Optimization of Chance-Constrained Submodular Functions

Benjamin Doerr, Carola Doerr, Aneta Neumann, Frank Neumann, Andrew M
Sutton

$\Pr[W(S) > B] \leq \alpha$ where $S \subseteq V$ and $W(S)$ is the sum of the random weights of the elements in S. Here, B is a deterministic constraint boundary and α is a tolerance level. The objective is thus to find a set that maximizes f subject to the probability that its stochastic weights violate the boundary is at most α.

Our focus is on *monotone submodular functions,* which are set functions characterized by the property that the function value cannot decrease by adding more elements. The optimization of the considered classes of submodular problems with deterministic constraints has already been investigated by (NWF78; LKG$^+$07). The theoretical contribution of this paper extends these results to the chance constrained setting.

Since the computation of the chance constraint is usually not efficiently feasible, we assume it is evaluated by using a surrogate function that provides an upper bound on the constraint violation probability $\Pr[\sum_{s \in S} W(s) > B]$. This upper bound ensures that the chance constraint is met if the surrogate provides a value of at most α. As surrogates, we use popular deviation inequalities such as Chebyshev's inequality and Chernoff bounds.

We show that using these surrogate functions, popular greedy algorithms are also applicable in the chance-constrained setting. In particular, we analyze the case of uniformly distributed weights with identical dispersion and show that both inequalities only lead to a loss of a factor of $1 - o(1)$ compared to the deterministic setting.

We complement our theoretical work with an experimental investigation on the influence maximization problem in social networks. This investigation empirically analyzes the behavior of the greedy approaches for various stochastic settings. In particular, it shows the effectiveness of using Chernoff bounds for large inputs if only a small failure rate in terms of α can be tolerated.

Related work

Submodular optimization has been studied for a wide range of different constraint types, see, for example, (KG14) and references mentioned therein. Many of the results on monotone submodular functions are based on simple greedy selection strategies that are able to achieve provably the best

Optimization of Chance-Constrained Submodular Functions

Benjamin Doerr,[1] Carola Doerr,[2] Aneta Neumann,[3] Frank Neumann,[3] Andrew M. Sutton[4]

[1]Laboratoire d'Informatique (LIX), CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
[2]Sorbonne Université, CNRS, LIP6, Paris, France
[3]Optimisation and Logistics, School of Computer Science, The University of Adelaide, Adelaide, Australia
[4]Department of Computer Science, University of Minnesota Duluth, Duluth, MN, USA

Abstract

Submodular optimization plays a key role in many real-world problems. In many real-world scenarios, it is also necessary to handle uncertainty, and potentially disruptive events that violate constraints in stochastic settings need to be avoided. In this paper, we investigate submodular optimization problems with chance constraints. We provide a first analysis on the approximation behavior of popular greedy algorithms for submodular problems with chance constraints. Our results show that these algorithms are highly effective when using surrogate functions that estimate constraint violations based on Chernoff bounds. Furthermore, we investigate the behavior of the algorithms on popular social network problems and show that high quality solutions can still be obtained even if there are strong restrictions imposed by the chance constraint.

Introduction

Many real-world problems optimization problems involve uncertain components such as the execution length of a job, the fraction of ore in a cartload of rocks, the probability of earthquakes, etc. Safe critical systems or expensive productions must limit the potential violation of constraints imposed by such stochastic components. Constraints that explicitly address the probability of violation are known as chance constraints. Chance-constrained optimization deals with optimizing a given problem under the condition that the probability of a constraint violation does not exceed a given threshold probability α.

In this work we study optimization under chance constraints for submodular functions. Submodular functions model problems where the benefit of adding components diminishes with the addition of elements. They form an important class of optimization problems, and are extensively studied in the literature (NW78; NWF78; Von10; KG14; BBKT17; LKG$^+$07; FHK17; HFWK19). However, to our knowledge, submodular functions have not yet been studied in the chance-constrained setting.

Our Results

Given a set V, we analyze the maximization of submodular functions $f : 2^V \to \mathbb{R}$ subject to linear chance constraints

possible approximation ratio in polynomial time, unless P=NP (NWF78). (FGN+19) recently showed that greedy approaches are also successful when dealing with non-monotone submodular functions. Furthermore, Pareto optimization approaches can achieve the same worst-case performance guarantees while performing better than greedy approaches in practice if the user allows for a sufficiently large time budget (QSYT17; QYZ15; QSY+17). (RNNF19) showed that the adaptation of greedy approaches to monotone submodular problems with dynamic constraints might lead arbitrarily bad approximation behavior, whereas a Pareto optimization approach can effectively deal with dynamic changes. Evolutionary algorithms for the chance-constrained knapsack problem, which constitutes a subclass of the chance-constrained submodular problems examined in this paper, have been experimentally investigated by (XHA+19).

The paper is structured as follows. Next, we introduce the class of submodular optimization problems and the algorithms that are subject to our investigations. Afterwards, we establish conditions to meet the chance constraints based on tail-bound inequalities. We present our theoretical results for chance-constrained submodular optimization for different classes of weights. Building on these foundations, we present empirical results that illustrate the effect of different settings of uncertainty on the considered greedy algorithms for the influence maximization problem in social networks. Finally, we finish with some concluding remarks.

