Global Approximation of Local Optimality: Nonsubmodular Optimization

Weili Wu[1†], Zhao Zhang[2*] and Ding-Zhu Du[1*]

[1*]Department of Computer Science, University of Texas at Dallas, Richardson, 75080, TX, USA.
[2]College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua, 321004, Zhejiang, China.

*Corresponding author(s). E-mail(s): hxhzz@sina.com; dzdu@utdallas.edu;
Contributing authors: weiliwu@utdallas.edu;
†These authors contributed equally to this work.

Abstract

In the study of information technology, one of important efforts is made on dealing with nonsubmodular optimizations since there are many such problems raised in various areas of computer and information science. Usually, nonsubmodular optimization problems are NP-hard. Therefore, design and analysis of approximation algorithms are important tasks in the study of nonsubmodular optimizations. However, the traditional methods do not work well. Therefore, a new method, the global approximation of local optimality, is proposed recently. In this paper, we give an extensive study for this new methodology.

Keywords: Global Approximation, Local Optimality, Nonsubmodular Optimization

1 Introduction

An important class of graph data is the social data, generated from the online social network (OSN). The social data have been growing rapidly over the internet through various OSNs, including online websites, such as Facebook, LinkedIn, ResearchGate, and messengers, such as Skype. The widespread use of them leads to an increasing interest in efficiently and correctly discovering important, useful, and implicit information. Efficient techniques for data harnessing about social data are crucial to applications across many domains, including public safety, environment management, election, and viral marketing.

In study of techniques for graph data harnessing, one of important efforts is made on dealing with nonsubmodular optimizations since there are many such problems raised in various research subjects on information discovery and data management. Let us mention a few examples.

Sharing and spreading various contents is growing importance of information diffusion in OSNs. To develop effective and efficient techniques for content processing, a content spread maximization problem is studied in [1], which is a nonsubmodular optimization. The target activation probability maximization [2] is another nonsubmodular optimization problem regarding information diffusion, which aims at influence towards certain target user.

There are two popular approaches to deal with the misinformation blocking. The first one is to cut connections at nodes or links (e.g., see [3, 4]). In [5], link deletion is considered and a rumor spread minimization problem is proposed, which is a nonsubmodular minimization problem. In [6], node-cuts are considered and a maximum protection problem is proposed, which is a nonsubmodular maximization problem. The second approach is to spread the truth to clarify misinformation [7, 8]. More optimization research problems are formulated along this line [9–13], which are also nonsubmodular optimizations.

The sentiment analysis has been widely used in mining information from OSNs for election prediction and decision making in business management [14–16]. The language corpora plays an important role in the sentiment analysis [17]. Previously, many larger corpora are created. The size of corpora effects the power and accuracy of analysis, however, it also brings redundancy. Consider such a tradeoff. A **language corpora extraction** problem is formulated in [18], which is nonsubmodular.

A very important application of social influence is the **viral marketing** where there exist various problem formulations based on various types of products. When influence tends to be of certain composed type, such as complementary relationship between products [19], multiple features of a product, the activity profit of multiple players [20], group influence [21], or hypergraph structure [22] is considered, corresponding optimization problem is nonsubmodular.

Actually, the submodular property is very important in optimization theory.

Definition 1 (Submodular and Monotone Properties)**.** *A function f over all subsets of a finite set V is said to be* **submodular** *if for any two subsets A and B, $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$. Moreover, f is* **monotone nondecreasing** *if for $A \subset B$, $f(A) \leq f(B)$.*

The submodular function is usually considered as a sort of discrete convex function since its unconstrained minimization can be solved in polynomial-time [23–25]. In the set function optimization, the analysis on submodular functions, especially on monotone nondecreasing submodular functions has been very well studied [26], so that although maximization and constrained minimization of submodular functions are often NP-hard, there exist many theoretical guaranteed approximation solutions for them [27–32].

However, for nonsubmodular optimizations, although many applications have appeared, no satisfied solution has been found. Based on the study of existing algorithms for nonsubmodular optimization, such as parameterized algorithms, sandwich methods, and submodular-supermodular algorithms, a new type of approximation algorithms, the global approximation of local optimality is proposed recently. In this paper, we give an extensive study on this new type of approximations.

2 What is Global Approximation of Local Optimality?

Before explain what is the global approximation of local optimality, let us first review previous methods for nonsubmodular optimization. They can be classified into three classes.

