PATENet: Pairwise Alignment of Time Evolving Networks

Shlomit Gur[1](✉) and Vasant G. Honavar[1,2]

[1] Huck Institutes of the Life Sciences, Pennsylvania State University,
University Park, PA 16802, USA
szgl80@psu.edu
[2] College of Information Sciences and Technology,
Pennsylvania State University, University Park, PA 16802, USA

Abstract. Networks that change over time, e.g. functional brain networks that change their structure due to processes such as development or aging, are naturally modeled by time-evolving networks. In this paper we present PATENet, a novel method for aligning time-evolving networks. PATENet offers a mathematically-sound approach to aligning time evolving networks. PATENet leverages existing similarity measures for networks with fixed topologies to define well-behaved similarity measures for time evolving networks. We empirically explore the behavior of PATENet through synthetic time evolving networks under a variety of conditions.

Keywords: Network science · Multilayer networks · Temporal alignment

1 Introduction

Network science has provided a variety of powerful tools for describing, representing, and analyzing a variety of real-world systems including social networks, the internet, functional brain networks, and biomolecular networks [9, 10]. While many of the tools and techniques of network science, e.g. topological analyses and network alignment, focus on networks with fixed topologies, the structure of networks that represent real-world systems change over time. Such networks are naturally modeled as time-evolving networks (TENs) [11, 14]. TENs can display dynamics on networks (where the network structure does not change over time, but the activity of the nodes does); dynamics of networks (where the activity of nodes does not change but the structure does); and dynamics of and on networks (where both the structure and activity change over time) [3]. The relatively young sub-field of TENs [9] has already yielded a substantial body of work, focusing primarily on models of time-evolving networks and the characterization of network dynamics [14]. However, there is limited work on methods for comparative analyses of TENs.

To motivate the underlying problem, consider experimental subjects who undergo functional magnetic imaging (fMRI) recordings of resting state brain activation at different points in time, e.g. in the context of a longitudinal study of changes in functional connectivity as a function of development, aging, or disease progression [16]. The resulting data from each subject are naturally represented as a temporally

P. Perner (Ed.): MLDM 2018, LNAI 10934, pp. 85–98, 2018.
https://doi.org/10.1007/978-3-319-96136-1_8

ordered sequence of functional connectivity networks. To complicate matters, it may not be straightforward to establish one-to-one correspondence between the recording times across subjects because of differences in the timing of recordings, missed recording sessions, etc. Furthermore, even in the case of subjects with recordings obtained at what appears to be matching time points e.g. age in years, because of differences in the onset and progression of development, aging, or disease, and the trajectories across subjects, the networks at the respective time points may not be comparable. With the exception of [15], which focuses on temporal registration of deforming meshes, to the best of our knowledge, there is no work on aligning (temporally) ordered sequences of networks (OSN). The most closely related body of work focuses on aligning ordered sequences of letters over a finite alphabet e.g. DNA or protein sequences [17], video frames [4], and clinical histories [13]. However, with the exception of methods for aligning sequences of letters [17], the methods used are ad hoc and are not supported by a sound mathematical rationale and hence lack precise mathematical characterization and are not amenable to generalization to other related problem domains.

Against this background, we focus on the problem of aligning a pair of OSNs. Specifically, we describe PATENet, a mathematically sound family of algorithms for aligning a pair of OSNs. PATENet requires as input, in addition to a pair of OSNs to be aligned, a measure of pairwise similarity of fixed topology networks, a monotonically increasing function, and a match threshold. It produces as output an optimal alignment of the given pair of OSNs. Specifically, PATENet generalizes the Smith-Waterman (SW) algorithm [17], a dynamic programming algorithm for aligning two ordered sequences of letters, given a pairwise measure of substitutability of letters and gap penalties. SW produces an optimal local alignment, i.e. aligned segments of the given pair of sequences with the largest cumulative similarity. Conceptually, adapting the SW algorithm to yield a mathematically sound algorithm for aligning a pair of OSNs is straightforward; we replace letters by networks, and replace pairwise substitutability of letters by a well-behaved measure of pairwise similarity of (fixed topology) networks. However, in order for this approach to yield both mathematically sound and practically useful algorithms for aligning OSNs, several challenges need to be addressed; there are a variety of measures of similarity or distance between networks that are tailored [6] to meet the needs of specific applications [7]. We need to adjust such measures so as to ensure that the algorithms that use them for aligning OSNs are mathematically well-behaved. In the current work we also show that the PATENet family of algorithms can be readily extended to align ordered sequences of elements other than networks, provided suitable and well-behaved measures of similarity between elements are available.

2 Preliminaries

We use $G = G(\mathcal{V}, \mathcal{E}_G)$ to denote a network, where $\mathcal{V}$ is its set of nodes and $\mathcal{E}_G$ is its set of edges. We define OSN $\mathcal{G}$ to be a sequence of n networks, $\mathcal{G} = \{G_1, G_2, \ldots, G_n\}$, where $\forall 1 \leq i \leq n \in \mathbb{N},\ G_i = (\mathcal{V}, \mathcal{E}_{G_i})$ denotes the i^{th} element of $\mathcal{G}$, which is a snapshot of a TEN at time t_i, and $\forall 1 < i \leq n \in \mathbb{N},\ t_{i-1} < t_i$. We use upper case letters, e.g. H, to denote matrices or networks, lower case letters, e.g. n, to denote scalars, and script letters, e.g. $\mathcal{V}$, to denote sets.

Definition 1. Let $G = (\mathcal{V}, \mathcal{E})$ and $G' = (\mathcal{V}, \mathcal{E}')$ be two networks with the same set of nodes $\mathcal{V}$, and respective sets of edges $\mathcal{E}$ and $\mathcal{E}'$ (either identical or different). A function $s(G, G')$, mapping two graphs to $[0, 1]$, is said to be a well-defined *unsigned* normalized network similarity measure (UNNSM) if it satisfies the following properties (adapted from [12]):

1. Identity property: $s(G, G') \leq s(G, G) = 1 \, \forall \, G, G'$.
2. Symmetry property: $s(G, G') = s(G', G) \, \forall \, G, G'$.
3. Minimum property: $s(G, G') \overset{|\mathcal{V}| \to \infty}{\to} 0$ where WLOG G is the complete network, and G' is the empty network (i.e. $\mathcal{E}^C = \mathcal{E}'$).

