國立臺灣大學電資學院電機工程學系

碩士論文

Department of Electrical Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master's Thesis

# 最大多重覆蓋問題的近似演算法

# Approximation algorithms for maximum multi-coverage problem

周志芳

Chih-Fan Chou

指導教授: 陳和麟 博士

Advisor: Ho-Lin Chen Ph.D.

中華民國 114 年 7 月

July, 2025

doi:10.6342/NTU202504145

# 國立臺灣大學碩士學位論文
# 口試委員會審定書
## MASTER'S THESIS ACCEPTANCE CERTIFICATE
## NATIONAL TAIWAN UNIVERSITY

最大多重覆蓋問題的近似演算法

Approximation algorithms for maximum multi-coverage problem

本論文係＿＿周志芳＿＿＿r12921088＿＿＿在國立臺灣大學電機工程學系完成之碩士學位論文，於民國114年7月24日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department of Electrical Engineering on 24.7.114 have examined a Master's thesis entitled above presented by Chih-Fan, Chou r12921088 candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

（指導教授 Advisor）

系主任 Director: 李建模

doi:10.6342/NTU202504145

# Acknowledgements

這篇論文的完成，首先我要感謝的是我的指導教授陳和麟老師。老師給予我自由選擇自己想要研究的主題，而每次與老師的討論中，老師也都會仔細地與我分析論文的定位與可研究的方向。不論我有什麼樣的問題，老師總是很有耐心地與我討論並給予建議。在撰寫論文的過程中，老師也提供了我許多寶貴的指導與協助。我也要感謝我的口試委員們——顏嗣鈞老師、于天立老師與呂學一老師——給予我指點與許多建設性的建議。

最後，我要感謝我的實驗室同學們，與我分享許多經驗，讓我能更快融入新的環境，也讓我的碩士生活更有方向。我也要感謝實習的夥伴們，陪伴我度過碩士最後一年的時光。我們一起讀書、一起吃飯、互相鼓勵與幫助，讓我能夠愉快且充實地度過碩士最後一年的日子。

周志芳

2025.08.07



doi:10.6342/NTU202504145

doi:10.6342/NTU202504145

# 摘要

集合覆蓋問題（Set Cover Problem, SCP）是一個經典的 NP-困難組合優化問題，具有重要的理論與實務價值。其變體如最大覆蓋與多重覆蓋問題，透過加入額外的約束與覆蓋需求，擴展了問題的應用範圍。我們的工作聚焦於線性上的多重覆蓋問題，目標是在有限點數下選擇點集，以最大化區間的完整多重覆蓋。此問題與兩個主要領域密切相關：資訊擴散領域—其中基於獨立級聯模型（Independent Cascade Model）與線性閾值模型（Linear Threshold Model）的基礎研究，促使影響力最大化與選舉控制等課題獲得廣泛關注；以及幾何最大覆蓋領域，此領域在區間與圓盤覆蓋問題上擁有強大的近似演算法與多項式時間求解法，應用範圍涵蓋設施選址、影像處理等。這些連結啟發我們研究連結兩者的廣義最大多重覆蓋問題。

本文針對最大點多重覆蓋問題（Maximum Points Multi-Cover, MPMC）處理三個主要議題。首先，對於區間重疊情形下的 MPMC 問題，我們提出了一種可達成全多項式時間近似方案（FPTAS）的演算法。其次，在排序區間的 MPMC 問題中，我們設計出一個多項式時間演算法，能達到 1/2 的近似保證。最後，針對一般的 MPMC 問題，我們設計了一個多項式時間演算法，其近似比可達 $(1-1/e)/2f-1$，其中 $f$ 表示所有區間中的最大覆蓋需求。

**關鍵字：**最大覆蓋、幾何覆蓋、多重覆蓋



doi:10.6342/NTU202504145

doi:10.6342/NTU202504145

# Abstract

The Set Cover Problem (SCP) is a classic NP-hard combinatorial optimization problem with significant theoretical and practical applications. Variants such as maximum coverage and multi-cover extend its scope by incorporating additional constraints and coverage requirements. Our work focuses on the multi-cover problem on a line, aiming to select limited points to maximize full multi-coverage of intervals. This problem is closely related to two key domains: information diffusion—where foundational models like the Independent Cascade and Linear Threshold models have spurred extensive research on influence maximization and election control—and geometric maximum coverage, a field with strong approximation results and polynomial-time algorithms for problems involving interval and disk coverage, with applications ranging from facility location to image processing. These connections motivate our study of generalized maximum multi-coverage problems bridging both areas.

In this paper, we address three main problems related to the Maximum Points Multi-



doi:10.6342/NTU202504145

Cover (MPMC) problem. First, we consider the MPMC problem on overlapping intervals. We propose an algorithm that yields a Fully Polynomial-Time Approximation Scheme (FPTAS) for this setting.Second, we study the MPMC problem on sorted intervals. For this case, we develop a polynomial-time algorithm that achieves a $1/2$-approximation guarantee.Third, we tackle the general MPMC problem by designing a polynomial-time algorithms.the algorithm achieves an approximation ratio to $(1 - 1/e)/(2f - 1)$, where $f$ denotes the maximum coverage requirement among all intervals.

