

Local Centrality Minimization with Quality Guarantees

Anonymous Author(s)*

ABSTRACT

Centrality measures, quantifying the importance of vertices or edges, play a fundamental role in network analysis. To date, triggered by some positive approximability results, a large body of work has been devoted to studying centrality maximization, where the goal is to maximize the centrality score of a target vertex by manipulating the structure of a given network. On the other hand, due to the lack of such results, only very little attention has been paid to centrality minimization, despite its practical usefulness.

In this study, we introduce a novel optimization model for local centrality minimization, where the manipulation is allowed only around the target vertex. We prove the NP-hardness of our model and that the most intuitive greedy algorithm has a quite limited performance in terms of approximation ratio. Then we design two effective approximation algorithms: The first algorithm is a highly-scalable algorithm that has an approximation ratio unachievable by the greedy algorithm, while the second algorithm is a bicriteria approximation algorithm that solves a continuous relaxation based on the Lovász extension, using a projected subgradient method. To the best of our knowledge, ours are the first polynomial-time algorithms with provable approximation guarantees for centrality minimization. Experiments using a variety of real-world networks demonstrate the effectiveness of our proposed algorithms: Our first algorithm is applicable to million-scale graphs and obtains much better solutions than those of scalable baselines, while our second algorithm is rather strong against adversarial instances.

ACM Reference Format:

Anonymous Author(s). 2018. Local Centrality Minimization with Quality Guarantees. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 13 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Among the many analytical tools that social network analysis [21] borrowed from graph theory, centrality measures play a fundamental role in a wide variety of analyses [12]. Centrality, which quantifies the *importance* of vertices or edges only based on the graph structure, has found many applications, including, e.g., identification of important users or connections in social networks, community detection [34], anomaly detection [28], to name a few.

Local centrality minimization is the problem of removing a few existing edges around a target vertex, so as to minimize its centrality score. A direct application can be found in the context of reducing

the visibility, or influence, of a targeted harmful user in a social network, without explicitly blocking the user account. The edges to be removed, identified by local centrality minimization, are the most important edges for the centrality (i.e., visibility or influence) of the target vertex. In this regard, another direct application is to keep satisfying influential users so that they are engaged in the platform. The degree of satisfaction of such influencers often depends on how actually influential they are in the platform, i.e., how much their content is consumed by the network. Local centrality minimization can be used for revealing the most important connections between the influencers and their followers, which are key in contributing to their visibility.

While a large body of work has been devoted to studying *centrality maximization* [12], much less attention has been paid to centrality minimization (see Section 2 for a brief literature survey). Generally speaking, the goal of centrality maximization is to maximize the centrality score of a target vertex by adding a limited number of edges to the network. In many reasonable optimization models, the objective function becomes monotone and submodular [14], and thus a simple greedy algorithm admits a $(1 - 1/e)$ -approximation [33]. This positive result makes the basis of various studies on centrality maximization [4, 7, 10, 11, 18, 29].

The lack of positive approximation results has instead limited the attention on the centrality minimization problem, especially in its local variant. Waniek et al. [41] introduced several optimization models for local centrality minimization under some specific objectives and constraints, investigated the computational complexity of their models, and devised some algorithms. Later, Waniek et al. [42] investigated local centrality minimization from a game-theoretic point of view. However, the work by Waniek et al. [41, 42] proposed only (exponential-time) exact algorithms and heuristics.

In this paper, we study the local centrality minimization problem, adopting the most well-established centrality measures called the *harmonic centrality* [12], which quantifies the importance of vertices based on the level of reachability from the other vertices. The harmonic centrality is known as an effective alternative to the closeness centrality [6], which was employed in Waniek et al. [41], in the sense that unlike the closeness centrality, it is well-defined even in the case where a graph is not strongly connected.

Boldi and Vigna [6] showed that among all the known centrality measures, only the harmonic centrality satisfies all the desirable axioms, namely the size axiom, density axiom, and score monotonicity axiom. Recently, Murai and Yoshida [32] theoretically and empirically demonstrated that among well-known centrality measures, the harmonic centrality is most stable (thus reliable) against the uncertainty of a given graph.

1.1 Paper contributions and roadmap

In this paper, we introduce a novel optimization model for local centrality minimization, where the harmonic centrality is employed as an objective function. Specifically, in our model, given a directed graph $G = (V, A)$, a target vertex $v \in V$, and a budget $b \in \mathbb{Z}_{>0}$,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXXX>

we aim to find a set of incoming edges of v with size (no greater than) b whose removal minimizes the harmonic centrality score of v , denoted by $h_G(v)$ (to be defined in Section 3).

For our optimization model, we first analyze the computational complexity. Specifically, we show that our model is NP-hard even on a very limited graph class (i.e., acyclic graphs), by constructing a polynomial-time reduction from the minimum k -union problem. Furthermore, we prove that the most intuitive greedy algorithm, which iteratively removes an incoming edge of v that maximally decreases the objective value, cannot achieve an approximation ratio of $o(|V|)$, while any reasonable algorithm has an approximation ratio of $O(|V|)$. This negative result motivates the design of algorithms that exploit the characteristics of our model.

We design two polynomial-time approximation algorithms. The first algorithm is a highly-scalable algorithm that has an approximation ratio of $\sqrt{2h_G(v)}$. We stress that as $\sqrt{2h_G(v)} = O(\sqrt{|V|}) = o(|V|)$, this approximation ratio is unachievable by the above greedy algorithm. Our algorithm first sorts the incoming neighbors of the target vertex v in the decreasing order of their harmonic centrality scores on a slightly modified graph, and then removes b incoming edges from the top- b vertices in the sorted list. To prove the approximation ratio, we scrutinize the relationship between the harmonic centrality scores of the target vertex and its incoming neighbors. In the end, we also prove the tightness of our analysis of the approximation ratio.

The second algorithm is a polynomial-time algorithm that has a bicriteria approximation ratio of $(\frac{1}{\alpha}, (\frac{1}{1-\alpha}, \epsilon))$ for any $\alpha \in (0, 1)$ and $\epsilon > 0$. That is, the algorithm finds a subset of incoming edges of the target vertex v with size at most b/α but attains the objective value at most the original optimal value times $\frac{1}{1-\alpha}$ plus ϵ . Therefore, the algorithm approximates the original optimal value while violating the budget constraint to some bounded extent. To design the algorithm, we first introduce a continuous relaxation of our model. To this end, we use the well-known extension of set functions, called the Lovász extension [26]. An important fact is that the objective function of our model is submodular, which guarantees that its Lovász extension is (not necessarily differentiable but) convex. Therefore, we can solve the relaxation (with an arbitrarily small error) using a projected subgradient method [3]. Once we get a fractional solution, we apply a simple probabilistic procedure and obtain a subset of incoming edges of the target vertex v .

Finally, our experiments on a variety of real-world networks show that our first algorithm is applicable to million-scale graphs and obtains much better solutions than those of scalable baselines, while our second algorithm is strong against adversarial instances.

In summary, our contributions are as follows:

- We study the *local harmonic centrality minimization problem*: We prove that it is NP-hard even on acyclic graphs, its objective function is submodular, and the most intuitive greedy algorithm cannot achieve $o(|V|)$ -approximation (Section 3).
- We devise a highly-scalable algorithm with an approximation ratio of $\sqrt{2h_G(v)}$, unachievable by the greedy algorithm.
- We then devise a bicriteria approximation algorithm that solves a continuous relaxation based on the Lovász extension, using a projected subgradient method (Sections 5 and 6).

To the best of our knowledge, ours are the first polynomial-time algorithms with provable approximation guarantees for centrality minimization.

2 RELATED WORK

In this section, we review related literature about centrality minimization and maximization, submodular minimization, and other relevant applications in social networks analysis.

Centrality minimization. The most related to the present paper is the work on centrality minimization by Waniek et al. [41, 42] whose goal is to provide a methodology that contributes to hiding individuals in social networks from centrality-based network analysis algorithms.

More specifically, Waniek et al. [41] introduced the following optimization model: Given a directed graph $G = (V, A)$, a target vertex $v \in V$, a budget $b \in \mathbb{Z}_{>0}$, and a set M of possible single-edge modifications, we are asked to select at most b actions in M so as to minimize the centrality score of v while satisfying some constraint on the influence of some vertices in the graph. As a centrality measure, they considered the degree centrality, closeness centrality, and betweenness centrality. They showed that except for the degree centrality case, the above model is NP-hard even when the constraint on the influence of vertices is ignored, and devised simple heuristics.

Waniek et al. [42] investigated local centrality minimization from a game-theoretic point of view. As a tool to analyze their game, they studied the following optimization model: Given a directed graph $G = (V, A)$, a target vertex $v \in V$, a hiding parameter δ , and a set M as above, we are asked to find a minimal subset of M guaranteeing that there are at least δ vertices having a centrality score greater than that of v . As a centrality measure, they again considered the above three. They showed that the model is 2-approximable for the degree centrality but is inapproximable within any logarithmic factor for the other two. Note that the above approximation is just for the size of the output rather than the ranking of the centrality score of v (or the centrality score of v).

Veremyev et al. [38] studied a global centrality minimization problem: Given an undirected graph $G = (V, E)$ with cost $c_e \in \mathbb{R}_{\geq 0}$ for each $e \in E$, a target vertex subset $S \subseteq V$, and a budget $b \in \mathbb{Z}_{>0}$, we are asked to find $F \subseteq E$ whose removal minimizes the centrality score of the target vertex subset S subject to the budget constraint $\sum_{e \in F} c_e \leq b$. The centrality score of a vertex subset is defined as a generalization of the centrality score of a vertex. As a centrality measure, they considered a quite general one, based on distance between vertices, which includes the harmonic centrality as a special case. They proved that the above model is NP-hard for any centrality measure included in the above, and as a by-product of the analysis, they also mentioned the NP-hardness of its local variant, which coincides with (the undirected-graph counterpart of) our proposed model. In the present paper, by focusing on the harmonic centrality, we prove that our model is NP-hard even on a very limited graph class (i.e., acyclic graphs). On a positive side, they presented an exact algorithm based on mathematical programming and greedy heuristics. Very recently, Liu et al. [25] addressed another global centrality minimization problem, where

the objective function is a centrality measure called the information centrality and the connectivity of the resulting graph is guaranteed.

