Network Majority on Tree Topological Network

Ravindra B. Bapat[*] Shinya Fujita[†] Sylvain Legay[‡] Yannis Manoussakis [§]
Yasuko Matsui[¶] Tadashi Sakuma[‖] Zsolt Tuza[**]

Abstract

Let $G = (V, E)$ be a graph, and $w : V \to \mathbb{Q}_{>0}$ be a positive weight function on the vertices of G. For every subset X of V, let $w(X) = \sum_{v \in G} w(v)$. A non-empty subset $S \subset V(G)$ is a *weighted safe set* if, for every component C of the subgraph induced by S and every component D of $G \setminus S$, we have $w(C) \geq w(D)$ whenever there is an edge between C and D.

In this paper we show that the problem of computing the minimum weight of a safe set is $\mathcal{NP}$-hard for trees, even if the underlining tree is restricted to be a star, but it is polynomially solvable for paths. Then we define the concept of a parameterized infinite family of "proper central subgraphs" on trees, whose polar ends are the minimum-weight connected safe sets and the centroids. We show that each of these central subgraphs includes a centroid. We also give a linear-time algorithm to find all of these subgraphs on unweighted trees.

1 Introduction

We can regard a network as a mature community on a large scale; more precisely, it consists of a collection of small communities with some mutual connections. In such a network, it is important to gain control of a "majority" so that we can control the network consensus. On the other hand, for those who are concerned about a network security, they would think that a network where we can easily get a majority is unstable and it has a risky structure in view of network vulnerability.

Motivated by these observations, we would like to give some appropriate definition for gaining a majority in a given network. As a network model, we here consider this problem on simple undirected graphs with some given weight on each vertex. Note that each weight on a vertex represents certain measure for importance in the network.

We use [4] for terminology and notation not defined here. Only finite, simple graphs are considered. For a graph $G = (V, E)$, let $\delta(G)$ be the minimum degree of G, and $\alpha(G)$ be the independence number of G. The order and size of G are denoted by n and m, respectively. The subgraph of G induced by a subset $S \subseteq V(G)$ is denoted by $G[S]$. When A and B are vertex-disjoint subgraphs of G, the set of edges that join some vertex of A and some vertex of B is denoted by $E(A, B)$.

Let $G = (V(G), E(G))$ be a graph and ω be a weighted function on $V(G)$ such that $\omega : V(G) \to \mathbb{R}_+$. For a vertex subset S of $V(G)$, let $\omega(S) := \sum_{v \in S} \omega(v)$. We often abuse notations for vertex subsets and subgraphs. So, for

[*]Indian Statistical Institute, 7, S.J.S. Sansanwal Marg New Delhi, 110016, India, E-mail address:rbb@isid.ac.in

[†]International College of Arts and Sciences, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan, E-mail address:shinya.fujita.ph.d@gmail.com

[‡]Laboratoire de Recherche en Informatique, University Paris-Sud, Orsay, Bât 650, 91405, Orsay cedex France, E-mail address:sylvain.legay@lri.fr

[§]Laboratoire de Recherche en Informatique, University Paris-Sud, Orsay, Bât 650, 91405, Orsay cedex France, E-mail address:Yannis.Manoussakis@lri.fr

[¶]Department of Mathematical Sciences, Faculty of Science, Tokai University, Japan, E-mail address:yasuko@tokai-u.jp

[‖]Systems Science and Information Studies, Faculty of Education, Art and Science, Yamagata University, 1-4-12 Kojirakawa, Yamagata 990-8560, Japan, E-mail address:sakuma@e.yamagata-u.ac.jp, the corresponding author

[**]Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Reáltanoda u. 13–15, Budapest, H-1053, Hungary, and Department of Computer Science and Systems Technology, University of Pannonia, Egyetem u. 10, Veszprém, H-8200, Hungary, E-mail address:tuza@dcs.uni-pannon.hu

a subgraph H of G, we write $\omega(H)$ for $\omega(V(H))$ (thus, $\omega(H) := \sum_{v \in V(H)} \omega(v)$).

If a connected subgraph H of G satisfies $\omega(G) \leqq 2\omega(H)$ then no one may object to considering that the subnetwork H plays a majority role in G. However, one might come up with the following natural question: Do we always need to get more than half weight for gaining the network majority?