**Keywords:** maximum coverage, geometric coverage, multi-cover



doi:10.6342/NTU202504145

# Contents




doi:10.6342/NTU202504145

doi:10.6342/NTU202504145

# List of Tables





doi:10.6342/NTU202504145

doi:10.6342/NTU202504145

# Chapter 1 Introduction

## 1.1 Introduction

The Set Cover Problem (SCP) is a fundamental combinatorial optimization problem in computer science and operations research. Given a universe $\mathcal{U}$ of elements and a collection $\mathcal{S}$ of subsets of $\mathcal{U}$, where each subset $S_i \in \mathcal{S}$ has an associated cost $c_i$, the objective is to find a minimum-cost subcollection $\mathcal{C} \subseteq \mathcal{S}$ whose union equals $\mathcal{U}$. This NP-hard problem has been extensively studied due to its theoretical significance and its wide range of applications, such as wireless sensor network deployment, facility location, and service planning.

In this paper, we focus on the problem of selecting points on a line to multi-cover a given set of intervals. Specifically, given several intervals on a line, our goal is to choose a set of points such that as many intervals as possible are covered multiple times. We will investigate several special cases of this general problem. The formal definition of the problem is presented as follows. Given a set $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ of $n$ intervals on a line, a set $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ of $m$ points on the same line, and a positive integer $K$, where each interval $I_i$ is associated with a covering requirement $q_i \in \mathbb{Z}^+$, the MPMC problem seeks a set $T$ of points on the line with $|T| = K$ that maximizes the number of fully covered intervals. An interval $I_i$ is considered fully covered if it contains at least $q_i$


doi:10.6342/NTU202504145

points from $T$, that is, if $|T \cap I_i| \geq q_i$. The objective is to determine such a set $T$ that maximizes the total number of intervals in $\mathcal{I}$ that are fully covered. In this work, we study certain subproblems and structural properties of the MPMC problem.

Several important variants of the set cover problem have also been studied. One such variant is the *maximum coverage problem*, where, given an integer $k$, the goal is to select $k$ subsets from $\mathcal{S}$ in order to maximize the number of covered elements in $\mathcal{U}$. Another extension is the *multi-cover* problem, in which each element $i \in \mathcal{U}$ has a coverage requirement $q_i$, and is considered fully covered only if it is covered by at least $q_i$ selected subsets. This generalization is particularly useful in practical scenarios where different elements may require different levels of coverage. For example, in wireless sensor networks, certain locations with higher population density may require more sensors to ensure reliable coverage.

## 1.2 Motivation

Our work is primarily related to two research areas: *information diffusion* and *geometric maximum coverage*.

In the domain of information diffusion, the foundational work by Kempe et al. [6] introduced the influence maximization problem under the Independent Cascade Model and Linear Threshold Model . They proposed a greedy algorithm leveraging the submodularity property of influence spread, achieving a $(1 - \frac{1}{e})$-approximation guarantee, which has since inspired extensive research on topics such as election control, influence maximization, and information diffusion in social networks. Building upon this, Wilder et al. [9] formalized the problem of election control through social influence. They considered a


doi:10.6342/NTU202504145

social network of voters where, under Independent Cascade Model, information about a candidate diffuses starting from a selected subset of voters. As the diffusion progresses, voters who receive the information update their preferences by promoting the candidate in their ranking. The goal is to select a fixed-size initial set of voters to maximize the candidate's chances of winning the election. For the constructive version of the problem, their algorithm achieves a $(1-1/e)/3$-approximation guarantee. In contrast, for the destructive version, they achieve a tighter approximation ratio of $(1-1/e)/2$.

Further advancing this line of work, Corò et al. [2] studied preference updates under the Linear Threshold Model . Their work generalized the classical Linear Threshold Model to better capture realistic scenarios where individuals may react differently to various pieces of information, allowing for more complex updates to their preference lists beyond simple one-position shifts. This generalization addresses limitations in previous models, which typically assumed uniform and minimal preference changes. They still achieve a $(1-1/e)/3$-approximation in a more general setting.

Another important concept related to information diffusion is the *majority illusion*, recently studied from an algorithmic perspective by Grandi et al. [4]. They investigated the problem of verifying and eliminating majority illusions in social networks. Given a network and a parameter $q \in [0,1]$ representing the fraction of nodes under illusion, they showed that deciding whether a labeling inducing a $q$-majority illusion exists is NP-complete for every rational $q \in (1/2,1]$, even on restricted graph classes such as planar graphs, bipartite graphs, and graphs with bounded maximum degree. Additionally, Venema-Los et al. [7] explored the relationships between weak and strict versions of majority illusion, as well as their interplay with local and global majority and majority opposition concepts.


doi:10.6342/NTU202504145

In the domain of information diffusion, many studies are based on the Linear Threshold Model and the Independent Cascade Model. However, both models rely on probabilities to describe the diffusion process. To gain a deeper understanding of models applicable to information diffusion or control, we further explore the field of geometric maximum coverage. This domain emphasizes geometric conditions, which may better reflect realistic scenarios. Moreover, these approaches do not incorporate probabilistic elements, offering alternative perspectives for information control.

In the domain of geometric maximum coverage, the max set coverage problem admits approximation within a factor of $(1 - 1/e)$ [5]. Among various topics in geometric set coverage, the following two papers are particularly relevant to our work. The first is the line-cover-disks problem, which, in one dimension, corresponds to the problem of selecting points to cover intervals. The second is the intervals-cover-points problem, which is closely related to our research and connected to certain subcases of our main problem. These connections and their implications will be elaborated in the following chapter.