Definition 2. Similarly, a function $s'(G, G')$, mapping two graphs to $[-1, 1]$, is said to be a well-defined *signed* normalized network similarity measure (SNNSM) if it satisfied the properties described in definition 1, with the minimum property adjusted to the signed range: $s'(G, G') \overset{|\mathcal{V}| \to \infty}{\to} -1$ (rather than 0).

For simplification purposes we assumed G and G' to have the same set of nodes $\mathcal{V}$. However, if $\mathcal{V}_G \neq \mathcal{V}_{G'}$, where $\mathcal{V}_G$ and $\mathcal{V}_{G'}$ denote the set of nodes of G and G', respectively, then $\mathcal{V} = \mathcal{V}_G \bigcup \mathcal{V}_{G'}$ for the definitions above.

2.1 The Smith-Waterman (SW) Sequence Alignment Algorithm

The SW algorithm is a local sequence alignment algorithm, designed to find pairs of segments with high cumulative degree of similarity between two sequences of amino acids (AAs), $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$. There are 22 AAs, and the similarity between every pair of AAs is specified by the entries of a ‘substitution matrix’ $SM \in \mathbb{R}^{22 \times 22}$. The SW algorithm uses dynamic programming to generate a ‘scoring matrix’ $H = H(A, B) \in \mathbb{R}^{(n+1) \times (m+1)}$, which is defined as follows:

$$\begin{aligned} &\forall 0 \leq i \leq n \in \mathbb{N} \;\; \forall 0 \leq j \leq m \in \mathbb{N}, \; H_{i,0} = H_{0,j} = 0 \\ &\forall 0 < i \leq n \in \mathbb{N} \, \forall 0 < j \leq m \in \mathbb{N}, \\ &H_{i,j} = max\left\{H_{i-1,j-1} + s\left(a_i, b_j\right), max_{1 \leq k \leq i}\left\{H_{i-k,j} - w_k\right\}, max_{1 \leq l \leq j}\left\{H_{i,j-l} - w_l\right\}, 0\right\} \end{aligned} \quad (1)$$

Where $s\left(a_i, b_j\right)$ is the similarity score between the two AAs $a_i \in A$ and $b_j \in B$, according to SM, and w_k is a value assigned to deletions or insertions of length k. Insertions and deletions refer to cases where an element (or a few) within one sequence is not aligned with an element (or a sequence of elements) within the paired sequence. The length of insertions and deletions is the number of consecutive insertions and/or deletions. $w_1 \in \mathbb{R}$ is referred to as ‘gap penalty’ and is the value assigned to a gap of length 1, and $w_k = f(w_1, k) \in \mathbb{R}$ is the penalty for a gap of length k, where $f(w_1, k)$ can be affine or linear, for example, in relation to w_1.

Let χ denote the maximum value in H, then $\chi : A \times B \rightarrow \mathbb{R}$ is the local alignment score between the two sequences A and B, and is used to reveal the best local alignment, by way of backtracing on H. Starting at a cell holding χ, backtracing is performed on H until a cell holding 0 is reached according to the following logic:

1. If $H_{i,j} = H_{i-1,j-1} + s\left(a_i, b_j\right)$, then a_i is aligned with b_j and the process continues from $H_{i-1,j-1}$.
2. Else if $H_{i,j} = H_{i-1,j} - w$, then a_i has no alignment in B, and the process continues from $H_{i-1,j}$.
3. Else if $H_{i,j} = H_{i,j-l} - w$, then b_j has no alignment in A, and the process continues from $H_{i,j-1}$.

The solution is not guaranteed to be unique; there could be multiple cells in H holding χ, in which case the backtracing process can be initiated at any of these cells, resulting in different, yet equally good, local alignments.

3 PATENet

In this paper we focus on aligning a pair of OSNs. To accommodate OSNs resulting from longitudinal recordings from subjects, we impose the following natural desiderata on the alignments returned by PATENet:

1. Preservation in the alignment of the relative order of elements within the sequences. E.g., if element 3 of the first sequence is aligned with element 5 of the second sequence, element 4 of the first sequence can be aligned only with elements in positions 6 or greater in the second sequence.
2. Accommodation of unaligned elements in both sequences (i.e. aligning two sequences of length n and m, respectively, should not force the alignment of $min(n, m)$ elements).
3. Accommodation of longitudinal gaps (e.g. time points existing in one sequence but missing in the other).

3.1 Alternative Substitution Matrix Construction

The SW algorithm requires a well-defined *SM*, holding both positive values for possible matches and negative values for non-matches. Furthermore, unlike in the case of AA sequences, where the sequence elements are drawn from a fixed alphabet, OSNs can contain arbitrary networks defined over a given set of vertices and edges. Hence, we will adapt existing network similarity measures to define pairwise similarity of elements (networks) in OSNs.

Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively. Let s be a well-defined UNNSM. Finally, let $0 < \varphi < 1 \in \mathbb{R}$ be a threshold on s, where $match\left(G_i, G_j'\right) = \begin{cases} 1, \, if \;\; \varphi \leq s\left(G_i, G_j'\right) \\ 0, \, if \;\; \varphi > s\left(G_i, G_j'\right) \end{cases}$ $\forall G_i \in \mathcal{G}$, $\forall G_j' \in \mathcal{G}'$, and let $\ell(x)$, ℓ: $[0, 1] \rightarrow [-1, 1]$ be a signed normalized monotonically increasing transform, with $\ell(\varphi) = 0$, $\ell(0) = -1$, and $\ell(1) = 1$. We propose

$$SM_{i,j} = \ell\left(s\left(G_i, G_j^{'}\right)\right) \forall 1 \leq i \leq n \in \mathbb{N} \ \forall 1 \leq j \leq m \in \mathbb{N} \tag{2}$$

to construct an 'alternative substitution matrix' $SM = SM(\mathcal{G}, \mathcal{G}') \in [-1, 1]^{n \times m}$.

For example, for $\alpha = \frac{\varphi}{1-\varphi}$ and $0.5 < \varphi < 1$, $\ell(x) = 1 - \log_\alpha[\alpha^2 + (1 - \alpha^2) \cdot x]$ is such a signed normalized monotonically increasing transform (proof omitted), and along with $DeltaCon\left(G_i, G_j^{'}\right)$ [12] as the well-defined UNNSM (by definition), can be used to construct an alternative SM. Another example includes $\tilde{s}(G, G') = (1 - NSSD(G, G'))$ as the well-defined UNNSM (proof omitted), where $NSSD(G, G')$ is the normalized sum squared difference, and $\tilde{\ell}(\tilde{x}) = \begin{cases} \frac{\tilde{x}-\varphi}{1-\varphi}, if\ \tilde{x} \geq \varphi \\ \frac{\tilde{x}-\varphi}{\varphi}, if\ \tilde{x} < \varphi \end{cases}$ as the signed normalized monotonically increasing transform (proof omitted).

