

Reconstructing irreducible links in temporal networks: which tool to choose depends on the network size

Original

Reconstructing irreducible links in temporal networks: which tool to choose depends on the network size / Nadini, Matthieu; Rizzo, Alessandro; Porfiri, Maurizio. - In: JOURNAL OF PHYSICS. COMPLEXITY. - ISSN 2632-072X. - ELETTRONICO. - 1:1(2020), p. 015001. [10.1088/2632-072X/ab6727]

Availability:

This version is available at: 11583/2833456 since: 2020-06-09T17:13:49Z

Publisher:

IOP Publishing

Published

DOI:10.1088/2632-072X/ab6727

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

IOP postprint/Author's Accepted Manuscript

"This is the accepted manuscript version of an article accepted for publication in JOURNAL OF PHYSICS. COMPLEXITY. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at <http://dx.doi.org/10.1088/2632-072X/ab6727>

(Article begins on next page)

PAPER • OPEN ACCESS

Reconstructing irreducible links in temporal networks: which tool to choose depends on the network size

To cite this article: Matthieu Nadini *et al* 2020 *J. Phys. Complex.* 1 015001

View the [article online](#) for updates and enhancements.

OPEN ACCESS



PAPER

Reconstructing irreducible links in temporal networks: which tool to choose depends on the network size

RECEIVED

25 September 2019

REVISED

27 November 2019

ACCEPTED FOR PUBLICATION

11 December 2019

PUBLISHED

29 May 2020

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Matthieu Nadini¹, Alessandro Rizzo^{2,3} and Maurizio Porfiri^{1,4}

¹ Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Brooklyn, NY, 11201, United States of America

² Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino, Italy

³ Office of Innovation, New York University Tandon School of Engineering, Brooklyn, NY 11201, United States of America

⁴ Department of Biomedical Engineering, New York University Tandon School of Engineering, Brooklyn, NY 11201, United States of America

E-mail: alessandro.rizzo@polito.it and mporfiri@nyu.edu

Keywords: activity, backbone network, disparity filter, fitness model, statistical filtering, time-varying

Abstract

Filtering information in complex networks entails the process of removing interactions explained by a proper null hypothesis and retaining the remaining interactions, which form the backbone network. The reconstructed backbone network depends upon the accuracy and reliability of the available tools, which, in turn, are affected by the specific features of the available dataset. Here, we examine the performance of three approaches for the discovery of backbone networks, in the presence of heterogeneous, time-varying node properties. In addition to the recently proposed evolving activity driven model, we extend two existing approaches (the disparity filter and the temporal fitness model) to tackle time-varying phenomena. Our analysis focuses on the influence of the network size, which was previously shown to be a determining factor for the performance of the evolving activity driven model. Through mathematical and numerical analysis, we propose general guidelines for the use of these three approaches based on the available dataset. For small networks, the evolving temporal fitness model offers a more reasonable trade-off between the number of links assigned to the backbone network and the accuracy of their inference. The main limitation of this methodology lies in its computational cost, which becomes excessively high for large networks. In this case, the evolving activity driven model could be a valid substitute to the evolving temporal fitness model. If one seeks to minimize the number of links inaccurately included in the backbone network at the risk of dismissing many links that could belong to it, then the temporal disparity filter would be the approach-of-choice. Overall, our contribution expands the toolbox of network discovery in the technical literature and should help users in choosing the right network discovery instrument, depending on the problem considered.

1. Introduction

Temporal networks have gained momentum in the scientific literature and now constitute a well-established mathematical framework to describe temporal interactions (links) among dynamical entities (nodes) in many real systems [1–3]. Temporal interactions are ‘reducible’, when they are uniquely explained by inherent node properties that do not require the dyadic nature of a network. They are ‘irreducible’ when they carry additional information that cannot be associated with the sole node properties [4].

This dichotomy calls for the design of network filtering algorithms to distinguish reducible from irreducible links, which form the so-called network backbone. For instance, when studying interactions in school or workplace, one may seek to remove accidental interactions and identify strong bounds that underlie friendship or strong work relationships [5]. Introducing reliable approaches for filtering reducible links and reconstructing the backbone is critical for the study of real temporal networks.

Various approaches have been proposed to discover the backbone, bringing to light salient structural features of complex systems. In [6], the authors propose to identify significant heterogeneities in the link weights to unveil a backbone that is invariant with respect to changes in the network scale. In [7], the paradigm of statistically validated networks is used to isolate links in bipartite topologies. The technique has been further applied in [8] to examine the evolution of motifs in mobile communications in Europe and in China. Information theory—notably the concept of maximum entropy—has been leveraged in [4] to establish an unbiased filtering and backbone detection algorithm for weighted networks. Other efforts have been pursued to discover the backbone in transport networks [9], identify the relationship between weights and topology in weighted networks [10], recover significant structures in the ensemble of US airports [11], study functional networks in the brain from magnetic resonance data [12], highlight cores of communities in bipartite networks [13], and determine topological structures that serve as superhighways for transport phenomena in complex systems [14].

The general idea of these approaches is to aggregate all the temporal interactions and compare the resulting weighted network against a suitable null model. If the observed link weight exceeds the model's prediction, the link is included in the backbone network, otherwise it is filtered out. This procedure is repeated for all the links to reconstruct the whole backbone.

Recent approaches contemplate the possibility that the generation of reducible links is mediated by time-varying processes [5, 15–20]. Compared to earlier studies, these approaches offer a more accurate description of the network evolution, which improves the backbone network detection. For example, the evolving activity driven model (EADM) [19] is an approach inspired by the paradigm of activity driven networks [21–23]. The key element of this methodology is the use of a node-specific activity to represent the individual propensity of generating links over time [21]. Within the EADM, activities are treated as time-varying, piece-wise constant functions. By letting individual properties (activities) vary over time, the number of false positives (links incorrectly identified as a part of the backbone network) are reduced with respect to alternative schemes that keep node properties as constant [19].

As suggested in [19] through preliminary numerical studies, the EADM is successful in the discovery of the backbone structure only for large networks, thereby failing to assist in the inference of irreducible links for small networks. This issue is called finite-size effect and it is well known in the scientific literature on network science. In general terms, the finite-size effect describes a limitation in the approach that could either relate to the capability of the method to provide reliable results or to its unsustainable computational cost. For example, mean-field theories are usually suitable to predict the evolution of diffusion processes on large networks [24, 25], but they may fail to describe aspects important at small scales. Conversely, other theories may yield successful results for small networks, such as those used for the inference of networks from time-series of individuals nodes [26, 27], while failing in large systems due to excessive computational costs.

In order to overcome the EADM finite-size effect, we introduce two new approaches: the evolving temporal fitness model (ETFM) and the temporal disparity filter (TDF). Both the approaches are extensions of existing methodologies for backbone discovery to incorporate the possibility of time-varying processes for the generation of reducible links. The ETFM extends the temporal fitness model [5], which offers a flexible framework for network discovery. Similar to the EADM, the ETFM assigns to each node in the network an activity, which is heterogeneously distributed. The TDF extends the disparity filter [6], which is currently the most popular approach to backbone detection and the first one to address heterogeneously weighted networks. Contrary to the EADM and ETFM, the TDF is not based on the use of activities, but it utilizes other node-specific properties, such as the node strength and degree [28, 29].

The EADM, ETFM, and TDF share the same three sequential steps: (i) determining the interval partition, (ii) estimating the models' parameters, piece-wise constant over successive intervals, and (iii) implementing a statistical filter, which removes all links explained by the null hypothesis and retains the backbone network. In principle, their performance may vary depending on the problem under consideration. Here, we focus on the influence of network size on backbone detection, which was shown to be a limitation of the EADM in [19]. Through this targeted effort, we aim at articulating the technical reasons why the EADM suffers from a finite-size effect and at determining when either the ETFM or TDF should be preferred.

First, we investigate the computational complexity of the EADM, ETFM, and TDF. The only methodology that is unfeasible for large networks is the ETFM, due to its excessive computational cost. Such a drawback is partially compensated by an improved capability to deal with small networks, whereby the ETFM estimates the activities in a maximum likelihood sense that allows for the study of small systems.

Next, we consider an artificial temporal network with a known backbone network following [19], toward comparing the performance of the EADM, ETFM, and TDF in reconstructing the backbone network. We find that the size of the artificial network strongly affects the performance of the approaches, measured in terms of true positives (links correctly included in the backbone network) and false positives (links incorrectly included in the backbone network). The TDF finds the least amount of false positive links among the three approaches, ensuring that all of them belong to the known backbone network. The main caveat of the TDF is that, in several cases, it

does not identify any true positive, thus failing to detect the known backbone network. The EADM has adequate performance in finding the known backbone network only when the network is large, while the ETFM can tackle the analysis of systems of any size. The main issue of the ETFM is that it detects several false positives when the known backbone network is present almost at any time step. In other words, the TDF detects the least amount of links (counted as the sum of false and true positives), the ETFM an intermediate value, and the EADM the largest amount. Further, when the network size is large, the ETFM and EADM detect the same backbone network. This is because, for large networks, the activity estimation of the ETFM is equivalent to the one of the EADM.

Finally, we detect the backbone network of two real systems freely available as part of the SocioPatterns project [30]. We study the Primary School and Workplace datasets, where nodes represent students (or teachers) and workers, respectively, interacting over successive days. We detect the backbone network in the entire school and workplace, but also in each class and department. The reconstructed backbone networks are coherent with expectation from the analysis of artificial networks: the TDF finds the least amount of links, the ETFM an intermediate value, and the EADM the largest number. Interestingly, we also observe that the EADM and ETFM detect the same backbone network when the real systems are sufficiently large.

Overall, we identify a strong finite-size effect in backbone detection, which suggests that care should be placed when selecting an approach. Further, the TDF minimizes the number of false positives, but often it does not detect any true positive, challenging the proper reconstruction of the backbone network. The ETFM is especially useful for small systems, offering a reasonable trade-off between the number of true and false positives detected. The EADM detects the same backbone network as the ETFM when the system is large, and should be preferred for the analysis of large systems due to its reduced computational burden.

The rest of the paper is organized as follows. In table 1, we summarize the acronyms used to indicate the new methodologies, the nomenclature of the article, and the parameters employed in the generation of the artificial networks. In section 2, we summarize the EADM and introduce the new approaches, namely, the ETFM (an extension of the temporal fitness model [5]) and the TDF (an extension of the disparity filter [6]). In section 3, we discuss our analytical and numerical findings for both artificial and real systems. Finally, in section 4, we draw our main conclusions and propose potential lines of further inquiry.

2. Significant links

In this section, we summarize the EADM to detect significant links [19] and introduce the ETFM, our extension of the temporal fitness model [5], and the TDF, our extension of the well-established disparity filter [6].

For all the approaches under consideration, we consider a temporal undirected network of N nodes that evolves in time over an observation window composed of $T \gg 1$ time steps, labeled by the discrete time index $t = 1, \dots, T$. At each time t , nodes interact among themselves and form a time-varying network of interactions, whose interconnection links are encoded in a binary adjacency matrix $A(t)$ that varies in time. The generic entry of $A(t)$ is defined as $A_{ij}(t)$, where $i, j = 1, \dots, N$ are the indices of the nodes. Through the paper, we adopt the following notation: given a quantity ' \cdot ', we define $\langle \cdot \rangle$ as its statistical average and $E[\cdot]$ as its expected value across multiple, independent realizations.

The EADM, ETFM, and TDF are similar in some aspects and different in others; a summary is provided in table 2. The approaches follow three sequential steps: (i) determining the interval partition, (ii) estimating model parameters, which are evolving over successive intervals, and (iii) designing a statistical filter, which removes all links explained by the null hypothesis and retains the backbone network. Details on the implementation of each of these steps and key differences among the methodologies are detailed in what follows.

2.1. Evolving activity driven model (EADM)

2.1.1. Interval partition

In our previous work [19], we demonstrated that dividing the overall observation window into successive intervals improves the accuracy of the reconstruction. Therein, we compared a naïve supervised partitioning method, which divides the time series at random, with an unsupervised one (the Bayesian Block (BB) method [31]). We found that the latter should be preferred for its performance over artificial network benchmarks, its freedom from the need of prior information about the interval partition, and its implementability on relatively short time series.

The BB method [31] employs maximum likelihood and marginal posterior functions to distinguish statistically significant features from random observational errors. It starts with considering the first time step $t = 1$, and then it adds and analyzes the subsequent time instants $t = 2, \dots, T$, approximating the time series in piecewise constant segments, where a homogeneous number of links is observed. The algorithm coherently builds on iterations from past time steps to efficiently compute future interval partition.

In our application, the BB method analyzes the total number of temporal links created in the entire network at time t

Table 1. Acronyms used to identify the methodologies under consideration, main variables of the article, and parameters employed in the generation of the artificial networks. The superscript ‘ts’ indicates that the variable is estimated from the time series of individual links.

Methodology	Acronym
Evolving activity driven model	EADM
Evolving temporal fitness model	ETFM
Temporal disparity filter	TDF
Variable	Description
N	Number of nodes in the system
T	Total number of discrete time steps comprising the whole observation window
I	Number of intervals used to partition the whole observation window
t, Δ	Indices for discrete time instants and intervals
i, j	Indices for the nodes in the system
$A_{ij}^{ts}(t)$	Generic entry of the adjacency matrix at time t
$\Omega^{ts}(t)$	Total number of temporal links created in the entire network at time t
$a_i(t)$	Activity value of node i at time t
$p_{ij}(t)$	Probability that node i interacts with node j at time t
$t_{in}(\Delta), \tau(\Delta)$	Initial time instant and length of the Δ th interval
$n(\Delta)$	Number of nodes that create at least a link in the Δ th interval
$s_i^{ts}(\Delta), \bar{s}_i^{ts}$	Number of links created by node i in the Δ th interval and in the whole observation window
$w_{ij}^{ts}(\Delta), \bar{w}_{ij}^{ts}$	Number of links exchanged between nodes i and j in the Δ th interval and in the whole observation window
$W^{ts}(\Delta)$	Total number of links generated in the entire network in the Δ th interval
$r_{ij}^{ts}(t), \bar{r}_{ij}^{ts}$	Normalized weight of links exchanged between nodes i and j at time t and in the whole observation window
\bar{k}_i^{ts}	Degree of node i in the whole observation window
α_{ij}	p-value of the EADM and ETFM
γ_{ij}	p-value of the TDF
N_E	Number of links created at least once in the whole observation window
β, β'	Univariate and multivariate threshold adopted to determine irreducible links
$\langle \cdot \rangle, E[\cdot]$	Statistical average and the expected value across multiple, independent realizations of a quantity ‘.’
Parameter	Description
λ_{ij}	Preponderance of the link between nodes i and j in the artificial backbone
$F(a), G(\lambda)$	Distributions of activity values and preponderance of link in the artificial backbone
a_{min}, λ_{min}	Lower cut-offs of the two distributions
ϵ	Fraction of links belonging to the artificial backbone
p	Autocorrelation parameter of the activity values in different intervals

Table 2. Summary of the three approaches under consideration, the evolving activity driven model (EADM), evolving temporal fitness model (ETFM), and temporal disparity filter (TDF).