In the first related work, Ran et al. [8] focused on the maximum multi-cover problem on a line, where intervals are used to cover points multiple times. They proposed a dynamic programming approach that achieves a 2-approximation and solves the problem in polynomial time.

The max coverage problem where points are used to cover intervals—where we want to find a set of $k$ points that intersect the maximum number of input intervals—is known as the partial interval hitting set problem. Damaschke [3] gave an $O(kn^2)$-time algorithm. Chung et al. [1] investigated maximum coverage problems involving lines covering disks under different constraints, including maximum coverage by $k$ arbitrary lines, $k$ parallel


doi:10.6342/NTU202504145

lines, and $k$ lines passing through a fixed point. They provided polynomial-time algorithms for two of these variants.

These geometric coverage problems have numerous applications. For example, in facility location, one may want to place linear facilities close to a set of demand regions to maximize service coverage. Similarly, in satellite deployment, placing $k$ satellites in different orbits to cover the largest number of population centers is a related problem. Moreover, such coverage models also find use in digital image processing tasks like edge detection, where lines or curves are used to detect edges by analyzing pixel arrangements.

These studies provide a rich theoretical foundation that motivates our investigation into coverage and influence problems in networks, particularly focusing on variants that extend classical models to more general and practical settings. In this paper, we primarily study the maximum multi-coverage problem, where points are used to cover intervals. Our work generalizes previous research and applies to both information diffusion and geometric maximum coverage contexts.

## 1.3 main result

In this paper, we address three main problems related to the Maximum Points Multi-Cover (MPMC) problem. Since when all the intervals are disjoint, the problem can be viewed as a knapsack problem where each item has a value of 1 and a weight equal to the coverage requirement. Given that the knapsack problem is known to be NP-complete, our problem is also NP-complete by reduction. Therefore, we focus on designing and analyzing approximation algorithms to tackle these problems effectively.

First, we consider the MPMC problem on overlapping intervals. We propose a dy-


doi:10.6342/NTU202504145

namic programming algorithm that yields a Fully Polynomial-Time Approximation Scheme (FPTAS) for this setting. Since the problem is known to be equivalent to the knapsack problem in certain special cases, we address it in a more general setting and develop a FPTAS, a well-established approximation approach for the knapsack problem.

Second, we study the MPMC problem on sorted intervals. For this case, we develop a polynomial-time algorithm that achieves a $\frac{1}{2}$-approximation guarantee. In the MPMC problem on sorted intervals, we establish a relationship with the Multi-Cover Points with Intervals (MCPI) problem studied by Yingli Ran et al. [8]. This connection provides valuable insights and highlights the significance of our problem, offering new perspectives for its solution.

Third, we tackle the general MPMC problem by designing two polynomial-time algorithms. the algorithm achieves an approximation ratio to $\frac{1-1/e}{2f-1}$, where $f$ denotes the maximum coverage requirement among all intervals. While the general maximum coverage problem is challenging to solve, we leverage the geometric structure to develop an approximation solution for this problem.

These results contribute new algorithmic insights and approximation guarantees for various variants of the MPMC problem.

Table 1.1: Comparison with Previous Work

| | **Line** | **Multi-Cover** | **restriction** | **Approximation** | **Time** |
|---|---|---|---|---|---|
| [5] | × | × | × | $(1-1/e)$ | $O\left(n^2\right)$ |
| [8] | ✓ | ✓ | Interval cover point | 2 | $O\left(k^6 m^4\right)$ |
| [1] | ✓ | × | × | × | $O\left(n^6\right)$ |
| theorem 1 | ✓ | ✓ | Overlapping Intervals | $(1-\varepsilon)$ | $O\left(n^3/\varepsilon\right)$ |
| theorem 2 | ✓ | ✓ | sorted intervals | 2 | $O\left(k^8 n^4\right)$ |
| theorem 4 | ✓ | ✓ | × | $\frac{1-1/e}{2f-1}$ | $O\left(n \log n\right)$ |


doi:10.6342/NTU202504145

# Chapter 2 Model

In this chapter, we introduce the problems addressed in this paper and highlight the distinctions among the various variants. We also provide detailed definitions for each problem considered.

In the first problem, The line represents attitudes toward a specific topic, ranging from positive to negative; the intervals represent the range of channels individuals are likely to watch based on these channel attitudes, and the points correspond to the channels themselves. We assume that individuals only watch channels whose content closely aligns with their own views. In reality, the number of available channels is limited. As a result, most people with extremely positive or negative attitudes toward a single topic tend to concentrate on only a few specific channels. This leads to groups of individuals with similar attitudes predominantly consuming content from the same set of channels. In this problem, we use groups as the input, where each group is associated with the number of intervals that can be covered. We modify the input format for this subproblem because a simpler input format can be used in this variant. Therefore, by using groups as the input, we demonstrate that in this setting it is possible to achieve a Fully Polynomial-Time Approximation Scheme (FPTAS) while covering a large number of intervals.

In the second problem, the line represents attitudes toward a specific topic, ranging


doi:10.6342/NTU202504145

from positive to negative. We assume that each individual only watches channels whose attitudes fall within a limited range that aligns with their own views. In other words, people tend to engage with content that matches their personal stance on the topic. Therefore, we focus on the scenario where each person watches channels within a restricted attitude spectrum.