Lemma 1. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively, and let s be a well-defined UNNSM, and $\ell\colon\ [0, 1] \to [-1, 1]$ be a signed normalized monotonically increasing transform, as described above. Then $\ell(s(G, G'))$, mapping two graphs to $[-1, 1]$, satisfies the properties of a well-defined SNNSM.*

Proof

Identity: $\ell(\mathrm{s(G, G')}) \leq \ell(\mathrm{s(G, G)}) = \ell(1) = \ell(\max\{\mathrm{s(G, G')}\}) = \max\{\ell(\mathrm{s(G, G')})\} = 1$ ∎

Symmetry: $\ell\left(s\left(G, G^{'}\right)\right) = \ell\left(s\left(G^{'}, G\right)\right)$ ∎

Minimum: $\ell\left(s\left(G, G^{'}\right)\right) \overset{|\mathcal{V}| \to \infty}{\to} \ell(0) = \ell\left(min\left\{s\left(G, G^{'}\right)\right\}\right) = min\left\{\ell\left(s\left(G, G^{'}\right)\right)\right\} = -1$ ∎

3.2 From SW to PATENet

The SW algorithm meets the first two desiderata of PATENet (preservation of temporal order and accommodation of possible unaligned elements in both sequences). To satisfy the third desideratum (accommodation of longitudinal gaps), we set the gap penalty to zero. Therefore, for an alternative SM, following the construction described above, the scoring matrix of PATENet $\tilde{H} = \tilde{H}\left(\mathcal{G}, \mathcal{G}^{'}\right) \in \mathbb{R}^{(n+1) \times (m+1)}$, hereafter referred to as 'OSN scoring matrix', is specified as follows:

$$\begin{aligned} &\forall 0 \leq i \leq n \in \mathbb{N} \ \forall 0 \leq j \leq m \in \mathbb{N},\ \tilde{H}_{i,0} = \tilde{H}_{0,j} = 0 \\ &\forall 0 < i \leq n \in \mathbb{N} \ \forall 0 < j \leq m \in \mathbb{N}, \\ &\tilde{H}_{i,j} = max\left\{\tilde{H}_{i-1,j-1} + SM_{i,j}, max_{1 \leq k \leq i}\left\{\tilde{H}_{i-k,j}\right\}, max_{1 \leq l \leq j}\left\{\tilde{H}_{i,j-l}\right\}, 0\right\} \end{aligned} \tag{3}$$

Lemma 2. *Let $\mathcal{G}$ and $\mathcal{G}^{'}$ be two OSNs with n and m elements, respectively. Let $\tilde{H} \in \mathbb{R}^{(n+1) \times (m+1)}$ be their OSN scoring matrix, then:*

(2.1) $\forall 1 \le i \le n \in \mathbb{N}$ *and* $\forall 1 \le j \le m \in \mathbb{N},\ \tilde{H}_{i-1,j} \le \tilde{H}_{i,j}$

(2.2) $\forall 1 \le i \le n \in \mathbb{N}$ *and* $\forall 1 \le j \le m \in \mathbb{N},\ \tilde{H}_{i,j-1} \le \tilde{H}_{i,j}$

Proof

(2.1) $\tilde{H}_{i,j} = max\left\{\tilde{H}_{i-1,j-1} + SM_{i,j}, max_{1 \le k \le i}\left\{\tilde{H}_{i-k,j}\right\}, max_{1 \le l \le j}\left\{\tilde{H}_{i,j-l}\right\}, 0\right\} \ge max_{1 \le k \le i}\left\{\tilde{H}_{i-k,j}\right\} \ge \tilde{H}_{i-1,j}$ ■

(2.2) $\tilde{H}_{i,j} = max\left\{\tilde{H}_{i-1,j-1} + SM_{i,j}, max_{1 \le k \le i}\left\{\tilde{H}_{i-k,j}\right\}, max_{1 \le l \le j}\left\{\tilde{H}_{i,j-l}\right\}, 0\right\} \ge max_{1 \le l \le j}\left\{\tilde{H}_{i,j-l}\right\} \ge \tilde{H}_{i,j-1}$ ■

Lemma 3. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively. Let $\tilde{H} \in \mathbb{R}^{(n+1)\times(m+1)}$ be their OSN scoring matrix, then:*

(3.1) $\forall 1 \le i \le n \in \mathbb{N}$ *and* $\forall 1 \le j \le m \in \mathbb{N},\ \underset{1 \le k \le i}{max}\left\{\tilde{H}_{i-k,j}\right\} = \tilde{H}_{i-1,j}$

(3.2) $\forall 1 \le i \le n \in \mathbb{N}$ *and* $\forall 1 \le j \le m \in \mathbb{N},\ \underset{1 \le l \le j}{max}\left\{\tilde{H}_{i,j-l}\right\} = \tilde{H}_{i,j-1}$

Proof

Intuitive based on Lemma 2 ■

Therefore, the OSN scoring matrix $\tilde{H}$ of PATENet is equivalent to:

$$\begin{aligned}&\forall 0 \le i \le n \in \mathbb{N}\ \forall 0 \le j \le m \in \mathbb{N},\ \tilde{H}_{i,0} = \tilde{H}_{0,j} = 0\\&\forall 0 < i \le n \in \mathbb{N}\ \forall 0 < j \le m \in \mathbb{N},\\&\tilde{H}_{i,j} = max\left\{\tilde{H}_{i-1,j-1} + SM_{i,j}, \tilde{H}_{i-1,j}, \tilde{H}_{i,j-1}, 0\right\}\end{aligned} \tag{4}$$

Lemma 4. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively. Let $\tilde{H} = \tilde{H}\left(\mathcal{G}, \mathcal{G}'\right) \in \mathbb{R}^{(n+1)\times(m+1)}$ and $\tilde{H}' = \tilde{H}\left(\mathcal{G}', \mathcal{G}\right) \in \mathbb{R}^{(m+1)\times(n+1)}$ be their OSN scoring matrices, and let SM and SM′ be the alternative substitution matrices of $\tilde{H}$ and $\tilde{H}'$, repectively. Then $\forall 1 \le i \le n \in \mathbb{N}$ and $\forall 1 \le j \le m \in \mathbb{N}$: (4.1) $SM_{i,j} = SM'_{j,i}$ and (4.2) $\tilde{H}_{i,j} = \tilde{H}'_{j,i}$.*

