Technical Report

February 3, 2018

Outline

1 Literature Review

1.1 Networks Review

2 Diffusion on networks

2.1 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 spread of infections with in a group of people, spread of information in a social network, failures in power supply on the grid, etc.

2.2 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.

3 Image Segmentation

3.1 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, however, lies in the segmentation of images using the random walk process on graph as explored in (Grady, 2006).

3.2 Ideas

a) To extend the idea of image segmentation using random walks on networks by accounting for long-range links in the network. The aim is to ascertain whether this approach yields image segmentation than the existing approach which does not involve long range interactions.

4 Systemic Risk and Contagion in Financial Systems

4.1 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 research 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-occurance or to guarantee minimal impact of the crisis etc.

4.2 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?

5 Laplacian centrality of an edge

5.1 Motivation

Centrality measures in networks have proved to be relevant tools in network analysis. They are indicators of the 'importance' of a given node or edge in a network. Though most work has been geared towards the study of importance of nodes (i.e degree, closeness, betweeness, subgraph, eigenvector, Laplacian centralities,etc.), interest in the study of edge centralities is now gaining ground with prominent work 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 its applicability in real-world for example identification of communities in networks, identifying strong relationships among people in social networks, etc.

5.2 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 laplcaian centrality of an edge to edge consensus, that is, addition or removal of edges so that the relative difference in centrality of edges is small.

g) Application to electric flow in circuits

6 Minimum-maximum cut

6.1 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.

6.2 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) Relationship between Laplacian energy and minimum cut of a network.

7 Noise on networks

7.1 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.

7.2 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?

8 Communicability in Networks

8.1 Motivation

Communicability in networks in a concept introduced by Estrada & Hatano (Estrada and Hatano, 2008) motivated by the fact that in 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:

8.2 Ideas

a) Communicability through an edge which is the sum of the total number of walks between all pairs of nodes in the network that go through the edge of interest.

b) Consider the impact of long range interactions to communicability in networks

c) Other possible real-world applications of communicability in networks

9 Literature Review

9.1 Complex Networks

Complex systems play an important role in our daily lives for instance in social, economic, science, technology among others. During his interview with San Jose Mercury News in January 2000, Stephen Hawking referred to the 21st century as a century of complexity. However,though complex systems consist of interconnected components, they display some properties that quite different from those of individual components. Due to the fact that complex systems play a vital role in our lives, it is necessary to be able to understand and predict the properties of these systems. One approach to realise the mentioned task is by use of network theory. First, as Estrada (Estrada et al., 2015) mentioned that complex networks are the skeletons of complex systems, we represent such systems by networks whose nodes (vertices) and links(edges) represent the components and the interactions among components respectively. For instance, a transportation system can be represented by network where nodes are cities or towns and the links are the roads, railways or flight routes. Second, is network analysis which entails studying the structure of the network from which the properties of the network (and the system) are drawn. An interestingly early historical application of network theory to the study of complex systems is the Käonigsberg bridge problem where Euler (**?**) solved the problem by reformulating problem in terms of a graph where vertices represent islands while edges represent the seven bridges joining any two islands. Work published by Leonhard Euler (Euler, 1976) is considered the genesis of the story of network theory. Due to the adverse increase in network size from just graphs of tens or hundreds of nodes which could easily be analysed by direct use of eye so as to ascertain the structure of the network to complex networks consisting of million or billion of nodes which call for advanced analytic approach that involves development of statistical methods to quantify such large networks. The statistical methods aid in answering questions such as how many nodes or edges should be removed for the network to break down?, what is the shortest path length of the network?, and many others.

9.2 characteristics of complex systems

1. Emergence
2. Self organisation
3. uncertainty

9.3 Networks in the Real World

In his work in (Newman, 2003), Newman considered a loose categorisation of networks: social networks, communication networks, technological networks, and biological networks.

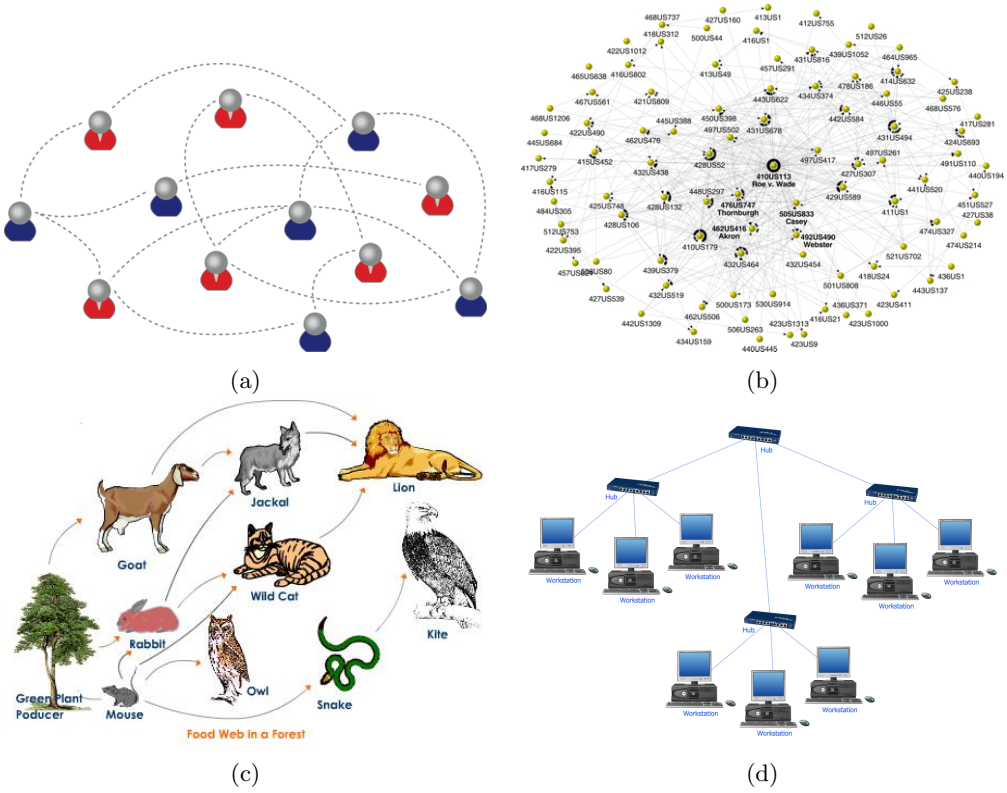
a. Social Networks Networks considered as social networks are ones whose nodes correspond to people or groups of people while the edges represent the interactions or relationship between them. For instance friendship networks such as facebook, twitter in which the interactions represent friendship ties among acquaintances, networks of intermarriages between families, social interaction networks which capture peoples' interactions through social activities or events, employee networks with companies, and many others. Some common networks that researchers have frequently experimented upon include: the Zachary karate network which consists of two communities centred at the administrator and instructor as a result of misunderstanding that prevailed with the karate club earlier on. The nodes in the network are the members of the club as the links represent interactions between members during non-club activities (Zachary, 1977). Other networks include the Dolphine network (Williams et al., 1993), terrorist network(Magouirk et al., 2008) among others .

b. Information networks: Information networks are also referred to as knowledge networks. Examples of networks under this category include: The world wide web which consists of billions of web pages as nodes that are linked together through links known as hyperlinks (Huberman, 2001). Another

network categorised as information networks are citation networks that are composed of nodes which are articles while directed link between two nodes written as $i \longrightarrow j$ indicate article i cites article j.

c. Technological networks: This category consist of networks made by man to aid in distribution or transfer of resources, services or commodities such as electricity, water, transportation services, and many others. Examples of such networks include the internet, transportation networks, power grids, to mention but a few.

d. Biological networks: Biological networks exists in areas related human and processes that take place with in the human body, animals and their ways of survival, chemistry. Such networks are the human brain network, protein-protein interaction network, network of metabolic path ways, ecological networks.



(a) (b) (c) (d)

Figure 1: Networks in real world: (a) A social network. (b) A citation network. (c) A food web. (d) Computer network. Source: (Internet)

10 Robustness of Complex Systems

A complex systems is considered robust if it can with stand failures or perturbations that is to say a system can still perform as expected even in circumstances of failure of one or more components in the system. Robustness of systems plays an important role in a number of fields for instance in Engineering, understanding robustness acts as a basis for designing communication, transportation systems, power grids that can perform basic operation despite failure of some system components. In biology,robustness explains why some mutations lead to diseases while others do not. For ecologists and environmental experts, robustness helps in predicting the failure of an ecosystem when faced with disruptive human behaviours. As mentioned earlier, the study of networks underlying complex systems provides insights about the properties and characteristics of the systems. Thus, Barabasi (Barabási, 2016) suggested that networks play a vital role in robustness of complex systems which implies that exploring robustness of the network reflects that of the system.

10.1 Robustness measures in networks

According to Ellens (Ellens and Kooij, 2013), robustness of a network as its ability to perform well when subject to failures or attacks. The attacks take on two forms namely: random attacks and targeted attacks. However, in order to tell whether a particular network is robust, there is need to a measure that quantifies the robustness. In the past, various robustness measures have been put forward by researchers (Sydney et al., 2008). Some of the measures are:

11 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).