Finally, we study the most general problem, which encompasses multiple cases of channel distributions with respect to the specific topic. Investigating this general problem is highly effective as it can model various real-world scenarios. Motivated by these observations, we formulate the following variants.

## 2.1 Problem Formulation : Maximum Points Multi-Cover (MPMC)

Given a set $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ of $n$ intervals on a line, a set $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ of $m$ points on the same line, and a positive integer $K$, where each interval $I_i$ is associated with a covering requirement $q_i \in \mathbb{Z}^+$, the *Maximum Points Multi-Cover* (MPMC) problem asks for selecting a set $T$ of points from $P$ with $|T| = K$ that maximizes the number of fully covered intervals. An interval $I_i$ is considered fully covered if it contains at least $q_i$ points from $T$, i.e., $|T \cap I_i| \geq q_i$. The objective is to determine such a set $T$ that maximizes the total number of intervals in $\mathcal{I}$ that are fully covered. In this paper, we focus on studying certain variants of the Maximum Points Multi-Cover (MPMC) problem. To facilitate this, we introduce and formally define several related variants that capture different aspects or restrictions of the original problem.


doi:10.6342/NTU202504145

## 2.2 Variants

**Definition 2.2.1** (Maximum Points Multi-Cover on fully Overlapping Intervals)**.** *Given a set of groups $\mathcal{G} = \{g_1, g_2, \ldots, g_n\}$ of exactly $n$ groups on a line, where each group $g_i$ consists of several overlapping intervals and is assigned a covering requirement $q_i \in \mathbb{Z}^+$. Note that different groups may include some of the same intervals. Specifically, each group $g_i$ is characterized by four parameters: $GI_i$, the number of intervals within the group; $Gq_i$, the largest covering requirement $q_i$ among its intervals; $Gl_i$, the leftmost endpoint of all intervals in the group; and $Gr_i$, the rightmost endpoint of all intervals in the group. If the covering requirement of a group is satisfied, then all intervals within that group are considered covered. If two groups overlap, the group with the higher covering requirement must contain more overlapping intervals at the same position than the other group. This is because the group with the higher covering requirement must be able to cover all intervals whose covering requirements are lower than its own. Additionally, we are allowed to place points at any positions on the line, and let $K$ be a positive integer. The* Maximum Points Multi-Cover on Fully Overlapping Intervals *problem asks for selecting a set $T$ of $K$ points on the line, i.e., $T$ is a set of points with $|T| = K$, so as to maximize the number of intervals that are fully covered. An interval $I_i$ is considered fully covered if it contains at least $q_i$ points from $T$, i.e., $|T \cap I_i| \geq q_i$. The objective is to determine such a set $T$ that maximizes the total number of intervals in $\mathcal{I}$ that are fully covered. We say an instance is* fully overlapping *if every pair of intervals that are not disjoint actually overlap. Two intervals $i$ and $j$ are said to be* overlapping *if they share the same endpoints, i.e., $l(i) = l(j)$ and $r(i) = r(j)$, where $l(\cdot)$ and $r(\cdot)$ denote the left and right endpoints of an interval, respectively.*


doi:10.6342/NTU202504145

**Definition 2.2.2** (Maximum Points Multi-Cover on Sorted Intervals)**.** *Given an instance of the MPMC problem where the set of intervals $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ is said to be* sorted *if both their left endpoints and right endpoints are in non-decreasing order. That is,*

$$\ell(I_1) \leq \ell(I_2) \leq \cdots \leq \ell(I_n) \quad \textit{and} \quad r(I_1) \leq r(I_2) \leq \cdots \leq r(I_n),$$

*where $\ell(I_i)$ and $r(I_i)$ denote the left and right endpoints of interval $I_i$, respectively.*


doi:10.6342/NTU202504145

# Chapter 3 Problems

## 3.1 Problem 1:MPMC on fully overlapped intervals

There exists an instance of the problem where all groups are disjoint. In this special case, the problem reduces to the classic knapsack problem, where each item has a value of number of intervals within the group and a weight corresponding to the covering requirement of fully covering each group. Consequently, since the knapsack problem is known to be NP-complete, this problem is also NP-complete.In this section, our main result is a Fully Polynomial-Time Approximation Scheme (FPTAS) algorithm. In MPMC on fully overlapped intervals problem, The line represents attitudes toward a specific topic, ranging from positive to negative; the intervals represent the range of channels individuals are likely to watch based on these channel attitudes, and the points correspond to the channels themselves. We assume that individuals only watch channels whose content closely aligns with their own views. In reality, the number of available channels is limited. As a result, most people with extremely positive or negative attitudes toward a single topic tend to concentrate on only a few specific channels. This leads to groups of individuals with similar attitudes predominantly consuming content from the same set of channels.

