Optim Lett (2017) 11:419–427
DOI 10.1007/s11590-015-0938-8

ORIGINAL PAPER

On positive-influence target-domination

Guangmo Tong[1] · Weili Wu[1] · Panos M. Pardalos[2,3] · Ding-Zhu Du[1,4]

Received: 27 March 2015 / Accepted: 22 August 2015 / Published online: 2 September 2015

Abstract Consider a graph $G = (V, E)$ and a vertex subset $A \subseteq V$. A vertex v is positive-influence dominated by A if either v is in A or at least half the number of neighbors of v belong to A. For a target vertex subset $S \subseteq V$, a vertex subset A is a positive-influence target-dominating set *for target set* S if every vertex in S is positive-influence dominated by A. Given a graph G and a target vertex subset S, the positive-influence target-dominating set (PITD) problem is to find the minimum positive-influence dominating set for target S. In this paper, we show two results: (1) The PITD problem has a polynomial-time $(1 + \log\lceil\frac{3}{2}\Delta\rceil)$-approximation in general graphs where Δ is the maximum vertex-degree of the input graph. (2) For target set S with $|S| = \Omega(|V|)$, the PITD problem has a polynomial-time $O(1)$-approximation in power-law graphs.

✉ Ding-Zhu Du
dingzhu.du@tdt.edu.vn; dzdu@utdallas.edu

Guangmo Tong
gxt140030@utdallas.edu

Weili Wu
weiliwu@utdallas.edu

Panos M. Pardalos
pardalos@ufl.edu

1 Department of Computer Science, University of Texas at Dallas, Richardson, TX 75080, USA

2 Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA

3 Laboratory of Algorithms and Technologies for Networks Analysis (LATNA), National Research University, Higher School of Economics, 20 Myasnitskaya st., Moscow 101000, Russia

4 Division of Algorithms and Technologies for Networks Analysis, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Keywords Positive-influence · Target-dominating · Social networks

1 Introduction

Consider a graph $G = (V, E)$ and a vertex subset $A \subseteq V$. A vertex v is *positive-influence dominated* by A if either v is in A or at least half the number of neighbors of v belong to A. The positive-influence dominating has been studied extensively in the literature [6,12,14] due to its applications in social networks.

A positive comment produces a positive-influence and a negative comment generates a negative-influence. Every person is influenced by his friends positively or negatively, and tends to adopt a behavior with more positive-influence. For example, more children begin smoking when parents are smoking [10]. This is the background of the concept of positive-influence dominating.

Suppose a company wants to advertise its product by sending out some free samples for generating positive-influence for dominating a group of target members in a social networks. For example, if the product is woman's clothing, then the target group should consist of only women. To dominate all target members by positive-influence, how can samples used be distributed to minimize the total number of free samples? This viral marketing problem can be formulated as the positive-influence target-dominating set (PITD) problem in the following way: given a graph G and a subset S of vertices, find the minimum vertex subset A such that every vertex in S is positive-influence dominated by A. Each vertex in S is called a *target* and S is called a target set. In this paper, we show two results:

1. The PITD problem has a polynomial-time $[\log \Delta + O(1)]$-approximation in general graph G where Δ is the maximum vertex-degree of G.
2. For target set S with $|S| = \Omega(|V|)$, the PITD problem has a polynomial-time $O(1)$-approximation in power-law graphs.

2 Preliminaries

Consider a finite set X and a function $f : 2^X \to \Re$. Function f is called a *submodular* function if for any two subsets A and B of X,

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B).$$

Function f is *monotone nondecreasing* if

$$A \subset B \Rightarrow f(A) \leq f(B).$$

A monotone nondecreasing submodular function f with $f(\emptyset) = 0$ is called a *polymatroid* function. The followings are well-known properties of submodular and monotone nondecreasing functions and polymatroid functions, which can be found in [8,11,13].

