

Permanent City Research Online URL: http://openaccess.city.ac.uk/15496/

Copyright & reuse
City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/ or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

Versions of research
The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

Enquiries
If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at publications@city.ac.uk.
Combined Banzhaf & Diversity Index (CBDI) for Critical Node Detection

Waqar Asif\textsuperscript{a,b,}\textsuperscript{*}, Hassaan Khaliq Qureshi\textsuperscript{a}, Muttukrishnan Rajarajan\textsuperscript{b}, Marios Lestas\textsuperscript{c}

\textsuperscript{a}National University of Sciences and Technology, H-12 Islamabad, Pakistan.
\textsuperscript{b}School of Engineering and Mathematical Sciences, City University, London, UK.
\textsuperscript{c}Department of Electrical Engineering, Frederick University, Nicosia, Cyprus.

Abstract

Critical node discovery plays a vital role in assessing the vulnerability of a computer network to malicious attacks and failures and provides a useful tool with which one can greatly improve network security and reliability. In this paper, we propose a new metric to characterize the criticality of a node in an arbitrary computer network which we refer to as the Combined Banzhaf & Diversity Index (CBDI). The metric utilizes a diversity index which is based on 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 shortest paths. The Banzhaf power index is inspired from the theory of voting games in game theory. The proposed metric is evaluated using analysis and simulations. The criticality of nodes in a network is assessed based on the degradation in network performance achieved when these nodes are removed. We use several performance metrics to evaluate network performance including the algebraic connectivity which is a spectral metric characterizing the connectivity robustness of the network. Extensive simulations in a number of network topologies indicate that the proposed CBDI index chooses more critical nodes which, when removed, degrade network performance to a greater extent than if critical nodes based on other criticality metrics were removed.

Keywords: Node criticality, Network vulnerability, Weighted node degree, Banzhaf power index, Algebraic connectivity.

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 \cite{1} 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 \cite{2}. Moreover, in \cite{3} it was observed that removal of 4\% of the nodes in a Peer to Peer Gnutella Network resulted in major fragmentation of the whole network. The node criticality problem in Peer to Peer and overlay networks was also addressed in \cite{4}. Finally, in \cite{5} it was shown that in a telecommunication network, the penetration of a virus can be prevented by removing a few critical nodes. Node criticality problem is also significant in network paradigms beyond computer networks. In road networks, for example, intersections which can be considered as nodes in a graph theoretic framework, might experience heavy traffic loads when in proximity to a major landmark. Identifying such critical nodes is significant when investigating possible extensions of the existing infrastructure \cite{6}. Likewise, in a social network of terrorist activists, the removal of a few critical nodes can paralyse the communication in the network, making the network ineffective \cite{7}.

Several studies have addressed the node criticality problem and various metrics have been proposed to characterize the criticality of nodes in a network. Among these metrics, the degree centrality metric \cite{8} is one of the most commonly used. In a simple undirected network, the degree centrality of a node is calculated as the number of its adjacent neighbours, whereas for a directed network, the metric, based on the direction of flow, is divided into the in-degree and out-degree centrality. The higher the degree, the more critical the node is assumed to be. Despite its simplicity, this metric does not take into account the geometrical characteristics of the network, which are known to highly affect performance and this has led to the consideration of the closeness centrality metric. Closeness Centrality \cite{8} utilizes the average geodesic distance between all nodes in the network. The node that has the highest closeness centrality value is the one which is placed in the geographical center of the network and it thus has

\textsuperscript{*}Corresponding author

\textit{Email addresses:} 09mscsewasif@seecs.edu.pk (Waqar Asif ), hassaan.khaliq@seecs.edu.pk (Hassaan Khaliq Qureshi), r.muttukrishnan@city.ac.uk (Muttukrishnan Rajarajan), eng.lm@frederick.ac.cy (Marios Lestas)
the shortest distance to all its neighbouring nodes. A distributed algorithm to find nodes with the highest closeness centrality value is presented in [9]. The global Clustering Coefficient metric [10] uses similar ideas to weigh each node's degree of participation in cluster formation thus characterizing its criticality.

The node criticality problem has also been viewed as an algebraic connectivity minimization problem, where the most critical nodes are the ones which minimize the algebraic connectivity of the network [11]. Since the solution of the optimization problem becomes computationally expensive to find as the size of the network increases, a number of suboptimal solutions have been proposed in literature [12],[13],[14]. Another set of approaches that exist in literature are based on the ability of nodes to fragment the network when removed. In [15], Neng et al. formulate two optimization models, namely the graph partitioning problem (GPP) and the critical node problem (CNP). They use GPP to identify nodes which, when removed result in the highest decrease in the sum of weights of the edges between disjoint sets and CNP to identify a set of nodes which result in the highest reduction in the pairwise connectivity of a network upon their removal. The proposed approaches have been shown to perform well in identifying critical nodes, however as the authors point out, the computational complexity of the proposed approaches increase significantly with the increase in the network size.