Approaches	Interval partition	Parameter estimation	Statistical filter
EADM	BB method [31]	Weighted configuration model [19, 32]	Poisson distribution [5, 19]
ETFM	BB method [31]	Maximum likelihood [5]	Poisson distribution [5, 19]
TDF	BB method [31]	Normalized weights [6]	Disparity filter [6]

$$\Omega^{ts}(t) = \sum_{i,j=1;i \leq j}^N A_{ij}^{ts}(t), \tag{1}$$

and returns the interval partition, which divides the overall time window T into I disjoint intervals indexed by $\Delta = 1, \dots, I$, containing approximately a uniform total number of connections [31]. The superscript ‘ts’ indicates that these variables are estimated from the time series of the individual links. From the knowledge of the interval partition, we determine the length, $\tau(\Delta)$, of the generic Δ th interval under the closure relation $\sum_{\Delta=1}^I \tau(\Delta) = T$.

2.1.2. Parameter estimation

In order to detect the backbone network, we should determine the extent by which each pair of nodes i and j will interact by simple chance. The EADM is based on a heterogeneously distributed parameter, the activity, which represents the propensity of generating links over time [21]. The probability that node i interacts with node j at time t is the product of the two activities [5, 19],

$$p_{ij}(t) = a_i(t)a_j(t). \quad (2)$$

Individual activities are estimated according to the weighted configuration model [19, 32].

Since the I intervals found by the BB method are disjoint, we can analyze each of them separately. For instance, we consider the generic Δ th time interval that starts at $t = t_{in}(\Delta)$ and ends at $t = t_{in}(\Delta) + \tau(\Delta) - 1$. The activity of node i at time $t \in [t_{in}(\Delta), t_{in}(\Delta) + \tau(\Delta) - 1]$ is computed through the following formula:

$$a_i(t) = \frac{s_i^{ts}(\Delta)}{\sqrt{2W^{ts}(\Delta)\tau(\Delta)}}, \quad (3)$$

where $s_i^{ts}(\Delta)$ and $W^{ts}(\Delta) \gg 1$ are the node strength and the total number of temporal links generated in the network in the Δ th interval, respectively. Specifically, $s_i^{ts}(\Delta)$ and $W^{ts}(\Delta)$ are computed from $A^{ts}(t)$ through

$$s_i^{ts}(\Delta) = \sum_{j=1}^N \sum_{t=t_{in}(\Delta)}^{t_{in}(\Delta)+\tau(\Delta)-1} A_{ij}^{ts}(t) \text{ and} \quad (4)$$

$$W^{ts}(\Delta) = \frac{1}{2} \sum_{i=1}^N s_i^{ts}(\Delta).$$

Once the activities are estimated according to equation (3), the probability in equation (2) can be computed as

$$p_{ij}(t) = \frac{s_i^{ts}(\Delta) s_j^{ts}(\Delta)}{2W^{ts}(\Delta)\tau(\Delta)}. \quad (5)$$

2.1.3. Statistical filter

The typical objective of a statistical filter is to test whether a specific link ij can be explained by the null hypothesis. If it were explained by the null model, it should be filtered out, otherwise it should be included in the backbone network. The same statistical test has to be applied to each of the N_E links that appear at least once over the entire observation window.

The EADM requires to compare the total number of connections between node i and node j , \bar{w}_{ij}^{ts} , observed in the entire observation window with its expected quantity, $E[\bar{w}_{ij}]$. The observed number of connections can be computed from $A_{ij}^{ts}(t)$ as

$$\bar{w}_{ij}^{ts} = \sum_{\Delta=1}^I w_{ij}^{ts}(\Delta) = \sum_{\Delta=1}^I \sum_{t=t_{in}(\Delta)}^{t_{in}(\Delta)+\tau(\Delta)-1} A_{ij}^{ts}(t). \quad (6)$$

The expected number of connections between nodes i and j in the time window T is calculated as the sum of the binomial random variables given in equation (2) [19]

$$E[\bar{w}_{ij}] = \sum_{t=1}^T p_{ij}(t). \quad (7)$$

The sum of non-identical binomial random variables in equation (7) is a Poisson binomial distribution, which is usually approximated by the Poisson distribution [33–35], due to a lack of closed-form expression.

The confidence that the observed weight, \bar{w}_{ij}^{ts} , could be explained by its corresponding expected weight, $E[\bar{w}_{ij}]$ in equation (7), is computed according to the cumulative function of the Poisson distribution

$$\alpha_{ij} \equiv 1 - \sum_{x=0}^{\bar{w}_{ij}^{ts}-1} P(x; E[\bar{w}_{ij}]), \quad (8)$$

where $P(x; E[\bar{w}_{ij}])$ indicates the Poisson distribution with random variable x and expected value $E[\bar{w}_{ij}]$. Equation (8) represents the p-value, α_{ij} , that the observed link ij can be explained by the null hypothesis of being extracted from a Poisson distribution. If the p-value is below a pre-defined threshold, then the link is deemed as significant and retained in the backbone network.

From equation (8), a total of N_E p-values are computed, one for each link created at least once in the overall network evolution. The interval partition begets better estimation of the expected number of links between nodes i and j , in equation (7), thereby improving on the estimation of the p-values through equation (8).

2.2. Evolving temporal fitness model (ETFM)

As detailed in table 2, the interval partition is carried using the BB method [31], thereby following section 2.1.1. Similarly, the statistical filter is equivalent to the one of the EADM, explained in section 2.1.3. Here, we detail the main difference between the EADM and ETFM, which pertains to the activity estimation.

Similar to the EADM, the probability that node i interacts with node j at time t is the product of the two individual activities, as given in equation (2). However, the ETFM estimates the activities in a maximum likelihood sense, based on the approach introduced in [5]. Specifically, given that the intervals returned from the BB method are disjoint, we can apply the maximum likelihood presented in [5] to each of the I intervals. The input parameters of the maximum likelihood are computed according to the EADM in equation (3).

For instance, in the Δ th interval, starting at $t = t_{\text{in}}(\Delta)$ and ending at $t = t_{\text{in}}(\Delta) + \tau(\Delta) - 1$, individual activities are assumed to be constant and estimated according to equation (3). Some nodes in the network, however, do not generate nor receive connections. These nodes have an activity equal to zero and they should be discarded in the maximum likelihood computation. Hence, from the set of all possible activities at time t , $\mathbf{a}(t) = (a_1(t), \dots, a_N(t))$, the maximum likelihood estimation is applied only to those entries with $a_i(t) > 0$. Assuming that there are $n(\Delta) \leq N$ of these nodes, we indicate the reduced set of optimal activities as $\mathbf{a}^*(t) = (a_1^*(t), \dots, a_{n(\Delta)}^*(t))$.

The estimation of the optimal activity values is carried similar to [5], whereby

$$\sum_{j=1; j \neq i}^{n(\Delta)} \frac{w_{ij}^{\text{ts}}(\Delta) - \tau(\Delta) a_i^*(t) a_j^*(t)}{1 - a_i^*(t) a_j^*(t)} = 0, \quad (9)$$

$$a_i^*(t) \in (0, 1), \quad i = 1, \dots, n(\Delta).$$

All the individual activities are constant (between zero and one) within the Δ th interval, and $w_{ij}^{\text{ts}}(\Delta)$ is given by equation (6). Once the optimal values for the activities are estimated from equation (9), they can be used to estimate the probability in equation (2).

2.3. Temporal disparity filter (TDF)

According to table 2 and similar to the EADM and ETFM, the interval partition is carried out through the BB method implemented on the time series of the individual links [31]. The parameter estimation and the statistical filter are different from the EADM and they are explained in the following.

2.3.1. Parameter estimation

In the TDF, parameters are estimated directly from the time series, without relying on any expected quantity like the activities for the EADM and ETFM. For instance, in the Δ th interval that starts at $t = t_{\text{in}}(\Delta)$ and ends at $t = t_{\text{in}}(\Delta) + \tau(\Delta) - 1$, a number of connections between node i and node j are observed, $w_{ij}^{\text{ts}}(\Delta)$, in equation (6). Similar to the disparity filter [6], two normalized quantities are computed by dividing $w_{ij}^{\text{ts}}(\Delta)$ either by the strength of node i , $s_i^{\text{ts}}(\Delta)$, or node j , $s_j^{\text{ts}}(\Delta)$, in equation (4), respectively, such that

$$r_{ij}^{\text{ts}}(t) = \frac{w_{ij}^{\text{ts}}(\Delta)}{s_i^{\text{ts}}(\Delta)}. \quad (10)$$

These quantities must satisfy the closure condition $\sum_j r_{ij}^{\text{ts}}(t) = 1 \forall t \in [t_{\text{in}}(\Delta), t_{\text{in}}(\Delta) + \tau(\Delta) - 1]$. Note that while $w_{ij}^{\text{ts}}(\Delta) = w_{ji}^{\text{ts}}(\Delta)$, in general $r_{ij}^{\text{ts}}(t) \neq r_{ji}^{\text{ts}}(t)$ since the strengths of nodes i and j can be different.

Over the entire time window T , the normalized weight, \bar{r}_{ij}^{ts} , is computed using a weighted time average starting from the results of equation (10), yielding

$$\bar{r}_{ij}^{\text{ts}} = \frac{1}{T} \sum_{\Delta=1}^I \frac{w_{ij}^{\text{ts}}(\Delta) \tau(\Delta)}{s_i^{\text{ts}}(\Delta)}. \quad (11)$$

Here, the longer is the interval, the greater is its contribution to the overall average. If only one interval is returned by applying the BB method to the time series, equation (11) reduces to the estimation of the classical disparity filter [6].

2.3.2. Statistical filter

The TDF statistical filter takes two quantities as input: the normalized weights computed according to equation (11) and the observed degree of either node i , \bar{k}_i^{ts} , or node j , \bar{k}_j^{ts} , over the entire network evolution. The node degree is computed from equation (6) as

$$\bar{k}_i^{\text{ts}} = \sum_{j=1}^N \mathcal{H}[\bar{w}_{ij}^{\text{ts}} - 1], \quad (12)$$

where \mathcal{H} is the Heaviside function that takes a value equal to one when $\bar{w}_{ij}^{\text{ts}} \geq 1$, and zero otherwise.

The TDF null hypothesis is that a node connects uniformly at random with all its neighbours. For node i with degree \bar{k}_i^{ts} , the TDF considers as a baseline that node i exchanges an equal number of connections with all its neighbors, which results in all normalized weights from equation (11) being equal to $1/\bar{k}_i^{\text{ts}}$. Irreducible links included in the backbone network have a normalized weight \bar{r}_i^{ts} greater than the expectation from the null hypothesis.

The advantage of the TDF with respect to the original disparity filter [6] is that it accounts for time-variations in node properties, through equation (10). The disparity filter in its original incarnation can be only applied to weighted networks. A p-value has to be associated with any link in the network. We use the same formula proposed for the disparity filter [6], namely,

$$\gamma_{ij} \equiv 1 - \left(\bar{k}_i^{\text{ts}} - 1 \right) \int_0^{\bar{r}_{ij}^{\text{ts}}} (1-x)^{\bar{k}_i^{\text{ts}}-2} dx, \quad (13)$$

where γ_{ij} is the p-value. In general, equation (13) may yield a different result if it is specialized to node i or node j , even for undirected networks. Therefore, it is necessary to consider an additional p-value for the undirected link ij , by focusing on node j rather than i , such that we compute γ_{ji} in lieu of γ_{ij} . If at least one of the two p-values γ_{ij} or γ_{ji} is lower than a pre-defined threshold, the undirected link ij is included in the backbone network. If only one interval is returned by applying the BB method to the time series, the TDF becomes equivalent to the disparity filter [6].

Similar to equations (7) and (8), the temporal partition is used to estimate the network weights and a p-value is associated with each link of the network that appeared at least once during the overall network evolution. Graphical representations of the ETFM and TDF are given in figure 1.

2.4. Multiple tests comparison

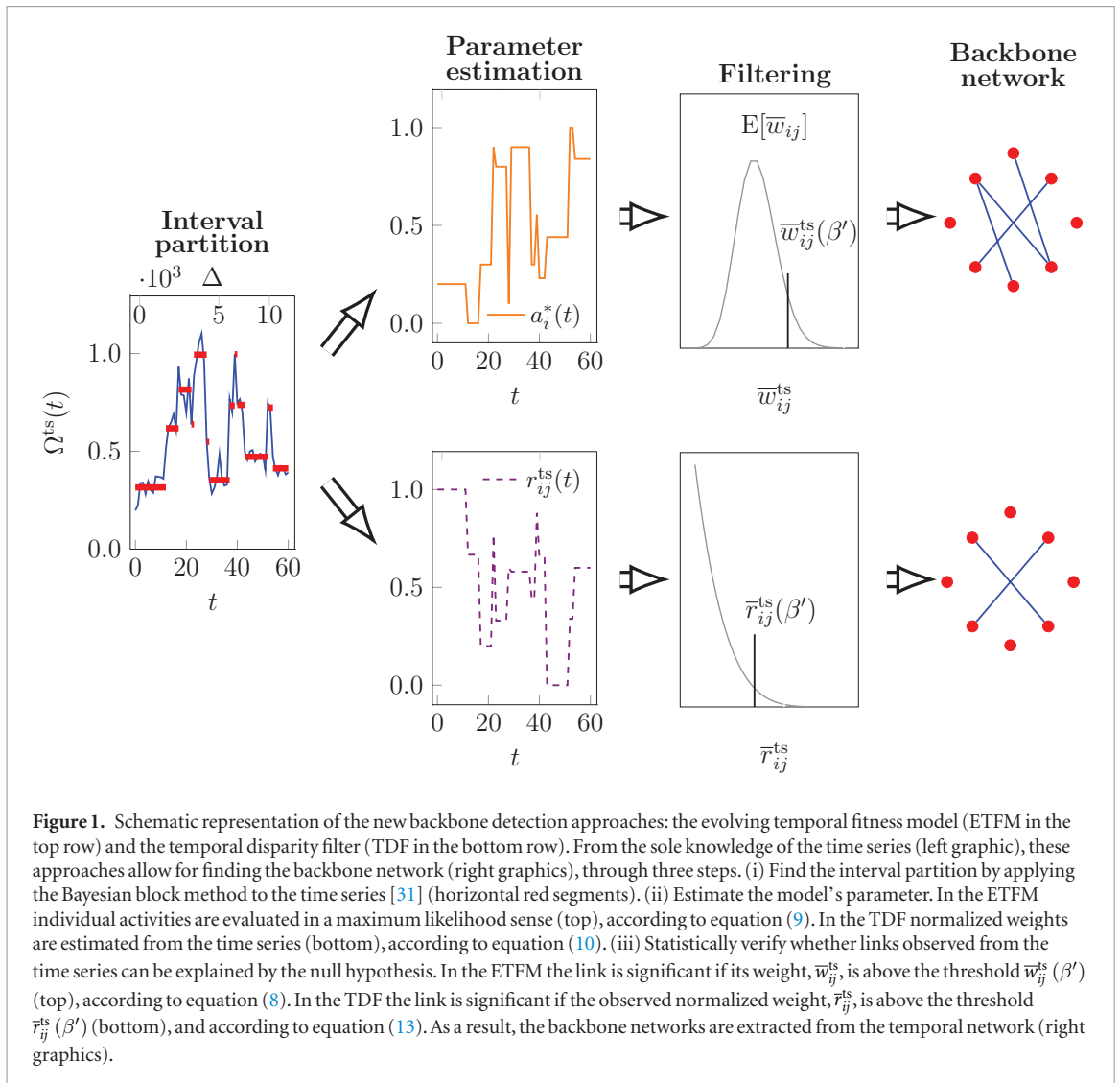
In all of the three approaches under consideration, a statistical test is carried out for all the N_E links observed at least once in the evolving network. The EADM and ETFM use equation (8), while the TDF employs equation (13). Since multiple hypotheses are tested, a multiple hypothesis test correction is required [36]. Here, we choose the Bonferroni correction [37]. It computes the multivariate threshold equal to $\beta' = \beta/N_E$, where β is the univariate threshold, usually set equal to 0.01. This correction assures that no false positives will be included in the backbone network with probability $1 - \beta$. An alternative to the Bonferroni correction might be the false discovery rate (FDR) [38], which is a less restrictive option because it ensures that the fraction of false positive is less than β .

