Vertex Fusion under Diameter Constraints*

Marc Comas and Maria J. Serna
ALBCOM Research group, LSI Dept. Universitat Politècnica Catalunya
{mcomas,mjserna}@lsi.upc.es

June 3, 2007

Abstract

Given a graph G=(V,E), a positive integer k and a positive integer d, we want find a subset V_k with k vertices such the graph obtained by identifying the vertices from V_k in G has diameter at most d. We prove that for every $d \geq 2$ the problem is NP-complete. For the case of trees we provide a polynomial time algorithm that exploits the relationship with the r-dominating set problem.

1 Introduction

The problem of augmenting a graph by adding new edges to reach a certain improvement of a graph or network parameter is a source of important problems on network reliability and fault tolerant computing. One of the fundamental parameters in communication networks is the graph diameter. Observe that the delay in sending a a mesage from a node to another one depends on the length of the route among them. Thus a graph with small diameter, even a small world graph with logarithmic diameter, will be preferred to another one with bigger diameter as the total delay in communication will be smaller.

The most discused augmentation problems in the literature are based on the notion of graph augmentation by the addition of edges under some constraints. In particular, for the case in which the constrained parameter is the diameter the problem is formulated as follows:

> **Augmentation under diameter constraints** (ADC). Given a graph G and a positive integer d, add a minimum number of edges to obtain a graph of diameter at most d.

*partially supported by the FET pro-actives Integrated Project 15964 (AEOLUS), and the Spanish projects TIN-2004-07925-C03-01 (GRAMMARS), TIN-2005-09198-C02-02 (ASCE), and TIN2005-25859-E. The work of the first author was partially supported by a FPI scholarship of the Spanish Ministry of Education and Science.

The ADC problem was shown to be NP-complete in [10, 4]. The complexity of the problem restricted to trees remains open, however several partial results are known. In the case of paths the optimal value can be determined up to an additive error constant term [6] as well as for cycles [1]. Other lower and upper bounds can be found in [1, 5]. For the case of trees a 2-approximation algorithm for even d was presented in [4]. A 8-approximation algorithm for odd d was provided in [8]. Finally a $(2 + 1/\delta)$- approximation algorithm for the case of d odd is due to [3].

We are interested in analyzing augmentation problems under other definitions of graph augmentation that make sense in modern overlay communication networks. In such a setting a new network is obtained by the superposition of two different networks with possible different communication technologies and thus the augmentation might have a more involved topology than just the buying of some additional communication links. Particular cases are augmentation by vertex fusion (vertex identification see formal definition later), by imposing a clique or by imposing a complete bipartite graph, and in general by imposing a particular network on a node subset.

In this paper we deal with the problem of augmentation under diameter constraint when augmentation means vertex fusion. This is the problem in which we assume that the communication technology of the overimposed network allow instantaneous communication between the nodes, therefore they can communicate at zero cost, this is equivalent to consider a network in which we fusion all of them in one node. The problem can be formalized as follows:

> **Vertex fusion under diameter constraints** (VFDC). Given a graph G and a positive integer d, fusion a minimum number of vertices to obtain a graph of diameter at most d.

We show that the VFDC problem is NP-hard for general graphs. In the positive side we show that the problem can be solved in polynomial time when the input graph is restricted to be a tree. Our results imply also consequence on the parameterized complexity of the problem for the natural parameterizations.

2 Definitions

We use standard terminology for graphs. Let $G = (V, E)$ be a graph. We denote the *distance between two vertices* v and w by $d(v, w)$, which is defined the minimum number of edges in a path joining them. We denote by $d(v, W)$ the minimum distance between v and the subset $W \subseteq V$, formally $d(v, W) = \min_{w \in W} d(v, w)$.

The *diameter* of a graph G, denoted by $diam(G)$, is the maximum distance between two vertices, $diam(G) = max_{v,w \in V} d(v, w)$. The *r-neighborhood* of a vertex v, denoted using the ball notation as $B_r(v)$, is the set formed by all vertices at distance at most r from v, i.e., $B_r(v) = \{w \in V; d(v,w) \leq r\}$. Extending the definition to sets, the *r-neighborhood of a set of vertices* $W \subseteq V$ is the set formed by all the vertices at distance at most r from W, i.e., $B_r(W) = \{v \in V; d(v,W) \leq r\} = \bigcup_{w \in W} B_r(w)$. In the same way, we introduce the *ball of radius r centered in an edge* $e = \{v_1, v_2\} \in E$ as $B_r(e) = B_r(v_1) \cup B_r(v_2)$.

Given a graph G and a subset of vertices $W \subset V$, the *vertex fusion* of W in G is $G/W = (V/R_W, E')$, where R_W is the equivalence relation defined by $v \sim w$ if v and $w \in W$, and E' is the edges set induced by the quotient after removing the multiple edges and loops.

Given a graph $G = (V,E)$ and a subset $W \subset V$, we say that W is a *d-fusion set* in G if G/W has diameter at most d.

