An Approximation Algorithm for the h-Hop Independently Submodular Maximization Problem and Its Applications

Wenzheng Xu, *Member, IEEE*, Hongbin Xie, Chenxi Wang, Weifa Liang, *Senior Member, IEEE*, Xiaohua Jia, *Fellow, IEEE*, Zichuan Xu, *Member, IEEE*, Pan Zhou, *Senior Member, IEEE*, Weigang Wu, *Member, IEEE, ACM*, and Xiang Chen, *Member, IEEE*

***Abstract*—This study is motivated by the maximum connected coverage problem (MCCP), which is to deploy a connected UAV network with given K UAVs in the top of a disaster area such that the number of users served by the UAVs is maximized. The deployed UAV network must be connected, since the received data by a UAV from its served users need to be sent to the Internet through relays of other UAVs. Motivated by this application, in this paper we study a more generalized problem – the h-hop independently submodular maximization problem, where the MCCP problem is one of its special cases with $h = 4$. We propose a $\frac{1-1/e}{2h+3}$-approximation algorithm for the h-hop independently submodular maximization problem, where e is the base of the natural logarithm. Then, one direct result is a $\frac{1-1/e}{11}$-approximate solution to the MCCP problem with $h = 4$, which significantly improves its currently best $\frac{1-1/e}{32}$-approximate solution. We finally evaluate the performance of the proposed algorithm for the MCCP problem in the application of deploying UAV networks, and experimental results show that the number of users served by deployed UAVs delivered by the proposed algorithm is up to 12.5% larger than those by existing algorithms.**

***Index Terms*—UAV communication networks, maximum connected coverage problem, connected sensor coverage problem, submodular function maximization, approximation algorithms.**

I. Introduction

IN THIS paper, we study an h-hop independently submodular maximization problem, which is defined later. We start by introducing two potential applications of the problem: one is to deploy a UAV communication network to serve people trapped in a disaster area, the other is to place a sensor network to monitor Points of Interest (PoIs) in an IoT network.

The first important application arises in the context of Unmanned Aerial Vehicles (UAV) networks. Wireless communication by leveraging the use of UAVs has attracted lots of attentions recently [13], [23], [43]. Unlike terrestrial communication systems, low-altitude UAV systems are more cost-effective by enabling on-demand operations, more swift and flexible for deployment and configuration [9], [14], [33], [36], [37], [41], [42]. Due to its maneuverability and flexibility, a UAV can act as an aerial base station (BS) by equipping with a lightweight base station device [8], [25]. It is expected that UAV networks consists of multiple UAVs are perfectly suitable for unexpected and temporary communication demands, such as natural disasters, traffic congestion, and concerts [3]. In addition, because of their high flying height, UAVs usually have higher Line-of-Sight (LoS) link opportunities with ground users, compared to terrestrial BSs [1]. Fig. 1 shows a UAV network in which four UAVs serve as aerial base stations to provide communication services to the trapped people in a disaster zone. With the help of the UAV network, the trapped people can send and receive critical voices, videos, and data to/from the rescue team, thereby saving their lives and reducing injuries. Our study is motivated by a fundamental *Maximum Connected Coverage Problem (MCCP)* [43] in a UAV network, which is to deploy K UAVs in the air to serve people in a disaster zone, such that the number of users served is maximized, subject to the constraint that the communication subnetwork induced by the K UAVs is connected. The rationale behind the connectivity constraint is that, the received data by a UAV from its served users need to be sent to a gateway UAV in the UAV network, where the

Manuscript received 9 March 2022; revised 1 August 2022; accepted 26 September 2022; approved by IEEE/ACM Transactions on Networking Editor B. Ji. The work of Wenzheng Xu was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62272328 and in part by the Double World-Class Project for Sichuan University under Grant 0082604151352. The work of Weifa Liang was supported by the City University of Hong Kong under Project CityU 9380137/CS. The work of Zichuan Xu was supported by NSFC under Grant 62172068 and Grant 61802048. The work of Pan Zhou was supported by the NSFC under Grant 61972448. The work of Weigang Wu was supported by the Guangdong Natural Science Foundation of China under Grant 2018B030312002. *(Corresponding author: Zichuan Xu.)*

Wenzheng Xu, Hongbin Xie, and Chenxi Wang are with the College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: wenzheng.xu3@gmail.com; xiehongbin@stu.scu.edu.cn; wangchenxi1998@stu.scu.edu.cn).

Weifa Liang and Xiaohua Jia are with the Department of Computer Science, City University of Hong Kong, Hong Kong, China (e-mail: weifa.liang@cityu.edu.hk; csjia@cityu.edu.hk).

Zichuan Xu is with the School of Software, Dalian University of Technology, Dalian 116024, China (e-mail: z.xu@dlut.edu.cn).

Pan Zhou is with the Hubei Engineering Research Center on Big Data Security, School of Cyber Science and Engineering, Huazhong University of Science & Technology, Wuhan 430074, China (e-mail: panzhou@hust.edu.cn).

Weigang Wu is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mail: wuweig@mail.sysu.edu.cn).

Xiang Chen is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: chenxiang@mail.sysu.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNET.2022.3210825, provided by the authors.

Digital Object Identifier 10.1109/TNET.2022.3210825

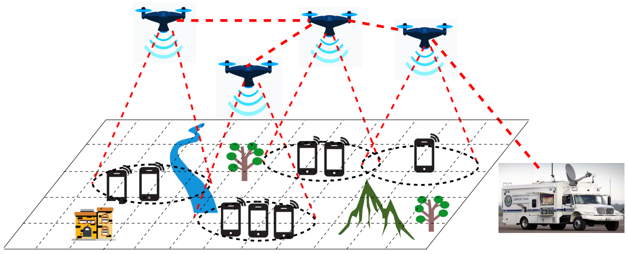


Fig. 1. A UAV network that provides communication services for ground users in a disaster area, where the network is connected to the Internet via an emergency communication vehicle.

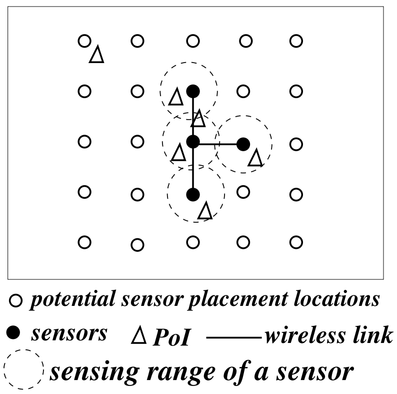


Fig. 2. An example of the connected sensor placement problem, where $K = 4$ sensors are deployed and five PoIs are monitored by the four sensors.

gateway UAV is connected to the Internet, with the help of an emergency communication vehicle or satellites, see Fig. 1.

We then focus on another application of the h-hop independently submodular maximization problem, which is to place K sensors at some strategic locations to monitor PoIs (Points of Interest) in an IoT network such that the number of PoIs monitored by the placed K sensors is maximized, subject to the constraint that the communication subnetwork induced by the K sensors is connected [15]. The rationale behind this connectivity constraint is that each sensor needs to send its sensing data to a base station directly or via the relays of other sensors. Fig. 2 shows an example of placing $K = 4$ sensors to monitor PoIs.

In addition to the aforementioned two applications, there are many other potential applications of the h-hop independently submodular maximization problem, including deploying wireless power chargers in wireless sensor networks [39], [40], placing wireless routers in wireless networks [20], choosing influential connected users in social networks [2], [17], [18], [24], [30], [32], [34], and so on.

Motivated by the aforementioned many applications, in this paper, we study a more generalized problem – the h-hop independently submodular maximization problem, which is briefly defined as follows. Given an undirected, connected graph $G = (V, E)$, let $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$ be a monotone function on the subsets of V, i.e., $f(A) \leq f(B)$ for any subsets A and B of V with $A \subseteq B$. In addition, given a positive integer $h \geq 1$, we say that f is *h-hop independently submodular* based on G if it meets the following two properties:

(i) *Submodularity*: $f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) + f(B)$ for any two subsets A and B of V with $A \subseteq B$, and any node $v \in V \setminus B$. The submodularity captures the property of *diminishing returns* in economics and many fields [6].