Parameterized Method: To deal with nonsubmodular optimization, one intends to measure how far the function differs from the submodularity. Motivated from this intension, several parameters are introduced and theoretical results for submodular optimization are extended to nonsubmodular optimization with those parameters. For example, the supermodular degree $\mathcal{D}_f^+$ for set function f is introduced in [33] and with this parameter, a greedy algorithm for monotone nondecreasing nonsubmodular maximization with k matroid constraints is proved to have performance ratio $\frac{1}{k(\mathcal{D}_f^+ +1)+1}$ [34]. Moreover, for monotone nondecreasing nonsubmodular maximization with size-constraint, a similar greedy algorithm is proved to have performance ratio $\frac{1}{\alpha}(1-e^{-\alpha\gamma})$ [35] where α is the curvature of objective set function f defined in [36] and γ is the submodularity ratio defined in [37]. These two results have been utilized in [11, 20], respectively. However, it is hard to find other significant application in the literature. In fact, in each specific nonsubmodular optimization problem in the real world problem, the parameter value often tends to give a performance ratio 0 for maximization and ∞ for minimization.

Sandwich Method: The sandwich method was initiated in [38] and later appeared in many publications [2, 5, 6, 9, 10, 13, 19, 21, 22, 39, 40]. To explain this method, suppose we face a problem $\max_{A\in\Omega} f(A)$ where Ω is a collection of subsets of 2^X and X is a finite set. Then the sandwich method works in the following way.

- **Input** a set function $f : 2^X \to R$.

- Initially, find two submodular functions u and l such that $u(A) \geq f(A) \geq l(A)$ for $A \in \Omega$. Then carry out following
  - Compute an α-approximation solution S_u for $\max_{A \in \Omega} u(A)$ and a β-approximation solution S_l for $\max_{A \in \Omega} l(A)$.
  - Compute a feasible solution S_o for $\max_{A \in \Omega} f(A)$.
- **Output** $S = \text{argmax}(f(S_u), f(S_o), f(S_l))$.

The solution S produced by the sandwich method satisfies the following:

$$f(S) \geq \max\left\{\frac{f(S_u)}{u(S_u)} \cdot \alpha, \frac{opt_l}{opt_f} \cdot \beta\right\} \cdot opt_f,$$

where opt_f (opt_l) is the objective function value of the minimum solution for $\max_{A \in \Omega} f(A)$ ($\max_{A \in \Omega} l(A)$).

Clearly, this performance ratio is really ugly, which cannot give a clean indication for the quality of approximation.

Seek Local Optimal Solution: Analysis in above two classes of algorithms indicates that the traditional approximation performance ratio cannot work well for nonsubmodular optimizations. Therefore, one's attention is moved to local optimal solutions and a class of algorithms are designed to stop at local optimal solutions. Submodular-supermodular algorithm [41], modular-modular algorithm [42], and iterated sandwich algorithm [43, 44] are examples in this class. To give further discussion, let us consider submodular-supermodular algorithm [41] as an example. This algorithm is designed based on following properties of set functions.

Theorem 1 (DS Decomposition [41, 42]). *Every set function f can be decomposed into the difference of two monotone nondecreasing submodular functions g and h, i.e., $f = g - h$.*

Theorem 2 (Sandwich Theorem [43]). *Consider a set function $f : 2^X \to R$. For any $A \subseteq X$, there exist two modular functions m_u and m_l such that $f(A) = m_u(A) = m_l(A)$ and $m_u(Y) \geq f(Y) \geq m_l(Y)$ for any $Y \in 2^X$. (A set function m is modular if both m and $-m$ are submodular.)*

The submodular-supermodular algorithm for unconstrained minimization $\min f(A)$ for $A \in 2^X$ is as shown in Algorithm 1.

To understand why this algorithm can reach a local minimal solution, let us first make clear what is a local minimal solution.

Definition 2 (Local Minimal Solution (Type 1)). *Consider a set function minimization problem $\min_{A \in \Omega} f(A)$ where Ω is the feasible domain. $A \in \Omega$ is called a* local minimal solution *if $A \in \Omega$ and for any $A' \in \Omega$, which is obtained from A by deleting an element or adding an element, $f(A') \geq f(A)$.*

Algorithm 1 Submodular-Supermodular Algorithm for Minimization.