To answer this question, Let us consider a weighted graph G with a weight function ω on $V(G)$, where we will always associate some given network $\mathcal{N}$ with (G, ω). (So we often identify/abuse notations (G, ω) and $\mathcal{N}$.) In view of graph topology, it may be natural to assume that the following three properties hold for $\mathcal{N}$:

(1) For any two vertices p, q in G, any communication between p and q is conducted on a path joining p and q in G.

(2) For a vertex subset S of G, when we consider the community associated with S on $\mathcal{N}$, as the consensus of S, S can block any communication for any two vertices on $V(G) \setminus S$ from two distinct components of $V(G) \setminus S$ by cuting off the paths on S joining them.

(3) For any two communities S_1, S_2 in G, S_1 and S_2 can form an alliance if and only if there is at least one safe way of communication (i.e. a path whose every vertex is in some communities which collude with either S_1 or S_2) between any pair of vertices in $V(S_1) \cup V(S_2)$.

For example, let us observe a weighted path $P_n = v_1 v_2 \dots v_{3n}$, where $\omega(v_i) = 1$ for all i. By taking a subpath $X = v_{n+1} v_{n+2} \dots v_{2n}$, we see that there is no component in $P \setminus V(X)$ whose weight sum exceeds the weight sum of X. Hence, under the above assumption, it would be appropriate for us to consider that X attains a majority role for any community on P. Hence we can conclude that the answer to the above question is negative. Moreover, to formulate our problem, we must consider the following basic question: How can we calculate the minimum weight of a subnetwork which attains a majority role for a given network? To answer this question, let us focus on a known concept called *safe sets*, which was introduced by Fujita, MacGillivray and Sakuma [9] for unweighted graphs. In this paper, we will generalize this concept to the weighted version in a natural manner and give some basic properties along this line.

A non-empty subset $S \subseteq V(G)$ is a *safe set* if, for every component C of $G[S]$ and every component D of $G \setminus S$, we have $|C| \geq |D|$ whenever $E(C, D) \neq \emptyset$. If $G[S]$ is connected, then S is called *a connected safe set*. The minimum cardinality among all safe sets (resp. connected safe sets) of G is called the *safe number* (resp. *connected safe number*) of G and is denoted by $\mathrm{s}(G)$ (resp. $\mathrm{cs}(G)$).

In this paper, we extend this concept on graphs in which each vertex has a positive weight. Formally, let $G = (V, E)$ be a graph, and $w : V \to \mathbb{Q}_{>0}$ be a positive weight function on the vertices of G. A non-empty subset $S \subset V(G)$ is a *weighted safe set* if, for every component C of the subgraph induced by S and every component D of $G \setminus S$, we have $w(C) \geq w(D)$ whenever $E(C, D) \neq \emptyset$. If $G[S]$ is connected, then S is called *a weighted connected safe set*. The minimum weight among all weighted safe sets (resp. connected safe sets) of (G, w) is called the *safe number* (resp. the *connected safe number*) of (G, w) and is denoted by $\mathrm{s}(G, w)$ (resp. $\mathrm{cs}(G, w)$).

As we mentioned before, the concept of (weighted) safe set can be thought as a suitable measure of network vulnerability, and hence it has some clear relation to other such graph invariants. For example, the *graph integrity*, a well studied measure [1, 2, 3, 14] of reliability of a graph network (G, w), is defined as

$$I(G) = \min_{S \subseteq V(G)} \big\{ w(S) + \max\{w(H) : H \text{ is a component of } G[V(G) \setminus S]\} \big\}.$$

From the definitions of the graph integrity and the safe number, we have the following:

Proposition 1. *For every graph network (G, w), the inequality $I(G) \leqq 2\,\mathrm{s}(G, w)$ holds. Furthermore, if a set $S (\subseteq V(G))$ attains the number $I(G)$ and the induced subgraph $G[S]$ is connected, then we also have the inequality $\mathrm{cs}(G, w) \leqq I(G) \leqq 2\,\mathrm{cs}(G, w)$.*