Centrality maximization. Centrality maximization has more actively been studied in the literature (e.g., [4, 7, 10, 11, 18, 29]), where the most related to ours is due to Crescenzi et al. [10]. They introduced the harmonic centrality maximization problem, where given a directed graph $G = (V, A)$, a target vertex $v \in V$, and a budget $b \in \mathbb{Z}_{>0}$, we are asked to insert at most b incoming edges of v so as to maximize the harmonic centrality score of v . Our proposed optimization model can be seen as a minimization counterpart of their problem. They proved that the problem is APX-hard, but devised a polynomial-time $(1 - 1/\epsilon)$ -approximation algorithm based on the submodularity of the objective function. Finally, we note that there is another class of problems also called centrality maximization, where the goal is to find $S \subseteq V$ that has the maximum group centrality score [1, 2, 4, 8, 18, 24, 27, 31, 35, 45], which is less relevant to the present paper.

Submodular minimization. Submodular minimization is one of the most well-studied problem classes in combinatorial optimization. Among the literature, the most related work is due to Svitkina and Fleischer [37]. They stated that a polynomial-time $(\frac{1}{\alpha}, \frac{1}{1-\alpha})$ -bicriteria approximation algorithm for submodular minimization with a cardinality upper bound (and thus for our proposed model) is possible for any $\alpha \in (0, 1)$, using techniques in Hayrapetyan et al. [17]. However, Hayrapetyan et al. [17] did not consider the problem: They addressed the problem called the minimum-size bounded-capacity cut, where the function in the constraint instead of the objective function is submodular. Therefore, the above statement is not trivial and even our proposed model should be handled in a formal way. Lovász extension has actively been used for developing novel network analysis algorithms (e.g., [22, 23, 36]).

Applications. Reducing the visibility or influence of target users in social networks has been studied in the context of influence minimization [20, 40, 43]. All of existing studies are based on some influence diffusion models such as the independent cascade model [15, 16] and the linear threshold model [19]. Unlike those, our model does not assume any influence diffusion model, but is just based on the network structure. Very recently, Fabbri et al. [13] and Coupette et al. [9] addressed the problem of reducing the exposure to harmful contents in social media networks.

On the other hand, identifying the users and/or connections that play a key role for user engagement in social networks has also attracted much attention. Bhawalkar et al. [5] initiated this kind of study from an optimization perspective. They invented a model that aims to find a group of users whose permanent use of the service guarantees user engagement as much as possible, and designed polynomial-time algorithms for some cases. Later, Zhang et al. [44] and Zhu et al. [46] introduced variants of the above model, and devised intuitive heuristics.

3 PROBLEM FORMULATION AND CHARACTERIZATION

In this section, we mathematically formulate our problem (Problem 1), and prove its NP-hardness (Theorem 1) and the submodularity of the objective function (Theorem 2). Finally, we show the quite limited performance of the greedy algorithm (Theorem 3).

Let $G = (V, A)$ be a directed graph (or *digraph* for short). Throughout the paper, we assume that digraphs are simple, that is, there exist neither self-loops nor multiple edges. For $F \subseteq A$, we define $G \setminus F$ as the subgraph of G that is constructed by removing all edges in F from G , i.e., $G \setminus F = (V, A \setminus F)$. For $v \in V$, we denote by $\rho(v)$ the set of incoming edges of v , i.e., $\rho(v) = \{(u, v) \in A \mid u \in V\}$. For $v \in V$, let $h_G(v)$ be the harmonic centrality score of v on a digraph G , i.e.,

$$h_G(v) = \sum_{u \in V \setminus \{v\}} \frac{1}{d_G(u, v)},$$

where $d_G(u, v)$ is the (shortest-path) distance from $u \in V$ to $v \in V$ on G , and by convention, $d_G(u, v) = \infty$ when v is not reachable from u . Note that, contrarily to other centrality measures, even in the case where a digraph is not strongly connected, the harmonic centrality is still well-defined (assuming by convention that $1/\infty = 0$). Intuitively, the harmonic centrality quantifies the importance of a given vertex v based on the level of reachability from the other vertices. The problem we tackle in this paper is formalized as follows:

Problem 1 (Local harmonic centrality minimization). *Given a digraph $G = (V, A)$, a target vertex $v \in V$, and a budget $b \in \mathbb{Z}_{>0}$, we are asked to find $F \subseteq \rho(v)$ with $|F| \leq b$ whose removal minimizes the harmonic centrality of $v \in V$, i.e., $f_{(G,v)}(F) := h_{G \setminus F}(v)$.*

By constructing a polynomial-time reduction from the NP-hard optimization problem called the *minimum k -union*, we can prove the following. The proof can be found in Appendix A.1.

Theorem 1. *Problem 1 is NP-hard even on acyclic graphs.*

We next show that the objective function $f_{(G,v)}$ of Problem 1 is submodular, which helps us design our bicriteria approximation algorithm in Section 5. Let S be a finite set. A set function $f: 2^S \rightarrow \mathbb{R}$ is said to be *submodular* if for any $X, Y \subseteq S$, it holds that

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y).$$

We prove the following in Appendix A.2:

Theorem 2. *For any $G = (V, A)$ and $v \in V$, the objective function $f_{(G,v)}$ of Problem 1 is submodular.*

Finally, we prove that the most intuitive greedy algorithm has a quite limited performance in terms of the approximation ratio. Specifically, we consider the algorithm that iteratively removes an incoming edge of the target vertex v that maximally decreases the harmonic centrality score of v , until it exhausts the budget. For reference, the pseudo-code is given in Algorithm 4 in Appendix A.3. This algorithm runs in $O(b|\rho(v)|(|V| + |A|))$ time. Note that, unlike many submodular maximization algorithms, the *lazy evaluation* technique [30] cannot be used to obtain a practically efficient implementation. The proof of the following is available in Appendix A.4.

Theorem 3. *The greedy algorithm has no approximation ratio of $o(|V|)$ for Problem 1, while any algorithm that outputs $F \subseteq \rho(v)$ with $|F| = b$ has an approximation ratio of $O(|V|)$.*

4 SCALABLE APPROXIMATION ALGORITHM

In this section, we present a highly-scalable $\sqrt{2h_G(v)}$ -approximation algorithm for Problem 1.

Algorithm 1: $\sqrt{2h_G(v)}$ -approximation algorithm for Problem 1

Input : $G = (V, A)$, $v \in V$, and $b \in \mathbb{Z}_{>0}$

Output: $F \subseteq \rho(v)$ with $|F| \leq b$

1 Sort the elements of $N_{\text{in}}(v)$ as $(w_1, \dots, w_{|\rho(v)|})$ so that

$$h_{G \setminus \rho(v)}(w_1) \geq \dots \geq h_{G \setminus \rho(v)}(w_{|\rho(v)|});$$

2 **return** $\{(w_1, v), \dots, (w_b, v)\}$;

4.1 Algorithm

Let $N_{\text{in}}(v)$ be the set of incoming neighbors of v , i.e., $N_{\text{in}}(v) = \{w \in V \mid (w, v) \in \rho(v)\}$. The intuition behind our algorithm is quite simple: As long as there exists a vertex $w \in N_{\text{in}}(v)$ that has a large harmonic centrality score, so does the target vertex v . This means that it is urgent to remove incoming edges of v that come from vertices having large harmonic centrality scores. Note that our algorithm and analysis consider the harmonic centrality scores on $G \setminus \rho(v)$ rather than G ; this is essential to obtain our approximation ratio. Specifically, our algorithm first sorts the elements of $N_{\text{in}}(v)$ as $(w_1, \dots, w_{|\rho(v)|})$ so that $h_{G \setminus \rho(v)}(w_1) \geq \dots \geq h_{G \setminus \rho(v)}(w_{|\rho(v)|})$ and just returns $\{(w_1, v), \dots, (w_b, v)\}$. For reference, the entire procedure is described in Algorithm 1.

The algorithm is highly scalable. Indeed, the time complexity of Algorithm 1 is dominated by the part of computing the harmonic centrality scores of vertices in $N_{\text{in}}(v)$, which just takes $O(|\rho(v)|(|V| + |A|))$ time. Therefore, the algorithm is asymptotically b times faster than the greedy algorithm (Algorithm 4).

4.2 Analysis

From now on, we analyze the approximation ratio of Algorithm 1. The following lemma demonstrates that the optimal value can be lower bounded using the maximum harmonic centrality score over the remaining incoming neighbors of v in the resulting graph:

Lemma 1. *Let F^* be an optimal solution to Problem 1 and N^* the vertex subset corresponding to F^* , i.e., $N^* = \{w \in V \mid (w, v) \in F^*\}$. Then it holds that*

$$f_{(G,v)}(F^*) \geq \frac{1}{2} \left(\max_{w \in N_{\text{in}}(v) \setminus N^*} h_{G \setminus \rho(v)}(w) + 1 \right).$$

PROOF. For any $w \in N_{\text{in}}(v) \setminus N^*$, we have

$$\begin{aligned} f_{(G,v)}(F^*) &= h_{G \setminus F^*}(v) = \sum_{u \in V \setminus \{v\}} \frac{1}{d_{G \setminus F^*}(u, v)} \\ &\geq \sum_{u \in V \setminus \{v\}} \frac{1}{d_{G \setminus F^*}(u, w) + 1} = 1 + \sum_{u \in V \setminus \{w, v\}} \frac{1}{d_{G \setminus F^*}(u, w) + 1} \\ &\geq 1 + \frac{1}{2} \sum_{u \in V \setminus \{w, v\}} \frac{1}{d_{G \setminus F^*}(u, w)} \geq \frac{1}{2} + \frac{1}{2} \sum_{u \in V \setminus \{w\}} \frac{1}{d_{G \setminus F^*}(u, w)} \\ &= \frac{1}{2} \left(h_{G \setminus F^*}(w) + 1 \right) \geq \frac{1}{2} \left(h_{G \setminus \rho(v)}(w) + 1 \right), \end{aligned}$$

where the first inequality follows from the triangle inequality of distance $d_{G \setminus F^*}$, the third equality follows from $d_{G \setminus F^*}(w, w) = 0$, and the second inequality follows from the fact that the addition of 1 to the denominator makes it at most twice the original. The

arbitrariness of the choice of $w \in N_{\text{in}}(v) \setminus N^*$ derives the statement. \square

On the other hand, the next lemma upper bounds the objective value of the output of Algorithm 1 using the harmonic centrality scores of the remaining incoming neighbors of v in the resulting graph:

Lemma 2. *Let F_{ALG} be the output of Algorithm 1 and N_{ALG} the vertex subset corresponding to F_{ALG} , i.e., $N_{\text{ALG}} = \{w \in V \mid (w, v) \in F_{\text{ALG}}\}$. Then we have*

$$f_{(G,v)}(F_{\text{ALG}}) \leq (|\rho(v)| - b) + \sum_{w \in N_{\text{in}}(v) \setminus N_{\text{ALG}}} h_{G \setminus \rho(v)}(w).$$