Input: a set function $f : 2^X \to R$ and its DS decomposition $f = g - h$ where g and h are submodular functions.
Output: A subset A of X.

```
choose a set A ⊆ X;
while A⁺ ≠ A do
    A ← A⁺;
    compute a lower bound modular function m_hl for h such that m_hl(A) = h(A);
    compute a minimum solution A⁺ for min_{Y∈2^X} [g(Y) − m_hl(Y)];
end while
return A.
```

Following property of modular lower bound plays an important role in establishment of local minimal solution.

Lemma 2.1. *Let $h : 2^X \to R$ be a submodular function. Let σ is a permutation of all elements in X such that $A = \{\sigma(1), ..., \sigma(\mathsf{A})\}$. Define the modular function m^σ_{hl} by*

$$m^\sigma_{hl}(Y) = h(\emptyset) + \sum_{\sigma(i)\in Y} \{h(S^\sigma_i) - h(S^\sigma_{i-1})\}$$

where $S^\sigma_i = \{\sigma(1), \sigma(2), ..., \sigma(i)\}$. Then m^σ_{hl} is a lower bound of h such that for any i, $h(S^\sigma_i) = m^\sigma_{hl}(S^\sigma_i)$. and especially, $h(A) = m^\sigma_{hl}(A)$.

Let Σ_A be a collection of at most $\max(| A |, | X \setminus A |)$ permutations σ of X such that $A = \{\sigma(1), ..., \sigma(| A |)\}$, $\sigma(| A |)$ goes over all elements of A, and $\sigma(| A | +1)$ goes over all elements of $X \setminus A$. In the submodular-supermodular algorithm, let m_{hl} go over all m^σ_{hl} for $\sigma \in \Sigma_A$. Then, the algorithm will stop at a local minimal solution. Therefore, we can state as follows.

Theorem 3. *The submodular-supermodular algorithm for unconstrained set function minimization stops at a local minimal solution if run each iteration in $O(n)$ time.*

There is an important question about the number of iterations. In [43, 44], it has been indicated that for some optimization problems, computing a local optimal solution is PLS-complete [45], that is, it is unlikely to have a polynomial-time algorithm to compute a local optimal solution. It follows that for submodular-supermodular algorithm, the number of iterations is unlikely to be bounded by a polynomial since they are working generally for a large class containing those optimization problems. Motivated from this fact, the global approximation of local optimality is proposed.

Global Approximation of Local Optimality: An algorithm is called a *global approximation of local optimality*, or *GL-approximation* for a brief name, if it can always produce an approximation solution within a guaranteed factor from some local optimal solution. Algorithm 2 is a GL-approximation obtained from modification of submodular-supermodular algorithm.

Algorithm 2 Modified Submodular-Supermodular Algorithm for Minimization.

Input: a set function $f : 2^X \to R$ and its DS decomposition$f = g - h$ where g and h are submodular functions.
Output: A subset A of X.

choose a set $A \subseteq X$; set $ok \leftarrow 1$;
while $ok = 1$ **do**
  for $\sigma \in \Sigma_A$ **do**
    compute a lower bound modular function m^σ_{hl} for h;
    compute a minimum solution A^+_σ for $\min_{Y \in 2^X} [g(Y) - m^\sigma_{hl}(Y)]$;
  end for
  $\sigma \leftarrow \mathrm{argmin}_{\sigma \in \Sigma_A} f(A^+_\sigma)$;
  $A^+ \leftarrow A^+_\sigma$
  if $f(A^+) \cdot (1 + \varepsilon) \geq f(A)$ **then**
    $ok \leftarrow 0$;
  end if
  $A \leftarrow A^+$;
end while
return A.

In the algorithm, Σ_A is a collection of at most $\max(\mid A \mid, \mid X \setminus A \mid)$ permutations σ of X such that $A = \{\sigma(1), ..., \sigma(\mid A \mid)\}$, $\sigma(\mid A \mid)$ goes over all elements of A, and $\sigma(\mid A \mid +1)$ goes over all elements of $X \setminus A$, and m^σ_{hl} is a modular lower bound of h, constructed by using permutation σ (see [42] for detail.)

For this algorithm, we can prove following result.