We show that a minimum safe set on weighted trees is also an appropriate indicator to express a central subgraph. We can obtain infinitely many scalings of the concept of "central subgraph", each of these includes a centroid. Moreover, if the vertex-weight function is uniform, then it is also included in a minimum safe set. The *betweenness centrality* of a vertex (an edge) is defined as the number of shortest paths that pass through that vertex (edge). In 1977, Linton [10] defined this concept and Girvan-Newman [11] extend the definition to the case of edges. Recently, the clustering of networks is in the focus of attention, many researches have proposed algorithms for it. Among them, some popular clustering algorithms typified by Girvan-Newman [11] tend to fail to extract communities with high betweenness centrality in a given network. (For example, some road traffic networks surely have such communities.) On the other hand, our concept of "central subgraphs" and algorithms to find them may be useful for extracting such communities on given networks. Newman [13] also introduced modularity clustering which is a major tool for detecting communities in a network. However, for finding maximum modularity, we need to spend much time. Actually, to the best of our knowledge, the fastest known algorithm for the modularity maximization problem on unweighted trees, developed by Dinha and Thai [6], runs in $O(n^5)$ steps. On the other hand, our algorithm can find any of the central subgraphs defined above in linear time. The substantial improvement (to optimal) in the order of running time indicates an advantage of our algorithm.

The paper is organized as follows.

In Section 2, we consider the time complexity of finding a minimum connected or non-connected safe set in a weighted tree. We show that this problem is $\mathcal{NP}$-hard even if the underlining tree is restricted to be a star. On the other hand, we construct a polynomial-time algorithm for finding a safe set of minimum weight on paths.

In Section 3, we define the concept of a parameterized infinite family of "proper central subgraphs" on trees, whose polar ends are the minimum-weight connected safe sets and the centroids. We show that each of these central subgraphs includes a centroid. We also give a linear-time algorithm to find all of these subgraphs on unweighted trees.

2 Complexity

2.1 $\mathcal{NP}$-completeness of Weighted Safe Set Problem

In this subsection, we consider the following decision problem:

CONNECTED VERTEX-WEIGHTED SAFE SET
INSTANCE: A connected graph $G = (V, E)$, a positive weight function $w : V \to \mathbb{Q}_{>0}$ on the vertex set V of G, and a positive rational number t.
QUESTION: Does there exist $S \subset V(G)$ with $w(S) \leq t$ such that $G[S]$ is connected and $w(S) \geq w(C)$ for every component C of $G \setminus S$?

We show the $\mathcal{NP}$-completeness of the above problem by a reduction from the following problem:

SUBSET SUM
INSTANCE: A finite set A, a size $s(a) \in \mathbb{Z}_{>0}$ for each $a \in A$, a positive integer I.
QUESTION: Is there a subset $A' \subseteq A$ such that the sum of the sizes of the elements in A' is exactly I?

The $\mathcal{NP}$-completeness of SUBSET SUM is well known.

Theorem 1 (Karp, 1972)**.** *The problem* SUBSET SUM *is $\mathcal{NP}$-complete.*

By using the above, we derive

Theorem 2. *The problem* CONNECTED VERTEX-WEIGHTED SAFE SET *is $\mathcal{NP}$-complete, even if the input graph is restricted to be a star.*

Proof of Theorem 2. Note that CONNECTED VERTEX-WEIGHTED SAFE SET clearly belongs to the class $\mathcal{NP}$. Let $T = (V, E)$ be a star defined by $V = \{c, u, v_1, \ldots, v_k\}$ and $E = \{cu, cv_1, \ldots, cv_k\}$. Let

$w : V \to \mathbb{Z}_{>0}$ be a positive integral weight function on the vertex-set V of T such that $w(c) = 1, w(u) = B, w(v_1) = a_1, \ldots, w(v_k) = a_k$, and $1 + \max\{a_i | i = 1, \ldots, k\} \leqq B \leqq \sum_{i=1}^{k} a_i$ hold.

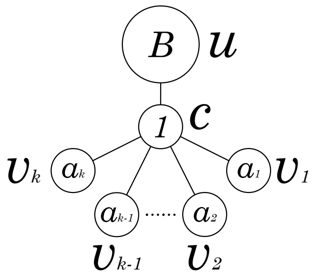

Figure 1: The star T and its vertex-weight w.