(ii) *h-hop independence*: $f(A) + f(B) = f(A \cup B)$ for any two non-empty subsets A and B of V with the minimum number of hops in G between any node in A and any node in B being at least h.

In this paper, we consider an *h-hop independently submodular maximization problem* in $G(V, E)$, which is to find a subset S of K nodes in V such that the value of $f(S)$ is maximized, subject to the constraint that the induced subgraph $G[S]$ of G by the nodes in S is *connected*, where K is a given positive integer with $1 \leq K \leq |V|$ and f is h-hop independently submodular. Notice that the MCCP problem for deploying a UAV network is a special case of the problem with $h = 4$, which will be shown in Section V.

There are several studies on special cases of the h-hop independently submodular maximization problem. For example, Garg [12] proposed a $\frac{1}{3+\epsilon}$-approximation algorithm for the problem when $h = 1$, where ϵ is a given constant with $0 < \epsilon \leq 1$. Notice that the submodular function f meets the additive property when $h = 1$, i.e., for any subset S of V, $f(S) = \sum_{v \in S} f(\{v\})$. Khuller *et al.* [18] proposed a $\frac{1-1/e}{12}$-approximation algorithm for the problem when $h = 3$, where e is the base of the natural logarithm. Yu *et al.* [39], [40] proposed a $\frac{1-1/e}{8(\lceil \frac{4}{\sqrt{3}}\alpha \rceil + 1)^2}$-approximation algorithm for the connected sensor placement problem (e.g., see Fig. 2), where $\alpha = \frac{r}{R}$ with $0 < r \leq R$, r and R are the sensing range and communication range of a sensor, respectively. It can be seen that the approximation ratio is a value between $\frac{1-1/e}{128}$ and $\frac{1-1/e}{32}$, as $0 < \alpha \leq 1$.

Notice that we recently devised a $\frac{1-1/e}{\sqrt{K}}$-approximation algorithm [35] for finding a set S with K nodes in a graph G such that a submodular function $f(S)$ is maximized, subject to that $G[S]$ is connected, where e is the base of the natural logarithm. This implies that the algorithm also delivers a $\frac{1-1/e}{\sqrt{K}}$-approximate solution to the problem considered in this paper. However, the approximation ratio $\frac{1-1/e}{\sqrt{K}}$ is small when K is large. In this paper, we consider the case that $\sqrt{K} \geq 2h + 3$, and propose an improved algorithm with an approximation ratio $\frac{1-1/e}{2h+3}$ for the problem, which is no less than $\frac{1-1/e}{\sqrt{K}}$.

A. Main Contributions

The main contributions of this paper are as follows.

(i) To the best of our knowledge, we are the first to introduce the h-hop independently submodular maximization problem, which generalizes many optimization problems arisen in different domains, such as the MCCP problem of deploying a UAV network to serve as many users as possible.

(ii) We propose a novel tree decomposition technique. With the help of the proposed technique, we devise a $\frac{1-1/e}{2h+3}$-approximation algorithm for the problem when $h \geq 2$. Consequently, the proposed algorithm delivers $\frac{1-1/e}{9}$ and $\frac{1-1/e}{11}$ approximate solutions to the problem, when $h = 3$ and $h = 4$, respectively, while the best approximation ratios so far for these two special cases with $h = 3$ and $h = 4$ are $\frac{1-1/e}{12}$ [18] and $\frac{1-1/e}{32}$ [39], [40], respectively.

(iii) We evaluate the performance of the proposed algorithm for the MCCP problem in the application of deploying UAV networks, and experimental results show that the number of users served by deployed UAVs in the solution delivered by the proposed algorithm is up to 12.5% larger than those by existing algorithms.

The rest of the paper is organized as follows. Section II introduces preliminaries and defines the problem. Section III proposes a $\frac{1-1/e}{2h+3}$-approximation algorithm for the h-hop independently submodular maximization problem, while Section IV shows the approximation ratio. Section V studies an application of the h-hop independently submodular maximization problem in UAV networks. Section VI evaluates the performance of the proposed algorithms. Section VII reviews related work, and Section VIII concludes the paper.

II. Preliminaries

A. Network Model

We consider an undirected, connected graph $G = (V, E)$, where V is the set of nodes and E is the set of edges. For any two nodes u and v in V, denote by $l(u, v)$ the minimum number of hops (i.e., edges) in G between nodes u and v. Also, for any two non-empty subsets A and B of V, denote by $l(A, B)$ the minimum number of hops between nodes in A and B, i.e., $l(A, B) = \min_{u\in A, v\in B}\{l(u, v)\}$.

We consider a nondecreasing submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, which meets the following three properties:

(i) $f(\emptyset) = 0$;

(ii) Monotonicity: $f(A) \leq f(B)$ for any two subsets A and B of V with $A \subseteq B$; and

(iii) Submodularity: $f(A\cup\{v\})-f(A) \geq f(B\cup\{v\})+f(B)$ for any two subsets A and B of V with $A \subseteq B$, and any node $v \in V \setminus B$.

B. Problem Definition

A function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$ is an *h-hop independently submodular function* in a graph $G = (V, E)$ if and only if (i) f is nondecreasing and submodular; and (ii) for any two non-empty subsets A and B of V, if the minimum number of hops between the nodes in A and the nodes in B is no less than h (i.e., $l(A, B) \geq h$), then $f(A) + f(B) = f(A \cup B)$, where $h \geq 1$ is a given positive integer.

In this paper, we consider *an h-hop independently submodular maximization problem*, which is defined as follows. Given an undirected, connected graph $G = (V, E)$, an h-hop independently submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, and a positive integer K, the problem is to find a set S of K nodes in V such that the value of $f(S)$ is maximized, subject to the constraint that the induced subgraph $G[S]$ by the nodes in S is connected.

We assume that the values of h and K satisfy the following relationship: $2h + 3 \leq \sqrt{K}$. The rationale behind the assumption is as follows. Xu *et al.* [35] devised a $\frac{1-1/e}{\sqrt{K}}$-approximation algorithm for finding a set S with K nodes in G such that a submodular function $f(S)$ is maximized, subject to that $G[S]$ is connected, where e is the base of the natural logarithm. This implies that the algorithm also delivers a $\frac{1-1/e}{\sqrt{K}}$-approximate solution to the problem considered in this paper. However, the approximation ratio $\frac{1-1/e}{\sqrt{K}}$ is small when K is large. Under the assumption that $2h + 3 \leq \sqrt{K}$, we will propose an improved algorithm with an approximation ratio $\frac{1-1/e}{2h+3}$ for the problem in this paper, which is no less than $\frac{1-1/e}{\sqrt{K}}$.

C. Quota Steiner Tree (QST) Problem

We define a *Quota Steiner Tree (QST) problem* [16]. Given an undirected graph $G = (V, E)$, a profit function $p : V \mapsto \mathbb{Z}^{\geq 0}$, a cost function $c : E \mapsto \mathbb{Z}^{\geq 0}$, and a positive integer (quota) q, the problem is to find a subtree T in G such that the cost of the T, i.e., $\sum_{e\in E(T)} c(e)$, is minimized, subject to the constraint that the profit sum of nodes in T is no less than q, i.e., $\sum_{v\in V(T)} p(v) \geq q$. Notice that there is a 2-approximation algorithm for the QST problem [12], [16], and the algorithm will be part of the solution to the problem in this paper.