Theorem 4. *Let f be a nonnegative function. Then above algorithm runs within $O(\frac{1}{\varepsilon} \ln \zeta)$ iterations and always ends at an approximation solution with objective function value within a factor of $(1 + \varepsilon)$ from a local optimal solution where ζ is the ratio of the maximum value and the minimum value of f.*

The GL-approximation provides a new tool to study nonsubmodular optimizations, which will introduce a lot of exploratory and innovative research opportunities.

A Remark Why we call the new-type of approximation as global approximation of local optimality? Let us look at theoretical development of nonlinear programming. For convex function minimization with or without constraints, a satisfied algorithm is usually expected to generate a sequence of points which is convergent to an optimal solution. However, such an expectation cannot realized for non-convex optimizations. Instead, one established another criterion, global convergence, which means that every cluster point is locally optimal, that is, every convergent subsequence converges to a local optimal point [46–48]. The name of global approximation of local optimality is an analogy of the global convergence.

3 Alternative Local Optimality

With DS decomposition, there exists another way to describe local optimal solutions. First, let us define subgradient of a submodular function.

Definition 3 (Subgradient of Submodular Function)**.** *Let h be a submodular function over $\Omega \subseteq 2^X$. For any $A \in 2^X$, define its subgradient at set A as follows:*

$$\partial h(A) = \{c \in R^X \mid h(Y) \geq h(A) + \langle c, Y - A \rangle \text{ for } Y \in \Omega\},$$

that is, $\partial h(A)$ consists of all linear functions $c : X \to R$ satisfying $h(Y) \geq h(A) + c(Y) - c(A)$ for $Y \in \Omega$ where $c(Y) = \sum_{y \in Y} c(y)$.

Let $f = g - h$ be a set function on $\Omega \subseteq 2^X$ and g and h submodular functions on $\Omega \subseteq 2^X$. If set A is a minimum solution for $\min_{Y \subseteq X} f(Y)$, then for any $Y \in \Omega$,

$$f(Y) = g(Y) - h(Y) \geq g(A) - h(A) = f(A).$$

Therefore, for any $c \in \partial h(A)$,

$$g(Y) - g(A) \geq h(Y) - h(A) \geq c(Y) - c(A),$$

for $Y \in \Omega$, that is, $c \in \partial g$. Hence,

$$\partial h(A) \subseteq \partial g(A).$$

This fact motivates the following definition.

Definition 4 (Local Minimal Solution (Type 2))**.** *Consider a set function minimization problem $\min_{A \in \Omega} f(A)$ where $\Omega \subseteq 2^X$ is the feasible domain. $A \in \Omega$ is called a* local minimal solution *if $\partial h(A) \subseteq \partial g(A)$.*

A continuous function is called a *DC function* if it can be represented as the difference of two convex functions. A discrete DC function is a DC function restricted in a discrete set in its definition domain. Every set function is a discrete DC function because every submodular set function can be extended to a convex function (Lovász extension) on domain $[0,1]^n$. Definition 4 is consistent with the definition of local minimum solution of discrete DC functions [49].

There exist several algorithms for minimization of discrete DC functions [49–51], which terminate at local minimum solutions. Are they polynomial-time algorithms? No answer exists in the literature. Following investigation provides with an answer and a potential research topic.

Lemma 3.1. *Let A be a local minimum solution defined in Definition 4. Then for any $Y \in \mathcal{U}$, $f(A) \leq f(Y)$ where*

$$\mathcal{U} = \{Y \in \Omega \mid \partial h(Y) \cap \partial g(A) \neq \emptyset\}.$$

Proof Choose $c \in \partial h(Y) \cap \partial g(A)$. Then

$$h(A) \geq h(Y) + (c(A) - c(Y)) \text{ and } g(Y) \geq g(A) + (c(Y) - c(A)).$$

Hence $h(Y) - h(A) \leq c(Y) - c(A) \leq g(Y) - g(A)$. Therefore, $f(Y) \geq f(A)$. □

Lemma 3.2 (Fujishige [52])**.** *A point $c \in R^X$ is an extreme point of $\partial f(A)$ if and only if there is a permutation σ for elements in X, i.e., $X = \{\sigma(1), \sigma(2), ..., \sigma(\mathsf{X})\}$, such that $A = \{\sigma(1), \sigma(2), ..., \sigma(\mathsf{A})\}$ and $c(\{\sigma(i)\}) = f(S_i) - f(S_{i-1})$ for $1 \leq i \leq \mathsf{X}$ where $S_0 = \emptyset$ and $S_i = \{\sigma(1), \sigma(2), ..., \sigma(i)\}$.*

