Computer Networks 104 (2016) 1–15

Contents lists available at ScienceDirect

Computer Networks

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

A Centrality Entropy Maximization Problem in Shortest Path Routing Networks

Vanniyarajan Chellappan[1,*], Krishna M. Sivalingam[1], Kamala Krithivasan

Department of Computer Science & Engineering, Indian Institute of Technology Madras, Chennai - 600036, India

ARTICLE INFO

Article history:
Received 21 June 2015
Revised 11 March 2016
Accepted 17 April 2016
Available online 30 April 2016

Keywords:
Traffic engineering
Topology design
Betweenness centrality
Entropy
Routing
OSPF

ABSTRACT

In the context of an IP network, this paper investigates an interesting case of the inverse shortest path problem using the concept of network centrality. For a given network, a special probability distribution, namely the *centrality distribution* associated with the links of a network can be determined based on the number of the shortest paths passing through each link. An entropy measure for this distribution is defined, and the inverse shortest path problem is formulated in terms of maximizing this entropy. We then obtain a centrality distribution that is as broadly distributed as possible subject to the topology constraints. A maximum entropy distribution signifies the decentralization of the network. An appropriate change in the weight of a link alters the number of the shortest paths that pass through it, thereby modifying the centrality distribution. The idea is to obtain a centrality distribution that maximizes the entropy. This problem is shown to be NP-hard, and a heuristic approach is proposed. An application to handling link failure scenarios in Open Shortest Path First routing is discussed.

1. Introduction

In the general context of network design, *a well designed topology is the basis for all stable networks* [1]. Two main design considerations for a good network topology design are: (i) reducing the single point of failures that can occur in the network; and (ii) reducing the hop count between any origin-destination (OD) pair. We investigate appropriate topology measurements based on the structural properties of the network, and study how these can be utilized to determine the maximally efficient topology. In this context, it is important to study the influence a node or link may have on the larger network based on its structural position in the topology. This will help in the identification of critical nodes and/or links in the network. A network is said to be highly centralized if some of its nodes or links are extremely critical to the operation of the network. Such a highly critical node or link conflicts with the design goal of eliminating single points of failure.

In this paper, we investigate a network-wide measurement called *network centrality* to determine the centralization of a network, as an instance of graph complexity measure [2]. The centrality distribution associated with the nodes or links of a network is determined based on the function of betweenness centrality values. A network wide measure is arrived by computing the entropy of the centrality distribution. We then formulate the network topology design problem as a problem of minimizing the centralization of the entire network or maximizing the entropy of the centrality distribution. We present a few interesting use cases of this proposed measure in the context of determining the efficiency of routing for a given topology.

Next, we study the inverse problem of determining the appropriate centrality distribution using suitable link weight setting techniques that maximize the entropy. This Centrality Entropy Maximization (CEM) problem is inspired by an earlier work called Network Entropy Maximization (NEM) [3], that connects the principle of maximum entropy with Internet Protocol (IP) routing. The CEM problem is shown to be NP-hard by reducing the known Open Shortest Path First (OSPF) optimal weight setting NP-hard problem [4]. We present a heuristic algorithm for the same. It is then shown, how this can be useful in handling link failure cases in OSPF networks. This paper consolidates and extends our previous work presented in [5,6].

The important contributions of this paper are summarized as follows: (i) studying the applicability of network centrality measure for network design problems; (ii) definition, proof of NP-hardness and a proposed heuristic solution for the centrality entropy maximization (CEM) problem; (iii) discuss various applications and use cases of the CEM framework including measuring the efficiency of routing in IP networks and understanding Braess Para-

* Corresponding author.
E-mail addresses: 003497@imail.iitm.ac.in, vanniaraj@gmail.com (V. Chellappan), skrishnam@iitm.ac.in, krishna.sivalingam@gmail.com (K.M. Sivalingam).
[1] India-UK Advanced Technology Centre of Excellence in Next Generation Networks, Systems and Services (IU-ATC), Chennai, India.

http://dx.doi.org/10.1016/j.comnet.2016.04.015
1389-1286/

dox in network routing; and (iv) an application of CEM for network topology design in tactical wireless networks.

The remainder of the paper is organized as follows. Section 2 presents the related work on entropy based measure of graph complexity and various centrality measures. Section 3 presents definitions along with some notations. Section 4 introduces the measure of *network centrality* and its variants. In Section 5, we show how the proposed network centrality measure can be applied to measure routing efficiency and detect Braess's paradox. Section 6 introduces the CEM problem and presents a heuristic approach to solve the same. Section 7 presents a use case of handling OSPF link failure case. Section 8 discusses the applicability of the proposed measure and the CEM framework to other interesting networking problems. Section 9 concludes the paper.

2. Related work

Centrality measures are often used in social networks to estimate the potential monitoring and control capabilities a person may have on communication flowing in the network. The concept of centrality has been extended to communication networks. Various centrality measures such as degree, closeness, and betweenness have been studied in the literature in order to analyze the internal topology of a given network [7]. These measures have been studied to quantify the influence of nodes or links on the dynamics of the entire network.

