Critical Node Identification for Accessing Network Vulnerability, A Necessary Consideration

By

Waqar Asif

120050615

Supervisors

Dr. Muttukrishnan Rajarajan

Dr. Marios Lestas

Dr. Hassaan Khaliq Qureshi

Dr. Veselin Rakocevic

A THESIS SUBMITTED TO THE SCHOOL OF ENGINEERING AND MATHEMATICAL SCIENCES, CITY, UNIVERSITY OF LONDON IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

(Oct 2016)

City, University of London
Northampton Square
London
EC1V 0HB
United Kingdom

T +44 (0)20 7040 5060

THE FOLLOWING CHAPTERS HAVE BEEN PUBLISHED AND HAVE BEEN REMOVED FROM THIS THESIS FOR COPYRIGHT REASONS:

pp. 31-37: Chapter 3. APL.
Published as: Asif, W., Qureshi, H.K., Iqbal, A. & Rajarajan, M. (2014) "On the Complexity of Average Path Length for Biological Networks and Patterns", *International Journal of Biomathematics,* 7(4).
doi: 10.1142/S1793524514500387

pp. 39-73: Chapter 4. VRAM.
Published as: Fiaz, M., Yousaf, R., Hanif, M., Asif, W. Qureshi, H.K. & Rajarajan, M. (2014), "Adding the Reliability on Tree Based Topology Construction Algorithms for Wireless Sensor Networks", *Wireless Personal Communication*, 74 (2), pp 989-1004.
doi: 10.1007/s11277-013-1334-2

Asif, W , Qureshi, H.K. & Rajarajan, M.(2013) "Variable Rate Adaptive Modulation (VRAM) for introducing Small-World model into WSNs," *47th Annual Conference on Information Sciences and Systems (CISS)*, pp.1-6. doi: 10.1109/CISS.2013.6552329

pp. 75-112: Chapter 5. CBDI.
Published as: Asif, W., Qureshi, H.K., Rajarajan, M. & Lestas, M. (2016) "Combined Banzhaf & Diversity Index (CBDI) for critical node detection." *Journal of Network and Computer Applications* 64, 76-88.
doi: 10.1016/j.jnca.2015.11.025

pp. 114-155: Chapter 6. Optimization based spectral partitioning for node criticality assessment.
Published as: Asif, W., Lestas, M,. Quereshi, H.K. & Rajarajan, M. (2016)"Optimization Based Spectral Partitioning for Node Criticality Assessment", *Journal of Network and Computer Applications* 75, 279-292.
doi: 10.1016/j.jnca.2016.09.003

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text. This dissertation contains fewer than 65,000 words including bibliography, footnotes, tables and equations and has fewer than 150 figures.

Waqar Asif

Acknowledgement

First And foremost, I am immensely thankful to Almighty Allah for letting me pursue and fulfil my dreams. Nothing could have been possible without His blessings.

I would like to thank my parents for their support throughout my educational career, specially during my Ph.D years, where their unending financial and moral support has helped me in completing this work. I would also like to thank my sisters, my elder brother and my wife for always being their for me at all times. I wish to express my gratitude to my academic advisors: Dr. Muttukrishnan Rajarajan, Dr. Marios Lestas, Dr. Hassaan Khaliq Qureshi and Dr. Veselin Rakocevic for guiding me throughout this work. I specially want to thank Dr. Marios Lestas for being more of a friend than a supervisor to me and guiding me through in every possible matter in life and for making my stay in Cyprus one of the best moments. I must acknowledge Dr. Arshad Ali and Dr. Christos Themistos for their sincere efforts and for selecting me as an exchange student for Cyprus. Here I would also like to thank everyone from the coordinating team of Erasmus Mundas Strong Ties who sponsored my stay in Cyprus and helped me groom both as a person and as a researcher.

List of Abbreviations

CBDI . . . . . Combined Banzhaf & Diversity Index

WSN . . . . . Wireless Sensor Network

CDS . . . . . Connected Dominating Set

APL . . . . . Average Path Length

VRAM . . . . Variable Rate Adaptive Modulation

NAW . . . . . Neighbour Avoiding Walk

HILPR . . . . Hybrid Interactive Linear Programming Rounding

QAM . . . . . Quadrature Amplitude Modulation

BFS . . . . . . Breath First Search

SNR . . . . . Signal to Noise Ratio

BER . . . . . Bit Error Rate

RF . . . . . . Radio Frequency

WFB . . . . . Wireless Flow Betweenness

CSMA/CA . Carrier Sense Multiple Access with Collision Avoidance

NS3 . . . . . . Network Simulator 3

OLSR . . . . Optimized Link State Routing

Cont . . . . . Controllability of complex networks

NUM . . . . . Network Utility Maximization

LHS . . . . . Left Hand Side

RHS . . . . . Right Hand Side

SSD . . . . . . Sum Squared Difference

NSSD . . . . Normalized Sum Squared Difference

SPNC . . . . Spectral Partitioning for Node Criticality

List of Notations/Symbols

$G(V, E)$ . . . . Graph with V vertices and E edges

l . . . . Average Path Length

$d(v, w)$ . . . . Geographic distance between v and w

N . . . . Number of nodes

$CC(v)$ . . . . . Closeness centrality of v

σ_{uv} . . . . Shortest path between u and v

$\kappa_{uv}(x)$ . . . . . Betweenness centrality of node x

F_{sal} . . . . . . maximum power flowing from s to d through node a

F_{sl} . . . . maximum power flowing from source s to destination l

$C_B(a)$ . . . . . Flow betweenness

λ_1 . . . . First Eigenvalue

v_t . . . . Eigenvector corresponding to node t

P_f . . . . Packet forwarding probability

$R(P_f, l)$ . . . . Packet delivery reliability

A . . . . Adjacency matrix

L . . . . Laplacian matrix

$G_d(v, w)$ . . . . Minimum distance between nodes v and w in graph G

P_b . . . . Probability of bit error

b . . . . Constellation size

M . . . . Modulation constellation

N_0 . . . . Noise power

γ_b . . . . average signal to noise ratio

E_s . . . . Symbol energy

E_b . . . . Bit energy

α . . . . Drain efficiency

d_{16} . . . . Distance covered by 16 QAM modulation

d_{64} . . . . Distance covered by 64 QAM modulation

$D_d(x)$ . . . . . Variation in link length of node x

T_x . . . . Transmission radius of node x

$N(x)$ . . . . . Neighbours of node x

$D_n(x)$ . . . . . Weighted node degree of node x

$H(x)$ . . . . . Diversity index of node x

I . . . . Set of all source destination pairs

$w(i, j)$ . . . . . Source destination pair node i and j

$Sp(w)$ . . . . . Set of nodes participating in the shortest path for pair w

C_k . . . . Banzhaf power index

$\delta(G)$ . . . . . . Diagonal degree matrix

$N(G)$ . . . . . Incidence matrix

D_{min} . . . . . . Minimum node degree

D_{max} . . . . . Maximum node degree

$P(u, v)$ . . . . Probability of having a link between u and v

ζ . . . . Probability of having an edge between any two nodes

ξ . . . . Ratio of long distance to short distance nodes

χ . . . . Distance between two nodes

a_l . . . . Column matrix

$\mu(L)$ . . . . . . Algebraic connectivity of L

L_o . . . . Laplacian matrix of the original network

h_k . . . . Column matrix of the node k being removed

x_k . . . . Boolean variable

$U_s(x_s)$ . . . . . Utility function assigned to the used x_s

$Lk(s)$ . . . . . Set of links associated with source s

$\vec{y_l}$ . . . . Unit vector associated with link l

n_l . . . . Cardinality of l

$\vec{i_s}$ . . . . Unit vector along the direction of x_s

Abstract

Timely identification of critical nodes is crucial for assessing network vulnerability and survivability. This thesis presents two new approaches for the identification of critical nodes in a network with the first being an intuition based approach and the second being build on a mathematical framework. The first approach which is referred to as the Combined Banzhaf & Diversity Index (CBDI) uses a newly devised diversity metric, that uses the variability of a node's attributes relative to its neighbours and the Banzhaf power index which characterizes the degree of participation of a node in forming the shortest path route. The Banzhaf power index is inspired from the theory of voting games in game theory whereas, the diversity index is inspired from the analysis and understanding of the influence of the average path length of a network on its performance. This thesis also presents a new approach for evaluating this average path length metric of a network with reduced computational complexity and proposes a new mechanism for reducing the average path length of a network for relatively larger network structures. The proposed average path length reduction mechanism is tested for a wireless sensor network and the results compared for multiple existing approaches. It has been observed using simulations that, the proposed average path

length reduction mechanism outperforms existing approaches by reducing the average path length to a greater extent and with a simpler hardware requirement.

The second approach proposed in this thesis for the identification of critical nodes is build on a mathematical framework and it is based on suboptimal solutions of two optimization problems, namely the algebraic connectivity minimization problem and a min-max network utility problem. The former attempts to address the topological aspect of node criticality whereas, the latter attempts to address its connection-oriented nature. The suboptimal solution of the algebraic connectivity minimization problem is obtained through spectral partitioning considerations. This approach leads to a distributed solution which is computationally less expensive than other approaches that exist in the literature and is near optimal, in the sense that it is shown through simulations to approximate a lower bound which is obtained analytically. Despite the generality of the proposed approaches, this thesis evaluates their performance on a wireless ad hoc network and demonstrates through extensive simulations that the proposed solutions are able to choose more critical nodes relative to other approaches, as it is observed that when these nodes are removed they lead to the highest degradation in network performance in terms of the achieved network throughput, the average network delay, the average network jitter and the number of dropped packets.