To address this problem, Thang et al. in [16] formulate two alternative optimization problems which use the pairwise connectivity measure of a network to identify a set of critical nodes or edges, which if removed result in the highest degradation in the networks pairwise connectivity. Moreover, [17],[18] and [19] use network partitioning concepts to assess the vulnerability of a network based on the size of the largest connected components after cascading failures occur. It has been shown that these approaches perform well in abstract models of interdependent networks which assume random interdependency between nodes. Finally, in [1], it is conjectured that partitioning of a network into two equal segments leads to the highest degradation in network performance thus motivating the consideration of the pairwise connectivity. The relevant critical node and link disruptor optimization problems are considered and the NP hardness of these problems is addressed by a heuristic method to which they refer to as HILPR.

The aforementioned metrics are based on topological properties of the network, which assess the criticality of a node without taking into consideration the flow paths of the active connections. The latter is accounted for in the Betweenness Centrality metric. Betweenness Centrality [20], determines the criticality of a node by estimating the contribution of each node in forming a shortest path route. A node that participates in maximum shortest path routes is considered as a highly critical node. The participation of a node in path formation is also accounted for in [21] where, a node is considered as critical, when it achieves the highest decrease in the rank of the routing matrix upon its removal from the network. The flow induced by the active connections is considered by Xuan et al. in [22] where, taking into account the traffic shockwave model which was earlier proposed by Wang Dianhai in [23], they identify as critical the nodes which when removed, result in the highest increase in average network congestion. A similar approach was also used by Yew-Yih et al. in [24] where the delayflow of the network is used as the performance metric with which the criticality is assessed.

Furthermore, in [25],[26], node criticality is assessed based on the resulting efficiency of the network after nodes are iteratively removed. The node that reports the highest reduction in efficiency upon its removal is referred to as the most critical. This approach suffers from the high computational complexity associated with the iterative procedure utilized to detect critical nodes. The problem is exacerbated by the fact that multiple node removal may also lead to maximum efficiency decrease. This problem is addressed in [27] where criticality is assessed not only based on the node removal but also on the removal of the associated paths.

In this work, based on our preliminary results in [28], we propose a criticality metric which is shown to be more successful in identifying nodes, the removal of which, significantly affects network operation. The metric encompasses three main node attributes: the weighted node degree, the variation in link length of the node from its neighbours and its contribution in forming shortest paths. Unlike previous proposals, which take into account the absolute node degree, in this proposal we consider the node degree weighted by the average common neighbours of the node with all its neighbours. The presence of common neighbours is an indication of the presence of path alternatives which undermine the criticality of a node. In addition, in order to account for long range links which cause nodes to act as relay nodes thus accommodating heavy traffic and becoming critical for the whole network operation, we introduce the notion of the variation in link length between neighbouring nodes. The diversity in the number of neighbours and the diversity in link lengths thus contribute to the criticality of a node and are used to form the diversity index. We then account for the contribution of each node in forming the routing paths by employing a new technique which is inspired by voting games in game theory. The metric emanating from this technique is known as the Banzhaf Power index. The combination of the latter with the diversity index yields the proposed criticality metric which we refer to as the Combined Banzhaf & Diversity Index (CBDI).

We evaluate the performance of the proposed metric using analysis and simulations. The evaluation is based on the degradation in performance reported when nodes selected using the criticality metric under consideration are removed from the network. We compare the proposed metric against other metrics that have been proposed in the literature, namely the Hybrid Interactive Linear Programming Rounding (HILPR) proposed in [1], the Control-
lability of complex networks (Cont) in [21] and the Degree Centrality, Betweenness Centrality, Closeness Centrality used in [8]. The Random Network Topology, the WaxMan Network Topology and the Small World Network Topology were considered in the simulation experiments and network performance was evaluated using a number of performance metrics which include the average node degree, the average path length, the number of isolated nodes, the network throughput, the average per packet delay, the average per packet jitter, the number of dropped packets and the algebraic connectivity. The latter, defined as the second smallest eigenvalue of the Laplacian of a network, serves as a connectivity robustness metric. It provides an analytical perspective as to why the proposed metric and its key features work effectively. Extensive simulations indicate that the proposed criticality metric in the considered scenarios is able to achieve a more severe degradation in network performance compared to other approaches, indicating that it is superior in characterizing the criticality of the network nodes.

The rest of the paper is organized as follows: in Section II we describe the proposed criticality metric, in Section III we elaborate on the algebraic connectivity of a network, in Section IV we evaluate its performance using simulations and finally in Section V we offer our conclusion of the paper.

2. Proposed Criticality Metric

As mentioned in the introduction, in this work, we propose a new criticality metric which is the combination of the Banzhaf power index and the diversity index. In this section, we explain the reasoning behind our design choices and formally define the diversity index and the Banzhaf power index. We then show how we combine the two to form the proposed criticality index.

2.1. Diversity Index

Diversity index is a measure of the variation of node properties between neighbouring nodes. We consider variation of two attributes of neighbouring nodes which are logically related to their criticality: the variability in link lengths and the variability in their list of neighbours. Increasing both the variability of link lengths and the variability in the list of neighbours implies greater node criticality. Below we give a detailed description of the two and explain how they are combined to form the diversity index.

2.1.1. Variation in link length

This attribute measures the variation in the length of the links between neighbouring nodes. A greater variation in link length certifies the existence of both long distance and short distance links. A node with the aforementioned property is capable of acting as a relay node between the nodes in proximity and the distant ones. This will aid neighbouring nodes in getting their data relayed to distant nodes and vice versa at a reduced network energy and time cost [29]. Since a node with a higher variation in link length has a higher probability of acting as a relay node hence, it is deemed as critical for information dissemination.