With the previous definitions we can define formally the problems we are interested in.

> **Vertex fusion problem** (VF). Given a graph G and two positive integers k and d. Does G have a d-fusion set with at most k vertices?
>
> **Min vertex fusion problem** (MVF). Given a graph G and a positive integer d. Compute the minimum k for which G has a d-fusion set with at most k vertices?

We alsconsider two parameterizations of the VF problem.

> d-**Vertex fusion problem** (d-VF). Given a graph G and an integer k. Does G have a d-fusion set with at most k vertices?
>
> k-**Vertex fusion problem** (k-VF). Given a graph G and two positive integers k and d. Does G have a d-fusion set with at most k vertices?

A *r-dominating set* of a graph G is subset $W \subset V$ for which the maximum distance from a vertex to W is at most r, this is $max_{v \in V} d(v,W) \leq r$. Or equivalently, $W \subset V$ for which $V = \bigcup_{w \in W} B_r(w)$. The associated problem is stated as follows:

> r-**dominating set problem**(r-DS). Given a graph G and an integer $k > 0$. Does G have an r-dominating set with at most k vertices?

Note that in the particular case that $r = 1$, the above formulation corresponds to the well known dominating set problem [7]. From [2, 9] we know that the r-dominating set is solvable in linear time in trees.

Now we introduce some additional notation for trees. Let $T = (V,E)$ be a tree. We define the *set of leaves* of T by $L(T)$ and if $W \subset V$ then the *subset of leaves* induced by W as $L(W,T) = W \cap L(T)$. Let $v \in V$ and let w be a neighbor of v in T. The subtree $T_v(w)$ is the subtree rooted at w in such a way that $v \notin T_v(w)$. Moreover we define the subtree $T_v^*(w)$ as the subtree induced by the vertex set $V(T_v(w)) \cup \{v\})$.

Finally, for any $r > 0$ and any subtree S of T, we use the notation $\beta_r(S)$ for the minimum number of balls of radius r that are necessary to cover S.

3 NP-Completeness

To prove the NP-hardness of the vertex fusion under diameter constrints problem, we show a strongest result, the NP-hardness of the d-VF problem for any $d \geq 2$. In the proof we modify the construction appeared in [4] used to prove the NP-hardness for edge augmentation problem.

Theorem 3.1. *The d-vertex fusion problem is NP-complete for any $d \geq 2$.*

Proof. Given a possible solution to the vertex fusion problem we can verify its correctness in polynomial time by calculating the distance from every vertex to every other vertex in G/V_k.

Therefore it remains to show that d-vc is an NP-hard problem. In this case we present a reduction from the dominating set problem (r-DS with $r = 1$). Assume that a value $d \geq 2$ has been fixed.

Given an instance of dominating set, consisting of a graph $G = (V, E)$ and an integer $k > 0$, we would reduce this problem to d-vertex fusion problem on a graph $G' = (V', E')$ and integer $k' = k + 1$,in such a way that there is a dominating set in G with k vertices if and only if we can obtain a d-fusion set with $k + 1$ vertices in G'.

We assume that the set of vertices in the original graph is $V = \{v_1, ..., v_n\}$. We have a different reduction for the cases where d is even and where d is odd.

Case 1: $d = 2r$ [*See the construction in Figure 1*]

Let $\{P_i = (V_i^P, E_i^P)\}_{1 \leq i \leq n}$ be a set of n paths with length $r - 1$ and let $v_i \in V$ be the first vertex of a path P_i in such a way that the vertex set V_i^p is formed by $V_i^p = \{p_i^1 = v_i, p_i^2, ..., p_i^{r-1}\}$ where p_i^j are new vertices for $j > 1$. Similarly, we consider a set of n paths $\{Q_i = (V_i^Q, E_i^Q)\}_{1 \leq i \leq n}$ with length r, having a common vertex b as their first vertex so that their vertex set is formed by $V_i^Q = \{q_i^1 = b, q_i^2, ..., q_i^r\}$ and a new vertex a.

Formally, we have $V' = V \cup \{v'_1, ..., v'_n\} \cup \{a\} \cup \bigcup_{i=1}^n (V_i^P \cup V_i^Q)$ as the vertex set in G'.

For the edges of G', we start with G joining its vertices to the paths P_i, a clique $K(V')$ on the vertex $V' = \{v'_1, ..., v'_n\}$, the paths Q_i and the vertex a. For any $1 \leq i \leq n$, we add the edges $\{w, v'_i\}$ for any $w \in B_1(v_i)$ and finally we connect a to b and a to every vertex v'_i.

and $E' = E \cup \{(a,b)\} \cup \bigcup_{i=1}^{n} (E_i^P \cup E_i^Q \cup \bigcup_{j=i}^{n} \{(v'_i, v'_j)\} \cup \bigcup_{w \in B_1^G(v_i)} \{(v'_i, v)\})$ as its edge set. See Figure 1 for the construction.