11.1 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 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 $G = (V, E)$ 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 ϕ_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, ϕ_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

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

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

$$\frac{d\boldsymbol{\phi}}{dt} = -C\mathbf{L}\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{2}$$

whose solution is

$$\boldsymbol{\phi}(t) = \boldsymbol{\phi}_0 \; e^{-C\mathbf{L}t}. \tag{3}$$

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

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

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

11.2 Equilibrium behaviour

As t goes to infinity, we have

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

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 $\mathbf{v^1} = \frac{1}{\sqrt{n}}[1, \cdots, 1]$, the eigenvector associated with $\lambda_i = 0$, is in the kernel of $\mathbf{L}$. We then have

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

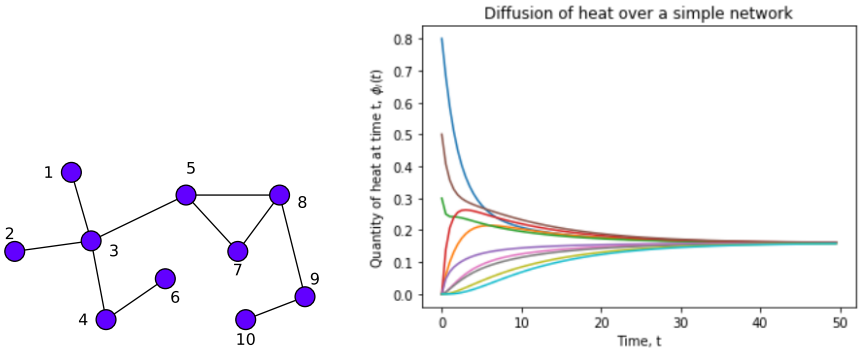
The quantity of heat $\phi_j(t)$ at any node j at time t is given by

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

At steady state, the value of $\boldsymbol{\phi}$ 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. 2(a). Suppose the quantity of heat at each node at time $t = 0$ is given by the vector $\boldsymbol{\phi}(0) = [0.3, 0.0, 0.8, 0.0, 0.5, 0.2, 0.0, 0.0, 0.0, 0.2]$, random values between 0 and 1. Let $C = 0.05$. Fig. 2(b) illustrates how heat spreads over the network in Fig. 2(a).*



(a) (b)

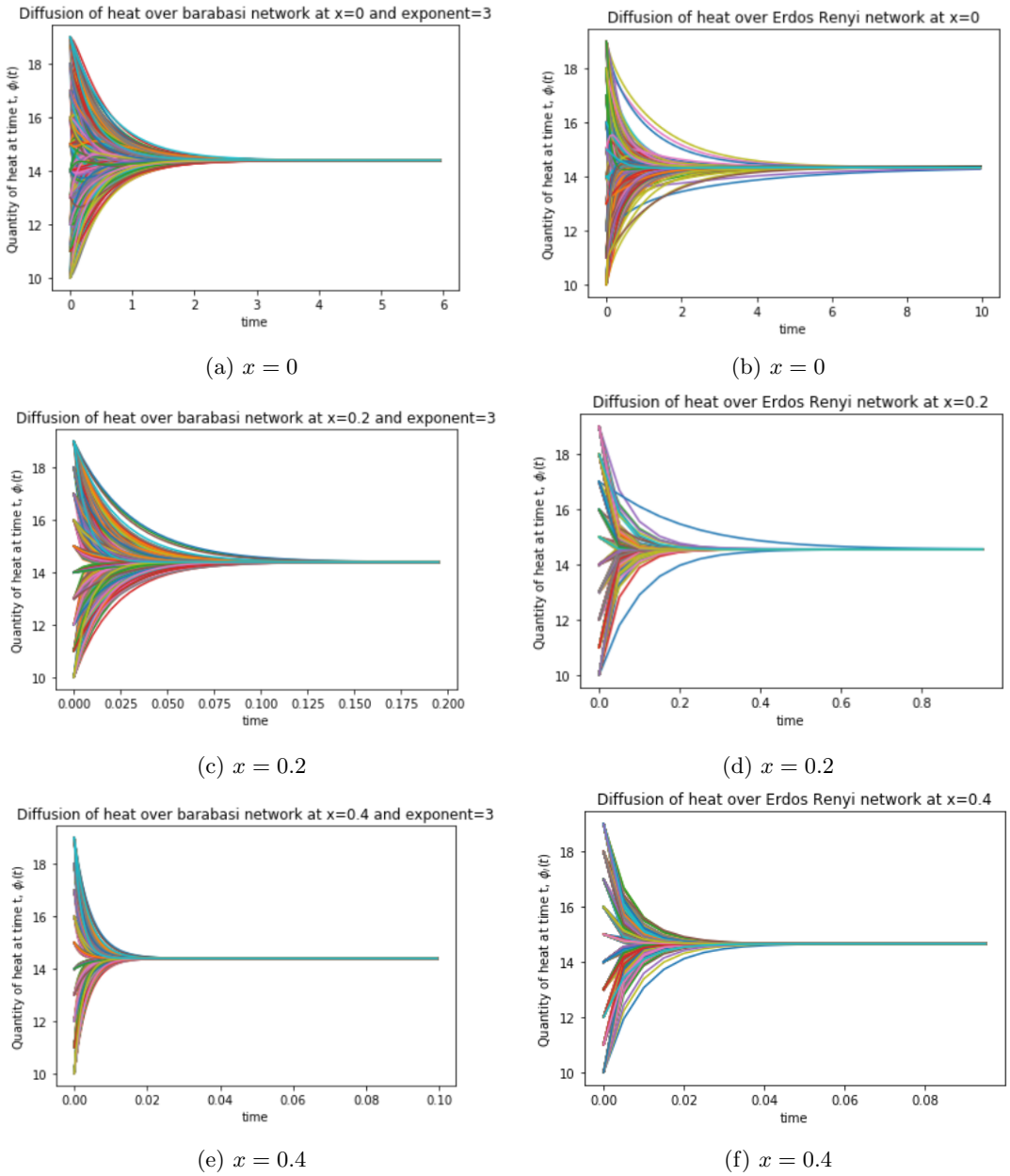
Figure 2: (b) is an illustration of the diffusion process over the network in (a).

From Fig.2,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 $\boldsymbol{\phi}_j(t)$ at each node tends to the equilibrium value of 0.2 which is attained at $t = 35$.

12 Impact of Structure on the rate of diffusion

The structure of a network basically means the way in which nodes are connected in the network. For instance, in a regular network each node is connected to equal number of nodes, for a star network one node is positioned in a way that all other nodes are connected to it. Let us consider two structures of networks that manifest in many artificial and real world networks. First, the Erdos-Renyi (ER) network

in which a pair of nodes is connected by random probability, p. The degree of nodes in ER network follow a Poisson distribution Erdos and Rényi (1960). Second, we consider the Barabasi-Albert(BA) network in which connection of nodes follows scale free power-law distribution, that is to say, the probability of finding a node with degree k decreases as the negative power of k. It therefore less likely to find a node with high degree (hub) compared to low degrees Barabási and Albert (1999); Estrada (2011) . Consider ER and BA networks with $n = 100$ and average degree $\bar{k} = 6$, we randomly assign quantities(range of 0 to 20) of heat to each node and allow diffusion to occur at different values of conductance x. After every time step t,we compute the quantities at each node as depicted in Figure 3.



(a) $x = 0$ (b) $x = 0$

(c) $x = 0.2$ (d) $x = 0.2$

(e) $x = 0.4$ (f) $x = 0.4$

Figure 3: Barabasi networks (left) and Erdos Renyi networks(right) of 1000 nodes and average degree of 6. Top row is illustration of diffusion process at $x = 0$ (i.e accounting for direct interactions only), middle row corresponds to $x = 0.2$ followed by $x = 0.4$ in the last row

From Fig 3 at the top row, we observe that at $x = 0$, equilibrium is reached faster for Barabasi-Alberto(BA) network (that is after about 4 time steps) compared to Erdos-Renyi(ER) network in which

equilibrium is reached after about 10 time steps. This is explained based on the fact that in BA networks there are more hubs compared to ER networks. These hubs tend to interact with a number of nodes with in the network thus fastening the diffusion process. On increasing x to 0.2, we observe a drastic drop in equilibrium time from 4 to 0.15 time steps and from 10 to 0.8 time steps for BA and ER networks respectively. It is important to note that drop in equilibrium time is relatively higher in ER than in BA and this is because of the few hubs in ER networks which aids a larger number of long range interactions than in BA networks. As x increases further to 0.4, equilibrium time further drops to 0.03 and 0.06 for BA and ER networks respectively.

13 Influence of Heterogeneity on Diffusion over network

The heterogeneity of a network is the irregularity characterised by the existence of a nodes with degree significantly larger than the average degree of the network Estrada (2010); Albert and Barabási (2002); Newman (2003). The quantification of heterogeneity is one the areas where tremendous research has been on going and various measures have been introduced Estrada (2010). Here, we consider heterogeneity in scale free networks with $n = 1000$ and average degree $\bar{k} = 20$ by varying power exponent, γ. For different conductances x, we assign initial quantities of heat to each of the 200 nodes with highest degree. Figure 4 illustrates how the average quantities of heat of the selected initial diffusion nodes varies with time.