Proof

(4.1) $\forall 1 \le \mathrm{i} \le \mathrm{n} \in \mathbb{N},\ \forall 1 \le \mathrm{j} \le \mathrm{m} \in \mathbb{N}:\ \mathrm{SM}_{\mathrm{i,j}} = \ell\left(\mathrm{s}\left(\mathrm{G_i}, \mathrm{G'_j}\right)\right) = \ell\left(\mathrm{s}\left(\mathrm{G'_j}, \mathrm{G_i}\right)\right) = \mathrm{SM'_{j,i}}$ ■

(4.2) For $j = i = 1$: $\tilde{H}_{1,1} = max\left\{\tilde{H}_{0,0} + SM_{1,1}, \tilde{H}_{0,1}, \tilde{H}_{1,0}, 0\right\} = max\left\{SM_{1,1}, 0\right\} = max\left\{SM'_{1,1}, 0\right\} = \tilde{H}'_{1,1}$.

For $j = 1$, $\forall 2 \le i \le n \in \mathbb{N}$, we can safely assume $\tilde{H}_{i-1,1} = \tilde{H}'_{1,i-1}$ for induction: $\tilde{H}_{i,1} = max\left\{\tilde{H}_{i-1,0} + SM_{i,1}, \tilde{H}_{i-1,1}, \tilde{H}_{i,0}, 0\right\} = max\left\{SM_{i,1}, \tilde{H}_{i-1,1}, 0\right\} = max\left\{SM'_{1,i}, \tilde{H}'_{1,i-1}, 0\right\} = \tilde{H}'_{1,i}$.

For $2 \le j = k \in \mathbb{N}$ and $i = 1$: $\tilde{H}_{1,k} = max\{\tilde{H}_{0,k-1} + SM_{1,k}, \tilde{H}_{0,k}, \tilde{H}_{1,k-1}, 0\} = max\{SM_{1,k}, \tilde{H}_{1,k-1}, 0\} = max\{SM_{1,k}, max\{SM_{1,k-1}, \tilde{H}_{1,k-2}, 0\}, 0\} = max\{SM_{1,k}, SM_{1,k-1}, \tilde{H}_{1,k-2}, 0\} = \cdots = max\{SM_{1,k}, SM_{1,k-1}, \ldots, SM_{1,2}, SM_{1,1}, 0\} \ldots = max\left\{SM'_{k,1}, SM'_{k-1,1}, \ldots, SM'_{2,1}, SM'_{1,1}, 0\right\} = \tilde{H}'_{k,1}$.

For $2 \le j = k \in \mathbb{N}$, $\forall 1 \le i \le n \in \mathbb{N}$, we can safely assume $\tilde{H}_{i,k-1} = \tilde{H}'_{k-1,i}$ as well as $\tilde{H}_{i-1,k} = \tilde{H}'_{k,i-1}$ and therefore also $\tilde{H}_{i-1,k-1} = \tilde{H}'_{k-1,i-1}$ for induction: $\tilde{H}_{i,k} = max\{\tilde{H}_{i-1,k-1} + SM_{i,k}, \tilde{H}_{i-1,k}, \tilde{H}_{i,k-1}, 0\} = max\left\{\tilde{H}'_{k-1,i-1} + SM'_{k,i}, \tilde{H}'_{k,i-1}, \tilde{H}'_{k-1,i}, 0\right\} = \tilde{H}'_{k,i}$ ∎

Lemma 5. *Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively. Let $\tilde{H} = \tilde{H}\left(\mathcal{G}, \mathcal{G}'\right) \in \mathbb{R}^{(n+1)\times(m+1)}$ be their OSN scoring matrix, and let SM be its alternative substitution matrix. Then the alignment score $\tilde{\chi} = max\{\tilde{H}\}$[1] is equivalent to $\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i, G'_j\right) \cdot SM_{i,j}\right]$, where $\rho\left(G_i, G'_j\right) = \begin{cases} 1, \textit{ if } \left(G_i, G'_j\right) \textit{are aligned with each other} \\ 0, \textit{otherwise} \end{cases}$.*

Proof

By definition of the SW algorithm, $\forall 1 \le i \le n \in \mathbb{N}, \forall 1 \le j \le m \in \mathbb{N}, \tilde{H}_{i,j} =$ the maximum similarity of two segments ending in G_i and G'_j. The similarity score of the alignment is the sum of similarity scores between every pair of aligned elements and weights of all insertions and deletions in the alignment. Since $w_1 = 0$ in PATENet, the weights of all insertions and deletions is always 0, leaving only the sum of similarity scores between every pair of aligned elements, which can be written as: $\tilde{\chi} = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i, G'_j\right) \cdot SM_{i,j}\right]$, where $\rho\left(G_i, G'_j\right) = \begin{cases} 1, \textit{ if } \left(G_i, G'_j\right) \textit{are aligned with each other} \\ 0, \textit{otherwise} \end{cases}$ ∎

3.3 OSN Alignment Score

Alignment of elements across a pair of OSNs may be informative by itself and reveal temporally-preserved similarities between the two OSNs. However, another concept worth borrowing from sequence alignment is that of the alignment score $\tilde{\chi} = max\{\tilde{H}\}$.

Lemma 6. *An OSN alignment score $\tilde{\chi} = max\{\tilde{H}\}$ satisfies properties that are similar to those of a well-defined UNNSM, except for the normalization-related upper bound. Identity property*[2]*: $\tilde{\chi}\left(\mathcal{G}, \mathcal{G}'\right) \le \tilde{\chi}(\mathcal{G}, \mathcal{G})\ \forall \mathcal{G}, \mathcal{G}'$; Symmetry property: $\tilde{\chi}\left(\mathcal{G}, \mathcal{G}'\right) = \tilde{\chi}\left(\mathcal{G}', \mathcal{G}\right)\ \forall \mathcal{G}, \mathcal{G}'$; Minimum property: $\tilde{\chi}\left(\mathcal{G}, \mathcal{G}'\right) \xrightarrow{|\mathcal{V}|\to\infty} 0$ where WLOG $\mathcal{G}$ is the complete OSN, and $\mathcal{G}'$ is the empty OSN (i.e. $\forall 1 \le i \le n \in \mathbb{N}, \forall 1 \le j \le m \in \mathbb{N}, \mathcal{E}^{C}_{G_i} = \mathcal{E}_{G'_j}$).*

---

[1] Notice that $\tilde{\chi} : \mathcal{G} \times \mathcal{G}' \to \mathbb{R}$.