We define the variation of link length as the average difference between the transmission radii of neighbouring nodes. We assume a graph $G = (V,E)$, where $V$ represents the set of Nodes and $E$ represents the set of Edges. Each node $x$ in $V$ is characterized by its transmission radius $T_x$. For each node $x$, the set of nodes which lie within the transmission range of $x$ is the set of its neighbours and is denoted by $N(x)$. The variation in link length of $x$ is denoted by $D_d(x)$ and is given by:

$$D_d(x) = \frac{1}{|N(x)|} \sum_{u \in N(x)} (T_x - T_u)$$

Figure 1: Node $N$ acts as a relay node between the two network partitions and thus has a higher variation in link length value compared to node $A$.

In order to demonstrate the way that the variation in link length characterizes the criticality of a node, we consider the example network of Fig 1. The links between nodes are drawn to scale so that longer link lengths on the diagram, indicate longer link lengths in the actual network. Nodes $A$ and $N$ in the considered network share the same node degree. However, node $N$ reports a larger value of the variation in link length metric, as it has both short and long length links. Node $A$ on the other hand, only has short length links resulting in a low variation in link length value. The removal of node $A$ partitions the network, however, it only isolates nodes $C$ and $M$. The removal of node $N$, on the other hand, partitions clusters 1 and 2 thus resulting in isolation of a far larger number of nodes. This demonstrates the higher criticality of node $N$ which is reflected in a higher value of variation in link length.

2.1.2. Weighted Node Degree

Node degree was used by Freeman in [8] for determining the criticality of a node. Despite the simplicity of the method it fails to take into consideration self loops and one hop reachability of neighbouring nodes which leads to overestimates of the node criticality. Therefore, in this
work, we avoid the consideration of these redundant paths by elaborating on the variability of the list of neighbours of neighbouring nodes, leading to the notion of weighted node degree. The weighted node degree takes values between 0 and 1, and increases as the number of common neighbours decreases. A greater number of common neighbours implies more one hop paths between neighbouring nodes which undermines the criticality of a node. The weighted node degree of a node is calculated as the product of the two metrics such that:

$$D_n(x) = \sum_{u \in N(x)} \frac{|N(u) \setminus N(x)|}{|N(u)|}$$

where \(\setminus\) denotes the set difference and \(|\cdot|\) denotes the cardinality of the set. So, the weighted node degree of a node \(x\) is calculated by summing the dissimilarity ratios of all of its neighbours. The dissimilarity ratio for a particular neighbour \(u\) is the ratio of number of neighbours of \(u\) which are not neighbours of \(x\) over the set of all neighbours of \(u\).

In order to highlight the methodology with which the weighted node degree determines the criticality of a node we use the example network of Fig 2. In this network, nodes \(A\) and \(D\) share the same node degree but a different weighted node degree. The weighted node degree of node \(A\) is lower than that of node \(D\) due to the link between nodes \(B\) and \(C\) which causes the neighbours of \(A\) to have one common link. This extra link adds to the redundancy of connections of node \(A\) and thus when node \(A\) is removed from the network, the resultant network is still connected through the link between nodes \(B\) and \(C\). On the other hand, the removal of node \(D\) completely isolates node \(E\) and hence partitions the network in two segments. This demonstrates that a node with a higher weighted node degree has a higher influence on the network upon its removal and is thus a more critical node.

Both the variation in link length and the weighted node degree of a node described above are used to calculate the diversity index of that node. The diversity index \(H(x)\) is defined as the product of the two metrics such that:

$$H(x) = D_a(x)D_n(x)$$

It follows from the discussion above that the greater the diversity index, the more critical a node is. The criticality of a node is further refined by weighing its participation in path formation. To this end, we use the Banzhaf power index which is described below.

2.2. Banzhaf power index

In game theory, different assumptions have led to different definitions for determining the importance of an agent in a game. One of the most prominent among these is the Banzhaf power index [30]. This index has been widely used primarily for the purpose of weighted voting games. In a voting game, each voter is assigned a weight and the coalition of these voters determines the outcome of the game. A game is considered as a winning game, if the sum of all the weights of the nodes in a coalition is greater than or equal to a predefined threshold weight. A node has a pivotal role if, its removal transforms a winning game into a loosing game. Nodes with the aforementioned property are called swing nodes. A node that acts as a swing node in maximum coalitions is the most critical node and is assigned the highest Banzhaf power index.

We adapt the above ideas in a communication network setting in order to characterize the criticality of nodes participating in the network. In the same way that weights are being used to select coalitions in a voting game setting, we use the link bandwidths in a communication network setting to select the nodes participating in shortest path formation. A coalition of nodes is considered as a winning coalition, if the path they form satisfies the bandwidth requirements of a particular source destination pair. We thus disregard links which cannot support these bandwidth requirements. Once a shortest path has been established, a node is called a swing node if it participates in the shortest path. The removal of a node that participates in maximum shortest path routes, will have a higher impact on network performance and is thus considered a critical node in the network. So, in analogy to the voting games setting, a node which acts as a swing node in maximum coalitions is the most critical node and is assigned the highest Banzhaf power index formally defined below.

