DISTANCES BETWEEN DIRECTED NETWORKS AND APPLICATIONS

Samir Chowdhury and Facundo Mémoli

Department of Mathematics,
The Ohio State University,
Columbus, OH 43210,
USA

ABSTRACT

Networks which show the relationships within and between complex systems are key tools in a variety of current scientific areas. A central aim in network analysis is to find a suitable metric for network similarity and comparison. We propose a definition for the space of all networks, and show that our definition leads to a natural and meaningful notion of distance between networks. We discuss the computational complexity involved in computing our network distance, and develop lower bounds by using invariants of networks that are significantly simpler to compute. By constructing a wide range of explicit examples, we show that these lower bounds are effective in distinguishing between networks. We describe multiple invariants and prove that all of them are stable in a quantitative sense.

Index Terms— Networks, motifs, metrics, distances, network signatures

1. INTRODUCTION

Networks which show the relationships within and between complex systems are key tools in a variety of current research areas. In the domain of bioinformatics, networks have been used to represent molecular activity [1], metabolic pathways [2], functional relations between enzyme clusters [3] and genetic regulation [4, 5]. Networks have been used as a natural tool for representing brain anatomy and function [6, 7]. Network-based methods also appear in data mining [8], where the goal is to extract patterns or substructures that appear with higher frequency than in a randomized network [9, 10, 11]. Whereas the aforementioned networks are typically studied as static objects, social networks [12, 13] and the World Wide Web [14, 15] are examples of dynamically evolving networks, and have also been studied extensively [16, 17].

A central aim in network analysis is to find a suitable metric for network similarity. Such a metric should be able to compare different networks and test for structural similarity,

and should be able to do so with reasonably low computational complexity. Throughout this paper, we aim to contribute to the body of mathematical tools for parsing network data. Related results in this direction include the framework of graph kernels, which may be computed between the nodes of a single graph [18] or between graphs [19]. More details can be found in [20]. Approaches using graph edit distance and Levenshtein distance are discussed in [21].

Our goals are to contribute the following: (1) a notion of distance between directed, weighted networks, (2) a list of network signatures, or *invariants*, that can be computed easily and used to test for dissimilarity between networks, and (3) efficiently computable lower bounds for the full notion of distance, based on these invariants.

A full version of the results announced here will appear elsewhere. Some complementary results appear in [22].

2. THE CONSTRUCTION OF THE METRIC

A recurring notion in our paper is that of the Hausdorff distance between two subsets of a metric space. Typically we will be comparing (finite) subsets of the real line, denoted $\mathbb{R}$. In this case, we denote the Hausdorff distance between $X, Y \subseteq \mathbb{R}$ by

$$d_{\mathcal{H}}^{\mathbb{R}}(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} |x - y|, \max_{y \in Y} \min_{x \in X} |x - y| \right\}.$$

For any set S we write $\mathrm{card}(S)$ for the cardinality of S. We denote by $F(S)$ the set of all finite subsets of S. Finally, we will denote the non-negative real numbers by $\mathbb{R}_+$.

Let X be a finite set, and let ω_X be a function from $X \times X$ to $\mathbb{R}$. By a 1-network, we will mean a pair (X, ω_X). We will denote the collection of all networks by $\mathcal{N}$.

We will often refer to the points of X as *nodes* and ω_X as the *weight function* of X. Pairs of nodes will often be called *edges*. The information contained in a network should be preserved when we relabel the nodes in a compatible way; we formalize this idea by introducing the notion of *isomorphism* of networks. To say (X, ω_X) and (Y, ω_Y) are *isomorphic* means that there exists a bijection $\varphi : X \to Y$ such that
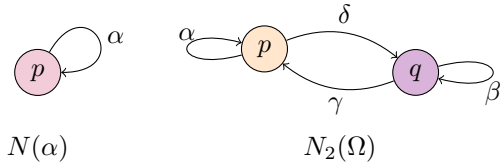
Fig. 1. Simple networks.

$\omega_X(x,x') = \omega_Y(\varphi(x), \varphi(x'))$ for all $x, x' \in X$. We will denote an isomorphism between networks by $X \cong Y$.