Lemma 1 *Function f is monotone nondecreasing submodular if and only if for any element $x \in X$, and two subsets A and B with $A \subset B$,*

$$\Delta_x f(A) \geq \Delta_x f(B)$$

where $\Delta_x f(A) = f(A \cup \{x\}) - f(A)$.

Lemma 2 *Suppose f is a monotone nondecreasing submodular function. Then for any constant c, $\min(c, f)$ is monotone nondecreasing submodular. Moreover, if f is a polymatroid function and $c > 0$ is a constant, then $\min(c, f)$ is also a polymatroid function.*

Lemma 3 *If f and g are two polymatroid functions, then $f + g$ is also polymatroid.*

Consider a polymatroid function f on 2^X and a nonnegative cost function $c : X \to \Re_+$. Define $c(A) = \sum_{v \in A} c(v)$. The following is called the *submodular cover problem*:

$$\begin{aligned} \min \quad & c(A) \\ \text{subject to} \quad & A \in \Omega(f) \end{aligned}$$

where $\Omega(f) = \{A \mid f(A) = f(X)\}$. There is a greedy approximation algorithm with a Theorem on its performance for the submodular cover problem which can found in [8,13].

Greedy Algorithm SC
$A \leftarrow \emptyset$;
while $f(A) < f(X)$ *do*
    choose $x = \text{argmax}_{x \in X} \frac{\Delta_x f(A)}{c(x)}$
    and set $A \leftarrow A \cup \{x\}$;
output A.

Theorem 4 *If f is a polymatroid integer function on 2^X, then Greedy Algorithm SC produces an approximation solution for the submodular cover problem within a factor of $1 + \ln \gamma$ from the optimal where $\gamma = \max_{x \in X} f(\{x\})$.*

3 In general graphs

Consider a graph $G = (V, E)$ and a set of targets, $S \subseteq V$. For any $v \in V$, denote $deg_A(v) = |\{(u, v) \mid u \in A \text{ or } v \in A\}|$. For any vertex subset A, define

$$f_S(A) = \sum_{v \in S} \min\left(\left\lceil \frac{1}{2} \cdot deg(v) \right\rceil, deg_A(v)\right).$$

Then we have the following.

Lemma 5 *For any given subset of vertices S, f_S is a polymatroid function.*

Proof We claim that $g(A) = deg_A(v)$ is a polymatroid function for any $v \in V$. Note that $g(\emptyset) = 0$. By Lemma 1, it is sufficient to show that for any vertex $x \in V$ and two vertex subsets A and B with $A \subset B$, $\Delta_x g(A) \geq \Delta_x g(B)$. We divide the proof into three cases.

Case 1. $v \in A$. Then $g(A) = g(A \cup \{x\}) = g(B) = g(B \cup \{x\}) = deg(v)$. Hence, $\Delta_x g(A) = \Delta_x g(B) = 0$.

Case 2. $v \in B \setminus A$. Then $g(B) = g(B \cup \{x\}) = deg(v)$. Hence, $\Delta_x g(A) \geq \Delta_x g(B) = 0$.

Case 3. $v \notin B$. If $v \neq x$, then we have

$$\begin{aligned}\Delta_x g(A) &= g(A \cup \{x\}) - g(A)\\ &= |(A \cup \{x\}) \cap N(v)| - |A \cap N(v)|\\ &\geq |(B \cup \{x\}) \cap N(v)| - |B \cap N(v)|\\ &= \Delta_x g(B),\end{aligned}$$

where $N(v)$ is the set of neighbors of v. If $v = x$, then we have

$$\begin{aligned}\Delta_x g(A) &= g(A \cup \{x\}) - g(A)\\ &= deg(v) - |A \cap N(v)|\\ &\geq deg(v) - |B \cap N(v)|\\ &= \Delta_x g(B).\end{aligned}$$

Now, by our claim and Lemmas 2 and 3, f_S is also a polymatroid function. □

Lemma 6 *A vertex subset A is a positive-influence dominating set for target S if and only if $f_S(A) = f_S(V)$.*