Now, let S be a dominating set in G, we want prove that $S \cup \{b\}$ is a d-fusion set in G'.

Consider the graph $G^* = G'/S \cup \{b\}$, and let v, w be two different vertices in G^*, we want prove that $d_{G^*}(v, w) \leq d$.

If v and w belongs in $V \cup V' \cup \{a, b\}$ then the distance between them in G^* can not be greater than 2 because their distance does not depend on the path's length and all the vertices in $V \cup V' \cup \{a, b\}$ are at distance at most one to $S \cup \{b\}$. So if we want consider the maximum distance in G^* we have to consider at least one vertex belongs in P_i or Q_i in G^*.

Consider that $v \in V_i^P$, then the furthest vertex in G^* is from v have to be the vertex q_j^r for some j or the vertex p_j^r for $j \neq i$. In the first case $d(v, q_j^r) = d(v, v_i) + r$ if $v_i \in S$ or $d(v, q_j^r) = d(v, v_i) + r + 1$ if $v_i \notin S$, but $d(v, v_i) \leq r - 1$ so $d(v, q_j^r) \leq d(v, v_i) + r + 1 \leq 2r$. And in the second case, $d(v, p_j^r) = d(v, v_i) + r - 1$ if $v_i, v_j \in S$, $d(v, p_j^r) = d(v, v_i) + r$ if $v_i \in S$ or $v_j \in S$ and $d(v, p_j^r) \leq d(v, v_i) + r + 1$ if $v_i, v_j \notin S$ so $d(v, p_j^r) \leq d(v, v_i) + r + 1 \leq 2r$. If we consider $v \in V_i^Q$, following the same way we arrive to $d(v, p_i^{r-1}) \leq 2r$ and $d(v, q_j^r) \leq 2r$. Note that in the worst case if we consider p_i^{r-1} and q_j^r for some i such that $v_i \notin S$ we have $d(p_i^{r-1}, q_j^r) = 2r$ then $S \cup b$ is a d-fusion set.

In the other direction, let S' be a d-fusion set in G'.

Suppose that $b \notin S'$ then if $k < n - 1$ $\exists i, j$ such that $V_i^P \cap S' = \emptyset$ and $V_j^Q \cap S' = \emptyset$, so the distance between p_i^{r-1} to q_j^r is at least the sum of the P_i's length plus one to go to S' and Q_j's length plus one to go to S', so $d(p_i^{r-1}, q_i^r) > 2r = d$, consequently $b \in S'$. Then we can write $S' = S \cup \{b\} = \{s_1, s_2, ..., s_l, b\}$. Note that if $k = n - 1$ then $V \setminus v$ for any vertex v is a dominating set in G, and if $k > n - 1$ V is a dominating set in G.

Let us consider a function t defined from $V' \cup \bigcup_{i=1}^{n} V_i^P$ to V where

$$\begin{aligned} v \in V_i^P &\longmapsto t(v) = v_i \\ v'_i \in V' &\longmapsto t(v'_i) = v_i \end{aligned}$$

be a function that convert a vertex in $V' \cup \bigcup_{i=1}^{n} V_i^P$ to a vertex in V. We claim that if $S \cup \{b\}$ is a d-fusion set in G' then $t(S) = \{t(s_1), t(s_2), ..., t(s_l)\}$ is a dominating set in G.

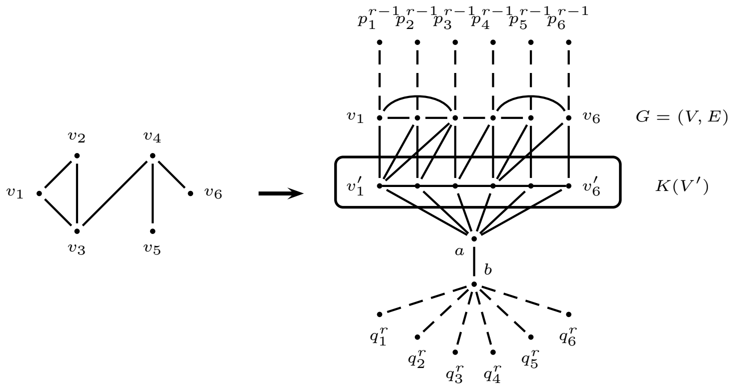

Figure 1: Transformation from Dominating set to Vertex fusion.*(even case)*

Suppose that $t(S) = \{t(s_1), t(s_2), ..., t(s_l)\}$ is not a dominating set in G, then exist $v_k \in V$ such that $d(v_k, t(S)) > 1$ in G' the relation is equivalent to exists $v_k \in V \subset V'$ such that $d(v_k, S) > 1$. We can consider one path Q_i such that $P_i \cap S = \emptyset$ then if we consider the distance between p_k^{r-1} and q_i^r we have $d(p_k^{r-1}, q_i^r) > r-1+1+r = 2r$ so $S \cup \{b\}$ is not a d-fusion set, consequently $t(S)$ have to be a dominating set in G.

