

# Technical Report

November 30, 2017

## Outline

### 1 Diffusion on networks

#### 1.1 Motivation

Diffusion on networks is one of the dynamic processes that occur on networks. Diffusion processes on networks are used to develop models for the study of real-world processes such as the spread of infections with in a group of people, spread of information in a social network, studying failures in power supply on the grid with in an area, etc.

#### 1.2 Ideas

- a) Diffusion on networks through direct interactions
- b) Accounting for long-range interaction in diffusion process on network
- c) k-path Laplacians, that is, the generalised Laplacian matrices. Considerations of Mellin and Laplace transforms of the k-path Laplacians to account for longrange interactions
- d) Illustration of Diffusion of heat on a lattice. First, considering direct interactions only and then accounting both direct and long-range interactions.

### 2 Image Segmentation

#### 2.1 Motivation

Image Segmentation is the problem of localising regions of an image relative to content. It aids in the extraction of objects of interest from an image. Image segmentation has applications in medical field, engineering field, computer science among others. A number of segmentation techniques have been developed based on graph theory concepts which include intelligent scissors, Normalised cut algorithm, random walker algorithm, etc. Our interest is in the segmentation of images using the random walk process on graph as put forward by Grady.

## 2.2 Ideas

- a) To extend the idea of image segmentation using random walks on networks by accounting for long-range links in the network. Our aim is to ascertain whether a better image segmentation can be obtained on including long-range links.

# 3 Systemic Risk and Contagion in Financial Systems

## 3.1 Motivation

The financial crisis of 2008 stimulated intense research in the area of financial systems. The use of complex networks in the study of financial systems is one of the avenues that have been used by a number of researchers in uncovering the root causes of the crisis, which financial institutions sparked off the cascaded failure leading to system break down, the impact of the network structure, strategies to either prevent the re-occurrence or to guarantee minimal impact of the crisis etc.

## 3.2 Ideas

- a) The concept of "too interconnected to fail", i.e. which measure (centrality) can we use to identify nodes that are considered too interconnected to fail.
- b) Can the centrality based on generalised degree centrality capture the node which is too interconnected to fail?
- c) How can we apply the concept of long-range interactions to financial networks? What does the long-range interactions mean?

# 4 Laplacian centrality of an edge

## 4.1 Motivation

Centrality measures in networks have proved to be relevant tools in network analysis. They are indicators of the 'importance' of nodes and edges in the networks. Though most work has been geared towards the study of importance of nodes (i.e. degree, closeness, betweenness, subgraph, eigenvector, Laplacian centralities, etc.), interest in the study of edge centralities is now gaining ground with prominent works of Girvan and Newman. Some of the known edge centralities include edge-betweenness, k-path edge centrality, among others. The motivation for the introduction of edge centrality measures lies in real-world applications in a wide range of context such as in identifying communities in networks, identifying strong relationships among people in social networks, etc.

## 4.2 Ideas

- a) Extend the concept of Laplacian centrality of nodes to edges, obtain the graph theoretical description of the edge centrality.
- b) What is the relation between Laplacian centrality of the node and that of an edge?
- c) Can we apply the concept of edge centrality to the minimum cut problem in graphs?
- d) Yang et.al, in his work "Air traffic network optimization via Laplacian energy maximization" highlights the change in Laplacian energy due to edge removal as a promising fair measure of network robustness. We consider comparing this measure with other measures of network robustness.

- e) How can we apply the edge laplacian centrality to the concept of robustness in targeted edge attacks
- f) Application of laplacian centrality of an edge to edge consensus, that is, addition or removal of edges so the relative difference in centrality of edges is small.
- g) Application to electric flow in circuits

## 5 Minimum-maximum cut

### 5.1 Motivation

The minimum cut of a network is the minimum weight or number of edges whose removal results into a disconnected network. The max-flow min-cut theorem states that in a flow network, the maximum amount of flow passing from the source to the sink is equal to the total weight of the edges in the minimum cut, i.e. the smallest total weight of the edges which if removed would disconnect the source from the sink. This theorem has applications in image segmentation, and in optimisation problems such as project selection by guiding the decision regarding the purchase of machines that can be used in a number of projects so as to maximise profits of the company.

### 5.2 Ideas

- a) Investigate whether there exists a unique feature or behaviour for the nodes around the minimum cut.
- b) Whether the minimum cut edges can be obtained by their Laplacian centrality.
- c) relation between Laplacian energy and the minimum cut of a network.

## 6 Noise on networks

### 6.1 Motivation

In systems made of interconnections of sensors, it's observed that the reported results from the systems are normally subject to errors due to noise. In order to obtain accurate results, we need to ascertain the effect of noise on the reported results. Since these systems can be represented as networks, we can then formulate this problem as one of studying the impact of noise on sensor network.

### 6.2 Ideas

- a) How do we define noise on networks? Could the noise be applied to nodes or edges?
- b) What could be the effect of noise on the structure of a network
- c) What insights do we draw from simulation of noise on networks by considering additive White Gaussian noise for instance?

## 7 Communicability in Networks

### 7.1 Motivation

Communicability in networks is a concept introduced by Estrada & Hatano based on the motivation that most real networks, communication between nodes does not necessarily follow shortest paths only

as normally assumed. Evidence shows that communication can occur along any paths which are not shortest path. Communicability between a pair of nodes accounts for all possible walks through which a given pair of nodes can communicate. Some of the applications of this concept include community detection in networks. We are looking forward to research possibilities in this area which include the following:

## **7.2 Ideas**

- a) Consider the impact of long range interactions to communicability in networks
- b) Other possible real-world application of communicability in networks

## Work Progress

### 8 Diffusion on networks

Diffusion is, among others, the movement of substance from a region of high concentration to a region of low concentration. Such substance include heat, gas, etc. (Newman, 2010).

Diffusion process over networks is one of the methods used in developing simple models that depict the spread of infections in a population, dissemination of information over social network for instance social network marketing, spread of heat over a conductor, among others. A number of models based on diffusion process have been developed and documented in (Estrada et al., 2011; Kasprzak, 2012; López-Pintado, 2008).

#### 8.1 Heat Diffusion Models

In this work, we consider spread of heat on a network. Recently, various models have been developed to depict the spread of heat diffusion process on networks which include selection of marketing candidates in social network marketing, data analysis and processing in which the observed data is considered as a sum of diffusion processes, dimensionality reduction and classification problems (Ma et al., 2008; Thanou et al., 2017; Belkin and Niyogi, 2003).

