Discrete Applied Mathematics 210 (2016) 150–163

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Component-cardinality-constrained critical node problem in graphs

M. Lalou [a,b,*], M.A. Tahraoui [b], H. Kheddouci [b]

[a] Département d'Informatique, Faculté des Sciences Exactes, Université de Bejaia, 06000 Bejaia, Algérie

[b] Université Claude Bernard Lyon 1 - UFR-Informatique, Lab LIRIS, 43 bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France

ARTICLE INFO

Article history:
Received 3 January 2014
Received in revised form 5 January 2015
Accepted 29 January 2015
Available online 4 March 2015

Keywords:
Critical nodes
Complexity
Vertex cover problem
Trees
Chordal graphs

ABSTRACT

An effective way to analyze and apprehend the structural properties of networks is to find their most critical nodes. This makes them easier to control, whether the purpose is to keep or to delete them. Given a graph, *Critical Node Detection* problem (CNDP) consists in finding a set of nodes, deletion of which satisfies some given connectivity metrics in the induced graph. In this paper, we propose and study a new variant of this problem, called *Component-Cardinality-Constrained Critical Node Problem* (*3C-CNP*). In this variant, we seek to find a minimal set of nodes, removal of which constrains the size of each connected component in the induced graph to a given bound. We prove the NP-hardness of this problem on a graph of maximum degree $\Delta = 4$, through which we deduce the NP-hardness of *CNP* (Arulselvan et al., 2009) on the same class of graphs. Also, we study *3C-CNP* on trees for different cases depending on node weights and connection costs. For the case where node weights and connection costs have non-negative values, we prove its NP-completeness. While, for the case where node weights (or connection costs) have unit values, we present a polynomial algorithm. Also, we study *3C-CNP* on chordal graphs, where we show that it is NP-complete on split graphs, and polynomially solvable on proper interval graphs.

1. Context and definitions

1.1. Theoretical definitions

All graphs considered in this paper are finite. For a graph $G = (V, E)$, we use V and E to denote, respectively, the vertex set and the edge set. Let $A \subseteq V$ be a subset of V. We use $G[A] = (A, E(A))$, where $E(A) = \{(u, v) \in E | u, v \in A\}$, to denote the subgraph of G induced by A. The subgraph induced by $V \setminus A$ is denoted $G[V \setminus A]$. If G is a directed graph, the edge (u, v) is an arc directed from u to v.

A path in G is a sequence of nodes $(v_1, v_2, \ldots, v_k)$ such that each pair (v_i, v_{i+1}) is an edge in E. Two nodes are connected if there is a path between them. For each pair $(u, v) \in V \times V$, the pairwise connectivity $p(u, v)$ is defined as follows:

$$p(u, v) = \begin{cases} 1, & \text{if } u \text{ and } v \text{ are connected,} \\ 0, & \text{otherwise.} \end{cases}$$

* Corresponding author at: Département d'Informatique, Faculté des Sciences Exactes, Université de Bejaia, 06000 Bejaia, Algérie.
E-mail addresses: mohammed.lalou@gmail.com (M. Lalou), mohammed-amin.tahraoui@univ-lyon1.fr (M.A. Tahraoui), hamamache.kheddouci@univ-lyon1.fr (H. Kheddouci).

http://dx.doi.org/10.1016/j.dam.2015.01.043
0166-218X/

Given a connected component h, we denote σ_h the cardinality of h. Also, we use $f(h)$ to denote the pairwise connectivity of h, $f(h) = \frac{\sigma_h(\sigma_h-1)}{2}$.

A general non-negative weight (resp. cost) may be associated with each node $v \in V$ (resp. edge $(u, v) \in E$). For undefined terms, we refer the reader to [10].

1.2. 3C-CNP definition

In networks, nodes have varying degrees of importance. For many years now, the goal has been to find which are the most important for the network. In the literature, the most important nodes have appeared under different names, and for various purposes according to the problem at hand, such as: most influential nodes [22], most vital nodes [13], most k-mediator nodes [25], key-player nodes [11], etc. Recently, research has focused particularly on finding the most critical nodes in networks. A node is *critical* if its failure or malicious behavior significantly degrades the network performances. Once identified, the critical node can be the focus of defensive monitoring for a positive fit, or of offensive attempts for a negative one.