Case 2: $d = 2r + 1$ [*See the construction in Figure 2*]

In this case the construction is similar to the construction before, but now we consider a set of vertices $B = \{b_1, ..., b_n\}$ connected to b and in this case the vertex set for every path Q_i is $V_i^Q = \{q_i^1 = b_i, q_i^2, ..., q_i^r\}$. Finally, we add to the edge set before G' the edges connecting every vertex b_i and an edge between b and every b_i.

Formally, the vertex set is defined by $V' = V \cup \{v'_1, ..., v'_n\} \cup \{a, b\} \cup \bigcup_{i=1}^n V_i^P \cup V_i^Q$ and the edge set is defined by $E' = E \cup \{(a, b)\} \cup \bigcup_{i=1}^n ((\{b, b_i\}) \cup E_i^P \cup E_i^Q \cup \bigcup_{j=i}^n \{(v'_i, v'_j), (b_i, b_j)\} \cup \bigcup_{w \in B_1^G(v_i)} \{(v'_i, v)\})$.

Let S be a dominating set in G, we want prove that $S \cup \{b\}$ is a d-fusion set in G'.

Consider the graph $G^* = G'/S \cup \{b\}$, and let v, w be two different vertices in G^*, we want prove that $d_{G^*}(v, w) \leq d$.

As in the case before, if we consider two vertices in $V \cup V' \cup \{a, b\} \cup B$ the distance between them can not be greater than 2, so we have to consider v or w are not belonging in $V \cup V' \cup \{a, b\} \cup B$.

Suppose that $v \in P_i$, if we consider $w \in V \cup V' \cup \{a, b\} \cup B$ their distance holds $d(v, w) \leq d(v, v_i) + 2$ the constant appearing in the second member is the cost necessary to go from v_i to a vertex in S and then go to w, so if we consider the vertex p_i^{r-1} we will have $d(v, w) \leq$

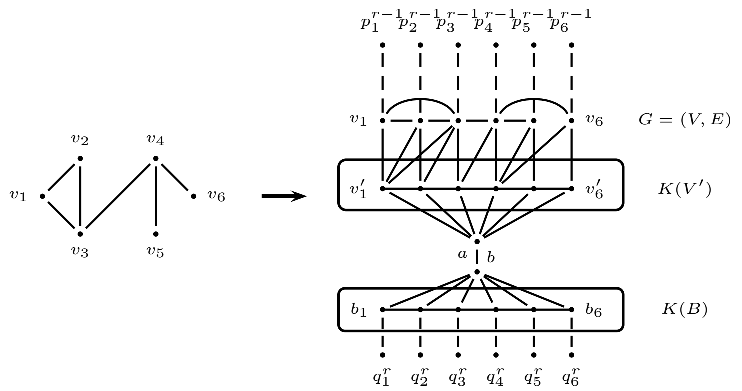

Figure 2: Transformation from Dominating set to Vertex fusion. *(odd case)*

$r+1$. If $w \in V_j^P$ then the distance between v and w is $d(v,w) \leq d(v,v_i)+d(w,v_j)+2$ if $v,w \notin S$ and $d(v,w) \leq d(v,v_i)+d(w,v_j)+1$ if v or $w \in S$ so we have $d(v,w) \leq d(v,v_i)+d(w,v_j)+2$ and this is maximum when consider p_i^{r-1} and p_j^{r-1} so $d(v,w) \leq 2r$. In the last case, if we consider $w \in Q_j$, we have $d(v,w) \leq d(v,v_i)+d(w,b_j)+2$ so maximizing the relation we obtain $d(v,w) \leq 2r+1$.

Suppose now that $v \in Q_i$ then if w belongs in $V \cup V' \cup \{a,b\} \cup B$ we have $d(v,w) \leq d(v,b_i)+2$ which implies $d(v,w) \leq r+2$ and if we consider w in Q_j their distance is given by $d(v,w) = d(v,b_i)+d(w,b_j)+1$, if we consider the vertices q_i^r and q_j^r we obtain $d(v,w) \leq 2r+1$. So $S \cup \{b\}$ is a d-fusion set.

In the other direction, suppose that S' is a d-fusion set in G'. We claim that exists $b^* \in S'$ for some $b^* \in \{b\} \cup B$.

Similarly to the example before, if $b^* \notin S'$ for any $b^* \in \{b\} \cup B$, then we can consider the distance between q_i^r and p_j^{r-1} for some j holding $V_j^P \cap S' = \emptyset$, so the distance between them have to be $d(q_i^r, p_j^{r-1}) \geq r+2+1+r-1 = 2r+2$.

Then we can write S' as $S \cup b^*$. With the same definition for the function t, we claim that $t(S) = \{t(s_1), ..., t(s_l)\}$ is a dominating set in G.