PROOF. On digraph $G \setminus F_{\text{ALG}}$, any $u \in V \setminus \{v\}$ satisfies either (i) there exists no (shortest) path from u to v or (ii) there exists $w(u) \in N_{\text{in}}(v) \setminus N_{\text{ALG}}$ that is contained in a shortest path from u to v , i.e., $d_{G \setminus F_{\text{ALG}}}(u, v) = d_{G \setminus F_{\text{ALG}}}(u, w(u)) + 1$. Let $V' \subseteq V \setminus \{v\}$ be the subset of vertices that satisfy the condition (ii). Then we have

$$\begin{aligned} f_{(G,v)}(F_{\text{ALG}}) &= h_{G \setminus F_{\text{ALG}}}(v) = \sum_{u \in V \setminus \{v\}} \frac{1}{d_{G \setminus F_{\text{ALG}}}(u, v)} \\ &= \sum_{u \in V' \setminus \{v\}} \frac{1}{d_{G \setminus F_{\text{ALG}}}(u, w(u)) + 1}. \end{aligned} \quad (1)$$

We see that the shortest path corresponding to $d_{G \setminus F_{\text{ALG}}}(u, w(u))$ does not contain v (and thus any edge in $\rho(v) \setminus F_{\text{ALG}}$). Otherwise there would exist $w'(u) \in N_{\text{in}}(v) \setminus N_{\text{ALG}}$ satisfying that $d_{G \setminus F_{\text{ALG}}}(u, w'(u)) < d_{G \setminus F_{\text{ALG}}}(u, w(u))$, which contradicts the fact that $w(u)$ is contained in a shortest path from u to v on $G \setminus F_{\text{ALG}}$. Hence, we have

$$d_{G \setminus F_{\text{ALG}}}(u, w(u)) = d_{G \setminus \rho(v)}(u, w(u)).$$

Combining this with the equality (1), we have

$$\begin{aligned} f_{(G,v)}(F_{\text{ALG}}) &= \sum_{u \in V' \setminus \{v\}} \frac{1}{d_{G \setminus \rho(v)}(u, w(u)) + 1} \\ &= (|\rho(v)| - b) + \sum_{u \in V' \setminus \{v\} \setminus (N_{\text{in}}(v) \setminus N_{\text{ALG}})} \frac{1}{d_{G \setminus \rho(v)}(u, w(u)) + 1} \\ &< (|\rho(v)| - b) + \sum_{u \in V' \setminus \{v\} \setminus (N_{\text{in}}(v) \setminus N_{\text{ALG}})} \frac{1}{d_{G \setminus \rho(v)}(u, w(u))} \\ &\leq (|\rho(v)| - b) + \sum_{w \in N_{\text{in}}(v) \setminus N_{\text{ALG}}} h_{G \setminus \rho(v)}(w), \end{aligned}$$

where the last inequality holds by the fact that any term $\frac{1}{d_{G \setminus \rho(v)}(u, w(u))}$ in the summation of the left-hand-side appears as a term in $h_{G \setminus \rho(v)}(w)$ for appropriate $w = w(u)$ in the right-hand-side. Therefore, we have the lemma. \square

We are now ready to prove our main theorem:

Theorem 4. *Algorithm 1 is a $2(|\rho(v)| - b)$ -approximation algorithm for Problem 1.*

PROOF. Here we use the notation that appeared in Lemmas 1 and 2. By the behavior of Algorithm 1, we have

$$\max_{w \in N_{\text{in}}(v) \setminus N_{\text{ALG}}} h_{G \setminus \rho(v)}(w) \leq \max_{w \in N_{\text{in}}(v) \setminus N^*} h_{G \setminus \rho(v)}(w).$$

Using Lemmas 1 and 2 together with this inequality, we have

$$\begin{aligned}
f_{(G,v)}(F_{\text{ALG}}) &\leq (|\rho(v)| - b) + \sum_{w \in N_{\text{in}}(v) \setminus N_{\text{ALG}}} h_{G \setminus \rho(v)}(w) \\
&\leq (|\rho(v)| - b) \left(1 + \max_{w \in N_{\text{in}}(v) \setminus N_{\text{ALG}}} h_{G \setminus \rho(v)}(w) \right) \\
&\leq (|\rho(v)| - b) \left(1 + \max_{w \in N_{\text{in}}(v) \setminus N^*} h_{G \setminus \rho(v)}(w) \right) \\
&\leq (|\rho(v)| - b) \left(1 + 2f_{(G,v)}(F^*) - 1 \right) = 2(|\rho(v)| - b)f_{(G,v)}(F^*),
\end{aligned}$$

which completes the proof. \square

Based on the theorem, we obtain the desired approximation ratio:

Corollary 1. *Algorithm 1 is a $\sqrt{2h_G(v)}$ -approximation algorithm for Problem 1.*

PROOF. For any instance that satisfies $b = |\rho(v)|$, Algorithm 1 outputs the trivial optimal solution (i.e., $\rho(v)$). Therefore, in what follows, we focus only on the instances with $b < |\rho(v)|$. Obviously the output of any algorithm for Problem 1 has an objective value at most $h_G(v)$. On the other hand, the optimal value is at least $|\rho(v)| - b$ because in the resulting digraph, there are still $|\rho(v)| - b$ incoming neighbors of v , each of which contributes exactly 1 to the objective value. Therefore, any algorithm (including Algorithm 1) for Problem 1 has an approximation ratio of $\frac{h_G(v)}{|\rho(v)| - b}$. By combining this with Theorem 4, the approximation ratio of Algorithm 1 can be improved to

$$\min \left\{ 2(|\rho(v)| - b), \frac{h_G(v)}{|\rho(v)| - b} \right\} \leq \sqrt{2h_G(v)},$$

which presents the statement. \square

It should be remarked that as $\sqrt{2h_G(v)} \leq \sqrt{2(|V| - 1)} = o(|V|)$, the approximation ratio is unachievable by the greedy algorithm. Finally, we conclude this section by showing that the analysis of the approximation ratio is tight up to a constant factor. The proof is available in Appendix B.1.

Theorem 5. *Algorithm 1 has no approximation ratio of $o(\sqrt{h_G(v)})$.*

5 BICRITERIA APPROXIMATION ALGORITHM

In this section, we present a polynomial-time $(\frac{1}{\alpha}, (\frac{1}{1-\alpha}, \epsilon))$ -bicriteria approximation algorithm ($\alpha \in (0, 1)$ and $\epsilon > 0$) for Problem 1. Our algorithm first solves a continuous relaxation of the problem and then applies a simple probabilistic procedure to the fractional solution to obtain the output.

5.1 Continuous relaxation

To obtain a continuous relaxation of Problem 1, we consider the well-known extension of set functions, called the Lovász extension [26]. For our objective function $f_{(G,v)}$, the Lovász extension $\widehat{f}_{(G,v)}: [0, 1]^{\rho(v)} \rightarrow \mathbb{R}$ is defined in the following way: Let $\rho(v) = \{e_1, \dots, e_{|\rho(v)|}\}$. For $\mathbf{x} \in [0, 1]^{\rho(v)}$, we relabel the elements of $\rho(v)$ so that $x_{e_1} \geq x_{e_2} \geq \dots \geq x_{e_{|\rho(v)|}}$, and construct a sequence of subsets $\emptyset = X_0 \subset X_1 \subset \dots \subset X_{|\rho(v)|} = \rho(v)$, where $X_i = \{e_1, \dots, e_i\}$

Algorithm 2: $(\frac{1}{\alpha}, (\frac{1}{1-\alpha}, \epsilon))$ -bicriteria approximation algorithm for Problem 1

Input : $G = (V, A)$, $v \in V$, and $b \in \mathbb{Z}_{>0}$

Output: $F \subseteq \rho(v)$

- 1 $\epsilon' \leftarrow (1 - \alpha)\epsilon$;
- 2 Solve Relaxation (using Algorithm 3 in Section 6) and obtain its ϵ' -additive approximate solution $\mathbf{x}^* \in [0, 1]^{\rho(v)}$;
- 3 Pick $p \in [\alpha, 1]$ uniformly at random;
- 4 **return** $\{e \in \rho(v) \mid x_e^* \geq p\}$;

for $i = 1, \dots, |\rho(v)|$. Based on these, we define the value of $\widehat{f}_{(G,v)}(\mathbf{x})$ as follows:

$$\begin{aligned}
\widehat{f}_{(G,v)}(\mathbf{x}) &= (1 - x_{e_1})f_{(G,v)}(\emptyset) \\
&\quad + \sum_{i=1}^{|\rho(v)|-1} (x_{e_i} - x_{e_{i+1}})f_{(G,v)}(X_i) + x_{e_{|\rho(v)|}}f_{(G,v)}(\rho(v)).
\end{aligned}$$

Observe that for any $F \subseteq \rho(v)$, it holds that $\widehat{f}_{(G,v)}(\mathbf{1}_F) = f_{(G,v)}(F)$, where $\mathbf{1}_F$ is an indicator vector of F , taking 1 if $e \in F$ and 0 otherwise. Therefore, $\widehat{f}_{(G,v)}$ is indeed an extension of $f_{(G,v)}$.

The Lovász extension can be defined on any (not necessarily submodular) set function. The Lovász extension is always continuous but not necessarily differentiable. An important fact is that the Lovász extension is convex if and only if the original set function is submodular [26]. Therefore, by Theorem 2, the Lovász extension $\widehat{f}_{(G,v)}$ of $f_{(G,v)}$ is convex.

Using $\widehat{f}_{(G,v)}$, we introduce our continuous relaxation as follows:

$$\begin{aligned}
\text{Relaxation: minimize } &\widehat{f}_{(G,v)}(\mathbf{x}) \\
&\text{subject to } \|\mathbf{x}\|_1 \leq b, \mathbf{x} \in [0, 1]^{\rho(v)}.
\end{aligned}$$

For convenience, we denote by C the feasible region of the problem, i.e.,

$$C := \left\{ \mathbf{x} \in \mathbb{R}^{\rho(v)} : \|\mathbf{x}\|_1 \leq b \text{ and } \mathbf{x} \in [0, 1]^{\rho(v)} \right\}.$$

From the above, we see that Relaxation is a non-smooth convex programming problem. We will present an algorithm for Relaxation and its convergence result in Section 6. In the remainder of this section, we assume that for $\epsilon' = (1 - \alpha)\epsilon$, we can compute, in polynomial time, an ϵ' -additive approximate solution for Relaxation, i.e., a feasible solution for Relaxation that has an objective value at most the optimal value plus ϵ' .