3. Results

Here, we discuss how finite-size effects influence the performance of the EADM, ETFM, and TDF, toward the ultimate objective of informing the selection of the most appropriate approach as a function of the size of the network.

Since the main goal of these approaches is to infer the backbone network in real datasets, we begin by briefly summarizing the features of the Primary School dataset under consideration, freely available as part of the Socio-Patterns project [30]. In appendix C, we analyze the Workplace dataset, also part of the SocioPatterns project [30], in a similar way.

Upon concluding the analysis of the Primary School dataset, we illustrate the features of the generative algorithm for the synthetic networks examined in this study. Our generative mechanism entails the creation of artificial networks with arbitrary distributions of both the activity of the network nodes and of the link weight in the network backbone. In the main text, we consider the case of a heterogeneous activity distribution and homogeneous backbone, similar to our previous work [19]. In appendix A, we contemplate three additional cases: (i) heterogeneous activity distribution and heterogeneous backbone, (ii) homogeneous activity distribution and homogeneous backbone, and (iii) homogeneous activity distribution and heterogeneous backbone.



Later, we study the computational complexity of the EADM, ETFM, and TDF, concluding that the ETFM is too demanding for application to large networks composed of several hundreds or more nodes, while the EADM and TDF are scalable approaches, suitable for large networks.

Next, we evaluate the error of the EADM and ETFM, by comparing expected with observed quantities. This analysis cannot be carried out for the TDF, because it does not compute any expected quantity and its filter takes as input only observed quantities (such as the node strength and degree). We determine that the ETFM optimizes the activities in a maximum likelihood sense and its results are always more accurate than those of the EADM. However, the error of the EADM decreases with the inverse of the size of the network so that the error is reasonably low for many applications on large systems.

In order to more precisely estimate how finite-size effects influence the EADM, ETFM, and TDF performance, we consider an artificial network, where the backbone network is a-priori known [19]. We find that the TDF brings to zero the false positives (links incorrectly identified as part of the backbone network). However, most of the links that truly belongs to the backbone network are not identified by the TDF. The ETFM offers a more flexible framework that allows for detecting most of the links belonging to the backbone network, without yielding an excess of false positives. The EADM has equivalent performance to the ETFM for networks with a hundred nodes or more, but performs poorly for smaller networks. Overall, we find that the TDF detects the least amount of true backbone links among the considered approaches. The ETFM finds an intermediate value and the EADM begets the largest number.

Finally, we illustrate our numerical arguments by studying the Primary School and Workplace datasets. We find that the TDF has the most restrictive null hypothesis, which allows for including in the backbone network only a few links. This approach should be used only if we must avoid false positives at all costs. The ETFM detects more links than the TDF and affords to study features of the backbone network that would have been lost otherwise, in agreement with prior studies [5]. When the network contains at least one hundred of nodes, the ETFM and EADM perform similarly, so that the EADM should be preferred from its reduced computational burden.

Table 3. Data summary of the Primary School dataset. The temporal networks evolve over successive time steps of length equal to either 5 or 15 min and all multiple links within that temporal resolution are removed. The ‘# temporal links’ column indicates the number of temporal links retained in the dataset for their respective temporal resolutions, in brackets. The ‘# static links’ column indicates the number of aggregated links, which corresponds to links appearing at least once in the overall temporal evolution. The ‘All’ dataset corresponds to the whole Primary School, the ‘Teachers’ dataset correspond to the interactions between teachers, and all the other datasets are named with the respective class-section in the Primary School. We always remove the time interval between the last connection of a day and the first one of the following day.

Data	# nodes	# temporal links [5 min]—[15 min]	# static links
1A	23	3295—2435	249
2A	23	3665—2549	246
3A	23	3933—2815	252
4A	21	3173—2228	209
5A	22	3556—2560	225
1B	25	6299—4137	299
2B	26	4193—2832	315
3B	22	4233—2906	229
4B	23	2033—1471	239
5B	24	3601—2543	270
Teachers	10	83—65	29
All	242	55 046—39 583	8317

Further, in appendix D, we compare the TDF with its basic version, the disparity filter, showing that the two approaches yield different predictions for the Primary School and Workplace datasets.

3.1. Data processing

Here, we describe in detail our approach to the analysis of the freely available Primary School dataset [30], a similar analysis is carried out for the Workplace dataset [30] in appendix C. Our approach is inspired by previous efforts [5, 19]. In the original dataset, person to person interactions are recorded every 20 seconds; however, in reality, interactions might last longer than this pre-defined temporal resolution. In order to remove redundancies in the collected data and confounds associated with a particular resolution [5, 19], we study two temporal resolutions of 5 and 15 min, and remove multiple connections at the same temporal resolution. We remove the time interval between the last connection of a day and the first one of the following day.

In order to explore the role of network size, we study different temporal networks from the Primary School dataset and detect a backbone network for each of them. Specifically, we focus on the temporal interactions occurring in each classroom (10 in total), the body of the teachers, and the entire school. A data summary is provided in table 3.

3.2. Generative mechanisms of artificial networks

We present a generative mechanism for temporal networks with known backbone network and arbitrary distribution of reducible and irreducible links. The mechanism extends the one introduced in [19], where temporal networks with heterogeneous activity distribution and homogeneous backbone are generated.

Artificial temporal networks are generated from the superposition of reducible and irreducible links. Reducible links are created according to equation (2) and evolve over the whole observation window composed by T time steps, partitioned into I intervals. Nodes have interval-dependent (piece-wise constant) activities $a_i(t)$, drawn from the activity distribution $F(a)$. Between two consecutive intervals, $t_1 \in [t_{in}(\Delta - 1), t_{in}(\Delta - 1) + \tau(\Delta - 1) - 1]$ and $t_2 \in [t_{in}(\Delta), t_{in}(\Delta) + \tau(\Delta) - 1]$, the activity values vary according to

$$a_i(t_2) = a_i(t_1)p + y(1 - p), \quad (14)$$

where p is an autocorrelation parameter and y a random number extracted from $F(a)$. Among all the links observed at least once, a fraction ϵ of them is arbitrarily assigned to the backbone network. The reponderance of a link between nodes i and j is λ_{ij} , such that if $\lambda_{ij} = 0$, the link appears according to the product of activity values, as in equation (2), while increasing λ_{ij} makes the backbone network easier to discover. The values of λ_{ij} are extracted from the distribution $G(\lambda)$. More details about the generative algorithm can be found in appendix A1.

In the main text, we consider the case of a heterogeneous activity distribution, $F(a) \sim a^{-2.1}$ and $a \in [a_{min}, 1]$, and homogeneous backbone network, $G(\lambda) \sim \delta(\lambda - \hat{\lambda})$, that is, $\lambda_{ij} = \hat{\lambda}$ for all link in the backbone network, where $\delta(\cdot)$ is the Dirac delta distribution. In appendix A, we examine three more cases: (i) heterogeneous activity distribution and heterogeneous backbone, (ii) homogeneous activity distribution and homogeneous backbone, and (iii) homogeneous activity distribution and heterogeneous backbone.

3.3. Computational complexity

All approaches determine the interval partition using the BB method, which requires a computational time of $\mathcal{O}(I_{\max}^2)$. Here, I_{\max} represents the number of times in which the total number of temporal links in the entire network at time t , as computed in equation (1), is different from what observed at $t + 1$, $\Omega(t) \neq \Omega(t + 1)$ for $t = 1, \dots, T - 1$, and \mathcal{O} is the Landau symbol. Similarly, the statistical filters in equations (8) and (13), share a similar computational cost that scales as $\mathcal{O}(N_E)$, where N_E is the number of links observed at least once in the temporal evolution.

The estimation of model parameters is different across the EADM, ETFM, and TDF. In the EADM, activities are estimated through equation (3), which scales linearly with the number of intervals, I , and the number of links created in each interval. The complexity of this approach is $\mathcal{O}(cNI)$, where $c \ll N$ for sparse networks and $c \sim N$ for dense networks. A similar reasoning applies for the estimation of the TDF's parameters in equations (10) and (11), yielding a complexity of $\mathcal{O}(cNI)$.

In the ETFM, activity values are optimized in a maximum likelihood sense, which requires to recursively solve equation (9). Hence, the ETFM computational cost depends on three factors: the number I of intervals, the number of nodes in the Δ th interval, $n(\Delta)$, and the number m of recursions. The computational cost grows linearly with the number of intervals I , since it has to be applied independently to each interval, and with m , which is the number of iterations for the maximum likelihood to converge. On the other hand, it is quadratic in $n(\Delta)$, because of equation (9). Overall, the ETFM complexity hits $\mathcal{O}(mIN^2)$ in the worst case scenario, while in the best case scenario, when only one interval is present, it is $\mathcal{O}(mN^2)$, like in the temporal fitness model [5].

By comparing the EADM and TDF complexities, $\mathcal{O}(cNI)$, with the ETFM complexity, $\mathcal{O}(mN^2I)$, we observe that the bottleneck is the size of the network. For large networks, the EADM and TDF are much faster procedures than the ETFM.

3.4. Accuracy

Our study involves the mathematical analysis of the model accuracy, which is verified through numerical simulations using the artificial networks described in section 3.2 and appendix A.1. In the main text, we focus on the case of heterogeneous activity distribution and homogeneous backbone network; in appendix A.2 we explore the case of heterogeneous backbone and in appendix A.3 the case of homogeneous activities.

The TDF statistical filter in equation (13) uses quantities that are computed directly from the time series, without relying on any expected value, which is otherwise needed for the EADM and ETFM. Therefore, the following analysis can be carried out only for the EADM and ETFM, since it involves comparison between expected and observed quantities. We focus on both microscopic and macroscopic quantities, including the observed node strength, s_i^{ts} , compared to its expected value, $E[s_i^{ts}]$; and the observed average strength, $\langle s^{ts} \rangle$, compared with its expected value, $E[\langle s^{ts} \rangle]$.

3.4.1. Mathematical analysis

The analysis of the ETFM is trivial, since, by definition, the maximum likelihood in equation (9) returns optimal activity values, $a_i^*(t)$ for $i = 1, \dots, N$, for an optimal representation of observed values. Hence, we expect that the ETFM provides the most accurate estimation of the backbone network. We focus on both global and local quantities, which we report here, and successively compare to the results of the EADM. In the ETFM, the expected individual strength of node i is computed by summing over the entire time-window and over all possible connections that node i may have

$$E[\bar{s}_i] = \sum_{j=1, j \neq i}^N \sum_{t=1}^T a_i^*(t) a_j^*(t). \quad (15)$$

In equation (15), we use the product of activities to be consistent with the probability that two nodes randomly connect, given in equation (2). The expected average strength is the simple average of equation (15) among all possible nodes in the network and reads

$$E[\langle \bar{s} \rangle] = \frac{1}{N} \sum_{i,j=1}^N \sum_{t=1}^T a_i^*(t) a_j^*(t). \quad (16)$$

The analysis of the EADM is more challenging. The node strength is a key quantity to compute the individual activities, in equation (3), to estimate the link weights, in equation (7), and to determine whether links are explained by our null model or they belong to the backbone network, in equation (8). From equation (7), we compute the expected individual strength for node i in the whole observation window by summing over all the possible connections that node i can make, excluding self-loops that are not allowed in the model. This reads

$$E [\bar{s}_i] = \sum_{j=1, j \neq i}^N \sum_{\Delta=1}^I \frac{s_i^{ts}(\Delta) s_j^{ts}(\Delta)}{2W^{ts}(\Delta)}. \tag{17}$$

By changing the order of the two summands and using the definition of the total number of links in the Δ th interval, $W^{ts}(\Delta)$ in equation (4), we obtain

$$E [\bar{s}_i] = \sum_{\Delta=1}^I s_i^{ts}(\Delta) \frac{\sum_{j=1, j \neq i}^N s_j^{ts}(\Delta)}{\sum_{k=1}^N s_k^{ts}(\Delta)}. \tag{18}$$

By summing and subtracting $s_i^{ts}(\Delta)$ from the numerator, we get

$$E [\bar{s}_i] = \sum_{\Delta=1}^I s_i^{ts}(\Delta) \left[1 - \frac{s_i^{ts}(\Delta)}{\sum_{k=1}^N s_k^{ts}(\Delta)} \right], \tag{19}$$

which can be rewritten as

$$E [\bar{s}_i] = \bar{s}_i^{ts} - \sum_{\Delta=1}^I \frac{[s_i^{ts}(\Delta)]^2}{\sum_{k=1}^N s_k^{ts}(\Delta)}, \tag{20}$$

where $E [\bar{s}_i]$ and \bar{s}_i^{ts} represent the expected and observed individual strengths, respectively. We observe that the EADM underestimates the observed strength, \bar{s}_i^{ts} , by a quantity that only depends on the self-loops. The contribution of self-loops is inversely proportional to the size of the network and scales as $1/N$. This result generalizes the proof in [39], based on the static configuration model, to temporal networks.

As a consequence, macroscopic quantities, such as the expected average strength, should become more accurate in the EADM as the size of the network increases. By averaging equation (20) for all nodes in the network we obtain

$$E [\langle \bar{s} \rangle] = \langle \bar{s}^{ts} \rangle - \frac{1}{N} \sum_{i=1}^N \sum_{\Delta=1}^I \frac{[s_i^{ts}(\Delta)]^2}{\sum_{k=1}^N s_k^{ts}(\Delta)}, \tag{21}$$

which scales as $1/N$. In other words, since the EADM error comes from self-loops only, the observed quantities, \bar{s}_i^{ts} and $\langle \bar{s} \rangle$, become equivalent to expected quantities if self-loops are included in both equation (20) and equation (21). We indicate these adjusted quantities as

$$E^{sl} [\bar{s}_i] = \bar{s}_i^{ts}, \quad E^{sl} [\langle \bar{s} \rangle] = \langle \bar{s} \rangle, \tag{22}$$

where the superscript ‘sl’ indicates that these sums include self-loops into equations (20) and (21).