Isomorphism aims to capture the notion that a network with n nodes is always distinguishable from a network with a different number of nodes. For networks of the same number of nodes, isomorphism requires that the weights be respected as well.

We provide several examples to illustrate our definitions.

Example 1. *Networks with one or two nodes will be very instructive in providing examples and counterexamples, so we introduce them now with some special terminology.*

- *Networks with one node: any network with exactly one node p can be specified by a number $\alpha \geqslant 0$. Since 1-networks with one node are very special objects, we reserve the notation $N(\alpha)$ for any of them. From our definition, it follows that networks $N(\alpha)$ and $N(\alpha')$ are isomorphic if and only if $\alpha = \alpha'$.*

- *Given 4 real numbers α, β, δ and γ let $\Omega = \left(\begin{smallmatrix} \alpha & \delta \\ \gamma & \beta \end{smallmatrix}\right)$. This matrix induces a 1-network $N_2(\Omega)$ over two nodes as in Figure 1. Any network with two nodes can be represented in this manner. Notice that networks $N(\alpha) \not\cong N_2(\Omega)$ since they have different numbers of nodes. By Proposition 4, it will turn out that given $\Omega, \Omega' \in \mathbb{R}^{2\times 2}$, $N_2(\Omega) \cong N_2(\Omega')$ if and only if there exists a permutation matrix P of size 2×2 such that $\Omega' = P\,\Omega\,P^T$.*

- *Any k-by-k matrix $\Sigma \in \mathbb{R}^{k\times k}$ induces a network on k nodes, which we refer to as $N_k(\Sigma)$. Notice that $N_k(\Sigma) \cong N_\ell(\Sigma')$ if and only if $k = \ell$ and there exists a permutation matrix P of size k such that $\Sigma' = P\,\Sigma\,P^T$.*

We wish to define a notion of distance on $\mathcal{N}$ that is compatible with isomorphism. A natural analog is the Gromov-Hausdorff distance defined between metric spaces [23]. To adapt that definition for our needs, we first introduce the definition of a correspondence.

2.1. Correspondences and the distortion map

Let $(X, \omega_X), (Y, \omega_Y)$ be two networks. A correspondence between these two networks is a set $R \subseteq X \times Y$ such that for each $x \in X$, there exists $y \in Y$ such that $(x, y) \in R$, and for each $y \in Y$, there exists $x \in X$ such that $(x, y) \in R$. The collection of all correspondences between X and Y will be denoted $\mathcal{R}(X, Y)$, abbreviated to $\mathcal{R}$ when the context is clear.

Example 2 (1-point correspondence). *Let X be a set, and let $\{p\}$ be the set with one point. Then there is a unique correspondence $R = \{(x, p) : x \in X\}$ between X and $\{p\}$.*

Let $(X, \omega_X), (Y, \omega_Y)$ be two networks and let $R \in \mathcal{R}$. The 1-distortion of R is given by

$$\text{dis}(R) := \max_{(x,y),(x',y')\in R} |\omega_X(x,x') - \omega_Y(y,y')|.$$

Then we define $d_{\mathcal{N}}(X, Y) := \frac{1}{2}\min_{R\in\mathcal{R}}\text{dis}(R)$.

Remark 3. *Two simple but important remarks are the following: (1) for any $X, Y \in \mathcal{N}$, $\mathcal{R}(X, Y) \neq \varnothing$, and (2) $d_{\mathcal{N}}(X, Y)$ is always bounded. Indeed, $X \times Y$ is always a valid correspondence between X and Y. So we have*

$$d_{\mathcal{N}}(X, Y) \leqslant \frac{1}{2}\text{dis}(X \times Y)$$
$$\leqslant \frac{1}{2}\left(\max_{x,x'}\omega_X(x,x') + \max_{y,y'\in Y}\omega_Y(y,y')\right) < \infty.$$

When both X and Y are networks with the same cardinality, we can use an alternate formulation of distance. Suppose $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ such that $\text{card}(X) = \text{card}(Y)$. Then define

$$\overline{d_{\mathcal{N}}}(X, Y) := \frac{1}{2}\min_{\varphi}\max_{x,x'\in X}\left|\omega_X(x,x') - \omega_Y\left(\varphi(x), \varphi(x')\right)\right|,$$

where $\varphi : X \to Y$ ranges over bijections from X to Y.

