STOCHASTIC-LAZIER-GREEDY ALGORITHM FOR MONOTONE NON-SUBMODULAR MAXIMIZATION

LU HAN
Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100190, P.R. China

MIN LI
School of Mathematics and Statistics
Shandong Normal University, Jinan 250014, P.R. China

DACHUAN XU
Department of Operations Research and Scientific Computing
Beijing University of Technology, Beijing 100124, P.R. China

DONGMEI ZHANG*
School of Computer Science and Technology
Shandong Jianzhu University, Jinan 250101, P.R. China

(Communicated by Wenxun Xing)

Abstract. The problem of maximizing a given set function with a cardinality constraint has widespread applications. A number of algorithms have been provided to solve the maximization problem when the set function is monotone and submodular. However, reality-based set functions may not be submodular and may involve large-scale and noisy data sets. In this paper, we present the Stochastic-Lazier-Greedy Algorithm (SLG) to solve the corresponding non-submodular maximization problem and offer a performance guarantee of the algorithm. The guarantee is related to a submodularity ratio, which characterizes the closeness to submodularity. Our algorithm also can be viewed as an extension of several previous greedy algorithms.

1. Introduction. The problem of maximizing a given set function with a cardinality constraint (i.e., set function maximization) has received widespread attention in many applications. For example, sensor placement [11, 12], influence maximization [9], sparse regression [15], active learning [5, 7], and document summarization [14]. In the set function maximization, we are given a ground set \(V\) of size \(n\) and a cardinality \(k\). A function \(f : 2^V \rightarrow \mathbb{R}\) is also given to measure the utility value over subsets of \(V\). We aim to find a subset \(S \subseteq V\) of a size at most \(k\) such that the utility value is maximized. The set function maximization is generally NP-hard even when the function is submodular [4, 18]. Many greedy algorithms were proposed to solve this problem [8, 10, 16, 17, 18, 20].

2010 Mathematics Subject Classification. Primary: 58F15, 58F17; Secondary: 53C35.
Key words and phrases. Set function maximization, cardinality constraint, non-submodular, monotone, greedy algorithm.

* Corresponding author: Dongmei Zhang.
The most fundamental and classical problem of the set function maximization is maximizing a monotone submodular function with a cardinality constraint (i.e., submodular maximization). A function \( f : 2^V \rightarrow \mathbb{R} \) is submodular, if for any subset \( V' \subseteq V \), the gain of the utility value by adding a new element to it is greater than the gain of the utility value of any its superset. We call the gain of the utility value of adding an element to a subset as the marginal contribution of the element to the subset. A function \( f : 2^V \rightarrow \mathbb{R} \) is monotone, if for any subset \( V' \subseteq V \), the utility value of itself is less than the utility value of any its superset. Submodular functions model many practical applications \([1, 3, 6, 13, 21]\), and the submodular maximization versatile in both artificial intelligence and machine learning. The most well-known algorithm for the submodular maximization is the Standard-Greedy Algorithm, which is proposed by Nemhauser et al. \([18]\). This algorithm starts from an empty set and iteratively adds an element, with largest marginal contribution, from the ground set to the current set until the size of the set reaches \( k \). The Standard-Greedy Algorithm obtains a \((1 - 1/e)\)-approximation guarantee, which is the optimal guarantee for the submodular maximization unless \( P=NP \) \([4]\), with \( O(nk) \) function evaluations. By exploiting the submodularity, Minoux \([16]\) gives the Lazy-Greedy Algorithm to accelerate the running time of function evaluations. Although the number of function evaluations of the Lazy-Greedy Algorithm is unsure, this algorithm leads speedup and performs better than the Standard-Greedy Algorithm in practice. Later, Mirzasoleiman et al. \([17]\) present a prominent Lazier-Greedy Algorithm. This algorithm starts from an empty set and iteratively adds an element, with largest marginal contribution, from a random subset to the current set until the size of the set reaches \( k \). The Lazier-Greedy Algorithm can obtain an approximation guarantee of \((1 - 1/e - \epsilon)\) in expectation with \( O(n \log(1/\epsilon)) \) times of function evaluations. It is notable that this algorithm is the first algorithm where its number of function evaluations is independent of the cardinality \( k \), and it runs much faster than the Lazy-Greedy Algorithm in practical scenarios.

