Positive influence dominating sets in power-law graphs

**Wei Zhang · Weili Wu · Feng Wang ·
Kuai Xu**

Abstract Finding the minimum Positive Influence Dominating Set (PIDS) is a problem arisen from the social network applications. The problem has been studied on general random graphs. However, the social networks is presented more precisely by power-law graphs. One of the most important properties of social networks is the power-law degree distribution. In this paper, we focus on the PIDS problem in power-law graphs and prove that the greedy algorithm has a constant approximation ratio. Simulation results also demonstrate that greedy algorithm can effectively select a small scale PIDS set.

W. Zhang (✉)
Faculty of Science, Xi'an Jiaotong University,
Xi'an 710049, China
e-mail: dc.zhangwei@gmail.com

W. Wu
Department of Computer Science, University of Texas at Dallas,
Richardson, TX 75080, USA
e-mail: weiliwu@utdallas.edu

F. Wang · K. Xu
Division of Mathematical and Natural Sciences,
Arizona State University, PO Box 37100,
Phoenix, AZ 85069, USA
e-mail: fwang25@asu.edu

K. Xu
e-mail: kuai.xu@asu.edu

1 Introduction

An online social network is a network composed of individuals who share the same interest and purpose which provides a powerful medium of communicating, sharing and disseminating information, and spreading influence beyond the traditional social interactions within a traditional social network setting. Online social networks have developed significantly in recent years. For example, online social network sites like Facebook, MySpace are among the most popular sites on the Internet. In fact, Facebook has surpassed Google to be the most popular website in the world in March 2010 (Dougherty 2010); online social networks have also raised special interest among commercial businesses, medical and pharmaceutical companies as a channel to influence the opinion of their customers (Bhattacharyya et al. 2011; Kayaalp et al. 2011; Rosen et al. 2011; Saravanan et al. 2011).

Most recent research has been focused on the understanding of the structure, traffic and properties of online social networks (Alan et al. 2007; Nazir et al. 2008; Aris et al. 2008; Schneider et al. 2009). Another major research focus is on the security and privacy in online social networks (Gross and Acquisti 2005; Baden et al. 2009). There are also studies about effectively utilizing online social networks to spread ideas and information within a group (David et al. 2003; Shishir Bharathi and Kempe 2007; Cha et al. 2009). In Wang et al. (2009), Positive Influence Dominating Set (PIDS) problem is introduced as follows: given a graph $G = (V, E)$, a PIDS is a subset D of V such that any node v in V is dominated by at least $\lceil \frac{d(v)}{2} \rceil$ nodes (that is, v has at least $\lceil \frac{d(v)}{2} \rceil$ neighbors) in D where $d(v)$ is the degree of node v. There are two requirements for PIDS: Firstly, every node not in D has at least half of its neighbors

in D, secondly every node in D also has at least half of its neighbors in D. Finding PIDS has important applications in social networks. For example, for college drinking problem, an intervention program needs to be launched to convert the drinker to abstainer. Due to the limitation of education resources, instead of choosing every binge student to participate in the intervention program, choose a PIDS of all the binge student to participate. Since each binge student will have more abstainer friends than binge friends, the binge student has a better chance to convert to abstainer. Another example can be the spread an opinion throughout a social network considering the bidirectional influence between two person and the possibility of either positive or negative impact a person can give to others. If a PIDS of all the people is convinced about the information, it is highly possible that the information is spread over the whole network since every person has more friends agreeing with the opinion. Wang et al. (2009) proposed a dominating set augment based algorithm to construct PIDS from a given network. Wang et al. (2011) proved the difficulty of PIDS problem in random graph and proposed a greedy algorithm with an approximation ratio of $H(\delta)$ where H is the harmonic function and δ is the maximum vertex degree of the graph representing a social network.