The set $\{c, u\}$ is clearly a connected safe set on (T, w). This set $\{c, u\}$ cannot be a minimum safe set on (T, w) if and only if there exists a subset $\Lambda \subseteq \{1, \ldots, k\}$ such that $B - 1 = \sum_{\lambda \in \Lambda} a_\lambda$ holds. Moreover, the set $\{c, u\}$ cannot be a minimum safe set on (T, w) if and only if there exists a connected safe set whose weight is at most B. Hence, by using the above gadget, we can reduce SUBSET SUM PROBLEM to CONNECTED VERTEX-WEIGHTED SAFE SET PROBLEM in a polynomial time, as follows:

Let $A = \{v_1, \ldots, v_m\}$ be an instance of SUBSET SUM, and let $s_i := 3s(v_i)$ for each $v_i \in A$. Set $w(v_i) = s_i$ for each $v_i \in A$. Set $B := 3I + 1$. Note that $\max\{s_i \mid i = 1, \ldots, m\} \leqq 3I = B - 1$ and $3 \leqq \min\{s_i \mid i = 1, \ldots, m\}$ hold. Set $k := m + 1$ and let v_{m+1} be an element outside of A such that $w(v_{m+1}) = B - 2$. Set $V := \{u, c, v_1, \ldots, v_k\}$ and $E := \{cu, cv_1, \ldots, cv_k\}$. Set $w(c) := 1, w(u) := B$. Set $t := B$. Note that any safe set X of the pair (T, w) with $w(X) \leqq B$ cannot contain the vertex v_{m+1}. Moreover we have that $B \leqq \sum_{i=1}^{k} s_i$. Hence the answer of SUBSET SUM for the instance is YES if and only if the answer of CONNECTED VERTEX-WEIGHTED SAFE SET for the instance graph $G := (V, E)$ is YES. $\square$

Actually, the problem CONNECTED VERTEX-WEIGHTED SAFE SET is $\mathcal{NP}$-complete on the following large class of graphs:

Corollary 1. *For an arbitrary given connected graph H, we have the following: The problem* CONNECTED VERTEX-WEIGHTED SAFE SET *is $\mathcal{NP}$-complete even if the input graph G is restricted to have a bridge e such that $G - e$ is a disjoint union of a star and the graph H.*

Proof of Corollary 1. Let all of V, E and w be as are defined in the last paragraph of the above proof of Theorem 2. Let h be a vertex of the graph H. Set $V' := V \cup V(H)$ and set $E' := E \cup E(H) \cup \{hv_{m+1}\}$. Reset $w(v_{m+1}) := (B - 2) - 0.1$. For every element v of V, set $w'(v) := w(v)$. For every vertex q of the graph H, set $w'(q) := \frac{1}{10|V(H)|}$. Let $G := (V', E')$. Then this $w' : V' \to \mathbb{Q}_{>0}$ be a positive weight function on the vertex-set of G such that the pair (T, w) has a safe set S with $w(S) \leqq B$ if and only if the pair (G, w') has a safe set S' with $w(S') \leqq B$. Hence the proof is complete. $\square$

2.2 Weighted safe set on paths

In this subsection, we consider the following optimization problem on paths.

WEIGHTED SAFE SET ON PATHS
INPUT: A path graph $P = (V, E)$ such that $|V| \geq 3$, and a positive weight function $w : V \to \mathbb{Q}_{>0}$.
OUTPUT: A safe set $S \subset V$ of minimum weight on P.

Theorem 3. *Finding a safe set of minimum weight is polynomial-time solvable on paths.*

To prove this theorem, we show that the problem of weighted safe set on a path is equivalent to finding a shortest weighted path on the acyclic digraph defined as follows: Let P be a path of n vertices, v_1, v_2, ..., v_n, with weights w_1, w_2, ..., w_n. For $1 \leq i \leq j \leq n$, we call $P_{i,j}$ the subpath of P consisting of the vertices v_i, v_{i+1}, ..., v_j.

