Reliability Engineering and System Safety 163 (2017) 46–56

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Effects of multi-state links in network community detection

Claudio M. Rocco[a], José Moronta[b], José E. Ramirez-Marquez[c,d,*], Kash Barker[e]

[a] Universidad Central de Venezuela, Caracas, Venezuela
[b] Universidad Simón Bolívar, Caracas, Venezuela
[c] Stevens Institute of Technology, Hoboken, NJ, USA
[d] Tecnologico de Monterrey, School of Engineering and Science, Guadalajara, Mexico
[e] School of Industrial and Systems Engineering, University of Oklahoma, USA

A R T I C L E I N F O

Keywords:
Multi-state
Networks
Communities
Similarity
Weights

A B S T R A C T

A community is defined as a group of nodes of a network that are densely interconnected with each other but only sparsely connected with the rest of the network. The set of communities (i.e., the network partition) and their inter-community links could be derived using special algorithms account for the topology of the network and, in certain cases, the possible weights associated to the links. In general, the set of weights represents some characteristic as capacity, flow and reliability, among others. The effects of considering weights could be translated to obtain a different partition. In many real situations, particularly when modeling infrastructure systems, networks must be modeled as multi-state networks (e.g., electric power networks). In such networks, each link is characterized by a vector of known random capacities (i.e., the weight on each link could vary according to a known probability distribution). In this paper a simple Monte Carlo approach is proposed to evaluate the effects of multi-state links on community detection as well as on the performance of the network. The approach is illustrated with the topology of an electric power system.

1. Introduction

A community in a network is generally defined as a subset of nodes "relatively densely connected to each other but sparsely connected to other dense groups in the network"[1]. For example, Fig. 1 depicts a network consisting of three communities [2]: Community A={1,2}, Community B={3,4,5}, and Community C={6,7,8,9}.

Community detection and analysis have expanded since the seminal work of Girvan and Newman [3], who identified the existence of communities in networks as an important network property. Porter et al. [1] highlight several applications, from technological systems (e.g., mobile phone communication) to social connections (e.g., social media) to biological systems (e.g., neural networks), among others. Recently, detecting communities in infrastructure networks has gained interest. In transportation networks, the identification of communities has enhanced planning in passenger traffic [4] and cargo traffic [5]. In electric power systems, the detection of communities has been used in innovative control procedures (e.g., area separations, identification of coherent generator, reactive power/voltage control) [6].

Recently, Rocco and Ramirez-Marquez [2] proposed an approach based on the use of community detection for identifying the most important nodes and links that "when removed from the network guarantee that a community is disconnected from the network." In this case, the vulnerability of the communities is evaluated considering the inter community links (ICL). For example, in Fig. 1, links between communities A and C or between B and C are considered critical and could be used for defining protection strategies or "for identifying potential weaknesses to disrupt or improve the network"[2].

The study of the partition in a network only requires an understanding of the topology of the network, or an un-weighted network. Recently Kim and Cho [7] analyzed the effects on the partition of a network when the information about network topology is incomplete, or where network topology is derived from partial observation. For networks where topology is known, Ramirez-Marquez et al. [8] recently analyzed the effects on community structures when one of several links in a network is lost due to failures, intentional attacks, or topology changes.

However, several authors (e.g., Mei et al. [6], Fan et al. [9]) have recognized that real networks contain weighted links. That is, network components have characteristics requiring a specific interpretation. For example, in a network of airports, the links should be weighted to represent the number of available seats on a connection [9], or in an electric power network, the transmitted reactive power between power substations should be considered as the weight for the corresponding

* Corresponding author at: School of Systems & Enterprises, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, United States.
E-mail address: jmarquez@stevens.edu (J.E. Ramirez-Marquez).

http://dx.doi.org/10.1016/j.ress.2017.02.004
Received 12 February 2016; Received in revised form 31 January 2017; Accepted 18 February 2017
Available online 20 February 2017
0951-8320/

Nomenclature

a	$(a_1, a_2, ..., a_m)^T$ network state vector denotes the state of all the links of the network
$\mathbf{b}_i$	$(b_{i1} = 0, b_{i2}, ..., b_{i\omega i} = M_i)^T$ the current state (capacity) of link i,
$\mathbf{p}_i$	$(p_{i1}, p_{i2}, ..., p_{i\omega})^T$ represents the probability associated with each state of link i
A	Set of links, $\lvert A\rvert = m$
$G(N,A,W)$	Network defined by N, A, and W
$\mathbf{F}_0$	$(f_{10}, f_{20}, ..., f_{m0})^T$ flow in the links using W_0
$\mathbf{F}_i$	$(f_{1i}, f_{2i}, ..., f_{mi})^T$ flow in the links using W_i, i=1, ...,$NSIMUL$
$FL_{\max i}$	i^{th} Maximum flow from source node s to terminal node t, i=1,...,$NSIMUL$
M_i	M_i equals the maximum capacity of component i
N	Set of nodes, $\lvert N\rvert = n$
$NSIMUL$	Number of samples
P_0	Set of communities in the reference partition
P_j	Partition corresponding to the set of weights F_j, j=1, ...,$NSIMUL$
SI(A,B)	similarity between two partitions A and B
SIM$_i$	i^{th} similarity between P_0 and P_i, i=1,...,$NSIMUL$
W	Set of link capacities, $\lvert W\rvert = m$
$\mathbf{W}_0$	$(w_{10}, w_{20}, ..., w_{m0})^T$ reference capacities defined for each link
W_i	i^{th} set of random capacities, i=1,...,$NSIMUL$
ω_j	Number of states of component j, j=1,...,$NSIMUL$

transmission line [6]. In general, weights are considered to be fixed values and could produce partitions that are different from those obtained when considering the network to be un-weighted. Moreover, the importance of links may be directly impacted by their weight.