Theorem 5. *If A is a local minimum solution of type 2, then A is a local minimum solution of type 1.*

Proof For any $x \in A$, consider permutation $X = \{\sigma(1), \sigma(2), ..., \sigma(\mathsf{X})\}$ such that $A = \{\sigma(1), \sigma(2), ..., \sigma(\mathsf{A})\}$ and $\sigma(\mathsf{A}) = x$. Define linear function c by $c(\{\sigma(i)\}) = h(S_i) - h(S_{i-1})$ for $1 \leq i \leq \mathsf{X}$ where $S_0 = \emptyset$ and $S_i = \{\sigma(1), \sigma(2), ..., \sigma(i)\}$. Then $c \in \partial h(A \setminus \{x\}) \cap \partial h$. Since $\partial h(A) \subseteq \partial g(A)$, we have $c \in \partial h(A \setminus \{x\}) \cap \partial g(A)$. By Lemma 3.1, $f(A) \leq f(A \setminus \{x\})$.

Similarly, we can show that for any $x \in X \setminus A$, $f(A) \leq f(A \cup \{x\})$. □

Corollary 3.3. *For any algorithm terminating at a local minimum solution of type 2, it is unlikely to run in polynomial-time.*

A research subject suggested by this corollary is to modify those algorithm into polynomial-time global approximations of local minimality.

4 Nonconvex Relaxation

Relaxation is a popular technique for design of approximation algorithms. Many traditional approximation algorithms are designed with LP relaxation, semi-definite relaxation, and convex relaxation. Now, nonconvex relaxation may become a new development for global approximation of local optimality, which will also introduce some interesting new research directions.

Let us consider a specific problem to make explanation. Consider the following problem [53]:

Problem 4.1 (Maximum Group Set Coverage with Composed Targets)**.** *Consider m groups $\mathcal{G}_1$, $\mathcal{G}_2$, ..., $\mathcal{G}_m$ of subsets of a finite set X. All elements in X are partitioned into subsets X_1, X_2, ... X_r, each called a* composed target*. Suppose that for every composed target X_t, each subset in any group $\mathcal{G}_i$ covers at most one elements in X_t. The problem is to select one subset from each group to cover the maximum total number of composed targets, where a composed target is said to be covered if all elements in the composed target are covered.*

This problem can be formulated as the following nonsubmodular maximization with linear constraints.

$$
\begin{aligned}
\max \quad & \sum_{t=1}^{\tau} \prod_{j \in X_t} y_j && (INLP)\\
\text{s.t.} \quad & y_j \le \sum_{i=1}^{m} \sum_{S: j \in S \in \mathcal{G}_i} x_{iS} \quad \forall j = 1, ..., n,\\
& \sum_{S: S \in \mathcal{G}_i} x_{iS} \le 1 \quad \forall i = 1, ..., m,\\
& y_j \in \{0,1\} \quad \forall j = 1, ..., n,\\
& x_{iS} \in \{0,1\} \quad \forall S \in \mathcal{G}_i \text{ and } i = 1, 2, ..., m.
\end{aligned}
$$

Its relaxation is a linear constrained nonlinear programming as follows.

$$
\begin{aligned}
\max \quad & \sum_{t=1}^{\tau} \prod_{j \in X_t} y_j && (NLP)\\
\text{s.t.} \quad & y_j \le \sum_{i=1}^{m} \sum_{S: j \in S \in \mathcal{G}_i} x_{iS} \quad \forall j = 1, ..., n,\\
& \sum_{S: S \in \mathcal{G}_i} x_{iS} \le 1 \quad \forall i = 1, ..., m,\\
& 0 \le y_j \le 1 \quad \forall j = 1, ..., n,\\
& 0 \le x_{iS} \le 1 \quad \forall S \in \mathcal{G}_i \text{ and } i = 1, 2, ..., m.
\end{aligned}
$$

Let (y_j^*, x_{iS}^*) be a $(1-\varepsilon)$-approximate for a local optimal solution (KKT-point) of (NLP). Let (x_{iS}, y_j) be a solution obtained from the following randomized rounding.