5.2 Algorithm

Let $\mathbf{x}^* \in [0, 1]^{\rho(v)}$ be an ϵ' -additive approximate solution for Relaxation, where $\epsilon' = (1 - \alpha)\epsilon$. Then our algorithm picks $p \in [\alpha, 1]$ uniformly at random and just returns $\{e \in \rho(v) \mid x_e^* \geq p\}$. For reference, the entire procedure is summarized in Algorithm 2.

5.3 Analysis

The following theorem gives the bicriteria approximation ratio of Algorithm 2:

Algorithm 3: Projected subgradient method for Relaxation**Input** : $x_0 \in C$ and some stopping condition**Output**: $x \in C$

```

1  $t \leftarrow 0$ ;
2 while the stopping condition is not satisfied do
3   Pick a stepsize  $\eta_t > 0$  and a subgradient  $\widehat{f}'_{(G,v)}(x_t)$  of
    $\widehat{f}_{(G,v)}$  at  $x_t$ ;
4    $x_{t+1} \leftarrow \text{proj}_C(x_t - \eta_t \cdot \widehat{f}'_{(G,v)}(x_t))$  and  $t \leftarrow t + 1$ ;
5 return  $x_t$ ;

```

Theorem 6. For any $\alpha \in (0, 1)$ and $\epsilon > 0$, Algorithm 2 is a polynomial-time $(\frac{1}{\alpha}, (\frac{1}{1-\alpha}, \epsilon))$ -bicriteria approximation algorithm for Problem 1.

PROOF. Let $F \subseteq \rho(v)$ be the output of Algorithm 2. The approximation ratio with respect to the size of F can be evaluated as follows:

$$\mathbb{E}[|F|] = \frac{1}{\alpha} \cdot \mathbb{E} \left[\sum_{e \in F} \alpha \right] \leq \frac{1}{\alpha} \cdot \mathbb{E} \left[\sum_{e \in F} x_e^* \right] \leq \frac{1}{\alpha} \sum_{e \in \rho(v)} x_e^* \leq \frac{b}{\alpha},$$

where the first inequality follows from $\alpha \leq p \leq x_e^*$ for any $e \in F$, the second inequality follows from the nonnegativity of x_e^* , and the third inequality follows from the first constraint in Relaxation.

Next we analyze the approximation ratio with respect to the quality of F . Let F^* be an optimal solution to Problem 1. As Relaxation is indeed a relaxation of Problem 1 and x^* is its ϵ' -additive approximate solution, we have $\widehat{f}_{(G,v)}(x^*) \leq f_{(G,v)}(F^*) + \epsilon'$. For convenience, we define $x_{e_0}^* = 1$ for an imaginary element e_0 . Let ℓ be the maximum number that satisfies $x_{e_\ell}^* \geq \alpha$. Then we have

$$\begin{aligned} \mathbb{E}[f_{(G,v)}(F)] &= \frac{\sum_{i=0}^{\ell-1} (x_{e_i}^* - x_{e_{i+1}}^*) f_{(G,v)}(X_i) + (x_{e_\ell}^* - \alpha) f_{(G,v)}(X_\ell)}{1 - \alpha} \\ &\leq \frac{\sum_{i=0}^{|\rho(v)|-1} (x_{e_i}^* - x_{e_{i+1}}^*) f_{(G,v)}(X_i) + x_{e_{|\rho(v)|}}^* f_{(G,v)}(\rho(v))}{1 - \alpha} \\ &= \frac{\widehat{f}_{(G,v)}(x^*)}{1 - \alpha} \leq \frac{f_{(G,v)}(F^*) + \epsilon'}{1 - \alpha} = \frac{f_{(G,v)}(F^*)}{1 - \alpha} + \epsilon, \end{aligned}$$

where the first equality follows from the random choice of p and the first inequality follows from the monotonicity of elements in x^* and the nonnegativity of $f_{(G,v)}$. Therefore, we have the theorem. \square

6 SOLVING RELAXATION

In this section, we present our algorithm for solving Relaxation.

6.1 Algorithm

Specifically, we design a projected subgradient method for Relaxation. The algorithm is an iterative method, where each iteration consists of two parts, i.e., the subgradient computation part and the projection computation part. The pseudo-code is given in Algorithm 3. All the details will be given later. The sequence generated by the algorithm is $\{x_t\}_{t \geq 0}$, while the sequence of function values generated by the algorithm is $\{\widehat{f}_{(G,v)}(x^t)\}_{t \geq 0}$. As the sequence of function values is not necessarily monotone, we are also interested

in the sequence of best-achieved function values at or before ℓ -th iteration, which is defined as

$$\widehat{f}_{\text{best}}^{(\ell)} = \min_{t=0,1,\dots,\ell} \widehat{f}_{(G,v)}(x_t).$$

Subgradient computation. From the definition of $\widehat{f}_{(G,v)}$, a subgradient $\widehat{f}'_{(G,v)}$ at $x_t \in C$ is given by

$$\widehat{f}'_{(G,v)}(x_t) = \sum_{i=1}^{|\rho(v)|} (f_{(G,v)}(X_i) - f_{(G,v)}(X_{i-1})) \mathbf{u}_{e_i}, \quad (2)$$

where \mathbf{u}_{e_i} is the $|\rho(v)|$ -dimensional vector that takes 1 in the element corresponding to e_i and 0 elsewhere. To compute the subgradient $\widehat{f}'_{(G,v)}(x_t)$, we need to sort the entries of x_t and compute $f_{(G,v)}(X_i)$ for all $i = 0, 1, \dots, |\rho(v)|$, which takes $O(|\rho(v)|(|V| + |A|))$ time.

Projection computation. For a given $x \in \mathbb{R}^{\rho(v)}$, it is not trivial how to compute the projection of x onto C because C is the intersection of the two sets $\{x \in \mathbb{R}^{\rho(v)} \mid \|x\|_1 \leq b\}$ and $\{x \in \mathbb{R}^{\rho(v)} \mid x \in [0, 1]^{\rho(v)}\}$. For simplicity, define $\text{Box}[0, 1] = \{x \in \mathbb{R}^{\rho(v)} \mid x \in [0, 1]^{\rho(v)}\}$. Let $\text{proj}_{\text{Box}[0,1]}(x)$ be the projection of x onto $\text{Box}[0, 1]$. Then by Lemma 6.26 in Beck [3], we have

$$\text{proj}_{\text{Box}[0,1]}(x) = (\min\{\max\{x_e, 0\}, 1\})_{e \in \rho(v)}.$$

Using this projection, we can give the projection of x onto C as follows:

Fact 1 (A special case of Example 6.32 in Beck [3]). Let $\text{proj}_C(x)$ be the projection of $x \in \mathbb{R}^{\rho(v)}$ onto C . Then we have

$$\text{proj}_C(x) = \begin{cases} \text{proj}_{\text{Box}[0,1]}(x) & \text{if } \|\text{proj}_{\text{Box}[0,1]}(x)\|_1 \leq b, \\ \text{proj}_{\text{Box}[0,1]}(x - \lambda^* \mathbf{1}) & \text{otherwise,} \end{cases}$$

where λ^* is any positive root of the nonincreasing function $\varphi(\lambda) = \|\text{proj}_{\text{Box}[0,1]}(x - \lambda \mathbf{1})\|_1 - b$.

In practice, we can compute the value of λ^* for $x := x_t - \eta_t \cdot \widehat{f}'_{(G,v)}(x_t)$ using binary search. Assume that the stepsize $\eta_t > 0$ is no greater than 1 for any iteration $t = 0, 1, \dots$, which is indeed the case of ours (specified later). As initial lower and upper bounds on λ^* , we can use 0 and $\max_{e \in \rho(v)} x_e$, respectively. From the fact that x_t is always contained in C and the definition of the subgradient (2), we see that

$$\begin{aligned} \max_{e \in \rho(v)} x_e &\leq \max_{e \in \rho(v)} x_t(e) + \eta_t \cdot \max_{i=1,\dots,|\rho(v)|-1} f_{(G,v)}(X_{i-1}) \\ &\leq 1 + f_{(G,v)}(\emptyset) \leq |V|, \end{aligned}$$

where $x_t(e)$ is the element of x_t corresponding to e . Therefore, the binary search finds λ^* in $O(|\rho(v)| \log(|V|/\delta))$ time with an additive error of $\delta > 0$. Note that any polynomial-time algorithm cannot recognize an additive error of $o(2^{-|V|^c})$ for constant c , due to its bit complexity. Hence, if we set $\delta = O(2^{-|V|^c})$, we can assume that the projection is exact, and the time complexity is still polynomial.

6.2 Convergence result

Let $L_{\widehat{f}_{(G,v)}} = \widehat{f}_{(G,v)}(\emptyset) (= f_{(G,v)}(\emptyset))$. Based on the convergence result of the projected subgradient method in Beck [3], which is reviewed in Appendix C.1, we present the convergence result of Algorithm 3 as follows:

Table 1: Real-world graphs used in our experiments.

Name	$ V $	$ A $	Stat. of in-degrees of targets
moreno-blogs	1,224	19,022	(337, 158.80, 101)
dimacs10-polblogs	1,224	33,430	(274, 143.45, 104)
librec-ciaodvd-trust	4,658	40,133	(361, 152.50, 100)
munmun-twitter-social	465,017	834,797	(174, 123.65, 100)
citeseer	384,054	1,744,590	(495, 190.70, 106)
youtube-links	1,138,494	4,942,297	(1311, 273.40, 100)
higgs-twitter-social	456,626	14,855,819	(1049, 280.20, 111)
soc-pokec-relationships	1,632,803	30,622,564	(316, 157.35, 101)

Theorem 7. Let Θ be an upper bound on the half-squared diameter of C , i.e., $\Theta \geq \max_{\mathbf{x}, \mathbf{y} \in C} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|^2$. Determine the stepsize η_t ($t = 0, 1, \dots$) as $\eta_t = \frac{\sqrt{2\Theta}}{L_{\widehat{f}(G,v)} \sqrt{t+1}}$. Let \widehat{f}^* be the optimal value of Relaxation.