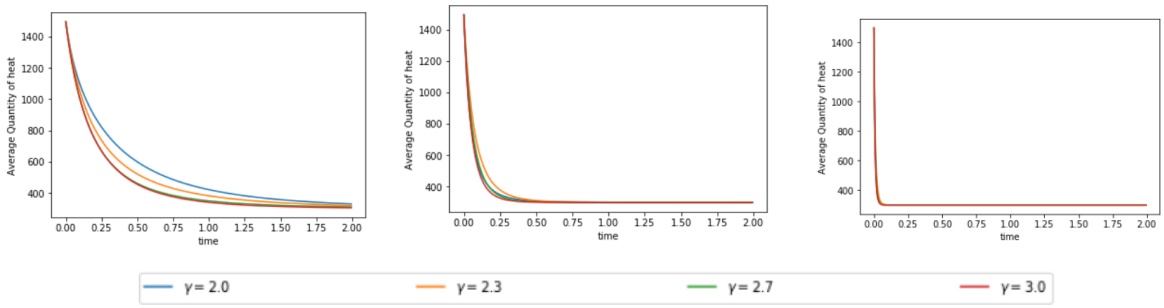


Figure 4: Plots of the average quantity of heat for 200 nodes with the highest degree centrality against time for 3 scale free networks having different values of the power exponent(2.0,2.3,2.7, and 3.0), n=1000 and average degree=6. The figures to the left, centre and right correspond to x values 0,0.1, and 0.3 respectively.

14 Impact of choice of Initial diffusion nodes on the diffusion process on networks

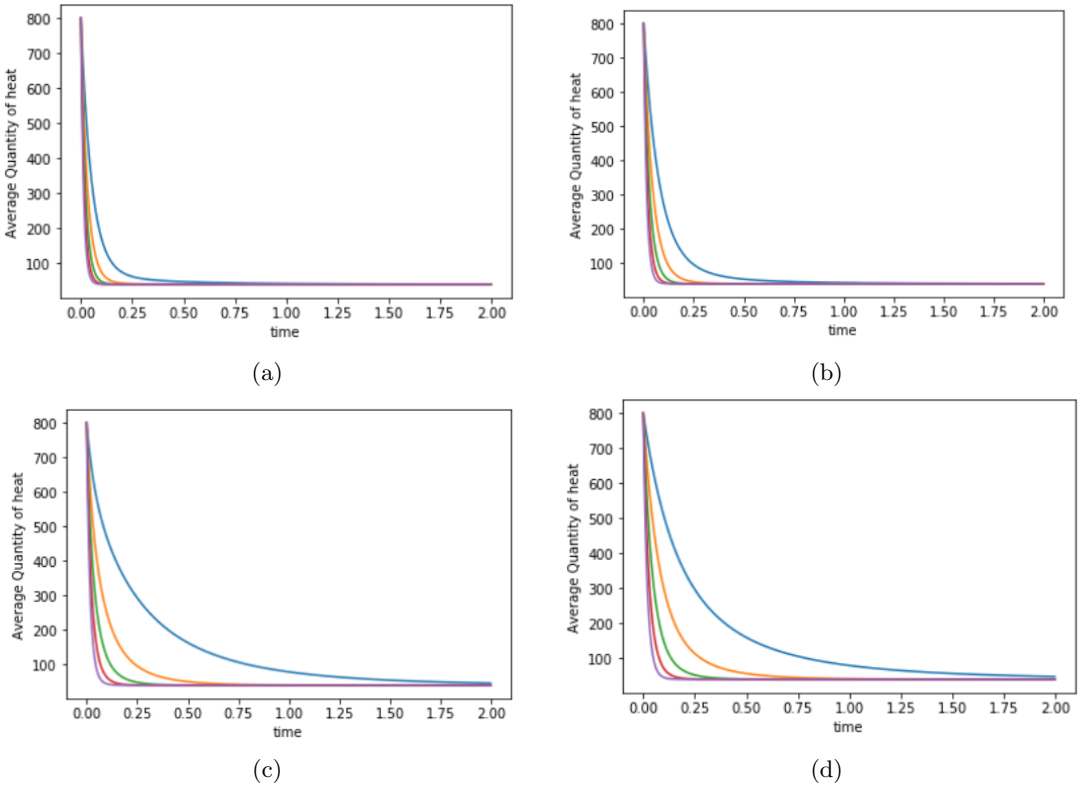


Figure 5: Results of the simulations for two networks. One is the Barabasi-Albert(BA) network and the other is Erdos-Renyi(ER) network, both of 100 nodes and average degree of 6. Taking 5 of the most important (by degree) nodes in the network from which diffusion is initiated by assigning certain quantities of heat to those nodes, the simulation of the diffusion is illustrated in plots (a) and (b) for the BA and ER networks respectively.

We observe that in BA network, heat spreads quite faster than in ER network that is to say by $t = 0.6$, all the 5 hubs in BA have levelled to equal amount while for ER, nodes still have un equal amounts of heat. This is because we assign quantities of heat to 5 hubs of a network and they quickly spread heat to other nodes compared to ER where there are relatively fewer hubs among the five chosen nodes to initiate the diffusion process.

14.1 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. There are various categories of directed networks which include:

Definition 1 (Weakly connected Digraph) *A directed graph is called weakly connected if replacing all of its directed edges with undirected edges produces a connected (undirected) graph.*

Definition 2 (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

$$\frac{d\boldsymbol{\phi}}{dt} = -C\mathbf{L}\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0. \tag{7}$$

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.

Definition 3 (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 $\sum_j a_{ij} = \sum_j a_{ji}, \forall i$.*

14.2 Diffusion and 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 $\lambda_1, \lambda_2, \cdots, \lambda_n$ in non-decreasing order.

14.2.1 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

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

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 $\boldsymbol{\phi}_0$ 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, $\boldsymbol{\phi}(t)$ at each node at a given time t is given by

$$\boldsymbol{\phi}(t) = \boldsymbol{\phi}_0 \, e^{-\mathbf{L}t}. \tag{9}$$

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

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

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

$$\lim_{t \longrightarrow \infty} = \boldsymbol{\phi}_0 \mathbf{v_r} \mathbf{v_l^T} \tag{11}$$

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 motivates the search for 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.

Proposition 1 *Consider a directed network $G = (V, E)$ that is strongly connected. Then the digraph G globally attains average equilibrium if and only if $\mathbf{1}^T \mathbf{L} = 0$.*

14.2.2 Equilibrium state for Balanced Graphs

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

Proposition 2 *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:*

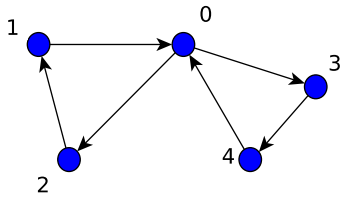
i) G is balanced,

ii) $\mathbf{v_l} = \mathbf{1}$ is the left eigenvector of the Laplacian of G associated with the zero eigenvalue, that is, $\mathbf{1}^T \mathbf{L} = 0$.

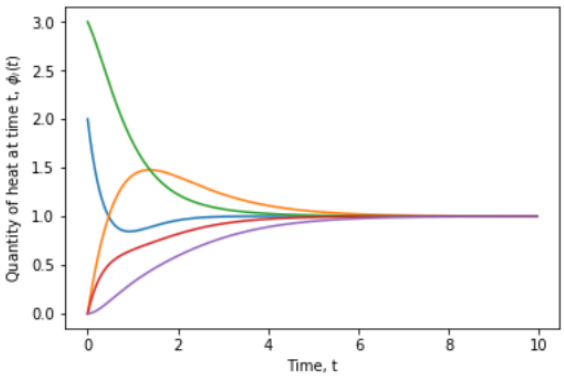
iii) $\sum_{i=1}^{n} u_i = 0, \forall x \in \mathbb{R}^n$ with $u_i = \sum_{j \in N_i} a_{i,j}(x_j - x_i)$.

Since for a balanced digraph $\mathbf{v_l}$ is an all ones vector, it therefore follows from Proposition 2 that at equilibrium, the value at all nodes in a balanced graph is the average of the initial values at all nodes.

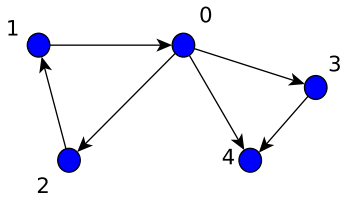
Example 2 *Let us consider two directed graphs, one is a balanced digraph and the other is not. We then assign initial quantities of heat to all nodes in the order 0 to 4 as in the vector $\phi_0 = [2, 0, 3, 0, 0]$ and set the diffusion coefficient, $C = 1$. We then obtain plots for diffusion on both graphs after a specific time t as shown in*



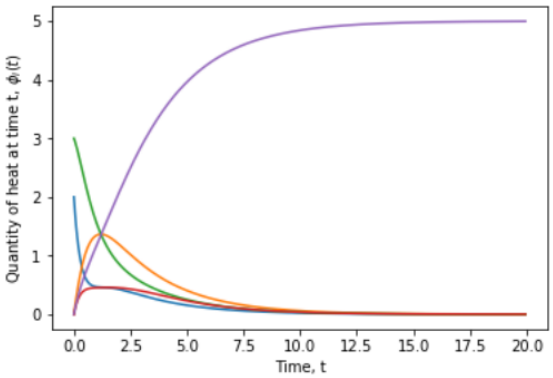
(a)



(b)



(c)



(d)

Figure 6: Diffusion over different categories of directed networks. (c) is an illustration of diffusion over weakly connected and unbalanced digraph in (a). (c) is an illustration of diffusion over strongly connected and balanced digraph (c).