D. Special Cases of the h-Hop Independently Submodular Maximization Problem

We here show that the problems studied in various applications [2], [12], [15], [16], [18], [20], [24], [39], [40], [43] in fact are special cases of the h-hop independently submodular maximization problem with different values of h, which are summarized in Table I and the value of $h(\alpha)$ is defined in Eq. (1).

$$h(\alpha) = \begin{cases} 2, & \text{if } 0 < \alpha \leq \dfrac{1}{2}, \\ 3, & \text{if } \dfrac{1}{2} < \alpha \leq \dfrac{\sqrt{2}}{2}, \\ 4, & \text{if } \dfrac{\sqrt{2}}{2} < \alpha \leq 1. \end{cases} \tag{1}$$

To verify our claim, we here prove that both the budgeted prize collecting Steiner tree problem [12], [16] and the budgeted connected dominating set problem [2], [18], [24] are special cases of the h-hop independently submodular maximization problem when $h = 1$ and $h = 3$, respectively. The proofs of the other problems listed in Table I are similar to the one in Section V, omitted.

We first consider the budgeted prize collecting Steiner tree problem [12], [16]. Given a graph $G = (V, E)$, a node profit function $f : V \mapsto \mathbb{Z}^{\geq 0}$, and a budget K, the problem is to find a subset S of V with no more than K nodes, such that the induced subgraph $G[S]$ by S is connected and the profit sum of the nodes in S, i.e., $\sum_{v\in S} f(v)$, is maximized. For any two non-empty subsets A and B of V with the minimum number of hops in G between any node in A and any node in B being at least $h = 1$, i.e., $A \cap B = \emptyset$, we have that $f(A \cup B) = \sum_{v\in A\cup B} f(v) = \sum_{v\in A} f(v) + \sum_{v\in B} f(v) = f(A) + f(B)$. Therefore, the problem is a special case of the h-hop independently submodular maximization problem when $h = 1$.

TABLE I
SPECIAL CASES OF THE h-HOP INDEPENDENTLY SUBMODULAR MAXIMIZATION PROBLEM

Problems	value of h	Remarks
Budgeted prize collecting Steiner tree problem [12], [16]	1	—
Budgeted connected dominating set problem [2], [18], [24]	3	—
Maximum Connected Coverage Problem [43]	$h(\alpha)$	$\alpha = \frac{r}{R}$, $0 < r \leq R$, r: communication range between a user and a UAV, R: communication range between two UAVs
Wireless charger network deployment problem [39], [40]	$h(\alpha)$	$\alpha = \frac{r}{R}$, $0 < r \leq R$, r: charging ranging of a wireless charger, R: communication range between two chargers
Wireless router network placement problem [20]	$h(\alpha)$	$\alpha = \frac{r}{R}$, $0 < r \leq R$, r: communication range between a user and a router, R: communication range between two routers
Connected sensor placement problem [15]	$h(\alpha)$	$\alpha = \frac{r}{R}$, $0 < r \leq R$, r: sensing range of a sensor, R: communication range between two sensors

We then consider the budgeted connected dominating set problem [2], [18], [24]. Given a graph $G = (V, E)$ and a subset S of V, a node v is *dominated* by S if v is contained in S or is a neighbor of a node in S. Denote by $f(S)$ the number of nodes dominated by S. Given a budget K, the problem is to find a subset S of V with no more than K nodes, such that the induced subgraph $G[S]$ by S is connected, and the number $f(S)$ of nodes dominated by S is maximized. For any two non-empty subsets A and B of V with the minimum number of hops in G between any node in A and any node in B being at least $h = 3$, i.e., $l(A, B) \geq 3$, it can be seen that no nodes in V are dominated by both A and B at the same time. Otherwise, suppose that there is a node v in V such that v is dominated by both A and B. Then, the minimum hop between A and v is no more than one, and the minimum hop between v and B is also no more than one, i.e., $l(A, v) \leq 1$ and $l(v, B) \leq 1$. This indicates that the minimum hop between any node in A and any node in B is $l(A, B) \leq l(A, v) + l(v, B) \leq 2$, which however contradicts the assumption that $l(A, B) \geq 3$. We conclude that no nodes are dominated by both A and B simultaneously. We thus know that the number of nodes dominated by the nodes in $A \cup B$ is the sum of the numbers of nodes dominated by the nodes in A and B, respectively, i.e., $f(A \cup B) = f(A) + f(B)$. Therefore, the budgeted connected dominating set problem is a special case of the h-hop independently submodular maximization problem when $h = 3$.

III. APPROXIMATION ALGORITHM

In this section, we propose a $\frac{1-1/e}{2h+3}$-approximation algorithm for the h-hop independently submodular maximization problem.

A. Basic Idea

The basic idea behind the proposed algorithm is that we assign profits to nodes in graph G in n different ways with $n = |V|$. We find a tree T_i in G with no more than K nodes so that the profit sum of nodes in T_i is maximized in each of the n profit assignments, by invoking the 2-approximation algorithm for the QST problem, where the QST problem here is to find a tree in G such that the number of nodes in the tree is minimized, subject to the constraint that the profit sum of nodes in the tree is at least a given quota q. The solution to the problem then is the set of nodes in one of the n found trees $T_1, T_2, \ldots, T_n$ such that the profit sum of nodes in the tree is maximized.

B. Approximation Algorithm

Given an undirected, connected graph $G = (V, E)$, an h-hop independently submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, and a positive integer K, let $V = \{v_1, v_2, \ldots, v_n\}$, where $n = |V|$. We assign profits to nodes in G with n different ways.

Denote by $p_i(v)$ the profit assigned to node $v \in V$ in G in the ith way with $1 \leq i \leq n$. This profit assignment proceeds as follows.

We start by assigning a profit $f(\{v_i\})$ to node v_i, i.e., $p_i(v_i) = f(\{v_i\})$. We then choose a node v in $V \setminus \{v_i\}$ with the maximum marginal profit $f(\{v, v_i\}) - f(\{v_i\})$ and assign node v the profit $p_i(v) = f(\{v, v_i\}) - f(\{v_i\})$, where ties are broken arbitrarily. The profit assignment procedure continues until each node in G is assigned a profit. The detailed profit assignment procedure is given in `Algorithm` 1.

Algorithm 1 Profit Assignment Procedure

Require: An undirected, connected graph $G = (V, E)$, an h-hop independently submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, and a starting node v_i
Ensure: the assigned profit $p_i(v)$ of each node $v \in V$ in the ith way
Assign profit $f(\{v_i\})$ to the starting node v_i, i.e., $p_i(v_i) = f(\{v_i\})$;
Let $D \leftarrow \{v_i\}$; /* the set of nodes assigned profits already*/
Let $U \leftarrow V \setminus D$;
while $U \neq \emptyset$ **do**
  Choose a node v in U with the maximum marginal profit $f(\{v\} \cup D) - f(D)$, i.e., $v = \arg\max_{v_j \in U}\{f(\{v_j\} \cup D) - f(D)\}$;
  Let $p_i(v) = f(\{v\} \cup D) - f(D)$;
  Let $D \leftarrow D \cup \{v\}$;
  Let $U \leftarrow U \setminus \{v\}$;
end while
return the assigned profit $p_i(v)$ of each node v in V.

Having assigned a profit $p_i(v)$ to each node $v \in V$ in the ith way, we find a tree T_i with no more than K nodes such that the profit sum of the nodes in T_i is maximized, based on the profit assignment. This problem however is NP-hard [16]. Denote by q_{opt} the optimal profit sum.

In the following, we find a quota Q_i by binary search with $Q_i \leq q_{opt}$, such that there are no more than K nodes in the tree delivered by the 2-approximation algorithm for the QST problem [12], [16] with quota Q_i, while there are more than K nodes in the found tree with quota $Q_i + 1$. We later show that the value of Q_i is no less than $\frac{1-1/e}{2h+3} \cdot OPT$ for some staring node v_i (e.g., see Eq. (2) in Section IV), where OPT is the optimal solution to the h-hop independently submodular maximization problem.