A natural question is whether $\overline{d_{\mathcal{N}}}$ and $d_{\mathcal{N}}$ agree on pairs of networks with the same cardinality. It turns out that $d_{\mathcal{N}}$ and $\overline{d_{\mathcal{N}}}$ agree on pairs of networks with two nodes. In general, these notions are different, as Remark 5 shows.

Proposition 4. *Suppose $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ such that $\text{card}(X) = \text{card}(Y) = 2$. Then we have $d_{\mathcal{N}}(X, Y) = \overline{d_{\mathcal{N}}}(X, Y)$. Furthermore, if $X = N_2(\left(\begin{smallmatrix} \alpha & \delta \\ \beta & \gamma \end{smallmatrix}\right))$ and $Y = N_2(\left(\begin{smallmatrix} \alpha' & \delta' \\ \beta' & \gamma' \end{smallmatrix}\right))$, then we have the explicit formula*

$$d_{\mathcal{N}}(X, Y) = \frac{1}{2}\min\left(\Gamma_1, \Gamma_2\right), \text{ where}$$
$$\Gamma_1 = \max\left(|\alpha - \alpha'|, |\beta - \beta'|, |\delta - \delta'|, |\gamma - \gamma'|\right),$$
$$\Gamma_2 = \max\left(|\alpha - \gamma'|, |\gamma - \alpha'|, |\delta - \beta'|, |\beta - \delta'|\right).$$

Remark 5. (The case $\text{card}(X) = \text{card}(Y)$). *Assume (X, ω_X) and (Y, ω_Y) are two 1-networks with the same cardinality. Then we have $d_{\mathcal{N}}(X, Y) \leqslant \overline{d_{\mathcal{N}}}(X, Y)$, but the inequality may be strict.*

It turns out that $\overline{d_{\mathcal{N}}}$ yields a legitimate notion of distance which is compatible with isomorphism between networks of the same cardinality. Because we would like to compare networks of different cardinalities, our main object of study is $d_{\mathcal{N}}$. The definition of $d_{\mathcal{N}}$ is sensible in the sense that it captures the notion of a distance:

6421

Fig. 2. The $d_{\mathcal{N}}$ distance between two one-node networks is simply $\frac{1}{2}|\alpha - \alpha'|$.

Theorem 6. *$d_{\mathcal{N}}$ is a pseudo-metric on $\mathcal{N}$ modulo isomorphism.*

Remark 7. *Notice that despite assuring that $X \cong Y$ implies that $d_{\mathcal{N}}(X,Y) = 0$, the theorem above does not preclude the possibility that there exist $X, Y \in \mathcal{N}$ non-isomorphic for which $d_{\mathcal{N}}(X,Y) = 0$. Consider the networks $N = N_1(1)$ and $N_2 = N_2(\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right))$ from Example 1. The two networks have different cardinalities, so they are not isomorphic. However, we claim that $d_{\mathcal{N}}(N, N_2) = 0$. To see this, note that by Example 2, the only correspondence is $\{(p,q),(p,r)\}$. Since $\omega_{N_2}(q,r) = \omega_{N_2}(r,q)$, we get $d_{\mathcal{N}}(N, N_2) = \frac{1}{2}|\omega_N(p,p) - \omega_{N_2}(q,r)| = \frac{1}{2}|1 - 1| = 0$. Notice that this example can be generalized in the following way: let $N_k = N_k(\mathbb{1}_{k \times k})$ be the network on k nodes with each edge weight equal to 1. Then $d_{\mathcal{N}}(N, N_k) = 0$ for all $k \in \mathbb{N}$, but clearly $N \ncong N_k$ whenever $k \geqslant 2$.*

Example 8. *Now we give some examples.*

- *For $\alpha, \alpha' \geqslant 0$ consider two 1-networks with one node each: $N(\alpha) = (\{p\}, \alpha)$ and $N(\alpha') = (\{p'\}, \alpha')$. Then by Example 2 there is only one correspondence, $R = \{(p,p')\}$ between these two networks, so that $dis(R) = |\alpha - \alpha'|$ and as a result $d_{\mathcal{N}}(N(\alpha), N(\alpha')) = \frac{1}{2}|\alpha - \alpha'|$.*