Many properties of the social networks have been revealed so far. One of the most important among them is the power-law distribution of the nodes degrees (Clauset et al. 2009), which means that for a positive integer i, let n_i be the number of nodes with degree i, then n_i is proportional to $i^{-\beta}$, where $\beta > 0$ varies for different application domain. To better describe this property of the social networks, Aiello et al. (2001) proposed a random graph model with given degree sequence which satisfies power-law distribution. Consider a graph G with two parameters α and β. Still we set n_i to be the number of nodes of degree i, then let $n_i = \lfloor e^\alpha/i^\beta \rfloor$ for $i = 1, 2 \ldots D$. Here D represents the maximum degree in the graph. The two parameters α and β roughly delineate the size and density of the graph respectively. We apply this model in our research of the PIDS problem. According to (Clauset et al. 2009), β varies from 2 to 3 in most of the real-life networks, so here we assume $\beta \in (2, 3)$. Cause that $n_i = e^\alpha/i^\beta \geq 1$, we get $D \leq e^{\alpha/\beta}$, here we set $D = e^{\alpha/\beta}$.

In this paper, we focus on PIDS problem in power-law graphs and show that the greedy algorithm has constant approximation ratio in power-law graphs. The constant approximation ratio is achieved by considering the node degree distribution in power-law graphs. When $\beta \sim 2.8$, the size of the PIDS constructed with the greedy algorithm is guaranteed to be at most 1.73 times the size of the optimal solution. We also carry out simulations over six social network topologies collected from a game application hosted by Facebook. We found that greedy algorithm

can effectively choose around 30% of all nodes into PIDS while the average ratio of positive neighbors over all neighbors is above 78%, which is significantly higher than the targeted 50%. This means the strategically chosen PIDS can effectively influence the whole network.

The rest of this paper is organized as follows. Section 2 describes the related work. Section 3 gives the greedy algorithm for PIDS problem. In Sect. 4, constant approximation ratio is achieved for the greedy algorithm in power-law graphs. Section 5 shows the simulation result of the greedy algorithm on facebook data. Section 6 concludes this paper and discusses our future plans.

2 Related work

Most of the current research in online social networks fall in three categories: one is to understand the properties and characteristics of online social networks, such as the work in Alan et al. 2007, Nazir et al. 2008, Aris et al. 2008, Schneider et al. 2009. Another major research focus is on the security and privacy in online social networks (Gross and Acquisti 2005; Baden et al. 2009). There are also studies about effectively utilizing online social networks to spread ideas and information within a group (David et al. 2003; Shishir Bharathi and Kempe 2007; Cha et al. 2009). Our work focuses on exploring how to utilize online social networks to help alleviate social problems in the physical world. The social problem is different from spreading ideas and information. The spread of ideas and information is uni-direction in that once a person in influence to adopt an idea or learn some information, he cannot revert to his original state by means of future influence. The positive or negative influence in social problems can flow in two directions, that is a positive individual can convert to a negative individual then can move back and forth between these two states multiple times. Another difference between our work and that of David et al. (2003), Shishir Bharathi and Kempe (2007) is that we find a set of individuals that guarantees the positive effect of education will be injected into the entire group. They focus on finding a subset of a pre-established size that maximize the spread of information, but not the subset regardless of size that infuses information to the whole group. Wang et al. (2009, 2011) are the two papers most closely to this paper. Wang et al. (2009) is the first one proposing the positive influence dominating set problem under the context of college drinking problem. The main idea of the proposed algorithm is to iteratively find 1-dominating set to augment PIDS set until all nodes have more positive neighbors than negative ones. Wang et al. (2011) proved the hardness of positive influence dominating set problem in random graph as APX-hard by providing an L-reduction to the vertex cover

problem in cubic graph and proposed a greedy algorithm with approximation ratio of $H(\delta)$ where H is the harmonic function and δ is the maximum vertex degree of the graph representing a social network. In this paper, we evaluate the performance of the greedy algorithm with simulations and prove its approximation ratio for power-law graphs.

3 Greedy algorithm

This algorithm has been proposed in Wang et al. (2011). We put it here to keep the completeness of the paper.