Publications

The following seven publications are part of this thesis:

0.1 Conference Publications:

1. Waqar Asif, Hassaan Khaliq Qureshi, Muttukrishnan Rajarajan, "Variable Rate Adaptive Modulation (VRAM) for introducing Small-World model into WSNs," in 47th Annual Conference on Information Sciences and Systems (CISS), 2013, pp,1-6.

2. Waqar Asif, Hassaan Khaliq, Muttukrishnan Rajarajan, Marios Lestas, "CBDI: Combined Banzhaf & diversity index for finding critical nodes," Global Communications Conference (GLOBECOM), 2014 IEEE, Austin, TX, 2014, pp. 758-763.

3. Waqar Asif, Marios Lestas, Hassaan Khaliq, Muttukrishnan Rajarajan, "Spectral partitioning for node criticality," 2015 IEEE Symposium on Computers and Communication (ISCC), Larnaca, 2015, pp. 877-882.

0.2 Journal Publications:

1. Mahzeb Fiaz, Roomana Yousaf, Maryam Hanif, Waqar Asif, Hassaan Khaliq Qureshi, Muttukrishnan Rajarajan, "Adding the Reliability on Tree Based Topology Construction Algorithms for Wireless Sensor Networks". Wireless Personal Communication, January 2014, Volume 74, Issue 2, pp 989-1004.

2. Waqar Asif, Hassaan Khaliq Qureshi, Adnan Iqbal, Muttukrishnan Rajarajan, "On the Complexity of Average Path Length for Biological Networks and Patterns", International Journal of Biomathematics 7.04 (2014): 1450038.

3. Waqar Asif, Hassaan Khaliq, Muttukrishnan Rajarajan, Marios Lestas, "Combined Banzhaf & Diversity Index (CBDI) for critical node detection." Journal of Network and Computer Applications 64 (2016): 76-88.

4. Waqar Asif, Marios Lestas, Hassaan Khaliq, Muttukrishnan Rajarajan, "Optimization Based Spectral Partitioning for Node Criticality Assessment", Journal of Network and Computer Applications 75 (2016): 279-292.

Table of Contents

List of Figures

Chapter 1

Introduction

The importance of Graph theory was first recognized by Euler in 1736. He used it to identify a suitable path with which a single person could pass through seven bridges in the city of Königsberg exactly once and return to the starting point [?]. Euler not only proved that such a path does exist, but also gave a general solution that could be applied to any arbitrarily arranged landmass and bridge structure. He also identified that the physical distance and the geographical locations of the bridges were not important for identifying the correct solution and what matters is the geometric position of the bridges.

A graph is a mathematical representation of a network which comprises of interconnected components known as nodes with the links between these nodes known as edges. A graph can be represented in a number of different ways: an undirected graph depicts no directional information to the connections whereas, a directed graph denotes the direction of flow of information through the links. Moreover, in binary graphs, the presence of an edge is denoted by a one and the absence of an edge is denoted by a zero, whereas, in a weighted graph the interconnection strength is quantified as weights of the links. Furthermore, the density of connections can range from fully connected graphs which are also referred to as completely connected graphs to very sparse graphs.

A graph can easily depict an abstraction of the reality and this is why graph theory methods have been widely used for understanding a wide range of systems. In a graph theoretic representation, network components are represented in terms of nodes and edges that connect these nodes. In a transport geography most networks have an obvious spatial foundation, namely the road and rail networks, which tend to be defined more by their links than by their nodes. This is not necessarily the case for all transportation networks. For instance, maritime and air networks tend to be more defined by their nodes than by their links since links are often not clearly defined. A telecommunication system can also be represented as a network, while its spatial expression can have limited importance and would actually be difficult to represent. Mobile telephone networks or the internet, possibly the most complex graphs to be considered, are relevant cases of networks having a structure that can be difficult to symbolize. However, cellular phones and antennas can be represented as nodes while the links could be individual phone calls. Routers, the core of the internet, can also be represented as nodes within a graph while the physical infrastructure between them, namely fiber optic cables, can act as links. Consequently, all transport/communication networks can be represented by graph theory in one way or the other.

Every graph differs from the other based on the attributes of its individual nodes and edges, where for example, attributes of a node comprise of the nodes location and the attributes of an edge incorporates its length and capacity. These individual components of a network influence their individuality upon the other thus enabling researchers to analyse carefully the characteristics of a network by only monitoring one set of components, either nodes or edges. The interconnectivity of these individual components defines the structure of a network and in the past a lot of research has been done in identifying prominent/vital network structures [**?**][**?**][**?**]. For example, in a multihop Wireless Sensor Network (WSN), nodes can be connected using various edges, each having a smaller length, compared to a conventional WSN, for the reduction in transmission energy consumption for the network. Nodes in a network are evaluated based on both their geographical location and the combined influence of all the edges that are connected to that node. The geographical location of a node helps approximate the traffic flow rate through nodes as it is established that a nodes close to the center of the network will experience a higher traffic flow compared to nodes close to the edge of a network. The later, on the other hand has its own importance, such as, a node with a higher number of edges is neighbour to a larger number of nodes in the network and therefore, it is critical for ensuring connectivity of the network. More elaboration of this phenomenon is explained later in this thesis.

The combined affect of both the aforementioned attributes defines the importance of a node in a network. Due to these attributes, there are a few nodes in a network which when removed result in disconnecting a chunk of the network and thus affect the performance of a network. These nodes are referred to as the articulation points.

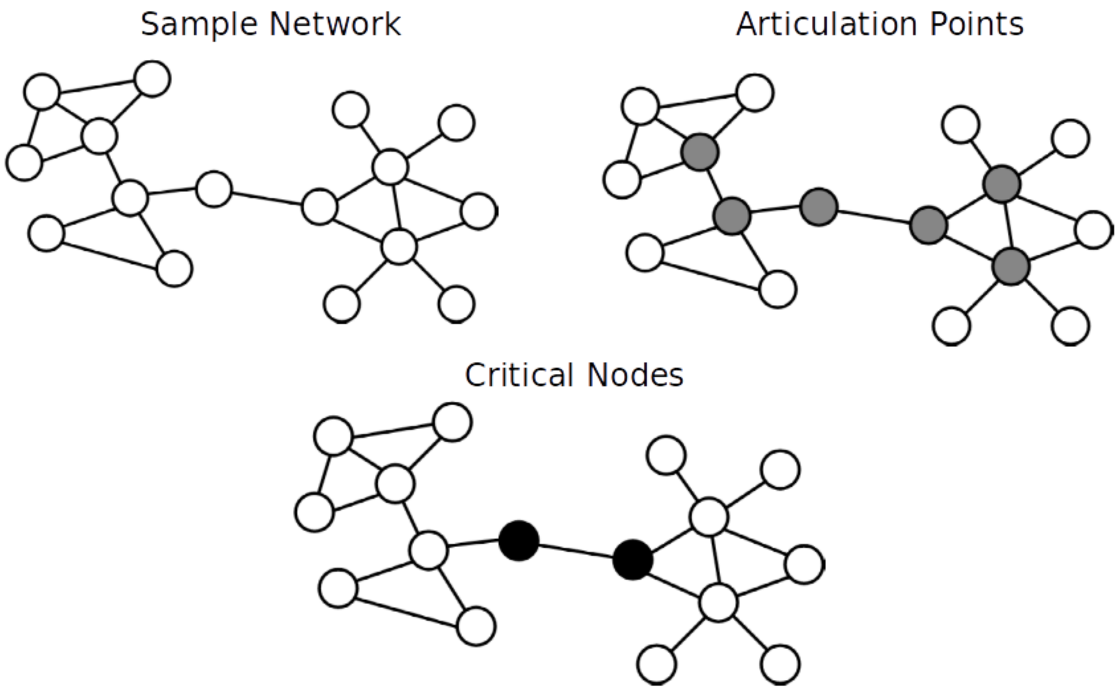

Figure 1.1: Difference in articulation points and critical node.

In the sample network of Fig **??**, we represent the articulation points in grey color. It is clear from the figure that, removal of any of these articulation points will render the network disconnected, where we use the term disconnected for a network in which every node is not accessible by every other node in the network. A few of these articulation points have shown to report a higher reduction in the performance of the network and these points which are coloured in black are referred to as the critical nodes of a network.

The aforementioned example is a special case in point where critical nodes are also articulation points, researchers have defined critical nodes in many different ways in literature, Wehmuth et al. in [?] refer to a node as critical based on how much affect it has on the notion of network centrality, a node that has a higher network centrality, is referred to as a critical node in a network. On the other hand, Ashwin et al. in [?] and Mario et al. in [?] refer to nodes as critical based on the affect they have on the pair-wise connectivity of the network upon there removal. Moreover, Nasiruzzaman et al. in [?] refer to a node as critical if it has maximum network traffic flow through it, the removal of such a node will reduce the network traffic flow to the minimum and is thus referred to as the most critical node in the network. In this thesis, we combine the above stated definitions and form a new definition that is used in this work. We refer to a node as critical if its removal has the highest influence on the performance of the network and we evaluate the performance of a network in both aspects, the connectivity of a graph and the network traffic flow that it is observing, therefore, according to the above stated definition, a node is referred to as critical if by its removal, the network observes maximum dis-connectivity and the network traffic flow observes the highest amount of decrease.

The identification of these critical nodes is vital for assessing the vulnerability of a network and this is the main motivation behind this thesis. The next section explains in detail the motivation.

1.1 Motivation

Evaluation of node criticality is significant in various complex networks. A few nodes in the network which are referred to as the critical nodes have been shown in literature to have a higher impact on the performance of a network [?][?][?]. This initiates the need for timely identification of these critical nodes for the purpose of timely rectifications and avoidance of any unexpected/unwanted network performance changes.