In order to properly assess the error of the EADM, we consider two metrics: the relative error and the normalized root mean squared error. The relative error determines whether our estimators are unbiased, while the normalized root mean squared error is needed to ensure that the variance of our estimators is bounded. The theoretical relative error and normalized root mean squared error of the EADM are obtained by properly combining equations (20)–(22)

$$\begin{aligned} RE &= \left| 1 - \frac{E[\langle \bar{s} \rangle]}{E^{sl}[\langle \bar{s} \rangle]} \right|, \\ NRMSE &= \sqrt{\frac{\sum_{i=1}^N [E^{sl}[\bar{s}_i] - E[\bar{s}_i]]^2}{NE^{sl}[\bar{s}_i]^2}}. \end{aligned} \tag{23}$$

3.4.2. Numerics

The mathematical analysis above suggests that the ETFM error is minimized, while the error of the EADM should decay as the inverse of the size of the network, $1/N$, and coincide with predictions from equation (23).

The numerical relative error and normalized root mean squared error are computed as follows:

$$\begin{aligned} RE &= \left| 1 - \frac{E[\langle \bar{s} \rangle]}{\langle \bar{s}^{ts} \rangle} \right|, \\ NRMSE &= \sqrt{\frac{\sum_{i=1}^N [\bar{s}_i^{ts} - E[\bar{s}_i]]^2}{N \langle \bar{s}^{ts} \rangle^2}}, \end{aligned} \tag{24}$$

where we substitute the expected quantities, $E^{sl}[\bar{s}_i]$ and $E^{sl}[\langle \bar{s} \rangle]$, in equation (23), with their observed values, \bar{s}_i^{ts} and $\langle \bar{s}^{ts} \rangle$. The remaining expected quantities in equation (24) can be either equations (15) and (16) for the ETFM numerics, or equations (20) and (21) for the EADM numerics.

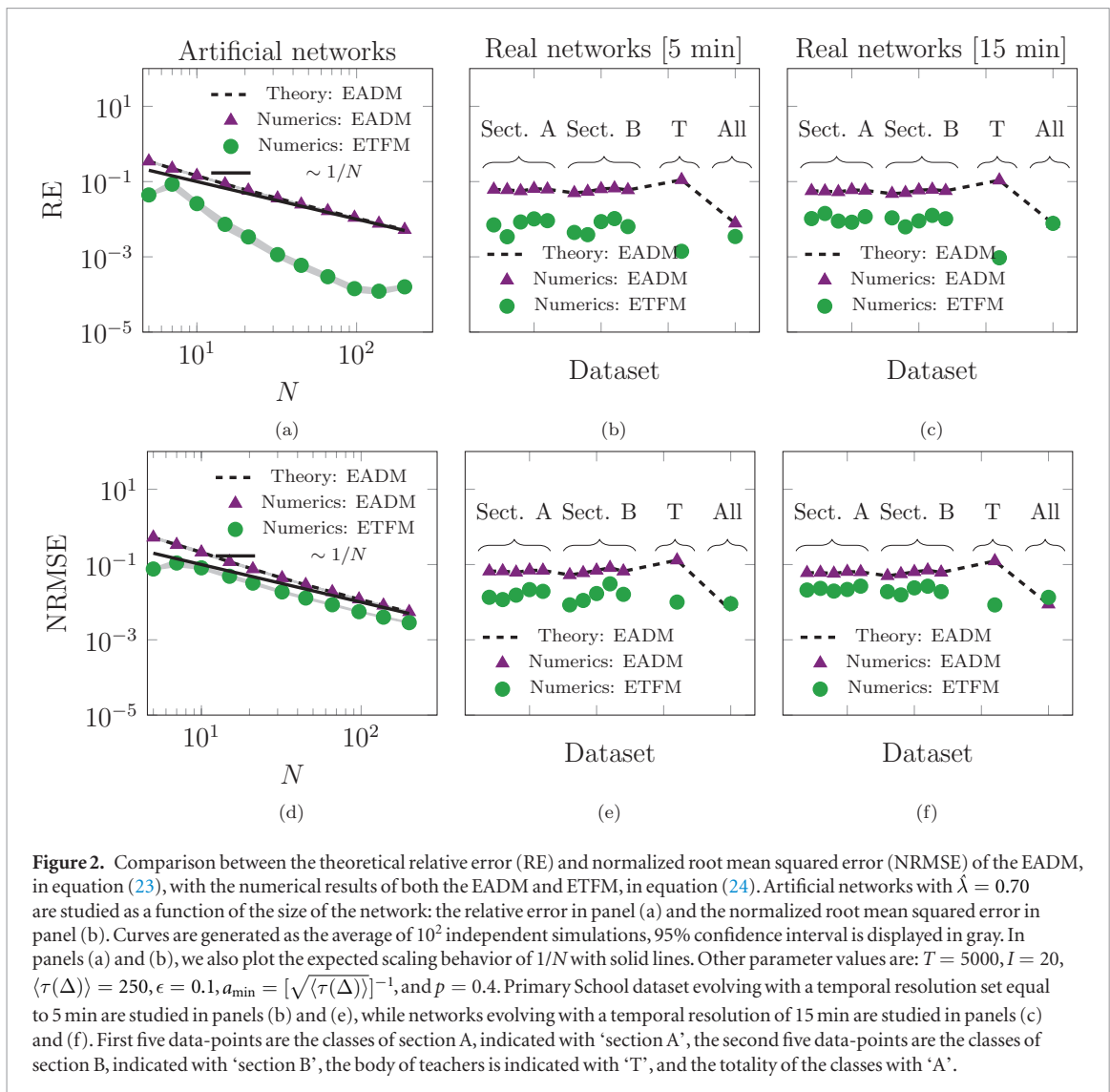


Figure 2. Comparison between the theoretical relative error (RE) and normalized root mean squared error (NRMSE) of the EADM, in equation (23), with the numerical results of both the EADM and ETFM, in equation (24). Artificial networks with $\hat{\lambda} = 0.70$ are studied as a function of the size of the network: the relative error in panel (a) and the normalized root mean squared error in panel (b). Curves are generated as the average of 10^2 independent simulations, 95% confidence interval is displayed in gray. In panels (a) and (b), we also plot the expected scaling behavior of $1/N$ with solid lines. Other parameter values are: $T = 5000, I = 20, \langle \tau(\Delta) \rangle = 250, \epsilon = 0.1, a_{\min} = [\sqrt{\langle \tau(\Delta) \rangle}]^{-1}$, and $p = 0.4$. Primary School dataset evolving with a temporal resolution set equal to 5 min are studied in panels (b) and (e), while networks evolving with a temporal resolution of 15 min are studied in panels (c) and (f). First five data-points are the classes of section A, indicated with 'section A', the second five data-points are the classes of section B, indicated with 'section B', the body of teachers is indicated with 'T', and the totality of the classes with 'A'.

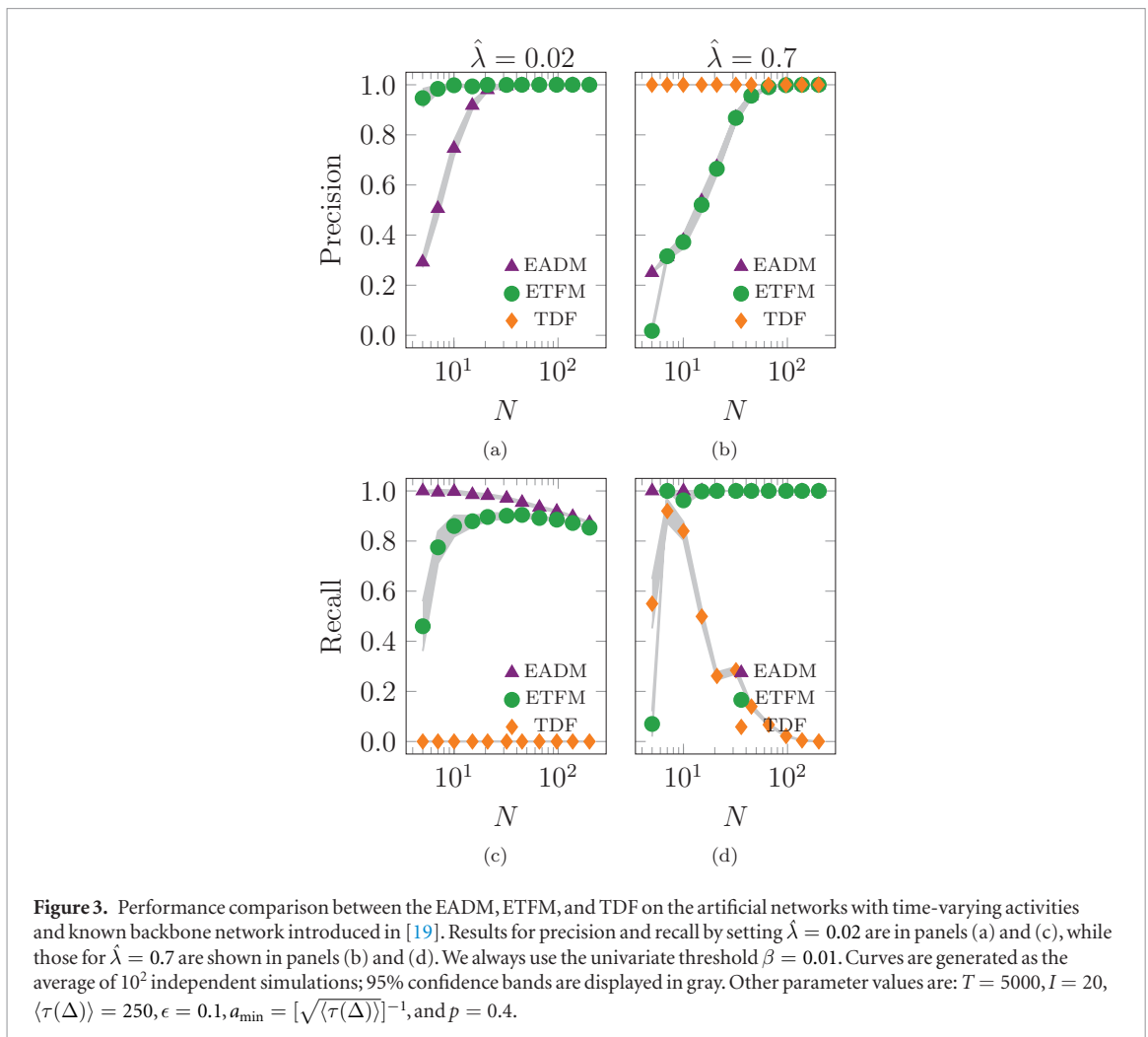
In our numerics, shown in figure 2, we consider artificial networks having a homogeneous backbone with $\hat{\lambda} = 0.7$, as introduced in appendix A.1, and real systems, as presented in table 3. The cases of artificial networks having either a homogeneous backbone with $\hat{\lambda} = 0.02$ or heterogeneous backbone are explored in appendix A.2. We determine that, for both artificial and real systems, our theoretical arguments hold. Specifically, we find that the theoretical error of the EADM comes from self-loops only and it scales as the inverse of the network size. Further, we show that the ETFM improves the EADM estimation, lowering the relative and normalized root mean squared errors. However, the ETFM estimation is challenged by the excessive preponderance of backbone network, in terms of the normalized root mean squared error. This is due to the fact that the maximum likelihood approach of the ETFM in equation (9) estimates the activity values without considering the presence of the backbone network.

Interestingly, we find that in both artificial and real systems, the relative EADM errors are about 1% when networks reach one hundred of nodes. This amount of error is acceptable for many practical applications. On the contrary, when the artificial and real systems are smaller, the relative EADM error may reach up to 10%, challenging its usability for many practical purposes.

3.5. Performance comparison on artificial networks

Going further in the comparison between the EADM, ETFM, and TDF, we examine the performance of the three approaches in detecting backbone networks as a function of the size of the network. We use the artificial network introduced in our prior work [19]. We analyze two different cases, resembling two realistic situations: one in which the backbone network is barely present, where we set $\hat{\lambda} = 0.02$, and another in which the backbone network appears almost at every time step, where we set $\hat{\lambda} = 0.7$. In this analysis, we set the univariate threshold β equal to 0.01 [15, 19], and then correct it according to the Bonferroni correction as detailed in section 2.4, to obtain the β' used to filter out links explained by the null models.

The performance of the three approaches is determined through the analysis of the algorithms' precision and recall [40]. The former corresponds to the ratio between the number of true positives (all the links discovered



that truly belong to the backbone network) divided by the number of all the detected links (true and false positives). The latter metric corresponds to the ratio between the true positives divided by the sum of true positives and false negatives (all the links in the irreducible backbone).

As shown in figures 3(a) and (c), when the backbone network is barely present, the TDF does not detect any link (recall is zero and precision is not defined). The ETFM outperforms the EADM, while it falls behind the EADM in terms of recall. In other words, the EADM detects more links than the ETFM and, some of them are false positive that belong to the known backbone network. When the size of the network reaches about one hundred nodes, the EADM and ETFM are equivalent. When the backbone network is present almost in every time step, the TDF has the highest precision among the three methodologies, but its recall goes rapidly to zero when the size of the network increases, as shown in figures 3(b) and (d). In this case, the EADM and ETFM perform similarly in terms of precision and recall, and they are accurate only when the size of the system reaches about one hundred nodes. If links in the backbone network appear in almost all time steps, they significantly modify the activity estimation in equations (3) and (9), which lowers the performance of the EADM and ETFM.

The analysis of artificial networks suggests that the TDF does not detect false positive, while often missing major parts of the backbone network. Failing to detect most of the backbone network may be due to two main aspects: (i) the use of the Bonferroni correction [37], which increases the methodology's precision and lowers its recall; and (ii) the heterogeneous distribution of activity values, whereby in appendix A.3 we determine an improvement in the number of true positives when dealing with a homogeneous activity distribution. The ETFM offers adequate performance in many situations, scoring poorly only when links in the backbone network are present at almost every time step. The EADM becomes equivalent to the ETFM when the size of the network reaches about one hundred nodes. Since the EADM is faster than the ETFM, it should be preferred for the study of large networks.

3.6. Network backbone discovery in real systems

Detecting the backbone in real systems yields more challenges than the study of artificial networks. The main limitation is that the backbone network is not known a-priori and performance can only be assessed through comparison. A possible way to circumvent such a limitation is to observe whether or not results from real systems are coherent with findings on artificial networks.