If $t(S)$ is not a dominating set in G, exist $v_k \in V$ for which $d(v_k, S) > 1$, so if we consider the distance between p_k^{r-1} and some q_j^r for which $V_j^Q \cap S = \emptyset$, we will have $d(p_k^{r-1}, q_j^r) > r-1+1+r+1 = 2r+1$ so $S \cup \{b^*\}$ is not a d-fusion set G' and consequently S is a dominating set in G.

□

As a consequence of the previous theorem and taking into account that the dominating set problem is W[1]-hard [], and that the construction in the proof of the previous theorem is a parameterized reduction, we have:

Corollary 3.2. *The vertex fusion problem under diameter constrains is NP-complete for any $d \geq 2$. The k-vertex fusion problem is W[1]-hard.*

4 Vertex fusion problem in trees

In the previous section we have shown that the problem is hard in general graphs. In this section we find an algorithm working in polynomial time for ths vertex fusion under diameter constrins for the case in which the input graph G is restricted to be a tree.

We start showing properties that relate the vertex fusion problem to the r-dominating set problem. Concretely, we use r-dominating sets to obtain bounds for d-fusion sets.

Proposition 4.1. *Let $T = (V, E)$ be a tree and r and k two positive integers. There is a $2r$-fusion set with k vertices in T if and only if there is an r-dominating set with k vertices in T.*

Proof. Obviously, if T has an r-dominating set S then S is a $2r$-fusion set in T.

For the opposite direction, assume that we have a $2r$-fusion set in T. We claim that we can construct an r-dominating set in T.

Let $V_k = \{v_1, v_2, ..., v_k\}$ be a $2r$-fusion set in T, if V_k is an r-dominating set in T we have finished so we have to analyze the case where V_k is not an r-dominating set in T.

Consider the following classification in V. Let $C_i = \{v_1^i, v_2^i, ..., v_{n_i}^i\} \subset V$ be the subset formed by the vertices in V with v_i as the nearest vertex in V_k. Formally we define the set C_i as $C_i = \{w \in V; d(w, v_i) \leq min_{j>i} d(w, v_j)$ and $d(w, v_i) < min_{j<i} d(w, v_j)\}$ (note that the second equation in the definition of C_i make that a vertex w belongs in just one component). In this way, we have $V = \bigcup_{i=1}^{k} C_i$. For clarity in the proof, we assume that the vertex v_1^i is one of furthest vertices in C_i to v_i. Let w_i be the vertex at distance r from v_1^i in the path between v_1^i to v_i, and consider the vertex $w_i = v_i$ if the class $C_i \setminus B_r(v_i)$ is empty.

Now we want prove, that the set formed by $\{w_1, ..., w_k\}$ is an r-dominating set in T. Assume that $v \in C_i$, we claim that $d(v, w_i) \leq r$ and consequently $v \in B_r(w_i)$.

Let $P(v_1^i, v_i)$ be the path between v_1^i and v_i and $P(v, v_i)$ the path between v and v_i. If $v \in P(v_1^i, v_i)$ then $d(v, w_i) \leq r$ because the distance between v_1^i and v_i is not greater than $2r$ and w_i is at distance r from v_1^i. So we have to consider $v \notin P(v_1^i, v_i)$. Obviously $P(v_1^i, v_i) \cap P(v, v_i) \neq \emptyset$ because $v_i \in$

$P(v_1^i, v_i) \cap P(v, v_i)$, so we consider the nearest vertex $v^* \in P(v_1^i, v_i) \cap P(v, v_i)$ to v.

We have two different cases to decompose the path $P(v_1^i, v_i)$:

1. $d(v_1^i, v_i) = d(v_1^i, w_i) + d(w_i, v^*) + d(v^*, v_i)$.
2. $d(v_1^i, v_i) = d(v_1^i, v^*) + d(v^*, w_i) + d(w_i, v_i)$.

In the first case, we have that $d(v, v_1^i) \leq 2r$, $d(v, w_i) = d(v, v^*) + d(v^*, w_i)$ and $d(v_1^i, v^*) = d(v_1^i, w_i) + d(w_i, v^*)$ so $2r \geq d(v, v_1^i) = d(v, v^*) + d(v^*, v_1^i) = d(v, v^*) + d(v^*, w_i) + d(w_i, v_1^i) = d(v, v^*) + d(v^*, w_i) + r = d(v, w_i) + r$ so this implies $d(v, w_i) \leq r$.

In the second case, we have that $d(v, v_i) = d(v, v^*) + d(v^*, w_i) + d(w_i, v_i)$ but $d(v_1^i, v_i) \geq d(v, v_i)$ so $d(v_1^i, v^*) + d(v^*, w_i) + d(w_i, v_i) \geq d(v, v^*) + d(v^*, w_i) + d(w_i, v_i)$ and then $r = d(v_1^i, w_i) \geq d(v, w_i)$.

