The Hardness and Approximation of the Densest $k$-Subgraph Problem in Parameterized Metric Graphs
Shih-Chia Chang, Li-Hsuan Chen, Ling-Ju Hung, Shih-Shun Kao, Ralf Klasing

To cite this version:
hal-03441672

HAL Id: hal-03441672
https://hal.archives-ouvertes.fr/hal-03441672
Submitted on 22 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The Hardness and Approximation of the Densest k-Subgraph Problem in Parameterized Metric Graphs

Shih-Chia Chang*, Li-Hsun Chen†, Ling-Ju Hung‡, Shih-Shun Kao§ and Ralf Klasing¶

*†Department of Business Administration, College of Management
National Taipei University of Business, Taipei, Taiwan
†‡ProBrand Technology, Inc., Taoyuan City 330, Taiwan
§¶CNRS, LaBRI, Université de Bordeaux, Talence, France
§Department of Computer Science and Information Engineering
National Cheng Kung University, Tainan, Taiwan

Abstract—A complete weighted graph $G = (V, E, w)$ is called $\Delta_{\beta}$-metric, for some $\beta \geq 1/2$, if $G$ satisfies the $\beta$-triangle inequality, i.e., $w(u, v) \leq \beta \cdot (w(u, x) + w(x, v))$ for all vertices $u, v, x \in V$. Given a $\Delta_{\beta}$-metric graph $G = (V, E, w)$, the $\Delta_{\beta}$-WEIGHTED DENSEST k-SUBGRAPH ($\Delta_{\beta}$-WDkS) problem is to find an induced subgraph $G[C]$ with exactly $k$ vertices such that the total edge weight of $G[C]$ is maximized. For $\beta = 1$, this problem, $\Delta$-WDkS, is known NP-hard and admits a $1/2$-approximation algorithms. In this paper, we show that for any $\beta > 1/2$, $\Delta_{\beta}$-WDkS is NP-hard. We also show how to modify any $\alpha$-approximation algorithm for $\Delta$-WDkS to obtain a $\delta_{\alpha, \beta}$-approximation algorithm for $\Delta_{\beta}$-WDkS with $\delta_{\alpha, \beta} > \alpha$ for every $\beta < 1$. Moreover, we prove that $\Delta_{\beta}$-WDkS can be approximated to within a factor $\frac{\beta}{\Delta_{\beta}}$ for any $\beta > \frac{1}{2}$.

I. INTRODUCTION

Various real-world systems can be modeled as graph-based representation. Many applications in social networks, communication networks, mobile ad hoc networks, World Wide Web (WWW) communities, bioinformatics are related to find a dense subgraph from a large graph [26]. In particular, on studying social networks, detecting cohesive subgroups is a very important task. It helps sociologists to understand the structures of networks. A cohesive subgroup can be defined as a complete graph (clique) [37]. However, it seems too restricted to consider a clique as a cohesive subgroup in real networks. The concept dense subgraph is a density-based clique relaxation model for defining cohesive subgraphs in social networks.

Given an undirected unweighted graph $G$, a densest $k$ subgraph of $G$ is an induced subgraph $G[C]$ of $G$ with exactly $k$ vertices such that the number of edges is maximized. If $G$ is a weighted graph, a densest $k$-subgraph of $G$ is an induced subgraph $G[C]$ of $G$ having exactly $k$ vertices satisfying that the total edge weight is maximized. The concept of densest $k$-subgraph is often used to define cohesive subgroups in a social network. In the following, we list the formal definition of the Densest k-Subgraph problem.

**DENSEST k-SUBGRAPH PROBLEM (DkS)**

**Input:** An undirected graph $G = (V, E, w)$, an integer $k > 0$.

**Output:** A vertex subset $C \subseteq V$, $|C| = k$ such that the number of edges in $G[C]$ is maximized.

**WEIGHTED DENSEST k-SUBGRAPH PROBLEM (WDkS)**

**Input:** An undirected weighted $G = (V, E, w)$, an integer $k > 0$.

**Output:** A vertex subset $C \subseteq V$, $|C| = k$ such that the total edge weight of $G[C]$ is maximized.

**Known results.** A densest $k$-subgraph is also called a $k$-cluster [25]. The problem of finding a densest $k$-subgraph in an undirected graph was introduced by Corneil and Perl [25]. It is a generalization of the maximum clique problem. The DkS problem is NP-hard on general graphs [25] and remains NP-hard on chordal graphs [25], bipartite graphs [25], planar graphs [32], even on graphs of maximum degree three [27]. Some exact exponential time algorithms were given for solving the DkS problem in general graphs [17], [18].