Let us consider a simple example illustrating the diffusion process on a simple network. Let  $G = (V, E)$  be a simple connected undirected graph with vertex set  $V$  and edge set  $E$ . Suppose we randomly select a few nodes (that is, sources) to which we assign specific amounts of heat as in vector  $\phi_0$ . With the heat diffusion coefficient,  $C \in [0, 1]$  which controls the rate of diffusion. When  $C$  tends to 0, heat transfer among nodes becomes difficult and as a result, heat does not spread to each of the nodes with in the network. However, as  $C$  tends to 1, heat spreads rapidly among nodes and thus, with out loss, heat is distributed to all nodes in the network.

At each time  $t$ , we obtain the quantities of heat at each node,  $\phi_t$ . The spread of heat is considered to occur following the edges connecting nodes, that is to say, through direct interactions.

The process of heat spread through out the network can therefore be modelled by

$$\frac{d\phi_i}{dt} = C \sum_j (\mathbf{A}_{ij} - \delta_{ij} k_i) \phi_j, \quad (1)$$

where  $\mathbf{A}$  is the adjacency matrix,  $k_i$  is the degree of node  $i$ , and  $\delta_{ij}$  is the Kronecker delta whose value is 1 if  $i = j$  and 0 otherwise. In matrix-vector notation, we have

$$\frac{d\phi}{dt} = -C\mathbf{L}\phi, \quad \phi(0) = \phi_0, \quad (2)$$

whose solution is

$$\phi(t) = \phi_0 e^{-C\mathbf{L}t}. \quad (3)$$

Alternatively, the solution can be expressed as a linear combination of eigenvectors of the Laplacian matrix. That is

$$\phi(t) = \sum_i \langle \phi(0), \mathbf{v}_i \rangle e^{-C\lambda_i t} \mathbf{v}_i,$$

where  $\lambda_i$ ,  $\mathbf{v}_i$  are respectively the eigenvalues and corresponding eigenvectors of the Laplacian matrix and  $\langle \phi(0), \mathbf{v}_i \rangle$  is simply the projection of  $\phi(0)$  onto the set of eigenvectors.

## 8.2 Equilibrium behaviour

As  $t$  goes to infinity, we have

$$\lim_{t \rightarrow \infty} e^{-C\lambda_i t} = \begin{cases} 0 & \text{if } \lambda_i > 0 \\ 1 & \text{if } \lambda_i = 0, \end{cases} \quad (4)$$

Asymptotically, the equilibrium state is completely determined by the kernel of  $\mathbf{L}$ . Since  $\sum_j \mathbf{L}_{ij} = 0$ , it is easy to see that  $\mathbf{v}^1 = \frac{1}{\sqrt{n}}[1, \dots, 1]$ , the eigenvector associated with  $\lambda_i = 0$ , is in the kernel of  $\mathbf{L}$ . We then have

$$\lim_{t \rightarrow \infty} \phi(t) = \langle \phi(0), \mathbf{v}^1 \rangle \mathbf{v}^1. \quad (5)$$

The quantity of heat  $\phi_j(t)$  at any node  $j$  at time  $t$  is given by

$$\lim_{t \rightarrow \infty} \phi_j(t) = \frac{1}{n} \sum_{i=1}^n \phi_i(0). \quad (6)$$

At steady state, the value of  $\phi$  converges to the same value at each of the nodes in the network, which is the average of the initial values at all of the nodes. This is because, as expected, neighboring nodes in the network will exchange heat until all nodes attain equal amounts of heat.

For better understanding of the heat diffusion model, let us consider the following simple example.

**Example 1** Let us consider diffusion of heat over the network in Fig. 1(a). Suppose the quantity of heat at each node at time  $t = 0$  is given by the vector  $\phi(0) = [0.3, 0.0, 0.8, 0.0, 0.5, 0.2, 0.0, 0.0, 0.0, 0.2]$ , random values between 0 and 1. Let  $C = 0.05$ . Fig. 1(b) illustrates how heat spreads over the network in Fig. 1(a).

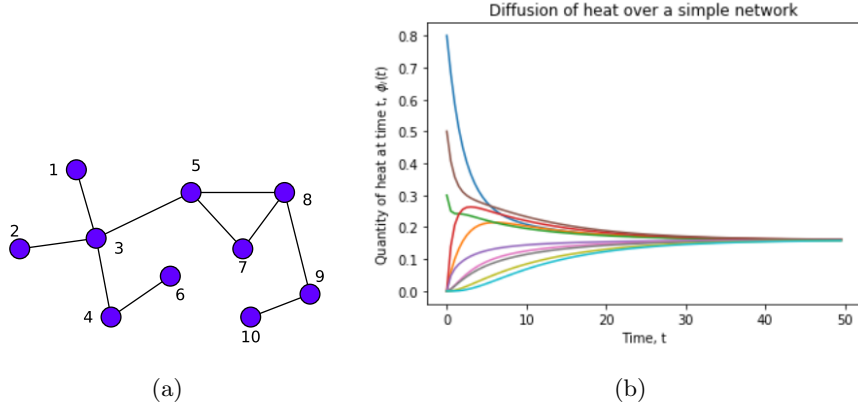


Figure 1: (b) is an illustration of the diffusion process over the network in (a).

We observe that at each time step  $t$ , nodes that initially have high amounts of heat (i.e 1, 3, 5, 6, and 10) exchange heat with adjacent nodes that initially had none or little amounts of heat. The latter gain heat from the former and eventually all nodes in the network have relatively equal amounts of heat. This explains the fact that as time  $t$  increases, the quantity of heat  $\phi_j(t)$  at each node tends to the equilibrium value of 0.2. From the fig, we can observe a tendency to equilibrium at around  $t = 35$ .

## 8.3 Diffusion on Directed Networks

A directed graph, also known as a Digraph or directed network, is one in which all the edges are directed from one vertex to another.

There are various complex systems whose skeleton can be captured by directed networks. Examples include ecological networks, power grids, transportation networks, communication networks, metabolic networks, gene regulatory networks, citation networks among others. It is therefore paramount to study how dynamic processes such as diffusion, consensus, occur on such networks.

