.586 Approximation of Dense-$\frac{n}{2}$-Subgraph *

Yinyu Ye and Jiawei Zhang

December 1999

Abstract

We again consider the DENSE-$\frac{n}{2}$-SUBGRAPH problem, i.e., determine a block of half nodes such that the weight within the subgraph induced by the block is maximized. We prove that a strengthened semidefinite relaxation with a mixed rounding technique yields a .586 approximations for the DENSE-$\frac{n}{2}$-SUBGRAPH problem. The previous best-known results for this problem are .25 using a simple randomization technique, .48 using a semidefinite relaxation, .5 using a linear programming relaxation or another semidefinite relaxation; and .519 using a mixed semidefinite relaxation and the simple randomization technique. In fact, a un-strengthened SDP relaxation provably yields no more than .5 approximation.

Key words. Dense-k-Subgraph, polynomial approximation algorithm, semidefinite programming.

*Department of Management Sciences, Henry B. Tippie College of Business, The University of Iowa, Iowa City, Iowa 52242. Research supported in part by NSF grants DMI-9908077 and DMS-9703490.

1 Introduction

Given an undirected graph $G = (V, E)$ and non-negative weights $w_{ij} = w_{ji}$ on the edges $(i, j) \in E$, the so-called DENSE-k-SUBGRAPH problem (DSP), i.e. determine a subset $S \subset V$ of k nodes such that the weight of the subgraph induced by S is maximized, see [7]. This problem can be formulated as:

$$
\text{(DSP)} \quad
\begin{array}{lll}
w^* := & \text{Maximize} & \frac{1}{4}\sum_{i<j} w_{ij}(1 + x_i + x_j + x_i x_j) \\
& \text{subject to} & \sum_{j=1}^{n} x_j = 2k - n, \\
& & x_j^2 = 1, \ j = 1, \ldots, n,
\end{array}
$$

In this paper, we consider the case of $k = n/2$.

The previous best-known results for this problem are .25 using a simple randomization technique, .48 using a semidefinite relaxation (Srivastav and Wolf [7]), .5 using a linear programming relaxation (Goemans [4]) or another semidefinite relaxation (Feige and Seltser [2]); and .519 using a mixed semidefinite relaxation and the simple randomization technique (Ye and Zhang [8]). In fact, Srivastav and Wolf also proved that the SDP relaxation used in their analysis yields at most .5 approximation of DENSE-$\frac{n}{2}$-SUBGRAPH.

Before present our results, let us introduce some notations. Let $\bar{X}$ be an optimal solution of (SDP). For a given $0 \le \theta \le 1$, let

$$
\alpha(\theta) = \min_{-1/3 < y \le 1} \frac{1 + \frac{6}{\pi}\arcsin(\theta y)}{1 + 3y}, \tag{1}
$$

$$
\beta(\theta) = 1 - \frac{2}{\pi}\arcsin(\theta) - \frac{1 - \frac{2}{\pi}\arcsin(\theta)}{n} + \min_{-1 \le y < 1} \frac{2}{\pi}\frac{\arcsin(\theta) - \arcsin(\theta y)}{1 - y}, \tag{2}
$$

and let θ^* and R^* be the optimizer and optimal value, respectively, to the following:

$$
\max_{0 \le \theta \le 1} R(\theta) := \frac{\alpha(\theta)(2\eta(\theta) - \eta(\theta)^2)}{1 + 4\eta(\theta) - \beta(\theta)(1 + 2\eta(\theta))}
$$

where

$$
\eta(\theta) = \frac{\sqrt{(9 - 5\beta(\theta))(1 - \beta(\theta))} - (1 - \beta(\theta))}{2(2 - \beta(\theta))}. \tag{3}
$$

Then, let $\alpha := \alpha(\theta^*)$, $\beta := \beta(\theta^*)$ and $\eta = \eta(\theta^*)$. Note that

$$
0 \le \alpha \le 1, \quad , 0 \le \beta \le 1, \quad \text{and} \quad 0 \le \eta \le 1/2.
$$