For the balanced graph in Fig. 6a, its Laplacian matrix $\mathbf{L}$,v_r *and* v_l *are respectively:*

$$\mathbf{L} = \begin{pmatrix} 2 & 0 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad v_r = v_l = \begin{pmatrix} 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \end{pmatrix}$$

The values for v_r *and* v_l *satisfy Theorem 1 as well as Proposition 2 and thus, equilibrium is attained at* $x^* = 1.0$ *which is the average of initial values* x_0.
On the other hand, for the unbalanced graph in Fig. 6c, we have the following matrices

$$\mathbf{L} = \begin{pmatrix} 3 & 0 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad v_r = \begin{pmatrix} 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \end{pmatrix}, \text{ and } v_l = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \\ 1.0 \end{pmatrix}$$

We observe that $Lv_r = 0$ *and* $v_l^T L = 0$. *However, the condition* $v_l^T v_r = 1$ *is not satisfied and therefore equilibrium cannot be attained as shown in Fig. 6. In addition, we observe that considering the structure, vertex* 4 *has only in coming edges which signifies that during the diffusion process, vertex* 4 *only receives heat from the immediate neighbours vertices* 0 *and* 3 *without giving out any due to lack of out going links. As a result, quantity of heat at vertex* 4 *keeps on increasing as shown in Fig. 6.*

14.3 Diffusion on network with long-range interactions

The 'classical' case considers diffusion over a network where a substance under consideration say heat flows along the edges of the network. However, long range interactions during diffusive processes on networks are evident in real world situations. Such interactions result into superdiffussion on networks which has been modelled by various models that include the random walks with Levy flights (RWLF), model based on fractional diffusion equation (FDE), and many more. Recently, an elegant model has been put forward by Estrada (Estrada et al., 2017) which accounts for longrange interactions by use of k-path Laplacian matrices resulting into a generalised diffusion process on networks.

However, long-range interactions over networks have been evident in real world situations, for instance in attaining consensus in networked multi-agent systems, an agent is influenced not only by the nearest neighbours but also by the non-nearest neighbours.

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 epidermic spread in networks (Estrada et al., 2011). Here, the long range interactions are considered to be nonrandom and depend on the social distances between individuals in the social network. Recently, work in (Estrada, 2012) 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 form its current location, such a diffusion process can be captured by replacing $\mathbf{L}$ in (2) by $\mathbf{L}_k$ in (17). That is

$$\frac{d\boldsymbol{\phi}}{dt} = -C\mathbf{L}_k\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0. \tag{12}$$

14.3.1 k-path Laplacian matrices, $\mathbf{L}_k$

The k-path Laplacian matrices are a natural generalisation of the combinatorial Laplacian of a graph. The entries of these matrices account for the existence of shortest paths of length k between pairs of nodes in the network.

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:

$$\mathbf{L}_k(i,j) = \begin{cases} -1 & \text{if } d_{i,j} = k, \\ \delta_k(i) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

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 13 for finite undirected graphs was extended for connected and locally finite infinite graphs as follows: Consider $\Gamma = (V, E)$ to be an indirected finite or infinite graph with vertices V and edges E. We assume that Γ 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 Γ, 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.

$$\delta_k(v) := \#\{w \in V : d(v, w) = k\}. \tag{14}$$

Since γ 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

$$\langle f, g\rangle = \sum_{v\in V} f(v)\overline{g(v)}, \quad f, g \in \ell^2(V) \tag{15}$$

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

$$e_v(w) := \begin{cases} 1 & \text{if w = v,} \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

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

$$(\mathbf{L}_k)(f) := \sum_{w \in V : d(v,w)=k} (f(v) - f(w)), \quad f \in C(V). \tag{17}$$

14.3.2 Generalisation of diffusion equation on Graphs

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. In accounting for indirect interaction, we consider the fact that longer paths contribute less compared to shorter ones. This is achieved by considering Equation 2 where $\mathbf{L}$ is any of the transformed k-path Laplacians given as

i) Social Distance Method

In this method, long range interactions are accounted for based on an analogy of the time value of money in quantitative finance. An example of application of this method is in (Estrada et al., 2011) where it is indicated that for effective modelling of how airborne or close contact infections are transmitted, a contact network capturing the pattern of human interactions is a requirement. These interactions are of two fold: close contacts which are frequent interactions among individual while casual or long range interactions are the non frequent encounters among individuals which facilitate the spread of infections.The latter are considered to be non-random. The casual contacts are accounted for by means of long range interactions which are considered to depend on the social distance between two individuals drawn from the structure of the social network of close contacts. The weight assigned to long range interaction between a pair of nodes i and j is given by

$$e_{i,j} = d_{i,j} \cdot x^{d_{i,j}-1} \tag{18}$$

where $d_{i,j}$ is the shortest path distance between i and j and x is a parameter called the conductance as it controls the extent to which casual contacts are allowed in the network. From Equation 18, we can evidently observe that the strength of the casual contact between two nodes reduces with decrease in social distance between them.

ii) The Laplace-transformed k-Laplacian

$$\tilde{\mathbf{L}}_{L,\lambda} = \mathbf{L} + \sum_{k=2}^{\infty} e^{-\lambda k} \mathbf{L}_k \tag{19}$$

iii) The factorial-transformed k-Laplacian

$$\tilde{\mathbf{L}}_{F,z} = \mathbf{L} + \sum_{k=2}^{\infty} \frac{z^k}{k!} \mathbf{L}_k \tag{20}$$

iv) The Mellin-transformed k-Laplacian

$$\tilde{\mathbf{L}}_{M,s} = \sum_{k=1}^{\infty} k^{-s} \mathbf{L}_k, \tag{21}$$

In (Estrada et al., 2017), it is shown that normal diffusion occurs only when $s > 3$. On the other hand, superdiffusion occurs when $1 < s < 3$ with superdiffusive exponent being $\kappa = \frac{2}{s-1}$, which leads to arbitrary values for $\kappa \in (1, \infty)$. The Mean Square Distance, $MSD \sim t^{\kappa}$. Normal diffusion is attained at $\kappa = 1$, superdiffusion is attained at $\kappa > 1$. However, a special type of diffusion known as Ballistic diffusion is characterised by the fact that at small times the particles are not hindered yet by collisions and diffuse very fast. it is attained at $\kappa = 2$.

14.4 Spectrum of the Generalised Laplacian matrix

The spectrum of the Laplacian matrix is the set of eigenvalues and their multiplicities Estrada et al. (2011). Let $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ be the eigenvalues of $L(x)$!h and their corresponding multiplicities $m(\lambda_1), m(\lambda_2), \cdots, m(\lambda_n)$. The spectrum of L is given by

$$S_p L = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n m(\lambda_1) & m(\lambda_2) & \cdots & m(\lambda_n) \end{pmatrix} .!h \tag{22}$$

Some analytical expressions for the spectra of the Laplacian matrix of some common simple networks are:

- Star, S_n: $S_p(L) = 0\ 1^{n-2}\ n$
- Path, P_n: $S_p(L) = 2 - 2\cos\left(\frac{\pi(j-1)}{n}\right)$

For generalised Laplacian matrix $L(x)$, the above expressions as a function of conductance x are given by

- Star, S_n: $S_p(L(x)) = 0\ (1+8x)^{n-2}\ n$
- Path, P_n: $S_p(L(x)) =$

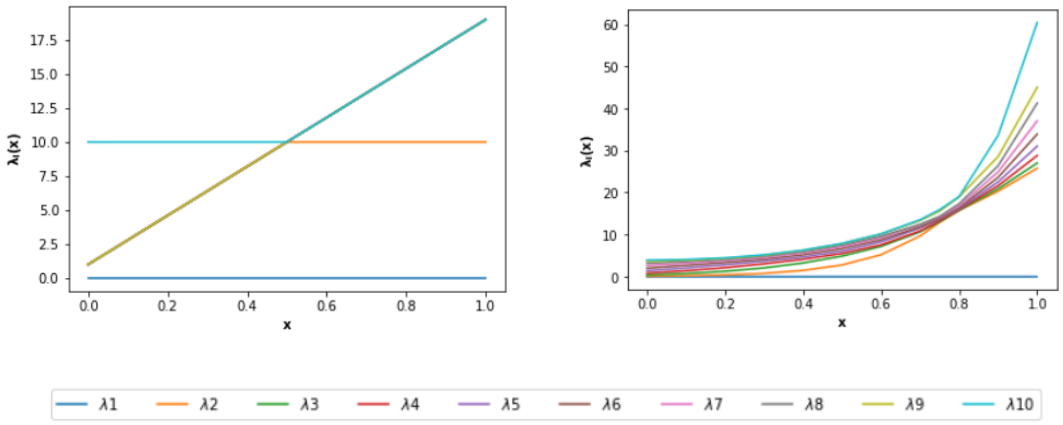


Figure 7: Illustrations of variation of eigenvalues with conductance x two networks. The first figure from the left a star network, S_{10} (in Fig. **??**) and the other is a path network. Both networks have 10 nodes each.

As discussed before, the eigenvalues $\lambda_i(x)$ of the Laplacian matrix $L(x)$ are an increasing function of x except for the smallest eigenvalue $\lambda_1(x) = 0$ which remains almost constant as observed in Figure 7. For the star network , both $\lambda_1(x) = 0$ and $\lambda_1(n) = 5$ remain constant the analytical expressions mentioned before. However, the rest of the eigenvalues increase linearly with increase in x values. At $x = 0.5$, we observe that all eigenvalues are equal with a value of 10. This is because at $x = 0.5$, all edges of the star network have a weight of 1 and thus, a complete network,K_n, whose spectrum is given by $S_p(L) = \{0\ n^{n-1}\}$. On the other hand, the path network follows a similar trend as that of the star network but the x value at which all eigenvalues (except $\lambda_i = 0$) are equal is relatively higher than 0.5 that is to say the point lies at $x = 0.7$.