Randomized Rounding: For each group $\mathcal{G}_i$, select one subset S (i.e., set $x_{iS} = 1$ and $x_{iS'} = 0$ for $S' \neq S$) with probability x_{iS}^* and no subset of group

$\mathcal{G}_i$ is selected with probability $1 - \sum_{S \in \mathcal{G}_i} x^*_{iS}$. Set $y_j = 1$ if element j appears in a selected subset, and $y_j = 0$ otherwise.

Du et al. [53] showed the following.

Theorem 6. *Let (y_j, x_{iS}) be an approximate solution obtained by the above randomized rounding. Then*

$$E\left[\sum_{t=1}^{\tau} \prod_{j \in X_t} y_j\right] \geq (1 - e^{-1})^{\alpha}(1 - \varepsilon) lopt$$

where lopt is the objective value of a local optimal solution of (NLP) and $\alpha = \max_{1 \leq t \leq \tau} \mathsf{X\_t}$.

There are two research directions raised by this result.

First, what is complexity of computing approximation of KKT-points? Vavasis [54] gave an algorithm for computing an ε-KKT of a box-constrained quadratic programming with $O(n^3(M/\varepsilon)^2)$ arithmetic operations. Ye [55] presented an algorithm for computing an ε-KKT of a linearly constrained quadratic programming with $O((n^3/\varepsilon)\log(1/\varepsilon) + n \log n)$ iterations. Can those works be extended to other objective functions, especially multilinear functions? It is an interesting problem. In fact, this is a research direction with a largely unexplored field. Research outcomes will become foundation for noncovex relaxation.

Second, what is the relationship between local optimal solutions of (INLP) and (NLP)? That is, what is the relationship between local optimal solutions of an integer programming and local optimal solutions of its nonconvex relaxation? So far, there is no research outcome in this research direction although it seems a fundamental problem related to nonconvex relaxation.

5 Conclusion

Design and analysis of GL-approximation is a new research direction in the study of nonsubmodular optimizations, with a lot of interesting open problems. With the increasing interests on nonsubmodular optimizations and their applications, one may pay more attention to theoretical foundation of this type of approximation algorithms. Therefore, our research on GL-approximations will not only contribute to the development of emerging components for building optimization theory, but also bring broader benefits in many technology developments. The discoveries on studying GL-approximation also promote the development of some fields in computer science and mathematics, such as social networks and nonlinear combinatorial optimization.

Declarations

- This research is supported in part by National Natural Science Foundation of China (*No.*U20A2068), Zhejiang Provincial Natural Science Foundation of China (*No.*LD19A010001) and Natural Science Foundation of USA (*No.*1907472).
- The authors declare that they have no conflict of interest.

References

[1] Wenguo Yang, Jianmin Ma, Yi Li, Ruidong Yan, Jing Yuan, Weili Wu, Deying Li: Marginal Gains to Maximize Content Spread in Social Networks. IEEE Trans. Comput. Soc. Syst. 6(3): 479-490 (2019).

[2] Yapu Zhang, Jianxiong Guo, Wenguo Yang, Weili Wu: Targeted Activation Probability Maximization Problem in Online Social Networks. IEEE Trans. Netw. Sci. Eng. 8(1): 294-304 (2021).

[3] M. Kimura, K. Saito, and H. Motoda: Blocking links to minimize contamination spread in a social network, ACM Trans. Knowl. Discov. Data, 3(2), pp. 9:1-23, Apr. 2009.

[4] H. Habiba, Y. Yu, T. Y. Berger-Wolf, and J. Saia: Finding spread blockers in dynamic networks, ser. SNAKDD'08, 2010, pp. 55-76.

[5] Ruidong Yan, Yi Li, Weili Wu, Deying Li, Yongcai Wang: Rumor Blocking through Online Link Deletion on Social Networks. ACM Trans. Knowl. Discov. Data 13(2): 16:1-26 (2019).

[6] Ling Gai, Hongwei Du, Lidong Wu, Junlei Zhu, Yuehua Bu: Blocking Rumor by Cut. J. Comb. Optim. 36(2): 392-399 (2018).

[7] R. M. Tripathy, A. Bagchi, and S. Mehta: A study of rumor control strategies on social networks, ser. CIKM'10. ACM, 2010.

[8] N. P. Nguyen, G. Yan, M. T. Thai, and S. Eidenbenz: Containment of misinformation spread in online social networks, ser. WebSci'12. ACM, 2012, pp. 213-222.