The importance of critical node identification was reignited when a Georgian woman in march 2011 disconnected 90% of Armenia from the access to the internet by accidentally sabotaging an optical fiber that was passing by her house [?]. She was scavenging for copper to sell as scrap when she came across this cable. Coincidently she had cut the only fiber cable that was connecting 3.2million Armenian people, thus depriving them from the access to the internet for continuous 5 hours. This highlighted the fact that despite the looks of the internet as depicted in Fig **??** the removal of a single critical node can have a high affect on the performance of the network.

The influence of critical nodes is not only limited to the internet, but it is also reflected in other fields such as a Peer to Peer Gnutella Network which reported a major network fragmentation after a removal of 4% of the most critical nodes of a network [?] and the North American power grid network which reported a 60% loss of network connectivity upon the removal of only 4% of the nodes in the network [?].

Similarly, in transportation networks [?], the need to identify critical nodes has increased with the ever increasing population which provokes the need for having a better and reliable network. These transportation networks, are prone both to the predictable human intervention and the unpredictable natural disasters such as hurricanes, floods and earthquakes. The more predictable human interventions can

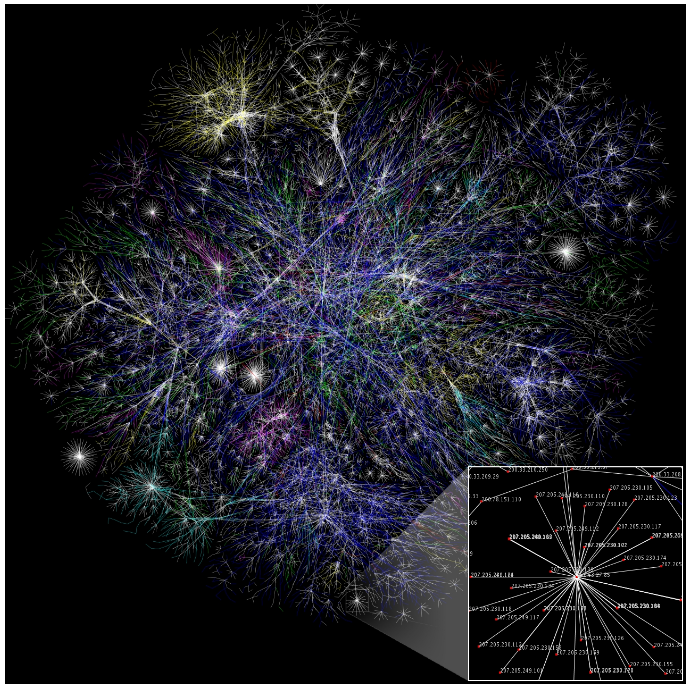

Figure 1.2: Partial map of the Internet based on the data found on January 15, 2005 [?].

cause network blockages due to two broadly defined reasons, either a regular network edge between two critical points is observing blockage due to limited link capacity or the transportation network observes occasional network blockages, as for example near a couple of famous touristic spots (critical nodes) in the network. In the former scenario, an efficient critical node detection algorithm will help in identifying such a critical link and thus to avoid the blockage, an alternate re-route can be designed that avoids the identified link. In the later scenario, the need of an extra special edge across all the critical nodes of the network can be avoided by efficiently re-routing traffic through edges that are not commonly used by various critical nodes of the network, thus avoiding network blockage. Regarding the the unpredictable nature of the natural disasters, proactive measures can also help in improving the pace of the rescue and recovery processes [?]. The proactive identification of critical nodes such as schools and hospitals can aid the rescuers to act quickly by using the shortest and the most affective paths and similarly for the reconstruction of an area that has been hit by a natural disaster such as a hurricane, the identification of critical nodes will aid the authorities in deciding as to which roads should be built first.

Likewise, in Telecommunication networks [?], the need to identify these critical nodes has never been more important than now. These days, with the introduction of smart phones for the purpose of increasing connectivity between people and for making life easier, we have also increased the risk of sabotaging our privacy by creating nodes (such as cellphones and tablets) that have all our vital information. A single virus that reaches our smart device can extract vital information such as our credit card details, our home address, can have access to our email and the list goes on. These viruses/bugs travel through our telecommunication network and thus to prevent the spread of such viruses, it is essential to timely identify the critical nodes and suppress their communication and thus avoid spreading these viruses and also maintain normal functionality for the rest of the customers [?]. Similarly, in biological networks [?], the detection of critical nodes can aid in neutralizing potentially harmful organisms such as bacteria and viruses. The interaction of protein with other proteins in a network can be represented in terms of graph theory and it is generally referred to as the protein-protein interaction network. These structures provide vital information for understanding biological structures and thus are widely used for designing drugs [?]. In particular, drugs are designed to affect the minimum cardinality set of proteins (the critical node set of a graph) whose removal will destroy the primal interaction and thus help neutralize the potentially harmful organism. The next section, explains in detail the problem statement of this thesis.

1.2 Problem Definition

Critical node identification plays a vital role in accessing network vulnerability and multiple approaches exist in literature that can be used for identifying critical nodes in a network. Some of the existing algorithms are based on intuition, whereas others are based on mathematical abstractions of networks of arbitrary topology and are thus characterized by properties which can be verified analytically prior to implementation. Most of these approaches either identify critical nodes based on the affect of a node on the traffic flow pattern of the network or they use the topological structure of the network to identify these critical nodes. To the best of our knowledge, no such algorithm exists in literature that identifies critical nodes based on both the topological structure and the traffic flow pattern of the network. To address this problem, this thesis proposes two metrics, the first is based on intuition and it uses a newly defined node diversity metric which incorporates the weighted node degree and the variation in link length capability of a node to address the topological properties of a network. The weighted node degree metric is a slight variant of the well know degree centrality metric [?], the major difference lies in the evaluation of the degree of a node based on the number of new nodes that are introduced by a particular node if it is accessible in the network. The variation in link length metric originates from [?]

which evaluates the affect of Average Path Length (APL) of a network. The variation in link length metric evaluates the diversity of a network by using the difference in path length that a node is maintaining, the intuition behind this approach is that a node that connects multiple nodes at varying distances is highly likely acting as a bridge node among various nodes in the network thus, it is probable that by removing that node the network will report a higher degradation in performance. The traffic flow pattern on the other side is incorporated in this critical node evaluation metric with the use of the Banzhaf power index, it is a slight variant of the well known betweenness centrality metric [?] and was previously used for weighted voting games.

The second metric is based on pure mathematical abstraction where we formulate the critical node identification problem in the form of an optimization problem where the objective is to identify a node that when removed has the highest impact on both, the algebraic connectivity of the network and the maximum traffic flow of the complete network. Here, the first part of the optimization problem deals with the topological properties of the network and the second part deals with the traffic flow of the network, thus addressing both sides of the problem. More detail on both these metrics are explained in Chapter ?? and ?? respectively.

1.3 Research Objectives

The objective of this research is to develop a new model that can correctly identify critical nodes in a network. The identified node, upon its removal, should:

- Increase the average path length of the network, thus increasing the time taken for nodes to communicate with each other.

- Reduce the Algebraic connectivity of the network, which means that the network is loosely connected and the removal of a few nodes will result in network partitioning. These few nodes that are holding the network together are the ones that will create bottleneck for the complete network.
- Increase network congestion and probability of collision thus reducing the network throughput and increasing per packet delay of the network.

Furthermore, the objectives include:

- The design of a distributed algorithm that can correctly identify the most critical node of a network without the need of a centralized monitoring body. This will aid in implementing this algorithm in complex networks such as, Wireless Sensor Network (WSN), Road networks, Communication networks and various other large sized complex network for the assessment of network vulnerability.
- The distributed critical node identification algorithm should be computationally less complex, thus increasing the possibility of its implementation in computationally complex networks.

1.4 Research Method

In order to identify the most critical node in a network, it is essential to identify the right metric that can comprehend the cumulative influence of most of the individual attributes of a node in a network. This thesis uses the node attributes in a network to evaluate various metrics. This selection is based on the consideration that the edge attributes are reflected in the node attributes of the nodes that are connected through that edge. A well known metric that reflects the node attributes of a network is known as the Average Path Length metric. The Average Path Length metric reflects the average time it takes a message to move from one node to any other node in the network. This thesis first emphasises on the existing approaches for estimating the Average Path Length of a network as it is known to be one of the major influencing factor for node criticality and then proposes a new approach that reduces the time complexity of calculating the Average Path Length (APL) of complex networks.

Later, this thesis considers a Wireless Sensor Network (WSN) as a special case example to highlight the affects of changing the Average Path Length of a network and proposes a new methodology for its reduction. The new methodology uses a Variable Rate Adaptive Modulation (VRAM) scheme on top of a Neighbour Avoiding Walk (NAW) mechanism for reducing the Average Path Length using the same transmission power. This helps in building an intuition based metric for the identification of critical nodes in a network. The intuition based critical node identification metric is referred to as the Combined Banzhaf & Diversity Index (CBDI). The Diversity index in CBDI originates from the Neighbour avoiding walk mechanism discussed in the APL reduction mechanism and the Banzhaf Power index in CBDI is a variant of the well known betweenness centrality metric. The CBDI mechanism is tested using simulations and it is shown to perform well in identifying critical nodes in a network.

The intuition based CBDI metric lacks in providing mathematical ground about the way that it works and for this a new critical node identification metric is proposed that is the resultant of the suboptimal solutions of two optimization problems. The critical node identification metric originating from these suboptimal solutions is tested through simulations and analysis. These suboptimal solutions have shown to perform well in identifying the most critical node in the network and they are used to formulate a critical node identification algorithm which is also among the contributions of this thesis.