14.5 The Heat Kernel

As discussed earlier, diffusion of heat on a graph can be modelled by the equation

$$\frac{d\phi}{dt} = -L\phi, \tag{23}$$

where L is either the Laplacian matrix or it's normalised version.

The heat kernel is the fundamental solution to the diffusion equation (23). It is obtained by exponentiating the Laplacian eigensystem and it is given by

$$\phi(t) = exp(-Lt) \tag{24}$$

It literally describes the flow of substance (heat) across edges (direct interactions) in the graph (Xiao et al., 2009).

When t tends to zero, the kernel behaviour can be obtained from the Taylor's expansion of (24) which is

$$e^{-\mathbf{L}t} = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}\mathbf{L}^k = \mathbf{I} - \mathbf{L}t + \frac{\mathbf{L}^2t^2}{2!} + \frac{\mathbf{L}^3t^3}{3!} + \cdots \tag{25}$$

Thus,

$$\lim_{t \longrightarrow 0}\left(e^{-\mathbf{L}t}\right) = \mathbf{I} - \mathbf{L}t. \tag{26}$$

It therefore implies that for t tending to zero, the heat kernel depends on the local connectivity structure of the graph. On the other hand, as t tends to infinity, we apply the spectral decomposition of Equation 24 which is

$$\phi(t) = \mathbf{V}e^{(-\Lambda t)}\mathbf{V}^{\mathbf{T}} = \sum_{i=0}^{n} e^{(-\lambda_i t)}v_i v_i^T \tag{27}$$

where λ_is are the eigenvalues of L in a non-decreasing order $0 = \lambda_1 \leq \lambda_2 \leq, \cdots \leq \lambda_n$ and v_i is the eigenvector corresponding to the eigenvalue λ_i. Thus,

$$\lim_{t \longrightarrow 0}\left(e^{-\mathbf{L}t}\right) = e^{(-\lambda_2 t)}v_2 v_2^T. \tag{28}$$

From Equation 28,its evident that for large t, the heat kernel behaviour is determined by the global structure of the graph.

14.5.1 The Heat Kernel Invariants

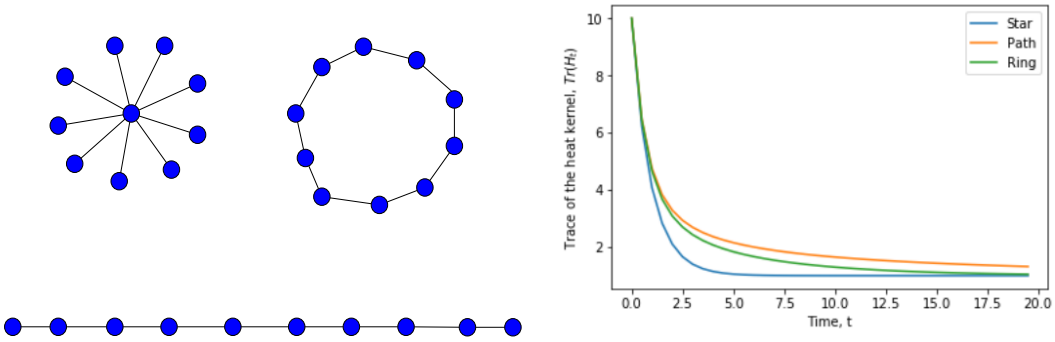
1 Heat Kernel Trace Like the trace of any other matrix, the trace of heat kernel is the summation of the main diagonal elements of the heat kernel matrix of a graph. It is therefore given by

$$Z(t) = Tr(h_t) = \sum_{i=1}^{|V|} e^{-\lambda_i t}, \tag{29}$$

where λ_i is the eigenvalue of the normalised Laplacian matrix Xiao et al. (2009). For a connected graph ($\lambda_1 = 0$), Equation 29 can be written as

$$Z(t) = 1 + e^{-\lambda_2 t} + e^{-\lambda_3 t} + \cdots + e^{-\lambda_N t} \tag{30}$$

Xiao Xiao et al. (2009) developed a potential application of the trace formula as a basis for distinguishing different graphs based on the shape of the curves obtained by plots of the trace of the heat kernel as a function of time. Let us consider 3 simple graphs namely a star, path and 2-regular graph of size 10. Fig shows the plot heat kernel trace against time for the three graphs.



(a) (b)

Figure 8: (8a) are the three graphs used for heat kernel analysis. (8b) plot of the heat kernel trace against time for star (blue), path (orange) and regular(green) graphs. From 8b, we observe that since the 3 graphs are different in terms of structure, the curves too take on different shapes. It is evident that since path and 2-regular graphs have similar topologies, their corresponding graphs are close to each other while for star graph, its curve has a deeper trough.

2 Zeta function The Zeta function associated with the Laplacian eigenvalues is obtained by exponentiating and summing the reciprocal of the non-zero Laplacian eigenvalues. It is thus defined by

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} \tag{31}$$

. Xiao Xiao et al. (2009) established a relationship between the zeta function and the heat kernel trace by use of the Mellin transform. For a function $f(t)$, it's Mellin transform is given by

$$F(s) = \int_0^\infty t^{s-1} f(t) dt. \tag{32}$$

Taking a function $f(t) = e^{-\lambda_i t}$, applying the Mellin transform gives

$$\lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} e^{-\lambda_i t} dt, \tag{33}$$

where $\Gamma(s)$ is the gamma function defined as

$$\Gamma(s) = \frac{0}{\infty} t^{s-1} e^{-t} dt \tag{34}$$

. On summation for all non-zero eigenvalues of the Laplacian, Equation 33 becomes

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \sum_{\lambda_i \neq 0} e^{-\lambda_i t} dt \tag{35}$$

The heat kernel trace can be also expressed as

$$Tr(h_t) = C + \sum_{\lambda_i \neq 0} e^{-\lambda_i t}, \tag{36}$$

where C is the multiplicity of zero eigenvalues of the Laplacian that is the number of connected components of a graph. Substituting Equation 36 into Equation 35 gives

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \{Tr(h_t) - C\} dt \tag{37}$$

. Thus the zeta function is related to the moments of the heat kernel trace. It is the moment generating function and thus a way of characterising the shape of the heat kernel trace.

3 Derivative of Zeta Function at the Origin The derivative or slope of the zeta function at the origin is another characterisation of the heat kernel trace second to the zeta function which measures it's shape.

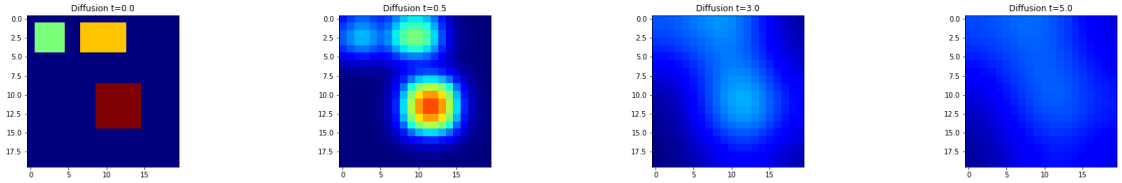
4 Heat Content

14.5.2 Simulations of diffusion on lattice

We consider a 2-dimensional discrete grid in which each point is connected to 8 of its nearest neighbours. Initially, we assign heat quantities to all the points on the grid and then we investigate how the diffusion process occurs and at each time t.

Let us take a 20 by 20 grid on which we assigned heat quantities of amounts 5, 7 and 10 to a few points and the rest are assigned zero. The diffusion on the lattice through direct interactions only as well as through both direct and indirect interactions by using the Social distance concept, Laplace and Mellin tranforms.

i) Diffusion on lattice through Direct interactions only



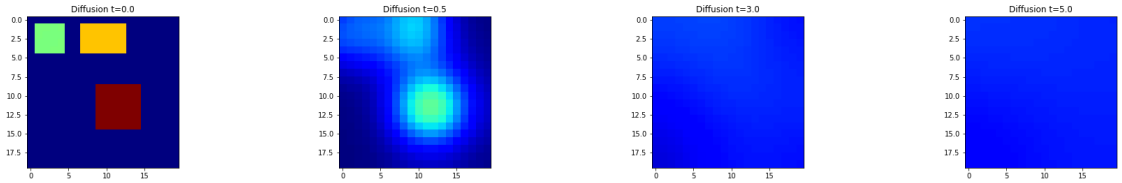
(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 9: Sample illustrations for progression of diffusion over a 20×20 lattice where specific regions are assigned particular heat quantities which in turn spread to other regions of low heat quantities through interactions along links in the network.

ii) Diffusion on lattice through both direct and indirect interactions. Considering a similar lattice with initial heat quantity assignments as in the directed case discussed before, we account for the diffusion process where interactions among nodes occurs through both direct and indirect interactions.