From P, we will construct the weighted digraph $G_P = (V(G_P), A(G_P))$ as follows:

$$V(G_P) = \{u_{i,j}, v_{i,j} \mid 1 \leq i \leq j \leq n, (i,j) \neq (1,n)\} \cup \{t_0, t_\infty\},$$

$$A(G_P) = \begin{array}{l} \{(u_{i,j}, v_{j+1,k}) \mid 1 \leq i \leq j < k \leq n, w(P_{i,j}) \geq w(P_{j+1,k})\} \cup \\ \{(v_{i,j}, u_{j+1,k}) \mid 1 \leq i \leq j < k \leq n, w(P_{i,j}) \leq w(P_{j+1,k})\} \cup \\ \{(t_0, u_{1,j}), (t_0, v_{1,j}) \mid j \in \{1, 2, ..., n-1\}\} \cup \\ \{(u_{i,n}, t_\infty), (v_{i,n}, t_\infty) \mid i \in \{2, ..., n\}\}, \end{array}$$

$\forall i, j,\ w(u_{i,j}) = w(P_{i,j}), w(v_{i,j}) = 0,$
$w(t_0) = w(t_\infty) = 0.$

Lemma 1. *There exists a bijection between the safe sets in P and the $t_0 - t_\infty$ paths in G_P.*

Proof of Lemma 1. Let Q be any directed $t_0 - t_\infty$ path in G_P. Q can be described by a set of pair $\{(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)\}$ such that Q is the path t_0, $\ldots$, $v_{j_{l-1}+1, i_l - 1}$, u_{i_l, j_l}, $v_{j_l+1, i_{l+1}-1}$ $\ldots$, t_∞. The fact that there is an arrow between $v_{j_{l-1}+1, i_l-1}$ and u_{i_l, j_l} imply that $w(P_{j_{l-1}+1, i_l-1}) \leq w(P_{i_l, j_l})$ and the fact that there is an arrow between u_{i_l, j_l} and $v_{j_l+1, i_{l+1}-1}$ imply that $w(P_{i_l, j_l}) \geq w(P_{j_l+1, i_{l+1}-1})$. These conditions satisfy the definition of components C of $G[S]$ and D of $G \setminus S$ weighted safe set S. Then, $\bigcup_{l=1}^{k} P_{i_l, j_l}$ is a weighted safe set of P. □

In this correspondence, a safe set of P is composed of components of the form $P_{i,j}$, and the property of being a safe set is translated to the condition that these components come from a directed path in G_P as described above.

Lemma 2. *The weight of a safe set in P is equal to the weight of its $t_0 - t_\infty$ path in G_P.*

A small example is shown in Figure 2.

Theorem 4. *For a given path P, we can construct the weighted directed graph $G_P = (V(G_P), A(G_P))$ in $O(n^3)$ time and find a minimum $t_0 - t_\infty$ path of G in $O(n^3)$ time, where n is the number of vertices of P.*

By using Eppstein's algorithm [8], we can enumerate all the paths of minimum vertex-weight in G_P. This yields

Corollary 2. *All safe sets of minimum weight can be enumerated in polynomial time delay on paths.*

3 Centroid and its generalization

In this section we deal with the centroids on trees and introduce some generalizations. Let T be a weighted tree. For every α in $[0,1]$, let $F_\alpha(T, w) := \{X \subseteq V \mid T[V \setminus X]$ has no component whose weight exceeds $\frac{\alpha w(V)}{2} + (1-\alpha) w(X)\}$, $s_\alpha(T, w) = \min\{w(X) \mid X \in F_\alpha(T, w)\}$, and $F_\alpha^{\min}(T, w) = \{X \in F_\alpha(T, w) \mid w(X) = s_\alpha(T, w)\}$. A member of $F_0^{\min}(T, w)$ is called a *minimum safe set* of (T, w), while a member of $F_1^{\min}(T, w)$ is called a *centroid* of (T, w). The order of every centroid of (T, w) is exactly one.

Proposition 2. *If $0 \leqq \alpha \leqq \beta \leqq 1$, then $F_\alpha(T, w) \subseteq F_\beta(T, w)$.*

Proof of Proposition 2. Let X be an arbitrary set in $F_\alpha(T, w)$.

If $w(X) \geqq \frac{w(V)}{2}$, then $\frac{\beta w(V)}{2} + (1-\beta) w(X) \geqq \frac{w(V)}{2} \geqq w(V \setminus X)$. Since the weight of any component of $T[V \setminus X]$ does not exceed $w(V \setminus X)$, by the definition of $F_\beta(T, w)$, X is in $F_\beta(T, w)$.