Proof Note that v is positive-influence dominated by A if and only if $\lceil \frac{1}{2} \cdot deg(v) \rceil \leq deg_A(v)$. Now, the lemma follows immediately from this fact. □

Theorem 7 *Let a target set S be given. Then there exists a polynomial-time greedy algorithm which produces an approximation solution for the PITD problem, within a factor of $1 + \ln\lceil \frac{3}{2}\Delta \rceil$ from the optimal where Δ is the maximum vertex degree of graph G.*

Proof By Lemma 6, $\Omega(f_S)$ is the set of positive-influence dominating sets for target S. Thus, finding the minimum positive-influence dominating set for target S is the submodular cover problem with polymatroid function f_S and constant cost function $c(u) = 1$ for every vertex u.

By Lemmas 5, 6 and Theorem 4, Greedy Algorithm SC produces an approximation solution within a factor of $1 + \ln \gamma$ from the optimal where $\gamma = \max_{u \in V} f_S(\{u\})$. Note that

$$f_S(\{u\}) = \sum_{v \in S} \min\left(\left\lceil \frac{1}{2} \cdot deg(v) \right\rceil, deg_{\{u\}}(v) \right).$$

In this sum, when $v = u$, the corresponding term is equal to $\lceil \frac{1}{2} \cdot deg(u) \rceil$ since by our definition of $deg_A(v)$, $deg_{\{u\}}(u) = deg(u)$. When $v \neq u$ and $(v, u) \in E$, the coresponding term is equal to 1. When $v \neq u$ and $(v, u) \notin E$, the corresponding term is equal to 0. Therefore,

$$\begin{aligned}
\gamma &= \max_{u \in V} f_S(\{u\}) \\
&\leq \max_{u \in V} \left(deg(u) + \left\lceil \frac{1}{2} deg(u) \right\rceil \right) \\
&= \max_{u \in V} \left\lceil \frac{3}{2} deg(u) \right\rceil \\
&\leq \left\lceil \frac{3}{2} \Delta \right\rceil .
\end{aligned}$$

□

4 In power law graphs

Next, we move our attention to power law graphs. A graph G is said to belong to class $C(\alpha, \gamma)$ if G has no isolated vertex and, for $k \geq 1$, the number of vertices with degree k is $\lfloor \frac{e^\alpha}{k^\gamma} \rfloor$. Clearly, the maximum vertex degree $\Delta = \lfloor e^{\alpha/\gamma} \rfloor$. For the number n of vertices and the number m of edges, we have

$$n = \sum_{k=1}^{\Delta} \left\lfloor \frac{e^\alpha}{k^\gamma} \right\rfloor \leq \sum_{k=1}^{\Delta} \frac{e^\alpha}{k^\gamma} < 2n,$$

and

$$m = \frac{1}{2} \cdot \sum_{k=1}^{\Delta} k \left\lfloor \frac{e^\alpha}{k^\gamma} \right\rfloor \leq \frac{1}{2} \cdot \sum_{k=1}^{\Delta} \frac{e^\alpha}{k^{\gamma-1}} < 2m.$$

Usually, an online social network varies as time changes, which can be formulated by a family of power law graphs with a fixed γ. In this family, m and n are varied when α is varied. Our study will emphase on a family of power law graphs with fixed $\gamma > 2$ because many social networks are such a family of power law graphs [1–5,9].

Let us first show some properties for such a family of power law graphs $C(\alpha, \gamma)$.

Lemma 8 *For any power law graph $G \in C(\alpha, \gamma)$ with fixed $\gamma > 2$, $n \geq c_1 \cdot m$ where c_1 is a constant depending on only γ.*

Proof It suffices to show that n/m is greater than a positive constant. Let $\zeta(\gamma)$ be the Riemann Zeta function, i.e., $\zeta(\gamma) = \sum_{k=1}^{\infty} \frac{1}{k^\gamma}$. Then

$$n/m \geq \frac{0.5 \sum_{k=1}^{\Delta} \frac{1}{k^\gamma}}{0.5 \sum_{k=1}^{\Delta} \frac{1}{k^{\gamma-1}}} > \frac{\zeta(\gamma)}{\zeta(\gamma - 1)}.$$

□

For any vertex subset A, denote $deg(A) = \sum_{v \in A} deg(v)$.