Betweenness centrality (BC) is one such graph theoretic concept that measures the degree to which a node or a link acts as an intermediary in the communication between every pair of nodes in the graph or topology. This measure of centrality is higher for certain nodes or links indicating that these nodes or links play a critical role. More precisely, the betweenness centrality of a node or link is determined by its occurrence in the shortest paths between pairs of nodes. There are different contexts in which *betweenness centrality* measure has been considered in a network [8–10]. The concept of Routing Betweenness Centrality (RBC) is introduced in [11], as a measure of the expected number of packets passing through a given node. A new edge betweenness centrality called *traffic-aware edge betweenness centrality* (TEBC) is defined in [12]. It is shown that TEBC can be used to influence and improve the performance of the shortest-path routing algorithm with respect to dynamic routing. Note that this metric is based on the fraction of traffic flow on an edge, and is used to re-balance the link's importance and lessen the problem of any bottleneck build-up on a link. More information on centrality related work can be found in [13,14].

The concept of node or link centrality has been extended further to the measurement of *network centrality* or *graph centrality* [7]. There are two distinct views in proposing such a graph- or network-wide measure. The first view leads to the development of measures of graph centrality based on the degree that all of its nodes or links are central. The alternative view leads to the development of measures of graph centrality based on the dominance of one node or link. We consider the first approach since it is more applicable in the network topology design problem.

We note that such a network-wide measure of centrality is also a measure of graph complexity. Graph complexity can be measured based on different structural features of the graph. For example, connectivity of a graph is measured based on node connectivity or link connectivity. The node or link connectivity is the smallest number of nodes or links whose removal results in a disconnected graph. This measure has been extended to measure the robustness of a network [15].

A taxonomy and overview of approaches to the measurement of graph complexity are presented in [2]. The taxonomy distinguishes between deterministic and probabilistic approaches. In the probabilistic approach, a probability distribution associated with the vertices or edges of a graph is determined based on the structural properties of the graph. Then, an entropy function is applied to the probability distribution to derive the measure of complexity. An entropy function measures how close a probability distribution is to being uniformly distributed or quantifies the *unevenness* of the probability distribution.

Shannon's entropy function [16] is one of the most commonly used entropy functions in measuring the complexity of graphs. In the context of an entropy function, the Principle of Maximum Entropy aims to determine a uniform or as broad a probability distribution as possible subject to the available constraints [17,18]. This principle has been used in solving some interesting networking problems [19,20]. The first work connecting the principle of maximum entropy with IP routing is called Network Entropy Maximization (NEM) [3].

This section summarized the related work. The next section presents the necessary definitions and notations.

3. Definitions and notations

This section provides necessary technical background, definitions and preliminaries of this paper. A network in its simplest form is a set of nodes or vertices joined together in pairs by edges or links. It can be represented as a directed graph $G = (V, E)$, where V is the set of vertices and E is the set of edges. An edge is labeled as (u, v) or simply uv, where $u, v \in V$. In a directed graph, $uv \neq vu$. For routing purpose, we assume that there are no self-loops, and the paths connecting any pair of vertices are loop-free. A graph is said to be *weighted* when we assign weight to each of its edges. Let $w : E \to \mathbb{R}_{\geq 0}$ be the weight function. If G is not provided with a weight function on the edges, we assume that each edge has unit weight. The weight is represented by $w_{u,v}$ for the link $(u, v) \in E$. A path in a graph is a finite sequence of edges which connect a sequence of vertices which are all distinct from one another. A graph is said to be connected when every pair of vertices is joined by a path.

Definition 1 (*Geodesic or Shortest Path*)**.** Given a connected weighted directed graph $G(V, E, w)$, associated with each edge $\langle u, v\rangle \in E$, there is a weight $w(u, v)$. The length of a path $p = \langle v_0, v_1, \ldots, v_k\rangle$ is the sum of the weights of its constituent edges:

$$w(p) = \sum_{i=1}^{k} w(v_{i-1}, v_i).$$

The length of the shortest path from u to v is defined by $\delta(u, v) = \min\{w(p) : p \text{ is a path from } u \text{ to } v\}$. $\delta(u, v)$ is called the distance between u and v. The path that realizes this distance is called the shortest path or geodesic.

There can be more than one shortest path between a pair of nodes.

Definition 2 ((*$\langle s, t\rangle$ induced subgraph*))**.** A sub-graph induced by the set of paths that connects the given source-destination pair $\langle s, t\rangle$. The union of paths that begin with s and end with t is called as the $\langle s, t\rangle$ induced subgraph, and is denoted by G_{st}.

Degree is a count of the number of edges incident upon a given node. The degree of a node is the simplest centrality measure of a node. It implies that the node with a higher degree has more incoming or outgoing paths, and hence critical to the entire network. Dividing it by the maximum possible degree $n-1$ gives us a normalized measure.

Definition 3 (*Closeness Centrality*)**.** As defined in [21], a node's closeness centrality is defined as the sum of the distances from