**Definition 1 (Strongly connected Digraph)** A digraph is called strongly connected if and only if any two distinct nodes of the graph can be connected via a path that respects the orientation of the edges of the digraph (*Saber and Murray, 2003*).

For a strongly connected digraph with atleast two distinct nodes and with no self loops, the diffusion process on this network can be modelled in a similar manner as its undirected counterpart by

$$\frac{d\phi}{dt} = -C\mathbf{L}\phi, \quad \phi(0) = \phi_0. \quad (7)$$

For undirected graph  $G$ , the graph Laplacian,  $L$ , is symmetric positive semi-definite. However, for directed graphs  $L$  is non-symmetric which implies that the diffusion on the former and latter graphs is not necessarily the same.

## 8.4 Equilibrium behaviour in Directed Network

In order to understand the process of attainment of steady state in networks, we need to study the spectral properties of graph Laplacian. Let  $G = (V, E)$  be a digraph with Laplacian  $L(G)$  with eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  in non-decreasing order.

### 8.4.1 Estimation of Eigenvalues of the Laplacian

Let  $d_{max}$  be the maximum node out-degree of  $G$ , then following from Gershgorin disk theorem, then all the eigenvalues of  $L(G)$  are located in the following disk

$$D(G) = \{z \in \mathbb{C} : |z - d_{max}| \leq d_{max}\} \quad (8)$$

with centre at  $z = d_{max} + 0j$  in the complex plane (*Saber and Murray, 2003*). Thus, for a strongly connected digraph  $G$ ,  $L$  has a zero eigenvalue  $\lambda_1 = 0$  and all the other non-trivial eigenvalues have non-negative real parts. Let us consider a strongly connected digraph  $G = (V, E)$ . Let  $\phi_0$  be the vector of quantities of heat at all nodes at  $t = 0$ ,  $C = 1$  be the diffusion coefficient. Similar to undirected case, the quantities of heat,  $\phi(t)$  at each node at a given time  $t$  is given by

$$\phi(t) = \phi_0 e^{-\mathbf{L}t}. \quad (9)$$

**Theorem 1 (Limit Theorem for Exponential Matrices)** Assume  $G$  is a strongly connected digraph with Laplacian  $L$  satisfying  $\mathbf{L}\mathbf{v}_r = \mathbf{0}$ ,  $\mathbf{v}_l^T \mathbf{L} = \mathbf{0}$ , and  $\mathbf{v}_l^T \mathbf{w}_r = 1$ . Then

$$R = \lim_{t \rightarrow +\infty} \exp(-Lt) = v_r v_l^T \text{ in } M_n, \quad (10)$$

where  $M_n$  denotes a set of square  $n \times n$  matrices,  $v_r$ , and  $v_l^T$  denote the right and left eigenvalues of  $L$  associated with eigenvalue  $\lambda_1 = 0$  (*Saber and Murray, 2003*).

From the theorem, we deduce that for a strongly connected digraph, equilibrium state can be attained and the quantity of heat at the nodes is given by

$$\lim_{t \rightarrow \infty} \phi(t) = \phi_0 \mathbf{v}_r \mathbf{v}_l^T \quad (11)$$

It is important to note that following Equation 11, any equilibrium value  $x^*$  can be attained such that  $x_i^* = x_j^*$  for all  $i, j$ . This therefore draws to find out which classes of digraphs attain equilibrium similar to that of undirected graphs where the value at each node is the average of the initial values at all nodes in the network.

### 8.4.2 Equilibrium state for Balanced Graphs

**Definition 2 (Balanced Graphs)** We say the node  $v_i$  of a digraph  $G = (V, E)$  is balanced if and only if its in-degree and out-degree are equal, that is,  $d_{out}(v_i) = d_{in}(v_i)$ . A graph  $G$  is balanced if and only if all its nodes are balanced, i.e.  $\sum_j a_{ij} = \sum_j a_{ji}, \forall i$ .

The proposition in (Saber and Murray, 2003) states that

Let  $G = (V, E)$  be a digraph with an adjacency matrix  $A = [a_{ij}]$  satisfying  $a_{ii} = 0, \forall i$ . Then, all the following statements are equivalent:

1.  $G$  is balanced,
2.  $\mathbf{v}_1 = \mathbf{1}$  is the left eigenvector of the Laplacian of  $G$  associated with the zero eigenvalue, that is,  $\mathbf{1}^T \mathbf{L} = 0$ .
3.  $\sum_{i=1}^n u_i = 0, \forall x \in \mathbb{R}^n$  with  $u_i = \sum_{j \in N_i} a_{i,j}(x_j - x_i)$ .

## 8.5 k-path Laplacian matrices

The  $k$ -path Laplacian matrices are a natural generalisation of the combinatorial Laplacian of a graph. The entries of these matrices are accounts of the existence of shortest paths of length  $k$  between two nodes.

The  $k$ -path laplacian matrix  $\mathbf{L}_k(k \leq d_{max})$  of a connected undirected graph  $G = (V, E)$  is defined as the square symmetric  $n \times n$  matrix whose entries are given by:

$$\mathbf{L}_k(i, j) = \begin{cases} -1 & \text{if } d_{i,j} = k, \\ \delta_k(i) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

where  $d_{i,j}$  is the shortest path distance between nodes  $i$  and  $j$ ,  $\delta_k(i)$  known as the  $k$ -path degree is the number of irreducible shortest paths of length  $k$  having node  $i$  as an endpoint.