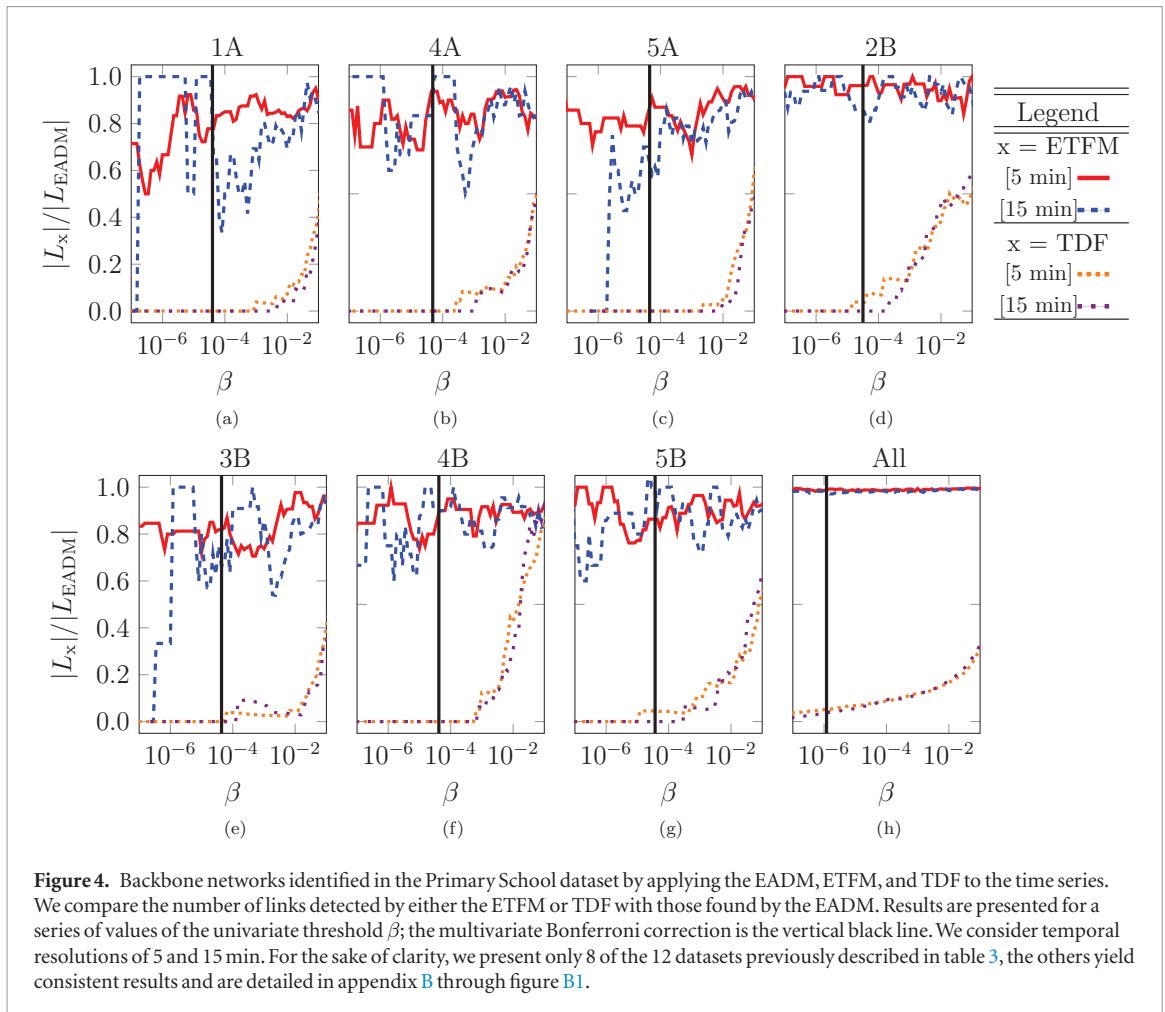


Figure 4. Backbone networks identified in the Primary School dataset by applying the EADM, ETFM, and TDF to the time series. We compare the number of links detected by either the ETFM or TDF with those found by the EADM. Results are presented for a series of values of the univariate threshold β ; the multivariate Bonferroni correction is the vertical black line. We consider temporal resolutions of 5 and 15 min. For the sake of clarity, we present only 8 of the 12 datasets previously described in table 3, the others yield consistent results and are detailed in appendix B through figure B1.

According to our results on artificial networks, the TDF has the most restrictive null hypothesis, which leads to the least number of links detected in the backbone network. The ETFM finds smaller backbone networks than the EADM when the networks under examination have only tens of nodes, while they tend to become equivalent for larger systems. To confirm this claim in real systems, we compare the set of irreducible links detected by the three methodologies, using the EADM as reference. We compare the sets of links discovered by the TDF, L_{TDF} , and by the ETFM, L_{ETFM} , with that of the EADM, L_{EADM} . The dimension of a set is indicated with $|\cdot|$. We use two metrics: the ratio of the number of links $|L_x|/|L_{\text{EADM}}|$ and the Overlap index [41], defined as

$$\text{Overlap}(L_{\text{EADM}}, L_x) = \frac{|L_{\text{EADM}} \cup L_x|}{\min(|L_{\text{EADM}}|, |L_x|)}, \quad (25)$$

where $x = \{\text{ETFM}, \text{TDF}\}$.

The ratio $|L_x|/|L_{\text{EADM}}|$ is useful to quantify whether the EADM detects more links than the ETFM and TDF. The Overlap coefficient in equation (25) measures the fraction of shared links between the two methodologies, thereby helping ascertain whether the ETFM and TDF find a subset of the links detected by the EADM. In order to properly compare the EADM with the ETFM, we consider a series of values for the univariate threshold β and indicate with a vertical black line the multivariate threshold β' , properly modified according to the Bonferroni correction.

In figure 4, we show that the TDF detects less links than the ETFM, which, in turn, finds less links than the EADM, as expected. The fact that the TDF detects the least amount of links is due to its more restrictive null hypothesis, which ensures that no false positives are identified, but it misses many links that belong to the backbone network. ETFM and EADM results are affected by the network size; but when the entire school (formed by a few hundreds of nodes) is analyzed, the two methodologies become equivalent. This result is in line with the analysis of artificial networks.

In figure 5, we show that, for most of the real datasets studied herein, the Overlap coefficient is close to one. This result implies that the ETFM and TDF share the same core of the backbone networks with the EADM. The few cases in which the Overlap coefficient is lower than one pertain to the comparison between the EADM and TDF, due to the difference in their null hypotheses.

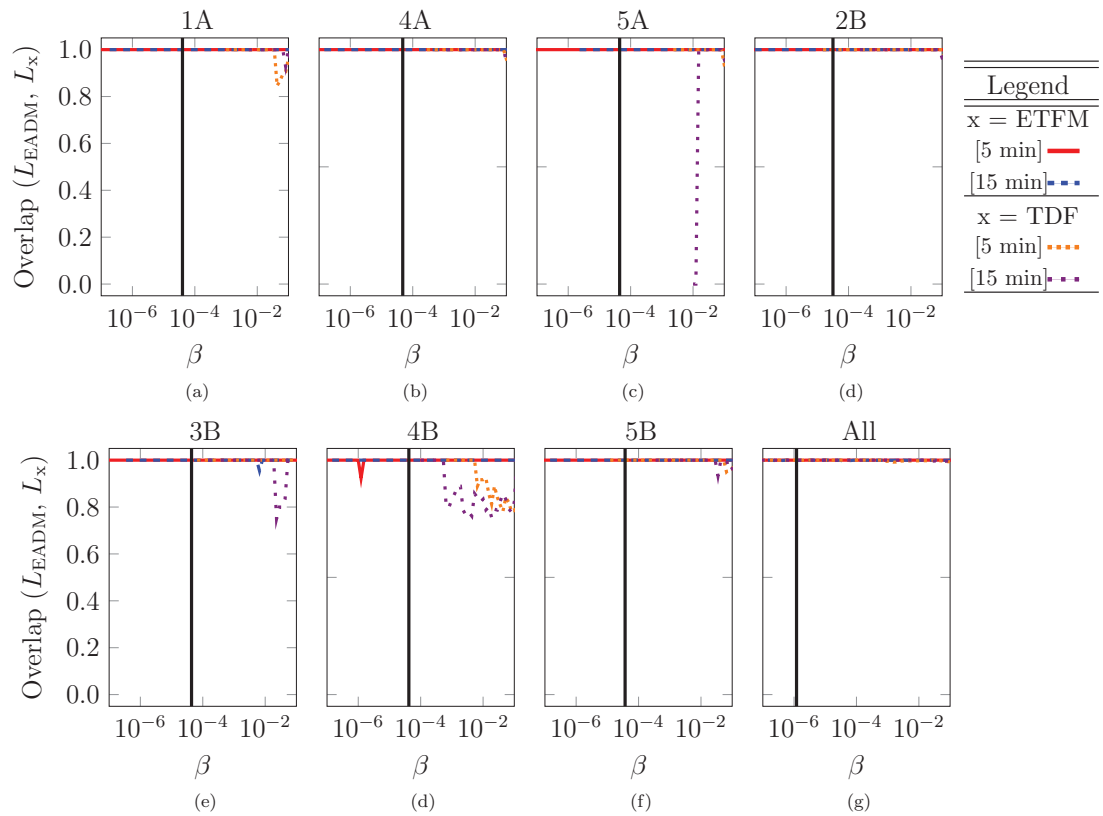


Figure 5. Fraction of irreducible links found by the EADM that are discovered by the ETFM and TDF in the Primary School dataset. Results are presented for a series of values by the univariate threshold β ; the multivariate Bonferroni correction is the vertical black line. We consider temporal resolutions of 5 and 15 min. For the sake of clarity, we present here only 8 of the 12 datasets previously described in table 3, the others yield consistent results and are detailed in appendix B through figure B2.

Table 4. Range of validity of each methodology (top row) and their main advantage and disadvantage (bottom row).

		Evolving ac- tivity driven network	Evolving temporal fit- ness model	Temporal disparity filter
Network size	Small		✓	✓
	Large	✓		✓
Advantage (Disadvantage)	Detection of many irreducible links (Some reducible links inaccurately ascribed to the backbone)	✓	✓	
	No reducible links is inaccurately identified as part of the backbone (Many of the irreducible links are not detected)			✓

By combining the results of figures 4 and 5 we draw two main conclusions. First, the TDF has the most restrictive null hypothesis and detects the smallest backbone network. Second, the ETFM finds a subset of the links detected by the EADM only when a classroom (composed of only tens of nodes) is studied, while it identifies the same backbone network when the entire school (composed of only hundreds of nodes) is analyzed. Since the EADM is faster than the ETFM, it should be used to discover backbone networks in larger systems.

In table 4, we synoptically present the take-home message of the comparisons conducted throughout this paper. While, the EADM should be used for large networks and ETFM for small networks, the TDF can be applied to any network. The EADM and ETFM can be successful in identifying many irreducible links, but this may come at the cost of some of the reducible links being inaccurately assigned to the backbone. The TDF allows to filter out almost all the reducible links, but it tends to miss most of the irreducible links.

4. Discussion

In this paper, we introduce two new approaches that account for time-varying individual properties: an extension of the temporal fitness model [5], the evolving temporal fitness model (ETFm), and an extension of the disparity filter [6], the temporal disparity filter (TDF). We critically compare the performance of these approaches with

the evolving activity driven model (EADM) [19], an approach we recently proposed to account for time-varying individual properties. Our analysis focuses on the network size, toward an improved understanding of finite-size effects in backbone detection.

Our study involves the analysis of both artificial and real systems. From the analysis of artificial networks, we find that the TDF proposes the most restrictive null hypothesis and the discovered backbone networks contain the smallest amount of links among the three methodologies. This approach ensures that no false positives (links incorrectly identified as part of the backbone network) are detected, but most of the times it misses links that truly belong to the backbone network. The ETFM offers a trade-off between the number of detected links and the likelihood that they belong to the backbone network. From our study, we determine that false positives are detected only when links in the backbone network are present in almost any time step and the size of the network is of the order of ten nodes. When the network contains hundreds of nodes, the number of false positives is reduced to zero. Interestingly, for networks with at least one hundred nodes, the ETFM has the same performance of the EADM. Given that the EADM has a smaller computational burden, it should be preferred to the ETFM in this case. Comparing the three approaches, we observe that the TDF has the most restrictive null hypothesis, and it detects the least amount of irreducible links. The ETFM identifies an intermediate number of irreducible links, which becomes equivalent to the EADM when the networks have at least one hundred nodes.

Our results on real systems confirm the insight gathered from artificial networks. When detecting the backbone network for the Primary School and Workplace datasets [30], we determine that the EADM, ETFM, and TDF detect a different number of irreducible links. In most cases, the TDF finds a subset of the links of the ETFM, which detects a subset of the links of the EADM, thereby validating the same core of the backbone network.

The size of the network plays a critical role on the proper determination of the backbone network. On the one hand, large systems can only be studied by the EADM and TDF. On the other hand, the ETFM requires a large computational power, but it offers the best trade-off between the number of detected links and the fraction of them that actually belongs to the backbone network.

In summary, we suggest to use the TDF if the objective of the study is to minimize the number of false positives, at the cost of failing to detect many irreducible links. Employing the Bonferroni correction, the TDF could detect most of the irreducible links only if the distribution of activity values is homogeneous, a scenario that is in contrast with several empirical observations [42–44]. Instead, if finding as many irreducible links as possible is the main goal, then the ETFM is preferable for small networks. When networks count several hundreds, thousands, or even more nodes, the ETFM and EADM become equivalent; thus, we propose to use the latter, since it is faster.

These methodologies are not free of limitations and our work identifies possible choices depending on the problem considered. For instance, an interesting line of research may entail improving the estimate of the interval partition, which now is common to all the nodes in the network. It is possible that each node in the network evolves according to its own pattern, which would result in considering different interval partitions, one for each node in the network.

The TDF can be applied to systems of any size, if the main objective of the study is to minimize the number of false positives. Instead, the other two approaches are highly affected by the size of the network. Real examples of human social systems in which the ETFM should be used are: classrooms and small departments (as done in our work), small organizations, bars, or small conferences. Instead, real examples of social systems that would benefit of the EADM are: entire schools or work environments (as in our work), entire universities, large companies, or large conferences. These approaches do not have to be limited to social systems but they can be applied to animal groups [45], the world wide web [46], financial markets [47], brain circuits [48], and many others.