[2] Notice that $\tilde{\chi}(G, G) = 1$ is not required, as the alignment score has no upper bound.

Proof

Based on Lemma 5, $\tilde{\chi}(\mathcal{G},\mathcal{G}') = max\{\tilde{H}\} = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i, G'_j\right)\cdot SM_{i,j}\right]$, where $\rho\left(G_i, G'_j\right) = \begin{cases} 1,\ if\ \left(G_i, G'_j\right) are\ aligned\ with\ each\ other \\ 0, otherwise \end{cases}$. Identity: $\tilde{\chi}(\mathcal{G},\mathcal{G}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\rho\left(G_i, G_j\right)\cdot SM_{i,j}\right] = \sum_{i=1}^{n}[1\cdot 1] = n$ and $\tilde{\chi}\left(\mathcal{G},\mathcal{G}'\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i, G'_j\right)\cdot SM_{i,j}\right] \le \sum_{i=1}^{n}[1\cdot 1] = n = \tilde{\chi}(\mathcal{G},\mathcal{G})$ ∎

Based on Lemma 4, if $\tilde{H} = \tilde{H}\left(\mathcal{G},\mathcal{G}'\right) \in \mathbb{R}^{(n+1)\times(m+1)}$ and $\tilde{H}' = \tilde{H}\left(\mathcal{G}',\mathcal{G}\right) \in \mathbb{R}^{(m+1)\times(n+1)}$, then $\forall 1 \le i \le n \in \mathbb{N}, \forall 1 \le j \le m \in \mathbb{N},\ \tilde{H}_{i,j} = \tilde{H}'_{j,i}$. Symmetry: $\tilde{\chi}\left(\mathcal{G},\mathcal{G}'\right) = max\{\tilde{H}\} = max_{1\le i\le n, 1\le j\le m}\{\tilde{H}_{i,j}\} = max_{1\le j\le m, 1\le i\le n}\left\{\tilde{H}'_{j,i}\right\} = max\{\tilde{H}'\} = \tilde{\chi}\left(\mathcal{G}',\mathcal{G}\right)$ ∎

Minimum: $\forall 1 \le i \le n \in \mathbb{N},\ \forall 1 \le j \le m \in \mathbb{N},\quad s\left(G_i, G'_j\right) \overset{|\mathcal{V}|\to\infty}{\to} 0 \Rightarrow \rho\left(G_i, G'_j\right) \overset{|\mathcal{V}|\to\infty}{\to} 0 \Rightarrow \tilde{\chi}(\mathcal{G},\mathcal{G}') = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i, G'_j\right)\cdot SM_{i,j}\right] \overset{|\mathcal{V}|\to\infty}{\to} \sum_{i=1}^{n}\sum_{j=1}^{m}[0] = 0$ ∎

We observe that PATENet can be used to extend the UNNSM used for constructing SM, into an unsigned normalized order-aware OSN similarity measure. Let $\mathcal{G}$ and $\mathcal{G}'$ be two OSNs with n and m elements, respectively. Let $\tilde{H} = \tilde{H}\left(\mathcal{G},\mathcal{G}'\right) \in \mathbb{R}^{(n+1)\times(m+1)}$ be the corresponding OSN scoring matrix, and let SM be its alternative substitution matrix. Let s be the well-defined UNNSM used for constructing SM, and $\rho\left(G_i, G'_j\right) = \begin{cases} 1,\ if\ \left(G_i, G'_j\right) are\ aligned\ with\ each\ other \\ 0, otherwise \end{cases}$.

Then $\mathscr{k}: \mathcal{G} \times \mathcal{G}' \to [0, 1]$ can be defined as

$$\mathscr{k}(\mathcal{G},\mathcal{G}') = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}[\rho(G_i, G'_j)\cdot s(G_i, G'_j)]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}\rho(G_i, G'_j), 1\right\}} \tag{5}$$

which is hereafter referred to as an ‘OSN similarity score’.[3]

Lemma 7. *An OSN similarity score $\mathscr{k}(\mathcal{G},\mathcal{G}')$ satisfies identity, symmetry and minimum properties, similar to those that hold for UNNSM. Identity property: $\mathscr{k}(\mathcal{G},\mathcal{G}') \le \mathscr{k}(\mathcal{G},\mathcal{G}) = 1\ \forall \mathcal{G},\mathcal{G}'$; Symmetry property: $\mathscr{k}(\mathcal{G},\mathcal{G}') = \mathscr{k}(\mathcal{G}',\mathcal{G})\ \forall \mathcal{G},\mathcal{G}'$; Minimum property: $\mathscr{k}(\mathcal{G},\mathcal{G}') \xrightarrow{|\mathcal{V}|\to\infty} 0$ where WLOG $\mathcal{G}$ is the complete OSN, and $\mathcal{G}'$ is the empty OSN (i.e. $\forall 1 \le i \le n \in \mathbb{N}, \forall 1 \le j \le m \in \mathbb{N}, \mathcal{E}_{G_i}^{C} = \mathcal{E}_{G'_j}$).*

---

[3] Notice that the OSN similarity score measures similarity in the context of the locally aligned segments of the sequences. That is, if OSNs $\mathcal{G}$ and $\mathcal{G}'$ have k elements aligned with average element-wise similarity of h, whether $k = min(n, m)$ or $k < min(n, m)$, $\mathscr{k}(\mathcal{G},\mathcal{G}') = h$. Additionally, if OSNs $\mathcal{G}$ and $\mathcal{G}'$ have one element aligned with element-wise similarity of 1.0, while OSNs $\mathcal{G}$ and $\mathcal{G}''$ have four elements aligned with each element-wise similarity being 0.9, $\mathscr{k}(\mathcal{G},\mathcal{G}') > \mathscr{k}(\mathcal{G},\mathcal{G}'')$ (but $\tilde{\chi}(\mathcal{G},\mathcal{G}') < \tilde{\chi}(\mathcal{G},\mathcal{G}'')$).