[9] Guangmo Amo Tong, Ding-Zhu Du, Weili Wu: On Misinformation Containment in Online Social Networks. NeurIPS 2018: 339-349.

[10] Ruidong Yan, Deying Li, Weili Wu, Ding-Zhu Du, Yongcai Wang: Minimizing Influence of Rumors by Blockers on Social Networks: Algorithms and Analysis. IEEE Trans. Netw. Sci. Eng. 7(3): 1067-1078 (2020).

[11] Jianming Zhu, Smita Ghosh, Weili Wu: Robust rumor blocking problem with uncertain rumor sources in social networks. World Wide Web 24(1): 229-247 (2021).

[12] Jianxiong Guo, Yi Li, Weili Wu: Targeted Protection Maximization in Social Networks. IEEE Trans. Netw. Sci. Eng. 7(3): 1645-1655 (2020).

[13] Jianxiong Guo, Tiantian Chen, Weili Wu: A Multi-Feature Diffusion Model: Rumor Blocking in Social Networks. IEEE/ACM Trans. Netw. 29(1): 386-397 (2021).

[14] Sinan Aral, and Dylan Walker. Creating social contagion through viral product design: A randomized trial of peer influence in networks. Management Science 57(9): 1623-1639 (2011).

[15] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. Social network analysis and mining for business applications. ACM Transactions on Intelligent Systems and Technology (TIST) 2(3): 22, (2011).

[16] Agnis Stibe. Exploring Social Influence and Incremental Online Persuasion on Twitter: A Longitudinal Study. In Mobile Web Information Systems, pp. 286-300. Springer International Publishing, 2014.

[17] Luobing Dong, Qiumin Guo, Weili Wu: Speech corpora subset selection based on time-continuous utterances features. J. Comb. Optim. 37(4): 1237-1248 (2019).

[18] Luobing Dong, Qiumin Guo, Weili Wu, Meghana N. Satpute: A semantic relatedness preserved subset extraction method for language corpora based on pseudo-Boolean optimization. Theor. Comput. Sci. 836: 65-75 (2020).

[19] Jianxiong Guo, Weili Wu: A Novel Scene of Viral Marketing for Complementary Products. IEEE Trans. Comput. Soc. Syst. 6(4): 797-808 (2019).

[20] Wenguo Yang, Jing Yuan, Weili Wu, Jianmin Ma, Ding-Zhu Du: Maximizing Activity Profit in Social Networks. IEEE Trans. Comput. Soc. Syst. 6(1): 117-126 (2019).

[21] Jianming Zhu, Smita Ghosh, Weili Wu: Group Influence Maximization Problem in Social Networks. IEEE Trans. Comput. Soc. Syst. 6(6): 1156-1164 (2019).

[22] Jianming Zhu, Junlei Zhu, Smita Ghosh, Weili Wu, Jing Yuan: Social Influence Maximization in Hypergraph in Social Networks. IEEE Trans.

Netw. Sci. Eng. 6(4): 801-811 (2019).

[23] M. Grötschel, L. Lovász, and A. Schrijver: Geometric Algorithms and Combinatorial Optimization. Springer-Verlag, 2nd edition, 1988.

[24] J. B. Orlin, A faster strongly polynomial time algorithm for submodular function minimization. Mathematical Programming, 118:237-251, 2009.

[25] A. Schrijver: A combinatorial algorithm minimizing submodular functions in strong polynomial time, J. Combinatorial Theory (B), 80: 346-355, (2000).

[26] Kazuo Murota, Discrete Convex Analysis, (SIAM, 2003).

[27] George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher, An analysis of approximations for maximizing submodular set functions - I. Mathematical Programming, 14(1): 265-294 (1978).

[28] Laurence A. Wolsey, An analysis of the greedy algorithm for the submodular set covering problem. Combinatorica, 2(4):385-393, (1982).

[29] U. Feige, V. Mirrokni, and J. Vondrák, Maximizing nonmonotone submodular functions, in Proceedings of the IEEE Foundations of Computer Science, 2007, pp. 461-471.

[30] J. Lee, V. Mirrokni, V. Nagarajan, and M. Sviridenko, Nonmonotone submodular maximization under matroid and knapsack constraints, in Proceedings of the ACM Symposium on Theory of Computing, 2009, pp. 323-332.

[31] M. Feldman, J. Naor, R. Schwartz: A unified continuous greedy algorithm for submodular maximization, IEEE FOCS 2011, pp. 570-579.