It has been shown that the DkS problems does not admit a Polynomial Time Approximation Scheme (PTAS) for general graphs under a complexity assumption [33]. There are PTASes given for graphs of minimum degree $\Omega(n)$ and dense graphs (of $\Omega(n^2)$ edges) when $k = \Omega(n)$ [3], stars of cliques [35] and interval graphs [39]. Many approximation algorithms were developed for the DkS problem on general graphs and special graphs. Feige et al. gave an approximation algorithm with approximation ratio $O(n^{\frac{1}{\alpha+\epsilon}})$, for some $\delta < \frac{1}{3}$ for the DkS problem on general graphs [28]. Bhaskara et al. improved the ratio to be $O(n^{1/4+\epsilon})$ for any $\epsilon > 0$ [7]. Asahiro et al. presented a simple greedy algorithm for this problem on general graphs.
and showed that the approximation ratio is $O(n/k)$ [4]. Chen et al. gave constant factor approximation algorithms for a large family of intersection graphs [19]. In [36], Liazzi et al. gave a $3$-approximation algorithm for chordal graphs. Backer and Keil gave a $\frac{3}{2}$-approximation algorithm for proper interval graphs and bipartite permutation graphs [5]. For WΔkS, it was shown NP-hard for metric graphs [40]. There are two approximation algorithms with approximation factors $4$ [40] and $2$ [29] for the WΔkS problem in metric graphs.

In this paper, we focus on solving the WΔkS problem in $\Delta_\beta$-metric graphs. A complete weighted graph $G = (V, E, w)$ is called $\Delta_\beta$-metric, for some $\beta \geq 1/2$, if $w(u, v) \geq 0$ for $u, v \in V$, and $G$ satisfies the $\beta$-triangle inequality, i.e., $w(u, v) \leq \beta \cdot (w(u, x) + w(x, v))$ for all vertices $u, v, x \in V$.

For $\beta = 1$, it defines the so-called metric graphs. The formal problem definition is listed in the following.

**Δβ-Weighted Densest k-Subgraph Problem (Δβ-WΔkS)**

**Input:** A $\Delta_\beta$-metric graph $G = (V, E, w)$, an integer $k > 0$.

**Output:** A vertex subset $S \subseteq V, |S| = k$ such that $w(S) = \sum_{u, v \in S} w(u, v)$ is maximized.

For $\beta = 1$, i.e., the input graph is a metric graph, we use $\Delta$-WΔkS to denote $\Delta_1$-WΔkS.

The design of approximation algorithms for the $\Delta$-WΔkS problem is related to the concept of stability of approximation for hard optimization problems [11], [15], [30], [31], [34]. It is similar to that of the stability of numerical algorithms. Suppose there is a small change in the specification (some parameters, characteristics) of the set of problem instances. It is of interesting to see that what the approximation ratio would be changed accordingly. We say an algorithm is stable if the change of the approximation ratio is small for every small change in the set of problem instances. There have been many research results on the concept of stability of approximation for solving fundamental hard optimization problems. E.g. in [1], [2], [6], [9]–[12], [38] it was shown that one can partition the set of all input instances of the Traveling Salesman Problem into infinitely many subclasses according to the degree of violation of the triangle inequality, and for each subclass one can guarantee upper and lower bounds on the approximation ratio. Similar studies demonstrated that the $\beta$-triangle inequality can serve as a measure of hardness of the input instances for other problems as well, in particular for the problem of constructing 2-connected spanning subgraphs of a given complete edge-weighted graph [13], and for the problem of finding, for a given positive integer $k \geq 2$, and an edge-weighted graph $G$, a minimum $k$-edge- or $k$-vertex-connected spanning subgraph [14], [16]. Moreover, $\beta$-triangle inequality is also applied to measure the hardness of several hub allocation problems [20]–[23].