It can be seen that the value of Q_i must be in the interval of $[f(\{v_i\}), f(V)]$. Specially, let lb and ub be the lower and upper bounds on Q_i, respectively. Initially, let $lb = f(\{v_i\})$ and $ub = f(V)$. Let $q = \lfloor \frac{lb+ub}{2} \rfloor$. We can find a tree T_q in G based on the profit assignment of the ith way so that the number of nodes in T_q is minimized, subject to the constraint that the profit sum of nodes in T_q is no less than q, by invoking the 2-approximation algorithm for the QST problem. Consider the number of nodes $|V(T_q)|$ in tree T_q. If $|V(T_q)| \leq K$, this implies that $q(= \lfloor \frac{lb+ub}{2} \rfloor)$ is no more than the value Q_i. In this case, let q become the updated lower bound on Q_i, i.e., $lb = q$. Otherwise ($|V(T_q)| > K$), this indicates that the value q is larger than Q_i, i.e., $q > Q_i$. Let q become the updated upper bound on Q_i, i.e., $ub = q$. The binary search will terminate when $ub = lb + 1$. Finally, the tree T_i can be found, by invoking the 2-approximation algorithm for the QST problem with a quota of $lb(= ub - 1)$.

The algorithm for the h-hop independently submodular maximization problem is presented in `Algorithm` 2.

Algorithm 2 Approximation Algorithm for the h-Hop Independently Submodular Maximization Problem

Require: An undirected, connected graph $G = (V, E)$, an h-hop independently submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, and a positive integer K.

Ensure: A set S of K nodes in G such that the value of $f(S)$ is maximized, subject to the constraint that the induced subgraph $G[S]$ is connected.

Let $S \leftarrow \emptyset$;
for $1 \leq i \leq n$ **do**
  Assign profits to nodes in V starting from node v_i by invoking `Algorithm` 1;
  Let $lb \leftarrow f(\{v_i\})$ and $ub \leftarrow f(V)$; /* lb and ub are the lower and upper bounds on the value of Q_i, respectively */
  while $lb + 1 < ub$ **do**
    Let $q \leftarrow \lfloor \frac{lb+ub}{2} \rfloor$; /* q is the quota in the QST problem */
    Find a tree T_q in G with the minimum number of nodes, subject to the constraint that the profit sum of nodes in T_q, i.e., $\sum_{v \in V(T_q)} p_i(v)$, is no less than quota q, by invoking the 2-approximation algorithm for the QST problem;
    if the number of nodes in T_q is no greater than K **then**
      Let $lb \leftarrow q$; /* the quota q is no more than Q_i */
    else
      Let $ub \leftarrow q$; /* the quota q is larger than Q_i */
    end if
  end while
  Let $q \leftarrow lb$, where $lb = ub - 1$;
  Find a tree T_i in G with the minimum number of nodes, subject to the constraint that the profit sum of nodes in T_i is no less than quota q, by invoking the 2-approximation algorithm for the QST problem. Notice that the number of nodes in T_i must be no greater than K.
  if $f(V(T_i)) > f(S)$ **then**
    Let $S \leftarrow V(T_i)$; /* find a better set of nodes */
  end if
end for
return set S.

IV. Analysis of the Approximation Algorithm

Denote by L_0 the set of nodes in an optimal solution to the problem. Then, $OPT = f(L_0)$. Also, denote by L_{h-1} the set of nodes such that the minimum number of hops in G between any node v in L_{h-1} and any node in L_0 is no more than $h-1$, but node v is not contained in L_0, i.e., $L_{h-1} = \{v \mid v \in V \setminus L_0,\ l(v, L_0) \leq h-1\}$, where $h \geq 1$, and $l(v, L_0)$ is the minimum number of hops between node v and nodes in L_0 in G. Let $L_h = V \setminus (L_0 \cup L_{h-1})$. It can be seen that the minimum number of hops between nodes in L_0 and nodes in L_h is no less than h, i.e., $l(L_0,\ L_h) \geq h$.

Consider a node v_i with the *maximum profit* in the optimal solution L_0, i.e., $v_i = \arg\max_{v \in L_0}\{f(v)\}$. Due to the submodularity of function f, we have $f(L_0) \leq \sum_{v \in L_0} f(v) \leq |L_0| \cdot f(v_i) = K \cdot f(v_i)$, where $K = |L_0|$. Then,

$$f(v_i) \geq \frac{f(L_0)}{K} = \frac{OPT}{K}. \quad (2)$$

Recall that in the ith **'for'** loop of `Algorithm` 2, we first assign a profit $p_i(v_i) = f(v_i)$ to node v_i, then assign profits to the other nodes in G greedily. Denote by D' the first K nodes in set $L_0 \cup L_{h-1}$ that have been assigned profits by the profit assignment procedure. Let $D' = \{v_i, v_1, v_2, \ldots, v_{K-1}\}$ with $i \notin \{1, 2, \ldots, K-1\}$. Denote by $p_i(D')$ the profit sum of nodes in D', i.e., $p_i(D') = \sum_{v \in D'} p_i(v)$.

Proof roadmap: In the rest, we first show that the profit sum of nodes in D' is no less than $(1 - 1/e) \cdot OPT$, i.e., $p_i(D') \geq (1-1/e)OPT$. We then prove that there is a tree T in G spanning the nodes in D', such that the number of nodes in T is no more than $(K-1)h+1$. The profit sum of nodes in T thus is no less than $(1-1/e)OPT$. We also show that tree T can be decomposed into no more than $2h+3$ subtrees such that the number of nodes in each subtree is no more than $\frac{K}{2}$. Then, there must have a subtree T' among the $2h+3$ subtrees such that the profit sum of nodes in T' is no less than $\frac{1}{2h+3}$ of the profit sum of nodes in T, i.e., $\sum_{v \in T'} p_i(v) \geq \frac{\sum_{v \in T} p_i(v)}{2h+3} \geq \frac{1-1/e}{2h+3} OPT$. Finally, a tree in G with no more than $2\frac{K}{2} = K$ nodes can be found such that the profit sum of nodes in the tree is no less than $\frac{1-1/e}{2h+3} \cdot OPT$, by invoking the 2-approximation algorithm for the QST problem in [12] and [16].

We start by showing that the profit sum of nodes in D' is no less than $(1 - 1/e) \cdot OPT$.

Lemma 1: Consider node v_i in the optimal solution L_0 with the maximum profit and the profit assignment procedure starting with node v_i. Let D' be the first K nodes in set $L_0 \cup L_{h-1}$ with the assigned profits. Then, $p_i(D') \geq (1 - 1/e) \cdot OPT$.

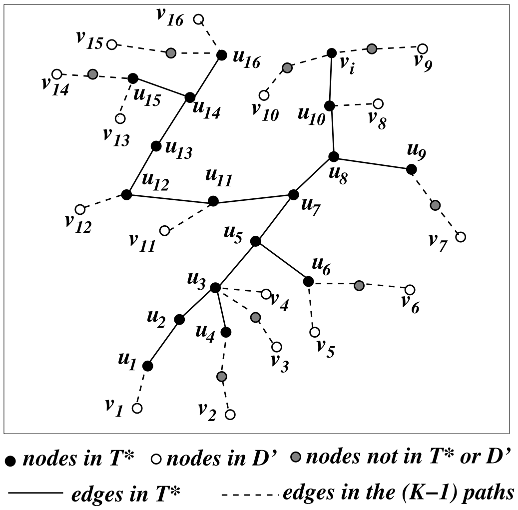


Fig. 3. The constructed tree T spanning the nodes in D', where $K = 17$ and $h = 3$. Notice that node v_i is contained by both T^* and D'.

Proof: The proof is contained in Section 1 of the supplementary file. □

A. The Existence of a Tree T With $(K-1)h+1$ Nodes That Spans All Nodes in D'

We then show that there is a tree T in G spanning the nodes in D' such that the number of nodes in T is no more than $(K-1)h+1$, which is less than Kh in [18].

Lemma 2: Given node $v_i \in L_0$ with the maximum profit and the profit function $p_i : V \mapsto \mathbb{Z}^{\geq 0}$, there is a tree T in G with no more than $(K-1)h+1$ nodes such that the profit sum of nodes in T, i.e., $\sum_{v\in V(T)} p_i(v)$, is no less than $(1-1/e)\cdot OPT$, where e is the base of the natural logarithm.