Several authors have analyzed the problem of uncertain weights and how the robustness of the initial partition (i.e., the partition derived without considering uncertainties) is affected (e.g., [10–13]). Recently Rocco et al. [14] have considered the effects of weight uncertainties on the robustness of the initial partition, the initial communities (i.e., before considering uncertainties), and the ICL.

However, in many real situations, networks behave as a system of multi-state components. In these networks, each link is characterized via a vector of known random capacities (i.e., the weight in each link could vary according to a known probability distribution). As discussed by Ramirez-Marquez and Rocco [15], "multi-state networks are ubiquitous and include among others, electric distribution networks, power networks, supply chain networks, and railway and transportation networks." In these networks, the performance of the network is usually assessed by quantifying its ability to successfully supply a flow, service, or process between two specific nodes defined as the source and terminal (s-t) nodes. In such a case, several effects could be analyzed: (i) the performance of the network (e.g., the maximum flow transported between nodes s and t could be a random variable with a probability distribution), (ii) the partition of the network (i.e., the community structure and the inter community links); and (iii) the effects of considering a multi-state behavior on a selected group of links (e.g., the ICL). While the global analysis of the performance of multi-state networks has been addressed by many other authors (e.g., Lisnianski and Levitin [16], Satitsatian and Kapur [17], Zio et al. [18]), to the author's knowledge, effects (ii) and (iii) above for multi-state networks have not yet been explored.

To motivate community detection and the effects of variation on link capacities in multi-state networks, consider the network of 20 nodes and 32 links in Fig. 2. Assume that the capacity of each link can take on {0,5,10,15,60}.

Fig. 3a shows the network partition when the capacity of each links is fixed at its maximum capacity whereas Fig. 3b shows the partition when the capacity of the link between node 1 and 6 is fixed at 10 units (community detection is performed using the Fast Modularity algorithm [19] available in the igraph 0.7.2 library of the R statistical coding environment). Note that the community with the black border is the only community common to both figures, that is, it is robust to the selected perturbation.

The contribution of this work is a novel approach to analyze and understand the effects on the performance of the network (the maximum s-t flow) of multi-state links when detecting network communities. The remainder of this paper is organized as follows: Section 2 describes background on (i) the multi-state network and the approach used to evaluate the maximum flow when multiple sources and/or multiple terminals are considered and (ii) some indices used to globally compare community structures. Section 3 describes the proposed approach. Section 4 provides results of experimentation on

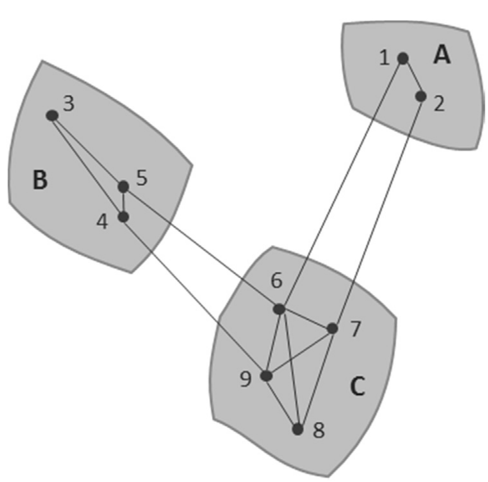

Fig. 1. An example partition of communities in a network [2].

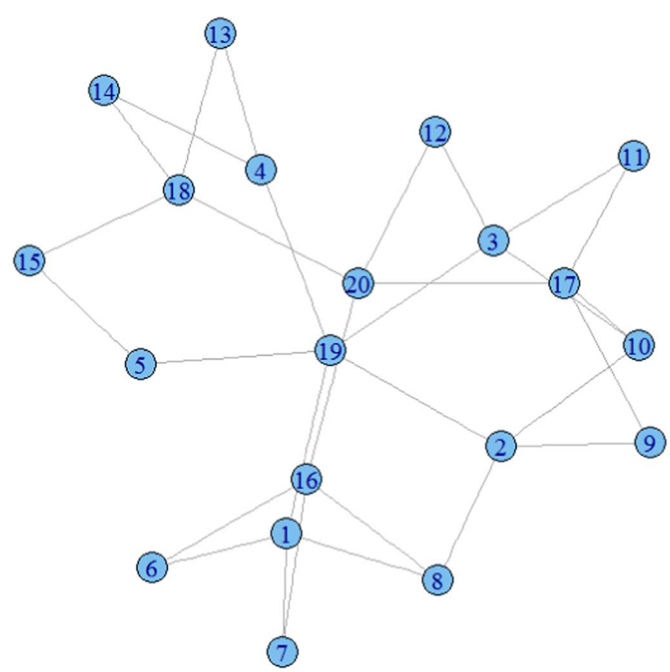

Fig. 2. An illustrative hypothetical network of 20 nodes, 32 links, and multi-state link capacities.