In many realistic situations, the submodular maximization problem may involve large-scale and noisy data sets. Hassidim and Singer \([8]\) propose the Stochastic-Standard-Greedy Algorithm, which is a stochastic variant of the well-known Standard-Greedy Algorithm, to deal the large-scale and noisy applications. In the Stochastic-Standard-Greedy Algorithm, a probability distribution \( D \) with expectation \( \mu \) is given. In each iteration, a value \( \xi \) is sampled from the distribution. Instead of iteratively adding an element with largest marginal contribution as in the Standard-Greedy Algorithm, this algorithm adds an element whose marginal contribution is at least a factor \( \xi \) of the largest marginal contribution (i.e., an element which is \( \xi \)-approximation of the largest marginal contribution). For the submodular maximization, the Stochastic-Standard-Greedy Algorithm obtains an approximation guarantee of \((1 - e^{-\mu})\).

For the set function maximization, it often involves non-submodular functions in real-world applications. We call the problem of maximizing a monotone non-submodular function with a cardinality constraint as the non-submodular maximization. Khanna et al. \([10]\) show that the Lazier-Greedy Algorithm provides an approximation guarantee for this problem. Recently, Qian et al. \([20]\) show that the Stochastic-Standard-Greedy Algorithm can give a \((1 - e^{-\mu \gamma_{\min}})\)-approximation guarantee for the non-submodular maximization, where \( \gamma_{\min} \) is related to a submodularity ratio that characterizes the closeness of the function to be submodular \([2]\).
Despite Qian et al. [20] show that the Stochastic-Standard-Greedy Algorithm can deal with the non-submodular maximization when the data sets are large-scaled and noisy. While, this algorithm just as the Standard-Greedy Algorithm, needs $O(nk)$ function evaluations. This fact stimulates us to design a new algorithm with lesser function evaluations. As our main contribution, we propose the Stochastic-Lazier-Greedy Algorithm (SLG), which is a stochastic variant of the less-function-evaluated Lazier-Greedy Algorithm, to solve the non-submodular maximization with large-scaled and noisy data sets in a speedier way. We also give the performance guarantee of the SLG, which is related to a submodularity ratio. For any function, its submodularity characterizes the closeness of itself to be submodular. It is worth mentioning that the SLG covers the Standard-Greedy Algorithm, the Lazier-Greedy Algorithm and the Stochastic-Standard-Greedy Algorithm as special cases.

The remainder structure of this paper is as follows. Sect. 2 describes the set function maximization and gives the definition of submodularity ratio of a function. Sect. 3 presents the SLG along with the analysis of its performance guarantee. Sect. 4 concludes the paper.

2. Preliminaries. In this section, we formally describe the function maximization. The definition of submodularity ratio is also shown in the preliminaries.

In the set function maximization, we are given a ground set $V = \{v_1, v_2, ..., v_n\}$ and a cardinality $k \in \{1, 2, ..., n\}$. A set function $f : 2^V \rightarrow \mathbb{R}$ is given to measure the utility value over subsets of $V$, where $\mathbb{R}$ denotes the set of reals. Our goal is to find a subset $S \subseteq V$ such that $|S| \leq k$ and the utility value is maximized.

The followings are some definitions.

- For any $S \subseteq V$, $v \in V$, denote $f_S(v)$ as the marginal contribution of $v$ to $S$ (i.e., $f_S(v) = f(S \cup \{v\}) - f(S)$).
- For any $S, T \subseteq V$, denote $f_S(T)$ as the marginal contribution of $T$ to $S$ (i.e., $f_S(T) = f(S \cup T) - f(S)$).
- Set function $f : 2^V \rightarrow \mathbb{R}$ is submodular: If for any $S, T \subseteq V$, we have
  \[
  f(S \cup T) + f(S \cap T) \leq f(S) + f(T),
  \]
  or equivalently, for any $S \subseteq T \subseteq V$ and $v \in V \setminus T$,
  \[
  f_T(v) \leq f_S(v),
  \]
  or equivalently, for any $S \subseteq T \subseteq V$,
  \[
  f(T) - f(S) \leq \sum_{v \in T \setminus S} f_S(v).
  \]

When the function $f$ in the set function maximization is monotone and submodular, the problem becomes the submodular maximization and when $f$ is monotone, the problem becomes the non-submodular maximization.