**Algorithm:** It is a knapsack-like algorithm, and the dynamic programming state


doi:10.6342/NTU202504145

transitions are defined as follows.

$$A[l,j] = \begin{cases} \min\left(A[l-1,j],\ A[l-1,j-V[l,k]] + W[l,k]\right) & \text{if } V[l,k] \leq j \\ A[l-1,j] & \text{otherwise} \end{cases}$$

The algorithm is an FPTAS algorithm. In this algorithm, we only consider intervals that can be covered. Specifically, if the covering requirement of interval $i$ exceeds the number of points within the range of $I_i$, we exclude this interval since it cannot be fully covered. We refer to position $i$ as the $i$-th overlapping interval from left to right. Let $L$ be a one-dimensional array, where $L[i]$ denotes the number of groups of overlapping intervals at position $i$. If multiple intervals with the same requirement overlap at position $i$, they are counted as a single group. For example, $L[i] = 2$ indicates that there are two groups of overlapping intervals at position $i$, even if there are five intervals in total (such as three intervals with requirement 3 and two intervals with requirement 5). The two-dimensional array $W$ is used to record the requirement value of each group, where $W[i,k]$ represents the requirement of the $k$-th group of overlapping intervals at position $i$. Similarly, $\bar{v}_i$ denotes the value obtained by rounding $v_i$ with respect to the parameter $b$. The term $V[i,k]$ represents the rounded number of intervals at position $i$ within the $k$-th group. The rounded value in $V[i,k]$ is computed as $\lceil \frac{\text{number of intervals at position } i \text{ within the } k\text{-th group}}{b} \rceil \times b$. The array $A$ functions as the dynamic programming (DP) table, storing optimal solutions for subproblems. Here, the integer $T$ denotes the sum of the values across all groups at each position., and $b$ is defined as $b = \left\lceil \frac{\epsilon \cdot \text{Max number of interval in the groups}}{2n} \right\rceil$, where $\epsilon$ is a given approximation parameter, $V_{\max}$ is the maximum value considered, and $n$ is the input size.

We iterate through each group of intervals, updating the dynamic programming table


doi:10.6342/NTU202504145

$A[i, j]$ at each step. Here, $A[i, j]$ represents the minimum total weight required to achieve a value of $j$ using only the first $i$ groups. For each group $i$, we consider whether to include the current group or not: If we do not include group $i$, the value remains $A[i-1, j]$. If we include group $i$ (and if $V[i, k] \leq j$), the value is updated as $A[i-1, j - V[i, k] + W[i, k]$. At each step, we take the minimum of these two options to ensure the optimal solution is maintained for every possible value $j$.

---

**Algorithm 1:** Set Multi-cover On Lines And Dots

---

**Input:** 1D array $L$, 2D arrays $A$, $V$, $W$, integers $b$, $T$

**for** $l \leftarrow 1;\ l \leq |L|;\ l = l + 1$ **do**
    **for** $k \leftarrow 1;\ k \leq L[l];\ k = k + 1$ **do**
        **for** $j \leftarrow b;\ j \leq T;\ j = j + b$ **do**
            $A[l, j] \leftarrow \min\left(A[l-1, j],\ A[l-1, j - V[l, k]] + W[l, k]\right)$;

$answer \leftarrow 0$;

**for** $i \leftarrow T;\ i \geq 0;\ i = i - b$ **do**
    **if** $A[|L|][i] \leq K$ **then**
        $answer \leftarrow i$;
        **break**;

**Output:** $answer$

---

**Theorem 1.** *The MPMC problem on overlapped intervals admits a fully polynomial-time approximation scheme (FPTAS), and can be solved exactly in pseudo-polynomial time.*

**Proof:** We use $\overline{\text{Value}}(\cdot)$ to denote the output value of the problem after rounding, and $\text{Value}(\cdot)$ to denote the output value before rounding. The notation OPT refers to the solution set of the original (before rounding) problem, while ALG denotes the solution set of the rounded (after rounding) problem. In first equation,ALG is the optimal solution in


doi:10.6342/NTU202504145

$\overline{\text{Value}}(\cdot)$.Therefore when we put ALG in $\overline{\text{Value}}(\cdot)$ the value of ALG will larger than OPT.

$$\overline{\text{Value}}(\text{ALG}) \geq \overline{\text{Value}}(\text{OPT}). \tag{1}$$

In second equation,OPT is the optimal solution in $\text{Value}(\cdot)$.Therefore when we put OPT in $\text{Value}(\cdot)$ the value of OPT will larger than ALG.

$$\text{Value}(\text{ALG}) \leq \text{Value}(\text{OPT}). \tag{2}$$

In the third equation, the rounded value is greater than the original value before rounding.

$$\text{Value}(\text{ALG}) \leq \text{Value}(\text{OPT}) \leq \overline{\text{Value}}(\text{OPT}). \tag{3}$$

In the fourth equation, ALG denotes the optimal solution set for the rounded problem. Therefore, the value of ALG in the rounded problem must be at least as large as the value of OPT in the rounded problem.

$$\text{Value}(\text{ALG}) \leq \overline{\text{Value}}(\text{ALG}). \tag{4}$$

In the fifth equation, since the difference between the value after rounding and the value before rounding is at most $b$, and at most $n$ items can be selected, the result of the original problem plus $nb$ is greater than or equal to the result after rounding.

$$\text{Value}(\text{ALG}) \leq \text{Value}(\text{ALG}) + nb. \tag{5}$$

In the sixth equation, we substitute the variable $b$ into the expression.

$$\text{Value}(\text{ALG}) \leq \text{Value}(\text{ALG}) + \frac{\epsilon}{2}\max(v_i). \tag{6}$$


doi:10.6342/NTU202504145

In the seventh equation, we note that each element can be selected, so the solution set can always include the item with the largest value.

$$\text{Value}(\text{ALG}) \leq \text{Value}(\text{ALG}) + \frac{\epsilon}{2}\text{Value}(\text{OPT}). \tag{7}$$