In Section II, we prove that for any $\beta > \frac{1}{2}$, the $\Delta$-WΔkS problem is NP-hard. In Section III, we show how to modify any $\alpha$-approximation algorithm for $\Delta$-WΔkS to obtain a $\delta_{\alpha,\beta}$-approximation algorithm for $\Delta_\beta$-WΔkS with $\delta_{\alpha,\beta} > \alpha$ for every $\beta < 1$. In Section IV, we show that a $\frac{1}{2}$-approximation algorithm given in [29] for solving the WΔkS problem in metric graphs can be applied to solve the $\Delta_\beta$-WΔkS problem for any $\beta > \frac{1}{2}$ and the approximation ratio is $\frac{2}{\beta}$. The concluding remarks are given in Section V.

In this section, we prove that for $\beta > \frac{1}{2}$, the $\Delta_\beta$-WΔkS is NP-hard. This shows that even in subclasses of metric graphs $\beta < 1$ (e.g., $\beta = \frac{1}{2} + \epsilon$ for any $0 < \epsilon < \frac{1}{2}$), the $\Delta_\beta$-WΔkS is still NP-hard.

**Theorem 1.** For any $\beta > \frac{1}{2}$, the $\Delta_\beta$-WΔkS problem is NP-hard.

**Proof.** We prove that the $\Delta_\beta$-WΔkS problem is at least as hard as the NP-hard problem, the ΔkS problem.

For an input graph $G = (V, E)$ of the ΔkS problem, construct a $\Delta_\beta$-metric graph $G' = (V, E, w)$ such that $w(u, v) = 2\beta$ if $(u, v) \in E$, otherwise $w(u, v) = 1$. It is easy to see that $G'$ is a $\Delta_\beta$-metric graph satisfying the $\beta$-triangle inequality for $\beta > \frac{1}{2}$. We show that the $\Delta_\beta$-WΔkS problem is as hard as the ΔkS problem.

Let $C$ be an optimal solution of the $\Delta_\beta$-WΔkS problem in $G'$ and $w(C) = 2\beta \cdot p + (\frac{k}{2}) - p$ i.e., $G'[C]$ has $p$ edges with weight $2\beta$ and $(\frac{k}{2}) - p$ edges with weight 1. Since the edge cost in $G'$ is either $2\beta$ or 1, we see that $G[C]$ has exactly $p$ edges. Suppose that there exists a vertex subset $D$ of size $k$ such that $G[D]$ has more than $p$ edges. It is easy to see that in $G'$, $w(D) > 2\beta \cdot p + (\frac{k}{2}) - p$, a contradiction. Thus, if $C$ is an optimal solution of the $\Delta_\beta$-WΔkS problem in $G'$, then $C$ is an optimal solution of the ΔkS problem in $G$. Notice that $w(u, v) \geq 0$ for $u, v \in V$ since $G$ satisfies the $\beta$-triangle inequality.

By the fact that the ΔkS problem is an NP-hard problem, this implies that the $\Delta_\beta$-WΔkS problem is also an NP-hard problem. This completes the proof.

**Remark 1.** Theorem 1 shows that the $\Delta_\beta$-WΔkS problem is already NP-hard on the class of $\Delta_\beta$-metric graphs where all the edge costs are in $\{1, 2\beta\}$.

**III. Using Δ-WΔkS Approximation Algorithms for Δ_β-WΔkS**

In this section we show how to modify any $\alpha$-approximation algorithm for $\Delta$-WΔkS to obtain a $\delta_{\alpha,\beta}$-approximation algorithm for $\Delta_\beta$-WΔkS with $\delta_{\alpha,\beta} > \alpha$ for every $\beta < 1$. The advantage of this approach is that any improvement on the approximation of $\Delta$-WΔkS automatically results in an
improvement of the approximation ratio for $\Delta_\beta$-WD$k$S. The idea of this approach is to reduce an input instance of $\Delta_\beta$-WD$k$S to an input instance of $\Delta$-WD$k$S by subtracting a suitable cost from all edges.