Then for all $t \geq 2$, it holds that

$$\widehat{f}_{\text{best}}^{(t)} - \widehat{f}^* \leq \frac{2(1 + \log 3)L_{\widehat{f}(G,v)} \sqrt{2\Theta}}{\sqrt{t+2}}.$$

The proof can be found in Appendix C.2. By this theorem and the above discussion of the time complexity, the following is almost straightforward:

Corollary 2. Let $\epsilon' > 0$. Set the stopping condition of Algorithm 3 as follows:

$$t \geq \left(\frac{2(1 + \log 3)L_{\widehat{f}(G,v)} \sqrt{2\Theta}}{\epsilon'} \right)^2 - 2.$$

Then, Algorithm 3 outputs, in polynomial time, an ϵ' -additive approximate solution for Relaxation.

7 EXPERIMENTAL EVALUATION

In this section, we thoroughly evaluate the performance of our proposed algorithms (i.e., Algorithms 1 and 2) using various real-world networks.

7.1 Setup

Instances. Table 1 lists real-world digraphs on which our experiments were conducted. All graphs were collected from the webpage of The KONECT Project¹. Note that self-loops and multiple edges were removed so that the graphs are made simple. For each graph, we randomly chose 20 vertices as target vertices among those having the in-degree at least 100. The last column of Table 1 gives the statistics of the in-degrees of the target vertices, i.e., the maximum, average, and minimum in-degrees. For each graph and each target vertex v , we vary the budget b in $\{\lfloor \frac{1}{4}|\rho(v)| \rfloor, \lfloor \frac{1}{2}|\rho(v)| \rfloor, \lfloor \frac{3}{4}|\rho(v)| \rfloor\}$.

Baselines. We employ the following baseline methods:

- Empty: This algorithm just outputs the empty set, thus presenting an upper bound on the objective function value (i.e., $h_C(v)$) of any feasible solution.
- Random: This algorithm randomly chooses b edges from $\rho(v)$. For each instance, this algorithm is run 100 times and the average objective value is reported.

¹<http://konect.cc/>

Table 2: Computation time (seconds) of the algorithms tested.

Name	b	Greedy	Algorithm 1	Algorithm 2
moreno-blogs	$\lfloor \frac{1}{4} \rho(v) \rfloor$	4.53	0.10	326.55
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	7.79	0.10	333.48
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	9.75	0.10	332.48
dimacs10-polblogs	$\lfloor \frac{1}{4} \rho(v) \rfloor$	4.40	0.12	393.72
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	7.59	0.13	416.21
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	9.55	0.13	404.52
librec-ciaodvd-trust	$\lfloor \frac{1}{4} \rho(v) \rfloor$	6.17	0.15	482.83
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	10.59	0.15	502.63
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	13.33	0.15	489.92
munmun-twitter-social	$\lfloor \frac{1}{4} \rho(v) \rfloor$	4.05	0.14	—
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	7.03	0.14	—
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	8.90	0.14	—
citeseer	$\lfloor \frac{1}{4} \rho(v) \rfloor$	—	4.58	—
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	—	4.57	—
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	—	4.57	—
youtube-links	$\lfloor \frac{1}{4} \rho(v) \rfloor$	—	130.56	—
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	—	130.54	—
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	—	132.19	—
higgs-twitter-social	$\lfloor \frac{1}{4} \rho(v) \rfloor$	—	188.84	—
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	—	187.92	—
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	—	187.68	—
soc-pokec-relationships	$\lfloor \frac{1}{4} \rho(v) \rfloor$	—	402.24	—
	$\lfloor \frac{1}{2} \rho(v) \rfloor$	—	392.63	—
	$\lfloor \frac{3}{4} \rho(v) \rfloor$	—	512.05	—

- Degree: This algorithm sorts the elements of $N_{\text{in}}(v)$ as $(w_1, \dots, w_{|\rho(v)|})$ so that $|\rho(w_1)| \geq \dots \geq |\rho(w_{|\rho(v)|})|$ and just returns $\{(w_1, v), \dots, (w_b, v)\}$.
- Greedy: Execute the greedy algorithm (Algorithm 4).

Machine specs and code. All experiments were conducted on Mac mini with Apple M1 Chip and 16 GB RAM. All codes were written in Python 3.9.

7.2 Performance of algorithms

Here we evaluate the performance of our algorithms. To this end, we run the algorithms together with the baselines for all graphs and budgets. For each graph and each budget, if the algorithm tested does not terminate within 1,200 seconds for the target vertex having the largest in-degree, we do no longer run the algorithm for the graph and the other budgets. Note that Algorithm 3 (in Algorithm 2) is run with stopping condition $t \geq 1000$ for scalability and initial solution $\mathbf{x}_0 = \mathbf{0}$.

The quality of solutions of the algorithms except for Algorithm 2 is illustrated in Figure 1. Due to space limitations, only the results for the budget $b = \lfloor \frac{1}{2}|\rho(v)| \rfloor$ are presented here. Although the trend is similar, the results for the other budget settings $b = \lfloor \frac{1}{4}|\rho(v)| \rfloor, \lfloor \frac{3}{4}|\rho(v)| \rfloor$ are given in Appendix D.1. As the solutions of Algorithm 2 may violate the budget constraint, it is unfair to compare those with the others in the same plots; thus, the solutions are evaluated later. In the plots in Figure 1, once we fix a value in the vertical axis, we can observe the cumulative number of solutions (i.e., targets) that attain the harmonic centrality score at most the fixed value. Therefore, we can say that an algorithm drawing a lower line has a better performance.

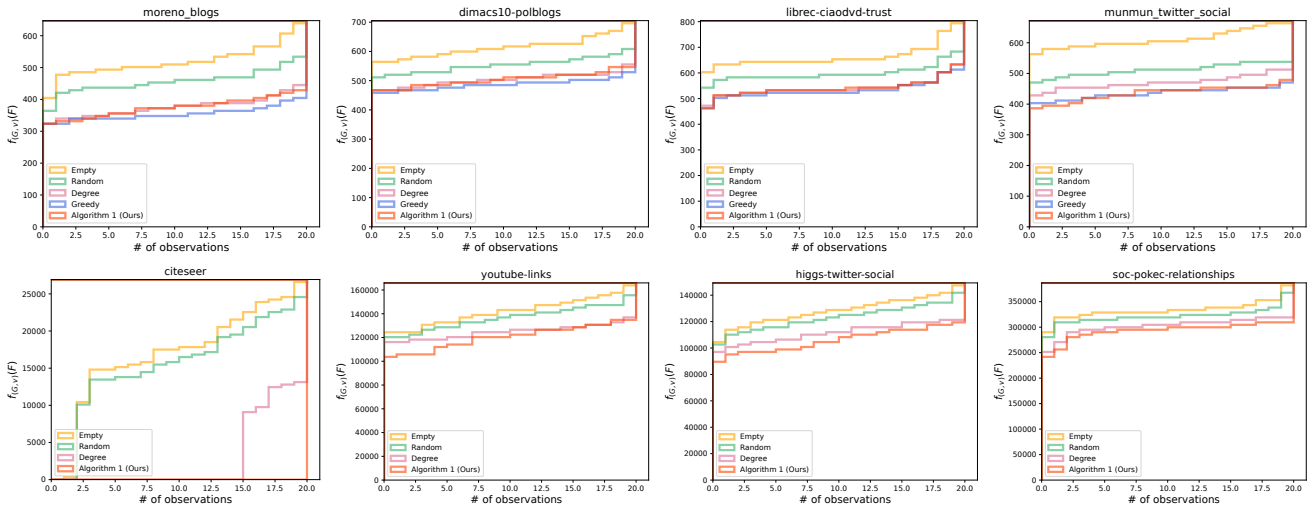


Figure 1: Quality of solutions of the algorithms (except for Algorithm 2) with $b = \lfloor \frac{1}{2} |\rho(v)| \rfloor$.

Table 3: Quality of solutions of Algorithm 2 with $\alpha = \frac{1}{3}, \frac{1}{2}$ compared with those of Algorithm 1 with $b = \lfloor \frac{1}{4} |\rho(v)| \rfloor$.

Name	$\alpha = \frac{1}{3}$		$\alpha = \frac{1}{2}$	
	Obj. val.	Size	Obj. val.	Size
moreno-blogs	1.12	0.14	1.13	0.12
dimacs10-polblogs	1.09	0.10	1.09	0.07
librec-ciaodvd-trust	1.09	0.06	1.09	0.03

As can be seen, Algorithm 1 outperforms the baselines. Indeed, Algorithm 1 is applicable to all graphs tested, thanks to its high scalability, and the quality of solutions is much better than that of the scalable baselines, i.e., Random and Degree. Most interestingly, for the graph *citeseer*, Algorithm 1 succeeds in reducing the harmonic centrality scores of all target vertices to the values relatively close to 0, although Degree fails to have centrality scores less than 9,000 for five target vertices. Note that this result does not contradict the fact that even an optimal solution has the harmonic centrality score no less than $|\rho(v)| - b$: The minimum objective value attained by Algorithm 1 is 54, rather than 0. For small graphs, Greedy performs slightly better than Algorithm 1; however, due to its heavy computation burden, Greedy is not applicable to *citeseer* and the larger graphs.

The detailed report of the computation time is found in Table 2, where the average computation time over all target vertices is reported. Note that the results for Random and Degree are omitted because Random is obviously quite fast and Degree records 0.00 seconds for all graphs and budgets. We remark that the computation time of Greedy grows roughly proportionally to the budget b , but that of Algorithm 1 remains almost the same for all settings of b .

Finally, the quality of solutions of Algorithm 2 (with $\alpha = \frac{1}{3}, \frac{1}{2}$), averaged over the target vertices, is reported in Table 3, where each of the objective value and the size of the solutions is a relative value compared with that of the solutions of Algorithm 1. As the rounding procedure of the algorithm contains randomness, we run

the procedure 100 times and report the average value. As can be seen, Algorithm 2 returns very small solutions that even do not exhaust the budget. However, this result does not contradict the guarantee given in Theorem 6; the objective value is not much worse than that of Algorithm 1 and the (expected) size of solutions is just upper bounded in Theorem 6.

To further examine the performance of Algorithm 2, we run it on the instance that appeared in the proof of Theorem 5, where we set $k = 50$ (and $b = 50$). Note that it is theoretically guaranteed that Algorithm 1 outputs a poor solution to the instance. On the other hand, Algorithm 2 with $\alpha = \frac{3}{4}$ always obtains the optimal solution among 100 rounding trials. Hence, we see that the algorithm is rather strong against adversarial instances, verifying the effectiveness of the theoretical approximation guarantee of Algorithm 2.