Consider a graph $G = (V, E)$, for any vertex subset $A \subseteq V$, let $n_A(v)$ denote the number of neighbors of v in A. For a vertex v with degree $deg(v)$ in G, we denote $h(v) = \lceil deg(v)/2 \rceil$. Now, define

$$f(A) = \sum_{v \in V} \min(h(v), n_A(v)).$$

Greedy Algorithm
$A \leftarrow \emptyset$;
while $f(A) < \sum_{v \in V} h(v)$ **do**
    choose $u \in V - A$ to maximize $f(A \cup \{u\})$
    and set $A \leftarrow A \cup \{u\}$;
output A.

4 Theoretical analysis

In this section, we analyze the performance of the greedy algorithm for the PIDS problem in power-law graphs. Firstly, we investigate the performance ratio of the greedy algorithm for a class of problems known as submodular-cover problems in power-law graphs. Then we present the approximation ratio of the greedy algorithm for the PIDS problem in power-law graphs.

4.1 Submodular-cover problems in power-law graphs

Many combinatorial problems can be formulated as the submodular-cover problem, such as the dominating set problem, the set-cover problem, and the hitting set problem. We will show later that the PIDS problem is also a submodular-cover problem. Firstly we give the formulation of the submodular-cover problem. Given a finite set I and a real function $g(\cdot)$ defined on the power set 2^I, a function $g(\cdot)$ is submodular if for any $A, B \subseteq I$,

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

A submodular-cover problem is a minimization problem of the following formulation:

$$\min |A|$$
$$s.t. \quad g(A) = g(I),$$

here g is normalized, monotone increasing and submodular. By normalized we mean $g(\emptyset = 0$. It is not hard to prove that the function $f(\cdot)$ we defined in Sect. 3 is normalized, monotone increasing and submodular. So the PIDS problem is also a submodular-cover problem.

Next we show that the degree distribution of the power-law graphs can improve the performance of the greedy algorithm while applied to submodular-cover problems. First, we introduce some notations that are needed in the rest of this paper.

Given $G = (V, E)$, a graph with maximum vertex degree D. Let $V_i: 1 \leq i \leq D$ be the set of vertices with degree i and $n_i = |V_i|$. For any integer $k \in [1, D]$, let $L(k)$ denote the set of vertices with degrees greater than or equal to k, that is $L(k) = \{v | deg(v) \geq k\} = \cup_{k \leq i \leq D} V_i$. Suppose OPT is the optimal solution of the submodular-cover problem in power-law graph and $|OPT| = opt$, t is the largest integer that satisfies $|L(t - 1)| \geq opt$, clearly $D \geq t \geq 2$. Let $M(k)$ denote the set of k nodes with the highest degrees in the graph.

Previous researches have shown the property of greedy algorithms for submodular-cover problems in random graph as stated in the following lemma (Du et al. 2010).

Lemma 1 *For a submodular-cover problem, let S^* be the optimal solution and S be the solution produced by greedy algorithm, then*

$$|S| \leq \sum_{x \in S^*} \sum_{i=1}^{f(x)} \frac{1}{i} \leq \sum_{x \in S^*} (1 + \ln(f(x))).$$

Here we assume that

$$f(x) = deg(x).$$

Remark Actually, in most submodular-cover examples such as the dominating set problem, the set-cover problem and the hitting set problem, the value of $f(x)$ equals to $deg(x)$. Hence, it is reasonable for us to assume that $f(x) = deg(x)$.

Then we have

$$|S| \leq \sum_{x \in S^*} (1 + \ln(deg(x))).$$

Here, we define a function g on V, for any subset A of V,

$$g(A) = \sum_{x \in A} (1 + \ln(deg(x)))/|A|.$$

$g(A)$ actually presents the mean value of $1 + \ln(deg(x))$ for all the nodes in A. Then Lemma 1 can be rewritten as $|S| \leq g(S^*) \cdot |S^*|$, which means that when applied to the submodular-cover problems, greedy algorithm has an approximation ratio of $g(S^*)$. In random graphs, the following inequality holds,

$g(S^*) \leq 1 + \ln(D)$.

However in power-law graphs, most of the vertices have small degrees while only a few vertices have very large degrees. This phenomenon will make the performance ratio of greedy algorithm in power-law graphs quite different from that in random graphs. The following theorem shows the approximation ratio of greedy algorithms for submodular-cover problem in power-law graphs.