Lemma 9 *Suppose G is a graph without isolated vertex and S is a target set. Let A be a positive-influence target-dominating set in G. Then* $deg(A) \geq 0.5 \cdot |S|$.

Proof Note that at least $|S|$ vertices are positive-influence dominated by A and each vertex has degree at least one. Thus, $deg(S) \geq |S|$. Since A is a positive-influence target-dominating set for target set S, each vertex $v \in S$ has at least half the number of neighbors in A. Hence, we have $deg(A) \geq 0.5deg(S) \geq 0.5|S|$. □

Lemma 10 *For any constant* $c > 0$ *and* $\gamma > 2$*, there exists a constant* $c_2 > 0$*, which depends on only c and* γ*, such that for any graph G in class* $C(\alpha, \gamma)$ *with fixed* $\gamma > 2$ *and any vertex subset A,*

$$deg(A) \geq cm \Rightarrow |A| \geq c_2 n.$$

Proof Let k_0 be the largest integer such that

$$|A| \leq \sum_{k=k_0}^{\Delta} \left\lfloor \frac{e^{\alpha}}{k^{\gamma}} \right\rfloor.$$

Then

$$cm \leq deg(A) \leq \sum_{k=k_0}^{\Delta} \frac{e^{\alpha}}{k^{\gamma-1}}. \tag{1}$$

Since $\sum_{k=1}^{\Delta} \frac{1}{k^{\gamma-1}} \to \zeta(\gamma-1)$ as $\Delta \to \infty$, there exists $\Delta_0 > 0$ such that for $\Delta > \Delta_0$, $\sum_{k=1}^{\Delta} \frac{1}{k^{\gamma-1}} \geq 0.5\zeta(\gamma-1)$, where Δ_0 depends on only γ. Note that for $1 \leq k \leq \Delta$, $\lfloor \frac{e^{\alpha}}{k^{\gamma}} \rfloor \geq 1$. Therefore, for $1 \leq k \leq \Delta$, $\lfloor \frac{e^{\alpha}}{k^{\gamma}} \rfloor > 0.5 \cdot \frac{e^{\alpha}}{k^{\gamma}}$. Thus, for $\Delta > \Delta_0$,

$$\begin{aligned} m &= \frac{1}{2} \sum_{k=1}^{\Delta} k \cdot \left\lfloor \frac{e^{\alpha}}{k^{\gamma}} \right\rfloor \\ &> \frac{1}{4} \sum_{k=1}^{\Delta} \frac{e^{\alpha}}{k^{\gamma-1}} \\ &\geq \frac{1}{8} \zeta(\gamma-1) e^{\alpha}. \end{aligned}$$

Choose $k_1 > 0$ such that

$$\sum_{k=k_1}^{\infty} \frac{1}{k^{\gamma-1}} < c\zeta(\gamma-1)/8.$$

Then k_1 depends on only c and γ. Moreover, for $\Delta > \Delta_0$,

$$cm > \sum_{k=k_1}^{\infty} \frac{e^{\alpha}}{k^{\gamma-1}} > \sum_{k=k_1}^{\Delta} \frac{e^{\alpha}}{k^{\gamma-1}}. \tag{2}$$

Comparing (2) with (1), we obtain $k_0 < k_1$. It follows that

$$|A| > \sum_{k=k_1}^{\Delta} \left\lfloor \frac{e^{\alpha}}{k^{\gamma}} \right\rfloor \geq \frac{1}{2} \sum_{k=k_1}^{\Delta} \frac{e^{\alpha}}{k^{\gamma}} \geq \frac{n}{2} \cdot \frac{\sum_{k=k_1}^{\Delta} \frac{1}{k^{\gamma}}}{\sum_{k=1}^{\Delta} \frac{1}{k^{\gamma}}}.$$