So, we have that $B_r(w_i) \supseteq C_i$ and then we finally have that $V = \bigcup_{i=1}^k C_i \subseteq \bigcup_{i=1}^k B_r(w_i)$, so $\{w_1, w_2, ..., w_k\}$ is an r-dominating set in T.

□

As a consequence of the previous result we get:

Corollary 4.2. *Let $T = (V, E)$ be a tree. If there is an r-dominating set in T with k vertices but there is not any $(r-1)$-dominating set with k vertices in T then there is an $(2r-1)$-fusion set or $2r$-fusion set S with k vertices in T for which T/S has minimum diameter.*

Proof. Let $S(T, k)$ be the minimum d for which a d-fusion set with k vertices exists in T.

If there exists a r-dominating set in T with k vertices then $S(T, k) \leq 2r$. In the case that an $r-1$-dominating set does not exist in T we have $2(r-1) < S(T, k)$. Thus we obtain that $2r - 1 \leq S(T, k) \leq 2r$. Therefore the minimum diameter for a d-vertex fusion problem is $S(T, k) = 2r - 1$ or $S(T, k) = 2r$. □

The algorithm differs again depending on whether d is even or odd. To deal with the case of odd d we need an additional definition. Let $T = (V, E)$ be a tree, $S \subset V$ is a *(k, r)-nice set* if and only if $\exists V_k$ that is a r-dominating set with k vertices in $T \setminus S$ such that $\forall s \in S$ $d(s, V_k) \leq r + 1$ and $\forall s_1, s_2 \in S$ $d(s_1, s_2) \leq 2r + 1$.

With the previous definition and taking into account the properties of r-dominating sets, we have the following result.

Proposition 4.3. *Let $T = (V, E)$ be a tree and let r and k be two positive integers. There is a $(2r+1)$-fusion set with k vertices in T if and only if there is a set $S \subset V$ that is (k, r)-nice.*

Proof. In one direction, if there is a r-nice subset S in T, then we have that

- $\forall v, w \in V \setminus S$, $d(v,w) \leq 2r$
- $\forall v \in V$ and $\forall w \in V \setminus S$, $d(v,w) \leq 2r+1$
- $\forall v, w \in S$, $d(v,w) \leq 2r+1$

so in all the cases we have that for any $v, w \in V$, $d(v,w) \leq 2r+1$. Similarly we have that for any $v, w \in V$, $d(v, V_k) + d(w, V_k) \leq 2r+1$.

The opposite direction is proved using an argument similar to the one in the proof of Proposition 4.1. Let V_k be the $(2r+1)$-fusion set that by hypothesis existis in T. Consider the classification of V into k classes $C_i = \{v_1^i, v_2^i, ..., v_{n_i}^i\} \subset V$ which is formed by the vertices in V with v_i as their nearest vertex in V_k. Formally we define the set C_i as $C_i = \{w \in V; d(w, v_i) \leq min_{j>i} d(w, v_j)$ and $d(w, v_i) < min_{j<i} d(w, v_j)\}$. Let v_1^j be one of the vertices in C_j furthest to v_j and consider the set $S_i = \{s \in C_i;\ d(v_1^i, v_i) = d(s, v_i) \geq r+1\}$. Finally, if $P(v_1^i, v_i)$ is the path between v_1^i and v_i consider the vertex w_i as the vertex in $P(v_1^i, v_i)$ at distance $r+1$ to v_1^i if $S_i \neq \emptyset$ and $w_i = v_i$ if $S_i = \emptyset$.

We claim that the set $S = \bigcup_{i=1}^k S_i$ is a (k,r)-nice set in T. To prove tha claim we first show that for any $v \in C_i \setminus S_i$, $d(v, w_i) \leq r$.

Let $v \in C_j \setminus S_j$ and consider the path $P(v, v_j)$ and v^* the nearest vertex to v_1^j in $P(v, v_j) \cap P(v_1^j, v_j)$, then we have to cases:

1. $d(v_1^j, v_j) = d(v_1^j, w_j) + d(w_j, v^*) + d(v^*, v_j)$.
2. $d(v_1^j, v_j) = d(v_1^j, v^*) + d(v^*, w_j) + d(w_j, v_j)$.

In the first case, we have that $d(v, v_1^j) \leq 2r+1$, $d(v, w_j) = d(v, v^*) + d(v^*, w_j)$ and $d(v_1^j, v^*) = d(v_1^j, w_j) + d(w_j, v^*)$ so $2r+1 \geq d(v, v_1^j) = d(v, v^*) + d(v^*, v_1^j) = d(v, v^*) + d(v^*, w_j) + d(w_j, v_1^j) = d(v, v^*) + d(v^*, w_j) + r + 1 = d(v, w_j) + r$ so this implies $d(v, w_j) \leq r$.