1.5 Contribution

In this thesis, a new mathematical model is presented that evaluates the Average Path Length of a tree structured network. This is an advancement upon the existing approaches that require tedious computation of all the possible paths in a network for the approximation of the average path length of a network. This contribution is also accompanied by a new approach for the reduction of the average path length of a network which can be used in various network scenarios for the introduction of small world network phenomenon into comparatively large networks.

The contribution of this thesis also incorporates the introduction of a network distributed critical node identification metric which is the outcome of the suboptimal solutions of two well known optimization problems. This thesis also presents a mathematical formulation that identifies the algebraic connectivity of the resultant network after critical nodes are removed from the network. Along with this, a deviation of the degree centrality metric is also proposed which is referred to as the weighted degree centrality metric and is shown through analysis and simulation that it is a better metric than the conventionally used degree centrality metric.

1.6 Thesis Structure

Chapter 1: Introduction This chapter provides information on the context of the research in hand along with the focus of the research work. It also highlights the aims and objectives of the research work.

Chapter 2: Related Work This chapter explains in detail the existing work that relates to the work in this thesis and highlights the deficiencies of the existing work that had initiated the need for this work.

Chapter 3: Average Path Length Calculation For Complex Tree Structures This chapter describes the conventional approaches used for calculating the APL of a complex topology and then proposes a simpler approach, that reduces the computational complexity of calculating the APL of a complex structure.

Chapter 4: Average Path Length and Network Performance This chapter highlights the affects of changing the APL of a network and in it we propose a new mechanism for reducing the APL.

Chapter 5: Intuition Based Critical Node Identification Approach This chapter elaborates on the importance of critical node detection and propose a new intuition based metric for identifying the most critical node in the network.

Chapter 6: Optimization Based Spectral Partitioning for Node Crit-

icality Assessment This chapter discusses the use of optimization theory for the identification of critical nodes in a network and with the aid of this theory we propose a new algorithm that can identify critical nodes in any arbitrary network.

Chapter 7: Conclusion and Future Work This chapter concludes this thesis and presents the future direction of work of this domain.

Chapter 2

Related Work

2.1 Introduction

A number of approaches have been proposed in literature for the identification of the critical nodes in a network. These approaches can be broadly categorized into two categories namely the connection based approach and the topology based approach, where the former uses the information flow pattern of the network to identify the most critical node and the later uses the topological structure of the network to identify the most critical node in the network.

2.2 Connection based schemes

The connection based approaches identify the criticality of a node based on the information flow pattern of a network. The information flow pattern in a network highlights the data flow rate through each individual node and also enables in identifying the node that can be the cause of a potential bottleneck in the network. Both of these parameters play a key role in identifying node criticality and a number of approaches exist in literature that use the information flow pattern of a network to identify the criticality of a node. This section highlights a few well known approaches that are later referred to as the connection based approaches in this thesis.

2.2.1 Average Path Length metric

The average path length metric is among the very widely used metrics and it is also referred to as the characteristic path length metric of a network [**?**]. This metric uses the sum of the shortest path of every node to every other node in the network for the identification of the most critical node in the network. In a graph $G = (V, E)$, where V is the set of vertices and E is the set of edges, the characteristic path length is defined by:

$$l = \frac{1}{N(N-1)} \sum_{v \in V} \sum_{w \neq v \in V} d(v, w) \tag{2.1}$$

where, $d(v, w)$ is the geodesic distance between v and w with $v, w \in V$, i.e., the cumulative distance of all the edges that lie in the shortest path between the two nodes and the factor $1/N(N-1)$ is the one over the total number of pairs of vertices. In such a network, a larger value of l represents a relatively larger time for the message to be disseminated inside a network whereas, a smaller value of l denotes a tightly bonded network where nodes are placed close to each other. The average path length metric identifies such a node as the most critical node which has the highest influence on the average path length of the complete network. It is easily relateable that the node with the shortest path length to every other node in the network, will have the highest influence on average path length of the network where, the average is computed using eq **??**.

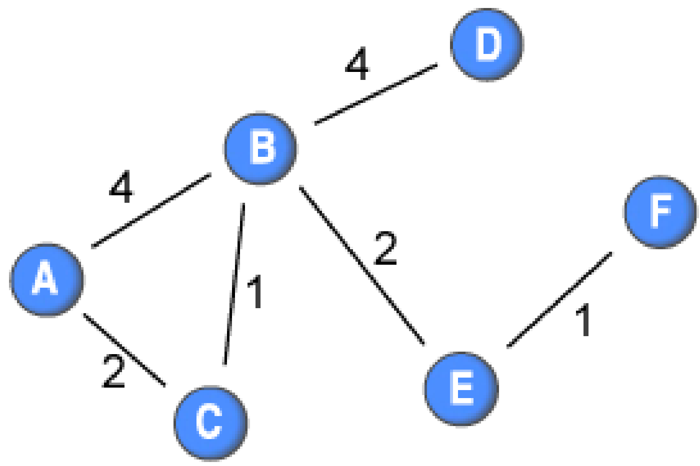

Figure 2.1: Weighted sample network for accessing shortest path length metric.

The sample network in Fig **??** shows an undirected weighted network of six nodes where, the weights represent the length of an edge. The shortest path length metric calculates the distance of every node to every other node in the network and then the node that has the shortest path length among the whole network is referred to as the most critical node in the network. The shortest path lengths for the sample network are shown in the symmetric matrix of table **??**.

It is clear from the table that node B has a shortest distance of 14 units from all the nodes in the network, thus it is assumed to be in the center of the network and the most accessible node in the network. Removing such a node will thus increase the average path length of the network and this makes it the most critical node in

Table 2.1: Path length matrix for Fig ??

	A	B	C	D	E	F
A	-	4	2	7	5	6
B	4	-	1	4	2	3
C	2	1	-	5	3	4
D	7	4	5	-	6	7
E	5	2	3	6	-	1
F	6	3	4	7	1	-

the network according to this average path length metric.

2.2.2 Closeness Centrality metric

Another well known approach that has been in consideration for a long time is the closeness centrality metric [?]. This metric identifies the criticality of a node by analysing the total distance of one node with all the nodes in the network and thus a node that has the lowest total distance and therefore is closer to all the nodes in the network is thus considered as critical in the network. The phenomenon behind the use of this metric is that a node closer to all the other nodes in the network will eventually have the highest network traffic flow through it, as it can reach maximum nodes in the network with the shortest distance. In order to calculate the closeness centrality of a node, researchers use the reciprocal of the total distance from a particular node to all other nodes in a network [?]:

$$CC(v) = \frac{1}{\sum_{u \in V} d(v,u)} \tag{2.2}$$

Unlike the average shortest path metric which is defined as the average distance of the whole network, the closeness centrality metric is a node specific metric, it identifies how close each individual node is to the rest of the network nodes. Fig ?? shows a sample network of 34 nodes where the criticality of the node is represented with both color and size of the node. A bright red node with the biggest size is considered as the most critical node in the network. As expected from the definition of the closeness centrality metric, the nodes close to the center of the network have a higher closeness centrality among all the nodes in the network.

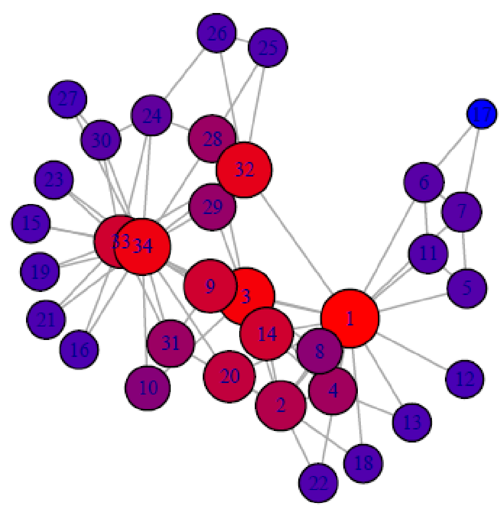

Figure 2.2: Sample network representing nodes with the highest closeness centrality [?].

2.2.3 Betweenness Centrality metric

The betweenness centrality metric was originally proposed by Freeman in his seminal paper [?] and since then it has been used by various researchers for identifying critical nodes in a network [?][?][?]. This is also a shortest path enumeration based metric and it identifies the most critical node based on the number of shortest paths that a node participates in, a node that participates in the highest number of shortest paths will have the highest influence on the performance of the network upon its removal and thus it is considered as the most critical node in the network. Let $\kappa_{uv}(x)$ denote the fraction of shortest paths between node u and v that pass through node x, then [?]:

$$\kappa_{uv}(x) = \frac{\sigma_{uv}(x)}{\sigma_{uv}} \tag{2.3}$$

The betweenness centrality of a vertex x is then defined as [?]:

$$BC(x) = \sum_{u \neq v \neq t \in V} \kappa_{uv}(x) \tag{2.4}$$

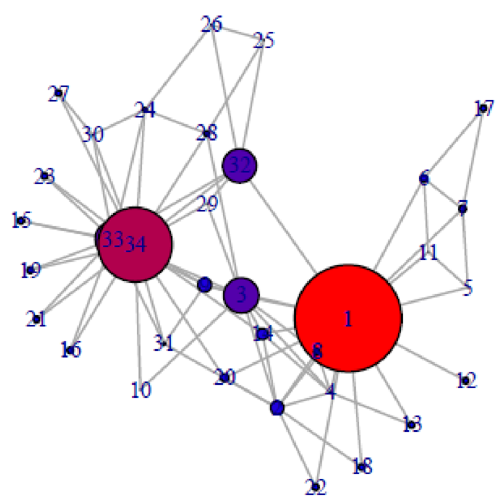

Figure 2.3: Sample network representing nodes with the highest betweenness centrality [?].