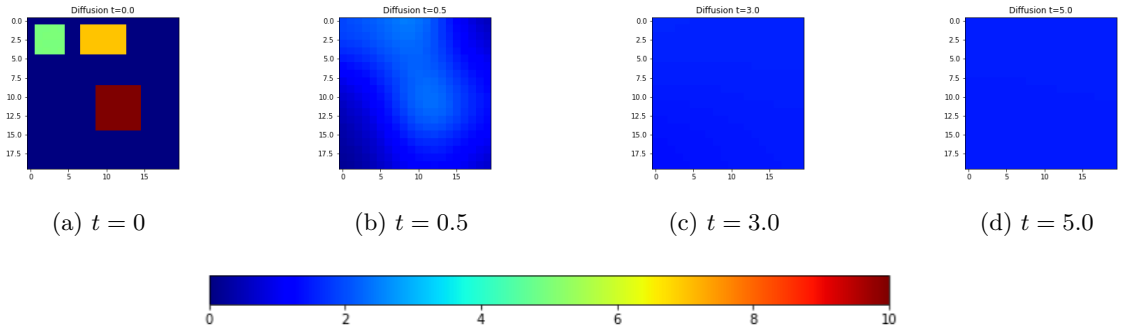
As discussed previously, long-range interactions can be accounted for by using various techniques as illustrated by the following illustrations.

a) Long range interactions using social distance Taking $x = 0.1$ and $x = 0.2$.



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

At $t = 0$, diffusion on the grid starts off with 3 strong regions having high quantities of heat as observed from figures 9, and 11 which correspond to diffusion with direct interactions only and diffusion with conductance $x = 0.1$ and 0.2 respectively. As the diffusion process continues, we see that at $t = 0.5$, strong heat points can still be spotted for direct interactions, relatively strong points in $x = 0.1$ and almost complete diffusion in $x = 0.2$. We can also observe that by $t = 2.0$, heat is uniformly distributed across the grid for $x = 0.1$ and $x = 0.2$. However, for the

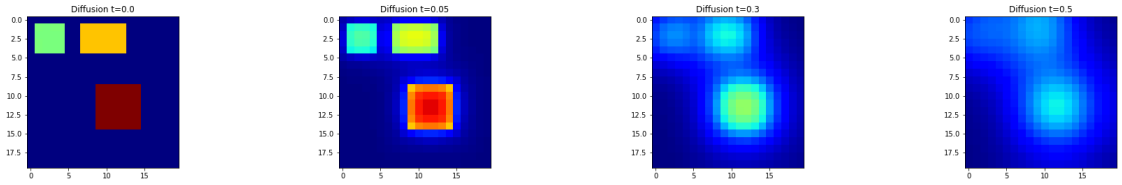


(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 11: Illustrations for diffusion over a grid with long-range interactions accounted for by the social distance technique. The upper row corresponds to diffusion with conductance $x = 0.1$ while lower row corresponds to results obtained $x = 0.2$. The intensity of heat follows a color grid where by red implies higher intensity followed by yellow and blue implies low heat intensities.

case of direct interactions (Fig.9d) diffusion is still ongoing and we can notice strong heat points at the centre of the grid. Following the sequences in the figures, we can conclude that as x (i.e increase in intensity of long range interactions), the diffusion process goes faster and equilibrium across the grid is reached faster as observed in the above simulations where for $x = 0.2$, $x = 0.1$ equilibrium is reached by $t = 3$ while for $x = 0$, equilibrium is not yet reached by then.

b) Longrange interactions using Laplace Transforms



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 12: Illustration for diffusion on the lattice with both direct and long range interactions. These longrange interactions are accounted for by Laplace transforms given by Equation 19 with $\lambda = 1$.

c) Longrange interactions using Mellin Transforms of the k-Laplacian matrices. For the case of

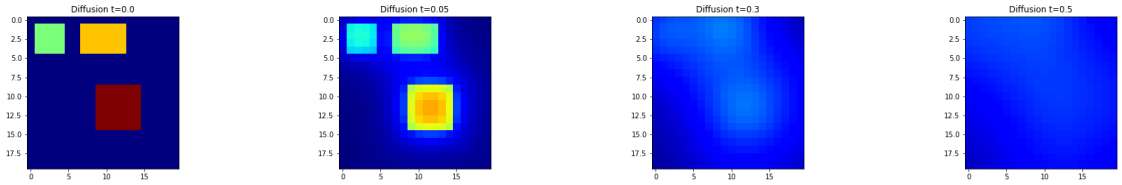


Figure 13

Mellin transforms, we observe that rate of diffusion is quite faster at $k = 2$(upper row) than at $k = 4$(lower row).This is due to the fact that at $k < 3$, super diffusion occurs, however normal diffusion occurs when $k > 3$ as shown in lower row of Fig.14.

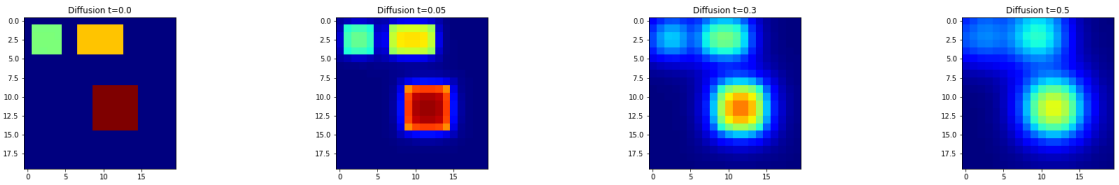


Figure 14: Illustrations of diffusion process on the lattice allowing for both direct and longrange interations. The latter are accounted for using Mellin transforms of the k-Laplacian matrices as in Equation. 21. The top row corresponds to the case for which $s = 2$ and the bottom row corresponds to case $s = 4$.

15 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.

15.1 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.

15.2 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

$$p(x) = \frac{1}{d_x} \sum_{y, y \in N(x)} p(y), \tag{38}$$

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

$$p(a) = 0 \tag{39}$$

$$p(b) = 1 \tag{40}$$

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

$$v(a) = 0 \tag{41}$$

$$v(b) = 1 \tag{42}$$

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

$$i_{x,y} = v(x) - v(y) \quad \forall x \neq y \neq b, \quad x, y \in E \tag{43}$$

Applying Kirchoff's law, we have

$$\sum_{x} \sum_{y} i_{x,y} = 0, \quad \sum_{y,x} \sum_{y} v(x) - v(y) = 0, \tag{44}$$

which gives

$$v(x) = \frac{1}{d_x} \sum_{y, y \in N(x)} v(y) \tag{45}$$

We observe that system of equation of the probabilities in (38,39,40) and that of the voltages (41,42,45) 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 $A = \{0, 1, 2, \cdots, N\}$ and let B be the set of boundary points $B = \{0, N\}$ and C be the set of interior points $C = \{1, 2, \cdots, N-1\}$. So the function $f(x)$ defined on A is harmonic if it satisfies the mean value property given as

$$f(x) = \frac{f(x+1) + f(x-1)}{2} \tag{46}$$

   From (38) and (45), we observe both functions p and v satisfy the mean value property in (46) 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 (39,40) and (41,42) 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.

15.3 The Random walker Algorithm

15.3.1 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

$$w_{i,j} = exp(-\beta(g_i - g_j)^2), \tag{47}$$

where β 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 (47). For instance, to capture colour, we replace $(g_i - g_j)^2$ with $\|g_i - g_j\|^2$ $\forall e_{i,j} \in E$ (Grady, 2006).

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 (Grady, 2006) 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.

15.3.2 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

$$\nabla^2 u = 0, \tag{48}$$

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

$$D[u] = \frac{1}{2}\int_{\Omega} |\nabla|^2 d\Omega, \tag{49}$$

for region Ω.
Let us consider an image represented as a graph G=(V,E) where nodes represent pixels while edges represent the neighbourhood among the pixels. The Combinatorial Laplacian matrix of the graph is defined as

$$L_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise,} \end{cases} \tag{50}$$

where $L_{i,j}$ is indexed by vertices v_i and v_j. The combinatorial formulation of the Dirichlet integral is

$$D[x] = \frac{1}{2}x^T L x = \sum_{e_{ij} \in E} w_{ij}(x_i - x_j)^2. \tag{51}$$

A combinatorial harmonic is a function x that minimises 51.

15.3.3 Implementation of Algorithm

1. Computing probabilities:

   With Graph $G = (V, E)$ representing the image to be segmented, partition the vertex set into two such that is V_m and V_u which are the seeded/marked and unseeded vertices respectively such that $V_m \bigcup V_u = V$ and $V_m \bigcap V_u = \emptyset$. Without loss of generality, we assume that the nodes in L are arranged such that the seeded nodes are first followed by the unseeded ones. Then equation 51 can be written as

   $$D[x_u] = \frac{1}{2}\begin{bmatrix} x_m^T x_m^T \end{bmatrix} \begin{bmatrix} L_m & B \\ B^T & L_u \end{bmatrix} \begin{bmatrix} x_m \\ x_u \end{bmatrix} = \frac{1}{2}(x_m^T L_m x_m + x_u^T B^T x_m + x_u^T L_u x_u), \tag{52}$$

   where x_m and x_u correspond to the potentials at the seeded and unseeded nodes respectively. Differentiating $D[x]$ with respect to x_u and then finding the critical points gives the following system of linear equations with $|V_u|$ unknowns

   $$L_u x_u = -B^T x_m. \tag{53}$$

   Let x_i^s be the probability of a random walker starting at node v_i first reaches label s. Then for each label s the probabilities with each unlabelled nodes first reach s is given by

   $$L_u x^s = -B^T m^s, \tag{54}$$

   and for all labels, we have

   $$L_u X = -B^T M, \tag{55}$$

   where X has K columns taken by each x^s and M has columns given by each m^s.

2. Assigning labels to unseeded nodes Having computed the probabilities for each unseeded nodes, we assign a label to each node for which the highest probability was obtained.