In the graph \(G = (V, E)\), \(I\) denotes the set of all source destination pairs \(w = (i, j), i, j \in V\). For each \(w \in I\), \(L(w)\) contains the set of nodes which constitute the shortest path route that fulfills the bandwidth requirements. A node \(k\) that belongs in \(L(w)\) acts as a swing node for the source destination pair \(w\). The Banzhaf power index for a node is the ratio between, the number of times a node acts as a swing node, over the total number of times all the nodes in \(V\) act as swing nodes. The Banzhaf power index is denoted by \(C_k\) and is given by:

$$C_k = \frac{\sum_{w \in I} |L(w)| - |L(w)\setminus k|}{\sum_{p \in V} \sum_{w \in I} (|L(w)| - |L(w)\setminus p|)}$$

2.3. Combined Banzhaf & Diversity Index (CBDI)

The proposed criticality metric is obtained by multiplying the diversity index and the Banzhaf Power Index as shown below:

$$CBDI(x) = C_k H(x)$$
The combination method used is a design parameter and we support our selection using simulations in section III. The metric is referred to as Combined Banzhaf & Diversity Index (CBDI) and refines the mechanism of critical node detection. According to this index, a node is critical not only if it participates in maximum shortest path routes but, if it is also prominent among its neighbours due to a higher variation in node attributes. The index, unlike previous approaches, is able to refine nodes which participate in the same number of shortest paths by differentiating between nodes which relay information from multiple inputs to multiple outputs and nodes which relay information from a single input to a single output. Further, it can identify nodes which can relay data to distant nodes thus having a high probability of experiencing heavy traffic. Finally, it is able to refine the information obtained by the node degree by excluding neighbouring nodes whose participation in path formation is not critical.

3. Algebraic Connectivity of a Network

Algebraic connectivity, also referred to as the Fiedler value, is a spectral metric defined as the second smallest eigenvalue of the Laplacian matrix of a network. Its significance stems from a theorem by Fiedler [31] which states that a network is disconnected, if and only if, the algebraic connectivity attains a value of zero. It has thus been conjectured that the algebraic connectivity can be used as a connectivity or robustness measure of the network in the sense, that the higher its value is, the more difficult it is to partition the network. Such a conjecture is supported by a number of theorems which offer insights towards this direction. The algebraic connectivity is related to the criticality of a node as it provides an analytical metric with which one can assess the degradation in network connectivity when the node is removed. In this section, we review key definitions and theorems pertinent to the algebraic connectivity concept which can provide insights in how key features of the proposed criticality metric identify nodes the removal of which lead to the network becoming more disconnected.

Let $G = (V,E)$ be a graph of $|V| = n$ nodes and $|E| = m$ edges. \( G \) is undirected then, \( A(G) = (a_{ij}) \) is the adjacency matrix of \( G \) with \( a_{ij} = 1 \) if nodes \( i \) and \( j \) share an edge \( z \in E \) and \( a_{ij} = 0 \) otherwise, for \( i,j \in V \). The diagonal degree matrix \( \delta(G) = diag(deg_1,deg_{i+1},...deg_n) \) is an \( n \times n \) matrix with the diagonal entry \( deg_i \) representing the degree of the node \( i \in V \) and all non-diagonal entries equal to zero. The Laplacian matrix for such an undirected graph \( G \) is an \( n \times n \) matrix, \( L(G) = \delta(G) - A(G) \). In case of a directed graph, the Laplacian matrix is represented by \( L(G) = N(G)N(G)^T \) where, \( N(G) \) denotes the incidence matrix [11]. The incidence matrix \( N(G) \) is an \( n \times n \) matrix with \( n_{ij} = 1 \) if an edge is directed from node \( i \) to \( j \), \( n_{ij} = -1 \) if the edge is directed from node \( j \) to \( i \) and \( n_{ij} = 0 \) otherwise. The Laplacian matrix \( L(G) \) of a graph is real, symmetric and non-negative semi-definite with all its eigenvalues being real and non-negative [32].

These eigenvalues are highly correlated with the connectivity of a graph and this relation is further elaborated in the following lemma [33].

**Lemma 1:** If \( 0 = \lambda_0(G) \leq \lambda_1(G) \leq ... \leq \lambda_{n-1}(G) \) are the eigenvalues of the Laplacian matrix \( L \) in an ascending order, then \( \lambda_1(G) > 0 \) if \( G \) is connected. Additionally, if \( \lambda_i(G) = 0 \) and \( \lambda_{i+1}(G) \neq 0 \), then \( G \) has exactly \( i + 1 \) disjoint connected components.

The zero row and column sum of the Laplacian matrix generates an eigenvalue of zero which is considered as the smallest eigenvalue \( \lambda_0 \) of the matrix. The aforementioned lemma indicates that if a graph is connected then the eigenvalue of zero will have a multiplicity of one whereas, if the eigenvalue of zero has a multiplicity of \( j \) then there are \( j \) disconnected components of the graph. Similar to the smallest eigenvalue, the largest eigenvalue \( \lambda_{n-1}(G) \) also has a multiplicity of 1 if the graph is connected [34]. The largest eigenvalue is upper bounded by the maximum degree \( D_{\text{max}} \) and lower bounded by \( \max(D(G),\sqrt{D_{\text{max}}(G)}) \) [34].