The definition of the $\gamma$-submodularity ratio, which is introduced by Das and Kempe [2], can demonstrate how close a function $f$ is to be submodular. Since the ratio will be used in the latter sections, we give its definition for the sake of completeness.

**Definition 2.1.** ($\gamma$-submodularity ratio) For a set function $f : 2^V \rightarrow \mathbb{R}$ with respect to a subset $S \subseteq V$ and a parameter $k \geq 1$, its submodularity ratio is

\[
\gamma_{S,k}(f) = \min_{T, L : L \subseteq S, |T| \leq k, T \cap L = \emptyset} \frac{\sum_{v \in T} f_L(v)}{f_L(T)}.
\]
When the set function $f$ is known, we use $\gamma_{S,k}$ as $\gamma_{S,k}(f)$ for short. From the definition of the $\gamma$-submodularity ratio, we always have that $\gamma_{S,k}(f) \leq 1$ for any $f$ with respect to any $S$ and $k$. Note that if $f$ is submodular, it is easy to see that $\gamma_{S,k}(f) = 1$ for any $S$ and $k$.

**Remark 1.** When the set function $f$ is non-submodular, we have that $\gamma_{S,k}(f) \leq 1$ for any $S$ and $k$.

3. The stochastic-lazier-greedy algorithm. The stochastic version of greedy algorithms are used to deal with the set function maximization that has large-scale and noisy data sets. In this section, We introduce the Stochastic-Lazier-Greedy Algorithm (SLG) for the non-submodular maximization with large-scale and noisy data sets and analyze the performance guarantee of the SLG.

The SLG is formally presented as Algorithm 1. In this algorithm, a distribution $\mathcal{D}$ with expectation $\mu$ is given. The algorithm starts with an empty set $S$. In each iteration, the algorithm first randomly samples a value $\xi \sim \mathcal{D}$ and a subset $R$ of size $s$ from $V \setminus S$, where the size $s$ is an input. Then, an element $v \in R$, which is $\xi$-approximation of the largest marginal contribution of the elements in $R$ to $S$, is added to $S$. When the size of $S$ is equal to $k$, the algorithm stops.

Stochastic-Lazier-Greedy

**Input:** a set function $f$, a cardinality $k$, a size $s$, and a distribution $\mathcal{D}$.

**Output:** a set $S$ with $k$ elements.

```plaintext
set $S \leftarrow \emptyset$
while $|S| < k$
do
set $\xi \leftarrow$ obtained by randomly sampling from $\mathcal{D}$
set $R \leftarrow$ a subset obtained by randomly sampling $s$ elements from $V \setminus S$
set $v \leftarrow$ an element from $R$, such that $f_S(v) \geq \xi \cdot \max_{u \in R} f_S(u)$
update $S \leftarrow S \cup \{v\}$
end while
return $S$
```

If we set $s = (n/k) \log(1/\epsilon)$, Algorithm 1 needs $O(n \log(1/\epsilon))$ function evaluations, since there are $k$ iterations and each iteration evaluates the function at most $(n/k) \log(1/\epsilon)$ times. In comparison with evaluations of $O(nk)$ of the previous Stochastic-Standard-Greedy Algorithm, the function evaluations of the SLG is lesser. The evaluations of the SLG is in linear-time of the size of the ground set and independent of the cardinality $k$.

Before the analysis of the guarantee of the SLG for non-submodular maximization, we introduce some notations. Let $S_i$ be the subset $S$ in the SLG after $i$ iterations. Denote $\xi_i$, $R_i$ and $v_i$ as the selected value, subset and element during the $i$th iteration of the algorithm, respectively. Thus, $S_i = \{v_1, v_2, ..., v_i\}$. Let $S^*$ denote the optimal subset for the non-submodular maximization, and let $OPT$ denote the utility value of $S^*$ (i.e., $OPT = f(S^*)$). It is clear that $|S^*| = k$. 