The concept of  $k$ -path Laplacians defined in Equation 12 for finite undirected graphs was extended for connected and locally finite infinite graphs as follows: Consider  $\Gamma = (V, E)$  to be an undirected finite or infinite graph with vertices  $V$  and edges  $E$ . We assume that  $\Gamma$  is connected and locally finite that is to say each vertex has only finitely many edges emanating from it. Let  $d$  be the distance metric on  $\Gamma$ , i.e.  $d(v, w)$  is the length of the shortest path from  $v$  to  $w$ , and let  $\delta_k(v)$  be the  $k$ -path degree of the vertex  $v$ , i.e.

$$\delta_k(v) := \#\{w \in V : d(v, w) = k\}. \quad (13)$$

Since  $\Gamma$  is locally finite,  $\delta_k(v)$  is finite for every  $v \in V$ . Denote by  $C(V)$  the set of all complex-valued functions on  $V$  and by  $C_0(V)$  the set of the complex-valued functions on  $V$  with finite support. Moreover, let  $\ell^2(V)$  be the Hilbert space of square-summable functions on  $V$  with inner product

$$\langle f, g \rangle = \sum_{v \in V} f(v) \overline{g(v)}, \quad f, g \in \ell^2(V) \quad (14)$$

In  $\ell^2(V)$  there is a standard orthonormal basis consisting of the vectors  $e_v, v \in V$ , where

$$e_v(w) := \begin{cases} 1 & \text{if } w = v, \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

Let  $\mathbf{L}_k$  be the following mapping from  $C(V)$  into itself:

$$(\mathbf{L}_k)(f) := \sum_{w \in V: d(v,w)=k} (f(v) - f(w)), \quad f \in C(V). \quad (16)$$



## 8.6 Diffusion on network with long-range interactions (generalisation of diffusion equation on Graphs)

In this case, we consider diffusion on a graph where by both direct and long range interactions are involved. One interesting study of long range interactions is by Estrada on modelling epidemic spread in networks. Here, the long range interactions are considered to be nonrandom and depend on the social distances between individuals in the social network. Recently, Estrada introduced a method of generalisation of the diffusion process on a given graph based on the  $k$ -path Laplacian operators  $\mathbf{L}_k$  in which we consider the fact that the diffusive particle at a given node can hop to not only its nearest neighbours (captured by the classical laplacian operator  $\mathbf{L}$ ) but to any other nodes in the graph with a probability that decays with the increase of the shortest path distance between the current node where the particle is residing to the one it will hop to. Thus, for a diffusive node hopping to other nodes separated at a distance  $k$  from its current location, such a diffusion process can be captured by replacing  $\mathbf{L}$  in (2) by  $\mathbf{L}_k$  in (16). That is

$$\frac{d\phi}{dt} = -C\mathbf{L}_k\phi, \quad \phi(0) = \phi_0. \quad (17)$$

We then consider the generalised diffusion on graph where interactions occur not only among neighbouring nodes but also among any other nodes with in the graph. This process follows from (2) where  $\mathbf{L}$  is any of the transformed  $k$ -path Laplacians given as

1. The Laplace-transformed  $k$ -Laplacian

$$\tilde{\mathbf{L}}_{L,\lambda} = \sum_{k=1}^{\infty} e^{-\lambda k} \mathbf{L}_k \quad (18)$$

2. The factorial-transformed  $k$ -Laplacian

$$\tilde{\mathbf{L}}_{F,z} = \sum_{k=1}^{\infty} \frac{z^k}{k!} \mathbf{L}_k \quad (19)$$

3. The Mellin-transformed  $k$ -Laplacian

$$\tilde{\mathbf{L}}_{M,s} = \sum_{k=1}^{\infty} \frac{1}{k^s} \mathbf{L}_k \quad (20)$$

## 8.7 Results

We considered a lattice where each node is connected to 8 nearest neighbours. We assigned specific heat quantities to a few selected nodes and then illustrated how the heat diffuses through the lattice based on direct interactions only and also based on both direct and long-range interactions among nodes. We observed that diffusion occurs much more faster by considering long-range interactions than when only direct interactions are accounted for.

## 9 Image segmentation

Image segmentation is defined as the problem of localising regions of an image relative to content such as image homogeneity. There are various algorithms that perform image segmentation based on specific content localisation such as intensity, color, texture, etc. It is important to note a good image segmentation algorithm must perform fast computation and editing, produce arbitrary segmentation with enough interaction as well as producing intuitive images.

Graph theory concepts are often applied in developing a number of image segmentation algorithms such as graph cuts algorithms, the intelligent scissor algorithm, the random walker algorithm, among others.

### 9.1 Image segmentation based on Random Walks

A random walk (also known as drunkard's walk) is a stochastic process that describes a path that consists of a succession of random steps on some mathematical space. An example of a random walk on a graph can be described as in the following way: Suppose a walker starts off a given node  $i$  and then moves to one of the neighbouring nodes of  $i$ , say  $j$ , with a uniform probability. Once at  $j$ , the walker again moves to the next node following a similar process until the stopping node. the sequence of these randomly visited nodes is referred to as the random walk.

There are various types of random walks, however, for the work we consider lattice random walkers that is to say random walkers that occur on on the lattice.

Here, we study the application of random walker on 2-d lattice to the segmentation of images.

### 9.2 Random Walks and Electric potential

Let's consider a graph  $G = (V, E)$  with two nodes,  $a$  and  $b$ , selected as traps with a score of 0 and 1 for hitting nodes  $a$  and  $b$  respectively. So for any other node  $x$ , the probability  $p(x)$  of reaching node  $b$  given that the walker started at node  $x$  is given by

$$p(x) = \frac{1}{d_x} \sum_{y, y \in N(x)} p(y), \quad (21)$$

where  $N(x)$  denotes the neighbours of  $x$ .

$$p(a) = 0 \quad (22)$$

$$p(b) = 1 \quad (23)$$

We then consider a graph  $G = (V, E)$  as an electrical network which consists of connected wires. The vertices represent the intersections of the wires in the circuit such that  $v_x$  is the voltage at a given vertex  $v \in V$ . On the other hand, the edges represent the resistors along the wires such that for each edge  $x, y \in E$  is a resistor of resistance  $r_{x,y}$ . Suppose we set the resistances  $r_{x,y}$  for any pairs of nodes in graph to 1 Ohm and set the voltages of vertices  $a$  and  $b$  to 0 and 1 respectively that is

$$v(a) = 0 \quad (24)$$

$$v(b) = 1 \quad (25)$$

By Ohm's law the current flowing through vertices  $x$  and  $y$  is given by

$$i_{x,y} = v(x) - v(y) \quad \forall x \neq y \neq b, \quad x, y \in E \quad (26)$$

Applying Kirchoff's law, we have

$$\sum_x i_{x,y} = 0, \quad \sum_{y, x \in E} v(x) - v(y) = 0, \quad (27)$$

which gives

$$v(x) = \frac{1}{d_x} \sum_{y \in N(x)} v(y) \quad (28)$$

We observe that system of equation of the probabilities in (21,22,23) and that of the voltages (24,25,28) follow the same law. That is to say they are both harmonic functions and satisfy the following properties

1. Mean value property Let us consider simpler example of one dimension case. Let  $A$  be the set of points  $A = \{0, 1, 2, \dots, N\}$  and let  $B$  be the set of boundary points  $B = \{0, N\}$  and  $C$  be the set of interior points  $C = \{1, 2, \dots, N-1\}$ . So the function  $f(x)$  defined on  $A$  is harmonic if it satisfies the mean value property given as

$$f(x) = \frac{f(x+1) + f(x-1)}{2} \quad (29)$$

From (21) and (28), we observe both functions  $p$  and  $v$  satisfy the mean value property in (29) since for any node other than  $a$  or  $b$ , the value of the function at that node is the average of the values of its neighbours.