Proof
WLOG, we assume $n \leq m$. Identity:

$$\mathscr{k}(\mathcal{G},\mathcal{G}) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\rho\left(G_i,G_j\right)\cdot s\left(G_i,G_j\right)\right]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\rho\left(G_i,G_j\right),1\right\}} = \frac{\sum_{i=1}^{n}\left[1\cdot s\left(G_i,G_i\right)\right]}{max\left\{\sum_{i=1}^{n}1,1\right\}} = \frac{n}{n} = 1 \text{ and}$$

$$\mathscr{k}(\mathcal{G},\mathcal{G}') = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i,G'_j\right)\cdot s\left(G_i,G'_j\right)\right]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}\rho\left(G_i,G'_j\right),1\right\}} \leq \frac{\sum_{i=1}^{n}[1\cdot 1]}{max\left\{\sum_{i=1}^{n}1,1\right\}} \leq \frac{n}{max\{n,1\}} = \frac{n}{n} = 1 = \mathscr{k}(\mathcal{G},\mathcal{G})$$

■

Symmetry:

$$\mathscr{k}(\mathcal{G},\mathcal{G}') = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i,G'_j\right)\cdot s\left(G_i,G'_j\right)\right]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}\rho\left(G_i,G'_j\right),1\right\}} = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n}\left[\rho\left(G'_j,G_i\right)\cdot s\left(G'_j,G_i\right)\right]}{max\left\{\sum_{j=1}^{m}\sum_{i=1}^{n}\rho\left(G'_j,G_i\right),1\right\}} = \mathscr{k}(\mathcal{G}',\mathcal{G})$$

■

Minimum: $\forall 1 \leq i \leq n \in \mathbb{N},\ \forall 1 \leq j \leq m \in \mathbb{N},$

$$s\left(G_i, G'_j\right) \xrightarrow{|\mathcal{V}|\to\infty} 0 \implies \rho\left(G_i, G'_j\right) \xrightarrow{|\mathcal{V}|\to\infty} 0 \implies \mathscr{k}(\mathcal{G},\mathcal{G}') = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\rho\left(G_i,G'_j\right)\cdot s\left(G_i,G'_j\right)\right]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}\rho\left(G_i,G'_j\right),1\right\}} \xrightarrow{|\mathcal{V}|\to\infty} \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}[0]}{max\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}0,1\right\}} = \frac{0}{1} = 0$$

■

4 Experiments

We now proceed to describe a set of experiments that explore the behavior of PATENet under a variety of conditions. Because "ground truth" alignments for real-world OSNs are unavailable, we generated synthetic OSNs for this purpose. Although PATENet has three user-defined parameters, we experimented with different match thresholds, while keeping the other two parameters (a well-defined UNNSM and a signed normalized monotonically increasing transform) constant, as they are more application- and domain-specific.

4.1 Empirical Design

We experimented with PATENet with a substitution matrix based on DeltaCon [12] and a logarithmic signed normalized monotonically increasing transform function ($\ell(x) = 1 - \log_{\alpha}[\alpha^2 + (1-\alpha^2)\cdot x]$ where $\alpha = \frac{\varphi}{1-\varphi}$ for $0.5 < \varphi < 1$). DeltaCon assesses node affinities similarity between two undirected networks with known node correspondence. It is a well-defined UNNSM (by definition).

To examine the robustness of PATENet to noise in the data, we corrupted one of the OSNs - containing otherwise identical subset of (in our experiments with synthetic data, six) elements in the OSNs to be aligned - with different levels of Gaussian noise added to the edge weights. Since PATENet uses a static match threshold, we also examined the interaction between the effect of noise on PATENet's performance and the choice of match threshold φ. We experimented with $\varphi = \{0.51, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90\}$ and Gaussian noise with $\mu = 0$ and $\sigma = \{0.1, 0.3, 0.5, \cdots, 3.9\}$.

Performance of alignment was evaluated using ‘goodness of alignment’, defined as the percentage of elements with known ground-truth match (based on the construction of the OSNs) that were aligned with their ground-truth matches.

4.2 Synthetic Data Generation

We constructed three sets of synthetic data: (1) random dynamic OSNs, (2) Barabasi-Albert (BA) [1] dynamic OSNs, and (3) Dorogovtsev-Mendes (DM) [5] dynamic OSNs. The BA and DM models describe evolving, rather than dynamic, networks, hence we adapted only the edge addition/removal portions of these models. For each dataset we examined three temporal conditions: linear, a single change in trend, and two changes in trend. The resulting OSNs consisted of 25 elements each, starting from an undirected random network with 50 nodes and a connectivity rate of $\sim$0.12 (141 edges out of possible 1225). We use O to denote such a 25-element OSN. We further experimented with the percent of edges added/removed from one element to the next in O, using one of four percentages: 1%, 2%, 4% or 8%.

Random dynamic OSNs were generated as follows: element 1 was generated using the Erdos-Renyi (ER) model [8]. In case of linear Os, edges were added at random to generate elements 2–25 (see Fig. 1A). Single trend change Os were generated by adding edges at random to generate elements 2–13, and then removing edges at random to generate elements 14–25 (see Fig. 1B). For Os with two trend changes, elements 2–9 were generated by adding edges at random, elements 10–17 were generated by removing edges at random, and elements 18–25 were generated by adding edges at random (see Fig. 1C).

BA and DM dynamic OSNs were generated in a manner similar to random dynamic OSNs, with the following changes: the BA model [2] with 50 nodes and $n = 3$ (resulting in 141 edges out of a possible 1225, similarly to the ER-based element 1) was used to generate element 1, and edges to be added were selected based on the corresponding model. Edge removal is done at random according to both models.

In any OSN O, 12 (roughly half) of the elements were selected at random and kept in order to make up a new OSN, denoted by M. Half (six) of the elements selected for M were then removed from O to generate a new OSN with 19 elements, denoted by O'. Consequently, any pair of OSNs (O', M) constructed according to the preceding procedure, shares six random elements, and $M \nsubseteq O'$ and $O' \nsubseteq M$ (see Fig. 1D). Gaussian noise (see Sect. 4.1) was added only to M prior to alignment.

4.3 Results

In all three synthetic datasets experiments revealed a similar relationship between the performance of PATENet, user-specified threshold φ and the added Gaussian noise (see Fig. 2). For lower values of φ, PATENet showed a high degree of noise tolerance, significantly outperforming random alignment over a broad range of Gaussian noise levels. As φ increased, so did PATENet’s susceptibility to noise, but for tolerable levels of noise, its performance was similar or better, as compared to PATENet with lower φ for the same noise level. We conclude that the choice of φ affects multiple aspects of the performance of PATENet in the presence of noise.

