Journal of Network and Computer Applications 64 (2016) 76–88

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Combined Banzhaf & Diversity Index (CBDI) for critical node detection

Waqar Asif [a,b,*], Hassaan Khaliq Qureshi [a], Muttukrishnan Rajarajan [b], Marios Lestas [c]

[a] National University of Sciences and Technology, H-12 Islamabad, Pakistan
[b] School of Engineering and Mathematical Sciences, City University, London, UK
[c] Department of Electrical Engineering, Frederick University, Nicosia, Cyprus

ARTICLE INFO

Article history:
Received 3 November 2014
Received in revised form
21 September 2015
Accepted 17 November 2015
Available online 5 February 2016

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

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.

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 (Shen et al., 2013) 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 (Liu et al., 2005). Moreover, in Krishna et al. (2002) 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 He et al. (2009). Finally, in Arulselvan et al. (2011) 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 (Narayanam, 2012). 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 (Krishnamoorthy and Deo, 1979).

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 (Freeman, 1977) 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 (Freeman, 1977) utilizes the average geodesic distance between all nodes in the network.

* Corresponding author at: School of Engineering and Mathematical Sciences, City University, London, UK.
E-mail addresses: 09mscsewasif@seecs.edu.pk (W. Asif), hassaan.khaliq@seecs.edu.pk (H.K. Qureshi), r.muttukrishnan@city.ac.uk (M. Rajarajan), eng.lm@frederick.ac.cy (M. Lestas).

http://dx.doi.org/10.1016/j.jnca.2015.11.025
1084-8045/

The node that has the highest closeness centrality value is the one which is placed in the geographical centre of the network and it thus has the shortest distance to all its neighbouring nodes. A distributed algorithm to find nodes with the highest closeness centrality value is presented in Wehmuth and Ziviani (2013). The global Clustering Coefficient metric (Costa et al., 2007) 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 (Mohar and Alavi, 1991). 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 the literature (Wei and Sun, 2011; Chen and Hero, 2011; Liu et al., 2014). Another set of approaches that exist in that literature is based on the ability of nodes to fragment the network when removed. Fan and Pardalos (2010) 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, Dinh et al. (2012) 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, Buldyrev et al. (2010), Huang et al. (2011) and Parshani et al. (2010) 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 Shen et al. (2013), 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 N-P 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 (Freeman, 1979) 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 Liu et al. (2011) 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 Zhang et al. (2011) where, taking into account the traffic shockwave model which was earlier proposed by Qu et al. (2001), 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 Cheng et al. (2001) where the delayflow of the network is used as the performance metric with which the criticality is assessed.

Furthermore, in Crucitti et al. (2003) and Taylor et al. (2006) 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 Liu et al. (2011) 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 Asif et al. (2014), 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 Shen et al. (2013), the *Controllability of complex networks* (*Cont*) in Liu et al. (2011) and the *degree centrality*, *betweenness centrality*, *closeness centrality* used in Freeman (1977). 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 2 we describe the proposed criticality metric, in Section 3 we elaborate on the algebraic connectivity of a network, in Section 4 we evaluate its performance using simulations and finally in Section 5 we offer our conclusion of the paper.