Apart from the smallest and the largest eigenvalues, the second smallest eigenvalue \( \lambda_1 \), which is also referred to as the Fiedler value, is of vital importance for determining the connectivity of a graph [31]. A higher order Fiedler value, which is strictly larger than zero, shows a connected graph whereas, a smaller value shows a weakly connected graph. It is lower bounded by the smallest eigenvalue of zero and upper bounded by the minimal nodal degree of the network. The minimal nodal degree defines the minimum number of links that if broken could possibly result in another disconnected component and hence, the bounds on the Fiedler value can be expressed as [32]:

\[
0 \leq \lambda_1(G) \leq \frac{n}{n-1}D_{\text{min}}(G)
\]

Here, \( D_{\text{min}} \) is the minimal nodal degree of an incomplete graph. The above inequality indicates that the algebraic connectivity can be used as a connectivity robustness measure. The smaller the \( D_{\text{min}} \) value, the easier it is for the network to become disconnected as fewer node removals are required to lead to network partitioning. As \( D_{\text{min}} \) decreases, so does the upper bound on the algebraic connectivity and one may thus conjecture that the easier it is for the network to become disconnected the more likely it is for the algebraic connectivity to attain a small value. Reversing the argument, one may conjecture that the smaller the algebraic connectivity value, the easier it is for the network to become disconnected. The use of the algebraic connectivity as a connectivity robustness measure can be further supported by the following lemma [32][31]:

**Lemma 2:** If there are two edge disjoint graphs with the same number of nodes \( G_a \) and \( G_b \), then \( \lambda_1(G_a) + \lambda_1(G_b) \leq \lambda_1(G_a \cup G_b) \).

**Corollary 1:** Likewise, if there are two graphs with the same number of nodes but different set of edges, such that \( G_a(V,E_a) \) and \( G_b(V,E_b) \) for \( E_a \subset E_b \) then the Fiedler value is non-decreasing and can be represented as, \( \lambda_1(G_a) \leq \lambda_1(G_b) \).
\( \lambda_1(G_k) \).

Corollary 1 suggests that the removal of edges from a network, which makes it easier for the network to become disconnected, leads to a decrease in the algebraic connectivity. Again reversing the argument one can conjecture that the smaller the algebraic connectivity value is, the easier it is for the network to become disconnected. The effect of removing edges from the network on the algebraic connectivity is captured by Corollary 1. Removal of nodes is also of primal importance as the criticality of a node is assessed by its impact on the network performance when it is removed. The following Lemma describes the effect of node removal on the algebraic connectivity of a network.

**Lemma 3:** If \( G_1 \) is the resultant graph after removal of \( k \) vertices along with all the adjacent edges, then:

\[
\lambda_1(G_1) \geq \lambda_1(G) - k
\]

The Lemma suggests that the removal of nodes decreases the lower bound on the algebraic connectivity. This means that by appropriate choice of the nodes, one can decrease the algebraic connectivity thus making it easier for the network to become disconnected. The above properties of the algebraic connectivity are now used to explain how the lower bound on the algebraic connectivity of the network decreases the algebraic connectivity thus making it easier for the network to become disconnected. The above indicates that when nodes with the same node degree but higher weighted node degree are removed then the algebraic connectivity of the network decreases. The weighted node degree can thus be used to refine the node degree concept and identify more critical nodes.

4. **Performance Evaluation**

In this section, we evaluate the performance of the proposed criticality index using simulations conducted on Matlab [35] and the Network Simulator (NS-3) [36]. We conduct a comparative study to investigate the performance of the proposed index against other approaches that have appeared in the literature: the Hybrid Interactive Linear Programming Rounding (HILPR) algorithm proposed in [1], the algorithm in [21] (Cont) which attempts to reduce the rank of the routing matrix and the node centrality metrics such as the betweenness centrality, closeness centrality and degree centrality metrics that are used in [20]. Among all criticality indices proposed in literature we have chosen the above as they contain some of the features included in our approach, namely the diversity, the node degree and the participation in shortest paths. In addition, they have been shown to outperform the other proposals in a number of scenarios. In each conducted simulation experiment, nodes participating in the network are assigned a criticality measure based on the criticality index under consideration. A fixed percentage of the most critical nodes are removed and the degradation in network performance is evaluated. The most effective criticality index is the one that leads to a greater degradation in performance.

In the first set of simulation experiments conducted on Matlab network performance is evaluated in terms of topological performance metrics, in the second set of simulation experiments network performance is evaluated in terms of the algebraic connectivity of the network and in the third set of simulation experiments conducted on the NS-3 simulator, network performance is evaluated in terms of the network centric performance metrics.

4.1. **Topology Based Evaluation**

In the first set of simulation experiments conducted on Matlab, we evaluate the ability of the proposed metric to choose critical nodes in terms of topology based performance metrics such as the Average Node Degree, the Average Path Length and the Number of Isolated Nodes. The Average Node Degree is the average number of neighbours of all nodes participating in the network. Small average node degree values imply smaller connectivity, therefore, the smaller the average node degree, the greater is the degradation in network performance. The Average Path length is obtained by calculating the average of all path lengths over all source destination paths in the network. High average path length in a network implies lack of critical nodes which can participate in shortest path routes. So, the higher the average path length, the greater is the degradation in network performance. Finally, the Number of Isolated Nodes are the nodes that have no connections with any other node in the network. High number of isolated nodes is undesirable as it implies greater network partitioning.