In the eighth equation, based on the previous equations, the value in the seventh equation is greater than the value of OPT in the original problem.

$$\text{Value}(\text{OPT}) \leq \text{Value}(\text{ALG}) + \frac{\epsilon}{2}\text{Value}(\text{OPT}). \tag{8}$$

In the ninth equation, we move $\frac{\varepsilon}{2} \cdot \text{Value}(\text{OPT})$ to the left-hand side of the eighth equation.

$$\left(1 - \frac{\epsilon}{2}\right) \text{Value}(\text{OPT}) \leq \text{Value}(\text{ALG}). \tag{9}$$

In the tenth equation, since $1 - \frac{\varepsilon}{2} \geq \frac{1}{1+\varepsilon}$, we obtain the following equation.

$$\frac{1}{1+\epsilon}\text{Value}(\text{OPT}) \leq \text{Value}(\text{ALG}). \tag{10}$$

## 3.2 Problem 2:MPMC on sorted intervals

To begin, we introduce the related problem of multi-covering points with intervals. There exists a strong connection between this problem and the Maximum Point Multi-Coverage (MPMC) problem on sorted intervals. Notably, we can leverage the algorithm developed by Yingli Ran et al. [8] to address our problem effectively. This connection enables us to adapt established techniques and results to our specific setting, facilitating a more efficient solution approach. In the MPMC on sorted intervals problem, the line represents attitudes toward a specific topic, ranging from positive to negative. We assume


doi:10.6342/NTU202504145

that each individual only watches channels whose attitudes fall within a limited range that aligns with their own views. In other words, people tend to engage with content that matches their personal stance on the topic. Therefore, we focus on the scenario where each person watches channels within a restricted attitude spectrum.

**Definition 3.2.1** (Multi-Cover Points with Intervals(MCPI))**.** *Given a set of $n$ **points** $\mathcal{P} = \{p_1, \ldots, p_n\}$ and $m$ **intervals** $\mathcal{I} = \{I_1, \ldots, I_m\}$ on a real line, along with two integer parameters $k \geq 1$ and $q \geq 1$ (the coverage requirement).*

*The objective is to select a subset $\mathcal{S} \subseteq \mathcal{I}$ of **exactly** $k$ **intervals** that maximizes the number of covered points, where a point $e \in \mathcal{P}$ is considered **covered** if it is contained in **at least** $q_e$ **intervals** from $\mathcal{S}$. In the **Lemma 4.1**, we first establish the relationship between the Multi-Cover Points with Intervals (MCPI) problem and the Maximum Points Multi-Cover (MPMC) problem on sorted intervals. Furthermore, we demonstrate that the techniques developed by Yingli Ran et al. [8] can be effectively applied to our problem.*

**Lemma 4.1** *MPMC on sorted intervals is equivalent to the problem of MCPI*

**Proof:** Consider an instance of MPMC on sorted intervals involving sorted intervals. We propose a transformation where we adjust the size of these intervals. Specifically, we can scale the intervals by either making them smaller or larger, while crucially maintaining the original sorted order of their start and end points. This scaling process is performed such that the existing solution set for the original MPMC on sorted intervals instance remains valid for the transformed instance. Let $T$ denote such a transformed problem instance, where each interval is normalized to a uniform size of $2\ell$.

Our goal is to select $k$ points from the set $P$. In this transformed setting $T$, there are


doi:10.6342/NTU202504145

at most $m$ candidate points available for selection. For each of these $m$ points, we extend it into an interval of length $2\ell$, specifically $[x - \ell, x + \ell]$, creating potential "covering intervals." Simultaneously, We represent the midpoint of each original interval as a "dot," which is the point to be covered in the Interval Multi-Cover on Points (IMCP) problem. This transformation effectively reduces the Maximum Points Multi-Cover (MPMC) problem on sorted intervals to an instance of the INTERVAL MULTI-COVER ON POINTS problem, where the aim is to cover these midpoints (dots) using the newly formed intervals of length $2\ell$.

Conversely, by applying the same method in reverse, we can reduce the problem of intervals covering points to the problem of points covering intervals, establishing a strong equivalence between these two formulations.

**Theorem 2.** *Given an instance of the MPMC ON SORTED INTERVALS problem, there exists a polynomial-time algorithm that achieves a 2-approximation ratio for the optimization variant.*

**Proof:** After applying the reduction from Lemma 4.1, Problem 2 becomes computationally equivalent to the MULTI-COVER POINTS WITH INTERVALS problem. According to the work of Yingli Ran et al [8], the latter admits a polynomial-time 2-approximation algorithm. Therefore, we can efficiently solve the reduced instance of Problem 2 using this algorithm, achieving the same 2-approximation guarantee.


doi:10.6342/NTU202504145

## 3.3 Problem 3:MPMC problem

### 3.3.1 MPMC Problem with Points Allowed Multiple Selections

MPMC problem is the most general problem, which encompasses multiple cases of channel distributions with respect to the specific topic. Investigating this general problem is highly effective as it can model various real-world scenarios. First, we consider the scenario where each point in $\mathcal{P}$ can be selected multiple times; in other words, we may imagine that each point in $\mathcal{P}$ is associated with a multiplicity $f$, representing $f$ overlapping copies of the point.