15.4 Possible application of Long range interactions to Image Segmentation

We considered a possible application of long range interaction to segmentation of image based on random walks. This was implemented by replacing the Combinatorial Laplacian matrix L in equation 55 with the the k-path Laplacian matrices to account for long range jumps of the random walker. However, the output obtained using this method was not a better image segmentation due to the fact that the long-range jumps do not take into account the localisation of objects in an image which results into less intuitive image segmentation.

16 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 (Freeman, 1978; Estrada and Rodriguez-Velazquez, 2005). 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 (Newman, 2001; Barrat et al., 2004; Opsahl, 2009; Opsahl et al., 2010). These standard centrality measures give information on either the local environment of a node (i.e degree centrality) or the global position of a 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 (Qi et al., 2012).

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 (Qi et al., 2012).

16.1 Laplacian energy of a network

Let $G = (V, E, W)$ be a simple undirected weighted network with the vertex set $V(G) = \{v_1, v_2, \cdots, v_n\}$, 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

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

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

Definition 5 (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*

$$E_L(G) = \sum_{i=1}^{n} x_i^2 + 2 \sum_{i<j} w_{i,j}^2. \tag{56}$$

Corollary 1 *If H is an arbitrary subgraph of a network G, then $E_L(H) \leq E_L(G)$. And equality holds if and only if $V(G) - V(H)$ is a set of isolated vertices.*

16.2 Laplacian centrality for a vertex

Definition 6 *If $G = (V, E, W)$ is a network with n vertices $\{v_1, v_2, \ldots, v_n\}$. Let G_i be the network obtained by deleting v_i from G. The Laplacian centrality is given by*

$$C_L(v_i, G) = \frac{(\Delta E)_i}{E_L(G)} = \frac{E_L(G) - E_L(G_i)}{E_L(G)} \tag{57}$$

For any vertex v, the denominator remains unchanged and from Corollary 1, 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 $w_i j$ copies of multiedges. For instance, for a 2-walk $v_1v_2v_3$ 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 $v_1, v_2, \ldots, v_n$. 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*

$$(\Delta E)_i = E_L(G) - E_L(G_i) = 4 \cdot NW_2^C(v_i) + 2 \cdot NW_2^E(v_i) + 2 \cdot NW_2^M(v_i). \tag{58}$$

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).

16.3 Laplacian Centrality of an edge

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, betweeness, subgraph, eigenvector, Laplacian centralities,etc.), interest in the study of edge centralities is now gaining ground with earlier works by Anthonisse (Anthonisse, 1971) and then following by prominent work by Girvan and Newman (Newman and Girvan, 2004). Some of the known edge centralities include edge degree, edge closeness (Ortiz Gaona et al., 2016), edge-betweenness which is calculated based on either shortest path distances or random walks as in (Newman and Girvan, 2004), k-path edge centrality Alahakoon et al. (2011), among others. The motivation for the introduction of edge centrality measures lies in real-world applications in a wide range of context such as community detection in networks (Newman and Girvan, 2004), identifying significant power lines, communication or transportation lines whose failure cause serious breakdown of the power, communication and transportation systems respectively (Ortiz Gaona et al., 2016), identification of strong relationships among people in social networks, etc.

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

$$E_L(G) = \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2, \tag{59}$$

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:

$$x'(v_i) = \begin{cases} x(v_i) - w_{v_1,v_2} & \text{if } v_i \sim e_{1,2}, \\ x(v_i) & \text{otherwise.} \end{cases} \tag{60}$$

The Laplacian energy of the subgraph is given by

$$E_L(H) = \sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1,v_2})^2 + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \tag{61}$$

Definition 7 (Laplacian centrality of an edge, $C_L(e)$) *The Laplacian centrality of an edge e, in Graph G is given by*

$$C_L(e) = \frac{E_L(G) - E_L(H)}{E_L(G)} = \frac{\Delta E_L}{E_L(G)} \tag{62}$$

From (62), 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 (59) and (61) is given by

$$\begin{aligned}
\Delta E &= \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - \Big(\sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1,v_2}^2)^2 + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \Big) \\
&= \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - \Big(\sum_{v_i \sim e_{1,2}} (x^2(v_i) - 2x(v_i) \cdot w_{v_1,v_2} + w_{v_1,v_2}^2) + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \Big) \\
&= 2 \sum_{v_i \sim e_{ij}} x(v_i).w_{v_1,v_2} - \sum_{v_i \sim e_{ij}} w_{1,2}^2 + 2 \cdot w_{v_1,v_2}^2 \\
&= 2 \sum_{v_i \sim e_{i,j}} (x(v_i) \cdot w_{v_1,v_2}) - 2 \cdot w_{v_1,v_2}^2 + 2 \cdot w_{v_1,v_2}^2 \\
&= 2 \sum_{v_i \sim e_{i,j}} x(v_i) \cdot w_{v_1,v_2} \\
&= 2 \sum_{v_i \sim e_{i,j}} w_{v_1,v_2} \sum_{u \in N(v_i)} w_{v_i,u} \\
&= 2 \cdot \sum_{v_1} w_{v_2,v_1} \sum_{u \in N(v_1); u \neq v_2} w_{v_1,u} + 2 \cdot \sum_{v_2} w_{v_1,v_2} \sum_{u \in N(v_2); u \neq v_1} w_{v_2,u} + 2 \cdot w_{v_1,v_2} \cdot w_{v_1,v_2} + 2 \cdot w_{v_2,v_1} \cdot w_{v_2,v_1}
\end{aligned}$$

$$\Delta E = 2 \cdot NW_2^U(v_1(E), v_2(M)) + 2 \cdot NW_2^U(v_1(E), v_2(M)) + 4 \cdot NW_2^C(v_1, v_2) \tag{63}$$

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 (63), 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 of interest.

Example 3 *Let us consider the weighted graph in Fig. 15. We compute the drop in Laplacian energy by First, we compute the Laplacian energy $E(G)$ of the graph as follows:*

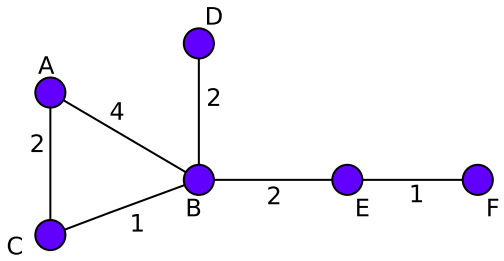


Figure 15

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

Table 1: Laplacian Centralities of edges

Edge(e)	$E(H); H = G - e$	$E(G) - E(H)$	$\Delta(E)$ by walks method in (63)
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 (63).

16.4 Laplacian Energy of Directed Networks

Though a lot of research has been inclined towards the study of undirected networks, however, directed networks are equally important as well. For instance, directed networks are used in the representation and study of structures of various real world networks such as communication networks, transportation networks, web-graphs among others.

The concept of Laplacian matrix and Laplacian energy of directed networks has been explored in various ways in Kissani and Mizoguchi (2010), Adiga and Smitha (2009).

Definition 8 (Skew Laplacian Energy of a simple, connected digraph) *Adiga and Smitha (2009) Let G be a simple (n, m) digraph with vertex set $V(G) = \{v_1, v_2, ..., v_n\}$ and arc set $\Gamma(G) \in V(G) \times V(G)$. The skew-adjacency matrix of G is the $n \times n$ matrix $S(G) = [a_{ij}]$ where $a_{ij} = 1$ whenever $(v_i, v_j) \in E(G), a_{ij} = -1$ whenever $(v_j, v_i) \in \Gamma(G), a_{ij} = 0$otherwise. Let $D(G) = diag(d_1, d_2, d_3, \cdots, d_n)$ the diagonal matrix with vertex degrees $d_1, d_2, \cdots, d_n$ for vertices $v_1, v_2, \cdots, v_n$. The skew Laplacian energy of a digraph G is defined as*

$$E_{SL}(G) = \sum_{i=1}^{n} \lambda_i^2 \tag{64}$$

where n is the order of G and $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of the Laplacian energy $L(G) = D(G) - S(G)$ of the digraph G.

Definition 9 (Laplacian Energy of a directed graph) *Kissani and Mizoguchi (2010) Let $A(G)$ be the adjacency matrix of a directed graph G whose entries are given as $A = (a_{ij})$, where $a_{ij} = 1$ whenever (v_i, v_j) is a directed edge and 0 otherwise. Let $D(G) = diag(d_1^{out}, d_2^{out}, \cdots, d_n^{out}$ be diagonal matrix with outdegree of the vertices $v_1, v_2, \cdots, v_n$. The Laplacian matrix $L(G) = D(G) - A(G)$. Then the Laplacian energy of G is defined as*

$$LE(G) = \sum_{i=1}^{n} \mu_i^2 \tag{65}$$

where n is the order of G and μ_i for $(i = 1, 2, \cdots, n)$ are the eigenvalues of $L(G)$.