Theorem 1 *For submodular-cover problems in power-law graphs with $2 < \beta < 3$, greedy algorithm generates a solution of size within a factor of $g(L(t))$ from the optimal solution, where $g(L(t)) \leq 1 + \left(\ln(t-1) + \frac{1}{\beta-1}\right) \left(1 - \frac{1}{t}\right)^{1-\beta}$.*

Proof As defined earlier, for a positive integer h, $M(h)$ is the set of h nodes with the highest degrees in the graph, so

$g(S^*) \leq g(M(opt))$.

Notice that $L(k)$ is the set of $|L(k)|$ nodes with the largest degrees and t is the largest integer that satisfies $|L(t-1)| \geq opt$, so $|L(t)| < opt$ and $|M(opt)| = opt > |L(t)|$, hence

$g(M(opt)) < g(L(t))$.

and

$$g(L(t)) = \frac{\sum_{x \in L(t)}(1 + \ln(deg(x)))}{|L(t)|}$$

$$= \frac{\sum_{i=t}^{D}(1 + \ln i)\frac{e^\alpha}{i^\beta}}{\sum_{i=t}^{D}\frac{e^\alpha}{i^\beta}}$$

$$= 1 + \frac{\sum_{i=t}^{D}\ln i \frac{e^\alpha}{i^\beta}}{\sum_{i=t}^{D}\frac{e^\alpha}{i^\beta}}$$

$$\leq 1 + \frac{\int_{t-1}^{D}\ln x \cdot x^{-\beta}dx}{\int_{t}^{D+1}x^{-\beta}dx}$$

$$= 1 + \frac{\ln x \cdot x^{1-\beta}|_{t-1}^{D} - \frac{1}{1-\beta}x^{1-\beta}|_{t-1}^{D}}{x^{1-\beta}|_{t}^{D+1}}.$$

Let P represent the second part of the above formula, then

$$P = \frac{\left(\ln(t-1) + \frac{1}{\beta-1}\right)(t-1)^{1-\beta} - \left(\ln D + \frac{1}{\beta-1}\right)D^{1-\beta}}{t^{1-\beta} - (D+1)^{1-\beta}}.$$

To simplify this expression, we consider function

$$q(x) = \left(\ln x + \frac{1}{\beta-1}\right)\left(1 - \frac{1}{x+1}\right)^{1-\beta}.$$

$$q'(x) = (x - (\beta-1)\ln x)\frac{1}{x^2}\left(1 + \frac{1}{\beta}\right)^{\beta-2}.$$

When $2 < \beta < 3$ and $x \geq 1$, $q'(x) > 0$. So $q(x)$ is monotone increasing and

$$\frac{\left(\ln(t-1) + \frac{1}{\beta-1}\right)(t-1)^{1-\beta}}{t^{1-\beta}} \leq \frac{\left(\ln D + \frac{1}{\beta-1}\right)(D)^{1-\beta}}{(D+1)^{1-\beta}}.$$

By the simple fact that if

$$\frac{a}{b} \leq \frac{c}{d},$$

then

$$\frac{a-c}{b-d} \leq \frac{a}{b},$$

thus

$$P \leq \frac{\left(\ln(t-1) + \frac{1}{\beta-1}\right)(t-1)^{1-\beta}}{t^{1-\beta}}.$$

So we have

$$g(S^*) \leq g(L(t)) \leq 1 + \left(\ln(t-1) + \frac{1}{\beta-1}\right)\left(1 - \frac{1}{t}\right)^{1-\beta}.$$

$\square$

Here we should notice that $\left(1 - \frac{1}{t}\right)^{1-\beta}$ is no bigger than 4 and tends to 1 quickly as t increases.

4.2 PIDS problem in power-law graphs

For the class of submodular-cover problems in power-law graphs, Theorem 1 shows that the greedy algorithm has a new performance ratio with a parameter t. In this section we show that for the PIDS problem, which is also a submodular-cover problem, the parameter t can be bounded. This yields a constant performance ratio for the greedy algorithm immediately.