**Algorithm 2:** To solve this problem, We use a greedy algorithm that scans through all points in $P$. During each iteration, it dynamically identifies the position that maximizes the coverage of currently uncovered intervals. For convenience of proof, we assume that $k$ is a multiple of $2f-1$. We select $k/(2f-1)$ points from the given set of points. In this problem, we assume that every interval can be fully covered.


doi:10.6342/NTU202504145

**Algorithm 2:** Greedy Placement of Points for Interval Multi-Cover

**Input:** $n$ intervals $I_1, I_2, \ldots, I_n$ on a line; $m$ points $P_1, P_2, \ldots, P_m$ on a line; $k$: total points to place; $q_e$: coverage requirement for each interval $e$

$f \leftarrow \max\{q_e \mid e = 1, \ldots, n\}$;

Sort all intervals by their right-end points in ascending order;

`answer` $\leftarrow \emptyset$;

**for** $i \leftarrow 1$ **to** $\frac{k}{2f-1}$ **do**

  `max_intervals` $\leftarrow 0$;

  `max_intervals_index` $\leftarrow -1$;

  **for** $j \leftarrow 1$ **to** $m$ **do**

    count $\leftarrow$ number of intervals intersecting point $P_j$;

    **if** *count* $>$ *max_intervals* **then**

      `max_intervals` $\leftarrow$ count;

      `max_intervals_index` $\leftarrow P_j$;

  `answer` $\leftarrow$ `answer` $\cup$ `2f-1 points on max_intervals_index`;

  Remove all intervals intersecting `max_intervals_index`;

**Output:** `answer`

**Lemma 3.3.1** (Submodularity of the Coverage Function)**.** *Let $\mathcal{I}$ be a collection of intervals on a line, and let*

$$f = \max_{I \in \mathcal{I}} requrie(I)$$

*denote the **maximum requirement** among all intervals, where $require(I)$ is the coverage requirement of interval $I$.*

*Let $S$ be a set whose elements are collections of $2f-1$ points, each collection corresponding to a point in $P$.*


doi:10.6342/NTU202504145

*Define the function $\mathcal{F} : \mathcal{D} \to \mathbb{N}$ such that for any $D \in \mathcal{S}$, $\mathcal{F}(D)$ equals the number of intervals* ***covered*** *by the points in $D$.*

*We say that a function $f$ is* ***submodular*** *if it satisfies a natural "diminishing returns" property: the marginal gain from adding a set to a set $D$ is at least as high as the marginal gain from adding the same set to a superset of $D$. Formally, a submodular function satisfies*

$$f(D \cup v) - f(D) \geq f(T \cup v) - f(T),$$

*for all sets $v$ and all pairs of sets $D \subseteq T$.*

*In our algorithm, since we place $2f-1$ points at each step and the maximum covering requirement is also $f$, we can cover all intervals that are covered by points placed at a point in P. Therefore, the function $F$ satisfies*

$$F(D \cup \{v\}) - F(D) \geq F(T \cup \{v\}) - F(T),$$

*Let $v \in \mathcal{S}$, and let $D$ and $T$ be subsets of $S$ such that $D \subseteq T \subseteq S$.*

*Given a smaller set of $D \subseteq S$ and a larger set $T \subseteq S$ with $D \subseteq T$, since $T$ already covers more intervals than $D$, adding the same $v \in \mathcal{S}$ to both sets results in a greater increase for $D$ than for $T$. This property shows that the function $F$ is submodular.*

*Furthermore, whenever we add a element in $S$ to the solution set, the value of $F$ never decreases; it always increases or remains the same. Thus, $F$ is also a monotone increasing function. Therefore, $F$ is a monotone submodular function.*

**Theorem 3** (Approximation for MPMC problem with a predetermined set of points P)**.** *Given an instance of the MPMC problem with requirement parameter $f$ and without a*



doi:10.6342/NTU202504145

*predetermined set of points $P$, there exists a polynomial-time algorithm that achieves a $\left(1 - \frac{1}{e}\right)/(2f - 1)$-approximation ratio for the optimization variant.*

**Proof:** Since each placement operation consumes $2f - 1$ points, the algorithm can select at most $k/(2f - 1)$ elements in $S$. Consequently, in the worst-case scenario, the coverage achieved by the algorithm is bounded by a factor of $1/2f - 1$ relative to the optimal solution (OPT).

In the subsequent phase, we employ a greedy selection strategy for $k/(2f - 1)$ elements in $S$. Given that the coverage function is monotone and submodular(**Lemma**: 3.3.1), the greedy algorithm guarantees a $(1 - 1/e)$-approximation for this subproblem.

Combining both approximation guarantees, the overall approximation ratio of the algorithm is:

$$\left(\frac{1}{f}\right) \times \left(1 - \frac{1}{e}\right) = \frac{1 - 1/e}{2f - 1}.$$

### 3.3.2 MPMC Problem

In MPMC problem , MPMC problem is the most general problem, which encompasses multiple cases of channel distributions with respect to the specific topic. Investigating this general problem is highly effective as it can model various real-world scenarios. In this section, we consider the scenario where points from the predefined set $\mathcal{P}$ cannot be selected multiple times.