- *Let $(X, \omega_X) \in \mathcal{N}$ be any network and let $N(\alpha) = (\{p\}, \alpha)$ be a network with just one node. Then, $d_{\mathcal{N}}(X, N(\alpha)) = \frac{1}{2} \max_{x,x' \in X} |\omega_X(x,x') - \alpha|$. Indeed, notice that as in the previous example, there exists a unique correspondence between X and $\{p\}$: the correspondence $R = \{(x,p), \ x \in X\}$. The distortion of R is the claimed quantity.*

The natural question now is the following: How do we extend these examples by computing $d_{\mathcal{N}}$ for large datasets? In practical applications we are interested in deciding when two networks are similar or different. This can be done using upper or lower bounds for $d_{\mathcal{N}}$, so we should try to judge the complexity of such calculations.

By Remark 5, we know that it is possible to obtain an upper bound on $d_{\mathcal{N}}$, in the case $\text{card}(X) = \text{card}(Y)$, by using $\overline{d_{\mathcal{N}}}$. Solving for $\overline{d_{\mathcal{N}}}(X,Y)$ reduces to minimizing the function $\max_{x,x' \in X} f(\varphi)$ over all bijections φ from X to Y. Here $f(\varphi) := \max_{x,x'} |\omega_X(x,x') - \omega_Y(\varphi(x), \varphi(x'))|$. However, this is an instance of an NP-hard problem known as the *bottleneck quadratic assignment problem* [24]. The structure of the

optimization problem induced by $d_{\mathcal{N}}$ is very similar to that of $\overline{d_{\mathcal{N}}}$ so it seems plausible that computing $d_{\mathcal{N}}$ would lead to NP-hard problems as well.

As such, one is confronted with a complicated computational problem. The next recourse is to find lower bounds for $d_{\mathcal{N}}(X,Y)$ that can be calculated in polynomial time. This will lead to the discussions in §2.2 and §2.3.

2.2. Invariants of networks

Often one wants to extract information out of networks in a perhaps lossy way. An extreme example would be to represent a given 1-network by a single real number. Intuitively, the number that we associate to two isomorphic networks should be the same. We define an $\mathbb{R}$-invariant of 1-networks to be a map $\iota : \mathcal{N} \to \mathbb{R}$ such that for any $X, Y \in \mathcal{N}$, if $X \cong Y$ then $\iota(X) = \iota(Y)$. In what follows, we will construct several maps and claim that they are invariants.

Example 9. *Define the diameter map to be the map* $\text{diam} : \mathcal{N} \to \mathbb{R}$ *given by* $(X, \omega_X) \mapsto \max_{x,x' \in X} |\omega_X(x,x')|$. *Then* diam *is an $\mathbb{R}$-invariant. An application of* diam *to Example 8 gives an upper bound on $d_{\mathcal{N}}(X,Y)$ in the following way:*

$$d_{\mathcal{N}}(X,Y) \leqslant d_{\mathcal{N}}(X, N(0)) + d_{\mathcal{N}}(N(0), Y)$$
$$= \frac{1}{2}(\text{diam}(X) + \text{diam}(Y)).$$

We will eventually state that our proposed invariants are *quantitatively stable*. This notion is made precise in §2.3 but for now, we introduce the terminology of *metric space valued invariants*. Let (V, d_V) be any metric space. A V-valued invariant is any map $\iota : \mathcal{N} \to V$ such that $\iota(X, \omega_X) = \iota(Y, \omega_Y)$ whenever $X \cong Y$. So diam is our first example of a metric space valued invariant for $V = \mathbb{R}$.