**Lemma 1** ([9]). Let $G$ be a $\Delta_\beta$-metric graph for $\frac{1}{2} \leq \beta < 1$. Let $c_{\min}$ and $c_{\max}$ be the minimum edge cost and maximum edge cost in $G$ respectively. Then $c_{\max} \leq \frac{2\beta^2}{1-\beta} \cdot c_{\min}$.

**Theorem 2.** Let $A$ be an approximation algorithm for $\Delta$-WD$k$S with approximation ratio $\alpha$, and let $\frac{1}{2} < \beta < 1$. Then, $A$ is an approximation algorithm for $\Delta_\beta$-WD$k$S with approximation ratio $\alpha + (1 - \alpha) \cdot \frac{(1-\beta)^2}{\beta^2}$.

**Proof.** Let $I = (G, \text{cost})$ be a problem instance of $\Delta_\beta$-WD$k$S, $\frac{1}{2} < \beta < 1$. Let $c = (1-\beta) \cdot 2 \cdot c_{\min}$ where $c_{\min}$ is the minimum edge cost in $G$. For all $e \in E(G)$, let $\text{cost}'(e) = \text{cost}(e) - c$. Then the WD$k$S instance $I' = (G, \text{cost}')$ still satisfies the triangle inequality: Let $x, y, z$ be the costs of the edges of an arbitrary triangle of $G$. Then $z \leq \beta \cdot (x + y)$ holds. Since $c = (1-\beta) \cdot 2 \cdot c_{\min} \leq (1-\beta) \cdot (x+y)$ it follows that $z \leq \beta \cdot (x+y) \leq x+y-c$ and thus $z-c \leq (x-c)+(y-c)$.

Furthermore we know that a $k$-subgraph is optimal for $I'$ if and only if it is optimal for $I$. Let $H_{\text{opt}}$ be an optimal $k$-subgraph for $I$. Let $H$ be the $k$-subgraph that is produced by the algorithm $A$ on the input $I'$. Then $\text{cost}'(H) \geq \alpha \cdot \text{cost}'(H_{\text{opt}})$ holds and thus

$$\text{cost}(H) - \left(\frac{k}{2}\right) \cdot c \geq \alpha \cdot (\text{cost}(H_{\text{opt}}) - \left(\frac{k}{2}\right) \cdot c).$$

This leads to

$$\text{cost}(H) \geq \alpha \cdot \text{cost}(H_{\text{opt}}) + (1-\alpha) \cdot \left(\frac{k}{2}\right) \cdot c$$
$$= \alpha \cdot \text{cost}(H_{\text{opt}}) + (1-\alpha) \cdot \left(\frac{k}{2}\right) \cdot (1-\beta) \cdot 2 \cdot c_{\min}$$
$$\geq \alpha \cdot \text{cost}(H_{\text{opt}}) + (1-\alpha) \cdot \left(\frac{k}{2}\right) \cdot (1-\beta) \cdot 2 \cdot \frac{1-\beta}{\beta^2} \cdot c_{\max} \quad (\text{by Lemma 1})$$
$$= \alpha \cdot \text{cost}(H_{\text{opt}}) + (1-\alpha) \cdot \frac{(1-\beta)^2}{\beta^2} \cdot \left(\frac{k}{2}\right) \cdot c_{\max}$$
$$\geq \alpha \cdot \text{cost}(H_{\text{opt}}) + (1-\alpha) \cdot \frac{(1-\beta)^2}{\beta^2} \cdot \text{cost}(H_{\text{opt}})$$
$$= \left(\alpha + (1-\alpha) \cdot \frac{(1-\beta)^2}{\beta^2}\right) \cdot \text{cost}(H_{\text{opt}})$$

which completes the proof. \qed

According to Theorem 2, we have the following corollary.

**Corollary 1.** For $\frac{1}{2} \leq \beta < 1$, Algorithm 1 is a $(\frac{1}{2} + \frac{(1-\beta)^2}{\beta^2})$-approximation algorithm for $\Delta_\beta$-WD$k$S.

Note that Corollary 1 provides a weaker approximation ratio than Theorem 3 in the next section.