For a given graph $G = (V, E)$ with $|V| = n$, let C with $|C| = opt$ be the minimum PIDS set. We claim that set C can not be too small. Because for all the nodes in a network, everyone of them has to see a certain number of PIDS nodes in the neighborhood. A relaxed constraint requires that each of them has to be connected to at least one node in PIDS. While on the other side, the whole PIDS can connect with at most $deg(C)$ such nodes where

$$deg(C) = \sum_{v \in C} deg(v).$$

Hence

$$deg(C) \geq n.$$

Lemma 2 *For power-law graph G with $2 < \beta < 3$, t is the largest integer that satisfies $|L(t-1)| \geq opt$. Then*

$$t - 1 \leq \sqrt[\beta-2]{\frac{\beta-1}{\beta-2}}.$$

Proof Based on the definition of function M, we have

$$deg(C) \leq deg(M(opt)) \leq \int_{t-1}^{D} \frac{e^{\alpha}}{x^{\beta-1}} dx.$$

In addition,

$$n \geq \sum_{i=1}^{D} |V_i| \geq \int_{1}^{D+1} \frac{e^{\alpha}}{x^{\beta}} dx.$$

Hence

$$\int_{t-1}^{D} \frac{e^{\alpha}}{x^{\beta-1}} \geq \int_{1}^{D+1} \frac{e^{\alpha}}{x^{\beta}} dx.$$

By simple integral operation we get

$$D^{2-\beta} - (t-1)^{2-\beta} \leq \theta(D+1)^{1-\beta} - \theta.$$

In which $\theta = \frac{\beta-2}{\beta-1}$. After eliminating and simplification we have

$$(t-1)^{\beta-2} \leq \frac{1}{\theta}$$

and

$$t - 1 \leq \sqrt[\beta-2]{\frac{\beta-1}{\beta-2}}$$

$\square$

This lemma together with Theorem 1 lead us directly to the following conclusion:

Theorem 2 *For the PIDS problem in power-law graphs, the greedy algorithm has a constant approximation ratio of*

$$1 + \left(\frac{1}{\beta-1} + \ln\gamma\right)\left(1 - \frac{1}{\gamma+1}\right)^{1-\beta}.$$

where $\gamma = \sqrt[\beta-2]{\frac{\beta-1}{\beta-2}}$.

Proof Theorem 1 already shows that for submodular-cover problem with the constraint $f(x) \leq deg(x)$, the performance ratio of the greedy algorithm is $g(L(t))$ and

$$g(L(t)) \leq 1 + \left(\ln(t-1) + \frac{1}{\beta-1}\right)\left(1 - \frac{1}{t}\right)^{1-\beta}.$$

Recall that in the proof of Theorem 1, we proved that

$$q(x) = \left(\ln x + \frac{1}{\beta-1}\right)\left(1 - \frac{1}{x+1}\right)^{1-\beta}$$
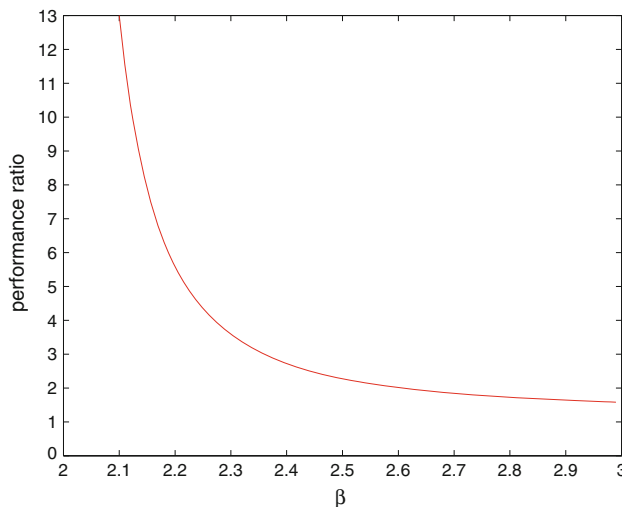