Since

$$\frac{\sum_{k=k_1}^{\Delta} \frac{1}{k^{\gamma}}}{\sum_{k=1}^{\Delta} \frac{1}{k^{\gamma}}} \to \frac{\sum_{k=k_1}^{\infty} \frac{1}{k^{\gamma}}}{\zeta(\gamma)} \quad \text{as } \Delta \to \infty,$$

we can choose $\Delta_1 \geq \Delta_0$ such that for $\Delta > \Delta_1$,

$$\frac{\sum_{k=k_1}^{\Delta} \frac{1}{k^{\gamma}}}{\sum_{k=1}^{\Delta} \frac{1}{k^{\gamma}}} \geq \beta_1 = 0.5 \frac{\sum_{k=k_1}^{\infty} \frac{1}{k^{\gamma-1}}}{\zeta(\gamma)}.$$

Here, Δ_1 depends on k_1 and γ, and hence depends on only c and γ. Therefore, for $\Delta > \Delta_1$,

$$|A| > n(\beta_1/2).$$

For $\Delta \leq \Delta_1$, we note $\Delta = e^{\lfloor \alpha/\gamma \rfloor}$ and hence $e^{\alpha} \leq (\Delta_1 + 1)^{\gamma}$. Therefore,

$$n \leq \sum_{k=1}^{\Delta} \frac{e^{\alpha}}{k^{\gamma}} \leq \beta_2 = \sum_{k=1}^{\Delta_1} \frac{(\Delta_1 + 1)^{\gamma}}{k^{\gamma}}.$$

Since $|A| \geq 1$, we have that for $\Delta \leq \Delta_1$,

$$|A| \geq n \cdot \frac{1}{\beta_2}.$$

Set $c_2 = \min(\beta_1/2, 1/\beta_2)$. Then

$$|A| \geq c_2 n.$$

Since β_1 and β_2 depend on only c and γ, c_2 depends on only c and γ. □

Theorem 11 *For any family of power-lar graphs $G = (V, E)$ in $C(\alpha, \gamma)$ with fixed $\gamma > 2$ and target set S with $|S| \geq \mu n$ for a constant $\mu > 0$, there exists a polynomial-time approximation algorithm for PITD problem, with constant performance ratio depending on only μ and γ.*

Proof Suppose A is a positive-influence target-dominating set for target set S. By Lemma 9, $deg(A) \geq 0.5|S|$. Since $|S| \geq \mu n$ and $n \geq c_1 m$ by Lemma 8, we have $deg(A) \geq 0.5|S| \geq 0.5\mu n \geq 0.5\mu c_1 \cdot m$. By Lemma 10, $|A| \geq c_2 n$ where c_2 depends on $0.5\mu c_1$ and γ, i.e., depends on μ and γ. Since $|A| \leq n$, the ratio of sizes of any two positive-influence target-dominating sets is upper bounded by $1/c_2$. This means that any positive-influence target-dominating set for target S is a $(1/c_2)$-approximation solution. □

Given a graph $G = (V, E)$ and a fraction r with $0 < r \leq 1$, the positive-influence partial-dominating set (PIPD) problem is to find the minimum vertex subset A such that at least a portion r of vertices are positive-influence dominated by A. With above approach, we can also show the following.