Consider the following construction: to each network (X, ω_X) assign the set $\text{spec}(X) := \{\omega_X(x,x') : x, x' \in X\}$. This set is called the *spectrum* of the network X. So for the networks in Example 1 we have $\text{spec}(N(\alpha)) = \{\alpha\}$, and $\text{spec}(N_2(\Omega)) = \{\alpha, \beta, \gamma, \delta\}$. This spec map is an invariant. Notice that spec is a map from $\mathcal{N}$ into finite subsets of $\mathbb{R}$, denoted $F(\mathbb{R})$. Since $F(\mathbb{R})$ can be regarded as a metric space by endowing it with the Hausdorff distance, spec is an example of a metric space valued invariant.

Another important construction is the following one which *localizes* spec: let $(X, \omega_X) \in \mathcal{N}$ and $x \in X$, and define $\text{spec}_X^{\text{out}}(x) := \{\omega_X(x,x') : x' \in X\} \subset \mathbb{R}$. For example, for the network $N = N_2(\Omega)$ in Figure 1, we have $\text{spec}_N^{\text{out}}(p) = \{\alpha, \delta\}$ and $\text{spec}_N^{\text{out}}(q) = \{\beta, \gamma\}$. Notice that $\text{spec}(X) = \bigcup_{x \in X} \text{spec}_X^{\text{out}}(x)$ for any network X, thus justifying the claim that this construction localizes spec.

Similarly, we define $\text{spec}_X^{\text{in}}(x) := \{\omega_X(x',x) : x' \in X\}$. Notice that one still has $\text{spec}(X) = \bigcup_{x \in X} \text{spec}_X^{\text{in}}(x)$ for any network X. The two local versions of spec do not necessarily coincide, as will be shown in the next example. Regardless,
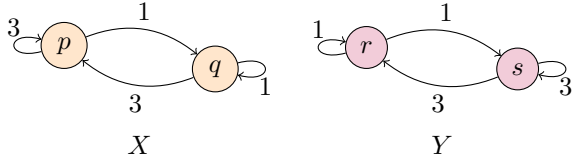
Fig. 3. Local spectra can distinguish between networks where global spectra fails.

Fig. 4. Lower-bounding $d_{\mathcal{N}}$ by using global spectra.

it turns out that the global and local versions of spec are all proper invariants of 1-networks.

Example 10. *Here is an application of the localized version of spec. Consider the networks in Figure 3. Notice that* $\mathrm{spec}(X) = \mathrm{spec}(Y) = \{1, 3\}$. *In particular,* $\mathrm{spec}_X^{\mathrm{out}}(p) = \{1, 3\}$ *and* $\mathrm{spec}_X^{\mathrm{out}}(q) = \{1, 3\}$. *On the other hand,* $\mathrm{spec}_Y^{\mathrm{out}}(r) = \{1\}$ *and* $\mathrm{spec}_Y^{\mathrm{out}}(s) = \{3\}$. *Thus* $X \not\cong Y$. *Notice that we also have* $\mathrm{spec}_X^{\mathrm{in}}(p) = \{3\}$, $\mathrm{spec}_X^{\mathrm{in}}(q) = \{1\}$ *and* $\mathrm{spec}_Y^{\mathrm{in}}(r) = \mathrm{spec}_Y^{\mathrm{in}}(s) = \{1, 3\}$. *So as invariants,* $\mathrm{spec}^{\mathrm{in}}$ *and* $\mathrm{spec}^{\mathrm{out}}$ *turn out to be independent.*

We emphasize that metric space valued invariants are useful because the *stability* of these invariants is easy to formalize.

2.3. Quantitative stability of invariants of networks

Let (V, d_V) be a given pseudo-metric space. The V-valued invariant $\iota : \mathcal{N} \to V$ is said to be *quantitatively stable* if there exists a constant $L > 0$ such that

$$d_V\big(\iota(X), \iota(Y)\big) \leqslant L \cdot d_{\mathcal{N}}(X, Y)$$

for all 1-networks X and Y. The least constant L such that the above holds for all $X, Y \in \mathcal{N}$ is the Lipschitz constant of ι and is denoted $\mathbf{L}(\iota)$.