2 The SDP Relaxation of (DSP)

Our SDP relaxation of (DSP) is

$$
\text{(DSDP)} \quad \begin{array}{lll} w_{SD} := & \text{Maximize} & \frac{1}{4}\sum_{0<i<j} w_{ij}(1 + X_{0i} + X_{0j} + X_{ij}) \\ & \text{subject to} & \sum_{j=1}^{n} X_{0j} = 0, \\ & & \sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij} = 0, \\ & & X_{jj} = 1, \quad j = 0, ..., n, \ X \succeq 0. \end{array} \tag{4}
$$

Here, the unknown $X \in R^{(n+1)\times(n+1)}$, row and column indexed from 0 to n. Note that this SDP relaxation is a strengthened version of the SDPs used in Srivastav and Wolf [7], where there is no constraint $\sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij} = 0$; and in Ye and Zhang [8], where there is no constraint $\sum_{j=1}^{n} X_{0j} = 0$.

The rounding algorithm for (DSP) is similar to Ye and Zhang [8]:

1. SDP Solving: Solve the problem (DSDP) to obtain a semidefinite matrix $\bar{X}$.

Repeat the following two steps for $[\epsilon^{-1} \log \epsilon^{-1}]$ times, where $\epsilon > 0$ is small constant as stated in Frieze and Jerrum [3], and output the best subgraph.

2. Randomized Rounding: Generate a vector

$$u \in N(0, \theta^* \bar{X} + (1 - \theta^*)I).$$

Then assign

$$\hat{x} = \text{sign}(u).$$

Select the block $S = \{i > 0 : \hat{x}_i = \hat{x}_0\}$. Let $\tilde{S} = S$.

3. Node Swapping:

• If $|\tilde{S}| > n/2$, then for each $i \in \tilde{S}$, let $\zeta(i) = \sum_{j\in\tilde{S}} w_{ij}$ and $\tilde{S} := \{i_1, i_2, ..., i_{|\tilde{S}|}\}$, where $\zeta(i_1) \geq \zeta(i_2) \geq ... \geq \zeta(i_{|\tilde{S}|})$. Then, remove node $i_{|\tilde{S}|}$ from $\tilde{S}$ and reassign $\tilde{S} := \{i_1, i_2, ..., i_{|\tilde{S}|-1}\}$. Repeat this swapping process till $|\tilde{S}| = n/2$.

• If $|\tilde{S}| \leq n/2$, arbitrarily add $n/2 - |\tilde{S}|$ nodes from outside of $\tilde{S}$ into $\tilde{S}$. ■

For any $U \subset V$, denote the total weights within the subgraph induced by U as $w(U)$, that is,

$$w(U) := \sum_{i<j,\ i\in U,\ j\in U} w_{ij}.$$

Then, the following lemma holds and is due to Srivastav and Wolf [7]:

Lemma 1 *The total weights of edges in the subgraph induced by $\tilde{S}$*

$$w(\tilde{S}) \geq \begin{cases} \frac{\frac{n}{2}(\frac{n}{2}-1)}{|S|(|S|-1)} w(S) & \text{if } |S| > n/2 \\ w(S) & \text{if } |S| \leq n/2 \end{cases}. \tag{5}$$

3 Analysis of the Approximation of (DSP)

In order to analyze the quality of bisection $\tilde{S}$, we consider three random variables:

$$w := w(S) = \frac{1}{4} \sum_{0<i<j} w_{ij}(1 + \hat{x}_0\hat{x}_i + \hat{x}_0\hat{x}_j + \hat{x}_i\hat{x}_j),$$

$$p := n - |S| = \frac{1}{2} \sum_{j>0} (1 - \hat{x}_0\hat{x}_j),$$

and

$$m := |S|(n - |S|) = \frac{n^2}{4} - \frac{(e^T\hat{x})^2}{4} = \frac{1}{4} \sum_{i>0,j>0} (1 - \hat{x}_i\hat{x}_j).$$

Lemma 2 *Our approximation method yields S satisfying the following three inequalities:*

$$E[w] \geq \alpha \cdot w_{SD} \geq \alpha \cdot w^*,$$

$$E[p] \geq \beta \cdot \frac{n}{2},$$

$$E[m] \geq \beta \cdot \frac{n^2}{4}.$$

Proof. The proof of the first and third inequalities is contained in [8]. To prove the second inequality, from Goemans and Williamson [5], Frieze and Jerrum [3], and Bertsimas and Ye [1], we have

$$\mathrm{E}[\hat{x}_0\hat{x}_j] = \frac{2}{\pi} \arcsin(\theta^* \bar{X}_{0j}), \quad j = 1, 2, \ldots, n. \tag{6}$$