In the second case, we have that $d(v, v_j) = d(v, v^*) + d(v^*, w_j) + d(w_j, v_j)$ but $d(v_1^j, v_j) > d(v, v_j)$ so $d(v_1^j, v^*) + d(v^*, w_j) + d(w_j, v_j) \geq d(v, v^*) + d(v^*, w_j) + d(w_j, v_j)$ and then $r + 1 = d(v_1^j, w_j) > d(v, w_j)$, consequently $r \geq d(v, w_j)$.

We can conclude that $W_k = \{w_1, w_2, ... w_k\}$ is an r-dominating set in $T \setminus S$ and consequently S is an r-nice set. □

As all the vertices in a (k,r)-nice set are contained in some ball centered at one edge, from now on we consider the tree T rooted at some edge e. We analyze properties of the vertices in this rooted tree that can imply their membership in a (k,r)-nice set.

Using the properties of r-dominating sets and the definition of niceness, we have:

Lemma 4.4. *Let $T = (V, E)$ be a tree rooted at some edge e, let r and k be positive integers. Assume that there is and edge ball $B_r(e)$ containing a (k, r)-nice set S in T. Consider an r-dominating set V_k in $T \setminus S$ then if V_k is an r-dominating set in $T \setminus (S \cup \{v\})$ for some $v \in B_r(e)$ then $S \cup \{v\}$ is a (k, r)-nice in T.*

Proof. Clearly, $\forall s_1, s_2 \in S \cup \{v\}$ $d(s_1, s_2) \leq 2r + 1$ because $S \cup \{v\} \subseteq B_r(e)$. Furthermore, $\forall s \in S \cup \{v\}$ $d(s, V_k) \leq r + 1$ due to the fact that $\forall s \in S$ $d(s, V_k) \leq r + 1$ and that $d(v, V_k) \leq r$ because V_k is an r-dominating set in $T \setminus S$. Therefore, we conclude that $S \cup \{v\}$ is a (k, r)-nice set in T. □

As a consequence of the previous results we have:

Corollary 4.5. *Let $T = (V, E)$ be a tree, let $e \in E$, and let r and k be two positive integers. If $S \subseteq V$ is a (k, r)-nice set in T contained in $B_r(e)$ then $S \cup L(B_r(e), T)$ is also a (k, r)-nice set in T.*

Proof. Let $T = (V, E)$ be a tree, let $e \in E$, and let r and k be two positive integers. If V_k is an r-dominating set in $T \setminus S$ then if we remove the leafs of T we have that V_k is an r-dominating set in $T \setminus (S \cup L(B_r(e), T))$. □

Corollary 4.6. *Let $T = (V, E)$ be a tree, let $e \in E$, and let $S \subseteq V$ be a (k, r)-nice set for some pair of positive integers k and r. If $v \in B_r(e)$ and $\beta_r(T_v^*(v_i)) = \beta_r(T_v(v_i)) + 1$ $\forall v_i$ son of v then $S \cup \{v\}$ is (k, r)-nice.*

Proof. Let V_k be an r-dominating set in $T \setminus S$, if we consider $T \setminus (S \cup \{v\})$ then $V_{k'} = V_k \cap \bigcup T_v(v_i)$ have at least $\sum \beta_r(T_v(v_i))$ vertices because if $V_{k'}$ have least vertices than $\sum \beta_r(T_v(v_i))$ then we have $\beta_r(T_v^*(v_i)) = \beta_r(T_v(v_i))$ for some son i. But we need just $\sum \beta_r(T_v(v_i))$ to r-dominate $\bigcup T_v(v_i)$ so by Lemma 4.4, $S \cup \{v\}$ is an r-nice set. □

Our next result shows that we always have a (k, r)-nice set without leaves in a subtree of T provided it exists in T.

Lemma 4.7. *Let $T = (V, E)$ be a tree, let $e \in E$, and let r and k be two positive integers. There is a (k, r)-nice set in T contained in $B_r(e)$ if and only if there is a (k, r)-nice set free of leaves in $T \setminus L(B_r(e), T)$ contained in $B_r(e)$.*

Proof. Let S be a (k, r)-nice set in T and let $e \in E$ be an edge such that $S \subseteq B_r(e)$, then if we consider the set $S' = S \setminus L(B_r(e), T)$ we have that S' is a (k, r)-nice set free of leafs in $T \setminus B_r(e)$. In the other direction, let S be a (k, r)-nice set free of leafs and let $S \subseteq B_r(e)$. Then if we add $L(B_r, T)$ to S by Corollary 4.5 we have that $S \cup L(B_r(e))$ is a (k, r)-nice set in T. □

Our next result provides the key result to select candidates for a (k, r)-nice set.