**Algorithm 3:** We use a greedy algorithm that scans through all points in $P$. During each iteration, it dynamically identifies the position that maximizes the coverage of currently uncovered intervals. For convenience of proof, we assume that $k$ is a multiple of


doi:10.6342/NTU202504145

$2f-1$. We simply modify the previous algorithm by selecting $k/(2f-1)$ points from the given set of points. In this problem, we assume that every interval can be fully covered.



doi:10.6342/NTU202504145

**Algorithm 3:** Greedy Placement of Points for Interval Multi-Cover

**Input:** $n$ intervals $I_1, I_2, \ldots, I_n$ on a line; $m$ points $P_1, P_2, \ldots, P_m$ on a line; $k$: total points to place; $q_e$: coverage requirement for each interval $e$

$f \leftarrow \max\{q_e \mid e = 1, \ldots, n\}$;

Sort all intervals by their right-end points in ascending order;

`selected_indexes` $\leftarrow \emptyset$;

**for** $i \leftarrow 1$ **to** $\frac{k}{2f-1}$ **do**

    `max_intervals` $\leftarrow 0$;

    `max_intervals_index` $\leftarrow -1$;

    **for** $j \leftarrow 1$ **to** $m$ **do**

        count $\leftarrow$ number of intervals intersecting point $P_j$;

        **if** *count* $>$ *max_intervals* **then**

            `max_intervals` $\leftarrow$ count;

            `max_intervals_index` $\leftarrow P_j$;

    `selected_indexes` $\leftarrow$ `selected_index` $\cup$ `max_intervals_index`;

    Remove all intervals intersecting `max_intervals_index`;

`answer_indexes` $\leftarrow \emptyset$;

**foreach** $i \in$ *selected_indexes* **do**

    `answer_indexes` $\leftarrow$ `answer_indexes` $\cup \{i\}$;

    **for** $j \leftarrow 1$ **to** $f - 1$ **do**

        **if** *point at position* $(i - j)$ *exists* **then**

            `answer_indexes` $\leftarrow$ `answer_indexes` $\cup$ {point at $(i - j)$}

        **if** *point at position* $(i + j)$ *exists* **then**

            `answer_indexes` $\leftarrow$ `answer_indexes` $\cup$ {point at $(i + j)$}

**Output:** `answer_indexes`


doi:10.6342/NTU202504145

**Lemma 3.3.2.** *If we select the point $i$ itself along with the $f-1$ points immediately to its left and the $f-1$ points immediately to its right, resulting in a total of $2f-1$ points, then all intervals intersecting point $i$ are guaranteed to be covered.*

*Proof.* Since each interval must can be covered, the number of points within any interval is at least equal to the interval's coverage requirement. By selecting the point $i$ itself along with the $f-1$ points immediately to its left and the $f-1$ points immediately to its right, we have chosen a total of $2f-1$ points around $i$. Therefore, any interval that covers point $i$ must contain enough selected points to satisfy its coverage requirement since maximum covering requirement is f, implying that all such intervals intersecting with point i are covered. □

**Theorem 4** (Approximation for MPMC problem)**.** *Given an instance of the MPMC problem with requirement parameter $f$, there exists a polynomial-time algorithm that achieves a $\left(1-\frac{1}{e}\right)/(2f-1)$-approximation ratio for the optimization variant.*

**Proof:** In the previous problem, *MPMC Problem with Points Allowed Multiple Selections*, we select $2f-1$ points in each iteration, and as a result, all intervals intersecting with the selected point are fully covered.

In **Lemma**: 3.3.2, we show that given a point $p$ on the line, selecting $2f-1$ points around $p$ guarantees full coverage of all intervals intersecting $p$.

Therefore, since the selection method in each iteration is the same as that in the previous problem, and the previous problem achieves an approximation ratio of $(1-1/e)2f-1$(theorem 3), this problem also attains an approximation ratio of $(1-1/e)/2f-1$.


doi:10.6342/NTU202504145

# References

[1] C. Chung, A. Vigneron, and H.-K. Ahn. Maximum coverage by $k$ lines. *Symmetry*, 16(2):206, 2024.

[2] F. Corò, E. Cruciani, G. D'Angelo, and S. Ponziani. Exploiting social influence to control elections based on positional scoring rules. *Information and Computation*, 289:104940, 2022.

[3] P. Damaschke. Refined algorithms for hitting many intervals. *Information Processing Letters*, 123:15–21, 2017. Available at Chalmers University website.

[4] U. Grandi, L. Kanesh, G. Lisowski, R. Sridharan, and P. Turrini. Identifying and eliminating majority illusion in social networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 5062–5069, 2023.

[5] D. S. Hochbaum and A. Pathria. Analysis of the greedy approach in problems of maximum k-coverage. *Naval Research Logistics (NRL)*, 45(6):615–627, 1998.

[6] D. Kempe, J. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD' 03)*, pages 137–146, New York, NY, USA, 2003. Association for Computing Machinery.


doi:10.6342/NTU202504145

[7] M. Los, Z. Christoff, and D. Grossi. On the graph theory of majority illusions. arXiv preprint arXiv:2304.02258, 2023. Last revised 11 Jul 2023.

[8] Y. Ran, Y. Xu, X. Liu, and X. Zhu. Approximation algorithm for the maximum interval multi-cover problem. In Z.-Z. Chen, X. Sun, and X. Wang, editors, *Combinatorial Optimization and Applications*, pages 30–50. Springer, Singapore, 2024.

[9] B. Wilder and Y. Vorobeychik. Controlling elections through social influence. arXiv preprint arXiv:1711.08615, 2017.


doi:10.6342/NTU202504145