Theorem 12 *For any family of powr-law graphs G in class $C(\alpha, \gamma)$ with fixed $\gamma > 2$, the PIPD problem has a polynomial-time approximation with a constant performance ratio depending on only γ and r.*

Proof Suppose that at least rn vertices are positive-influence dominated by A and each vertex has degree at least one. Let B denote the set of vertices positive-influence dominated by A. Then $|B| \geq rn$ and $deg(B) \geq rn$. Since each vertex $v \in B$ has at least a half number of neighbors in A, we have $deg(A) \geq 0.5rn$. By Lemma 8, $deg(A) \geq 0.5rn \geq 0.5rc_1 \cdot m$. By Lemma 10, $|A| \geq c_2 n$, where c_2 depends on only r and γ. This means that every feasible solution for the PIPD problem is a $(1/c_2)$-approximation. □

5 Discussion

Note that the positive-influence dominating set (PIDS) problem in [6,7] is exactly the PITD problem in case $S = V$ and the PIPD problem in case $r = 1$. Therefore, all lower bound results in [6,7] can be extended to the PITD problem and the PIPD problem. Thus, we have the following.

Theorem 13 (a) *In general graphs, the PITD (and the PIPD) problem has no polynomial-time $(0.5 - \varepsilon)\ln n$-approximation for any $\varepsilon > 0$ unless $NP \subseteq DTIME(n^{O(\log\log n)})$.*
(b) *In power-law graphs, both the PITD and the PIPD problems are NP-hard.*

This lower-bound indicates that the result in Theorem 7 is almost the best possible. However, it is an open problem whether there exists a PTAS for the PITD (or the PIPD) problem.

Acknowledgements Authors wish to thank referees for their insightful comments.

References

1. Aiello, W., Chung, F., Lu, L.: A Random Graph Model for Massive Graphs. In: Proceedings of the thirty-second annual ACM symposium on Theory of computing (STOC'2000), pp. 171–180. ACM press, New York, NY, USA (2000)
2. Aiello, W., Chung, F., Lu, L.: Random Evolution in Massive Graphs. In Handbook of Massive Data Sets. Kluwer Academic Publishers, Boston (2001)
3. Barabasi, A., Albert, R., Jeong, H.: Scale-free characteristics of random networks: the topology of the world-wide web. Phys. A **281**, 69–77 (2000)

4. Barabasi, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. Phys. A Stat. Mech. Appl. **311**(3–4), 590–614 (2002)
5. Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. SIAM Rev. **51**, 661–703 (2007)
6. Dinh, T.N., Shen, Y., Nguyen, D.T., Thai, M.T.: On the approximability of positive influence dominating set in social networks. J. Combin. Optim. (JOCO) (2012). doi:10.1007/s10878-012-9530-7
7. Dinh, T.N., Nguyen, D.T., Thai, M.T.: Cheap, easy, and massively effective viral marketing in social networks: truth or fiction? In: Proceedings of ACM Conference on Hypertext and Social Media (Hypertext) (2012)
8. Ding-Zhu, Du, Ko, Ker-I, Hu, X.: Design and Analysis of Approximation Algorithms. Springer, Berlin (2011)
9. Ferrante, A., Pandurangan, G., Park, K.: On the Hardness of Optimization in Power Law Graphs. In: COCOON 2007. Lecture Notes in Computer Science, vol. 4598, pp. 417–427. Springer, Berlin (2007). doi:10.1007/978-3-540-73545-8_41
10. Hill, K.G., Hawkins, J.D., Catalano, R.F., Abbott, R.D., Guo, J.: Family influences on the risk of daily smoking initiation. J. Adolesc. Health **37**(3), 202–210 (2005). http://www.sciencedirect.com/science/article/B6T80-4GWPPTF-5/2/3548286c9c2228ce3c7afaa4db17830c
11. Schrijver, A.: Combinatorial Optimization. Springer, Berlin. ISBN 3-540-44389-4 (2003)
12. Wang, F., Camacho, E., Xu, K.: Positive influence dominating set in online social networks. In: COCOA '09, pp. 313–321. Springer, Berlin (2009)
13. Wolsey, L.A.: An analysis of the greedy algorithm for submodular set covering problem. Combinatorica **2**(4), 385–393 (1982)
14. Zhang, H., Dinh, T.N., Thai, M.T.: Maximizing the spread of positive influence in online social networks. In Proceedings of the IEEE Int Conference on Distributed Computing Systems (ICDCS) (2013)