In addition, both functions have the same values at the boundary that is at nodes  $a$  and  $b$ .

2. Maximum principle A harmonic function  $f(x)$  defined on  $A$  is said to satisfy the maximum principle if minimum and maximum values are attained only at the boundary. For functions  $v$  and  $p$ , we can observe from (22,23) and (24,25) respectively that the maximum principle holds.
3. Uniqueness Principle

If  $f(x)$  and  $g(x)$  are harmonic functions on  $A$  such that  $f(x) = g(x)$  on  $B$ , then  $f(x) = g(x)$  for all  $x$ . Following from previous discussions,  $v(x)$  and  $p(x)$  follow the uniqueness principle which implies that the two functions are equal. This therefore implies that the probability of a random walker reaching a given labelled node  $y$  given that the walker started at node  $x$  is equal to the electric potential developed at node  $x$  when node  $y$  is set to a potential of 1 volt. It is this similarity between random walks and electric potentials on graphs that was applied in the development of image segmentation using random walker algorithm as we will explore in the next sections.

## 9.3 The Random walker Algorithm

### 9.3.1 Problem Formulation

First, we represent the image as a weighted undirected graph (discrete object) where each pixel is represented as a node and each node connected by weighted edges to either 4 or 8 of its nearest neighbours. The real-valued edge weights can be captured by various weighting functions. However, in this work we use a Gaussian weighting function which maps changes in pixel intensities to the edge weights as given by

$$w_{i,j} = \exp(-\beta(g_i - g_j)^2), \quad (30)$$

where  $\beta$  is a free parameter,  $g_i$  is the intensity at pixel  $i$ . To capture colour, filter coefficients, texture or any other desired image features, we can modify (30). For instance, to capture colour, we replace  $(g_i - g_j)^2$  with  $\|g_i - g_j\|^2 \forall e_{i,j} \in E$ .

Having represented the image as a graph and given user-defined seeds that represent regions of belonging to the desired objects, we obtain an interactive image segmentation by assigning each unseeded node or pixel to the seed to which a random walker starting at the unseeded node first reaches a particular seed. The random walker is biased by the edge weights to avoid crossing sharp intensity gradients for respect of boundary objects.

Suppose we have an image that we would like to segment into  $K$  regions or objects that is to say a  $k$ -way image segmentation. First, seeds are selected by user to represent the desired  $k$  regions. The main task involves obtaining a  $k$ -tuples of probabilities with which a random walker starting at each unseeded

node reaches the seeds. Each unseeded node is then assigned a label for the seed for which the largest probability was obtained. One approach to this problem would be simulation of the random walker process though unfortunately, the method would be infeasible for certain image segmentation problems of interest. Alternatively, work in [references ] indicates that the probability with which a random walker first reaches a given seed can be found as a solution to the Combinatorial Dirichlet problem with boundary conditions at the location of the seed points with the seed in question set to 1 as the rest of the seeds are set to 0. With this approach, we can then analytically compute the desired probabilities as we will discuss in the next subsections.

### 9.3.2 Combinatorial Dirichlet Problem

Dirichlet problem is the problem of finding a harmonic function subject to its boundary conditions. A harmonic function is a function that satisfies the Laplace equation

$$\nabla^2 u = 0, \quad (31)$$

for a field  $u$ . The harmonic function that satisfies the boundary conditions minimises the Dirichlet integral

$$D[u] = \frac{1}{2} \int_{\Omega} |\nabla|^2 d\Omega, \quad (32)$$

for region  $\Omega$ .

### 9.3.3 Implementation of Algorithm

## 10 Laplacian Centrality of weighted Networks

The centrality of a node is a measure of how important or central a node is with in a network. A variety of centrality measures have been introduced for undirected unweighted networks based on various definitions of importance of a node. These include: degree, closeness, betweenness, eigenvector and subgraph centralities. Standard centrality measures i.e degree, closeness, and betweenness were extended to weighted networks due to the fact that weighted networks provide more information about the network and therefore measures applied to these networks are of great importance. These standard centrality measures give information on either the local environment of a node (i.e degree centrality) or the global position of the node in the network (i.e closeness, betweenness and subgraph centralities). This implies that information about the intermediate (between local and global) environment of a node cannot be captured by any of the standard centralities, yet, such information is very useful in the study of real-world networks. For instance, quantifying the relative importance of a particular actor in a social network. It is for this reason that a new type of centrality known as the Laplacian centrality was introduced by Qui, et al.

With Laplacian centrality measure, the importance of a node is determined by the ability of the network to respond to the deactivation of the node from the network. In other words, it is a measure of the relative drop of Laplacian energy in the network due to the removal (or deactivation) of the node from the network. The drop of Laplacian energy with respect to node  $v$  is determined by the number of 2-walks that  $v$  participates in the network.

### 10.1 Laplacian energy of a network

Let  $G = (V, E, W)$  be a simple undirected weighted network with the vertex set  $V(G) = \{v_1, v_2, \dots, v_n\}$ , edge set  $E$ , where each edge  $e = (v_i, v_j)$  is attached with a weight  $w_{ij}$ . If there is no edge between  $v_i$  and  $v_j$ ,  $w_{i,j} = 0$ . In addition,  $w_{i,i} = 0$  and  $w_{i,j} = w_{j,i}$ . We define

$$\mathbf{W}(\mathbf{G}) = \begin{pmatrix} 0 & w_{1,2} & \dots & w_{1,n} \\ w_{2,1} & 0 & \dots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \dots & 0 \end{pmatrix} \text{ and } \mathbf{X}(\mathbf{G}) = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n \end{pmatrix},$$

where  $x_i$  is the sum-weight of vertex  $v_i$  given by  $x_i = \sum_{j=1}^n w_{i,j} = \sum_{u \in N(v_i)} w_{v_i,u}$ , where  $N(v_i)$  is the neighborhood of  $v_i$ .