The evaluation was conducted using three different network topologies in an area of 1000 x 1000m². The Random Network Topology assumes \( x \) and \( y \) coordinates of the nodes which are uniformly distributed in the area under consideration. The number of nodes were chosen in the range of 10 - 80 and among them 90% of the nodes were assumed to have a constant transmission range equal to 300m whereas, some randomly selected 10% of nodes...
were assigned a transmission range of 450m in order to enable long distance links [37]. In the WaxMan Network Model [38], the probability that a connection is established between any two randomly distributed nodes \( u, v \) in the network \( P(u, v) \) depends on the distance \( d \) between the nodes as shown below:

\[
P(u, v) = \alpha e^{-d/bL}
\]  

(8)

where \( 0 < \alpha < 1 \) and \( b \leq 1 \) are constants and \( L \) is the maximum distance between any two nodes. As \( \alpha \) increases, the probability of having edges between two nodes increases, whereas, with the increase in \( b \), the ratio of long distance to short distance edges increases. In our simulations, we fix, the total number of nodes to 80 and consider a constant value of \( b = 0.5 \). In order to analyze the effect of node density on the performance of the network, we vary the value of \( \alpha \) from 20 – 80%. Finally, in the Small World Network Model [39], \( N \) nodes form a one-dimensional lattice with each node placed uniformly on the boundary of a circle. Each node in the network forms a direct connection with its \( k^{th} \) nearest neighbours, where \( k \) is a constant and it represents the edge connectivity of the network. In this network topology, a network size varying from 20 – 80 was considered, with a fixed edge connectivity of \( k = 2 \). In addition, 10% of the edges are randomly re-wired to introduce the long range links in the network. These long range links reduce the average path length between the nodes.

In each of these topologies, the criticality metric was evaluated by removing the selected critical nodes from the network and then measuring the network performance. In order to reduce the variance of the obtained results, each simulation experiment was repeated 50 times and the values presented, are averages over all obtained outputs. We assume a fluid flow model of the network and the bandwidth of each node is randomly selected according to a uniform distribution with a maximum value of 2Gbits/sec. Information sources are assumed to be non-responsive and their data rate is chosen from a uniform distribution in the range 0-2Gbits/sec. In each experiment, the performance of the reference network (we refer to it as the original network) is evaluated and then compared with the performance of the network when 20% of the total nodes are removed. The nodes which are removed are the ones which have been assigned the highest criticality value according to the criticality index under investigation.

In Fig 3 for each network topology we show the average node degree values obtained in the original network and compare it with the values obtained when the most critical nodes are removed using the three criticality metrics under investigation. For the Random Network Topology, and the Small World Topology, the average node degree is plotted against the number of nodes within the network. In the WaxMan Topology, the average node degree is plotted against the parameter \( \alpha \) of the model which is a measure of the edge density within the network. The greater the value of \( \alpha \), the greater is the edge density and thus the number of edges. We observe that in all cases, the proposed CBDI criticality metric achieves a larger reduction in the average node degree, a strong indication of a greater degradation in network performance. This implies that the nodes removed using the proposed CBDI metric are more critical. The highest impact of our approach compared to the others is observed in the Small World Topology whereas the smallest impact is reported in the Random Network Topology. It is worth noting that in the Random Network Topology as the number of nodes increases, so does the average node degree. This is expected due to the increase in node density. A similar pattern is observed in the WaxMan Network Topology, however, the increase rate is smaller. For the Small World topology, the average node degree is fairly constant with increasing number of nodes due to the nature of the model which assumes a constant value for the average node degree equal to 2.

In Fig 4, for each considered network topology, we show the Average Path Length reported in the original network and the network resulting from the removal of the critical nodes. The critical nodes are chosen using the proposed criticality metric and the other two metrics under consideration. Higher Average Path Length values are desirable, when removing critical nodes, as they imply the removal of nodes which participate in shortest paths. We observe that the proposed metric, is able to slightly increase the average path length in the WaxMan and Random Network Topologies, at high \( \alpha \) and number of node values respectively. This is expected due to a higher variability in node attributes when increasing the node density. In the Small World Network almost zero path length values are reported by the CBDI metric due to the large number of isolated
nodes that it creates. This is highlighted below.

Finally in Fig 5 we show the number of isolated nodes reported in each of the network topologies under consideration. The number of isolated nodes is shown for increasing values of the number of nodes and $\alpha$ in the original network and when the critical nodes have been removed using the considered criticality metrics. The results demonstrate the superiority of the proposed metric, especially in the case of the Random Network Topology and the Small World topology. In all three topologies, the removal of critical nodes using the proposed CBDI criticality metric yields a larger number of isolated nodes implying a severe degradation in network performance. Increasing number of isolated nodes suggests that the network becomes increasingly intermittent in nature. It is worth noting that, in the Random Network Topology and the WaxMan Network Topology, as the number of nodes and $\alpha$ increase, the isolated nodes decrease. This is expected due to the fact that an increase in the node or edge density makes isolation of nodes more improbable. On the other hand, in the case of the Small World Topology as the number of nodes increases, so does the number of isolated nodes. This is due to the fact that in this topology the average node degree is fixed, which means that as the number of nodes increases, the number of nodes removed also increases which renders more nodes to become isolated. The fact that the node degree is originally fixed yields zero isolated nodes in the original network, as shown in Fig 5.