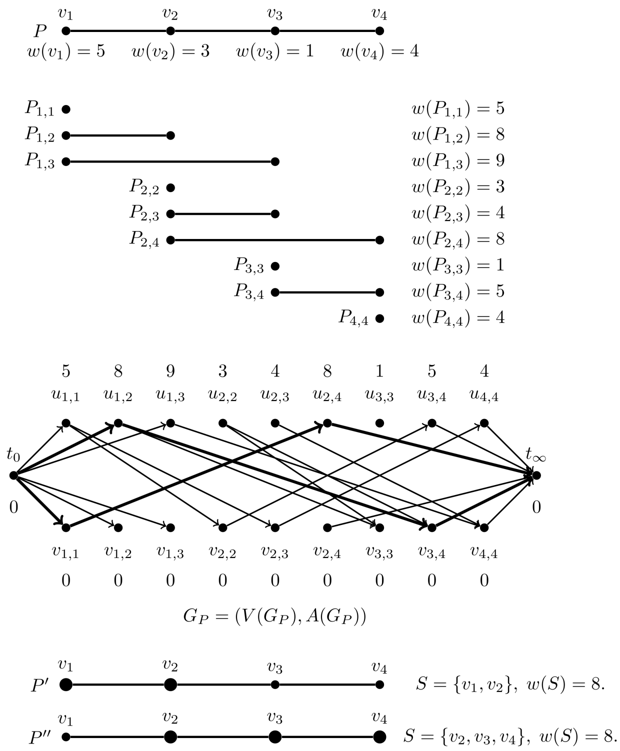

Figure 2: A path $P, P_{i,j}$ and the corresponding directed graph G_P. Each shortest $t_0 - t_\infty$ path in G_P corresponds to a minimum safe set S in P, and vice versa.

If $w(X) < \frac{w(V)}{2}$, then $\frac{\alpha w(V)}{2} + (1-\alpha)w(X) \leqq \frac{\beta w(V)}{2} + (1-\beta)w(X)$ holds, and hence, again by the definition of $F_\beta(T, w)$, X is in $F_\beta(T, w)$. $\square$

Corollary 3. *If $0 \leqq \alpha \leqq \beta \leqq 1$, then $s_\alpha(T, w) \geqq s_\beta(T, w)$.*

Corollary 4. *For every real number $\alpha \in [0, 1]$, $s_\alpha(T, w) \leqq s_0(T, w) \leqq \frac{w(V)}{2}$ holds.*

Proposition 3. *Let $T = (V, E)$ be a tree with n vertices, and $w : V \to \mathbb{R}_{>0}$ be a positive weight function on the vertices of T. If the pair (T, w) has at least two distinct centroids $u, v \in V$, then uv is an edge of T and the set $\{u, v\}$ is the set of all centroids of T.*

Proof of Proposition 3. Let P denote the path from u to v on T. Let p be the vertex on P adjacent to u, q the vertex on P adjacent to v. And let T_p denote the component of $T \setminus up$ containing p and v, T_q the component of $T \setminus qv$ containing q and u. Since u and v are centroids of T, both $w(V(T_p)) \leqq \frac{w(V)}{2}$ and $w(V(T_q)) \leqq \frac{w(V)}{2}$ hold. If $V(P) \setminus \{u, v\} \neq \emptyset$, then $V = V(T_p) \cup V(T_q)$ and $V(P) \setminus \{u, v\} \subseteq V(T_p) \cap V(T_q)$ and $w(V(P) \setminus \{u, v\}) > 0$ hold. Hence, we have $w(V) < w(V(T_p)) + w(V(T_q)) \leqq w(V)$, which is a contradiction. $\square$

Theorem 5. *Let $T = (V, E)$ be a tree with n vertices, $w : V \to \mathbb{R}_{>0}$ be a positive weight function on the vertices of T, and α be an arbitrary positive real number in $[0, 1)$. Let u be a centroid of (T, w). If an element S in $F_\alpha^{\min}(T, w)$ does not contain the centroid u, then S must contain the other centroid v, $w(S) = \frac{w(V)}{2}$ holds, and $V \setminus S$ is also an element of $F_\alpha^{\min}(T, w)$. Furthermore, in this case, $\{S, V \setminus S\} \subseteq F_0^{\min}(T, w)$ and $s_\alpha(T, w) = s_0(T, w)$.*