The following lemma is essential for the analysis of the approximation guarantee. The proof of this lemma does not involve the property of monotone or submodular.

**Lemma 3.1.** For the subset $S_i$, which is obtained after $i$ iterations of the Stochastic-Lazier-Greedy Algorithm, we have

$$E[f(S_{i+1}) - f(S_i) \mid S_i] \geq \frac{(1 - \epsilon) \mu}{k} \sum_{v \in S^* \setminus S_i} f_S(v).$$

**Proof.** From the SLG, we know that subset $R_{i+1}$ is obtained by randomly sampling $s$ elements from $V \setminus S_i$ in the $(i+1)$th iteration. Thus,

$$\Pr[R_{i+1} \cap (S^* \setminus S_i) = \emptyset] = \left(1 - \frac{|S^* \setminus S_i|}{|V \setminus S_i|}\right)^s \leq e^{-s|S^* \setminus S_i|} \leq e^{-\frac{s|S^* \setminus S_i|}{k}}.$$  

The first inequality holds by $1 + x \leq e^x$ for any $x$. Therefore,

$$\Pr[R_{i+1} \cap (S^* \setminus S_i) \neq \emptyset] \geq 1 - e^{-\frac{s|S^* \setminus S_i|}{k}} \geq \left(1 - e^{-\frac{s|S^* \setminus S_i|}{k}}\right) \frac{|S^* \setminus S_i|}{k}.$$  

The last inequality can be understood from a diagrammatic figure. We draw $1 - e^{-(s/n)x}$ as a function of $x$ in Figure 1. From the figure, we have the following equality

$$a \frac{|S^* \setminus S_i|}{k} = 1 - e^{-\frac{s|S^* \setminus S_i|}{k}},$$

i.e.,

$$a = (1 - e^{-\frac{s}{n}}) \cdot \frac{|S^* \setminus S_i|}{k}.$$  

Since the function $1 - e^{-(s/n)x}$ is a concave function of $x$, we have that

$$a \leq 1 - e^{-\frac{s}{n}} |S^* \setminus S_i|.$$  

(1)

Combining (1) and (2), it is clear that the inequality holds. In the SLG, we set $s = n/k \log(1/\epsilon)$, thus

$$\Pr[R_{i+1} \cap (S^* \setminus S_i) \neq \emptyset] \geq (1 - \epsilon) \frac{|S^* \setminus S_i|}{k}.$$  

(3)
Let \( S'_i = R_{i+1} \cap (S^* \setminus S_i) \). An uniformly random element in \( S'_i \) is as in \( S^* \setminus S_i \), since \( R_{i+1} \) is equally likely to contain each element in \( S^* \setminus S_i \). Thus,

\[
\mathbb{E}[f(S_{i+1}) - f(S_i) \mid S_i] 
\geq \Pr[S'_i \neq \emptyset] \cdot \mathbb{E}[f(S_{i+1}) - f(S_i) \mid S_i : S'_i \neq \emptyset] 
= \Pr[S'_i \neq \emptyset] \cdot \mathbb{E} \left[ \max_{u \in R_{i+1}} \xi_{i+1} f_{S_i}(u) \mid S_i : S'_i \neq \emptyset \right] 
\geq \Pr[S'_i \neq \emptyset] \cdot \mathbb{E} \left[ \max_{u \in S'_i} \xi_{i+1} f_{S_i}(u) \mid S_i : S'_i \neq \emptyset \right] 
\geq \Pr[S'_i \neq \emptyset] \cdot \frac{\xi_{i+1}}{|S^* \setminus S_i|} \sum_{v \in S^* \setminus S_i} f_{S_i}(v).
\]  

(4)

Taking the expectation of \( \xi_{i+1} \) over both sides of (4), we have that

\[
\mathbb{E}[f(S_{i+1}) - f(S_i) \mid S_i] \geq \Pr[S'_i \neq \emptyset] \cdot \frac{\mu}{|S^* \setminus S_i|} \sum_{v \in S^* \setminus S_i} f_{S_i}(v).
\]  

(5)

Combining (3) and (5), we complete the proof of this lemma.