Acknowledgments

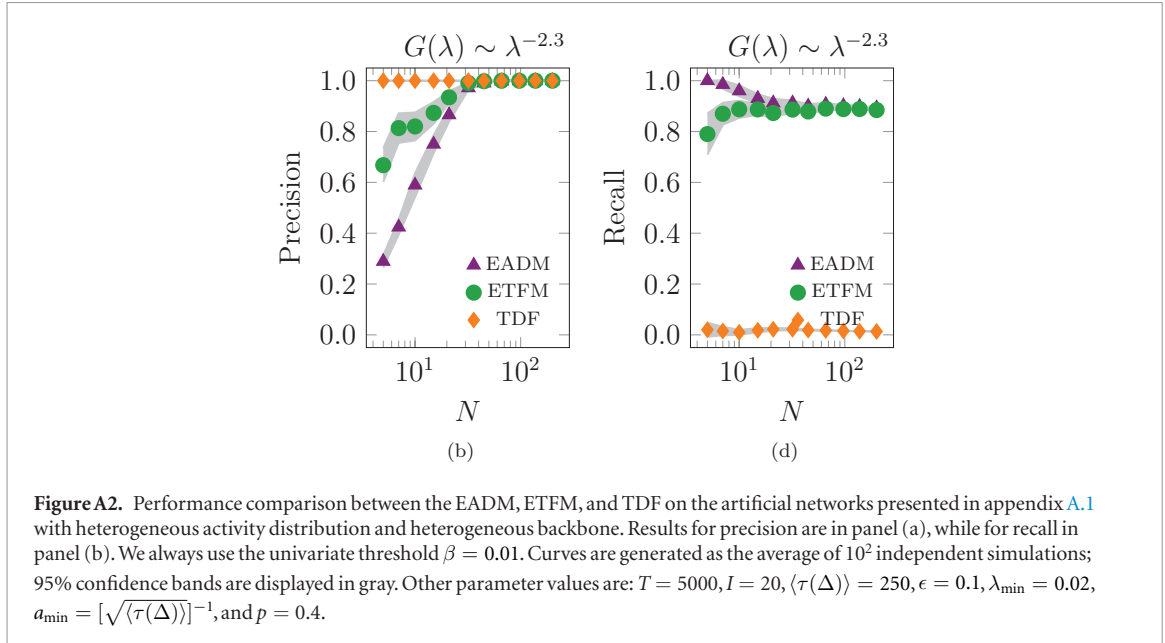
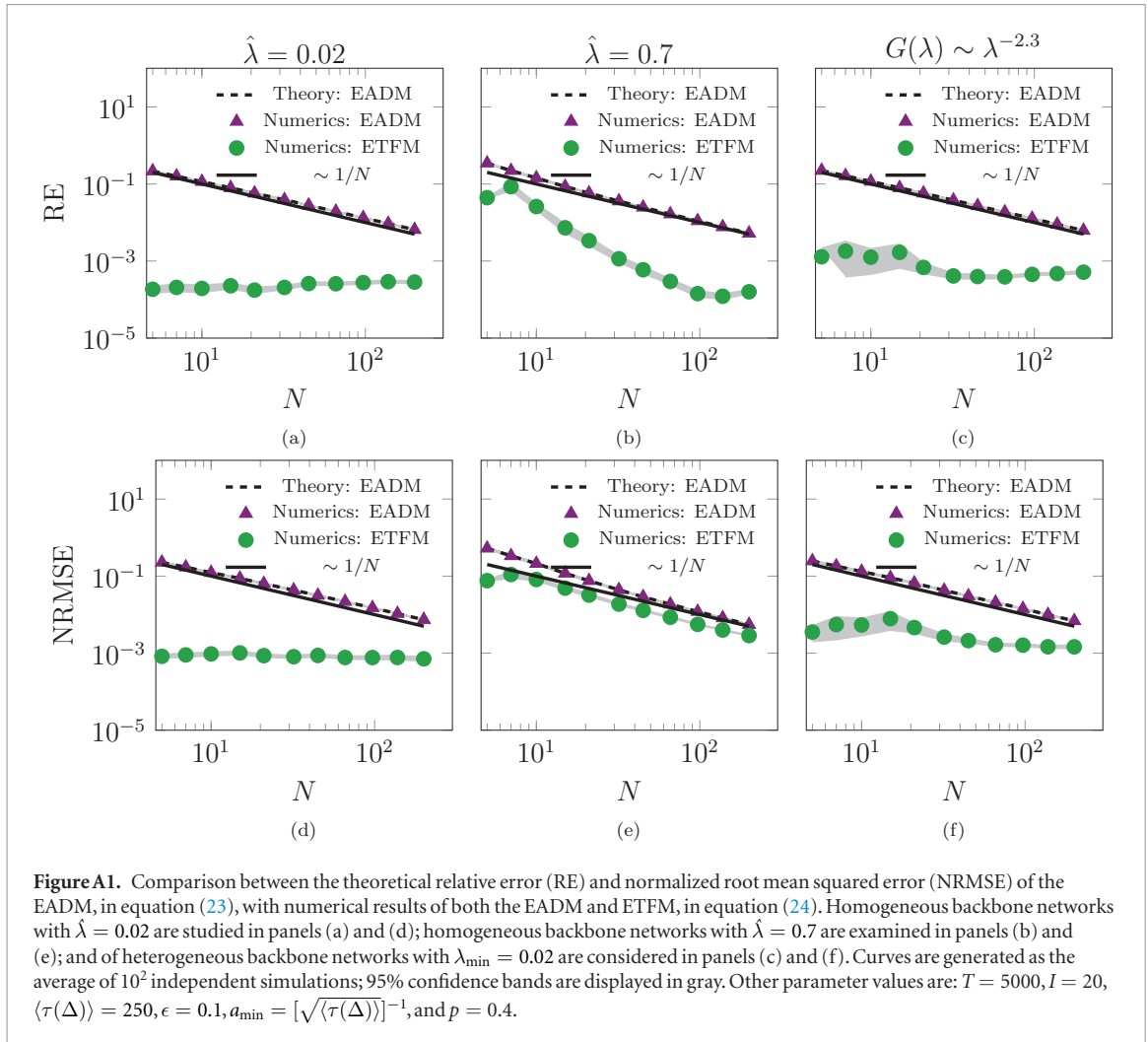
The authors acknowledges financial support from the National Science Foundation under Grant No. CMMI-1561134. AR acknowledges financial support from Compagnia di San Paolo, Italy, and the Italian Ministry of Foreign Affairs and International Cooperation, within the project ‘Mac2Mic’, ‘Macro to Micro: uncovering the hidden mechanisms driving network dynamics’.

Code availability

Python 2.7 codes are freely available at [49].

Data availability statement

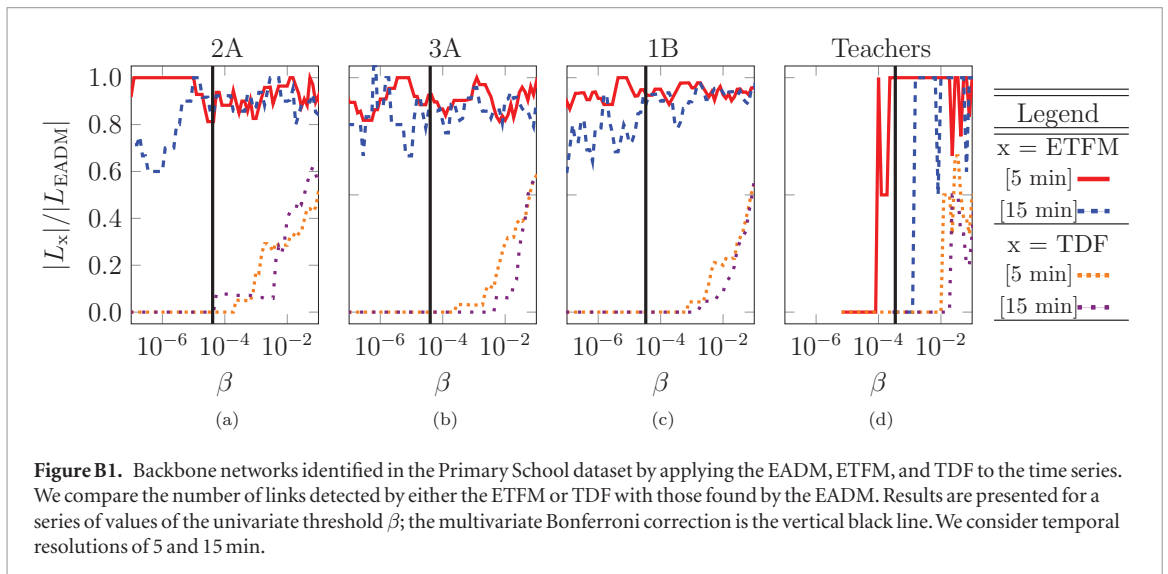
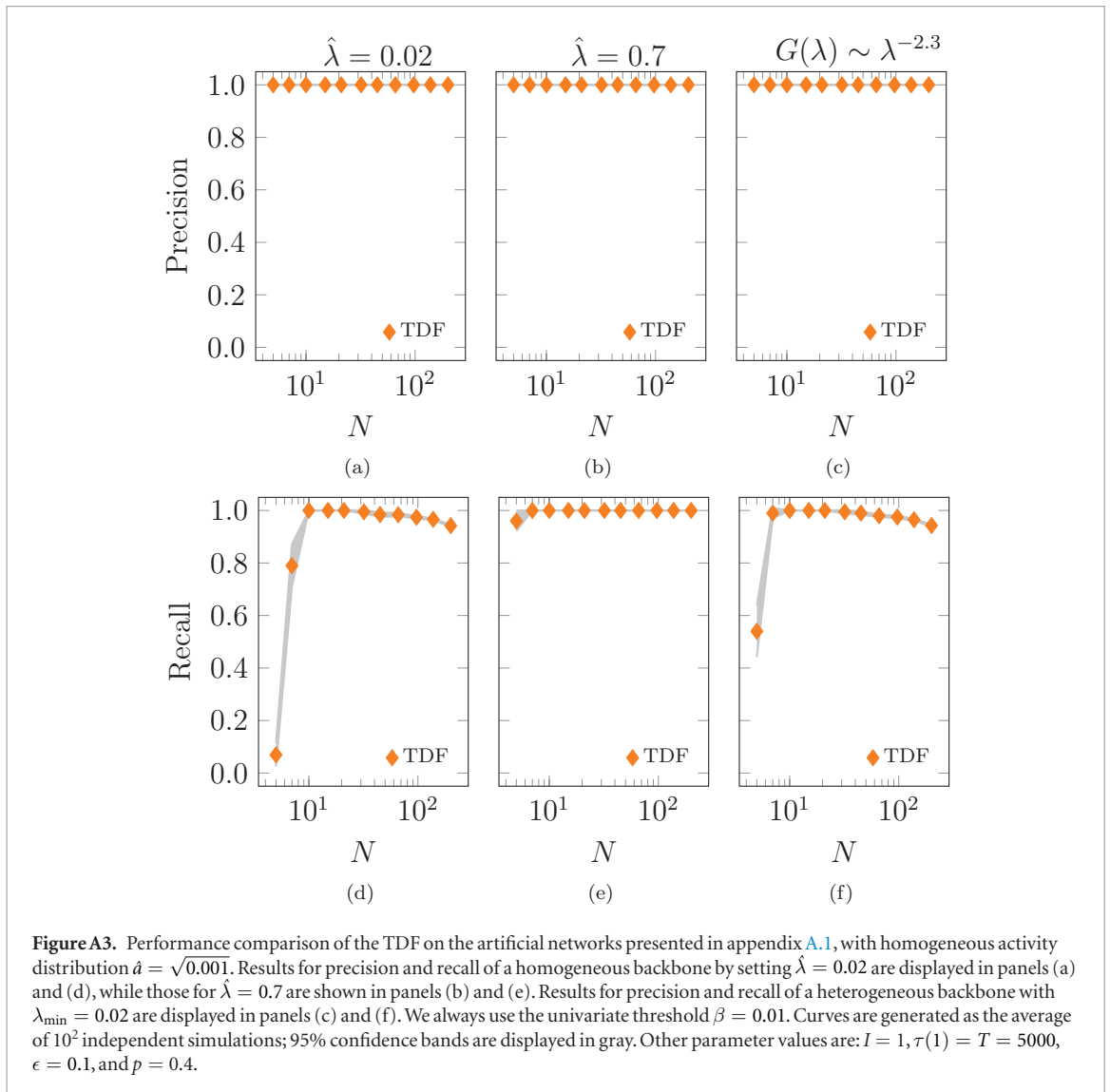
The data that support the findings of this study are openly available at [50] (Primary School dataset) and [51] (Workplace dataset).



Appendix A. Artificial networks

A.1. Generation of artificial temporal networks with arbitrary distributions of activity values and links in the backbone

We present here an extension of the generative algorithm proposed in our previous work [19]. Specifically, we introduce a mechanism to create a temporal network with known backbone network and arbitrary distribution



of reducible and irreducible links. The distribution of reducible links is regulated by the activity distribution $F(a)$, while the distribution of irreducible links is controlled by the preponderance of the backbone network, $G(\lambda)$. The case of heterogeneous activity distribution $F(a) \sim a^{-2.1}$, with $a \in [a_{\min}, 1]$, and homogeneous backbone network $G(\lambda) \sim \delta(\lambda - \hat{\lambda})$ is already described in our previous work [19]. The procedure is as follows.

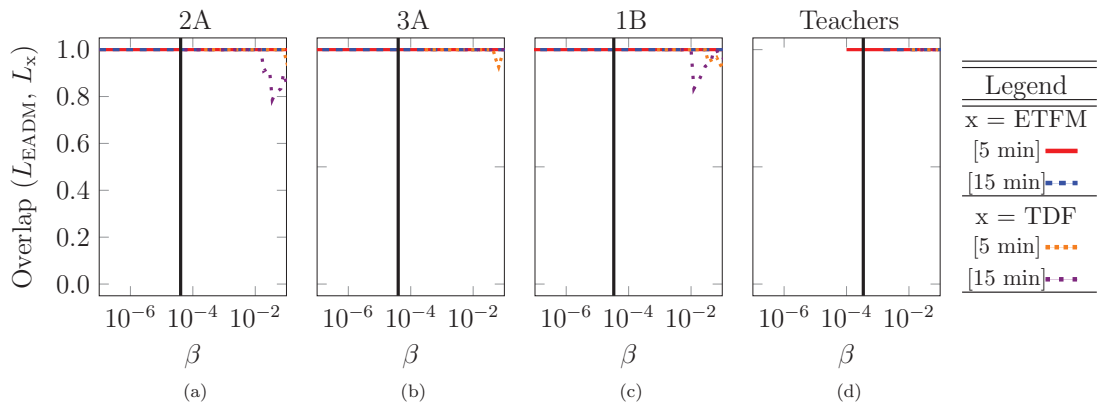


Figure B2. Fraction of irreducible links found by the EADM that are discovered by the ETFM and TDF in the Primary School dataset. Results are presented for a series of values by the univariate threshold β ; the multivariate Bonferroni correction is the vertical black line. We consider temporal resolutions of 5 and 15 min.