Proof of Theorem 5. Let u be a centroid of (T, w). Suppose that an element S in $F_\alpha^{\min}(T, w)$ does not contain the centroid u. If S does not contain any centroid of (T, w), then the maximum cardinality among all components of $T[V \setminus S]$ is strictly more than $\frac{w(V)}{2}$, which contradicts the fact that $s_\alpha(T, w) \leqq \frac{w(V)}{2}$. Hence S contains a centroid v. In this case, by removing the edge uv from T, the resulting graph has two components, each of whose weight is equal to $\frac{w(V)}{2}$. It means that $\frac{\alpha w(V)}{2} + (1-\alpha)w(S) = \frac{w(V)}{2}$, and hence either $\alpha = 1$ or $w(S) = \frac{w(V)}{2}$ holds. Since $\alpha < 1$, we have that $s_\alpha(T,w) = w(S) = \frac{w(V)}{2}$. And hence $s_\alpha(T,w) = s_0(T,w)$ holds. $\square$

Corollary 5. *Let $T = (V,E)$ be a tree with n vertices, $w : V \to \mathbb{R}_{>0}$ be a positive weight function on the vertices of T, and α be an arbitrary positive real number in $[0,1)$. If $s_\alpha(T,w) < \frac{w(V)}{2}$, then, for every real number $\beta \in [\alpha, 1)$ and for every member S of $F_\beta(T,w)$, S contains all the centroids of (T,w).*

Let $T = (V,E)$ be a tree with n vertices. For every real number $\alpha \in [0,1]$, let us define $F_\alpha(T) = \{X \subseteq V \mid$ *Every component of $T[V \setminus X]$ has at most $\frac{\alpha n}{2} + (1-\alpha)|X|$ vertices* $\}$, $s_\alpha(T) = \min\{|X| \mid X \in F_\alpha(T)\}$, and $F_\alpha^{\min}(T) = \{X \in F_\alpha(T) \mid |X| = s_\alpha(T)\}$.

Corollary 6. *Let $T = (V,E)$ be a tree with n vertices. If a set S in $F_\alpha^{\min}(T)$ does not contain a centroid of T, $F_\alpha^{\min}(T)$ has another element S' such that $S \cap S' = \emptyset$ and $|S| = |S'| = \frac{n}{2}$. Consequently, in this case, $s_\alpha(T) = s_0(T)$ holds.*

Corollary 7. *Let $T = (V,E)$ be a tree with at least 5 vertices, and α be an arbitrary real number in $[0,1)$. Every set in $F_\alpha^{\min}(T)$ contains all of the centroids of T.*

Let $G = (V,E)$ be a graph of order n. For each subset X of V, let $N(X)$ denote the *open neighborhood* $\{y \in V \setminus X \mid \exists x \in X$ such that $xy \in E(G)\}$ of X.

CONNECTED SAFE SET CONTAINING A SPECIFIED VERTEX (CSSV)
INPUT: A tree T with at least two vertices, a vertex p of T and a nonnegative real number α at most 1
OUTPUT: A set S in $F_\alpha(T)$ whose cardinality is minimum subject to $p \in S$
INITIALIZATION
Set $S = \{p\}$ ## $T[S]$ is always connected
For each vertex u in $V \setminus \{p\}$, find the unique edge uv such that v is a vertex on the path from u to p on T. Let T_u be the component of $T - uv$ containing u and set $c(u) = |V(T_u)|$.
MAIN LOOP
Repeat
  Let x be a vertex in $N(S)$ such that $c(x) = \max\{c(u) : u \in N(S)\}$
  Set $S := S \cup \{x\}$
until $\max\{c(u) : u \in N(S)\} \leqq \frac{\alpha n}{2} + (1-\alpha)|S|$
RETURN S

Theorem 6. *Let $T = (V,E)$ be a tree with at least 5 vertices, α be an arbitrary real number in $[0,1]$, and p be an arbitrary vertex of T. Algorithm* **CSSV** *finds a set S in $F_\alpha(T)$ whose cardinality is minimum subject to $p \in S$. Further, it can be implemented to run in time $O(n)$.*

Proof of Theorem 6. The following statements are easy to prove:

- The whole list of the pairs $\{(v, c(v)) \mid v \in V\}$ can be calculated in time $O(n)$;
- the subgraph induced by S is connected.