Let

$$\alpha' = \min_{-1 \leq y < 1} \frac{2}{\pi} \frac{\arcsin(\theta^*) - \arcsin(\theta^* y)}{1 - y}$$

and noting that

$$\sum_{j>0} \bar{X}_{0j} = 0,$$

we derive

$$\begin{aligned}
\mathrm{E}[n - |S|] &= \frac{1}{2}\sum_{j>0}\left(1 - \frac{2}{\pi}\arcsin(\theta^* \bar{X}_{0j})\right) \\
&= \frac{1}{2}\sum_{j>0}\left(1 - \frac{2}{\pi}\arcsin(\theta^*) + \frac{2}{\pi}\arcsin(\theta^*) - \frac{2}{\pi}\arcsin(\theta^* \bar{X}_{0j})\right) \\
&\geq \frac{1}{2}\sum_{j>0}\left(1 - \frac{2}{\pi}\arcsin(\theta^*) + \alpha'(1 - \bar{X}_{0j})\right) \\
&= \left(1 - \frac{2}{\pi}\arcsin(\theta^*) + \alpha'\right) \cdot \frac{n}{2} \\
&\geq \beta \cdot \frac{n}{2}.
\end{aligned}$$

■

Consider a new random variable

$$z(\eta, \gamma) := \frac{w}{w^*} + 2\eta\gamma\frac{2p}{n} + \gamma\frac{4m}{n^2}, \tag{7}$$

where η is given by (3) and

$$\gamma = \frac{\alpha}{1 + 4\eta - \beta(1 + 2\eta)}. \tag{8}$$

By lemma 2, we have

$$\mathrm{E}[z(\eta, \gamma)] \geq \alpha + 2\eta\gamma\beta + \gamma\beta \quad \text{and} \quad z(\eta, \gamma) \leq 1 + 4\eta\gamma + \gamma(1 - \eta)^2.$$

Now we prove the final lemma:

Lemma 3 *If random variable $z(\eta, \gamma)$ fulfills its expectation, i.e., $z(\eta, \gamma) \geq \alpha + (2\eta + 1)\gamma\beta$, then*

$$w(\tilde{S}) \geq \frac{n-2}{n-1} \cdot R^* \cdot w^*,$$

where recall that

$$R^* = \frac{\alpha(2\eta - \eta^2)}{1 + 4\eta - \beta(1 + 2\eta)}.$$

Proof. Suppose

$$w(S) = \lambda w^* \quad \text{and} \quad |S| = \delta n,$$

which from (7) and $z(\eta, \gamma) \geq \alpha + (2\eta + 1)\gamma\beta$ implies that

$$\begin{aligned} \lambda &\geq \alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma(1-\delta) - 4\gamma\delta(1-\delta) \\ &= \alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma + 4\delta\gamma(\eta - 1) + 4\delta^2\gamma. \end{aligned}$$

Consider the case $1/2 \leq \delta \leq 1$. Applying Lemma 1, we see that

$$\begin{aligned} w(\tilde{S}) &\geq \frac{\frac{n}{2}(\frac{n}{2} - 1)}{\delta n(\delta n - 1)} \cdot w(S) \\ &= \frac{n-2}{n - \frac{1}{\delta}} \cdot \frac{1}{4\delta^2} \cdot w(S) \\ &\geq \frac{n-2}{n-1} \cdot \frac{1}{4\delta^2} \cdot w(S) \\ &= \frac{n-2}{n-1} \cdot \frac{\lambda}{4\delta^2} \cdot w^* \\ &\geq \frac{n-2}{n-1} \cdot \frac{\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma + 4\delta\gamma(\eta - 1) + 4\delta^2\gamma}{4\delta^2} \cdot w^* \\ &= \frac{n-2}{n-1} \cdot \left(\frac{\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma}{4\delta^2} + \frac{\gamma(\eta - 1)}{\delta} + \gamma \right) \cdot w^* \\ &\geq \frac{n-2}{n-1} \cdot \left(\gamma - \frac{\gamma^2(1-\eta)^2}{\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma} \right) \cdot w^* \\ &:= \frac{n-2}{n-1} \cdot R_1 \cdot w^*. \end{aligned}$$