Note that by identifying a non-constant quantitatively stable V-valued invariant ι, we immediately obtain a lower bound for the $d_{\mathcal{N}}$ distance between any two 1-networks (X, ω_X) and (Y, ω_Y). Furthermore, given a finite family $\iota_\alpha : \mathcal{N} \to V$, $\alpha \in A$, of non-constant quantitatively stable invariants, we may obtain the following lower bound for the distance between networks X and Y:

$$\big(\max_{\alpha \in A} \mathbf{L}(\iota_\alpha)\big)^{-1} \max_{\alpha \in A} d_V(\iota_\alpha(X), \iota_\alpha(Y)) \leqslant d_{\mathcal{N}}(X, Y).$$

It is often the case that computing $d_V(\iota(X), \iota(Y))$ is substantially simpler than computing the $d_{\mathcal{N}}$ distance between X and Y (which leads to a possibly NP-hard problem).

Example 11. *It is easy to verify that* $\mathbf{L}(\mathrm{diam}) = 2$, *thus, for all networks X and Y, we have* $d_{\mathcal{N}}(X, Y) \geqslant \frac{1}{2}\big|\mathrm{diam}(X) - \mathrm{diam}(Y)\big|$. *For example, for the networks* $X = N_2(\left(\begin{smallmatrix} 1 & 5 \\ 2 & 4 \end{smallmatrix}\right))$ *and* $Y = N_k(\mathbb{1}_{k \times k})$ *(the all-ones matrix) we have* $d_{\mathcal{N}}(X, Y) \geqslant \frac{1}{2}|5 - 1| = 2$, *for all* $k \in \mathbb{N}$.
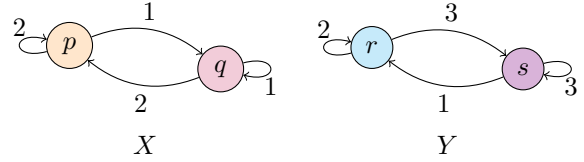
Now we state a proposition regarding the stability of global and local spectrum invariants. These will be of particular interest for computational purposes.

Proposition 12. *Let* $\mathrm{spec}_{\bullet}$ *refer to either the* out *or* in *version of local spectrum. Then, for all* $X, Y \in \mathcal{N}$ *we have*

$$d_{\mathcal{N}}(X, Y) \geqslant \frac{1}{2} \min_{R \in \mathcal{R}} \max_{(x,y) \in R} d_{\mathcal{H}}^{\mathbb{R}}(\mathrm{spec}_X(x), \mathrm{spec}_Y(y))$$
$$\geqslant \frac{1}{2} d_{\mathcal{H}}^{\mathbb{R}}(\mathrm{spec}(X), \mathrm{spec}(Y)).$$

Corollary 13. $\mathbf{L}(\mathrm{spec}_{\bullet}) = \mathbf{L}(\mathrm{spec}) = 2$. *Here* $\mathrm{spec}_{\bullet}$ *denotes either of the local versions of* spec.

Example 14. *Consider the networks in Figure 4. By Corollary 13, we may calculate a lower bound for* $d_{\mathcal{N}}(X, Y)$ *by simply computing the Hausdorff distance between* $\mathrm{spec}(X)$ *and* $\mathrm{spec}(Y)$, *and dividing by 2. In this example,* $\mathrm{spec}(X) = \{1, 2\}$ *and* $\mathrm{spec}(Y) = \{1, 2, 3\}$. *Thus* $d_{\mathcal{H}}^{\mathbb{R}}(\mathrm{spec}(X), \mathrm{spec}(Y)) = 1$, *and* $d_{\mathcal{N}}(X, Y) \geqslant \frac{1}{2}$.

Computing the lower bound involving local spectra requires solving a bottleneck linear assignment problem over the set of all correspondences between X and Y. This can be solved in polynomial time. The second lower bound stipulates computing the Hausdorff distance on $\mathbb{R}$ between the (global) spectra of X and Y – a computation which can be carried out in (smaller) polynomial time as well.

3. DISCUSSION

We introduced a model for the space of all networks, and defined a notion of isomorphism between any two networks. Then we proposed a notion of distance compatible with this notion of isomorphism and verified that our definition actually does induce a pseudometric. Next, we took on the practical question of estimating and computing this distance. Direct computation of $d_{\mathcal{N}}(X, Y)$ appears to lead to NP-hard problems in general.