Proof: Recall that $D' = \{v_i, v_1, v_2, \ldots, v_{K-1}\}$ and $p_i(D') \geq (1-1/e)\cdot OPT$ by Lemma 1. We construct a tree T in G spanning all nodes in D' such that T contains no more than $(K-1)h+1$ nodes, based on profit function $p_i(\cdot)$.

Since L_0 is the optimal solution, the induced subgraph $G[L_0]$ by the nodes in L_0 is connected. Denote by T^* a spanning tree in $G[L_0]$, assuming that the cost of each edge is one. Notice that v_i is in L_0 and each node in D' is contained in $L_0 \cup L_{h-1}$, where L_{h-1} is the set of nodes such that the minimum number of hops in G between any node v in L_{h-1} and any node in L_0 is no greater than $h-1$, but node v is not contained in L_0. Then, it can be seen that there is a path P_k in G between any node v_k in $D'\setminus\{v_i\}$ and a node u_k in L_0 such that the number of edges in P_k is no more than $h-1$.

A tree T can be constructed, which is the union of T^* and the $K-1$ found paths, i.e., $T = T^* \bigcup(\bigcup_{k=1}^{K-1} P_k)$. Fig. 3 illustrates such a tree, where $K = 17$, $h = 3$, $L_0 = V(T^*) = \{v_i, u_1, u_2, \ldots, u_{16}\}$, $D' = \{v_i, v_1, v_2, \ldots, v_{16}\}$, and the number of edges in each path P_k is no more than $h-1 = 2$.

The number of edges in T is no more than $|E(T)| = |E(T^*)| + \sum_{k=1}^{K-1} |E(P_k)| \leq K-1+(K-1)\cdot(h-1) = (K-1)h$. The number of nodes in T thus is no greater than $(K-1)h+1$. The lemma then follows. □

B. A Novel Tree Decomposition

We now show that there is a subtree T' in G with no more than $\lfloor \frac{K}{2} \rfloor$ nodes such that the profit sum of nodes in T' is no less than $\frac{1-1/e}{2h+3}\cdot OPT$. Following Lemma 2, there is a tree T in G with no more than $(K-1)h+1$ nodes such that the profit sum $\sum_{v\in V(T)} p_i(v)$ of the nodes in T is no less than $(1-1/e)\cdot OPT$. We here propose a novel tree decomposition technique that decomposes T into no more than $2h+3$ subtrees such that the number of nodes in each subtree is no greater than $\lfloor \frac{K}{2} \rfloor$, where $2h+3 < 4h$ for any integer h with $h \geq 2$. Then, there must be a subtree T' among the $2h+3$ subtrees such that the profit sum of nodes in T' is no less than $\frac{1}{2h+3}$ of the profit sum of nodes in T, i.e., $\sum_{v\in T'} p_i(v) \geq \frac{\sum_{v\in T} p_i(v)}{2h+3} \geq \frac{1-1/e}{2h+3} OPT$.

1) Tree Decomposition Procedure: We show the tree decomposition procedure when K is odd. Then, $\lfloor \frac{K}{2} \rfloor = \frac{K-1}{2}$. On the other hand, $\lfloor \frac{K}{2} \rfloor = \frac{K}{2}$ when K is even. The procedure with the case that K is even is omitted, due to its similarity with the case that K is odd.

Recall that tree T is the union of a spanning tree T^* in $G[L_0]$ and $K-1$ paths $P_1, P_2, \ldots, P_{K-1}$, where the number of edges in P_k is no more than $h-1$ with $1 \leq k \leq K-1$, see Fig. 3. Without loss of generality, we further assume that $P_1, P_2, \ldots, P_{K-1}$ are edge-disjoint. Otherwise, the paths with edge-sharing can be converted to edge-disjoint paths, by duplicating the shared edges.

Let node $v_i \in L_0$ be the root of tree T. Denote by T_v the subtree of T rooted at node v for any node $v \in T$, and denote by $w(T_v)$ the number of edges in T_v. We decompose tree T by a Depth-First Search (DFS) starting from node v_i, until the number of edges in the residual tree is no more than $\frac{K-1}{2} - 1 = \frac{K-3}{2}$. The detailed tree decomposition procedure is given as follows.

Assume that v is the node being visited by the DFS. If the number of edges in tree T_v is no more than $\frac{K-3}{2} - 1$, i.e., $w(T_v) \leq \frac{K-3}{2} - 1$, nothing is done and the tree decomposition procedure continues; otherwise ($w(T_v) \geq \frac{K-3}{2}$ as $w(T_v)$ is an integer), a tree will be decomposed from T as follows. We later show that node v must be contained in tree T^* by Lemma 3 in Section IV-B.2, where T^* is a spanning tree in $G[L_0]$ by the optimal solution L_0.

Assume that node T_v has n_v children $v'_1, v'_2, \ldots, v'_{n_v}$. Denote by tree T'_l the union of edge (v, v'_l) and subtree $T_{v'_l}$ rooted at a child v'_l, i.e., $T'_l = (v, v'_l) \cup T_{v'_l}$, where $1 \leq l \leq n_v$.

Following the work in [31], the n_v subtrees $T'_1, T'_2, \ldots, T'_{n_v}$ can be partitioned into, say $n' (\geq 2)$, groups $g_1, g_2, \ldots, g_{n'}$ such that the number of edges of subtrees in each group is no more than $\frac{K-3}{2}$ (i.e., $\sum_{T'_l \in g_j} w(T'_l) \leq \frac{K-3}{2}$ for each j with $1 \leq j \leq n'$), while the number of edges in the subtrees of *any two groups* is larger than $\frac{K-3}{2}$ (i.e., $\sum_{T'_l \in g_j \cup g_{j'}} w(T'_l) > \frac{K-3}{2}$ for each pair of j and j' with $1 \leq j, j' \leq n'$ and $j \neq j'$). For example, Fig. 4(a) shows that tree T_{u_3} rooted at u_3 consists of four subtrees, and these four subtrees are partitioned into $n' = 2$ groups, where $K = 17$ and $\frac{K-3}{2} = 7$. Also, it can be seen that the numbers of edges in the subtrees of groups g_1 and g_2 are 6 and 3, respectively. Then, $w(g_1) = 6 \leq \frac{K-3}{2} = 7$

and $w(g_2) = 3 \le \frac{K-3}{2} = 7$, while $w(g_1) + w(g_2) = 6 + 3 = 9 > \frac{K-3}{2} = 7$.

For each group g_j with $1 \le j \le n'$, denote by n_j^* the number of edges in $g_j \cap T^*$, where T^* is a spanning tree in graph $G[L_0]$. For example, consider two groups g_1 and g_2 in Fig. 4(a). It can be seen that $n_1^* = 3$ and $n_2^* = 0$.

A tree T_j'' is decomposed from T by distinguishing into two cases. **Case (i):** the number of edges in the subtrees of a group g_j is no less than $\frac{K-3}{2} - n_j^* - (h-2)$, i.e.,

$$w(g_j) = \sum_{T_l' \in g_j} w(T_l') \ge \frac{K-3}{2} - n_j^* - (h-2). \tag{3}$$

A tree T_j'' in **Case (i)** is constructed, which is the union of the subtrees in group g_j, as each subtree in g_j contains the root v of T_v. Finally, the edges in T_j'' are removed from T. For example, Fig. 4(a) shows that the number of edges in the subtrees of group g_1 is $w(g_1) = 3 + 3 = 6 > \frac{K-3}{2} - n_1^* - (h-2) = \frac{17-3}{2} - 3 - (3-2) = 3$. Tree T_j'' thus is the union of the subtrees in group g_1, see Fig. 4(a).