The betweenness centrality of a node measures the control of a node on the overall communication in a network, and it is therefore used to identify critical nodes in a network. A higher centrality index indicates that a node lies on a large number of shortest path routes and thus by its removal the network will face a greater decrease in the average network traffic flow rate. Fig **??** shows in red the nodes that participate in maximum shortest path routes in a network of 34 nodes. The size of the node reflects its importance in the network and therefore the node with the largest size and the brightest red color is referred to as the most critical node in the network based on the betweenness centrality metric.

2.2.4 Ego centrality metric

A slight variant of the betweenness centrality is the ego centrality metric [**?**][**?**]. The ego centrality metric was designed for a special class of graphs that are known as the centred graphs [**?**], these graphs are in a star structure and thus restrict nodes from either having a direct link with the neighbour or a path of 2 hops between any two nodes in the network. The ego centrality metric takes benefit of this graph structure and determines the criticality of a node based on the number of times a node participates in forming this two hop path between any two nodes. This definition is in line with the previously defined betweenness centrality metric but the major difference lies in the network structure type. As the ego centrality metric was mainly defined for the star network, thus the maximum length between two nodes of a graph cannot exceed two hop counts [**?**].

2.2.5 Network traffic flow metric

Nasiruzzaman et al. [?] on the other hand believe that it is not necessary that all real life networks use the shortest path routes to relay traffic/messages. Instead they propose a new metric which is build on the phenomenon that the traffic flow pattern is a better estimation metric for the evaluation of critical nodes in a network. Therefore, their proposed approach considers a node to be critical if it observes a higher traffic flow through it. Let F_a be the net maximum power flowing through node a in the network with source node $s \in S$ and load node $l \in L$. Then F_{sal} is defined as:

$$F_{sal} = \sum_{s \in S} \sum_{l \in L} F_a^{sl} \tag{2.5}$$

where $s \neq l \neq a$. Also let, F_n be the net maximum power flowing through the network with the source node $s \in S$ and load node $l \in L$, then F_{sl} is defined as:

$$F_{sl} = \sum_{s \in S} \sum_{l \in L} F_n^{sl} \tag{2.6}$$

The ratio of these two powers could be used to measure the importance of a node and this ratio is called the flow betweenness. This flow betweenness is defined as:

$$C_B(a) = \frac{F_{sal}}{F_{sl}} \tag{2.7}$$

With this approach, a node that has a relatively higher flow betweenness in the network is then considered as the most critical node in the network.

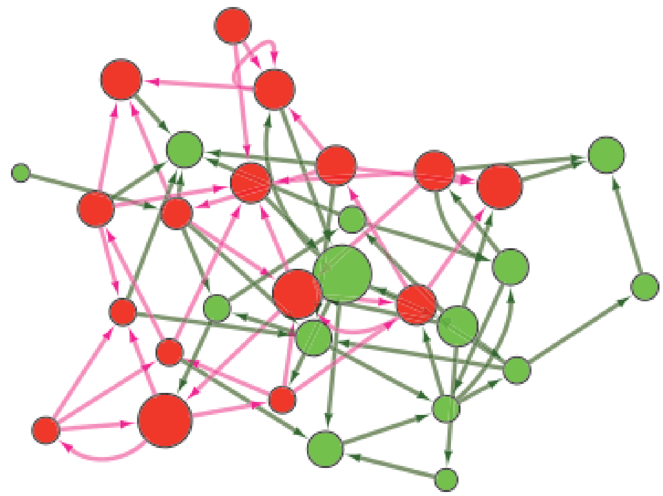

Figure 2.4: Sample network representing nodes with the rank of each node in the network [?].

2.2.6 The Rank matrix

Another approach that exists in literature is the rank matrix approach [?][?]. This matrix uses the traffic that passes through a node to form a $V \times V$ matrix to evaluate the most critical nodes in the network. Unlike previous approaches, this metric identifies a set of critical nodes whereas, the previously stated approaches can identify a single most critical node in the network. In this approach, the minimum number of nodes that report a full rank of $V \times V$ matrix of the network are reported as the most critical nodes. Here the $V \times V$ matrix represents the traffic on the link between nodes in the network. Fig **??** represents a network with multiple nodes where, the size of a node represents the degree whereas, the color represents the criticality. Nodes in red are the ones that have the highest influence on the rank of the network. As per the rank matrix, the node that has the highest degree and has the most influence on the rank of the matrix is considered as the most critical node in the network, this means that the larger red nodes are the most critical nodes in the network.

2.3 Topology based schemes

The topology of a network refers to the arrangement of nodes and their interconnection through edges. These topology based schemes have always been of keen interest to researchers when it comes to networks where there is no relative traffic flow such as, social networks. To tackle this phenomenon of network structure and to understand node criticality based on this structure, various approaches have been proposed in the literature. This subsection highlights a few of these approaches that are highly relevant to the work presented.

2.3.1 Degree Centrality metric

Among all the topology based approaches, the most widely used approach is the degree centrality metric [**?**][**?**]. The degree centrality metric as obvious from the name, uses the degree of a node to identify the most critical node of a network. The node that reports the highest node degree is thus referred to as the most critical node of the network. The key idea here is that, a node with a higher node degree is neighbours to more nodes in the network and thus by removing that particular node a higher number of non-neighbouring nodes will loose connection with each other. Fig **??** shows a graph of 34 nodes with the size and color of a node representing the criticality of a node in a network based on the degree centrality metric. It is evident from the definition of the degree centrality metric that the most critical node will lie in the center of the network and that is also depicted in Fig **??**.

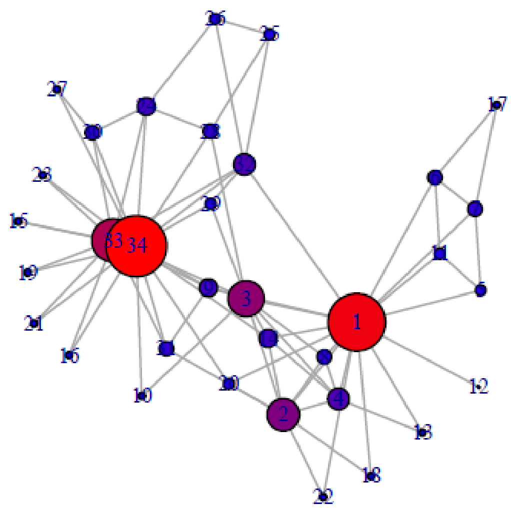

Figure 2.5: Sample network representing nodes with the degree of each node in the network [**?**].

2.3.2 Bonachich metric

Bonachich et al. in [**?**] improvised on the degree centrality by proposing a new power measure and then connecting it with a modified degree centrality measure to obtain a better centrality metric. Bonachich metric is build on the phenomenon that the neighbours of a node play a vital role in determining its importance in a network. A node whose neighbours are connected to less neighbours makes the particular node more powerful as it is likely that the node under consideration is the reason that its neighbours are connected to multiple nodes in the network. Therefore, a node whose neighbours are less connected makes that node more powerful as, by its removal the neighbouring nodes will lose connectivity. On the other hand, if you are connected to more neighbouring nodes then this makes you more central and less powerful thus, the identification of node criticality requires a trade off between node power and centrality. Fig **??** shows a network of 34 nodes with the colors and size of the nodes representing the criticality of a node. It is worth mentioning that the modifications that Bonachich et al. proposed identifies a different node as compared to the one pointed out by the degree centrality metric.

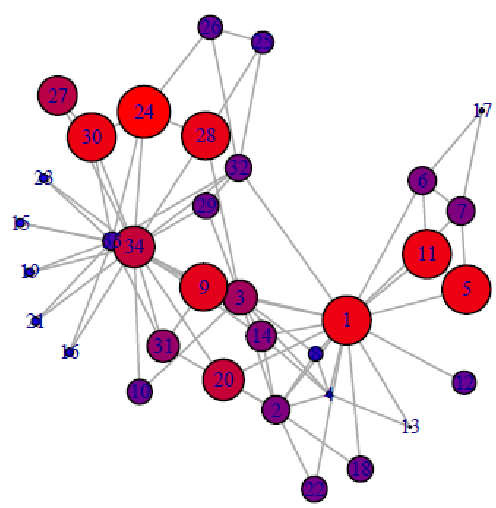

Figure 2.6: Sample network representing nodes with the highest bonachich centrality [**?**].

2.3.3 Eigenvector Centrality metric

Another approach that exists in literature is the eigenvector centrality [**?**][**?**]. Eigenvector centrality is a measure of the influence of a node in a network. It assigns relative scores to all the nodes in a network based on the concept that connections to high scoring nodes contributes more to the score of a node when compared to equal connections of low scoring nodes. For a given graph $G = (V, E)$ the adjacency matrix

$A = (a_{x,t})$ will have $a_{x,t} = 1$ if node x is linked to node t and zero otherwise. The centrality score of node x is defined as:

$$v_x = \frac{1}{\lambda} \sum_{t \in M(x)} v_t \tag{2.8}$$

where, v_t is the eigenvector of the node t that belongs to the neighbour set $M(x)$ of node x with λ being a constant. With a small rearrangement this can be rewritten in a vector notation as the eigenvector equation:

$$Av = \lambda v \tag{2.9}$$

In general, there will be many different eigenvalues λ for which an eigenvector solution exists but considering the additional requirement of all positive eigenvectors only the highest eigenvalue reports the desired result.