**IV. A $\frac{1}{2\beta^3}$-APPROXIMATION ALGORITHM FOR ALL $\beta > \frac{1}{2}$**

In [29], a $\frac{1}{2\beta^3}$-approximation algorithm was given for solving the WD$k$S problem in metric graphs. We list this algorithm in Algorithm 1. In this section, we show that Algorithm 1 can be applied to solve the $\Delta_\beta$-WD$k$S problem for any $\beta > \frac{1}{2}$ and the approximation ratio is $\frac{1}{2\beta}$. It means that the algorithm can be applied to solve the problem not only restricted to the input graph being a metric graph but also in a graph belonging to a super graph class of metric graphs.

**Algorithm 1** Approximation algorithm for $\Delta_\beta$-WD$k$S $(G, w)$

1: Initially, $C := \emptyset$
2: while $|C| \leq k-2$ do
3: Select $(u, v)$ such that $w(u, v)$ is of maximum weight in $G$;
4: $C := C \cup \{u, v\}$;
5: Remove all edges incident to $u$ or $v$ in $G$;
6: end while
7: if $k$ is odd then
8: Add an arbitrary vertex to $C$.
9: end if
10: return $C$.

**Theorem 3.** For $\beta \geq \frac{1}{2}$, the $\Delta_\beta$-WD$k$S problem can be approximated to within a factor $\frac{1}{2\beta^3}$ in $O(n^2 + k^2 \log k)$ time.

**Proof.** Let $C_k$ be the solution returned by Algorithm 1 for the $\Delta_\beta$-WD$k$S. Let $C_k^*$ be an optimal solution of the $\Delta_\beta$-WD$k$S problem in $G$. Let $e = (u, v)$ be the edge of maximum weight in $G$ and let $G' = G[V \setminus \{u, v\}]$. Let $C_{k-2}$ be the approximation solution on $G'$ returned by Algorithm 1. Assume that $C_{k-2} = C_k \setminus \{u, v\}$. Let $C_{k-2}^*$ be an optimal solution on $G'$. The proof is by induction on $k$. If $k = 2$, we see that

$$w(C_k^*) = w(x, y) \leq w(u, v)$$

(since $w(u, v)$ is of maximum weight in $G$)
$$= w(C_2)$$
$$\leq 2\beta \cdot w(C_2).$$

Thus $\frac{w(C_2)}{w(C_2^*)} \geq \frac{1}{2\beta}$. The theorem is true.

Suppose that $k = 3$. Let $C_3^* = \{x, y, z\}$. We see that
Corollary 2. The approximation ratio $\frac{1}{2\beta}$ of Algorithm 1 is asymptotically tight.

Proof. We give an example to show that the approximation ratio $\frac{1}{2\beta}$ of Algorithm 1 is asymptotically tight. The example can be constructed by the following steps:

1) Construct a graph $G$ of $n = 4h$ vertices, consisting of a left half $G_L$ and a right half $G_R$. Let $k = 2h$.
2) The weights in $G$ are constructed as follows:

$$w(C_3') = w(x, y) + w(y, z) + w(z, x)$$

$$\leq 3 \cdot w(u, v)$$

(since $(u, v)$ is of maximum weight)

$$\leq w(u, v) + 2 \cdot \beta \cdot (w(u, t) + w(t, v))$$

(by $\beta$-triangle inequality)

$$\leq 2 \beta (w(u, v) + w(u, t) + w(t, v))$$

(by $\beta \geq \frac{1}{2}$)

$$= 2 \beta \cdot w(C_3').$$

Thus, $w(C_3') \geq \frac{1}{2\beta}$. The theorem is true for $k \leq 3$.

Suppose that the theorem is true for $k = 2$. Now we prove it for $k$. Notice that $(u, v)$ is a maximum weight edge in $G$. There are three cases.

Case 1: $u, v \in C_k$. Let $e = (u, v)$.
Case 2: $u \in C_k$ and $v \notin C_k$. Arbitrary pick $x \in C_k$ and let $e = (u, x)$.
Case 3: $u, v \notin C_k$. Arbitrary pick $x, y \in C_k$ and let $e = (x, y)$.