**Definition 3 (Weighted Laplacian matrix)** *The Laplacian matrix of a weighted network  $G$  is the matrix  $\mathbf{L}(\mathbf{G}) = \mathbf{X}(\mathbf{G}) - \mathbf{W}(\mathbf{G})$ .*

**Definition 4 (Laplacian Energy of a network)** *Let  $G = (V, E, W)$  be a weighted network on  $n$  vertices and  $\lambda_1, \lambda_2, \dots, \lambda_n$  be the eigenvalues of its Laplacian matrix. The Laplacian energy of  $G$  is defined as*

$$E_L(G) = \sum_{i=1}^n \lambda_i^2.$$

As networks become larger, computing eigenvalues of the Laplacian matrix becomes very hard. We therefore, use the entries of the Laplacian matrix rather than its eigenvalues to compute the Laplacian energy of a network as given by Theorem 2.

**Theorem 2** *For any network  $G = (V, E, W)$  on  $n$  vertices whose vertex sum-weights are  $x_1, x_2, \dots, x_n$  respectively, we have*

$$E_L(G) = \sum_{i=1}^n x_i^2 + 2 \sum_{i < j} w_{i,j}^2. \quad (33)$$

## 10.2 Laplacian centrality of node

If  $G = (V, E, W)$  is a network with  $n$  nodes  $V = \{v_1, v_2, \dots, v_n\}$ . Let  $G_i$  be the network obtained by deleting  $v_i$  from  $G$ . The Laplacian centrality is given by

$$C_L(v_i, G) = \frac{(\Delta E)_i}{E_L(G)} = \frac{E_L(G) - E_L(G_i)}{E_L(G)} \quad (34)$$

For any vertex  $v$ , the denominator remains unchanged and from Corollary ??, we can tell that  $E_L(G) - E_L(G_i)$  is non-negative. We then focus on obtaining the expression for  $(\Delta E)_i$ . In order to obtain the graph theoretical descriptions of Laplacian centrality, we will study the  $k$ -walks (discussed in Chapter 2) for the weighted graph, specifically, for  $k = 2$ . For better understanding of the weighted network concept, we represent a weighted network as an unweighted multigraph network by replacing each edge  $e = (v_i, v_j)$  with  $w_{ij}$  copies of multiedges as shown in Fig. ??(??). For instance, for a 2-walk  $v_1 v_2 v_3$  in a weighted network, the number of 2-walks in its corresponding unweighted network is  $w_{v_1, v_2} w_{v_2, v_3}$ .

**Theorem 3** Let  $G = (V, E, W)$  be a weighted network of  $n$  vertices  $v_1, v_2, \dots, v_n$ . Let  $G_i$  be the network obtained by deleting vertex  $v$  from  $G$ , then the drop of Laplacian energy with respect to  $v_i$  is

$$(\Delta E)_i = E_L(G) - E_L(G_i) = 4 \cdot NW_2^C(v_i) + 2 \cdot NW_2^E(v_i) + 2 \cdot NW_2^M(v_i). \quad (35)$$

where  $NW_2^C(v_i)$ ,  $NW_2^E(v_i)$ , and  $NW_2^M(v_i)$  are closed 2-walks containing vertex  $v_i$ , non-closed 2-walks with vertex  $v_i$  as one of the end points and non-closed 2-walks with vertex  $v_i$  as the middle point (Qi et al., 2012).

## 10.3 Laplacian Centrality of an edge

The identification of the most central nodes in a network is an important concept that has applications in the control of spread of an epidemic, faster dissemination of information with in a social networks, etc. On the other hand, we believe that it is also vital to identify how central an edge is with in the network. this therefore prompts us to study the Laplacian centrality centrality of an edge and define a graph theoretical expression for the drop in energy when an edge is removed from the network.

Similar to laplacian centrality of a node, we define the Laplacian centrality of an edge as the drop in Laplacian energy when an edge is removed from a network. Let us consider an undirected weighted network  $G = (V, E)$ . The Laplacian energy of  $G$  is given by

$$E_L(G) = \sum_{i=1}^n x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2, \quad (36)$$

where  $x(v_i) = \sum_{j \in N(v_i)} w_{i,j}$ .

On removing an arbitrary edge  $e_{1,2}$ , without loss of generality, assume  $H = G - e_{1,2}$ . Let  $N(v_i)$  be the neighborhood of vertex  $v_i$  in  $G$ ,  $v_i \in e_{i,j}$  represent that edge  $e_{i,j}$  is incident to vertex  $v_i$  in  $G$ , and  $x'(v_i)$  be the corresponding sum-weight of the vertex  $v_i$  in  $H$ . We have:

$$x'(v_i) = \begin{cases} x(v_i) - w_{v_1, v_2} & \text{if } v_i \sim e_{1,2}, \\ x(v_i) & \text{otherwise.} \end{cases} \quad (37)$$

The Laplacian energy of the subgraph is given by

$$E_L(H) = \sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1, v_2})^2 + \sum_{v_i \not\sim e_{1,2}} x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2 - 2 \cdot w_{v_1, v_2}^2 \quad (38)$$

**Definition 5 (Laplacian centrality of an edge,  $C_L(e)$ )** The Laplacian centrality of an edge  $e$ , in Graph  $G$  is given by

$$C_L(e) = \frac{E_L(G) - E_L(H)}{E_L(G)} = \frac{\Delta E_L}{E_L(G)} \quad (39)$$

From (39), we observe the denominators remain the same in computing the Laplacian centrality for edges. This then directs our attention to obtaining the graph theoretical descriptions of the drop in the Laplacian energy when a given node is removed from the graph.