Chance-Constrained Submodular Functions

Given a set $V = \{v_1, \ldots, v_n\}$, we consider the optimization of a monotone submodular function $f \colon 2^V \to \mathbb{R}_{\geq 0}$. A function is called monotone iff for every $S, T \subseteq V$ with $S \subseteq T$, $f(S) \leq f(T)$ holds. A function f is called submodular iff for every $S, T \subseteq V$ with $S \subseteq T$ and $x \notin T$ we have

$$f(S \cup \{x\}) - f(S) \geq f(T \cup \{x\}) - f(T).$$

We consider the optimization of such a monotone submodular function f subject to a chance constraint where each element $s \in V$ takes on a random weight $W(s)$. Precisely, we are considering constraints of the type

$$\Pr[W(S) > B] \leq \alpha.$$

where $W(S) = \sum_{s \in S} W(s)$ is the sum of the random weights of the elements and B is the given constraint bound. The parameter α quantifies the probability of exceeding the bound B that can be tolerated.

It should be noted that for the uniform distribution, the exact joint distribution can, in principle, be computed as convolution if the random variables are independent. There is also an exact expression for the Irwin-Hall distribution (JKB95) which assumes that all random variables are independent and uniformly distributed within $[0, 1]$. However, using these approaches may not be practical when the number of chosen items is large.

Greedy Algorithms

We consider in this work the performance of greedy algorithms for the optimization of chance constrained submodular functions. Our first greedy algorithm (GA, see Algorithm 1) starts with an empty set and subsequently adds in each iteration an element with the largest marginal gain that does not violate the chance constraint. It ends when no further element can be added. Algorithm 1 was already investigated by (NWF78) in the deterministic setting. Note that the computation of the probability $\Pr[W(S) > B]$ can usually not be computed exactly and we make use of a surrogate $\widehat{\Pr}[W(S) > B] \leq \alpha$ on this value (see line 5 of Algorithm 1). Since we use upper bounds for the constraint violation probability, we are guaranteed that the constraint is met whenever our surrogate $\widehat{\Pr}$ is at most α.

Our second greedy algorithm is the generalized greedy algorithm (GGA), and is listed in Algorithm 2. The GGA extends the GA to the case in which the elements have different expected weights. It has previously been used in the deterministic setting (KMN99; LKG+07). The algorithm starts with the empty set. In each iteration, it adds an element whose ratio of the additional gain with respect to the submodular function f and the expected weight increase $E[W(S \cup \{v\}) - W(S)]$ of the constraint is maximal while still satisfying the chance constraint. The algorithm terminates if no further element can be added. At this point, it compares this constructed greedy solution with each of the n solutions consisting of a single element, and returns the solution with the maximal f-value subject to the surrogate function is at most α. Note that we are using the exact calculation for $\Pr[W(v) > B]$ when considering a single element in line 9. Lines 9 and 10 of Algorithm 2 are needed in cases where large items of high profit exist, see (KMN99; LKG+07) for more details.

Concentration Bounds

We work with two different surrogates, which are concentration bounds of Chernoff and Chebyshev type. Such bounds are frequently used in the analysis of randomized algorithms (MR95). All bounds are well-known and can be found, e.g., in (Doe18).

Theorem 1 (Multiplicative Chernoff bound)**.** *Let* $X_1, \ldots, X_n$ *be independent random variables taking values in* $[0, 1]$. *Let* $X = \sum_{i=1}^{n} X_i$. *Let* $\epsilon \geq 0$. *Then*

$$\Pr[X \geq (1 + \epsilon)E[X]] \leq \left(\frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^{E[X]} \quad (1)$$

$$\leq \exp\left(-\frac{\min\{\epsilon^2, \epsilon\}E[X]}{3} \right). \quad (2)$$

For $\epsilon \leq 1$, *(2) simplifies to*

$$\Pr[X \geq (1 + \epsilon)E[X]] \leq \exp\left(-\frac{\epsilon^2 E[X]}{3} \right). \quad (3)$$

For our experimental investigations, we work with equation (1), whereas equation (3) is used through our theoretical analysis. Note that equation (3) gives the weaker bound. Therefore, our theoretical results showing approximation guarantees also hold when working with equation (1). Chernoff bounds are very useful when requiring very small values of α. For larger values of α, e.g. $\alpha = 0.1$, we often get

---

Algorithm 1: Greedy Algorithm (GA)

input: Set of elements V, budget constraint B, failure probability α.

1   $S \leftarrow \emptyset$;
2   $V' \leftarrow V$;
3   **repeat**
4     $v^* \leftarrow \arg\max_{v \in V'}(f(S \cup \{v\}) - f(S))$;
5   **if** $\widehat{\Pr}[W(S \cup \{v^*\}) > B] \leq \alpha$ **then**
6       $S \leftarrow S \cup \{v^*\}$;
7       $V' \leftarrow V' \setminus \{v^*\}$;
8   **until** $V' \leftarrow \emptyset$;
9   **return** S;

Algorithm 2: Generalized Greedy Algorithm (GGA)

input: Set of elements V, budget constraint B, failure probability α.