Fig. 1 The value of the constant guarantee of greedy algorithm for PIDS problem on power-law graphs

is monotone increasing. Combine with the result of Lemma 2 that

$$t - 1 \leq \sqrt[\beta-2]{\frac{\beta-1}{\beta-2}},$$

it is clear that

$$g(L(t)) \leq 1 + \left(\frac{1}{\beta-1} + \ln\gamma\right)\left(1 - \frac{1}{\gamma+1}\right)^{1-\beta}.$$

where $\gamma = \sqrt[\beta-2]{\frac{\beta-1}{\beta-2}}$. $\square$

The value of this performance ratio function is illustrated in Fig. 1. For the social network of Portland city introduced in (Eubank et al. 2004) with $\beta \sim 2.8$, the guarantee value would be close to 1.73.

Remarks this constant performance ratio also holds for the general dominating set problem in power-law graphs.

5 Simulation result

To evaluate the effect of the proposed greedy algorithm, we also carry out simulation over social network topologies retrieved from real online social networks. As identified in Willinger et al. (2010), one challenge to online social network research is the size of the network which is usually billions of users. To do experiment, we can only grab a subset of the online social network topology. In order to generate representative social network topologies, we collect the data from one of the popular gaming applications, Fighter's Club (FC) (Nazir et al. 2008) on Facebook social networking site. The FC game has attracted over 3.4
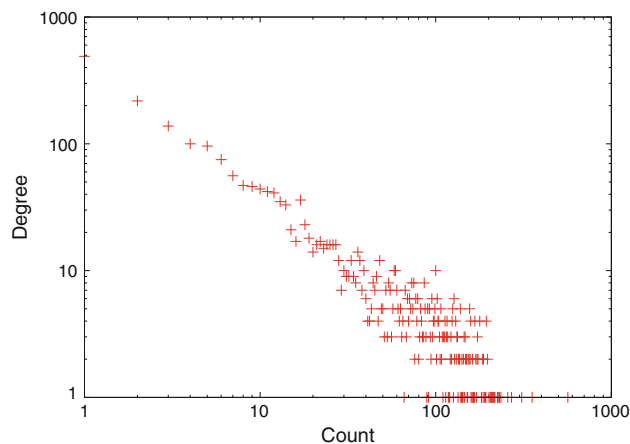
Fig. 2 Node degree distribution of an gaming application of 2,334 users on facebook

Table 1 Positive influence dominating set size

Network size	Avg. degree	PIDS size	PIDS percentage (%)	Avg. positive degree percentage (%)
2,334	28.42	683	29.26	79.05
1,825	30.25	501	27.45	79.03
797	12.36	273	34.25	78.63
697	13.62	238	34.14	81.53
543	7.58	250	46.04	82.51

Fig. 3 The impact of prune on PIDS size

Fig. 4 The impact of prune on the degree of positive neighbors

million Facebook users since it was initially launched on June 2007. The gaming application records the players as well as their IP addresses. We generate five different social network topologies from facebook users who play FC games together by choosing subsets of players from same network, that is, the players having the same 24-bit network address in their internet addresses. These five social networks are of size 543, 697, 797, 1825, 2334 respectively and all demonstrate power-law distribution. Figure 2 illustrates the distribution of the node degree in the online community of size 2,334. It is clear that the node degree in this Facebook application community follows a power-law distribution, similar to the observations in prior studies (Barabsi and Bonabeau 2003).