[32] Z. Svitkina and L. Fleischer, Submodular approximation: Sampling-based algorithms and lower bounds, SIAM J. Comput. 40: 1715-1737 (2011).

[33] U. Feige and R. Izsak, Welfare maximization and the supermodular degree, in ACM ITCS, 2013, pp. 247-256.

[34] M. Feldman and R. Izsak, Constrained monotone function maximization and the supermodular degree, in ACM-SIAM SODA, 2014.

[35] A.A. Bian, J.M. Buhmann, A. Krause, S. Tschiatschek: Guarantees for greedy maximization of non-submodular functions with applications, arXiv:1703.02100 (2017).

[36] A. Das, D. Kempe: Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. arXiv:1102.3975 (2011).

[37] W.J. Sviridenko, M.J. Vondrk: Optimal approximation for submodular and supermodular optimization with bounded curvature. Math. Oper. Res. 42(2), 1197-1218 (2017).

[38] Wei Lu, Wei Chen, Laks V.S. Lakshmanan, From competition to complementarity: comparative influence diffusion and maximization, Proc. the VLDB Endowsment, 9(2): 60-71 (2015).

[39] Wei Chen, Tian Lin, Zihan Tan, Mingfei Zhao, Xuren Zhou, Robus influence maximization, KDD'16, San Francisco, CA, USA, 2016.

[40] Zhefeng Wang, Yu Yang, Jian Pei, and Enhong Chen, Activity maximization by effective information diffusion in social networks, arXiv: 1610.07754v1 [cs.SI] 25 Oct 2016.

[41] M. Narasimhan and J. Bilmes, A submodular-supermodular procedure with applications to discriminative structure learning, In Proc. UAI (2005).

[42] R. Iyer and J. Bilmes, Algorithms for Approximate Minimization of the Difference between Submodular Functions, In Proc. UAI (2012).

[43] Weili Wu, Zhao Zhang, and Ding-Zhu Du: Set function optimization. J Oper Res Soc China 7(2):183-193 (2019).

[44] Chuangen Gao, Shuyang Gu, Ruiqi Yang, Jiguo Yu, Weili Wu, Dachuan Xu: Interaction-aware influence maximization and iterated sandwich method. Theor. Comput. Sci. 821: 23-33 (2020).

[45] A. A. Schäfer and M. Yannakakis, Simple local search problems that are hard to solve, SIAM J. Comput., 20 (1991), pp. 56-87.

[46] D.-Z. Du, P.M. Pardalos, and Weili Wu, Mathematical Theory of Optimization, Kluwer Academic Publishers, 2001.

[47] Ding-Zhu Du and Xiang-Sun Zhang: A convergence theorem of Rosen's gradient projection method, Mathematical Programming, 36 (1986) 135-144.

[48] Ding-Zhu Du and Xiang-Sun Zhang: Global convergence of Rosen's gradient projection method, Mathematical Programming, 44 (1989) 357-366.

[49] T. Maehara, N. Marumo, K. Murota: Continuous relaxation for discrete DC programming. In: Proceedings of the 3rd International Conference on Modelling, Computation and Optimization in Information Systems and Management Sciences, pp. 181-190 (2015).

[50] Chenchen Wu, Yishui Wang, Zaixin Lu, Panos M. Pardalos, Dachuan Xu, Zhao Zhang, Ding-Zhu Du: Solving the degree-concentrated fault-tolerant spanning subgraph problem by DC programming. Math. Program. 169(1): 255-275 (2018).

[51] T. Maehara, K. Murota: A framework of discrete DC programming by discrete convex analysis. Math. Program. Ser. A 152, 435-466 (2015).

[52] S. Fujishige: Submodular Functions and Optimization, Annals of Discrete Mathematics, volume 58. (Elsevier Science, 2005).

[53] Hongwei Du, Zhao Zhang, Zhenhua Duan, Cong Tian, Ding-Zhu Du, Full view camera sensor coverage and group set coverage, Proceedings of 15th EAI International Wireless Internet Conference, Dallas, USA, 2022.

[54] S.A. Vavasis: Nonlinear Optimization: Complexity Issues, Oxford Science, New York, 1991.

[55] Yinyu Ye: On the complexity of approximating a KKT point of quadratic programming, Math. Program. 80: 195-211 (1998).