1   $S \leftarrow \emptyset$;
2   $V' \leftarrow V$;
3   **repeat**
4     $v^* \leftarrow \arg\max_{v \in V'} \frac{f(S \cup \{v\}) - f(S)}{E[W(S \cup \{v\})] - W(S)]}$;
5   **if** $\widehat{\Pr}[W(S \cup \{v^*\}) > B] \leq \alpha$ **then**
6       $S \leftarrow S \cup \{v^*\}$;
7       $V' \leftarrow V' \setminus \{v^*\}$;
8   **until** $V' \leftarrow \emptyset$;
9   $v^* \leftarrow \arg\max_{\{v \in V; \Pr[W(v) > B] \leq \alpha\}} f(v)$;
10   **return** $\arg\max_{Y \in \{S, \{v^*\}\}} f(Y)$;

---

better estimates when working with a variant of Chebyshev's inequality. As we are only interested in the probability of exceeding a given constraint bound, we consider a one-sided Chebyshev inequality (also known as Cantelli's inequality), which estimates the probability of exceeding the expected value taking into account the variance of the considered random variable.

Theorem 2 ((One-sided) Chebyshev's inequality). *Let X be a random variable with expected value $E[X]$ and variance $\text{Var}[X] > 0$. Then, for all $\lambda > 0$,*

$$\Pr[X \geq E[X] + \lambda] \leq \frac{\text{Var}[X]}{\text{Var}[X] + \lambda^2}. \tag{4}$$

Chance Constraint Conditions

We now establish conditions to meet the chance constraint. We start by considering the Chernoff bound given in equation (3).

Lemma 1. *Let $W(s) \in [a(s) - \delta, a(s) + \delta]$ be independently chosen uniformly at random. If*

$$(B - E[W(X)]) \geq \sqrt{3\delta k \ln(1/\alpha)},$$

where $k = |X|$, then $\Pr[W(X) > B] \leq \alpha$.

Proof. Every item has an uncertainty of δ. Instead of considering $W(s) \in [a(s) - \delta, a(s) + \delta]$ chosen uniformly at

random, we can consider $W'(s) \in [0, 2\delta]$ chosen uniformly at random and have $W(s) = a(s) - \delta + W'(s)$. For a selection X with $|X| = k$ elements, we can therefore write $W(X) = E[W(X)] - \delta k + \sum_{x \in X} W'(X)$.

We have $E[W'(X)] = \delta k$. We consider the probability for exceeding this expected value by $\epsilon \delta k$. We set $\epsilon = (B - E[W(X)])/(\delta k)$ which implies $\epsilon \delta k + E[W(X)] = B$.

We investigate

$$\Pr[W(X) > B] = \Pr[W'(X) > \epsilon \delta k + k \delta].$$

Note that if $\epsilon = (B - E[W(X)])/(\delta k) > 1$ then $\Pr[W(X) > B] = 0$ as all weights being maximal within their range would not exceed the bound B. For $\epsilon \leq 1$, we get

$$\Pr[W(X) > B] = \Pr[W'(X) > \epsilon \delta k + k \delta]$$
$$\leq \exp\left(-\frac{\epsilon^2 k \delta}{3}\right)$$

using equation (3). In order to meet the chance constraint, we require

$$\exp\left(-\frac{\epsilon^2 k \delta}{3}\right) \leq \alpha$$
$$\iff -\frac{\epsilon^2 k \delta}{3} \leq \ln(\alpha)$$
$$\iff \epsilon^2 k \delta \geq 3\ln(1/\alpha)$$
$$\iff \epsilon^2 \geq (3\ln(1/\alpha))/(k\delta).$$

This implies that $\epsilon \geq \sqrt{(3\ln(1/\alpha))/(k\delta)}$ meets the chance constraint condition according to the considered Chernoff bound. Setting $\epsilon = (B - E[W(X)])/(\delta k)$ leads to

$$(B - E[W(X)])/(\delta k) \geq \sqrt{(3\ln(1/\alpha))/(k\delta)}$$
$$\iff (B - E[W(X)]) \geq \sqrt{3\delta k \ln(1/\alpha)},$$

which completes the proof. $\qquad \square$

Based on Chebyshev's inequality, we can obtain the following condition for meeting the chance constraint.

Lemma 2. *Let X be a solution with expected weight $E[W(X)]$ and variance $\text{Var}[W(X)]$. If*

$$B - E[W(X)] \geq \sqrt{\frac{(1 - \alpha)\text{Var}[W(X)]}{\alpha}}$$

then $\Pr[W(X) > B] \leq \alpha$.

Proof. We have

$$\frac{\text{Var}[W(X)]}{\text{Var}[W(X)] + (B - E[W(X)])^2} \leq \alpha$$
$$\iff \text{Var}[W(X)] \leq \alpha(\text{Var}[W(X)] + (B - E[W(X)])^2)$$
$$\iff (1 - \alpha)\text{Var}[W(X)] \leq \alpha(B - E[W(X)])^2$$
$$\iff (B - E[W(X)])^2 \geq \frac{(1 - \alpha)\text{Var}[W(X)]}{\alpha}$$

This together with Lemma 2 implies that the chance constraint is met if

$$B - E[W(X)] \geq \sqrt{\frac{(1 - \alpha)\text{Var}[W(X)]}{\alpha}}$$

holds. $\qquad \square$

Uniform IID Weights

We first study the case that all items have iid weights $W(s) \in [a - \delta, a + \delta]$ ($\delta \leq a$). For this case we prove that the greedy algorithm with the Chernoff bound surrogate achieves a $(1 - o(1))(1 - 1/e)$ approximation of the optimal solution for the deterministic setting when $B = \omega(1)$. This extends the same bound for the deterministic setting by (NWF78) to the chance-constrained case.

