w'_{ij}}{\sqrt{\sum_{i<j} (w_{ij})^2} \sqrt{\sum_{i<j} (w'_{ij})^2}} \]

where \( x \) encodes both the filtering method (ST, DP, ECM) and the significance level \( \alpha \) and the sums run only on the pairs of nodes present in both backbones.
FIG. S6. Weight distributions of significant and non-significant edges for a fixed filtering level ($\alpha = 10^{-4}$) and $\Delta = 1$ min, for the ST filter (left column), the DP filter (middle column) and the ECM filter (right column).
FIG. S7. Clustering coefficient and share of intra-class edges among the significant edges, as a function of the number of such edges, for the various filters. In addition to DP, ECM and ST we consider here the simplest filtering retaining the edges with largest weight, and a random filter selecting edges at random. $\Delta = 1$ min.
FIG. S8. Number of significant triads vs. the significance level \( \alpha \), for different data sets and temporal resolutions.
FIG. S9. Share of triangles with a given number of dyadic significant ties, vs. the filtering level $\alpha$, for various data sets and temporal resolutions.

FIG. S10. Share of significant triads among the triangles composed by three significant ties. $\Delta = 1$ min.