The following lemma shows that in each iteration the selected element can effectively reduce the gap of utility value to the optimal subset.

**Lemma 3.2.** For subset \( S_i \), which is obtained after \( i \) iterations of the Stochastic-Lazier-Greedy Algorithm, we have

\[
\mathbb{E}[f(S_{i+1}) - f(S_i) \mid S_i] \geq \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} (OPT - f(S_i)).
\]

**Proof.** From Lemma 3.1, the following inequalities holds for \( S_i \).

\[
\begin{align*}
\mathbb{E}[f(S_{i+1}) - f(S_i) \mid S_i] & \geq \frac{(1 - \epsilon)\mu}{k} \sum_{v \in S^* \setminus S_i} f_{S_i}(v) \\
& \geq \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} \cdot f_{S_i}(S^* \setminus S_i) \\
& = \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} (f(S^* \cup S_i) - f(S_i)) \\
& \geq \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} (f(S^*) - f(S_i)) \\
& = \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} (OPT - f(S_i)).
\end{align*}
\]

The second inequality is by the definition of the \( \gamma \)-submodularity ratio. The third inequality is by the monotonicity of \( f \).

Define \( \gamma_{\min} \) as \( \min_{S_i \mid |S_i| = k-1} \gamma_{S_i,k} \). Now we are ready to show the performance guarantee, which is achieved by the SLG, for the non-submodular maximization.

**Theorem 3.3.** For the non-submodular maximization, the Stochastic-Lazier-Greedy Algorithm achieves a \( (1 - e^{-\mu \gamma_{\min}} - \epsilon) \)-approximation guarantee in expectation.

**Proof.** Taking expectation of \( S_i \) for Lemma 3.2, we have

\[
\mathbb{E}[f(S_{i+1})] - \mathbb{E}[f(S_i)] \geq \frac{(1 - \epsilon)\mu \gamma_{S_i,k}}{k} (OPT - \mathbb{E}[f(S_i)]) \\
\geq \frac{(1 - \epsilon)\mu \gamma_{\min}}{k} (OPT - \mathbb{E}[f(S_i)]).
\]
The last inequality is by the definition of $\gamma_{\text{min}}$. By adding $E[f(S_i)]$ on both sides of the inequality, we have that

$$E[f(S_{i+1})] \geq \frac{(1 - \epsilon)\mu\gamma_{\text{min}}}{k} \cdot \text{OPT} + \left(1 - \frac{(1 - \epsilon)\mu\gamma_{\text{min}}}{k}\right)E[f(S_i)].$$

Therefore, we obtain the following claim, which can be proved by induction.

**Claim.** For any $i \in \{1, 2, ..., k\}$,

$$E[f(S_i)] \geq \left(1 - \left(1 - \frac{(1 - \epsilon)\mu\gamma_{\text{min}}}{k}\right)^i\right) \cdot \text{OPT}.$$

Thus, we have

$$E[f(S_k)] \geq \left(1 - \left(1 - \frac{(1 - \epsilon)\mu\gamma_{\text{min}}}{k}\right)^k\right) \cdot \text{OPT} \geq \left(1 - e^{-(1-\epsilon)\mu\gamma_{\text{min}}}\right) \cdot \text{OPT} \geq \left(1 - e^{-\mu\gamma_{\text{min}} - \epsilon}\right) \cdot \text{OPT}.$$ 

This completes the proof of this theorem. □

As aforementioned, our SLG includes several greedy algorithms as special cases.

**Remark 2.** The followings are all the greedy algorithms covered by the SLG.

- **The Standard-Greedy Algorithm**
  When $D$ is the Bernoulli distribution with expectation $\mu = 1$ and $s = n$, the SLG reduces to the Standard-Greedy Algorithm.

- **The Lazier-Greedy Algorithm**
  When $D$ is the Bernoulli distribution with expectation $\mu = 1$, the SLG reduces to the Lazier-Greedy Algorithm.

- **The Stochastic-Standard-Greedy Algorithm**
  When $s = n$, the SLG reduces to the Stochastic-Standard-Greedy Algorithm.