4.2. Algebraic Connectivity Evaluation

In section III, we have argued that the algebraic connectivity can be used as a robustness metric for the connectivity of the network. In this section, we use the algebraic connectivity as the performance metric, to show that the proposed criticality metric and key constituents such as the weighted node degree and the variation in link length outperform other proposals which have been proposed in literature. The evaluation has been simulative with the experiments conducted on Matlab. The Random Network Topology was considered. The number of nodes in the considered area were chosen in the range 20 – 80, and 10% of the most critical nodes were removed each time.

We first compare the proposed weighted node degree against the node degree metric. The weighted node degree aims at refining the criticality assessment of the normal degree metric by taking into account one hop paths which are identified by the existence of common neighbours. The reported algebraic connectivity values for various number of nodes are shown in Fig 6 for the original network, for the network when 10% of the most critical nodes are removed according to the weighted node degree metric and when they are removed according to the degree centrality metric. We observe that the weighted node degree achieves the most significant reduction in the algebraic connectivity value. Since the algebraic connectivity is a connectivity robustness metric it follows that the weighted node degree renders the network more susceptible to network partitioning indicating that it is more successful in identifying the most critical nodes.

We next use the algebraic connectivity to compare the variation in link length against the closeness centrality metric and the betweenness centrality metric. The variation in link length uses local information (neighbouring link length information) to identify nodes which are likely to act as relay nodes thus accommodating a large number
of active connections. The number of active connections at a node is also considered by the betweenness centrality metric, which, however, requires full network information in order to calculate all shortest paths in the network. In Fig 7 we show the reduction in the algebraic connectivity achieved by the closeness centrality, the betweenness centrality and the proposed variation in link length. We observe that the betweenness centrality and the variation in link length achieve the most severe reduction in the algebraic connectivity indicating that they are able to best assess the criticality of the nodes. In addition, we note that despite the fact that the variation in link length requires only local information its performance is comparable to that of betweenness centrality.

Finally, we use the algebraic connectivity to compare the performance of the proposed CBDI metric against the performance of the HILPR scheme and the Cont Scheme. The results are shown in Fig 8. We observe that the CBDI scheme removes nodes which reduce the algebraic connectivity to the greatest extent, indicating that the connectivity of the network is mostly affected. The Cont scheme achieves a smaller reduction, while the HILPR scheme leads to a slight increase in the algebraic connectivity despite the node removal.

4.3. Network Centric Evaluation

Our final set of experiments aim at evaluating the performance of the proposed criticality metric in more realistic network scenarios. We conduct the simulation experiments on the NS-3 Simulator (NS-3) [36] and evaluate the network performance using network centric performance criteria such as the total throughput, the average per packet delay, the average per packet jitter and the number of packet drops. In all the simulations we use the Random Network Topology to evaluate the performance of the proposed criticality metric against metrics such as: Cont [21], HILPR [1], Degree centrality, closeness centrality and betweenness centrality [8].

The evaluation was conducted on a wireless adhoc network of 100 nodes which were uniformly distributed in an area of $1500 \times 1500m^2$ thus forming a Random Network Topology. Each node was equipped with a 802.11b transceiver with a transmit power of 7.5 dbm. 15% of them had the option of transmitting at a power of $1.5 \times 7.5 dbm$ [37] thus forming long range communication links. The degradation in signal strength as a function of the distance covered was represented by the Friis loss propagation model. A randomly selected set of 20 source/sink pairs initiate the communication in the network by transmitting packets at a rate of 2.048 Kb/s each. Packet based transmission was assumed with the packet size set to 64 byte packets. Routing paths within the network are formed using the OLSR routing protocol [40]. All measurements are obtained in the interval $100 - 300$ seconds after the start of the simulation. This provides sufficient time for the OLSR algorithm to converge to its equilibrium state. The degradation in network performance is evaluated after 10% of the most critical nodes are removed from the network. This process is repeated 10 times with the results averaged to decrease the stochastic uncertainty of the obtained results.

We first compare the performance of the proposed criticality metric against the HILPR and Cont Algorithms in terms of the throughput achieved. The throughput is defined as the total number of packets delivered to their destinations within the network per unit time. The main goal of any network configuration is to maximize the achieved
throughput. In Fig 9 we show the achieved throughput as a function of time when 10% of the nodes are removed using the three metrics under consideration. We observe that the CBDI algorithm reports the largest decrease in the achieved throughput relative to the original network before node removal. This demonstrates that the proposed algorithm is successful in choosing more critical nodes. The decrease in average throughput observed at certain periods of time is due to long range link enabled nodes attempting to transmit at that time. Since their transmission power is higher, they attempt to reserve a larger portion of the common communication medium, thus increasing the probability of collisions and leading to throughput degradation.

We next use the achieved throughput as the performance metric in order to compare key components of the proposed CBDI metric against similar approaches which exist in the literature. We first compare the weighted node degree metric against the degree centrality metric. The proposed weighted node degree refines the degree centrality metric by considering as more critical, nodes which have small number of common neighbours with their neighbours. Smaller number of common neighbours indicates smaller number of one hop path alternatives when the node is removed. So, upon removal of a node with high criticality, it is easier for the network to become disconnected thus increasing the probability of reporting a smaller throughput. This is in fact what is reported by the simulation results presented in Fig 10. When removing nodes identified as critical using the weighted node degree, a larger degradation in throughput is achieved compared to node removal using the degree centrality metric. This demonstrates the superiority of the weighted node degree metric.