For each social network, we first apply greedy algorithm on the topology to find PIDS. The simulation result is illustrated in Table 1. Network size is the total number of nodes in the social network, average degree is the average degree over all the nodes in the social network. PIDS size is the number of nodes chosen into PIDS. PIDS percentage is the ratio of PIDS size over network size. Positive degree of a node is the number of its PIDS neighbors. Average positive degree percentage is the average of the ratio of positive degree over node degree. PIDS percentage and
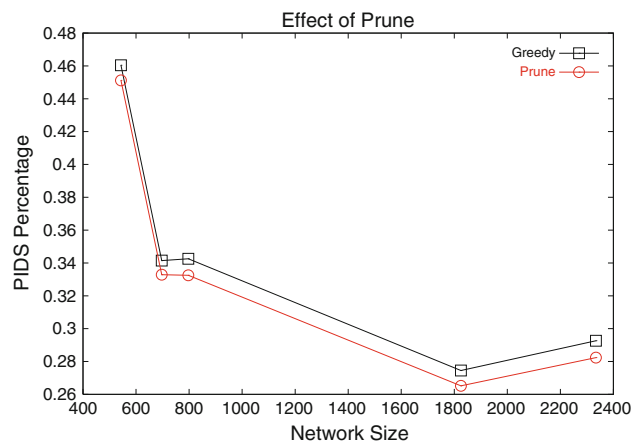
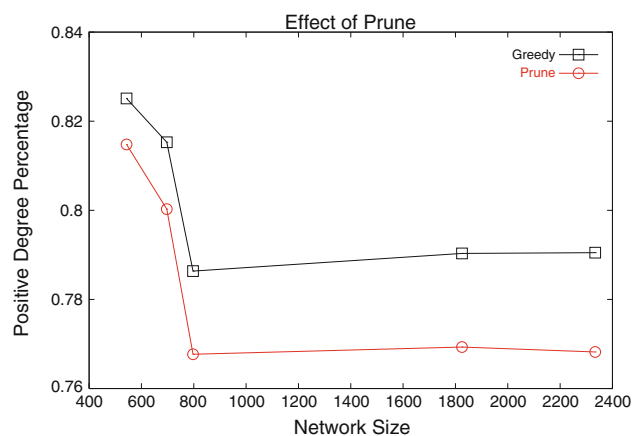average positive degree percentage measure the performance of the greedy algorithm. The lower the PIDS percentage, the higher the average positive degree percentage, the better the algorithm performance. As can be seen, for a social network including 2334 players and average degree of 28.42, 29.26% is chosen into PIDS and in average, each node has 79.05% of its neighbors in PIDS, which is much better than the targeted 50% since each node has more positive influence neighbors means the higher possibility that it will convert to have positive impact to its neighbors. This also motivate us to prune the constructed PIDS to further reduce the size of PIDS in order to reduce the cost. The prune is implemented as follows: for each node u in PIDS, if removing node u from PIDS leaves every node in the original topology positive influenced by the remaining PIDS nodes, then remove node u from PIDS. Figure 3 illustrates the effect of pruning. With pruning, the PIDS percentage of all five social networks drops by close to 1%. We also did experiment of pruning node with minimum degree first but the result is very close to randomly choose a prunable node to prune first. Figure 4 illustrates the
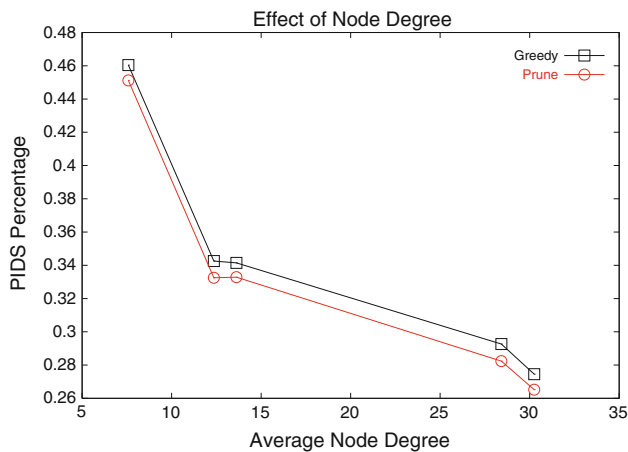
Fig. 5 Percentage of PIDS and degree impact

impact of prune process on the degree of positive neighbors of a node. *x*-axis represents number of nodes in a social network, *y*-axis represents the average ratio of positive neighbor degree over node degree of the nodes. It is clear that as less nodes (around 1%) are chosen into PIDS, average positive neighbor degree of nodes drops correspondingly (between 1 and 2.2%).

Figure 5 illustrates the impact of node degree on PIDS size. *x*-axis is the average node degree of the five social networks, *y*-axis is the ratio of PIDS nodes over all nodes. It demonstrates that as the node degree which presents the network density increases, the ratio of PIDS nodes decreases.