From Definition 9, the following theorems hold:

Theorem 4 *Let G be a directed graph with vertex degrees $d_1^{out}, d_2^{out}, \cdots, d_n^{out}$. Then If G is a simple directed graph, then*

$$LE(G) = \sum_{i=1}^{n} (d_i^{out})^2 \tag{66}$$

If G is a symmetric directed graph (i.e a graph in which each edge is bidirected) then

$$LE(G) = \sum_{i=1}^{n} d_i^{out}(d_i^{out} + 1). \tag{67}$$

Theorem 5 *If G is a disconnected directed graph with components $G_1, G_2, \cdots, G_n$,*

$$LE(G) = \sum_{i=1}^{n} LE(G_i). \tag{68}$$

16.5 Laplacian Centrality of a node in directed network

Earlier on in this chapter, for simple undirected networks, we defined Laplacian Centrality of a vertex as the relative drop in Laplacian energy when the vertex is removed from the network. Similarly, we adopt the same definition for vertices in directed networks. Based on Definition 9, the drop in Laplacian energy of a simple directed graph $G = (V, E)$ due to the removal of vertex v is given by

$$\begin{aligned}
\Delta LE &= LE(G) - LE(G - v) \\
&= \sum_{i=1}^{n} (d_i^{out})^2 - \Big(\sum_{i \in N(v_{in})} (d_i^{out} - 1)^2 + \sum_{i \notin N(v_{in}); i \neq v} (d_i^{out})^2 \Big) \text{ where } N(v_{in}) = \{i \in N(v) | e_{iv} \in E\} \\
&= (d_v^{out})^2 + \sum_{i \in N(v_{in})} (2d_i^{out} - 1)
\end{aligned}$$

16.6 Edge Centrality based Edge reversal

Instead of measuring the importance of an edge based on the drop in energy on edge removal as discussed for undirected networks, we consider the importance of an edge in a directed network as the relative change in energy on edge reversal.

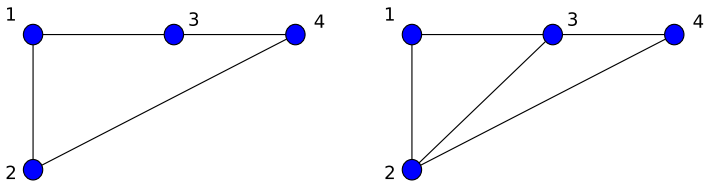
16.7 Laplacian Energy as a fair measure of robustness of network

Apart of its application in identifying the most important nodes and edges that is the Laplacian centrality, the Laplacian energy of a network has been recently identified as a measure of robustness of network (Yang et al., 2016). Robustness of networks is defined as the response of a network on edge or node removal or addition. The study of robustness of networks plays a vital role in design of systems, understanding the performance and stability of systems such as ecological systems, technological systems, biological system

among others. For instance, in air transportation sector, it is paramount to design air traffic networks that will ensure robustness when one or more flights are removed or added to the network thereby posing a need for an effective measure of robustness in such networks. The algebraic connectivity which is the second smallest eigenvalue (λ_2) of the Laplacian matrix of a network is one of the most common measure of robustness in networks (Jamakovic and Van Mieghem, 2008; Byrne et al., 2005). Unfortunately, the algebraic connectivity captures only the global information about the connectivity of a network which implies that some changes in network through addition or removal of edges may not be captured as the algebraic connectivity remains constant. However, work in (Yang et al., 2016) introduces the Laplacian energy of a network as a fair and effective measure of robustness compared to the algebraic connectivity. This is so because the Laplacian centrality captures the local information of the network thus a change in network structure by edge or node removal or addition is captured by a change in Laplacian energy. This therefore justifies Laplacian energy as a fair and effective measure of robustness compared to algebraic connectivity.

16.8 Comparison of Laplacian Energy and Algebraic Connectivity as measures of robustness of networks

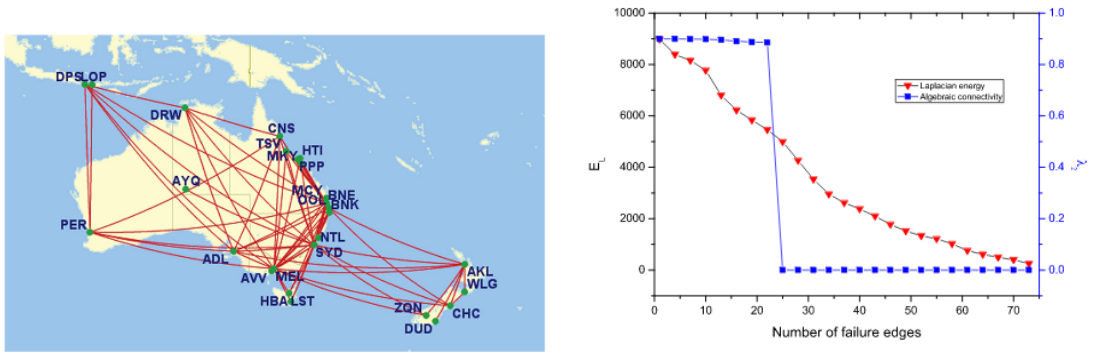
Let us consider a simple example in which we illustrate how the both the algebraic connectivity and Laplacian energy capture the changes in networks due to edge addition or removal.



(a) (b)

Figure 16: The simple network in (a) has algebraic connectivity $\lambda_2 = 2$ and Laplacian energy of 24. Adding a new edge,$e_{2,3}$ to form a network in (b) with algebraic connectivity $\lambda_2 = 2$ and Laplacian energy 36. We observe that on a change in network through edge addition, the algebraic connectivity remains constant while the Laplacian energy changes to reflect a change in network structure.

Further more, let us consider the comparison between the two measures on real work network that is the air traffic network of Jet-star Asia Airway among Indonesia, Australia, and New Zealand shown in (Fig. 17a).



(a) (b)

Figure 17: (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 λ_2 against the number of randomly failed edges from the or Jetstar Asia Airway (Indonesia, Australia, and New Zealand). Source: (Yang et al., 2016)

From the Fig. (17b), 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

C. Adiga and M. Smitha. On the skew laplacian energy of a digraph. In *Int. Math. Forum*, volume 4, pages 1907–1914, 2009.

T. Alahakoon, R. Tripathi, N. Kourtellis, R. Simha, and A. Iamnitchi. K-path centrality: A new centrality measure in social networks. In *Proceedings of the 4th workshop on social network systems*, page 1. ACM, 2011.

R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

J. M. Anthonisse. The rush in a directed graph. *Stichting Mathematisch Centrum. Mathematische Besliskunde*, (BN 9/71):1–10, 1971.

A.-L. Barabási. *Network science*. Cambridge university press, 2016.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

A. Barrat, M. Barthelemy, 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.

R. Byrne, J. Feddema, and C. Abdallah. Algebraic connectivity and graph robustness. *Sandia National Laboratories, Albuquerque, New Mexico*, 87185, 2005.

K. C. Das. The laplacian spectrum of a graph. *Computers & Mathematics with Applications*, 48(5): 715–724, 2004.

P. De Meo, E. Ferrara, G. Fiumara, and A. Ricciardello. A novel measure of edge centrality in social networks. *Knowledge-based systems*, 30:136–150, 2012.

C. H. Edwards and D. E. Penney. *Differential equations and boundary value problems*, volume 2. Prentice Hall, 2004.

W. Ellens and R. E. Kooij. Graph measures and network robustness. *arXiv preprint arXiv:1311.5064*, 2013.

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 N. Hatano. Communicability in complex networks. *Physical Review E*, 77(3):036111, 2008.

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.

E. Estrada, E. Hameed, N. Hatano, and M. Langer. Path laplacian operators and superdiffusive processes on graphs. i. one-dimensional case. *Linear Algebra and its Applications*, 523:307–334, 2017.

L. Euler. Leonhard euler and the königsberg bridges. *Scientific American*, 189(1):66–70, 1953.

L. Euler. The solution of a problem relating to the geometry of position. 1976.

L. C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

J. Gower. A modified leverrier-faddeev algorithm for matrices with multiple eigenvalues. *Linear Algebra and its Applications*, 31:61–70, 1980.

L. Grady. Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 28(11):1768–1783, 2006.

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.

B. A. Huberman. The laws of the web, 2001.

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&imgrc=_. [Online; accessed 2016-03-29].

A. Jamakovic and P. Van Mieghem. On the robustness of complex networks by using the algebraic connectivity. *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, pages 183–194, 2008.

R. Kasprzak. Diffusion in networks. *Journal of Telecommunications and Information Technology*, pages 99–106, 2012.

P. Kissani and Y. Mizoguchi. Laplacian energy of directed graphs and minimizing maximum outdegree algorithms. *Kyushu University Institutional Repository*, 2010.

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.

J. Magouirk, S. Atran, and M. Sageman. Connecting terrorist networks. *Studies in Conflict & Terrorism*, 31(1):1–16, 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.

M. E. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.

Nexus. Zachary's karate club [karate], 2012. URL 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.

R. Ortiz Gaona, M. Postigo Boix, and J. L. Melus Moreno. Centrality metrics and line-graph to measure the importance of links in online social networks. *International Journal of New Technology and Research*, 2(12):20–26, 2016.

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.

A. Sydney, C. Scoglio, P. Schumm, and R. E. Kooij. Elasticity: topological characterization of robustness in complex networks. In *Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Sytems*, page 19. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

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.

J. A. Williams, S. M. Dawson, and E. Slooten. The abundance and distribution of bottlenosed dolphins (tursiops truncatus) in doubtful sound, new zealand. *Canadian journal of zoology*, 71(10):2080–2088, 1993.

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

B. Xiao, E. R. Hancock, and R. C. Wilson. Graph characteristics from the heat kernel trace. *Pattern Recognition*, 42(11):2589–2606, 2009.

C. Yang, J. Mao, and P. Wei. Air traffic network optimization via laplacian energy maximization. *Aerospace Science and Technology*, 49:26–33, 2016.

W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):452–473, 1977.