4. **Conclusion.** In this paper, we concentrate on solving the non-submodular maximization with large-scale and noisy data sets. Our main contribution is to propose the SLG, which generalizes several previous greedy algorithms. We also provide an approximation guarantee for the SLG. The number of function evaluations in the SLG is in linear-time and independent on the cardinality. Comparing with the recent Stochastic-Standard-Greedy Algorithm, our algorithm requires less computational cost. We do believe that the SLG makes an essential step for solving large-scale and noisy non-submodular maximization.

**Acknowledgments.** The first and third authors are supported by Natural Science Foundation of China (No. 11531014). The second author is supported by Natural Science Foundation of Shandong Province (No. ZR2019MA032) of China. The third and fourth authors are supported by Natural Science Foundation of China (No. 11871081).

**REFERENCES**

[1] A. Dasgupta, R. Kumar and S. Ravi, Summarization through submodularity and dispersion, *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, (2013), 1014–1022.

[2] A. Das and D. Kempe, Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection, *Proceedings of the 28th International Conference on International Conference on Machine Learning*, (2011), 1057–1064.
[3] K. El-Arini and C. Guestrin, Beyond keyword search: Discovering relevant scientific literature, Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2011), 439–447.

[4] U. Feige, A threshold of $\ln n$ for approximating set cover, Journal of the ACM, 45 (1998), 634–652.

[5] D. Golovin and A. Krause, Adaptive submodularity: Theory and applications in active learning and stochastic optimization, Journal of Artificial Intelligence Research, 42 (2011), 427–486.

[6] R. Gomes and A. Krause, Budgeted nonparametric learning from data streams, Proceedings of the 27th International Conference on International Conference on Machine Learning, (2010), 391–398.

[7] A. Guillory and J. Bilmes, Active semi-supervised learning using submodular functions, Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence, (2011), 274–282.

[8] A. Hassidim and Y. Singer, Robust guarantees of stochastic greedy algorithms, Proceedings of the 34th International Conference on Machine Learning, (2017), 1424–1432.

[9] D. Kempe, J. Kleinberg and É. Tardos, Maximizing the spread of influence through a social network, Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2003), 137–146.

[10] Khanna, E. Elenberg, A. Dimakis, S. Negahban and J. Ghosh, Scalable greedy feature selection via weak submodularity, Artificial Intelligence and Statistics, (2017), 1560–1568.

[11] A. Krause, H. B. McMahan, C. Guestrin and A. Gupta, Robust submodular observation selection, Journal of Machine Learning Research, 9 (2008), 2761–2801.

[12] A. Krause, A. Singh and C. Guestrin, Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies, Journal of Machine Learning Research, 9 (2008), 235–284.

[13] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen and N. Glance, Cost-effective outbreak detection in networks, Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2007), 420–429.

[14] H. Lin and J. Bilmes, Multi-document summarization via budgeted maximization of submodular functions, Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, (2010), 912–920.

[15] A. Miller, Subset Selection in Regression, Second edition. Monographs on Statistics and Applied Probability, 95. Chapman & Hall/CRC, Boca Raton, FL, 2002.

[16] M. Minoux, Accelerated greedy algorithms for maximizing submodular set functions, Optimization Techniques, 7 (1978), 234–243.

[17] B. Mirzasoleiman, A. Badanidiyuru, A. Karbasi, J. Vondrák and A. Krause, Lazier than lazy greedy, Proceedings of the 29th AAAI Conference on Artificial Intelligence, (2015), 1812–1818.

[18] G. L. Nemhauser and L. A. Wolsey, Best algorithms for approximating the maximum of a submodular set function, Mathematics of Operations Research, 3 (1978), 177–188.

[19] G. L. Nemhauser, L. A. Wolsey and M. L. Fisher, An analysis of approximations for maximizing submodular set functions - I, Mathematical Programming, 14 (1978), 265–294.

[20] C. Qian, Y. Yu and K. Tang, Approximation guarantees of stochastic greedy algorithms for subset selection, International Joint Conferences on Artificial Intelligence Organization, (2018), 1478–1484.

[21] R. Sipos, A. Swaminathan, P. Shivashwamy, and T. Joachims, Temporal corpus summarization using submodular word covering, Proceedings of the 21st ACM International Conference on Information and Knowledge Management (2012), 754–763.