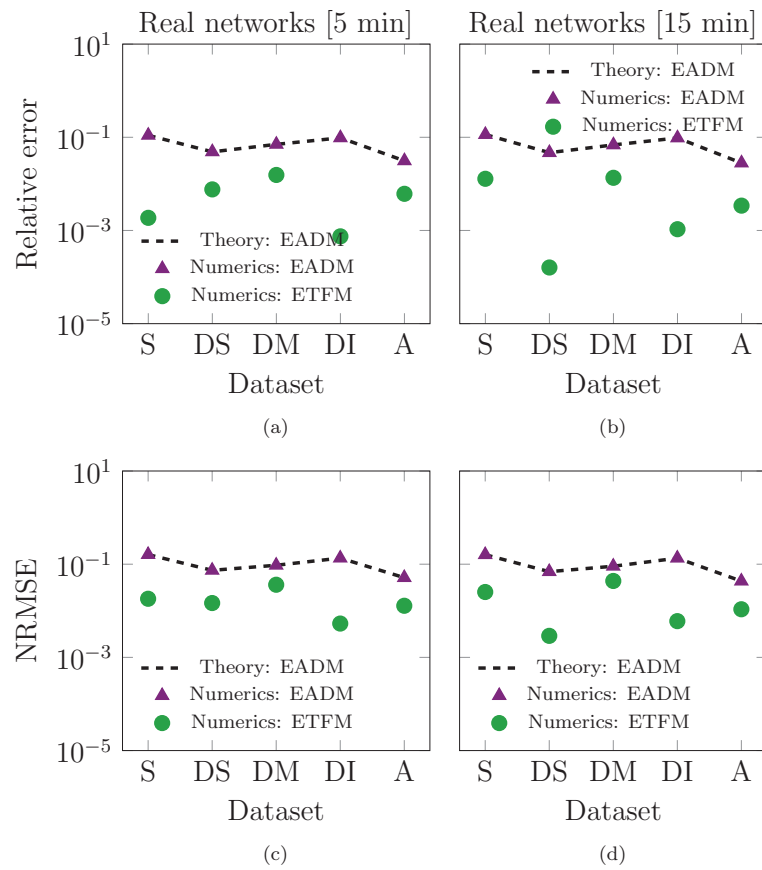
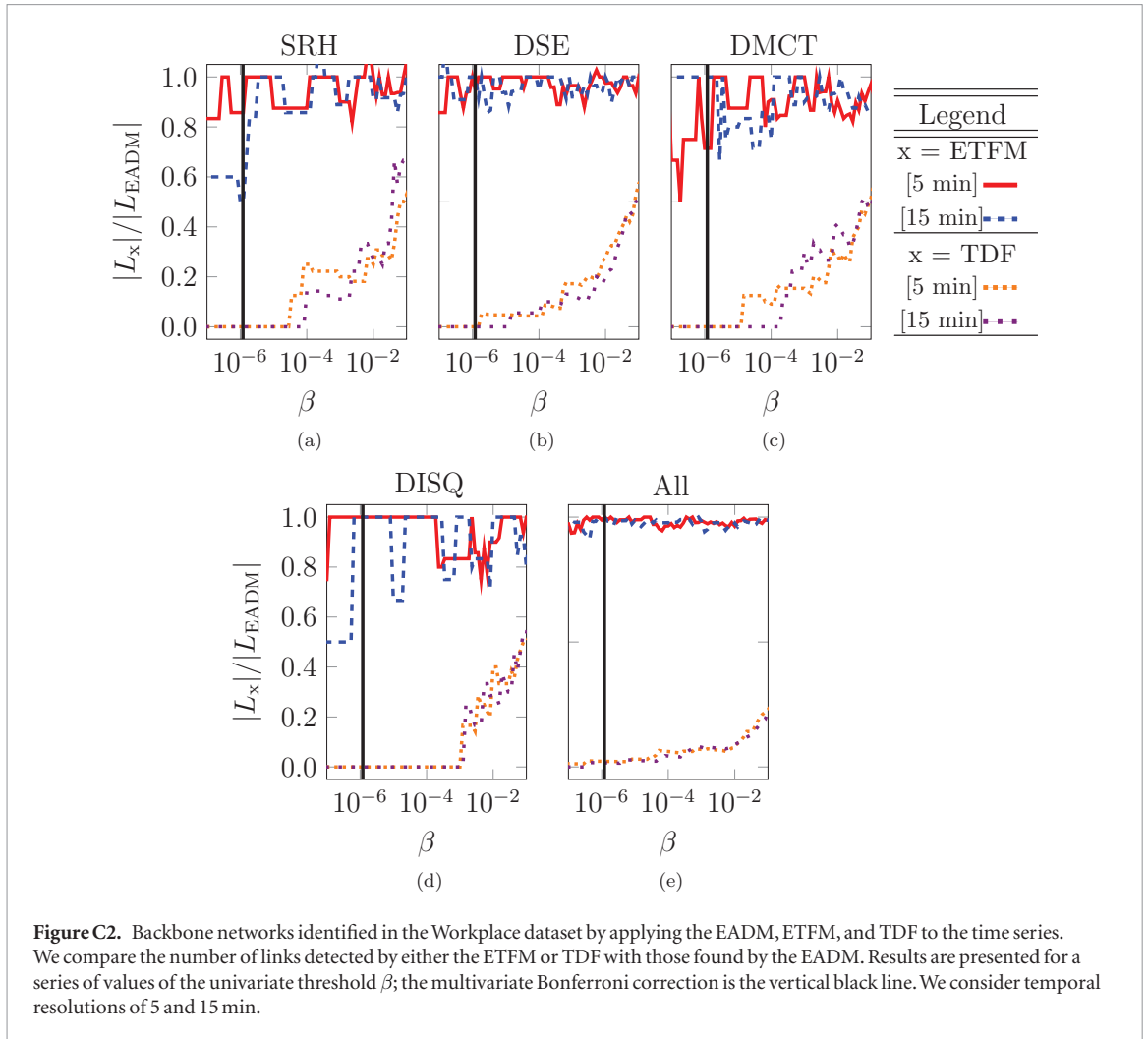


Figure C1. Comparison between the theoretical relative error (RE) and normalized root mean squared error (NRMSE) of the EADM, in equation (23), with the numerical results of both the EADM and ETFM, in equation (24). The Workplace dataset evolving with a temporal resolution set equal to 5 min are studied in panels (a) and (c), while systems evolving with a temporal resolution of 15 min are examined in panels (b) and (d). The ‘S’ label stands for SRH, the ‘DS’ for DSE, the ‘DM’ for DMCT, the ‘DI’ for DISQ, and the ‘A’ for the entire dataset.

- (i) We generate a temporal network unfolding over an observation window of length T , partitioned into I consecutive intervals. We manually set the initial time of the first interval at the beginning of the time series, $t_{in}(1) = 1$, and randomly draw without replacement the $I - 1$ initial time of the remaining intervals in $\{2, \dots, T\}$, which we sort as $t_{in}(2) < t_{in}(3) < \dots < t_{in}(I)$. The generic interval Δ has length $\tau(\Delta)$, and the average length of the intervals is $\langle \tau(\Delta) \rangle = T/I$.
- (ii) Each node in the network has a time-varying, piece-wise constant, individual activity, $a_i(t)$, which is extracted according to the following procedure.



- For the first interval, $\Delta = 1$, we extract N activities, one for each node in the network, from the activity distribution $F(a)$. These values are held constant in the interval.
- When $2 \leq \Delta \leq I$, we may observe correlations between the activities in the present interval and the previous one, that is, between $[t_{\text{in}}(\Delta - 1), t_{\text{in}}(\Delta - 1) + \tau(\Delta - 1) - 1]$ and $[t_{\text{in}}(\Delta), t_{\text{in}}(\Delta) + \tau(\Delta) - 1]$, respectively. Activities vary according to $a_i(t_2) = a_i(t_1)\rho + \gamma(1 - \rho)$, where γ is a random number extracted from $F(a)$, and ρ is an autocorrelation parameter.

(iii) In the whole observation window $[1, T]$, we generate a temporal network where both reducible and irreducible links are present.

- Reducible links are created according to equation (2).
- Once reducible links are formed, we appoint a fraction ϵ of them that appeared at least once in the overall temporal evolution as the irreducible links. For each of these irreducible links ij of the backbone network, we record the time instants where the adjacency matrix is zero and with probability λ_{ij} , we create a temporal link in that time instant, that is, we switch A_{ij} to 1. The parameter λ_{ij} , extracted from $G(\lambda)$, measures the preponderance of the link during the whole observation window, such that larger values of λ_{ij} favor the association between the link and the backbone network.

A.2. Results for artificial temporal networks with heterogeneous activity distribution

As in the main text, we consider the case of heterogeneous activity distribution $F(a) \sim a^{-2.1}$ with $a_{\text{min}} = [\sqrt{\tau(\Delta)}]^{-1}$. The backbone networks considered in figure A1 are three: (i) a homogeneous backbone with a low value of λ , $G(\lambda) \sim \delta(\lambda - 0.02)$ in panels (a) and (d); (ii) a homogeneous backbone with a high

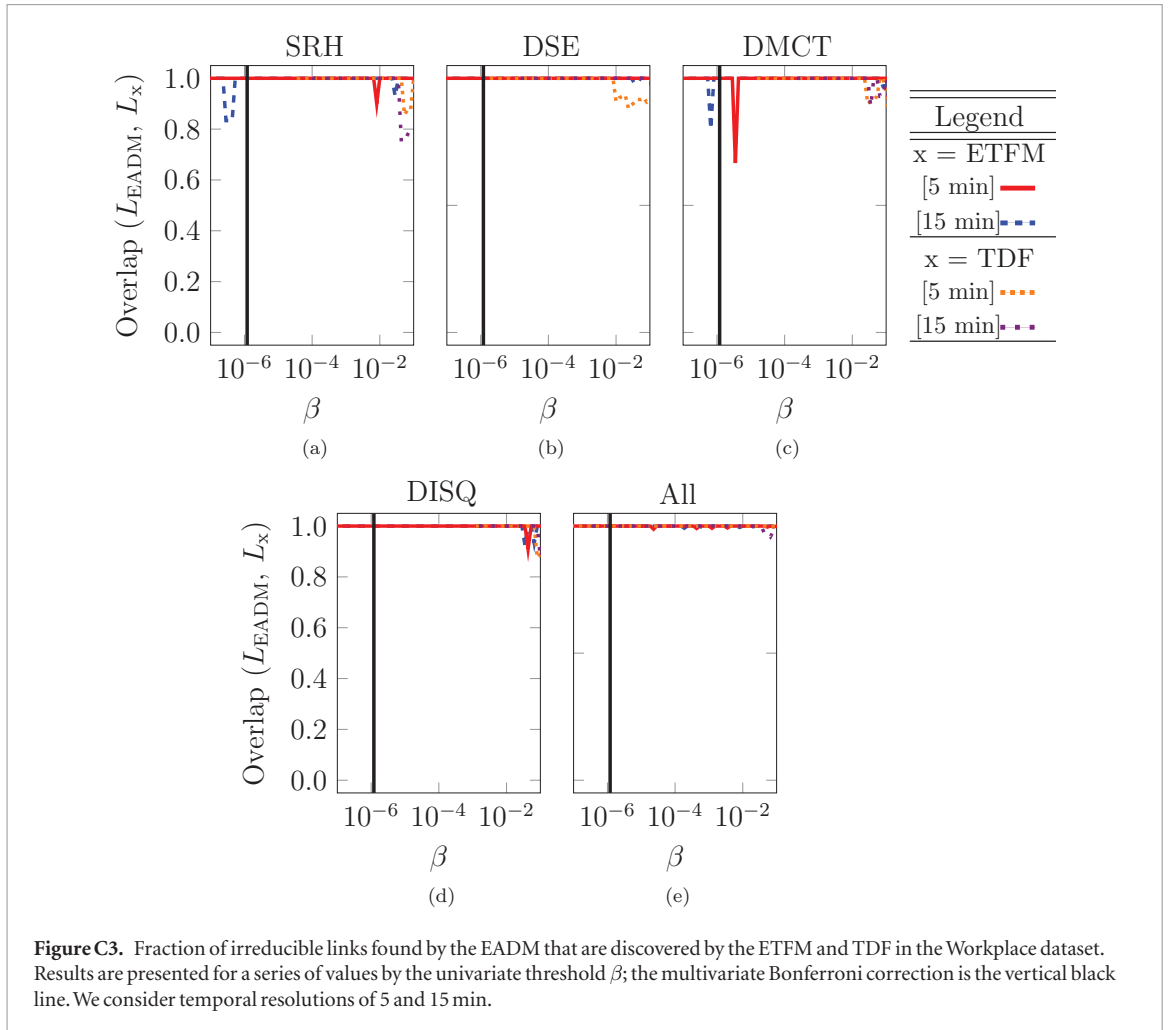


Table C1. Data summary of the Workplace dataset. Similar to the Primary school dataset in table 3, we consider two temporal resolutions (in brackets). All multiple links within a single time step are removed and the '# temporal links' column reports the retained links in the temporal network. The '# static links' column indicates the number of aggregated links. We indicate the various departments with capital letters, while the 'All' label represent the entire Workplace. We do not consider the SFLE department, since only 4 nodes are present. We always remove the time interval between the last connection of a day and the first of the following day.

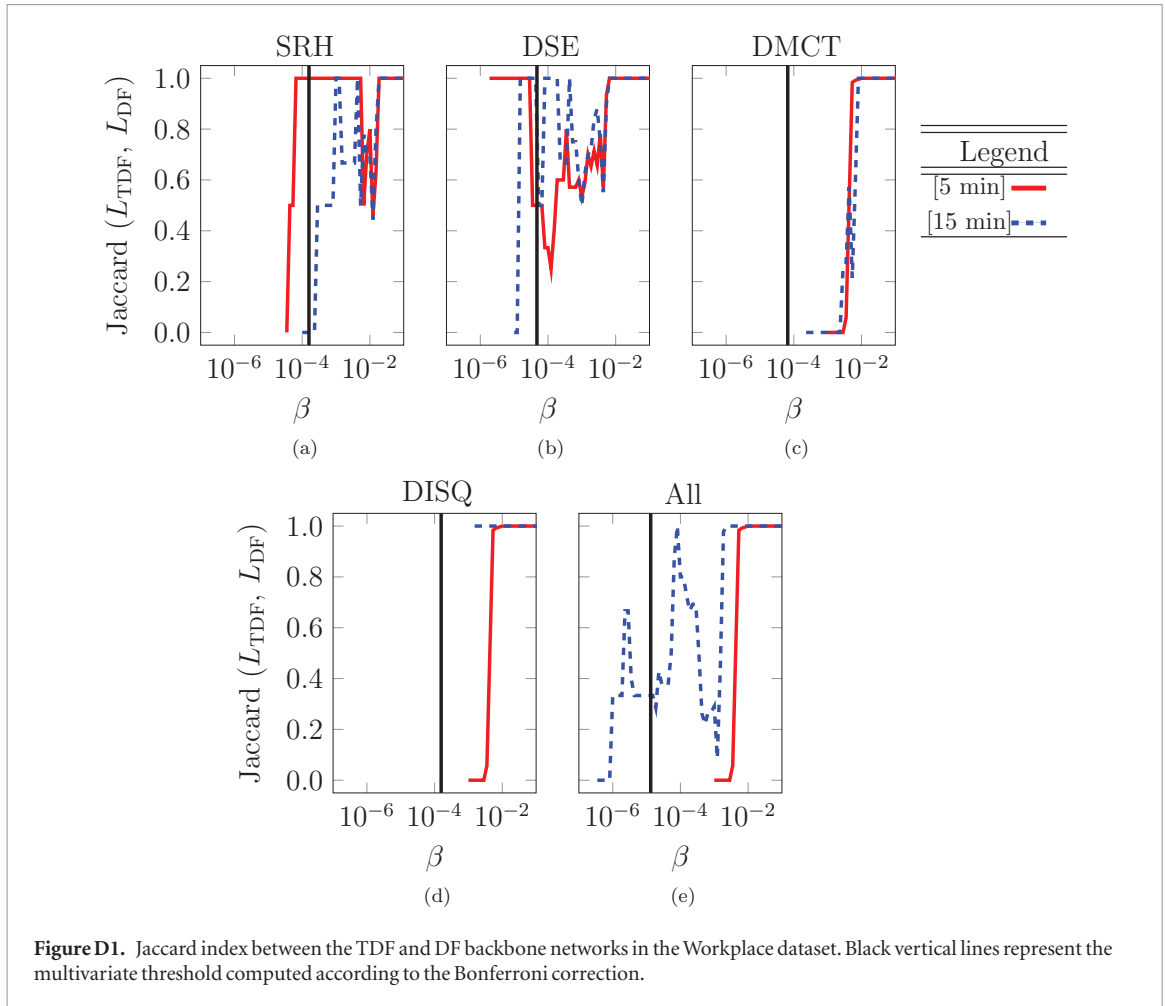
Data	# nodes	# temporal links [5 min]—[15 min]	# static links
SRH	13	856—614	63
SFLE	4	9—8	2
DSE	34	1007—787	212
DMCT	26	561—453	151
DISQ	15	418—327	64
All	92	3538—2737	755

value of λ , $G(\lambda) \sim \delta(\lambda - 0.70)$ in panels (b) and (e); and (iii) a heterogeneous backbone $G(\lambda) \sim \lambda^{-2.3}$ and $\lambda \in [0.02, 1]$ in panels (c) and (f). We find that the ETFM provides better estimations than the EADM, whose error scales as the inverse of the size of network. Further, the ETFM is challenged by the excessive preponderance of the backbone network, whereby $\hat{\lambda} = 0.70$ yields a normalized root mean squared error that is comparable to the one of the EADM.