The last inequality follows from simple calculus that

$$\delta = \frac{\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma}{2\gamma(1-\eta)} = \frac{1}{2(1-\eta)}$$

(the second equality above is from (8) that $\gamma = \alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma$), yields the minimal value for

$$\frac{\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma}{4\delta^2} + \frac{\gamma(\eta - 1)}{\delta} + \gamma$$

in the interval $[1/2, 1]$.

Consider the case $0 \leq \delta \leq 1/2$. Applying Lemma 1, we see that

$$w(\tilde{S}) \geq w(S)$$

$$
\begin{aligned}
&\geq \quad (\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma + 4\delta\gamma(\eta - 1) + 4\delta^2\gamma) \cdot w^* \\
&\geq \quad \left(\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma - \gamma(1 - \eta)^2\right) \cdot w^* \\
&:= \quad R_2 \cdot w^*.
\end{aligned}
$$

The last inequality follows from simple calculus that

$$\delta = \frac{1 - \eta}{2}$$

yields the minimal value for

$$\alpha + (2\eta + 1)\gamma\beta - 4\eta\gamma + 4\delta\gamma(\eta - 1) + 4\delta^2\gamma$$

in the interval $[0, 1/2]$.

Finally, our choice for γ in (8) makes

$$R_1 = R_2 = R^*,$$

which proves the lemma. ■

Note that η given by (3) maximizes

$$\frac{\alpha(2\eta - \eta^2)}{1 + 4\eta - \beta(1 + 2\eta)}$$

for any given $0 \leq \beta \leq 1$ and $\alpha > 0$.

Finally, we have the main result:

Theorem 1 *The worst case performance ratio of the approximation algorithm for the DENSE-$n/2$-SUBGRAPH is at least $.586$ for sufficiently large n.*

Proof. For sufficiently large n, we verified that for $\theta = .89$, $\alpha(\theta) \geq .7368$ and $\beta(\theta) \geq .9621$, therefore, $\eta(\theta) = .1737$ and

$$R^* = \frac{\alpha(2\eta - \eta^2)}{1 + 4\eta - \beta(1 + 2\eta)} \geq .5866.$$

Following the rest of proof of Frieze and Jerrum [3], we have the theorem proved. ■

References

[1] D. Bertsimas and Y. Ye, "Semidefinite relaxations, multivariate normal distributions, and order statistics," *Handbook of Combinatorial Optimization (Vol. 3)*, D.-Z. Du and P.M. Pardalos (Eds.) pp. 1-19, (1998 Kluwer Academic Publishers).

[2] U. Feige and M. Seltser, "On the densest k-subgraph problem," Technical report, Department of Applied Mathematics and Computer Science, The Weizmann Institute, Rehovot, September 1997.

[3] A. Frieze and M. Jerrum, "Improved approximation algorithms for max k-cut and max bisection," *Proc. 4th IPCO Conference* (1995) 1-13.

[4] M. X. Goemans, "Mathematical programming and approximation algorithms," Lecture given at the Summer School on Approximate Solution of Hard Combinatorial Problems, Udine, September 1996.

[5] M. X. Goemans and D. P. Williamson, " Improved approximation algorithms for Maximum Cut and Satisfiability problems using semidefinite programming," *Journal of ACM* 42 (1995) 1115–1145.

[6] Y. E. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization Methods and Software* 9 (1998) 141-160.

[7] A. Srivastav and K. Wolf, "Finding dense subgraphs with semidefinite programming," in *Approximation Algorithms for Combinatorial Optimization*, K. Jansen and J. Rolim (Eds.) (1998) pp. 181-191.

[8] Y. Ye and J. Zhang, ".602 Approximation of the Complement of Min-Bisection and .519 Approximation of Dense-$\frac{n}{2}$-Subgraph," *Manuscript*, November 1999.

[9] U. Zwick, "Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to Max-Cut and other problems," in *31th STOC* (1999) 679-687.