8 CONCLUSION

In this study, we have introduced a novel optimization model for local centrality minimization and designed two effective approximation algorithms. The first algorithm (Algorithm 1) is a highly-scalable $\sqrt{2}h_G(v)$ -approximation algorithm. We stress that this approximation ratio is unachievable by the most intuitive greedy algorithm (Algorithm 4). The second algorithm (Algorithm 2) is a polynomial-time $(\frac{1}{\alpha}, (\frac{1}{1-\alpha}, \epsilon))$ -bicriteria approximation algorithm. To the best of our knowledge, ours are the first polynomial-time algorithms with provable approximation guarantees for centrality minimization. Experiments using a variety of real-world networks demonstrate the effectiveness of our proposed algorithms.

Our work opens up several interesting problems. Can we design a polynomial-time algorithm that has an approximation ratio better than that of Algorithm 1 or a bicriteria approximation ratio better than that of Algorithm 2? Another interesting direction is to study Problem 1 with a more capable setting. For example, it would be valuable to have a target vertex subset (rather than a single target vertex) and aim to minimize its group harmonic centrality score, as in the literature on global centrality minimization [38].

REFERENCES

- [1] E. Angriman, R. Becker, G. D'Angelo, H. Gilbert, A. van der Grinten, and H. Meyerhenke. Group-harmonic and group-closeness maximization – Approximation and engineering. In *Proceedings of ALENEX '21*, pages 154–168, 2021.
- [2] E. Angriman, A. van der Grinten, A. Bojchevski, D. Zügner, S. Günnemann, and H. Meyerhenke. Group centrality maximization for large-scale graphs. In *Proceedings of ALENEX '20*, pages 56–69, 2020.
- [3] A. Beck. *First-Order Methods in Optimization*. MOS–SIAM Series on Optimization. SIAM, 2017.
- [4] E. Bergamini, P. Crescenzi, G. D'angelo, H. Meyerhenke, L. Severini, and Y. Velaj. Improving the betweenness centrality of a node by adding links. *ACM Journal of Experimental Algorithmics*, 23:1–32, 2018.
- [5] K. Bhawalkar, J. Kleinberg, K. Lewi, T. Roughgarden, and A. Sharma. Preventing unraveling in social networks: The anchored k -core problem. *SIAM Journal on Discrete Mathematics*, 29(3):1452–1475, 2015.
- [6] P. Boldi and S. Vigna. Axioms for centrality. *Internet Mathematics*, 10(3–4):222–262, 2014.
- [7] M. Castaldo, C. Catalano, G. Como, and F. Fagnani. On a centrality maximization game. *IFAC-PapersOnLine*, 53(2):2844–2849, 2020.
- [8] C. Chen, W. Wang, and X. Wang. Efficient maximum closeness centrality group identification. In *Proceedings of the 27th Australasian Database Conference*, pages 43–55, 2016.
- [9] C. Coupette, S. Neumann, and A. Gionis. Reducing exposure to harmful content via graph rewiring. In *Proceedings of KDD '23*, pages 323–334, 2023.
- [10] P. Crescenzi, G. D'angelo, L. Severini, and Y. Velaj. Greedily improving our own closeness centrality in a network. *ACM Transactions on Knowledge Discovery from Data*, 11(1):1–32, 2016.
- [11] G. D'Angelo, M. Olsen, and L. Severini. Coverage centrality maximization in undirected networks. In *Proceedings of AAAI '19*, pages 501–508, 2019.
- [12] K. Das, S. Samanta, and M. Pal. Study on centrality measures in social networks: A survey. *Social Network Analysis and Mining*, 8:1–11, 2018.
- [13] F. Fabbri, Y. Wang, F. Bonchi, C. Castillo, and M. Mathioudakis. Rewiring what-to-watch-next recommendations to reduce radicalization pathways. In *Proceedings of The Web Conference '22*, page 2719–2728, 2022.
- [14] S. Fujishige. *Submodular Functions and Optimization*, volume 58 of *Annals of Discrete Mathematics*. Elsevier, 2005.
- [15] J. Goldenberg, B. Libai, and E. Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 12:211–223, 2001.
- [16] J. Goldenberg, B. Libai, and E. Muller. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Academy of Marketing Science Review*, 9(3):1–18, 2001.
- [17] A. Hayrapetyan, D. Kempe, M. Pál, and Z. Svitkina. Unbalanced graph cuts. In *Proceedings of ESA '05*, pages 191–202, 2005.
- [18] V. Ishakian, D. Erdős, E. Terzi, and A. Bestavros. A framework for the evaluation and management of network centrality. In *Proceedings of SDM '12*, pages 427–438, 2012.
- [19] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of KDD '03*, pages 137–146, 2003.
- [20] E. Khalil, B. Dilkina, and L. Song. Cuttingedge: Influence minimization in networks. In *Proceedings of Workshop on Frontiers of Network Analysis: Methods, Models, and Applications*, pages 1–13, 2013.
- [21] D. Knoke and S. Yang. *Social Network Analysis*, volume 154 of *Quantitative Applications in the Social Sciences*. SAGE Publications, 2020.
- [22] A. Konar and N. D. Sidiropoulos. Exploring the subgraph density-size trade-off via the Lovász extension. In *Proceedings of WSDM '21*, pages 743–751, 2021.
- [23] A. Konar and N. D. Sidiropoulos. The triangle-densest- k -subgraph problem: Hardness, Lovász extension, and application to document summarization. In *Proceedings of AAAI '22*, pages 4075–4082, 2022.
- [24] H. Li, R. Peng, L. Shan, Y. Yi, and Z. Zhang. Current flow group closeness centrality for complex networks? In *Proceedings of The Web Conference '19*, pages 961–971, 2019.
- [25] C. Liu, X. Zhou, A. N. Zehmakan, and Z. Zhang. A fast algorithm for moderating critical nodes via edge removal. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–14, 2023. In press.
- [26] L. Lovász. Submodular functions and convexity. In A. Bachem, B. Korte, and M. Grötschel, editors, *Mathematical Programming: The State of the Art*, pages 235–257. Springer, 1983.
- [27] A. Mahmoody, C. E. Tsourakakis, and E. Upfal. Scalable betweenness centrality maximization via sampling. In *Proceedings of KDD '16*, pages 1765–1773, 2016.
- [28] A. Maulana and M. Atzmueller. Many-objective optimization for anomaly detection on multi-layer complex interaction networks. *Applied Sciences*, 11(9):4005, 2021.
- [29] S. Medya, A. Silva, A. Singh, P. Basu, and A. Swami. Group centrality maximization via network design. In *Proceedings of SDM '18*, pages 126–134, 2018.
- [30] M. Minoux. Accelerated greedy algorithms for maximizing submodular set functions. In *Proceedings of the 8th IFIP Conference on Optimization Techniques*, pages 234–243, 2005.
- [31] S. Mumtaz and X. Wang. Identifying top- k influential nodes in networks. In *Proceedings of CIKM '17*, pages 2219–2222, 2017.
- [32] S. Murai and Y. Yoshida. Sensitivity analysis of centralities on unweighted networks. In *Proceedings of The Web Conference '19*, pages 1332–1342, 2019.
- [33] G. L. Nemhauser and L. A. Wolsey. Best algorithms for approximating the maximum of a submodular set function. *Mathematics of Operations Research*, 3(3):177–188, 1978.
- [34] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69:026113, 2004.
- [35] L. Pellegrina. Efficient centrality maximization with Rademacher averages. In *Proceedings of KDD '23*, pages 1872–1884, 2023.
- [36] S. S. Rangapuram, T. Bühler, and M. Hein. Towards realistic team formation in social networks based on densest subgraphs. In *Proceedings of WWW '13*, pages 1077–1088, 2013.
- [37] Z. Svitkina and L. Fleischer. Submodular approximation: Sampling-based algorithms and lower bounds. *SIAM Journal on Computing*, 40(6):1715–1737, 2011.
- [38] A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao. Finding critical links for closeness centrality. *INFORMS Journal on Computing*, 31(2):367–389, 2019.
- [39] S. A. Vinterbo. Privacy: A machine learning view. *IEEE Transactions on Knowledge and Data Engineering*, 16(8):939–948, 2004.
- [40] X. Wang, K. Deng, J. Li, J. X. Yu, C. S. Jensen, and X. Yang. Efficient targeted influence minimization in big social networks. *World Wide Web*, 23(4):2323–2340, 2020.
- [41] M. Waniek, T. P. Michalak, M. J. Wooldridge, and T. Rahwan. Hiding individuals and communities in a social network. *Nature Human Behaviour*, 2(2):139, 2018.
- [42] M. Waniek, J. Woźnica, K. Zhou, Y. Vorobeychik, T. P. Michalak, and T. Rahwan. Hiding from centrality measures: A Stackelberg game perspective. *IEEE Transactions on Knowledge and Data Engineering*, 2023. In press.
- [43] L. Yang, Z. Li, and A. Gíua. Influence minimization in linear threshold networks. *Automatica*, 100:10–16, 2019.
- [44] F. Zhang, Y. Zhang, L. Qin, W. Zhang, and X. Lin. Finding critical users for social network engagement: The collapsed k -core problem. In *Proceedings of AAAI '17*, 2017.
- [45] P. Zhao, Y. Li, H. Xie, Z. Wu, Y. Xu, and J. C. Lui. Measuring and maximizing influence via random walk in social activity networks. In *Proceedings of DASFAA '17*, pages 323–338, 2017.
- [46] W. Zhu, C. Chen, X. Wang, and X. Lin. K -core minimization: An edge manipulation approach. In *Proceedings of CIKM '18*, pages 1667–1670, 2018.

A OMITTED CONTENTS FROM SECTION 3

A.1 Proof of Theorem 1

To prove the theorem, we formally introduce the following problem:

Problem 2 (Minimum k -union). *Given a ground set $U = \{e_1, \dots, e_n\}$, a set system $\mathcal{S} = \{S_1, \dots, S_m\} \subseteq 2^U$, and $k \in \mathbb{Z}_{>0}$, we are asked to find $J \subseteq \{1, \dots, m\}$ with $|J| = k$ that minimizes $|\bigcup_{j \in J} S_j|$. Without loss of generality, it can be assumed that $\bigcup_{j \in \{1, \dots, m\}} S_j = U$.*

Fact 2 (Theorem 1 in Vinterbo [39]). *Problem 2 is NP-hard.*

PROOF OF THEOREM 1. We construct a polynomial-time reduction from Problem 2 to Problem 1. Take an arbitrary instance of Problem 2: $U = \{e_1, \dots, e_n\}$, $\mathcal{S} = \{S_1, \dots, S_m\} \subseteq 2^U$, and $k \in \mathbb{Z}_{>0}$. We make an instance (i.e., a gadget) of Problem 1, i.e., $G = (V, A)$, $v \in V$, and $b \in \mathbb{Z}_{>0}$, as follows:

- $G = (V, A)$ such that
 - $V = \{v_0, v_{S_1}, \dots, v_{S_m}, v_{e_1}, \dots, v_{e_n}\}$,
 - $A = \{(v_{S_j}, v_0) \mid j = 1, \dots, m\} \cup \{(v_{e_i}, v_{S_j}) \mid e_i \in S_j, i = 1, \dots, n, j = 1, \dots, m\}$;
- $v = v_0$;
- $b = m - k$.