We corroborate the results in the main text regarding precision and recall of the EADM, ETFM, and TDF, by analyzing the case of a heterogeneous backbone network with $G(\lambda) \sim \lambda^{-2.3}$ and $\lambda \in [0.02, 1]$, in figure A2. For small systems, the highest precision is given by the TDF, then by the ETFM, and finally by the EADM; with respect to recall, we find the opposite scenario. For large systems, all methodologies yield the same precision; the recalls of EADM and ETFM are similar, and they are both much higher than the TDF.

A.3. Results for artificial temporal networks with homogeneous activity distribution

To shed some light into the reasons for the small recall in the TDF, we consider the case a homogeneous activity distribution, $F(a) \sim \delta(a - \sqrt{0.001})$. Activities are time invariant as we consider a single interval, $I = 1$, with



length equal to the whole observation window, that is, $\tau(1) = T = 5000$. The backbone networks considered in figure A3 are three: (i) a homogeneous backbone with a low value of λ , $G(\lambda) \sim \delta(\lambda - 0.02)$ in panels (a) and (d); (ii) a homogeneous backbone with a high value of λ , $G(\lambda) \sim \delta(\lambda - 0.70)$ in panels (b) and (e); and (iii) a heterogeneous backbone $G(\lambda) \sim \lambda^{-2.3}$ with $\lambda \in [0.02, 1]$ in panels (c) and (f). In all the cases, both precision and recall shown close to one, suggesting that the low recall shown in figures 3 and A2 may be due to heterogeneity in the activity distribution. Another reason for the low recall observed in figures 3 and A2 may be the use of the Bonferroni correction [37], which increases the precision of the methodology and lowers its recall.

Appendix B. Primary school dataset

B.1. Significant links

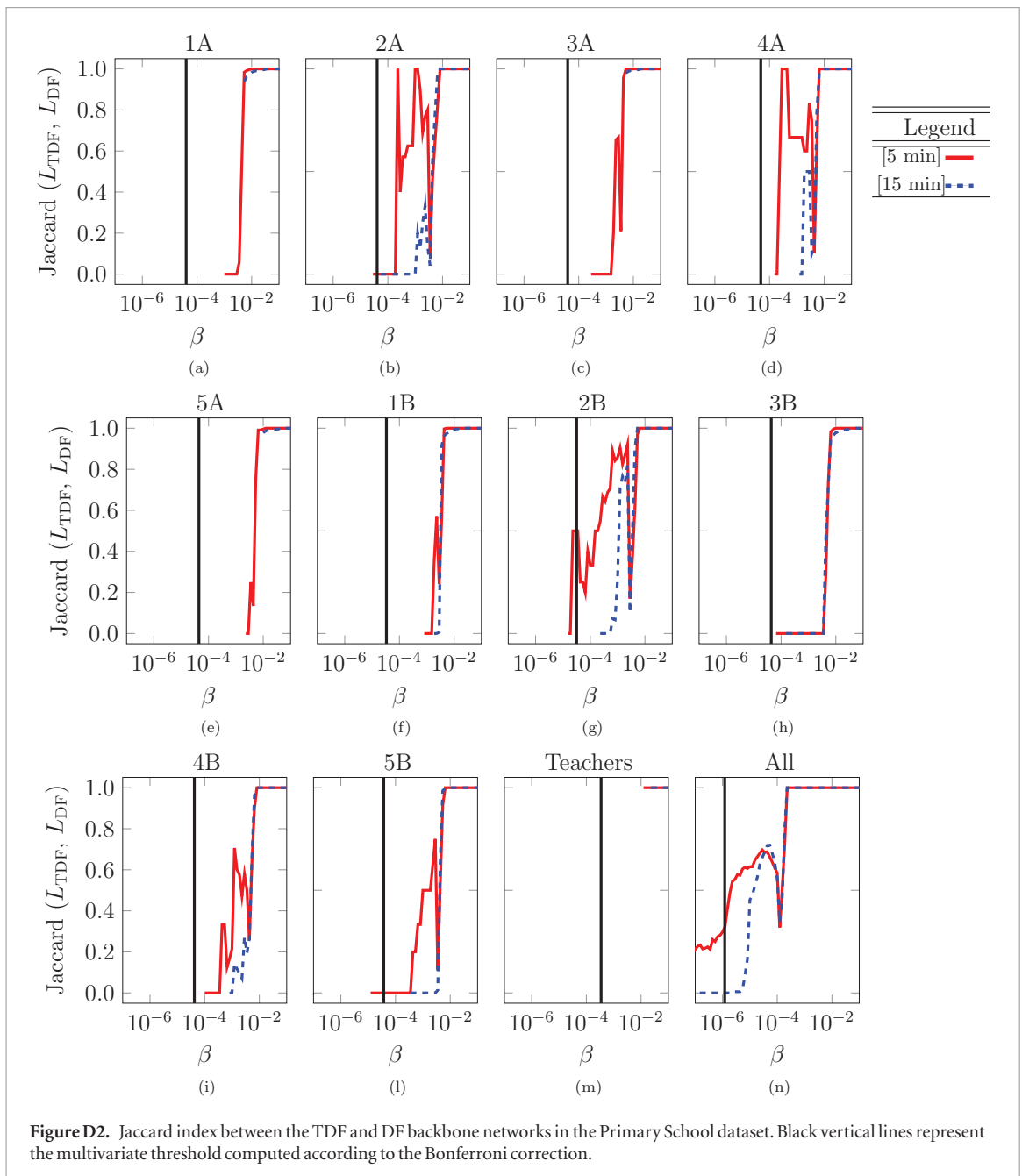
In figure B1, we report the missing results of figure 4 in the main text concerning the fraction of links detected by the ETFM or TDF over the EADM, $|L_x|/|L_{\text{EADM}}|$.

B.2. Overlap

In figure B2, we report the missing results of figure 5 in the main text concerning the Overlap index between either the ETFM or TDF and the EADM, $\text{Overlap}(L_{\text{EADM}}, L_x)$.

Appendix C. Workplace dataset

Following the main text, we analyze another real network from the Sociopattern project [30]: the Workplace dataset. We provide a data summary in table C1; analyze the EADM and ETFM accuracy in figure C1; compare the number of links detected by the ETFM or TDF with those found by the EADM in figure C2; and display the Overlap coefficient according to equation (25) in figure C3. All the results support the claims in the main text: (i) the TDF has the most restrictive null hypothesis and finds the smallest amount of links in the backbone network; (ii) the ETFM detects less links than the EADM when the system is composed of few tens of nodes, and (iii) the ETFM and EADM yield equivalent results when the system reaches one hundred nodes.



Appendix D. Temporal disparity filter and disparity filter: a comparison

In the case of both the Workplace, in figure D1, and Primary School, in figure D2, datasets, the TDF finds a different backbone network than the disparity filter [6]. We use the Jaccard index [52] defined as

$$J(L_{\text{TDF}}, L_{\text{DF}}) = \frac{|L_{\text{TDF}} \cap L_{\text{DF}}|}{|L_{\text{TDF}} \cup L_{\text{DF}}|}, \quad (\text{D.1})$$

where L_{DF} represents the set of irreducible links detected by the disparity filter. A Jaccard index equal to one implies that the two methodologies yield the same backbone network, while a lower value indicates a difference in the two detected backbone networks.

ORCID iDs

Matthieu Nadini  <https://orcid.org/0000-0003-4542-7481>

Alessandro Rizzo  <https://orcid.org/0000-0002-2386-3146>

Maurizio Porfiri  <https://orcid.org/0000-0002-1480-3539>

References

- [1] Holme P 2015 *Eur. Phys. J. B* **88** 234
- [2] Holme P and Saramäki J 2012 *Phys. Rep.* **519** 97–125
- [3] Masuda N and Lambiotte R 2016 *A Guide to Temporal Networks* vol 4 (Singapore: World Scientific)
- [4] Gemmetto V, Cardillo A and Garlaschelli D 2017 (arXiv:1706.00230)
- [5] Kobayashi T, Takaguchi T and Barrat A 2019 *Nat. Commun.* **10** 220
- [6] Serrano M Á, Boguná M and Vespignani A 2009 *Proc. Natl Acad. Sci.* **106** 6483–8
- [7] Tumminello M, Micciche S, Lillo F, Piilo J and Mantegna R N 2011 *PLoS One* **6** e17994
- [8] Li M X et al 2014 *New J. Phys.* **16** 083038
- [9] Dai L, Derudder B and Liu X 2018 *J. Transp. Geogr.* **69** 271–81
- [10] Yan X, Jeub L G, Flammini A, Radicchi F and Fortunato S 2018 *Phys. Rev. E* **98** 042304
- [11] Dianati N 2016 *Phys. Rev. E* **93** 012304
- [12] Eguiluz V M, Chialvo D R, Cecchi G A, Baliki M and Apkarian A V 2005 *Phys. Rev. Lett.* **94** 018102
- [13] Bongiorno C, London A, Micciché S and Mantegna R N 2017 *Phys. Rev. E* **96** 022321
- [14] Wu Z, Braunstein L A, Havlin S and Stanley H E 2006 *Phys. Rev. Lett.* **96** 148702
- [15] Musciotto F, Marotta L, Piilo J and Mantegna R N 2018 *Palgrave Commun.* **4** 92
- [16] Curme C, Tumminello M, Mantegna R N, Stanley H E and Kenett D Y 2015 *Quant. Finance* **15** 1375–86
- [17] Challet D, Chicheportiche R, Lallouache M and Kassibrakis S 2018 *Adv. Complex Syst.* **21** 1850019
- [18] Mazzarisi P, Barucca P, Lillo F and Tantari D 2020 *Eur. J. Oper. Res.* **281** 50–65
- [19] Nadini M, Bongiorno C, Rizzo A and Porfiri M 2020 *Nonlinear Dyn.* **99** 855–78
- [20] Maracaccioli R and Livan G 2019 *Nat. Commun.* **10** 745
- [21] Perra N, Gonçalves B, Pastor-Satorras R and Vespignani A 2012 *Sci. Rep.* **2** 469
- [22] Nadini M, Rizzo A and Porfiri M 2018 *IEEE Trans. Netw. Sci. Eng.* (<https://doi.org/10.1109/TNSE.2018.2885483>)
- [23] Nadini M, Sun K, Ubaldi E, Starnini M, Rizzo A and Perra N 2018 *Sci. Rep.* **8** 2352
- [24] Pastor-Satorras R, Castellano C, Van Mieghem P and Vespignani A 2015 *Rev. Mod. Phys.* **87** 925
- [25] Allard A and Hébert-Dufresne L 2019 *Phys. Rev. X* **9** 011023
- [26] Hecker M, Lambeck S, Toepfer S, Van Someren E and Guthke R 2009 *Biosystems* **96** 86–103
- [27] Markowetz F and Spang R 2007 *BMC Bioinform.* **8** S5
- [28] Barrat A, Barthelemy M, Pastor-Satorras R and Vespignani A 2004 *Proc. Natl Acad. Sci.* **101** 3747–52
- [29] Newman M E 2004 *Phys. Rev. E* **70** 056131
- [30] www.sociopatterns.org
- [31] Scargle J D, Norris J P, Jackson B and Chiang J 2013 *Astrophys. J.* **764** 167
- [32] Serrano M Á and Boguná M 2005 *AIP Conf. Proc.* **776** 101–7
- [33] Barbour A and Eagleson G 1983 *Adv. Appl. Probab.* **15** 585–600
- [34] Steele J M 1994 *Am. Math. Mon.* **101** 48–54
- [35] Le Cam L et al 1960 *Pac. J. Math.* **10** 1181–97
- [36] Shaffer J P 1995 *Annu. Rev. Psychol.* **46** 561–84
- [37] Hochberg Y and Tamhane A 1987 *Multiple Comparison Procedure* (New York: Wiley)
- [38] Benjamini Y and Hochberg Y 1995 *J. R. Stat. Soc. B* **57** 289–300
- [39] Newman M E J 2010 *Networks: an Introduction* (Oxford: Oxford University Press)
- [40] Manning C D, Raghavan P and Schütze H 2008 *Introduction to Information Retrieval* (Cambridge: Cambridge University Press)
- [41] Vijaymeena M and Kavitha K 2016 *Mach. Learn. Appl. Int. J* **3** 19–28
- [42] Perra N, Baronchelli A, Mocanu D, Gonçalves B, Pastor-Satorras R and Vespignani A 2012 *Phys. Rev. Lett.* **109** 238701
- [43] Perra N, Balcan D, Gonçalves B and Vespignani A 2011 *PLoS One* **6** e23084
- [44] Rizzo A, Pedalino B and Porfiri M 2016 *J. Theor. Biol.* **394** 212–22
- [45] Krause J et al 2002 *Living in Groups* (Oxford: Oxford University Press)
- [46] Rosenfeld L and Morville P 2002 *Information Architecture for the World Wide Web* (Sebastopol, CA: O'Reilly Media, Inc.)
- [47] Easley D et al 2010 *Networks, Crowds, and Markets* vol 8 (Cambridge: Cambridge University Press)
- [48] LeDoux J E 2000 *Annu. Rev. Neurosci.* **23** 155–84
- [49] <https://github.com/matnado/Backbone-Detection>
- [50] <http://sociopatterns.org/datasets/primary-school-temporal-network-data/>
- [51] <http://sociopatterns.org/datasets/contacts-in-a-workplace/>
- [52] Jaccard P 1901 *Bull. Soc. Vaudoise Sci. Nat.* **37** 547–79