Theorem 3. *Let $a > 0$ and $0 \leq \delta < a$. Let $W(s) \in [a - \delta, a + \delta]$ be chosen uniformly at random for all s. Let $\epsilon(k) = \frac{\sqrt{3\delta k \ln(1/\alpha)}}{a}$ and k^* be the largest integer such that $k + \epsilon(k) \leq k_{opt} := \lfloor B/a \rfloor$.*

Then the first k^* items chosen by the greedy algorithm satisfy the chance constraint and are a $1 - (1/e) \exp(\frac{1+\epsilon(k)}{k^*+1+\epsilon(k)})$-approximation. For $B = \omega(1)$, this is a $(1 - o(1))(1 - 1/e)$-approximation.

Proof. Let $k_{opt} = \lfloor B/a \rfloor$ be the number of elements that are contained in an optimal solution OPT in the case that the weights are deterministic and attain the value a.

Having produced a solution with k elements following the greedy procedure, we have obtained a solution X where

$$f(X) \geq (1 - (1 - 1/k_{opt})^k) \cdot f(OPT)$$

due to an inductive argument given by (NWF78).

We now give a lower bound on k using Chernoff bound as a surrogate. Let X be a set of selected items containing $k = |X|$ elements and $E[X] = \sum_{x \in X} a(x)$ be its expected weight, δ be the uncertainty common to all items.

Since all items have the same expected weight a, we have $E[W(X)] = ak$. Using Lemma 1, the chance constraint is met if $(B - ak) \geq \sqrt{3\delta k \ln(1/\alpha)}$. We have $k_{opt} = \lfloor B/a \rfloor$ for the number of elements that could be added if the weights were deterministic. So any k with $k + \frac{\sqrt{3\delta k \ln(1/\alpha)}}{a} \leq k_{opt}$ is feasible when using the Chernoff bound.

Let

$$k^* = \max \left\{ k \mid k + \frac{\sqrt{3\delta k \ln(1/\alpha)}}{a} \leq k_{opt} \right\}. \quad (5)$$

Then

$$k_{opt} < (k^* + 1) + \frac{\sqrt{3\delta(k^* + 1)\ln(1/\alpha)}}{a} =: \beta(k^*).$$

Let X^* be a solution with k^* elements constructed by the greedy algorithm. Using the well-known estimate $(1 + x) \leq e^x$, we bound $f(X^*)$ from below by

$$(1 - (1 - 1/k_{opt})^{k^*}) \cdot f(OPT)$$

$$\geq \left(1 - \left(1 - \frac{1}{\beta(k^*)} \right)^{k^*} \right) \cdot f(OPT)$$

$$\geq \left(1 - \exp\left(-\frac{k^*}{\beta(k^*)} \right) \right) \cdot f(OPT)$$

$$= \left(1 - \exp\left(-\frac{k^*}{k^* + 1 + \epsilon(k^* + 1)} \right) \right) \cdot f(OPT)$$

$$= \left(1 - \frac{1}{e} \exp\left(\frac{1 + \epsilon(k^* + 1)}{k^* + 1 + \epsilon(k^* + 1)} \right) \right) \cdot f(OPT).$$

When $k^* = \omega(1)$, the $\exp(\cdot)$ expression is $(1 + o(1))$, yielding the asymptotic part of the claim. $\square$

For comparison, we now determine what can be obtained from using a surrogate based on Chebyshev's inequality. This bound is weaker for small values of α, but can be better for larger values of α (depending on the other constants involved).

We observe that $\mathrm{Var}[W(X)] = |X| \cdot \delta^2/3$. Defining $\tilde{\epsilon}(k) = \frac{\sqrt{(1-\alpha)k\delta^2}}{\sqrt{3\alpha}a}$ and replacing equation (5) by

$$k^* = \max \{k \mid k + \tilde{\epsilon} \leq k_{opt}\}$$

our proof above yields the following theorem.

Theorem 4. *Let $a > 0$ and $0 \leq \delta < a$. Let $W(s) \in [a - \delta, a + \delta]$ be chosen uniformly at random for all s. Let $\tilde{\epsilon}(k) = \frac{\sqrt{(1-\alpha)k\delta^2}}{\sqrt{3\alpha}a}$ and k^* be the largest integer such that $k + \tilde{\epsilon}(k) \leq k_{opt} := \lfloor B/a \rfloor$.*

Then the first k^* items chosen by the greedy algorithm satisfy the chance constraint and are a $1 - (1/e) \exp(\frac{1+\tilde{\epsilon}(k)}{k^*+1+\tilde{\epsilon}(k)})$-approximation. For $B = \omega(1)$, this is a $(1 - o(1))(1 - 1/e)$-approximation.

Note that the main difference between the Chernoff bound and Chebyshev's inequality lies in the confidence level of α that needs to be achieved as the equation using Chernoff only increases logarithmically with $1/\alpha$, whereas the one based on Chebyshev's inequality increases with the square root of $1/\alpha$.

We note that, in principle, Chebyshev's inequality does not require that the items are chosen independently. We can use Chebyshev's inequality and the approach above whenever we can compute the variance.

Uniform Weights with the Same Dispersion

We now consider the case that the items may have different random weights $W(s) \in [a(s) - \delta, a(s) + \delta]$. However, we still assume the weights are chosen independently and uniformly at random. We also assume that the uncertainty δ is the same for all items.