Case (ii): the number of edges in all subtrees of each group g_j is no more than $\frac{K-3}{2} - n_j^* - (h-2)$, i.e., $w(g_j) < \frac{K-3}{2} - n_j^* - (h-2)$ for each j with $1 \le j \le n'$. For example, consider node u_5 in Fig. 4(b), where the number of edges in the subtrees of group g_1 (or g_2) is 4, while the value of $\frac{K-3}{2} - n_j^* - (h-2)$ is $\frac{17-3}{2} - 1 - (3-2) = 5$. The n' groups $g_1, g_2, \ldots, g_{n'}$ are sorted in non-increasing order of the numbers of their edges in T^*. Without loss of generality, assume that $n_1^* \ge n_2^* \ge \cdots \ge n_{n'}^*$, where $n_j^* = |E(g_j \cap T^*)|$ with $1 \le j \le n'$. Notice that node v is contained in each subtree $g_j \cap T^*$ with $1 \le j \le n'$, as v is contained in each subtree of g_j and v is in T^*.

A tree T_j'' in **Case (ii)** is constructed from T as follows. First, let T_j'' be the union of the subtrees in group g_1. Then, we duplicate the edges in $g_2 \cap T^*$ and add the edges to T_j''. Notice that T_j'' is connected after adding the edges in $g_2 \cap T^*$, since node v is contained in both $g_1 \cap T^*$ and $g_2 \cap T^*$. Finally, recall that for each node v_k in $D' \setminus \{v_i\}$, there is a path P_k between v_k and a node u_k in T^* such that the number of edges in P_k is no more than $h-1$. We continue adding a path P_k of a node v_k in $(D' \setminus \{v_i\}) \cap g_2$ to T_j'' as long as the number of edges in T_j'' is no more than $\frac{K-3}{2}$. For example, Fig. 4(b) illustrates such a tree, where the edge in $g_2 \cap T^*$ is (u_5, u_6), and path P_5 consisting of only edge (v_5, u_6) is added to T_j''.

Having constructed tree T_j'', the edges in T_j'' except the edges in $g_2 \cap T^*$ are removed from T, see Fig. 4(c) for the residual tree of T after the tree decomposition in **Case (ii)**.

Denote by $\mathcal{T}$ the set of the decomposed subtrees from T by the tree decomposition procedure. For example, Fig. 4(d) shows that seven subtrees are obtained through the tree decomposition of tree T. It can be seen that the number of edges of each tree in $\mathcal{T}$ is no more than $\frac{K-3}{2}$. Then, the number of nodes of each tree is no greater than $\frac{K-3}{2} + 1 = \frac{K-1}{2} \le \frac{K}{2}$.

2) Property of a Node v With $w(T_v) \ge \frac{K-3}{2}$ in Tree T:

Lemma 3: Consider a node v in the tree decomposition of T. If the number of edges in the subtree T_v rooted at v is no less than $\frac{K-3}{2}$, i.e., $w(T_v) \ge \frac{K-3}{2}$, then v must be contained in tree T^*, where T^* is a spanning tree in $G[L_0]$.

Proof: The proof is contained in Section 2 of the supplementary file. □

3) Bound the Number of Decomposed Subtrees in $\mathcal{T}$:

Lemma 4: Assume that $\sqrt{K} \ge 2h+3$. Then, the tree T in G with no more than $(K-1)h+1$ nodes can be decomposed into no more than $2h+3$ subtrees, such that the number of nodes in each subtree is no more than $\frac{K}{2}$. Then, there is a subtree T' among the $2h+3$ subtrees with no more than $\frac{K}{2}$ nodes such that the profit sum of nodes in T' is no less than $\frac{1-1/e}{2h+3} \cdot OPT$, i.e., $|V(T')| \le \frac{K}{2}$ and $\sum_{v \in T'} p_i(v) \ge \frac{1-1/e}{2h+3} \cdot OPT$.

Proof: It can be seen that the number of nodes of each subtree in $\mathcal{T}$ is no greater than $\frac{K}{2}$. We show that the number of subtrees in $\mathcal{T}$ is no more than $2h+3$. Recall that, before splitting any subtree off from tree T, T consists of a spanning tree T^* in $G[L_0]$ and $K-1$ paths $P_1, P_2, \ldots, P_{K-1}$, see Fig. 3(a). It can be seen that the numbers of edges in T^* and T are $K-1$ and $(K-1)h$, respectively, by Lemma 2.

Let $\mathcal{T} = \{T_1'', T_2'', \ldots, T_x'', T_{x+1}''\}$ be the set of decomposed subtrees of T by the tree decomposition procedure, where the number of subtrees in $\mathcal{T}$ is $x+1$ and x is a nonnegative integer.

Following the tree composition procedure, some edges of T^* will be removed when decomposing each subtree T_j'' from T, where $1 \le j \le x+1$. The set of the removed edges can be represented as $(E(T_j'') \cap E(T^*)) \setminus E(T^*)$. Let $n_j^* = |((E(T_j'') \cap E(T^*)) \setminus E(T^*)|$. It can be seen that $\sum_{j=1}^{x+1} n_j^* \le K-1$, as any edge in T^* will not be contained in other subtrees once it has been removed from T. Especially, we have

$$\sum_{j=1}^{x} n_j^* \le \sum_{j=1}^{x+1} n_j^* \le K-1, \tag{4}$$

where T_{x+1}'' is the last decomposed subtree.

We show that the number of removed edges from T after decomposing each subtree T_j'' is no less than $\frac{K-3}{2} - n_j^* - (h-2)$, i.e.,

$$w(T_j'' \setminus T) \ge \frac{K-3}{2} - n_j^* - (h-2). \tag{5}$$

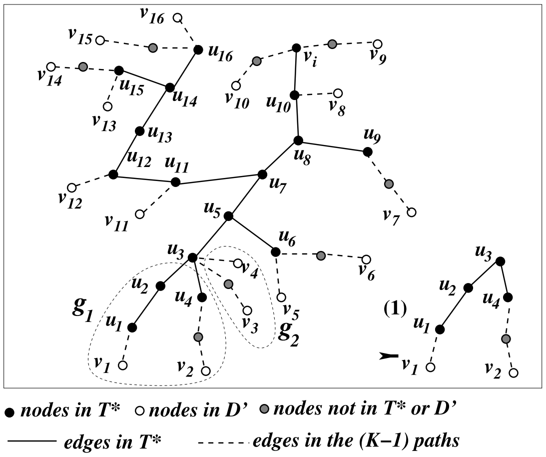
Assume that a node v in T is being visited by the DFS in the tree decomposition procedure when T_j'' is decomposed. Then, T_j'' is a subtree of T_v. Subtree T_j'' may be obtained in either Case (i) or Case (ii) of the tree decomposition procedure, see Fig. 4. For Case (i) (see Fig. 4(a)), we have

$$w(T_j'' \setminus T) = w(T_j'') \ge \frac{K-3}{2} - n_j^* - (h-2), \text{ by Ineq. (3).} \tag{6}$$

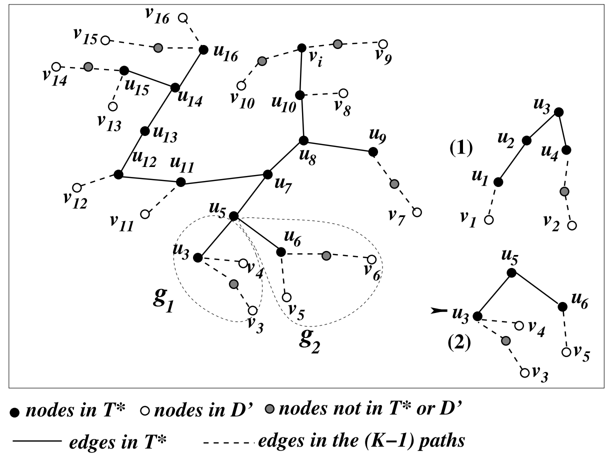
On the other hand, assume that T_j'' is obtained by Case (ii) in the tree decomposition procedure. It can be seen that the number of edges in T_j'' is at least $\frac{K-3}{2} - (h-2)$, i.e.,