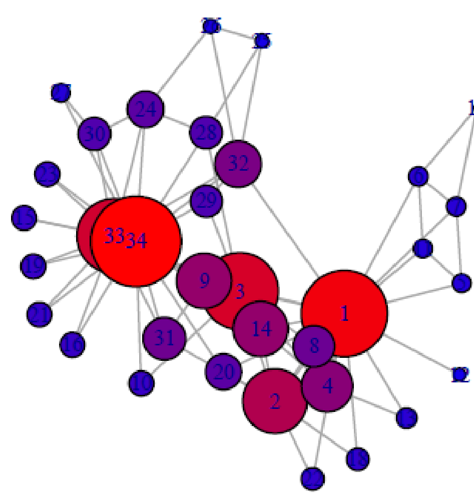

Figure 2.7: Sample network representing nodes with the highest eigencentrality. [?].

Fig **??** reflects the most critical node in the network with aid of size and color. the brightest red coloured node with the biggest size is represented as the most critical node in the network.

2.3.4 The HILPR metric

In the Hybrid Interactive Linear Programming Rounding (HILPR) algorithm [**?**] Yilin et al. propose a different approach of defining node criticality based on the pairwise connectivity of the resultant network after the node removal. They emphasise that the node pair whose removal leads to the most balanced disconnected components and ensures the non-existence of giant components will result in the highest degradation in the performance of the network and thus should be ranked as the most critical node of the network. A similar approach is followed in the GREEDY Critical Node Detection Problem approach (GREEDY-CNDP) [**?**] and the $\beta - disruptor$ approach [**?**], both of which propose an efficient algorithm to minimize pairwise connectivity upon removal of k nodes from the network. Another approach that exists in literature is the algebraic connectivity metric [**?**][**?**][**?**], which is also the focus of this work and is explained in detail in later chapters. The phenomenon here is that the algebraic connectivity is known to be a well defined connectivity metric for a network, therefore to identify a critical node, it is vital to identify the node that reports the highest reduction in the algebraic connectivity of the network. The node that reports the highest reduction in the algebraic connectivity of the network will thus be identified as the most critical node of the network.

Chapter 3

Average Path Length Calculation For Complex Tree Structures

3.1 Introduction

Among all the critical node identification approaches highlighted in this thesis, one of the key approach that exist in literature is the Average Path Length (APL) metric. APL is defined as the mean of the shortest path lengths between all pair of vertices and it represents the closeness and consequently, how quickly information transfer takes place in a network [?] and this makes it as one of the key metric in evaluating the criticality of a node in a network. A node is deemed critical if by its removal the Average Path Length of a network increases and as a resultant information transfer takes a longer time. Most real world networks unexpectedly have short average path lengths, as popularized by six degree of freedom play. This property is known as

Chapter 4

Average Path Length and Network Performance

4.1 Introduction

Average Path Length (APL), node criticality and network reliability are three closely connected features. A network that has a higher APL would have nodes located at a farther distance and thus a few selected nodes will be responsible for relaying most of the data through the network, making these nodes critical. Moreover, the larger the distance between nodes (a larger APL) the lower will be the network reliability [?]. This chapter focuses on elaborating on the connection between the APL of a network and network reliability and also proposes a new approach for reducing the APL of a network which is adopted from the Small-World phenomenon used in Social Networks. In the end this chapter highlights how the APL of a network affects network

Chapter 5

Intuition Based Critical Node Identification Approach

5.1 Introduction

Critical node discovery is an important process for understanding network vulnerability. A node is deemed as critical, if it plays a vital role in maintaining network performance and by removing that node, the overall performance deteriorates and in some cases leads to network partitioning [?] which is highly undesirable. Evaluating the criticality of nodes is significant in various complex networks. In Wireless Sensor Networks (WSNs) employing geographical routing, for example, malicious attack or malfunction of a few beacon nodes leads to fallacious node discovery for the remaining nodes in the network, thus jeopardizing the stable operation of the routing protocol [?]. Moreover, in [?] it was observed that removal of 4% of the nodes in a Peer to Peer

Chapter 6

Optimization Based Spectral Partitioning for Node Criticality Assessment

6.1 Introduction

The identification of critical nodes is vital for accessing network vulnerability and security [?]. The failure of a few critical nodes can have an adversarial effect on network performance varying from slight degradation in the Quality of Service up to the complete breakdown of the network [?]. The significance of critical nodes has been highlighted in a number of examples most of which are explained in chapter ?? & ??. Some of these algorithms are based on intuition, whereas others are based on mathematical abstractions of networks of arbitrary topology and are thus characterized by

Chapter 7

Conclusion and Future Work

This chapter provides the conclusion of the work presented in this thesis along with a brief overview of a future direction of research.

7.1 Conclusion

The change in APL of a network upon removal or addition of a node is among the key considerations when dealing with critical node identification or accessing network vulnerability. This thesis, addressed this critical node identification problem by first identifying the parameters that affect the APL of a network, which is shown in literature to be computed using the complete knowledge of the network, where each nodes computes its distance from every other node in the network. This tedious approach is replaced by a much simpler and computationally less expensive approach in this thesis where, the network is broken down into branch and leaf nodes and then the APL approximated. The proposed approach when tested against existing approaches using extensive simulations was shown to outperform existing approaches in terms of computational complexity.

The APL of a network, as established in this thesis, plays a major role in approximating the average time it takes for a message to be transferred throughout the network and extensive research has been done in order to reduce the APL of a network. Existing approaches achieve this goal by either equipping nodes with special high power antennas that would cover a longer distance or equip nodes with special directed antennas for a directed beam forming or rely on the addition of a dedicated wire, connecting different nodes for increasing the connectivity of the network. This thesis eliminated this requirement of adding special hardware by proposing a new Variable Rate Adaptive Modulation (VRAM) scheme, that changes the modulation schemes to achieve long distance communication. The proposed approach reduces the APL of a network and improves the communication between nodes. It was observed here that, the proposed approach reported an average improvement of 41% in reducing the APL of a network and it reported an average 21% increase in the average node degree of the network when compared to existing approaches. The increase in average node degree is evidence to the fact that more nodes are directly connected with each other and thus message sharing among distant nodes in a network will take a shorter amount of time. This led to the deduction that a critical node will be the one that would increase the APL of a network upon its removal and will therefore increase the time taken for communication between nodes in the network. Numerous approaches exist in literature that work on identifying critical nodes in a network.

A number of approaches that exist in literature mainly deal with the geographic location of nodes or the networks traffic flow pattern to identify these critical nodes of a network whereas, this thesis proposes two approaches, the first being an intuitive approach, that identifies critical nodes of a network (nodes that result in the highest decrease in the performance of the network upon their removal), based on a newly defined diversity index which is combined with an existing Banzhaf power index approach. The newly defined diversity index comprises of the diversity in the link length capability of a node and is referred to as the variation in link length metric and the diversity in weights of the node degree which is referred to as the weighted node degree. The combined affect of the diversity index and the banzhaf power index has been reported in this thesis to outperform existing approaches in identifying critical nodes in a network. The identification of these critical nodes will aid in timely adaptation of the network so that their influence on the performance of the network can be mitigated. The proposed approach reported a 7% and 18% percent increase in average path length of the random and WaxMan network topology respectively took place and a total paths elimination for the small world network topology when the identified critical nodes were removed from the network. Furthermore, it was also reported that a prominent decrease of $13\%, 28\%$ and 68% took place in the average node degree for the random, WaxMan and Small World network topologies respectively which means that the identified/removed node was connected to multiple node in the network thus breaking multiple connections upon its removal. The proposed approach also outperformed existing approaches in terms of increasing the number of isolated nodes in a network. The increase in number of isolated nodes is the evidence to the fact that the network has been partitioned into multiple disconnected components. The proposed approach reported a $150\%, 400\%$ increase in the number of isolated nodes in the random and Small World network topologies respectively and a small increase of 0.8% decrease in the number of isolated nodes in a WaxMan network topology. This small increase is negligible when it comes to really large networks. The proposed approach was also tested against other approaches for analysing the connectivity and the affect it had on the performance of the network and it was observed that the critical nodes identified by the proposed approach when removed from the network result in a decrease in the algebraic connectivity of a network by 58% whereas, as for the performance of the network, the throughput of the network degraded by 22% for the random network topology and it was backed by an increase in the average delay, average jitter and average number of dropped packet by $33\%, 45\%$ and 75% respectively. These all are evidence to the fact that the identified critical node is vital for maintaining normal network functionality and upon its removal the network undergoes major performance degradations.

In order to justify the claims made in the aforementioned work, a mathematical framework was also built which uses suboptimal solutions for two optimization problems, namely the algebraic connectivity minimization and the network utility maximization problem. The resultant solution of these optimization problems when used to identify critical nodes in a network has been shown to outperform existing work in identifying critical nodes in a network. In this thesis, a lower bound on the algebraic connectivity is also calculated that identifies the affect on the algebraic connectivity of the network when a certain node is removed from the network. The critical node identified using this mathematical abstraction resulted in a reduction in the algebraic connectivity of the network by 22% which denotes that the network is loosely connected and the removal of a few more nodes can easily partition the network. It is also reported in this thesis that, the loosely connected network formed after removing the identified critical node results in a bottleneck close to the center of the network which increases the network congestion, reduces network throughput by 16% and increases the average per packet delay, the average number of dropped packets and the average jitter experienced in the network by 6%, 4% and 6% receptively. The proposed approach is complimented in this thesis with a distributed implementation that is computationally less complex and can be implemented in complex networks. It was observed that the proposed approach reduces the average computation time of a network by 36% when compared to existing approaches in the network. These statistics clearly state that the proposed approach outperforms existing approaches in identifying critical nodes in a network and that these nodes when removed result in a higher degradation in performance of the network, thus, in order to maintain normal network functionality, it is necessary to timely identify these critical nodes and take appropriate measures.