Let $a_{\max} = \max_{s \in V} a(s)$. We assume that $\frac{a_{\max} + \delta - B}{2\delta} \leq \alpha$ holds. This means that every single item fulfills the chance constraint. Note that items that would not fulfill this condition could be filtered out in a preprocessing step as they can

not be part of any feasible solution. Furthermore, we assume that $B = \omega(1)$ grows with the input size. The following theorem extends the results of Theorem 3 by (LKG$^+$07) to the chance-constrained setting.

Theorem 5. *For all $s \in V$ let $W(s) \in [a(s) - \delta, a(s) + \delta]$ with $a(s) > 0$ and $0 \leq \delta < \min a(s)$. If $\frac{a_{\max} + \delta - B}{2\delta} \leq \alpha$ and $B = \omega(1)$, then the solution obtained by the Generalized Greedy algorithm GGA using Chernoff bound* (3) *as surrogate for the chance constraint is a* $(1/2 - o(1))(1 - 1/e)$-*approximation.*

Proof. Let S be the greedy solution constructed by the GGA (Algorithm 2) and let v^* be the element with the largest function value. Let T be the solution produced by the generalized greedy algorithm in the deterministic setting.

We know that any solution X with $E[W(X)] \leq B - \sqrt{3\delta k \ln(1/\alpha)}$ is feasible. Furthermore, every single item is feasible as $\frac{a_{\max} + \delta - B}{2\delta} \leq \alpha$. Let $\hat{B} = B - \sqrt{3\delta k \ln(1/\alpha)}$. Let $S' = X_L \subseteq T$ be the subset of T consisting of the first L elements exactly as chosen when working with the deterministic bound B. Note that we have $f(S) \geq f(S')$ as f is monotone. If $S' = T$, we know $\arg\max_{Y \in \{S, \{v^*\}\}} f(Y)$ is a $(1/2)(1 - 1/e)$-approximation.

Furthermore, let $v^* \in V \setminus S'$ be the the element with the largest function value. Note that v^* meets the constraint bound based on our assumptions. Let $z = X_{L+1} \setminus X_L \in T$ be the first element of S not added when working with the chance constraint. If $v^* = z$, then we have $E[W(S)] + a(v^*) \geq \hat{B}$. As v^* is the single element with the largest function value, we have $f(v^*) \geq f(z)$.

Let x be the element added in iteration i to the partial solution X_{i-1} in order to obtain $X_i \subseteq S'$. Furthermore, let OPT be an optimal solution for the deterministic setting. Then following (LKG$^+$07) and using the expected cost value instead of the deterministic ones, we have

$$f(X_i) - f(X_{i-1}) \geq \frac{a(x)}{c(OPT)} \cdot (f(OPT) - f(X_{i-1}))$$

which by induction gives

$$f(X_i) \geq \left[1 - \prod_{k=1}^{i}\left(1 - \frac{a(k)}{B}\right)\right] \cdot f(OPT).$$

Every element added to S' would have also been chosen when working with the deterministic bound B. Furthermore, the single element v^* providing the largest possible profit is also accepted due to our assumption on B and we have $f(v^*) \geq f(z)$.

We have

$$f(S') + f(z) \geq \left[1 - \prod_{k=1}^{L+1}\left(1 - \frac{a(k)}{B}\right)\right] \cdot f(OPT),$$

where we take $a(L + 1)$ to be $a(z)$. Furthermore, we have

$$E[W(S')] + a(z) = a(z) + \sum_{s \in S'} a(s) \geq \hat{B}$$
$$\geq B - \sqrt{3\delta L \ln(1/\alpha)}$$

as z is the first element of S not added under the chance constraint. This implies

$$f(S) + f(v^*) \geq f(S') + f(v^*) \geq f(S') + f(z)$$

$$\geq \left[1 - \prod_{k=1}^{L+1}\left(1 - \frac{a(s)}{B}\right)\right] f(OPT)$$

$$\geq \left[1 - \prod_{k=1}^{L+1}\left(1 - \frac{B - \sqrt{3\delta(L+1)\ln(1/\alpha)}}{(L+1)B}\right)\right] f(OPT)$$

$$\geq \left[1 - \left(1 - \frac{1}{L+1} + \frac{\sqrt{3\delta\ln(1/\alpha)}}{\sqrt{L+1} \cdot B}\right)^{L+1}\right] f(OPT)$$

Again, we assume that α and δ are constants and $B = \omega(1)$ is growing with the input size. This implies $f(S) + f(v^*) \geq (1 - o(1))(1 - 1/e)$ and therefore $\max_{Y \in \{S, \{v^*\}\}} f(Y) \geq (1/2 - o(1))(1 - 1/e) \cdot f(OPT)$. $\square$

Using Chebyshev's inequality with $\text{Var}[W(X)] = |X| \cdot \delta^2/3$ and replacing $\sqrt{3\delta(L+1)\ln(1/\alpha)}$ by $\sqrt{\frac{(1-\alpha)(L+1)\delta^2}{3\alpha}}$ we obtain the following theorem.

Theorem 6. *In the situation of Theorem 5, if $\frac{a_{\max} + \delta - B}{2\delta} \leq \alpha$ and $B = \omega(1)$, then the solution obtained by the Generalized Greedy algorithm using Chebyshev's inequality as surrogate for the chance constraint is a* $(1/2 - o(1))(1 - 1/e)$-*approximation.*

Experimental Investigations

We examine our greedy algorithms on the submodular influence maximization problem (ZV16; QSYT17; LKG$^+$07). We study the impact of Chebyshev's inequality and Chernoff bounds for the chance-constrained optimization problem in the context of various stochastic parameter settings.