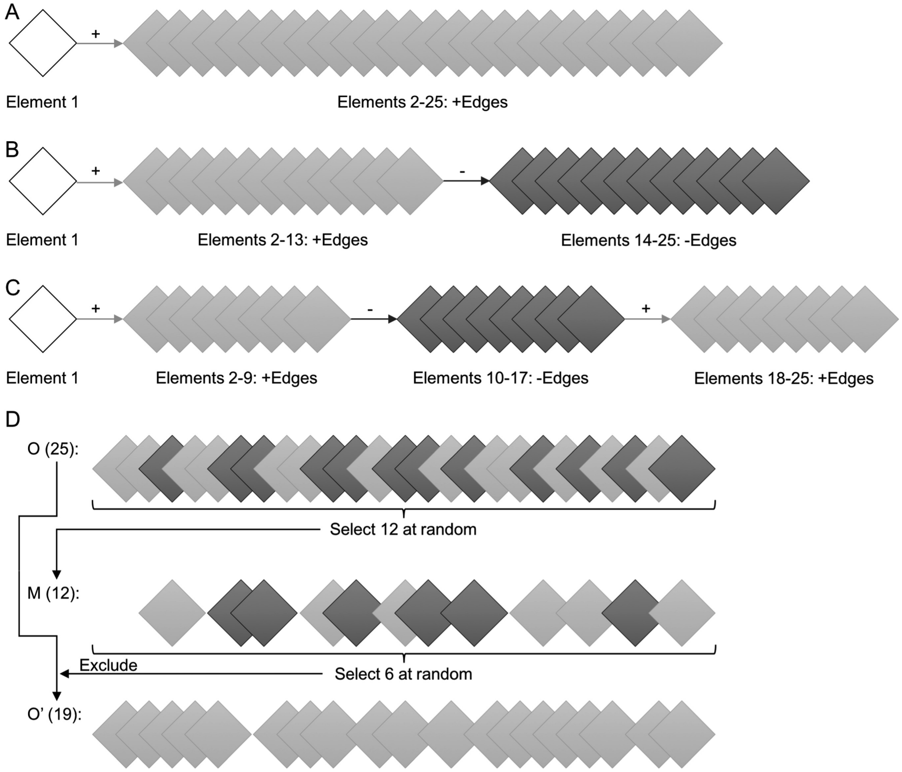

Fig. 1. Synthetic data generation. (A–C) Generation of a linear OSN (A), a single trend change OSN (B) and an OSN with 2 trend changes (C). All OSNs Os (A–C) consist of 25 elements, the first element (white) being a random graph (ER for random dynamics OSNs and BA for BA and DM OSNs). Rectangles represent elements, with light gray indicating increase trend (edges added between elements) and dark gray indicating decrease trend (edges removed between elements); + edges are added between elements; - edges are removed between elements. (D) Generation of OSNs M and O' from OSN O. In dark gray are the elements selected according to the corresponding description in the text. Starting from 25 elements in O, 12 elements are selected at random to create M, six of which are removed from the copy of O to O' (resulting in 19 elements in O').

5 Discussion

5.1 Additional Considerations and Future Directions

In real world OSN data, e.g. those derived from longitudinal studies of functional brain connectivity networks, at present, there are no effective approaches to estimating the noise level in the data. Our results demonstrate a tradeoff between PATENet's resistance to noise and performance with low levels of noise as a function of the choice of match threshold. Hence, in practical settings, it might be worth exploring a probabilistic combination of different match thresholds.

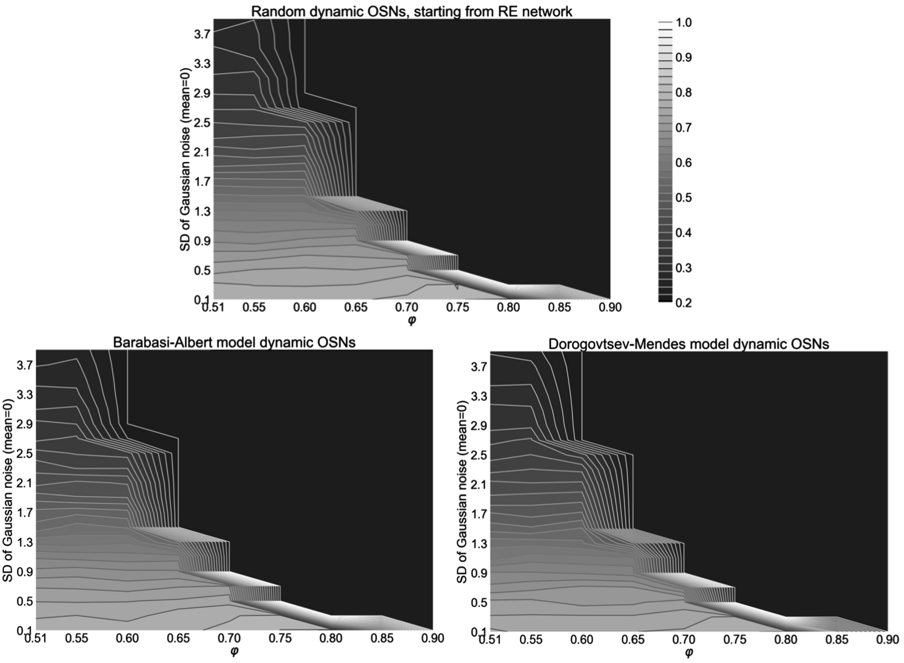

Fig. 2. Effect of noise and match threshold on PATENet's performance. Goodness of alignment of PATENet with the synthetic data as a function of added Gaussian noise (vertical axis) and match threshold φ (horizontal axis). Goodness of alignment was normalized and averaged across all four percentages and three temporal conditions (12 conditions overall), as their patterns were comparable. Top: random dynamic OSNs, starting from RE network. Bottom left: BA dynamic OSNs. Bottom right: DM dynamic OSNs. The same color bar is used in all three plots, ranging from the average performance of random alignment (comparable in all three datasets) to perfect alignment (1.0).

Some natural directions include PATENet as an OSN kernel, to use in classification and regression problems where the input to the classifier is an OSN. Possible applications include assigning subjects to different categories (e.g. normal development, accelerated development, retarded development) based on the observed development from longitudinal studies. Another natural direction for future work is to extend PATENet to align multiple OSNs (as opposed to a pair of OSNs). The resulting multi-sequence variant of PATENet can also be used to cluster OSNs.