7.2 Future Direction of Research

This thesis proposes solutions to two major problems, the first being the estimation of the average path length of a network and the second being the identification of critical nodes in a network. In the first problem, the underlying assumption is of reducing an arbitrary network into a tree structure and it uses a simple tree structure of a single stem with only one set of leaf nodes, the elimination of this assumption will lead to multiple open problems and this can be used as a future direction of work.

The second problem that is being addressed by this thesis initiates its own set of problems that can be addressed in the future. The first being that, as a conventional approach the vulnerability of a network is estimated for a particular instance when the most critical node is removed from the network, but in real life scenarios, most of the networks have a recovery mechanism with the aid of which they adapt and change the network structure to regain maximum network utility. A new direction of research in this domain would be of finding a particular critical node, who's removal will have an impact that cannot be recovered by the network with the aid of a conventional recovery mechanism.

Along with this, another open problem that originates from this thesis is the design of an efficient distributed critical node identification mechanism that is scalable to large sized networks. The proposed approach uses a flooding approach for identifying critical nodes in a network and it is believed that an efficient approach if used would reduce the computation time to a greater extent and thus will be appropriate for use in large sized networks. This thesis also proposes a distributed mechanism for identification of critical nodes but it relies only on the sign of the Fiedler vector elements that correspond to each node in the network. It was observed during this thesis that, the magnitude of the Fiedler values decreases as one moves to a node close to the center of the network (the sign cut region) and it increases as one moves away from this region. As a future direction of research, the use of this change in magnitude along with the change in sign would help in forming a more efficient distributed critical node identification mechanism.

Bibliography

[1] Centrality metrics. *https://cs.hse.ru/data/2015 /05/14/1098547089 /4. Centrality Metrics.pdf*.

[2] Maple homepage. *http://www.maplesoft.com*.

[3] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. Wireless sensor networks: a survey. *Computer networks*, 38(4):393–422, 2002.

[4] Réka Albert, István Albert, and Gary L Nakarado. Structural vulnerability of the north american power grid. *Physical review E*, 69(2):025103, 2004.

[5] K.M. Alzoubi, P.J. Wan, and O. Frieder. Distributed heuristics for connected dominating sets in wireless ad hoc networks. *Journal of Communications and Networks*, 4(1):22–29, 2002.

[6] J.S. Andrade Jr, H.J. Herrmann, R.F.S. Andrade, and L.R. da Silva. Apollonian networks: Simultaneously scale-free, small world, euclidean, space filling, and with matching graphs. *Physical review letters*, 94(1):18702, 2005.

[7] Ashwin Arulselvan, Clayton W Commander, Lily Elefteriadou, and Panos M Pardalos. Detecting critical nodes in sparse graphs. *Computers & Operations Research*, 36(7):2193–2200, 2009.

[8] Ashwin Arulselvan, Clayton W Commander, Oleg Shylo, and Panos M Pardalos. Cardinality-constrained critical node detection problem. In *Performance models and risk management in communications systems*, pages 79–91. Springer, 2011.

[9] W. Asif, M. Lestas, H. K. Qureshi, and M. Rajarajan. Spectral partitioning for node criticality. In *Symposium on Computers and Communication (ISCC)*, pages 877–882. IEEE, 2015.

[10] W. Asif, H.K. Qureshi, M. Rajarajan, and M. Lestas. Cbdi: Combined banzhaf and diversity index for finding critical nodes. In *IEEE Global Communications Conference (GLOBECOM)*, pages 758–763, Dec 2014.

[11] Waqar Asif, Hassaan Khaliq Qureshi, and Muttukrishnan Rajarajan. Variable rate adaptive modulation (vram) for introducing small-world model into wsns. In *47th Annual Conference on Information Sciences and Systems*, pages 1–6. IEEE, 2013.

[12] Giorgio Ausiello, Donatella Firmani, et al. Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, 59(3):275–288, 2012.

[13] David Bader, Kamesh Madduri, et al. Parallel algorithms for evaluating centrality indices in real-world networks. In *International Conference on Parallel Processing, ICPP*, pages 539–550. IEEE, 2006.

[14] Abhik Banerjee, Rachit Agarwal, Vincent Gauthier, Chai Kiat Yeo, Hossam Afifi, and FB Lee. A self-organization framework for wireless ad hoc networks as small worlds. *IEEE Transactions on Vehicular Technology*, 61(6):2659–2673, 2012.

[15] Alexander Bertrand and Marc Moonen. Distributed computation of the fiedler vector with application to topology inference in ad hoc networks. *Signal Processing*, 93(5):1106–1117, 2013.

[16] Vladimir Boginski and Clayton W Commander. Identifying critical nodes in protein–protein interaction networks. *Clustering challenges in biological networks*, pages 153–167, 2009.

[17] Phillip Bonacich. Power and centrality: A family of measures. *American journal of sociology*, pages 1170–1182, 1987.

[18] Phillip Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555–564, 2007.

[19] U. Brandes. A faster algorithm for betweenness centrality*. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[20] Ulrik Brandes. On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2):136–145, 2008.

[21] Pin-Yu Chen and Alfred O Hero. Local fiedler vector centrality for detection of deep and overlapping communities in networks. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1120–1124. IEEE, 2014.

[22] Ed H Chi, Peter LT Pirolli, James E Pitkow, Rich Gossweller, Jock D Mackinlay, and Stuart K Card. Usage based methods of traversing and displaying generalized graph structures, January 21 2003. US Patent 6,509,898.

[23] Paul Christiano, Jonathan A Kelner, Aleksander Madry, Daniel A Spielman, and Shang-Hua Teng. Electrical flows, laplacian systems, and faster approximation of maximum flow in undirected graphs. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 273–282. ACM, 2011.

[24] Richard Church and M Paola Scaparra. Analysis of facility systems reliability when subject to attack or a natural disaster. In *Critical Infrastructure*, pages 221–241. Springer, 2007.

[25] F. Clad, A. Gallais, and P. Mérindol. Energy-efficient data collection in wsn: A sink-oriented dynamic backbone. In *IEEE International Conference on Communications (ICC)*, 2012.

[26] S. Cui, A.J. Goldsmith, and A. Bahai. Energy-efficiency of mimo and cooperative mimo techniques in sensor networks. *Journal on Selected Areas in Communications*, 22(6):1089–1098, 2004.

[27] Nair Maria Maia de Abreu. Old and new results on algebraic connectivity of graphs. *Linear algebra and its applications*, 423(1):53–73, 2007.

[28] Inderjit S Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 269–274. ACM, 2001.

[29] Thang N Dinh, Ying Xuan, My T Thai, Panos M Pardalos, and Taieb Znati. On new approaches of assessing network vulnerability: hardness and approximation. *IEEE/ACM Transactions on Networking*, 20(2):609–619, 2012.

[30] Jürgen Drews. Drug discovery: a historical perspective. *Science*, 287(5460):1960–1964, 2000.

[31] Pradeep Dubey and Lloyd S Shapley. Mathematical properties of the banzhaf power index. *Mathematics of Operations Research*, 4(2):99–131, 1979.

[32] Leonhard Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 8:128–140, 1741.

[33] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[34] Miroslav Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975.

[35] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[36] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.

[37] Linton C Freeman. Centered graphs and the structure of ego networks. *Mathematical Social Sciences*, 3(3):291–304, 1982.

[38] A. Fronczak, P. Fronczak, and J.A. Hołyst. Average path length in random networks. *Physical Review E*, 70(5):056110, 2004.

[39] Arpita Ghosh and Stephen Boyd. Growing well-connected graphs. In *45th Conference on Decision and Control*, pages 6605–6611. IEEE, 2006.

[40] P.J. Giabbanelli. Impact of complex network properties on routing in backbone networks. In *GLOBECOM Workshops (GC Wkshps)*, pages 389–393. IEEE, 2010.

[41] P.J. Giabbanelli, D. Mazauric, and J.C. Bermond. On the average path length of deterministic and stochastics recursive networks. *Complex Networks*, pages 1–12, 2011.

[42] Gene H Golub. Some modified matrix eigenvalue problems. *Siam Review*, 15(2):318–334, 1973.

[43] Willem H Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its applications*, 226:593–616, 1995.

[44] X. Han, X. Cao, E.L. Lloyd, and C.C. Shen. Fault-tolerant relay node placement in heterogeneous wireless sensor networks. *IEEE Transactions on Mobile Computing*, 9(5):643–656, 2010.

[45] Yuan He, Hao Ren, Yunhao Liu, and Baijian Yang. On the reliability of large-scale distributed systems–a topological view. *Computer Networks*, 53(12):2140–2152, 2009.

[46] Petter Holme, Beom Jun Kim, Chang No Yoon, and Seung Kee Han. Attack vulnerability of complex networks. *Physical Review E*, 65(5):056109, 2002.

[47] B.D. Hughes. Random walks and random environments: Volume 2: Random environments (vol 2). 1996.

[48] Muhammad Imran, Mohamed A Alnuem, Mahmoud S Fayed, and Atif Alamri. Localized algorithm for segregation of critical/non-critical nodes in mobile ad hoc and sensor networks. *Procedia Computer Science*, 19:1167–1172, 2013.

[49] A Jamakovic and Piet Van Mieghem. On the robustness of complex networks by using the algebraic connectivity. In *NETWORKING 2008 Ad Hoc and Sen-*

sor Networks, Wireless Networks, Next Generation Internet, pages 183–194. Springer, 2008.

[50] Frank Kelly. Charging and rate control for elastic traffic. *European transactions on Telecommunications*, 8(1):33–37, 1997.

[51] Frank P Kelly, Aman K Maulloo, and David KH Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, pages 237–252, 1998.

[52] Valdis Krebs. Uncloaking terrorist networks. *First Monday*, 7(4), 2002.

[53] Stefan S Krishna, P Krishna Gummadi, and Steven D Gribble. A measurement study of peer-to-peer file sharing systems. *Multimedia Computing and Networking, San Jose, CA, USA*, 2002.