The Influence Maximization Problem

The influence maximization problem (IM) involves finding a set of most influential users in a social network. IM aims to maximize the spread of influence through a social network, which is the graph of interactions within a group of users (KKT03). A selection of users encounters a cost and the problem of influence maximization has been studied subject to a deterministic constraint that limits the cost of the user selection.

Formally, the problem investigated in our experiments is defined as follows. Given a directed graph $G = (V, E)$ where each node represents a user, and each edge $(u, v) \in E$ has assigned an edge probability $p_{u,v}$ that user u influences user v, the aim of the IM problem is to find a subset $X \subseteq V$ such that the expected number of activated nodes $E[I(X)]$ of X is maximized. Given a cost function $c : V \to \mathbb{R}^+$ and a budget $B \geq 0$, the corresponding submodular optimization problem under chance constraints is given as

$$\arg\max_{X \subseteq V} E[I(X)] \text{ s.t. } \Pr[c(X) > B] \leq \alpha.$$

A more detailed description of the influence maximization problem on social networks is available at (LKG$^+$07;
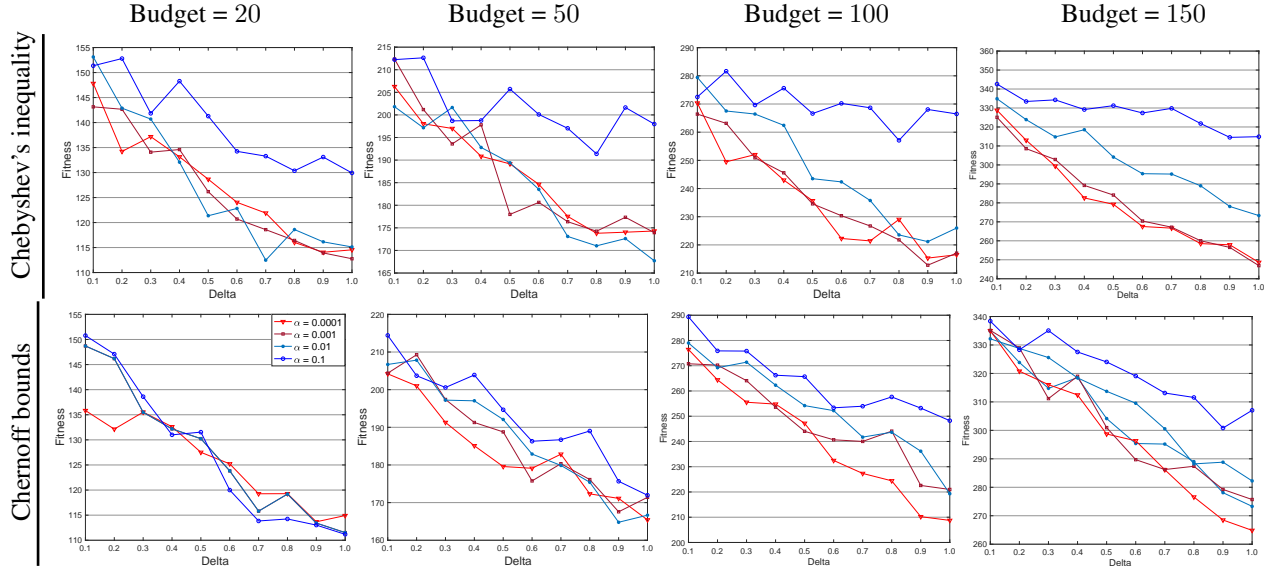
Figure 1: Function value for budgets $B = 20, 50, 100, 150$ (from left to right) using Chebyshev's inequality (top) and Chernoff bound (bottom) for $\alpha = 0.1, 0.01, 0.001, 0.0001$ with all the expected weights 1.
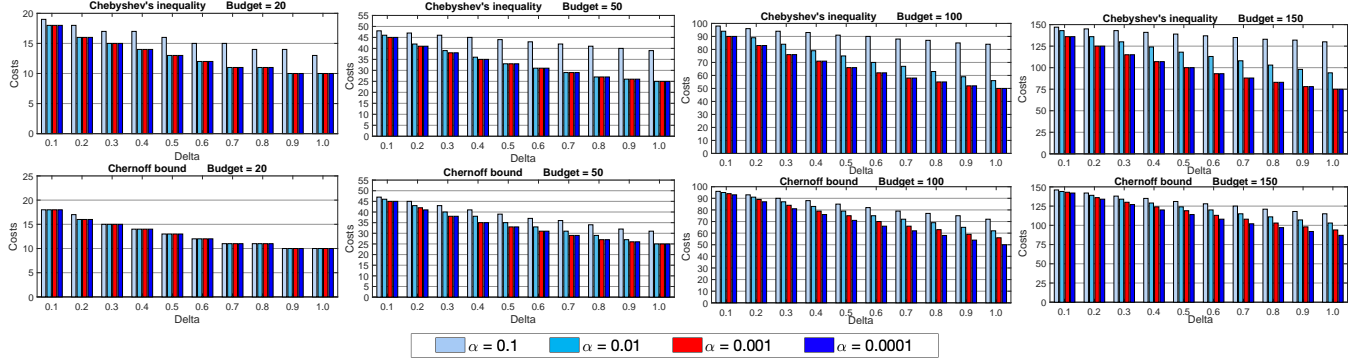

Figure 2: Maximal cost values for budgets $B = 20, 50, 100, 150$ (from left to right) using Chebyshev's inequality (top) and Chernoff bound (bottom) for $\alpha = 0.1, 0.01, 0.001, 0.0001$ with uniform expected weights set to 1.