Next we prove the ratio $w(C_k) / w(C_k') \geq \frac{1}{2\beta}$.

$$w(C_k) \leq w(e) + 2(k - 2) \cdot w(u, v) + w(C_{k-2})$$

$$\leq w(u, v) + 2(k - 2) \cdot w(u, v) + 2 \beta \cdot w(C_{k-2})$$

(by induction hypothesis)

$$\leq w(u, v) + 2 \sum_{t \in C_{k-2}} \beta \cdot (w(u, t) + w(v, t))$$

$$+ 2 \beta \cdot w(C_{k-2})$$

(by $\beta$-triangle inequality)

$$\leq 2 \beta \cdot (w(u, v) + \sum_{t \in C_{k-2}} (w(u, t) + w(v, t)))$$

$$+ 2 \beta \cdot w(C_{k-2})$$

(by $\beta \geq \frac{1}{2}$)

$$= 2 \beta \cdot w(C_k).$$

Thus, we obtain that $w(C_k) / w(C_k') \geq \frac{1}{2\beta}$. This shows that $\Delta_\beta$-WDS problem can be approximated to within a factor $\frac{1}{2\beta}$.

It is not hard to see that a straightforward implementation of Algorithm 1 is $O(kn^2)$. It was proved in [29] that by applying a linear time selection algorithm [8] and a heap data structure [24], Algorithm 1 can be executed in $O(n^2 + k^2 \log k)$ time. This completes the proof. □

Corollary 2. The approximation ratio $\frac{1}{2\beta}$ of Algorithm 1 is asymptotically tight.

Proof. We give an example to show that the approximation ratio $\frac{1}{2\beta}$ of Algorithm 1 is asymptotically tight. The example can be constructed by the following steps:

(a) Identify a perfect matching of the $2h$ vertices in $G_L$, and give each of the edges of the matching weight $2\beta$. All other edges in $G_L$ have weight 1.
(b) All edges in $G_R$ have weight $2\beta$.
(c) All edges between $G_L$ and $G_R$ have weight 1.

It is not hard to see that $G$ is a $\Delta_\beta$-metric graph. An optimal solution of the $\Delta_\beta$-WDS problem in $G$ can be obtained by selecting all vertices in $G_R$. We have $OPT = \frac{k}{2} \cdot 2\beta$. If Algorithm 1 chooses all vertices of $G_L$ into the solution, the solution returned will be $APX = \frac{k}{2} + \frac{k}{2} \cdot (2\beta - 1)$. This implies

$$\frac{APX}{OPT} = \frac{\frac{k}{2} + \frac{k}{2} \cdot (2\beta - 1)}{\frac{k}{2} \cdot 2\beta}$$

$$= \frac{1}{2\beta} + \frac{2\beta - 1}{2\beta} \cdot \frac{k}{2}$$

$$\leq \frac{1}{2\beta} + \frac{1}{k - 1}$$

$$\approx \frac{1}{2\beta} \quad \text{(since } k = \frac{n}{2}).$$

This shows that the approximation ratio $\frac{1}{2\beta}$ of Algorithm 1 is asymptotically tight even when the edge weights have only two distinct values. □

V. Concluding remarks

In this paper, we prove that for $\beta > \frac{1}{2}$, the $\Delta_\beta$-WDS problem is NP-hard. It implies that for $\frac{1}{2} < \beta < 1$ (subclasses of metric graphs), the $\Delta_\beta$-WDS problem is still NP-hard. We show that a $\frac{1}{4}$-approximation algorithm given for solving the WDS problem in metric graphs can be applied to solve the $\Delta_\beta$-WDS problem for any $\beta > \frac{1}{2}$ and its approximation ratio is $\frac{1}{2\beta}$. It is of interest to see that whether $\Delta_\beta$-WDS problem can be approximated to within a factor better than $\frac{1}{2\beta}$ for any $\beta$, especially for $\beta < 1$. Moreover, it is also of interest to know whether the $\Delta_\beta$-WDS problem has a PTAS. If not, we must show that there exists a function $r(\beta)$ such that to approximate the $\Delta_\beta$-WDS problem to within a factor $r(\beta)$ is NP-hard.

References