[54] Mukkai S Krishnamoorthy and Narsingh Deo. Node-deletion np-complete problems. *SIAM Journal on Computing*, 8(4):619–625, 1979.

[55] Marios Lestas, Andreas Pitsillides, Petros Ioannou, and George Hadjipollas. A new estimation scheme for the effective number of users in internet congestion control. *IEEE/ACM Transactions on Networking (TON)*, 19(5):1499–1512, 2011.

[56] Yu-Shuai Li, Da-Zhong Ma, Hua-Guang Zhang, and Qiu-Ye Sun. Critical nodes identification of power systems based on controllability of complex networks. *Applied Sciences*, 5(3):622–636, 2015.

[57] Donggang Liu, Peng Ning, and Wenliang Du. Detecting malicious beacon nodes for secure location discovery in wireless sensor networks. In *25th IEEE International Conference on Distributed Computing Systems*, pages 609–619, 2005.

[58] Hao Liu, Xianghui Cao, Jianping He, Peng Cheng, Jiming Chen, and Youxian Sun. Distributed identification of the most critical node for average consensus. *IFAC Proceedings Volumes*, 47(3):1843–1848, 2014.

[59] Hao Liu, Xianghui Cao, Jianping He, Peng Cheng, Chunguang Li, Jiming Chen, and Youxian Sun. Distributed identification of the most critical node for average consensus. *Transactions on Signal Processing*, 63(16):4315–4328, 2015.

[60] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[61] László Lovász and József Pelikán. On the eigenvalues of trees. *Periodica Mathematica Hungarica*, 3(1):175–182, 1973.

[62] MathWorks. Matlab 2013. *http://www.mathworks.com /support /sysreq /sv-r2013a/.*

[63] Consolee Mbarushimana and Alireza Shahrabi. Comparative study of reactive and proactive routing protocols performance in mobile ad hoc networks. In *21st International Conference on Advanced Information Networking and Applications Workshops*, volume 2, pages 679–684. IEEE, 2007.

[64] Russell Merris. Laplacian matrices of graphs: a survey. *Linear algebra and its applications*, 197:143–176, 1994.

[65] S. Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

[66] Bojan Mohar and Y Alavi. The laplacian spectrum of graphs. *Graph theory, combinatorics, and applications*, 2:871–898, 1991.

[67] Ramasuri Narayanam. A game theoretic approach to identify critical components in networked systems. In *SRII Global Conference*, pages 515–524. IEEE, 2012.

[68] ABM Nasiruzzaman and HR Pota. Critical node identification of smart power system using complex network framework based centrality approach. In *North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2011.

[69] Mark EJ Newman. A measure of betweenness centrality based on random walks. *Social networks*, 27(1):39–54, 2005.

[70] NSNAM. Network simulator 3. *https://www.nsnam.org/*.

[71] Daniel P Palomar and Mung Chiang. Alternative distributed algorithms for network utility maximization: Framework and applications. *IEEE Transactions on Automatic Control*, 52(12):2254–2269, 2007.

[72] Charles Pandana and KJ Ray Liu. Robust connectivity-aware energy-efficient routing for wireless sensor networks. *IEEE Transactions on Wireless Communications*, 7(10):3904–3916, 2008.

[73] The Opte Project. Originally from the english wikipedia; licensed under cc by 2.5 via wikimedia commons, 2006.

[74] H.K. Qureshi, S. Rizvi, M. Saleem, S.A. Khayam, V. Rakocevic, and M. Rajarajan. Poly: A reliable and energy efficient topology control protocol for wireless sensor networks. *Computer Communications*, 2011.

[75] O Rojo, R Soto, and H Rojo. Bounds for sums of eigenvalues and applications. *Computers & Mathematics with Applications*, 39(7):1–15, 2000.

[76] S. Schnettler. A structured overview of 50 years of small-world research. *Social Networks*, 31(3):165–178, 2009.

[77] M.Á. Serrano and M. Boguná. Topology of the world trade web. *Physical Review E*, 68(1):015101, 2003.

[78] Gaurav Sharma and Ravi Mazumdar. Hybrid sensor networks: a small world. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pages 366–377. ACM, 2005.

[79] Z. Shelby, C. Pomalaza-Raez, H. Karvonen, and J. Haapola. Energy optimization in multihop wireless embedded and sensor networks. *International journal of wireless information networks*, 12(1):11–21, 2005.

[80] Yilin Shen, Nam P Nguyen, Ying Xuan, and My T Thai. On the discovery of critical links and nodes for assessing network vulnerability. *IEEE/ACM Transactions on Networking (TON)*, 21(3):963–973, 2013.

[81] Yilin Shen, Nam P Nguyen, Ying Xuan, and My T Thai. On the discovery of critical links and nodes for assessing network vulnerability. *IEEE/ACM Transactions on Networking*, 21(3):963–973, 2013.

[82] L. Sitanayah, K.N. Brown, and C.J. Sreenan. Fault-tolerant relay deployment for k node-disjoint paths in wireless sensor networks. In *Wireless Days (WD), IFIP*, pages 1–6. IEEE, 2011.

[83] M.E. Smoot, K. Ono, J. Ruscheinski, P.L. Wang, and T. Ideker. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27(3):431–432, 2011.

[84] M. Soltan, I. Hwang, and M. Pedram. Modulation-aware energy balancing in hierarchical wireless sensor networks. In *3rd International Symposium on Wireless Pervasive Computing,ISWPC*, pages 355–359. IEEE, 2008.

[85] Daniel A Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.

[86] Daniel A Spielman and Shang-Hua Teng. Spectral partitioning works: Planar graphs and finite element meshes. *Linear Algebra and its Applications*, 421(2):284–305, 2007.

[87] E. Stai, V. Karyotis, and S. Papavassiliou. Socially-inspired topology improvements in wireless multi-hop networks. In *International Conference on Communications Workshops (ICC)*, pages 1–6. IEEE, 2010.

[88] S.H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[89] Godfried T Toussaint. The relative neighbourhood graph of a finite planar set. *Pattern recognition*, 12(4):261–268, 1980.

[90] Mario Ventresca and Dionne Aleman. Efficiently identifying critical nodes in large complex networks. *Computational Social Networks*, 2(1):1, 2015.

[91] Chetan Kumar Verma, Bheemarjuna Reddy Tamma, BS Manoj, and Ramesh Rao. A realistic small-world model for wireless mesh networks. *Communications Letters, IEEE*, 15(4):455–457, 2011.

[92] P.J. Wan, K.M. Alzoubi, and O. Frieder. Distributed construction of connected dominating set in wireless ad hoc networks. In *Twenty-First Annual Joint Conference of the Computer and Communications Societies, INFOCOM*, volume 3, pages 1597–1604. IEEE, 2002.

[93] P. Wang and I.F. Akyildiz. Spatial correlation and mobility-aware traffic modeling for wireless sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 19(6):1860–1873, 2011.

[94] Takamitsu Watanabe and Naoki Masuda. Enhancing the spectral gap of networks by node removal. *Physical Review E*, 82(4):046102, 2010.

[95] D. Watts and S. Strogatz. The small world problem. *Collective Dynamics of Small-World Networks*, 393:440–442, 1998.

[96] D.J. Watts. Networks, dynamics, and the small-world phenomenon 1. *American Journal of Sociology*, 105(2):493–527, 1999.

[97] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-worldnetworks. *nature*, 393(6684):440–442, 1998.

[98] Bernard M Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, 1988.

[99] WT Webb and R. Steele. Variable rate qam for mobile radio. *IEEE Transactions on Communications*, 43(7):2223–2230, 1995.

[100] Klaus Wehmuth and Artur Ziviani. Distributed location of the critical nodes to network robustness based on spectral analysis. In *7th Latin American Network Operations and Management Symposium (LANOMS)*, pages 1–8. IEEE, 2011.

[101] Peng Wei and Dengfeng Sun. Weighted algebraic connectivity: An application to airport transportation network. In *Proceedings of the 18th IFAC World Congress, Milan, Italy*, 2011.

[102] J. Wu, M. Cardei, F. Dai, and S. Yang. Extended dominating set and its applications in ad hoc networks using cooperative communication. *IEEE Transactions on Parallel and Distributed Systems*, 17(8):851–864, 2006.

[103] X Xiaochun and W Xiaoyan. A simple rate control algorithm for maximizing total user utility. In *Proceedings of the International Conference on Communication Technology*, pages 135–138, 2003.

[104] Ning Xu, Sumit Rangwala, Krishna Kant Chintalapudi, Deepak Ganesan, Alan Broad, Ramesh Govindan, and Deborah Estrin. A wireless sensor network for structural monitoring. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 13–24. ACM, 2004.

[105] P Yu and H Van de Sompel. Networks of scientific papers. *Science*, 169:510–515, 1965.

[106] Z. Yuanyuan, X. Jia, and H. Yanxiang. Energy efficient distributed connected dominating sets construction in wireless sensor networks. In *Proceedings of the international conference on Wireless communications and mobile computing*, pages 797–802. ACM, 2006.

[107] Z. Zhang, L. Chen, S. Zhou, L. Fang, J. Guan, and T. Zou. Analytical solution of average path length for apollonian networks. *Physical Review E*, 77(1):017102, 2008.

[108] Zhongzhi Zhang, Lili Rong, and Chonghui Guo. A deterministic small-world network created by edge iterations. *Physica A: Statistical Mechanics and its Applications*, 363(2):567–572, 2006.

[109] Z. Zinonos, V. Vassiliou, C. Ioannou, and M. Koutroullos. Dynamic topology control for wsns in critical environments. In *4th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–5. IEEE, 2011.