Now we will prove the correctness of the algorithm by reductio ad absurdum. Let S be the output of the algorithm. Suppose that there exists a member S_1 of $F_\alpha(T)$ such that $p \in S_1$ and $|S_1| < |S|$ hold. Then there must exist a leaf v of $T[S]$ such that $v \in S \cap S_1$ and that $T[V \setminus \{v\}]$ has a component X such that $S \cap V(X) = \emptyset$ and $|V(X)| \geqq \lfloor \frac{\alpha n}{2} + (1-\alpha)|S| \rfloor$ hold. Actually, if $T[S]$ has no such a leaf v, then, for every leaf x of $T[S]$ such that $x \in S \setminus S_1 (\neq \emptyset)$ holds, $S_0 := S \setminus \{x\}$ will be also a member of $F_\alpha(T)$, which contradicts the fact that the values $c(x)$ of vertices x added to S are monotonically decreasing in our algorithm. On the other hand, if such a leaf v of $T[S]$ exists, then the weight of a unique vertex q of X adjacent to v is strictly more than a weight of any leaf x of $T[S]$ such that $x \in S \setminus S_1 (\neq \emptyset)$ holds. More precisely, $c(x) < c(q) < c(v)$ holds, in this case. Hence, according to the rule of our algorithm, we must add q to the set S before we add x to the set, which is again a contradiction. □

Corollary 8. *If we set p to be a centroid of a tree T and set $\alpha := 0$, then the algorithm* **CSSV** *outputs a minimum safe set of T.*

Corollary 9. *For every tree T and for every two real numbers α, β such that $0 \leqq \alpha \leqq \beta \leqq 1$, there exist a member S_α of $F_\alpha^{\min}(T)$ and a member S_β of $F_\beta^{\min}(T)$ such that $S_\beta \subseteq S_\alpha$ holds.*

References

[1] M. Atici, C. Ernst, "On the Range of Possible Integrities of Graphs $G(n, k)$", *Graphs and Combinatorics* 27:475-485, 2011.

[2] KS. Bagga, L.W. Beineke, W.D. Goddard, M.J. Lipman and R.E. Pippert, "A survey of integrity", *Discrete Applied Mathematics* 37/38:13-28, 1992.

[3] D. Benko, C. Ernst, and D. Lanphier, "Asymptotic bounds on the integrity of graphs and separator theorems for graphs", *SIAM Journal on Discrete Mathematics* 23(1):265-277, 2009.

[4] G. Chartrand, L. Lesniak and P. Zhang, *Graphs and Digraphs*, 5th ed., Chapman and Hall, London, 2011.

[5] E. W. Dijkstra, "A note on two problems in connexion with graphs", *Numerische Mathematik* 1:269-271, 1959.

[6] T. N. Dinha and M. T. Thai, "Towards optimal community detection: from trees to general weighted networks." *Internet Mathematics* 11(3):181-200, 2015.

[7] P.G. Drange, M. S. Dregi and P. van't Hof, "On the computational complexity of vertex integrity and component order connectivity", In: Algorithms and Computation; Proceedings of the 25th International Symposium ISAAC 2014; LNCS, vol. 8889, pp. 285-297. Springer, Heidelberg (2014).

[8] D. Eppstein, "Finding the k Shortest Paths", *SIAM J. Computing* 28(2):652-673, 1998.

[9] S. Fujita, G. MacGillivray and T. Sakuma, "Safe set Problem on Graphs", Bordeaux Graph Workshop 2014 Enseirb-Matmeca and LaBRI, Bordeaux, France, 19-22 November, 2014 (Abstracts pp. 71-72).

[10] F. Linton, "A set of measures of centrality based on betweenness", *Sociometry* 40: 35-41, 1977.

[11] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks", *Proceedings of the National Academy of Sciences of the United States of America* 99, 7821-7826, 2002.

[12] B. Korte and J. Vygen, *Combinatorial Optimization*, 5th ed., Springer-Verlag Berlin Heidelberg, 2011.

[13] M. E. J. Newman, "Modularity and community structure in networks", *Proceedings of the National Academy of Sciences of the United States of America* 103 (23): 8577-8696, 2006.

[14] A. Vince, "The integrity of a cubic graph", *Discrete Applied Mathematics* 140:223-239, 2004.