Clearly, G is acyclic. Let $F^* \subseteq \rho(v_0) (= \{(v_{S_j}, v_0) \mid j = 1, \dots, m\})$ be an optimal solution to this gadget, and let $J^* \subseteq \{1, \dots, m\}$ be the index set that satisfies $F^* = \{(v_{S_j}, v_0) \mid j \in J^*\}$. We see that $|J^*| = b$. Let $\bar{J}^* = \{1, \dots, m\} \setminus J^*$. Then, the objective value of F^* , i.e., the harmonic centrality of v_0 after the removal of F^* , can be evaluated as $(m - b) + \frac{1}{2} \left| \bigcup_{j \in \bar{J}^*} S_j \right|$. Therefore, as F^* is optimal, $\left| \bigcup_{j \in \bar{J}^*} S_j \right|$ is minimized. Noticing that $|\bar{J}^*| = m - b = m - (m - k) = k$, we see that \bar{J}^* is an optimal solution to the instance of Problem 2. \square

A.2 Proof of Theorem 2

Before proving the theorem, we mention a fundamental fact of the submodularity. It is well known that the submodularity is equivalent to the diminishing marginal return property [14], which can be written as follows: For any $X \subset Y \subseteq S$ and any $p \in S \setminus Y$, it holds that

$$f(X \cup \{p\}) - f(X) \geq f(Y \cup \{p\}) - f(Y).$$

PROOF OF THEOREM 2. Let $G = (V, A)$ be any digraph and $v \in V$ be any vertex. From the above fact, it suffices to show that for any $E \subset F \subseteq \rho(v)$ and $e \in \rho(v) \setminus F$,

$$f_{(G,v)}(E \cup \{e\}) - f_{(G,v)}(E) \geq f_{(G,v)}(F \cup \{e\}) - f_{(G,v)}(F)$$

holds. From the definition of $f_{(G,v)}$, we have

$$\begin{aligned} & f_{(G,v)}(E \cup \{e\}) - f_{(G,v)}(E) \\ &= h_{G \setminus (E \cup \{e\})}(v) - h_{G \setminus E}(v) \\ &= \sum_{u \in V \setminus \{v\}} \left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus E}(u, v)} \right). \end{aligned}$$

For each $e \in \rho(v)$, we define

$$U_G(v, e) = \{u \in V \setminus \{v\} \mid$$

there exists a path from u to v on G and

all shortest paths from u to v on G contain $e\}$.

Noticing that $d_{G \setminus (E \cup \{e\})}(u, v) = d_{G \setminus E}(u, v)$ for all $u \in V \setminus U_{G \setminus E}(v, e)$, we have

$$\begin{aligned} & \sum_{u \in V \setminus \{v\}} \left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus E}(u, v)} \right) \\ &= \sum_{u \in U_{G \setminus E}(v, e)} \left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus E}(u, v)} \right). \end{aligned}$$

Applying a similar argument, we also have

$$\begin{aligned} & f_{(G,v)}(F \cup \{e\}) - f_{(G,v)}(F) \\ &= \sum_{u \in U_{G \setminus F}(v, e)} \left(\frac{1}{d_{G \setminus (F \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus F}(u, v)} \right). \end{aligned}$$

Combining the above two equalities, we have

$$\begin{aligned} & \left(f_{(G,v)}(E \cup \{e\}) - f_{(G,v)}(E) \right) - \left(f_{(G,v)}(F \cup \{e\}) - f_{(G,v)}(F) \right) \\ &= \sum_{u \in U_{G \setminus E}(v, e)} \left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus E}(u, v)} \right) \\ &\quad - \sum_{u \in U_{G \setminus F}(v, e)} \left(\frac{1}{d_{G \setminus (F \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus F}(u, v)} \right) \\ &\geq \sum_{u \in U_{G \setminus E}(v, e)} \left(\left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus E}(u, v)} \right) \right. \\ &\quad \left. - \left(\frac{1}{d_{G \setminus (F \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus F}(u, v)} \right) \right) \\ &= \sum_{u \in U_{G \setminus E}(v, e)} \left(\frac{1}{d_{G \setminus (E \cup \{e\})}(u, v)} - \frac{1}{d_{G \setminus (F \cup \{e\})}(u, v)} \right) \geq 0, \end{aligned}$$

where the first inequality follows from $U_{G \setminus E}(v, e) \subseteq U_{G \setminus F}(v, e)$ and $d_{G \setminus (F \cup \{e\})}(u, v) \geq d_{G \setminus F}(u, v)$ for any $u \in V$, and the second equality follows from $d_{G \setminus E}(u, v) = d_{G \setminus F}(u, v)$ for any $u \in U_{G \setminus E}(v, e)$. Therefore, we have the theorem. \square

We wish to remark that the above theorem can also be proved using the fact that the objective function of the harmonic centrality maximization problem introduced in Crescenzi et al. [10] is submodular. However, if we employed such an approach, the notation would be more complex. Therefore, we proved it from scratch, based on the definition of the submodularity.

A.3 Pseudo-code of the greedy algorithm

See Algorithm 4.

Algorithm 4: Greedy algorithm for Problem 1

Input : $G = (V, A)$, $v \in V$, and $b \in \mathbb{Z}_{>0}$

Output: $F \subseteq \rho(v)$ with $|F| \leq b$

- 1 $F \leftarrow \emptyset$;
 - 2 **while** $|F| < b$ **do**
 - 3 Find $e \in \operatorname{argmin}\{f_{(G,v)}(e) \mid e \in \rho(v)\}$;
 - 4 $F \leftarrow F \cup \{e\}$ and $G \leftarrow G \setminus \{e\}$;
 - 5 **return** F ;
-

A.4 Proof of Theorem 3

PROOF. We first show that any algorithm that outputs $F \subseteq \rho(v)$ with $|F| = b$ has an approximation ratio of $O(|V|)$. Let $G = (V, A)$, $v \in V$, $b \in \mathbb{Z}_{>0}$ be an arbitrary instance of Problem 1. If the optimal value is equal to 0, then $b = |\rho(v)|$ holds, which means that the above algorithm can also achieve the optimal value. If the optimal value is greater than 0, then it is at least 1, while any algorithm can achieve the objective value at most $h_G(v) \leq |V| - 1$. Therefore, the above algorithm has an approximation ratio of $O(|V|)$.

Next we show that the greedy algorithm has no approximation ratio of $o(|V|)$. To this end, for any integer $k \geq 2$, we construct an instance of Problem 1 as follows: Let $G = (V, A)$ be a digraph. The vertex set V consists of two vertices n_L, o_L , and size- k subsets $N_R = \{n_R^1, \dots, n_R^k\}$ and $O_R = \{o_R^1, \dots, o_R^k\}$, in addition to one vertex v' . The edge set A is defined as the union of the following three sets:

$$\begin{aligned} & \{(n_L, v'), (n_R^i, v') \mid i = 1, \dots, k\}, \\ & \{(o_L, n_L)\}, \\ & \{(o_R^i, n_R^j) \mid i = 1, \dots, k, j = 1, \dots, k\}. \end{aligned}$$

Then we employ v' as target vertex v and k as budget b .

From now on, we analyze the performance of Algorithm 4 for this instance. In the first iteration, the removal of (n_L, v) decreases the objective value by $3/2$, while the removal of any incoming edge from N_R decreases the objective value by 1. Therefore, the algorithm removes (n_L, v) . In the later iterations, the removal of any edge decreases the objective value by 1, and thus the algorithm removes $k - 1$ edges from N_R to v and then terminates. We see that the objective value of the output is $1 + k/2$. On the other hand, consider the following feasible (actually optimal) solution to the instance, i.e., all k edges from N_R to v . Then the objective value of this solution is $1 + 1/2$. Thus, the approximation ratio of the algorithm is lower bounded by

$$\frac{1 + k/2}{1 + 1/2} = \Omega(|V|),$$

meaning that Algorithm 4 has no approximation ratio of $o(|V|)$ for Problem 1. \square

B OMITTED CONTENTS FROM SECTION 4

B.1 Proof of Theorem 5

PROOF. For any integer $k \geq 2$, we construct an instance of Problem 1 as follows: Let $G = (V, A)$ be a digraph. The vertex set V consists of size- k subsets $N_L = \{n_L^1, \dots, n_L^k\}$, $N_R = \{n_R^1, \dots, n_R^k\}$, $O_R = \{o_R^1, \dots, o_R^k\}$ and size- $(k(k-1))$ subset $O_L = \{o_L^1, \dots, o_L^{k(k-1)}\}$, in addition to one vertex v' . The edge set A is defined as the union of the following three sets:

$$\begin{aligned} & \{(n_L^i, v'), (n_R^i, v') \mid i = 1, \dots, k\}, \\ & \{(o_L^i, n_L^j) \mid i = 1, \dots, k(k-1), j = 1, \dots, k, \lceil i/(k-1) \rceil = j\}, \\ & \{(o_R^i, n_R^j) \mid i = 1, \dots, k, j = 1, \dots, k\}. \end{aligned}$$

Then we employ v' as target vertex v and k as budget b .