Given a graph, Critical Node Detection problem (*CNDP*) consists in finding a set of nodes, deletion of which results in a connectivity disruption in the induced graph. The level of disconnectivity in the induced graph is modeled as a metric to satisfy once the nodes have been deleted. In the literature, many connectivity metrics have been considered, including: maximize the number of connected components [29,30], reduce overall pairwise connectivity to some certain value [15], generate the maximum number of components with minimum difference on their cardinality [4,32], bound the largest component size to a given threshold [29], maximize (resp. minimize) the number of small (resp. large) components [31].

Identification of the most critical nodes has many applications in a number of fields, including network risk management [5], network vulnerability assessment [15,28], social network analysis [18,24], biological molecule studies [9,33], network immunization [4,23] and network communication [12,3].

The theoretical formulation for studying the combinatorial aspect of *CNDP* was first defined by Arulselvan et al. (2009) [4]. This definition was derived from the study explored by Borgatti (2006) [11], where he studied the *key-players* in networks using the pairwise connectivity metric. Arulselvan et al. presented two variants of the *CNDP*, namely *CNP* [4] and *CC-CNP* [6]. They proved their complexity on general graphs. Also, they derived mathematical formulations, and presented some heuristic algorithms for solving them in general graphs.

Since then, many studies and variants have been presented depending on the connectivity metric to be checked in the induced graph. Among recent works on this subject, we can cite the work of Di Summa et al. (2011) [32], in which they studied the *CNP* variant on trees taking into account the case of nonnegative weight for nodes and cost for connections. Also, Addis et al. (2013) [31] considered *CNP* on special classes of graphs, namely split graphs, bipartite graphs, complement bipartite graphs and graphs of bounded treewidth. Shen et al. [29,30] examined two new variants of *CNDP*, namely *MaxNum* and *MinMaxC*, which aim, respectively, to maximize the number of connected component, and minimize the size of the maximal component, in the induced graph. They studied both variants on trees, series parallel graphs and k-hole graphs. In [15,14], Dinh et al. presented a new formulation of *CNP*, called β-vertex separator problem. They studied its complexity and inapproximability on general graphs, and proposed a pseudo-approximation algorithm, as well as a heuristic approach to solve it on general graphs. Similarly, Y. Shen et al. [28,27] provided also complexity analysis for *CNP* (called it *CND*) on general graphs, unit disks and power-law graphs.

Below we give the definition of the first two alternatives of the *CNDP*, namely *CNP* and *CC-CNP*:

1. **CNP** [32,4]: given a graph G and an integer k, find a set of k nodes, deletion of which results in a maximum fragmentation of G by minimizing pairwise connectivity in the induced graph.
2. **CC-CNP** [6]: given a graph G, find the minimum set of nodes, deletion of which limits the connectivity indices of nodes, in the induced graph, to a given bound.

In this paper, we introduce and study a new variant of Critical Node Detection problem, called *3C-CNP* for *Component-Cardinality-Constrained Critical Node Problem*. Given a graph $G = (V, E)$, a non-negative cost c_{ij} is specified for each edge $(v_i, v_j) \in E$, a weight w_i is specified for deleting each node $v_i \in V$, and an integer $L > 0$ is given as input. Given a set of nodes $A \subseteq V$, deletion of which induces a residual graph $G[V \setminus A]$ whose connected components have the following sets of nodes: $h_1, h_2, \ldots, h_p$. The *3C-CNP* seeks to find A such that:

$$\textit{minimize} \sum_{v_i \in A} w_i \tag{1}$$

$$\textit{s.t.}\ \forall \sigma_k \in G[V \setminus A], \quad \sum_{v_i, v_j \in h_k} c_{ij} \leq L. \tag{2}$$

Thus, *3C-CNP* consists in finding a set of nodes $A \subseteq V$ of minimal total weight, such that the total connection costs of each connected component, in the induced graph $G[V \setminus A]$, is no more than L.

In the unit case, where $w_i = c_{ij} = 1$, *3C-CNP* asks for deleting as few nodes as possible, such that the pairwise connectivity of each connected component is at most L. (*i.e.*, minimizing the cardinality of A, rather than the total weight, such that the pairwise connectivity, rather than the total connection cost, is no more than L).