We constructed multiple quantitatively stable invariants, with examples illustrating their behavior, and quantified their stability. Even though not expounded in this paper, there exists an algorithm of complexity $O(n^2 \times m^2)$ that uses local spectra to compute a lower bound for $d_{\mathcal{N}}$.

6423

4. REFERENCES

[1] Roded Sharan and Trey Ideker, "Modeling cellular machinery through biological network comparison," *Nature biotechnology*, vol. 24, no. 4, pp. 427–433, 2006.

[2] Ron Y Pinter, Oleg Rokhlenko, Esti Yeger-Lotem, and Michal Ziv-Ukelson, "Alignment of metabolic pathways," *Bioinformatics*, vol. 21, no. 16, pp. 3401–3408, 2005.

[3] Hiroyuki Ogata, Wataru Fujibuchi, Susumu Goto, and Minoru Kanehisa, "A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters," *Nucleic acids research*, vol. 28, no. 20, pp. 4021–4028, 2000.

[4] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon, "Network motifs in the transcriptional regulation network of escherichia coli," *Nature genetics*, vol. 31, no. 1, pp. 64–68, 2002.

[5] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon, "Network motifs: simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[6] Olaf Sporns, *Networks of the Brain*, MIT press, 2011.

[7] Mikail Rubinov and Olaf Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[8] Takashi Washio and Hiroshi Motoda, "State of the art of graph-based data mining," *Acm Sigkdd Explorations Newsletter*, vol. 5, no. 1, pp. 59–68, 2003.

[9] Diane J Cook and Lawrence B Holder, "Substructure discovery using minimum description length and background knowledge," *Journal of Artificial Intelligence Research*, pp. 231–255, 1994.

[10] Akihiro Inokuchi, Takashi Washio, and Hiroshi Motoda, "Complete mining of frequent patterns from graphs: Mining graph data," *Machine Learning*, vol. 50, no. 3, pp. 321–354, 2003.

[11] Kenichi Yoshida, Hiroshi Motoda, and Nitin Indurkhya, "Graph-based induction as a unified learning framework," *Applied Intelligence*, vol. 4, no. 3, pp. 297–316, 1994.

[12] Ravi Kumar, Jasmine Novak, and Andrew Tomkins, "Structure and evolution of online social networks," in *Link mining: models, algorithms, and applications*, pp. 337–357. Springer, 2010.

[13] David Easley and Jon Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, New York, NY, USA, 2010.

[14] Dennis Fetterly, Mark Manasse, Marc Najork, and Janet Wiener, "A large-scale study of the evolution of web pages," in *Proceedings of the 12th international conference on World Wide Web*. ACM, 2003, pp. 669–678.

[15] Alexandros Ntoulas, Junghoo Cho, and Christopher Olston, "What's new on the web?: the evolution of the web from a search engine perspective," in *Proceedings of the 13th international conference on World Wide Web*. ACM, 2004, pp. 1–12.

[16] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005, pp. 177–187.

[17] Mark EJ Newman, "The structure and function of complex networks," *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

[18] Risi Kondor and John Lafferty, "Diffusion kernels on graphs and other discrete input spaces," in *ICML*, 2002, vol. 2, pp. 315–322.

[19] Alexander J Smola and Risi Kondor, "Kernels and regularization on graphs," in *Learning theory and kernel machines*, pp. 144–158. Springer, 2003.

[20] S V N Vishwanathan, Nicol N Schraudolph, Risi Kondor, and Karsten M Borgwardt, "Graph kernels," *The Journal of Machine Learning Research*, vol. 11, pp. 1201–1242, 2010.

[21] Xinbo Gao, Bing Xiao, Dacheng Tao, and Xuelong Li, "A survey of graph edit distance," *Pattern Analysis and applications*, vol. 13, no. 1, pp. 113–129, 2010.

[22] Samir Chowdhury and Facundo Mémoli, "Metric structures on networks and applications," *2015 53nd Annual Allerton Conference on Communication, Control, and Computing*, to appear.

[23] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Birkhäuser Boston Inc., 1999.

[24] Panos M Pardalos and Henry Wolkowicz, Eds., *Quadratic assignment and related problems*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 16. American Mathematical Society, Providence, RI, 1994.