$$w(T_j'') \ge \frac{K-3}{2} - (h-2). \tag{7}$$

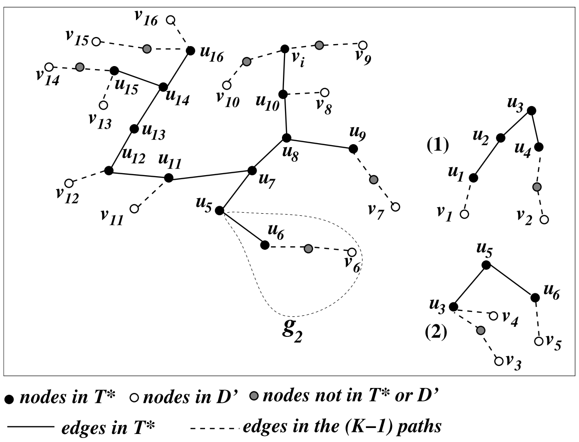
Otherwise $(w(T_j'') < \frac{K-3}{2} - (h-2))$, we have $w(T_j'') \le \frac{K-3}{2} - (h-2) - 1 = \frac{K-3}{2} - (h-1)$. We then can add another path $P_{k'}$ of a node $v_{k'}$ in $(D' \setminus \{v_i\}) \cap g_2$, such that the number of edges in T_j'' is at most $\frac{K-3}{2} - (h-1) + |E(P_{k'})| \le \frac{K-3}{2}$,



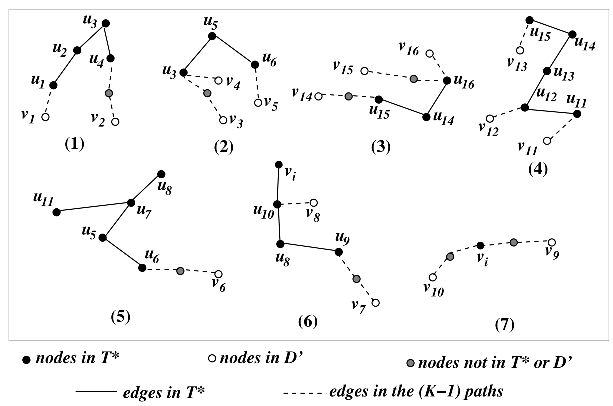
(a) **Case (i)**: the number of edges in the subtrees of group g_1 i.e., 6, is no less than $\frac{K-3}{2} - n_1^* - (h-2) = 3$, where $K = 17$, $n_1^* = |E(g_1 \cap T^*)| = 3$, and $h = 3$



(b) **Case (ii):** the number of edges in the subtrees of any group g_j, i.e., 4, is no more than $\frac{K-3}{2} - n_j^* - (h-2) = 5$, where $K = 17$, $n_j^* = |E(g_j \cap T^*)| = 1$ with $1 \le j \le 2$, and $h = 3$



(c) The residual tree of T after the tree decomposition in **Case (ii)**, and the edge (u_5, u_6) in $g_2 \cap T^*$ is not removed



(d) The obtained 7 subtrees after the tree decomposition, where $7 \le 2h+3 = 9$, and the number of edges in each subtree is no greater than $\frac{K-3}{2} = 7$

Fig. 4. The execution illustrations of the tree decomposition procedure.

as the number of edges in $P_{k'}$ is no more than $h-1$. This however contradicts the construction of tree T_j''.

It can be seen that the set of edges removed from T after decomposing subtree T_j'' by Case (ii) is $E(T_j'') \setminus E(g_2 \cap T^*)$, since the edges in $E(g_2 \cap T^*)$ are not removed from T in the tree decomposition, see Fig. 4(b) and Fig. 4(c). We then have

$$\begin{aligned} w(T_j'' \setminus T) &= w(T_j'') - |E(g_2 \cap T^*)| \\ &\ge \frac{K-3}{2} - (h-2) - |E(g_2 \cap T^*)|, \text{ by Ineq. (7)} \\ &= \frac{K-3}{2} - n_2^* - (h-2), \text{ as } n_2^* = |E(g_2 \cap T^*)| \\ &\ge \frac{K-3}{2} - n_1^* - (h-2), \text{ as } n_1^* \ge n_2^*. \end{aligned} \tag{8}$$

Combining Ineq. (6) and Ineq. (8), Ineq. (5) holds.

Since there are $(K-1)h$ edges in tree T initially, the number of edges removed from T after decomposing the first x subtrees is no greater than $(K-1)h$. We thus have

$$\begin{aligned} (K-1)h &\ge \sum_{j=1}^{x} w(T_j'' \setminus T) \\ &= \sum_{j=1}^{x} (\frac{K-3}{2} - n_j^* - (h-2)), \text{ by Ineq. (5)} \\ &= (\frac{K-1}{2} - (h-1)) \cdot x - \sum_{j=1}^{x} n_j^* \\ &\ge (\frac{K-1}{2} - (h-1)) \cdot x - (K-1), \text{ by Ineq. (4)}. \end{aligned} \tag{9}$$

Then,

$$\begin{aligned} 2(h+1) &\ge (1 - \frac{2(h-1)}{K-1}) \cdot x \\ &\ge (1 - \frac{2(h-1)}{(2h+3)^2 - 1}) \cdot x, \\ &\quad \text{by the assumption that } \sqrt{K} \ge 2h+3. \end{aligned} \tag{10}$$

By re-arranging Ineq. (10), we have

$$x \le 2h + 3 - \frac{5h+7}{2h^2 + 5h + 5}. \tag{11}$$

Since x is an integer, we have

$$x \leq 2h + 2. \tag{12}$$

Then, the number of subtrees in $\mathcal{T}$ is $x + 1 \leq 2h + 3$. For example, Fig. 4(d) shows that seven subtrees are obtained after the tree decomposition of tree T, $|\mathcal{T}| = 7 \leq 2h + 3 = 9$, and the number of edges of each subtree is no more than $\frac{K-3}{2} = 7$. The lemma then follows. □

C. Analysis of the Approximation Ratio

Lemma 5: Given node $v_i \in L_0$ with the maximum profit, assign profits to nodes in G with profit function $p_i : V \mapsto \mathbb{Z}^{\geq 0}$. Then, the 2-approximation algorithm for the QST problem in [12] and [16] can find a tree in G with no more than K nodes such that the profit sum of nodes in the tree is no less than a quota q if $q \leq \lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$. Equivalently, if the algorithm in [12] and [16] delivers a tree with more than K nodes, then the quota q is larger than $\lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$.

Proof: Following Lemma 4, there is a tree T' in G with no more than $\frac{K}{2}$ nodes such that the profit sum of nodes in T' is no less than $\lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$, as the profit sum is an integer. Therefore, tree T' is a feasible solution to the QST problem when the quota $q \leq \lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$. Then, the optimal solution to the QST problem with a quota q contains no more than $\frac{K}{2}$ nodes. We thus conclude that the tree delivered by the 2-approximation algorithm for the QST problem [12], [16] contains no more than $2 \cdot \frac{K}{2} = K$ nodes. The lemma then follows. □

We finally analyze the approximation ratio of the proposed approximation algorithm by the following theorem.

Theorem 1: Given an undirected, connected graph $G = (V, E)$, an h-hop independently submodular function $f : 2^V \mapsto \mathbb{Z}^{\geq 0}$, and a positive integer K with $K \leq |V|$, Then, there is an approximation algorithm, `Algorithm` 2, for the h-hop independently submodular maximization problem, which delivers a $\frac{1-1/e}{2h+3}$-approximate solution, where h is a given positive integer with $h \geq 2$, and e is the base of the natural logarithm. In addition, the time complexity of the algorithm is $O(n^3 T_c(f(V)) + n^4 \log n \log f(V))$, where $n = |V|$, $T_c(f(V))$ is the time for computing the value of $f(V)$.