6 Conclusion

In this paper, we study the performance of greedy algorithm for general submodular-cover problem in power-law graphs then prove that the greedy algorithm for PIDS problem in power-law graphs has constant approximation ratio. Simulation results show that the greedy algorithm can effectively choose a small set of PIDS which can positively influence the whole network. Future work includes further improvement of the approximation ratio of the greedy algorithm by considering the structure property in addition to degree distribution in power-law graphs and study variations of PIDS problem such as defining new mathematical model to capture the impact between two nodes.

References

Aiello W, Chung F, Lu L (2001) A random graph model for power-law graphs. Exp Math 10:53–66

Alan M, Massimiliano M, Peter G, Bobby B (2007) Measurement and analysis of online social networks. ACM SIGCOMM IMC. doi:10.1145/1298306.1298311

Aris A, Ravi K, Mohammad M (2008) Influence and correlation in social networks. ACM SIGKDD. doi:10.1145/1401890.1401897

Baden R, Bender A, Spring N, Bhattacharjee B, Starin D (2009) Persona: an online social network with user-defined privacy. ACM SIGCOMM. doi:10.1145/1592568.1592585

Barabsi A, Bonabeau E (2003) Scale-free networks. Sci Am 288:60–69

Bhattacharyya P, Garg A, Wu S (2011) Analysis of user keyword similarity in online social networks. SOCNET. doi:10.1007/s13278-010-0006-4

Cha M, Mislove A, Gummadi K (2009) A measurement-driven analysis of information propagation in the flickr social network. WWW, pp 721–730

Clauset A, Shalizi C, Newman M (2009) Power-law distributions in empirical data, SIAM Rev. doi:10.1137/070710111

David K, Jon K, Eva T (2003) Maximizing the spread of influence through a social network. ACM SIGKDD, pp 137–146

Dougherty H (2010) Facebook Reaches Top Ranking in US. Hitwise Intelligence. http://weblogs.hitwise.com/heather-dougherty/2010/03/facebook_reaches_top_ranking_i.html (accessed 15 March 2010)

Du D, Ko K, Hu X (2010) Design and analysis of approximation algorithms. Manuscripts

Eubank S, Kumar V, Marathe M, Srinivasan A, Wang N (2004) Structural and algorithmic aspects of massive social networks. ACM SIAM, pp 718–727

Gross R, Acquisti A (2005) Information revelation and privacy in online social networks. ACM workshop on Privacy in the electronic society, pp 71–80

Kayaalp M, Özyer T, Özyer S (2011) A mash-up application utilizing hybridized filtering techniques for recommending events at a social networking site. SOCNET. doi:10.1007/s13278-010-0010-8

Nazir A, Raza S, Chuah C (2008) Unveiling facebook: a measurement study of social network based applications. ACM SIGCOMM IMC. doi:10.1145/1452520.1452527

Rosen D, Barnett G, Kim J (2011) Social networks and online environments: when science and practice co-evolve. SOCNET. doi:10.1007/s13278-010-0011-7

Saravanan M, Prasad G, Karishma S, Suganthi D (2011) Analyzing and labeling telecom communities using structural properties. SOCNET. doi:10.1007/s13278-011-0020-1

Schneider F, Feldmann A, Krishnamurthy B, Willinger W (2009) Understanding online social network usage from a network perspective. ACM SIGCOMM IMC, pp 35–48

Shishir Bharathi M, Kempe D (2007) Competitive influence maximization in social networks. WINE. doi:10.1007/978-3-540-77105-0-31

Wang F, Camacho E, Xu K (2009) Positive influence dominating set in online social networks. LNCS, pp 313-321

Wang F, Du H, Camacho E, Xu K, Lee W, Shi Y, Shan S (2011) On positive influence dominating sets in social networks. Theor Comput Sci. doi:10.1016/j.tcs.2009.10.001

Willinger W, Rejaie R, Torkjazi M, Valafar M, Maggioni M (2010) Research on online social networks: time to face the real challenges. SIGMETRICS Perform. Eval. Rev. doi:10.1145/1710115.1710125