Lemma 4.8. *Let T be a tree , let $e \in E$, and let r and k be two positive integers. Let $v \in B_r(e)$ be such that $\beta_r(T_v^*(v_i)) = \beta_r(T_v(v_i)) + 1$ for any $v_i \in V$ that is a son of v. Then, there is a (k,r)-nice set free of leaves in T if and only if there is a (k',r)-nice set free of leaves in $T \setminus T(v)$, for $k' = k - \sum \beta_r(T_v(v_i))$.*

Proof. In one direction, let S be a (k,r)-nice set free of leaves in T then by Corollary 4.6 $S \cup \{v\}$ is n (k,r)-nice set free of leaves in T, which implies that $T \setminus (S \cup \{v\})$ has an r-dominating set V_k, but $\sum \beta_r(T_v(v_i))$ is the number of vertices necessary to cover $\bigcup T_v(v_i)$ so we can cover with $k - \sum T_v(v_i)$ vertices $(T \setminus T_v) \setminus (S \cup \{v\})$ so $(S \cup \{v\}) \setminus T_v$ is a (k',r)-nice set free of leaves in $T \setminus T_v$ for $k' = k - \sum \beta_r(T_v(v_i)$.

In the other direction, if S is an (k',r)-nice set free of leaves in $T \setminus T_v$ for $k' = k - \sum \beta_r(T_v(v_i))$, then $S \cup \{v\}$ is a (k,r)-nice set in T because we can cover $\bigcup T_v(v_i)$ with $\sum \beta_r(T_v(v_i))$ vertices. □

Now we can state here the main result

Theorem 4.9. *The vertex fusion problem under diameter constraints when the input graph is restricted to be a tree can be solved by a polynomial time algorithm.*

Proof. By Corollary 4.2 we can bound the minimum possible diameter for some r between $2r$ and $2r+1$.

From Lemma 4.7 we can find (if it exists) a (k,r)-nice set free of leaves in $T \setminus L(B_r(e),T)$ for some $e \in E$. To find it, we use Lemma 4.8 decreasing the size of the tree in every step. Finally and depending on whether a (k,r)-set exist or not in T, by Proposition 4.3 we can decide if the minimal diameter is $2r+1$ or $2r$.

The complexity of the algorithm is dominated by the complexity of computing the r-dominating set in all the subtrees of $T \setminus B_r(e)$ for every $e \in E$. This step has a cost $O(n^2)$ for a tree rooted on an edge e and thus, a total cost $O(n^3)$ when performing the computation for every $e \in E$. As we have to repeat the process using Lemma 4.8 at most k times, the total complexity of the algorithm is $O(kn^3)$. □

5 Conclusions and further results

We have shown that the VFDC problem is NP-hard in general and provided a polynomial time algorithm when the input is restricted to be a tree. Our algorithm is based on the computation of an adequate r-dominating set. We suspect that our algorithm can also be extended to other graph classes for which the dominating set problem can be solved in polynomial time.

Concerning parameterized complexity two parameter arise naturally, the graph diameter d and the size of the fusion set k, this leads to the definition of the d-**vertex fusion problem** and of the k-**vertex fusion problem**. Theorem 3.1 implies that the d-**vertex fusion problem** is NP-hard for $d \geq 2$ and thus not likely to be fixed parameter tractable. The reduction in Theorem 3.1 shows also that the size of the d-fusion set is $r + 1$, therefore taking into account that the k-DS problem is W[1]-hard, the k-**vertex fusion problem** is not likely to be fixed parameter tractable.

There remain several open problems related to vertex fusion an distances, for example the complexity of augmentation problem under radius or eccentricity constraints.

References

[1] N. Alon, A. Gyárfás, and M. Ruszinkó. Decreasing the diameter of bounded degree graphs. *J. Graph Theory*, 35:161–172, 2000.

[2] R. Chandrasekaran and A. Daughety. Location on tree networks: P-center and n-dispersion problems. Discussion Papers 357R, Northwestern University, Center for Mathematical Studies in Economics and Management Science, December 1978. available at http://ideas.repec.org/p/nwu/cmsems/357r.html.

[3] Victor Chepoi, Bertrand Estellon, Karim Nouioua, and Yann Vaxès. Mixed covering of trees and the augmentation problem with odd diameter constraints. *Algorithmica*, 45(2):209–226, 2006.

[4] Victor Chepoi and Yann Vaxès. Augmenting trees to meet biconnectivity and diameter constraints. *Algorithmica*, 33(2):243–262, 2002.

[5] P. Erdös, A. Gyárfás, and M. Ruszinkó. How to decrease the diameter of triangle-free graphs. *Combinatorica*, 18(4):493–502, 1998.

[6] Chung F.R.K. and Garey M.R. Diameter bounds for altered graphs. *J. Graph Theory*, 8:511–534, 1984.

[7] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.

[8] Toshimasa Ishii, Shigeyuki Yamamoto, and Hiroshi Nagamochi. Augmenting forests to meet odd diameter requirements. In *ISAAC*, pages 434–443, 2003.

[9] O. Kariv and S. L. Hakimi. An algorithmic approach to network location problems. I. The p-centers. 37(3):513–538, December 1979.

[10] A. A. Schoone, H. L. Bodlaender, and J. van Leeuwen. Diameter increase caused by edge deletion. *J. Graph Theory*, 11(3):409–427, 1987.