KKT15; ZV16; QSYT17). Note though, that these works all study the case in which the cost of adding a user is deterministic.

We consider two types of cost constraints matching the settings investigated in the theoretical analysis. In the first type, the expected weight of all nodes is 1, i.e. $a(v) = 1$, for all $v \in V$ whereas in the second setting the expected cost of a node v is given by $a(v) = deg(v) + 1$. Here $deg(v)$ denotes the degree of v in the given graph G. We examine GA for the first type and GGA for the second type of instances. The chance constraint for both settings requires that the probability of exceeding the given bound B is at most α. We investigate different reliability thresholds given by α together with a range of δ values which determine the amount of uncertainty. To evaluate the influence maximization problem in the context of surrogate chance constraints, we employ a graph obtained from social news data, with simple settings collected from a real-world data set obtained from the social news aggregator

Digg. The Digg dataset (HL12) contains stories submitted to the platform over a period of a month, and identification (IDs) of users who voted on the popular stories. The data consist of two tables that describe friendship links among users and the anonymized user votes on news stories (HL12; RA15). We use the preprocessed data with 3523 nodes and 90244 edges, and estimated edge probabilities from the user votes based on the method in (BBM12).

Uniform Chance Constraints

We consider the results for the greedy algorithm based on Chebyshev's inequality and the greedy algorithms based on Chernoff bounds subject to the uniform distribution with all the expected weights 1. Figure 1 shows the results of influence spread maximization for the GA based on Chebyshev's inequality (first row) and the GA based on Chernoff bounds (second row) for budgets $B = 20, 50, 100, 150$. For the experimental investigations of our GA, we consider all combinations of $\alpha = 0.1, 0.01, 0.001, 0.0001$, and

Figure 3: Function values for budgets $B = 100$ (left) and $B = 500$ (right) using Chebyshev's inequality and Chernoff bound for $\alpha = 0.1, 0.01, 0.001, 0.0001$ with degree dependent random weights.
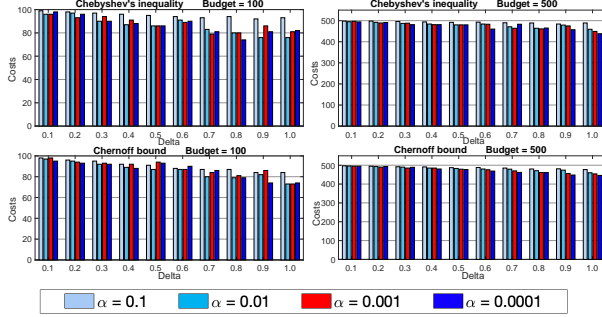

Figure 4: Maximal cost values for budget $B = 100$ (left) and $B = 500$ using Chebyshev's inequality (top) and Chernoff bound (bottom) for $\alpha = 0.1, 0.01, 0.001, 0.0001$ with degree dependent random weights.

$\delta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$.

We compare the results in terms of the fitness achieved at each α and δ level for budgets $B = 20, 50, 100, 150$. For the deterministic setting, where B elements can be used, the obtained function values for $B = 20, 50, 100, 150$ are $161.42, 214.78, 287.56$, and 345.84, respectively. For the chance-constrained problem, Figure 1 shows that the GA using Chernoff chance estimates has a better performance/higher fitness values than the GA using Chebyshev's inequality among the considered δ values in most of the cases. It should be noted that higher budget values have a larger influence on the separation between the two surrogate constraints. In these cases, the Chernoff-based GA performs consistently better, and this is the result of the good asymptotic behavior of the Chernoff bound (Che52; MU05).

We also compare the results in terms of the expected cost of the solution obtained by the GA based on Chebyshev's inequality and Chernoff bound for budgets $B = 20, 50, 100, 150$ as the tail inequalities limit the maximum cost. Note that cost in this case is the same as the number of items picked by GA. The results are shown in Figure 2. The GA using the Chernoff bound is able to include more elements. The results show that the GA using the Chernoff bound allows for solutions with a larger number of elements if the budget B is high and α is small, for example $B = 150$, and for $\alpha = 0.001, 0.0001$. The GA using Chebyshev's inequality has a better performance in the case of high α

values, $\alpha = 0.1, 0.01$ on the examined problem.

Non-Uniform Chance Constraints

We now consider the GGA for the setting where the expected weights are not uniform but depend on the degree of the node of the graph. We consider budgets $B = 100, 500$ keeping the other parameters the same as in the uniform case. Figure 3 and 4 show the function and cost values obtained by the GGA based on Chernoff bounds for the combinations of α and δ. We do not observe a clear difference between using Chebyshev's inequality or Chernoff bound for the different combinations of δ and α which is due to a relatively small number of elements present in the solutions. The function values of the approaches are overall decreasing for each α with increasing value of δ. Due to the stochastic behavior of the evaluation, the obtained fitness values show a jagged, irregular course with increasing value of δ.