5.2 Generalizations

The empirically-demonstrated version of PATENet is limited to the case where the elements of the OSNs are undirected networks with pre-specified correspondence between nodes in each element of one OSN and nodes in each element of any other OSN to be aligned with it. It would be interesting to explore variants of PATENet that can work with OSNs consisting of directed graphs, graphs with both directed and undirected edges, or colored graphs (with multiple types of nodes and/or edges), etc. It would also be interesting to consider variants of PATENet that can work in settings where the correspondence between nodes in each element of one OSN and nodes in each element of any other OSN to be aligned with it is not specified, but instead needs to be established based on some node similarity criteria [18].

Furthermore, while in this paper we have focused on the pairwise alignment of OSNs, the PATENet algorithm can be further generalized to work with ordered sequences of arbitrary elements (instead of networks) so long as we can specify a well-behaved unsigned normalized similarity measure between such elements.

6 Conclusion

Networks that change over time, e.g. functional brain networks that change their structure due to processes such as development or aging, are naturally modeled by TENs. Longitudinal measurements of such TENs are naturally represented as OSNs, where each network in the sequence represents a static snapshot of the TEN at a specific time of observation. In this paper we proposed PATENet, a novel algorithm for optimal local alignment of a pair of OSNs. The algorithm requires three user-defined inputs in addition to a pair of OSNs to be aligned: a well-defined UNSSM, a signed normalized monotonically increasing transform, and a match threshold. We showed how PATENet can be used to compute an alignment score, as well as a similarity score, for a pair of aligned OSNs.

Our experiments using PATENet to align synthetic OSNs produced using different generative models of OSNs with their noise corrupted counterparts show that: at lower match thresholds, PATENet displays a high degree of noise tolerance, significantly outperforming random alignment over a broad range of noise levels; at higher match thresholds (more stringent match criteria), PATENet shows increased susceptibility to noise.

PATENet offers a mathematically sound approach to aligning OSNs, which is amenable to being generalized along a number of dimensions, e.g. OSNs consisting of directed networks, labeled networks, or even ordered sequences of other types of elements.

Acknowledgments. This project was supported in part by the National Center for Advancing Translational Sciences, National Institutes of Health through Grant UL1 TR000127 and TR002014, by the National Science Foundation, through Grant SHF 1518732, the Center for Big Data Analytics and Discovery Informatics at Pennsylvania State University, the Edward Frymoyer Endowed Professorship in Information Sciences and Technology at Pennsylvania State University and the Sudha Murty Distinguished Visiting Chair in Neurocomputing and Data Science funded by the Pratiksha Trust at the Indian Institute of Science [both held by Vasant Honavar]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors. We thank Sanghack Lee for helpful discussions during the course of this work.

References

1. Albert, R., Barabasi, A.L.: Topology of evolving networks: local events and universality. Phys. Rev. Lett. **85**(24), 5234–5237 (2000). https://doi.org/10.1103/PhysRevLett.85.5234
2. Barabasi, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999). https://doi.org/10.1126/science.286.5439.509
3. Bassett, D.S., Sporns, O.: Network neuroscience. Nat. Neurosci. **20**(3), 353–364 (2017). https://doi.org/10.1038/nn.4502
4. Caspi, Y., Irani, M.: Spatio-temporal alignment of sequences. IEEE Trans. Pat. Anal. Mach. Int. **24**(11), 1409–1424 (2002). https://doi.org/10.1109/TPAMI.2002.1046148
5. Dorogovtsev, S.N., Mendes, J.F.F.: Scaling behaviour of developing and decaying networks. Europhys. Lett. **52**(1), 33–39 (2000). https://doi.org/10.1209/epl/i2000-00400-0
6. Elzinga, C.H.: Distance, similarity and sequence comparison. In: Blanchard, P., Bühlmann, F., Gauthier, J.A. (eds.) Advances in Sequence Analysis: Theory, Method, Applications. LCRSP, vol. 2, pp. 51–73. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-04969-4_4
7. Emmert-Streib, F., Dehmer, M., Shi, Y.: Fifty years of graph matching, network alignment and network comparison. Inf. Sci. **346**, 180–197 (2016). https://doi.org/10.1016/j.ins.2016.01.074
8. Erdos, P., Renyi, A.: On random graphs I. Publ. Math. Debrecen **6**, 290–297 (1959)
9. Holme, P.: Modern temporal network theory: a colloquium. Eur. Phys. J. B **88**(9), 234–263 (2015). https://doi.org/10.1140/epjb/e2015-60657-4
10. Holme, P., Saramki, J.: Temporal networks. Phys. Rep. **519**(3), 97–125 (2012). https://doi.org/10.1016/j.physrep.2012.03.001
11. Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A.: Multilayer networks. J. Complex Netw. **2**(3), 203–271 (2014). https://doi.org/10.1093/comnet/cnu016
12. Koutra, D., Vogelstein, J.T., Faloutsos, C.: DELTACON: a principled massive-graph similarity function. In: Proceedings of the 2013 SIAM International Conference on Data Mining, pp. 162–170. SIAM (2013). https://doi.org/10.1137/1.9781611972832.18
13. Lee, W.N., Das, A.K.: Local alignment tool for clinical history: temporal semantic search of clinical databases. In: AMIA Annual Symposium Proceedings, pp. 437–441 (2010)
14. Li, A., Cornelius, S.P., Liu, Y.Y., Wang, L., Barabasi, A.L.: The fundamental advantages of temporal networks. Science **358**(6366), 1042–1046 (2017). https://doi.org/10.1126/science.aai7488
15. Luo, G., Cordier, F., Seo, H.: Spatio-temporal segmentation for the similarity measurement of deforming meshes. Vis. Comput. **32**(2), 243–256 (2016). https://doi.org/10.1007/s00371-015-1178-8
16. Madhyastha, T., Peverill, M., Koh, N., McCabe, C., Flournoy, J., Mills, K., King, K., Pfeifer, J., McLaughlin, K.A.: Current methods and limitations for longitudinal fMRI analysis across development. Dev. Cogn. Neurosci. (2017). https://doi.org/10.1016/j.dcn.2017.11.006
17. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. **147**(1), 195–197 (1981). https://doi.org/10.1016/0022-2836(81)90087-5
18. Towfic, F., Greenlee, M.H.W., Honavar, V.: Aligning biomolecular networks using modular graph kernels. In: Salzberg, S.L., Warnow, T. (eds.) WABI 2009. LNCS, vol. 5724, pp. 345–361. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04241-6_29