PHYSA: 17435 Model 3Gsc pp. 1–13 (col. fig: NIL)

ARTICLE IN PRESS

Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Link Influence Entropy

Priti Singh, Abhishek Chakraborty, B.S. Manoj *

Department of Avionics, Indian Institute of Space Science and Technology, Thiruvananthapuram, Kerala 695547, India

HIGHLIGHTS

- We propose a new metric, Link Influence Entropy (LInE), that can capture link/node influences and influence stability of nodes in a network.
- LInE can identify nodal importance based on link influence which degree based centrality or betweenness centrality cannot differentiate.
- It is found that spatial wireless networks and regular grid networks, respectively, demonstrate the lowest and highest LInE values.
- LInE has been extensively studied on various example real-world and arbitrary networks and proved its novelty.
- Usefulness of LInE is highlighted in estimating the influence stability of nodes in a dynamic network.

ARTICLE INFO

Article history:
Received 4 June 2015
Received in revised form 6 June 2016
Available online xxxx

Keywords:
Complex networks
Network entropy
Average path length
LInE
Influence stability
Wireless mesh networks

ABSTRACT

In this paper we propose a new metric, Link Influence Entropy (LInE), which describes importance of each node based on the influence of each link present in a network. Influence of a link can neither be effectively estimated using betweenness centrality nor using degree based probability measures. The proposed LInE metric which provides an effective way to estimate the influence of a link in the network and incorporates this influence to identify nodal characteristics, performs better compared to degree based entropy. We found that LInE can differentiate various network types which degree-based or betweenness centrality based node influence metrics cannot. Our findings show that spatial wireless networks and regular grid networks, respectively, have lowest and highest LInE values. Finally, performance analysis of LInE is carried out on a real-world network as well as on a wireless mesh network testbed to study the influence of our metric as well as influence stability of nodes in dynamic networks.

1. Introduction

The characteristic influence of each node or link determines the behavior of a network. Depending on how and where a node is placed, nodal influence varies with different networks. Recent studies on complex networks have unveiled unique characteristics which led to the emergence of newer network models such as small-world networks and scale-free networks. Small-world and scale-free network models are differentiated based on variation in the average path length, average clustering coefficient, and degree distribution [1].

Network entropy which is used to analyze complex networks is defined as a measure of uncertainty present in a network [2]. Shannon's entropy definition [2] is defined as $H(\cdot) = -\sum p(\cdot)\log(p(\cdot))$, where, $H(\cdot)$ is the average information present in a network, $p(\cdot)$ is the probability of occurrence of a state, and the logarithmic function of inverse of $p(\cdot)$ gives the self information contained in a state.

* Corresponding author.
E-mail addresses: priti@isac.gov.in (P. Singh), abhishek2003slg@ieee.org (A. Chakraborty), bsmanoj@ieee.org (B.S. Manoj).

http://dx.doi.org/10.1016/j.physa.2016.08.019
0378-4371/

Please cite this article in press as: P. Singh, et al., Link Influence Entropy, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.08.019

In Ref. [3], entropy of degree sequence was measured in order to understand heterogeneity in a scale-free network. Entropy is found to be the maximum for homogeneous or regular networks, whereas, the minimum entropy is observed for most of the heterogeneous networks [3]. Furthermore, authors of Ref. [4] studied entropy of the degree distribution to understand heterogeneity and robustness in the context of scale-free networks. In Ref. [5], network entropy was obtained based on an ensemble of different network configurations for Erdős–Rényi (ER) network model [6]. Different configurations in Ref. [5] are obtained by changing the network size and the link probability. On the other hand, authors of Ref. [7] determined network connectivity entropy to determine the degree of nodal or link connectivity even after removal of a few links that may create sub-graphs or isolated nodes in the network [7].

In Ref. [8], average information of the most important node in a network is measured with network centrality entropy. Network centrality entropy also measures heterogeneity of complex networks when various path distribution, that is number of shortest paths from a node in a bidirectional network, is concerned. In Ref. [9], the authors analyzed cyclic based entropy performance with respect to the degree based entropy in the context of different real-world networks. Authors of Ref. [10] identified equal degree nodes with automorphism partition which also efficiently calculates heterogeneity and complexity of a network. On the contrary, in Ref. [11], authors used entropy measure to study network evolution with Darwinian principles of variation and preferential selection.

A K-shell decomposition approach along with Susceptible-Infected-Recovered (SIR) and Susceptible-Infected-Susceptible (SIS) models was applied in Ref. [12] to identify a key spreader node in a complex network. In Ref. [13], nodal importance was evaluated based on the dynamical influence of a node. Dynamical influence, which is a centrality measure, plays a key role identifying a node's dynamic state which influences the collective behavior of a system. Authors in Ref. [14] proposed a semi-local centrality measure to design an efficient node ranking method in the context of undirected networks where mechanisms such as cascading, spreading, and synchronization are highly affected by a tiny fraction of influential nodes. In Ref. [15], influential nodes in a community were identified on the basis of influence scope maximization. The influence scope maximization was measured by deploying K-medoid clustering algorithm along with information transfer probability between any node-pair in a network. A centrality measure based on Dempster–Shafer evidence theory was proposed in Ref. [16] where the measure trades-off between degree centrality and nodal strength in a weighted network. In Ref. [17], on the other hand, authors captured edge-to-edge relationships based on the pseudo-inverse Laplacian of a network that unveils the global properties of a network. The derived edge-to-edge relationships help in understanding the network dynamics when edge influence is concerned.

In this paper, we propose Link Influence Entropy (LInE), a metric which measures influence of links in a network, and thus, effectively understand the influence as well as characteristics of nodes present in the network. LInE can find several real-world applications such as identifying bridge links in a network, discover hub-nodes in a scale-free network, or finding out traffic bottleneck of a network.

Rest of this paper is structured as follows. In Section 2, we discuss about features of LInE along with its application in the context of dynamic networks. Section 3 reports LInE performance when various network models are concerned. Moreover, LInE performance is evaluated in the context of a real-world network as well as in a testbed deployment of wireless mesh networks. We conclude this paper in Section 4.

2. Link influence entropy

Influence or importance of a link in a network varies based on its location in the network. Presence of long-ranged links (LLs) or normal links (i.e., short links) influence the way information is traversed in a network. Addition or removal of a certain link from a network affects the Average Path Length (APL). Link Influence Entropy (LInE) measures influence of each link in a network by comparing APL values with and without the link in a network. Details of LInE metric are presented in what follows.

2.1. LInE metric

Network APL, defined as hop-distance between a node-pair averaged over a network, is estimated as $APL = 2 \times \sum_{i \neq j} d(i,j)/(n \times (n-1))$, where, $d(i,j)$ is the path length in terms of number of hops between node-pair (i,j), and n is the total number of nodes in the network. In this paper, we consider network with bi-directional links, hence, the equation incorporates 2 in order to consider bi-directionality. A network with lower APL is more efficient than a network with higher APL due to its efficiency in the context of low end-to-end delay in data transmission and better quality of service (QoS). Behavior of LInE for a link is primarily influenced by network APL and location of links in a network. LInE of a network is quantified by Eq. (1),

$$H = -\sum_{i \neq j} p_{ij} \log p_{ij}, \tag{1}$$

where H is LInE of a network, and p_{ij} is the probability of influence of a link between node-pair (i,j), realized by Eq. (2).

$$p_{ij} = \frac{|APL - APL_{\overline{ij}}|}{\sum_{i \neq j} |APL - APL_{\overline{ij}}|}. \tag{2}$$