We next compare the proposed variation in link length metric against the betweenness centrality and the closeness centrality metrics. All three approaches aim at identifying nodes which accommodate the largest number of active connections. However, the closeness centrality and the betweenness centrality metrics use global network information whereas, the variation in link length utilizes local information only to achieve the same thing. The throughput achieved for the original network and when nodes are removed according to the various metrics are shown in Fig 11. The closeness centrality and the variation in link length achieve significant reduction in the throughput achieved. It is really striking to note that the betweenness centrality metric reports similar throughput to the original network prior to node removal. A possible explanation is the existence of alternative paths which upon node removal continue to render the network, ensuring high network throughput.

We now conduct similar experiments, aiming to compare the proposed criticality metric and its key constituents against other approaches, using other performance metrics. The delay experienced by packets in transit is an important network attribute which describes its performance. Low delays are preferable. In wireless ad hoc networks, such as the one considered in this study, delays are due to a number of reasons: network congestion resulting in queuing delays, poor channel behaviour resulting in re-transmissions and contention resulting in large vacant medium delay times due to the CSMA/CA mechanism. In this work, we consider the average per packet delay as the performance metric. This is calculated by dividing the total number of delays observed with the number of delays transmitted throughout the simulation time. In Fig 12 we show the time evolution of the average per packet delay reported in the original network and when nodes are removed according to a number of proposed criticality met-
metrics including the proposed criticality metric and its key constituents. We observe that the proposed CBDI metric is able to bring a major degradation in performance as the average per packet delay increases significantly when nodes are removed. In addition, the weighted node degree does not on average increase the per packet delays however, it does manage to outperform the degree centrality metric which reports smaller per packet delay values. When the variation in link length is now compared to the closeness centrality and the betweenness centrality metrics we observe that they eventually exhibit similar behaviour by decreasing the per packet delays compared with the original network. However, what is important is that despite individual constituent elements not always outperforming other proposals, when combined, achieve a significant degradation in network performance.

We next consider the average per packet delay jitter as the performance metric. The latter is calculated by dividing the total delay jitter observed throughout the simulation experiment with the total number of transmitted packets. The delay jitter is calculated as the variation in packet reception times at the receiver. Increasing delay jitter values indicate increasing congestion within the network, so small delay jitter values are preferable. In Fig 13, we show the time evolution of the average per packet jitter observed when nodes are removed based on a) CBDI, Cont and HILPR, b) Degree Centrality, Weighted Node Degree, c) Closeness Centrality, Betweenness Centrality, Variation in Link Length.

which is shown to decrease the average delay jitter relative to the original network and the closeness centrality metric. However, as mentioned above, despite individual elements, such as the variation in link length, not outperforming other proposals, when these are combined, cause the proposed CBDI metric to cause major degradation in network performance.

Finally, we consider the total number of dropped packets as the performance metric. High number of dropped packets in the network due to buffer overflow, is a strong indication of congestion. When nodes are removed from the network, the number of available paths decreases and the remaining paths are forced to accommodate all traffic. This makes them more vulnerable to congestion. When critical nodes are removed, congestion is expected to be more severe and the number of dropped packets is thus higher. The results of the conducted simulation experiments are shown in Fig 14. We observe that during the whole simulation time the proposed CBDI scheme is able to bring a major increase in the number of dropped packets compared to HILPR and Cont Algorithms. The other two algorithms report packet drops similar to the ones reported prior to node removal. Fig 14 also highlights the superiority of the weighted node degree relative to the degree centrality metric. Both cause the number of packets dropped to increase, however the increase achieved by the weighted node degree is higher. The picture for the variation in link length is different. While the variation in link length leads to an increase in the number of dropped packets, the closeness centrality metric reports an even higher number. The betweenness centrality metric in fact reports a slight decrease in the number of dropped packets. This is consistent with the throughput performance analysis an-
analyzed earlier. Despite the fact that the closeness centrality exhibits superior performance relative to the variation in link length metric, the superiority of the other constituent elements of the proposed criticality metric render it to be more successful than other metrics proposed in literature. In addition, as mentioned before, the variation in link length requires only local information whereas the closeness centrality requires full network information thus increasing the implementation complexity.

5. Conclusions

In this work, we highlight the contribution of critical nodes in network operation and demonstrate how the network reacts when these critical nodes are affected. We propose a new criticality index which is based on the diversity of node attributes within the network and the participation of each node in forming shortest path routes. We evaluate the performance of the proposed metric under various network topologies using multiple performance metrics and observe that the proposed metric outperforms existing approaches by showing a greater degradation in network performance when the critical nodes, selected using this index, are removed from the network.

References


J.-q. LIU, J.-l. LAN, J.-x. WU, Evaluating the vital network nodes based on node estranging, Computer Engineering & Science 3 (2011) 004.


M. Fiedler, Algebraic connectivity of graphs, Czechoslovak Mathematical Journal 23 (2) (1973) 298–305.