Following a similar procedure in the proof for Theorem 3, the drop in Laplacian energy which the difference between (36) and (38) is given by

$$\begin{aligned}
\Delta E &= \sum_{i=1}^n x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2 - \left( \sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1,v_2}^2)^2 + \sum_{v_i \not\sim e_{1,2}} x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \right) \\
&= \sum_{i=1}^n x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2 - \left( \sum_{v_i \sim e_{1,2}} (x^2(v_i) - 2x(v_i) \cdot w_{v_1,v_2} + w_{v_1,v_2}^2) + \sum_{v_i \not\sim e_{1,2}} x^2(v_i) + 2 \sum_{i < j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \right) \\
&= 2 \sum_{v_i \sim e_{i,j}} x(v_i) \cdot w_{v_1,v_2} - \sum_{v_i \sim e_{i,j}} w_{1,2}^2 + 2 \cdot w_{v_1,v_2}^2 \\
&= 2 \sum_{v_i \sim e_{i,j}} (x(v_i) \cdot w_{v_1,v_2}) - 2 \cdot w_{v_1,v_2}^2 + 2 \cdot w_{v_1,v_2}^2 \\
&= 2 \sum_{v_i \sim e_{i,j}} x(v_i) \cdot w_{v_1,v_2} \\
&= 2 \sum_{v_i \sim e_{i,j}} w_{v_1,v_2} \sum_{u \in N(v_i)} w_{v_i,u} \\
&= 2 \cdot \sum_{v_1} w_{v_2,v_1} \sum_{u \in N(v_1); u \neq v_2} w_{v_1,u} + 2 \cdot \sum_{v_2} w_{v_1,v_2} \sum_{u \in N(v_2); u \neq v_1} w_{v_2,u} + 2 \cdot w_{v_1,v_2} \cdot w_{v_1,v_2} + 2 \cdot w_{v_2,v_1} \cdot w_{v_2,v_1} \\
\Delta E &= 2 \cdot NW_2^U(v_1(E), v_2(M)) + 2 \cdot NW_2^U(v_1(E), v_2(M)) + 4 \cdot NW_2^C(v_1, v_2) \tag{40}
\end{aligned}$$

where

$NW_2^U(v_1(E), v_2(M))$  is the number of non-closed walks of length 2 with vertex  $v_2$  as the middle point and vertex  $v_1$  as an end point.

$NW_2^U(v_1(E), v_2(M))$  is the number of non-closed walks of length 2 with vertex  $v_1$  as the middle point and vertex  $v_2$  as an end point.

$NW_2^C(v_1, v_2)$  is the number of closed walks of length 2 between vertices  $v_1$  and  $v_2$ .

From (40), we can easily tell the energy drop can be obtained by taking into account the immediate neighbourhood around the edge, that is, the nearest neighbours of the two nodes that are incident to the edge in question.

**Example 2** Let us consider the weighted graph in Fig. 2. We compute the drop in Laplacian energy by First, we compute the Laplacian energy  $E(G)$  of the graph as follows:

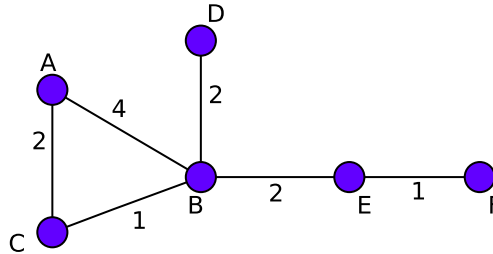


Figure 2

$$\begin{aligned}
E_L(G) &= \sum_{i=1}^n x_i^2 + 2 \sum_{i < j} w_{i,j}^2 \\
&= (6^2 + 3^2 + 9^2 + 3^2 + 1^2 + 2^2) + 2(4^2 + 2^2 + 1^2 + 2^2 + 2 + 1^2) \\
&= 200
\end{aligned}$$

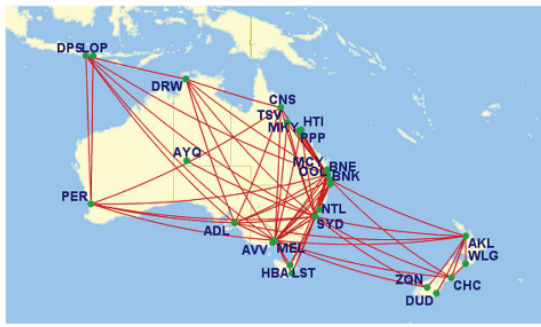
Table 1: Laplacian Centralities of edges

Edge(e)	$E(H); H = G - e$	$E(G) - E(H)$	$\Delta(E)$ by walks method in (40)
a,c	164	$200 - 164 = 36$	$2(2) + 2(8) + 4(4) = 36$
b,c	176	$200 - 176 = 24$	$2(2 + 2 + 4) + 2(2) + 4(1) = 24$
a,b	80	$200 - 80 = 120$	$2(8) + 2(4 + 8 + 8) + 2(16) = 120$
b,d	156	$200 - 156 = 44$	$2(0) + 2(4 + 8 + 2) + 4(4) = 44$
b,e	152	$200 - 158 = 88$	$2(2) + 2(4 + 8 + 2) + 4(4) = 48$
e,f	192	$200 - 192 = 8$	$2(0) + 2(2) + 4(1) = 8$

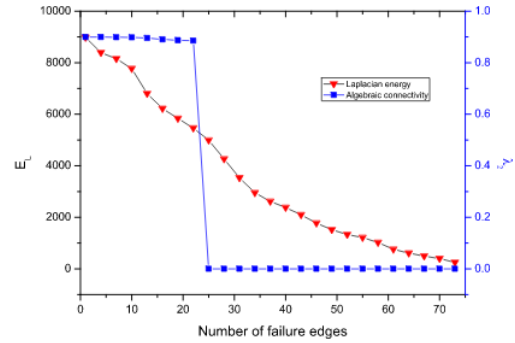
From Table 1, we observe the drop in energy computed by the difference between Laplacian energy of the graph  $G$  and that of the subgraph,  $H$  obtained on removing edge  $e$  is equal to that obtained using closed and non-closed walks in (40).

#### 10.4 Laplacian Energy as a fair measure of robustness of network

With the fast increasing usage of air transport in most regions of the world, there is a pressing need to design robust air traffic networks that will ensure robustness when one or more flights are removed or added to the network. The second smallest eigenvalue of the Laplacian matrix of air traffic networks,  $\lambda_2$ , also known as the algebraic connectivity is a common measure of robustness in networks. Unfortunately, the algebraic connectivity captures only the global information about the connectivity of a network. Recently, the laplacian energy was introduced as another measure of robustness that is considered fair and effective compared to the algebraic connectivity. This is so because the Laplacian centrality captures the local information of the network (Yang,etal, 2015). In order to illustrate the relationship between Laplacian energy and air traffic network robustness, a real air traffic network of Jet-star Asia Airway among Indonesia, Australia, and New Zealand (Fig. 3a ) was investigated.