From now on, we analyze the performance of Algorithm 1 for this instance. For any $n_L^i \in N_L$ ($i = 1, \dots, k$), we have $h_{G \setminus \rho(v)}(n_L^i) =$

Algorithm 5: Projected subgradient method

Input : $x_0 \in C$ and some stopping condition

Output: $x \in C$

```

1 while the stopping condition is not satisfied do
2   Pick stepsize  $\eta_t > 0$  and subgradient  $f'(\mathbf{x}_t)$  of  $f'$  at  $\mathbf{x}_t$ ;
3    $\mathbf{x}_{t+1} \leftarrow \text{proj}_C(\mathbf{x}_t - \eta_t \cdot f'(\mathbf{x}_t))$  and  $t \leftarrow t + 1$ ;
4 return  $\mathbf{x}_t$ ;

```

$k - 1$. For any $n_R^i \in N_R$ ($i = 1, \dots, k$), we have $h_{G \setminus \rho(v)}(n_R^i) = k$. Therefore, the output of Algorithm 1 is the set of all k edges from N_R to v , which has an objective value of $k + k(k-1)/2$. On the other hand, consider the following feasible (actually optimal) solution to the instance, i.e., the set of all k edges from N_L to v . Then the objective value of this solution is $k + k/2$. Thus, the approximation ratio of the algorithm is lower bounded by

$$\frac{k + k(k-1)/2}{k + k/2} = \frac{1 + (k-1)/2}{1 + 1/2} = \Omega(\sqrt{|V|}),$$

meaning that Algorithm 1 has no approximation ratio of $o(\sqrt{|V|})$ for Problem 1. Noticing that $h_G(v) \leq |V| - 1$, we have the statement. \square

C OMITTED CONTENTS FROM SECTION 6

C.1 Convergence result in Beck [3]

For self-containedness, we review the convergence result of the projected subgradient method in Beck [3], on which the convergence result of Algorithm 3 is based. Note that for simplicity, the description is appropriately specialized to our setting: For example, we consider the objective function f on \mathbb{R}^n rather than a general vector space.

Let $f: \mathbb{R}^n \rightarrow (-\infty, \infty]$ and $C \subseteq \mathbb{R}^n$. Consider the following problem:

$$\begin{aligned} \text{(P): minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in C. \end{aligned}$$

Assumption 1 (Assumption 8.7 in Beck [3]). *The following hold:*

- (1) $f: \mathbb{R}^n \rightarrow (-\infty, \infty]$ is proper closed and convex;
- (2) $C \subseteq \mathbb{R}^n$ is nonempty closed and convex;
- (3) The interior of the (effective) domain of f contains C ;
- (4) The set of optimal solutions is nonempty, and the optimal value is denoted by f^* .

We consider the projected subgradient method for (P), which is summarized in Algorithm 5. Note that we denote by $\text{proj}_C(\mathbf{x})$ the projection of $\mathbf{x} \in \mathbb{R}^n$ onto C . The sequence generated by the algorithm is $\{\mathbf{x}_t\}_{t \geq 0}$, while the sequence of function values generated by the algorithm is $\{f(\mathbf{x}^t)\}_{t \geq 0}$. As the sequence of function values is not necessarily monotone, we are also interested in the sequence of best-achieved function values at or before ℓ -th iteration, which is defined as

$$f_{\text{best}}^\ell = \min_{t=0,1,\dots,\ell} f(\mathbf{x}_t).$$

If C is compact, then there exists a constant $L_f > 0$ for which $\|f'(\mathbf{x})\| \leq L_f$ for all $f'(\mathbf{x}) \in \partial f(\mathbf{x})$ and $\mathbf{x} \in C$, where $\partial f(\mathbf{x})$ is

the subdifferential of f at \mathbf{x} . Based on this, we have the following convergence result of Algorithm 5:

Fact 3 (A special case of Theorem 8.30 in Beck [3]). *Suppose that Assumption 1 holds and assume that C is compact. Let Θ be an upper bound on the half-squared diameter of C , i.e., $\Theta \geq \max_{\mathbf{x}, \mathbf{y} \in C} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|^2$. Determine the stepsize η_t ($t = 0, 1, \dots$) as $\eta_t = \frac{\sqrt{2\Theta}}{L_f \sqrt{t+1}}$. Then for all $t \geq 2$, it holds that*

$$f_{\text{best}}^t - f^* \leq \frac{2(1 + \log 3)L_f \sqrt{2\Theta}}{\sqrt{t+2}}.$$

C.2 Proof of Theorem 7

PROOF. The proof is almost straightforward from Fact 3 in Appendix C.1. By defining $\widehat{f}_{(G,v)}(\mathbf{x}) = \infty$ for any $\mathbf{x} \in \mathbb{R}^{\rho(v)} \setminus C$, we have an extended real-valued function $\widehat{f}_{(G,v)}: \mathbb{R}^{\rho(v)} \rightarrow (-\infty, \infty]$. Then Relaxation becomes a special case of (P) in Appendix C.1. Let us confirm the validity of Assumption 1 in Appendix C.1 for Relaxation:

- (1) From the definition of $\widehat{f}_{(G,v)}$ (and $f_{(G,v)}$), $\widehat{f}_{(G,v)}$ does not attain $-\infty$ and $\widehat{f}(\mathbf{x}) < \infty$ for any $\mathbf{x} \in C \neq \emptyset$, which means that $\widehat{f}_{(G,v)}$ is proper. As $\widehat{f}_{(G,v)}$ is continuous over C and the

(effective) domain of $\widehat{f}_{(G,v)}$ (i.e., C) is closed, by Theorem 2.8 in Beck [3], $\widehat{f}_{(G,v)}$ is closed. As mentioned above, $\widehat{f}_{(G,v)}$ is convex by the submodularity of $f_{(G,v)}$.

- (2) This is trivial.
- (3) As the boundary of C is not contained in the interior of the (effective) domain of $\widehat{f}_{(G,v)}$, this does not hold. However, according to the analysis in Beck [3], this assumption is used only for guaranteeing the subdifferentiability of the objective function over the feasible region (i.e., the subdifferentiability of $\widehat{f}_{(G,v)}$ over C in our case), which is clearly valid.
- (4) From the boundedness of C and the piecewise-linearity of $\widehat{f}_{(G,v)}$, we see that the set of optimal solutions to Relaxation is nonempty.

Finally, C is obviously compact. Therefore, we have the theorem. \square

D OMITTED CONTENTS FROM SECTION 7

D.1 Counterparts of Figure 1

See Figures 2 and 3.

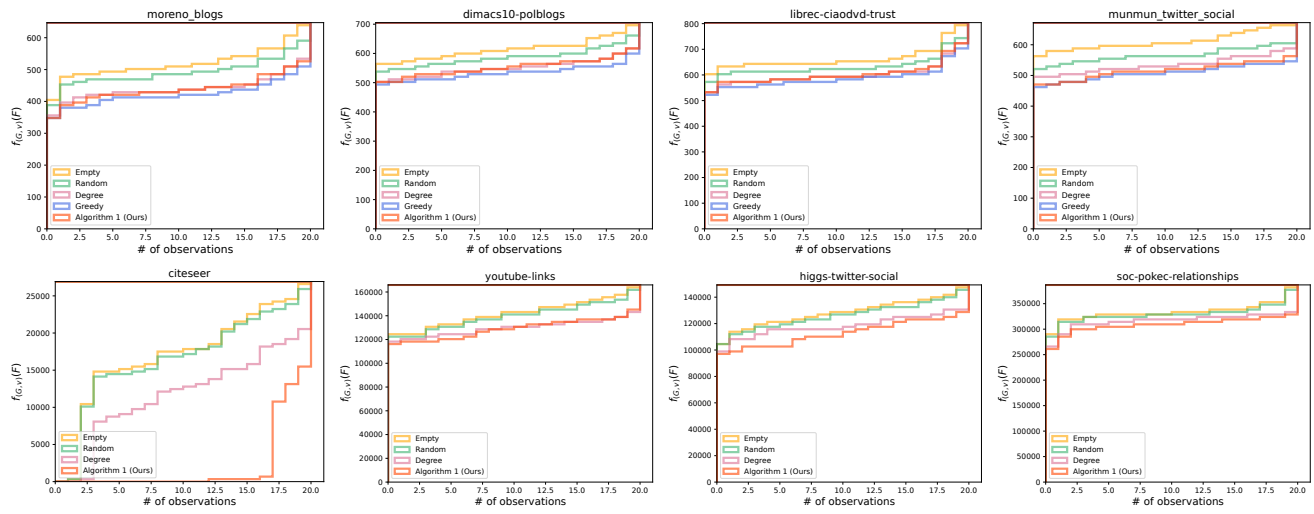


Figure 2: Quality of solutions of the algorithms (except for Algorithm 2) with $b = \lfloor \frac{1}{4} |\rho(v)| \rfloor$.

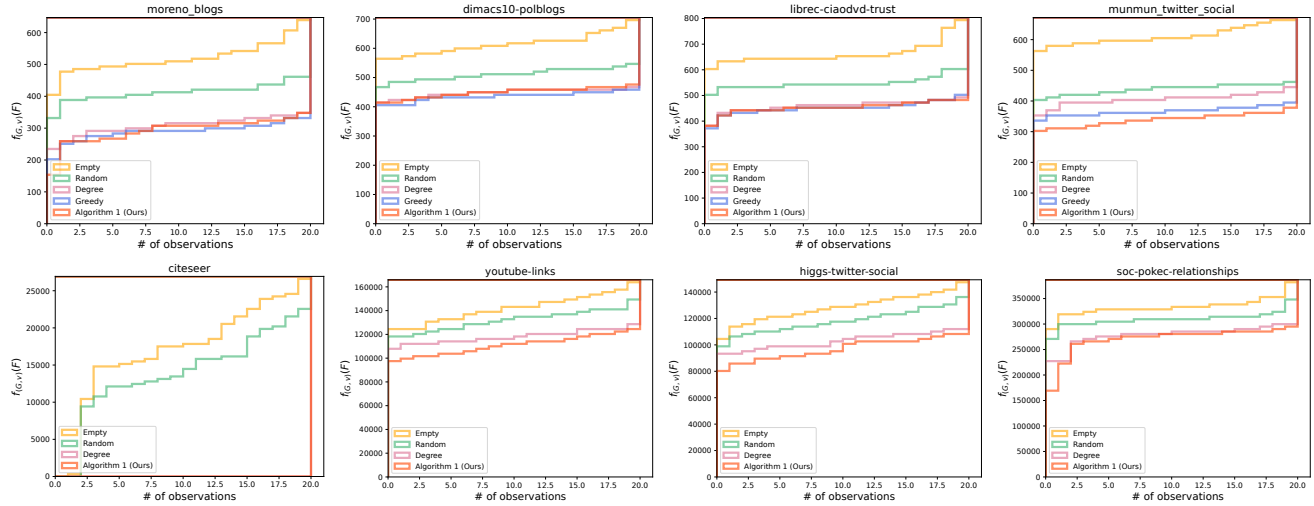


Figure 3: Quality of solutions of the algorithms (except for Algorithm 2) with $b = \lfloor \frac{3}{4} |\rho(v)| \rfloor$.