Conclusion

Chance constraints play a crucial role when dealing with stochastic components in constrained optimization. We presented a first study on the optimization of submodular problems subject to a given chance constraint. We have shown that popular greedy algorithms making use of Chernoff bounds and Chebyshev's inequality provide provably good solutions for monotone submodular optimization problems with chance constraints. Our analysis reveals the impact of various stochastic settings on the quality of the solution obtained. Furthermore, we have provided experimental investigations that show the effectiveness of the considered greedy approaches for the submodular influence maximization problem under chance constraints.

Acknowledgements

References

[BBKT17] Andrew An Bian, Joachim M. Buhmann, Andreas Krause, and Sebastian Tschiatschek, *Guarantees for*

greedy maximization of non-submodular functions with applications, Proceedings of the 34th International Conference on Machine Learning, ICML 2017, 2017, pp. 498–507.

[BBM12] Nicola Barbieri, Francesco Bonchi, and Giuseppe Manco, *Topic-aware social influence propagation models*, IEEE Conference on Data Mining, IEEE Computer Society, 2012, pp. 81–90.

[Che52] Herman Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. **23** (1952), no. 4, 493–507.

[Doe18] Benjamin Doerr, *Probabilistic tools for the analysis of randomized optimization heuristics*, CoRR **abs/1801.06733** (2018).

[FGN+19] Tobias Friedrich, Andreas Göbel, Frank Neumann, Francesco Quinzan, and Ralf Rothenberger, *Greedy maximization of functions with bounded curvature under partition matroid constraints*, The 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 2019, pp. 2272–2279.

[FHK17] Moran Feldman, Christopher Harshaw, and Amin Karbasi, *Greed is good: Near-optimal submodular maximization via greedy optimization*, Proceedings of the 2017 Conference on Learning Theory, PMLR, 07–10 Jul 2017, pp. 758–784.

[HFWK19] Chris Harshaw, Moran Feldman, Justin Ward, and Amin Karbasi, *Submodular maximization beyond non-negativity: Guarantees, fast algorithms, and applications*, Proceedings of the 36th International Conference on Machine Learning, PMLR, 09–15 Jun 2019, pp. 2634–2643.

[HL12] Tad Hogg and Kristina Lerman, *Social dynamics of Digg*, EPJ Data Science **1** (2012), no. 1, 5.

[JKB95] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan, *Continuous univariate distributions. Vol. 2*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., 1995. MR 1326603

[KG14] Andreas Krause and Daniel Golovin, *Submodular function maximization*, Tractability: Practical Approaches to Hard Problems, Cambridge University Press, 2014, pp. 71–104.

[KKT03] David Kempe, Jon M. Kleinberg, and Éva Tardos, *Maximizing the spread of influence through a social network*, Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2003, ACM, 2003, pp. 137–146.

[KKT15] ———, *Maximizing the spread of influence through a social network*, Theory of Computing **11** (2015), 105–147.

[KMN99] Samir Khuller, Anna Moss, and Joseph Naor, *The budgeted maximum coverage problem*, Information Processing Letters **70** (1999), no. 1, 39–45.

[LKG+07] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne M. VanBriesen, and Natalie S. Glance, *Cost-effective outbreak detection in networks*, Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2007, ACM, 2007, pp. 420–429.

[MR95] Rajeev Motwani and Prabhakar Raghavan, *Randomized algorithms*, Cambridge University Press, 1995.

[MU05] Michael Mitzenmacher and Eli Upfal, *Probability and computing: Randomized algorithms and probabilistic analysis*, Cambridge University Press, 2005.

[NW78] G. L. Nemhauser and L. A. Wolsey, *Best algorithms for approximating the maximum of a submodular set function*, Mathematics of Operations Research **3** (1978), no. 3, 177–188.

[NWF78] George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher, *An analysis of approximations for maximizing submodular set functions - I*, Math. Program. **14** (1978), no. 1, 265–294.

[QSY+17] Chao Qian, Jing-Cheng Shi, Yang Yu, Ke Tang, and Zhi-Hua Zhou, *Subset selection under noise*, Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 2017, pp. 3563–3573.

[QSYT17] Chao Qian, Jing-Cheng Shi, Yang Yu, and Ke Tang, *On subset selection with general cost constraints*, International Joint Conference on Artificial Intelligence, IJCAI 2017, 2017, pp. 2613–2619.

[QYZ15] Chao Qian, Yang Yu, and Zhi-Hua Zhou, *Subset selection by Pareto optimization*, Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS 2015, 2015, pp. 1774–1782.

[RA15] Ryan A. Rossi and Nesreen K. Ahmed, *The network data repository with interactive graph analytics and visualization*, Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI 2015, AAAI Press, 2015, pp. 4292–4293.

[RNNF19] Vahid Roostapour, Aneta Neumann, Frank Neumann, and Tobias Friedrich, *Pareto optimization for subset selection with dynamic cost constraints*, The 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, AAAI Press, 2019, pp. 2354–2361.

[Von10] Jan Vondrák, *Submodularity and curvature: The optimal algorithm*, RIMS Kôkyûroku Bessatsu **B23** (2010), 253–266.

[XHA+19] Yue Xie, Oscar Harper, Hirad Assimi, Aneta Neumann, and Frank Neumann, *Evolutionary algorithms for the chance-constrained knapsack problem*, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2019, ACM, 2019, pp. 338–346.

[ZV16] Haifeng Zhang and Yevgeniy Vorobeychik, *Submodular optimization with routing constraints*, Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI 2016, AAAI Press, 2016, pp. 819–826.