(a)



(b)

Figure 3: (a) is the air traffic network route map for Jetstar Asia Airway (Indonesia, Australia, and New Zealand) in 2015, (b) is a plot of Laplacian energy  $E_L$  and algebraic connectivity  $\lambda_2$  against the number of randomly failed edges from the or Jetstar Asia Airway (Indonesia, Australia, and New Zealand).

From the Fig. (3b), we observe that values for both the laplacian energy and algebraic connectivity decreases with removal of edges from the network. However, on removal of 20 to 30 edges, the algebraic



connectivity abruptly drops from 0.9 to close to 0 (that is in only one instance) which signifies a disconnection in the network that is, more than one connected component in the network. on the other hand though, the laplacian curve indicates gradual degradation of the network robustness on removal of 20 to 30 edges of the network. The ability of the Laplacian energy measure to capture the change from one connected component to more connected components in much more instances makes it an effective measure for network robustness over algebraic connectivity.

## References

- R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.
- M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.
- K. C. Das. The laplacian spectrum of a graph. *Computers & Mathematics with Applications*, 48(5):715–724, 2004.
- C. H. Edwards and D. E. Penney. *Differential equations and boundary value problems*, volume 2. Prentice Hall, 2004.
- P. Erdos and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.
- E. Estrada. Quantifying network heterogeneity. *Physical Review E*, 82(6):066102, 2010.
- E. Estrada. *The structure of complex networks: theory and applications*. OUP Oxford, 2011.
- E. Estrada. Path laplacian matrices: introduction and application to the analysis of consensus in networks. *Linear Algebra and its Applications*, 436(9):3373–3391, 2012.
- E. Estrada. Introduction to complex networks: structure and dynamics. In *Evolutionary Equations with Applications in Natural Sciences*, pages 93–131. Springer, 2015.
- E. Estrada and J. A. Rodriguez-Velazquez. Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103, 2005.
- E. Estrada, F. Kalala-Mutombo, and A. Valverde-Colmeiro. Epidemic spreading in networks with non-random long-range interactions. *Physical Review E*, 84(3):036110, 2011.
- E. Estrada, P. Knight, et al. *A first course in network theory*. Oxford University Press, USA, 2015.
- L. Euler. Leonhard euler and the königsberg bridges. *Scientific American*, 189(1):66–70, 1953.
- J. Gower. A modified leverrier-faddeev algorithm for matrices with multiple eigenvalues. *Linear Algebra and its Applications*, 31:61–70, 1980.
- I. Gribkovskaia, Ø. Halskau, and G. Laporte. The bridges of königsberg—a historical perspective. *Networks*, 49(3):199–203, 2007.
- I. Gutman and O. E. Polansky. *Mathematical concepts in organic chemistry*. Springer Science & Business Media, 2012.
- J. M. Harris, J. L. Hirst, and M. J. Mossinghoff. *Combinatorics and graph theory*, volume 2. Springer, 2008.

- Internet. Network. URL [https://www.google.com/search?q=internet+network&client=ubuntu&hs=xZp&channel=fs&gl=za&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiZsrLOueTMAhVMCcAKHS\\_8AysQ\\_AUIBygB&biw=1215&bih=900#channel=fs&gl=za&tbm=isch&q=computer+network&imgsrc=..](https://www.google.com/search?q=internet+network&client=ubuntu&hs=xZp&channel=fs&gl=za&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiZsrLOueTMAhVMCcAKHS_8AysQ_AUIBygB&biw=1215&bih=900#channel=fs&gl=za&tbm=isch&q=computer+network&imgsrc=..). [Online; accessed 2016-03-29].
- R. Kasprzak. Diffusion in networks. *Journal of Telecommunications and Information Technology*, pages 99–106, 2012.
- R. Kondor and J.-P. Vert. Diffusion kernels. *kernel methods in computational biology*, pages 171–192, 2004.
- J. Lafferty and G. Lebanon. Diffusion kernels on statistical manifolds. *Journal of Machine Learning Research*, 6(Jan):129–163, 2005.
- D. López-Pintado. Diffusion in complex social networks. *Games and Economic Behavior*, 62(2):573–590, 2008.
- H. Ma, H. Yang, M. R. Lyu, and I. King. Mining social networks using heat diffusion processes for marketing candidates selection. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 233–242. ACM, 2008.
- S. Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.
- J. J. Molitierno. *Applications of combinatorial matrix theory to Laplacian matrices of graphs*. CRC Press, 2012.
- F. K. Mutombo. *Long-range interactions in complex networks*. PhD thesis, University of Strathclyde, 2012.
- M. Newman. *Networks: an introduction*. OUP Oxford, 2010.
- M. E. Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132, 2001.
- M. E. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- Nexus. Zachary’s karate club [karate], 2012. URL [http://nexus.igraph.org/api/dataset\\_info?id=1&format=html](http://nexus.igraph.org/api/dataset_info?id=1&format=html). [Online; accessed 2016-04-25].
- T. Opsahl. *Structure and evolution of weighted networks*. PhD thesis, Queen Mary, University of London, 2009.
- T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social networks*, 31(2):155–163, 2009.
- T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.
- X. Qi, E. Fuller, Q. Wu, Y. Wu, and C.-Q. Zhang. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194:240–253, 2012.
- X. Qi, R. D. Duval, K. Christensen, E. Fuller, A. Spahiu, Q. Wu, Y. Wu, W. Tang, C. Zhang, et al. Terrorist networks, network energy and node removal: a new measure of centrality based on laplacian energy. *Social Networking*, 2(01):19, 2013.
- R. O. Saber and R. M. Murray. Agreement problems in networks with directed graphs and switching topology. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 4, pages 4126–4132. IEEE, 2003.
- D. Thanou, X. Dong, D. Kressner, and P. Frossard. Learning heat diffusion graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 3(3):484–499, 2017.
- X. F. Wang and G. Chen. Complex networks: small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, 3(1):6–20, 2003.
- D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393(6684):440–442, 1998.

R. J. Wilson. *An introduction to graph theory*. Pearson Education India, 1970.