Proof: Consider node $v_i \in L_0$ in the optimal solution with the maximum profit, and a profit function $p_i : V \mapsto \mathbb{Z}^{\geq 0}$. It can be seen that $ub = lb{+}1$ when `Algorithm` 2 terminates, where ub and lb are the upper and lower bounds on the value of $\lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$. Also, the algorithm in [12] and [16]for the QST problem delivers a tree with no more than K nodes when the quota $q = lb$, while it delivers a tree more than K nodes when the quota $q = ub$. Then, $ub > \lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil$ by Lemma 5. We thus have $ub \geq \lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil + 1$, due to that the value of ub is an integer. Therefore, $lb \geq \lceil \frac{1-1/e}{2h+3} \cdot OPT \rceil \geq \frac{1-1/e}{2h+3} \cdot OPT$. That is, the tree delivered by the algorithm for the QST problem with quota $q = lb(\geq \frac{1-1/e}{2h+3} \cdot OPT)$ in [12] and [16] contains no more than K nodes. Therefore, the approximation ratio of `Algorithm` 2 is $\frac{1-1/e}{2h+3}$.

The time complexity analysis of `Algorithm` 2 is contained in Section 3 of the supplementary file. □

V. Application of the h-Hop Independently Submodular Maximization Problem in UAV Networks

In this section, we show that the proposed algorithm for the h-hop independently submodular maximization problem is applicable to solve optimization problems by providing improved solutions to these problems. Particularly, we study its application for the MCCP problem, and show that the proposed algorithm delivers an improved $\frac{1-1/e}{11}$-approximate solution to the problem.

We first briefly describe the application scenario of the MCCP problem [43]. Assume that there are m ground users to be served in a disaster area. Denote by L and W the length and width of the area, respectively. We need to deploy a given number of K UAVs to serve the m users.

Assume that the K UAVs hover at the same altitude H_{uav}, which is the optimal altitude for the maximum coverage from the sky [1], [43], e.g., $H_{uav} = 300\ m$. Denote by R the communication range between any two UAVs at altitude H_{uav}. Also, denote by r' the communication range between a ground user and a UAV at altitude H_{uav}. Notice that r' is no greater than R [15]. Let $r = \sqrt{r'^2 - H_{uav}^2}$. Assume that a UAV hovers at a location with its coordinate (x_i, y_i, H_{uav}). Then, it can be seen that the coverage of the UAV is a disk that centers at location $(x_i, y_i, 0)$ with radius r, i.e., the set of points with coordinates $(x, y, 0)$ such that $(x - x_i)^2 + (y - y_i)^2 \leq r^2$. Thus, the users in the disk can communicate with the UAV directly. Let $\alpha = \frac{r}{R}$. Then, $0 < \alpha \leq 1$, as $r \leq r' \leq R$.

The set V of potential UAV hovering locations are constructed as follows. The plane at altitude H_{uav} is divided into equal size squares with a given side length δ, where $0 < \delta \leq R$, and R is the communication range of UAVs, assuming that R is divisible by δ, e.g., $\delta = \frac{R}{2}$. Also, assume that both the length L and width W of the disaster area are divisible by δ. Thus, the plane is divided into $n = \frac{L}{\delta} \times \frac{W}{\delta}$ grids. Denote by $v_1, v_2, \ldots, v_n$ the center locations of the n grids. The set of the potential UAV placement locations is $V = \{v_1, v_2, \ldots, v_n\}$.

Denote by $d(v_i, v_j)$ the Euclidean distance between two locations v_i and v_j in V. A graph $G = (V, E)$ then is constructed, where there is an edge (v_i, v_j) in E between two locations v_i and v_j in V if their Euclidean distance $d(v_i, v_j)$ is no greater than the communication range R, i.e., $E = \{(v_i, v_j) \mid v_i, v_j \in V, i \neq j, d(v_i, v_j) \leq R\}$. Given K UAVs, recall that *the maximum connected coverage problem* in G is to find a set S of K hovering locations in G for placing the K UAVs such that the number of ground users served by the K placed UAVs is maximized, subject to the constraint that the induced subgraph $G[S]$ by the nodes in S is connected.

Yu *et al.* [39], [40] recently proposed a $\frac{1-1/e}{8(\lceil \frac{4}{\sqrt{3}\alpha} \rceil + 1)^2}$-approximation algorithm for the problem, where $\alpha = \frac{r}{R}$. It can be seen that the approximation ratio is between $\frac{1-1/e}{128}$ and $\frac{1-1/e}{32}$, as $0 < \alpha \leq 1$.

We show that the proposed algorithm in Section III for the h-hop independently submodular maximization problem can deliver a $\frac{1-1/e}{2h(\alpha)+3}$-approximate solution to the maximum connected coverage problem, where the value $h(\alpha)$ is determined by the value of α and $\alpha = \frac{r}{R}$, see Eq. (1) in Section II-D.

For a subset S of V, denote by $f(S)$ the number of users served by the UAVs placed at locations in S. We show that function $f(S)$ is $h(\alpha)$-hop independently submodular by the following lemma.

Lemma 6: Function f is $h(\alpha)$-hop independently submodular, where the value of $h(\alpha)$ is defined in Eq. (1), and α is a given constant with $0 < \alpha \leq 1$.

Proof: The proof is contained in Section 4 of the supplementary file. □

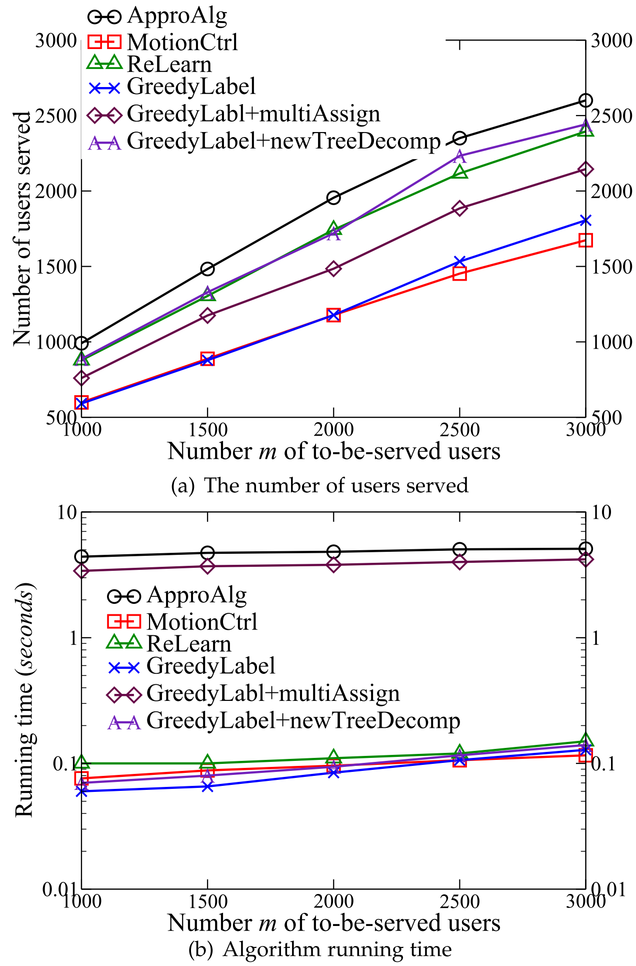
VI. Performance Evaluation

In this section, we evaluate the performance of the proposed algorithm through experimental simulations. We also study the impact of important parameters on the algorithm performance, including the number m of to-be-served users, the number K of UAVs, the communication range R between two UAVs, and the communication range r of a ground user.

A. Experimental Environment Settings

We consider an application of the problem for deploying a connected UAV network to serve ground users in a disaster area. Consider a disaster area of $3 \times 3\ km^2$ square [43], in which 500 to 3,000 users are located, where the human density follows the fat-tailed distribution, i.e., many people are located at a small portion of places while a few people are located at other places [28]. The number of deployed UAVs K varies from 10 to 50. Then, the approximation ratio of the proposed algorithm is $\frac{1-1/e}{11}$, where e is the base of the natural logarithm. We assume that each UAV hovers at altitude $H_{uav} = 300\ m$ [1]. The communication range R between any two UAVs is 600 m, while the communication range r' between a user and a UAV is 500 m [43]. Then, $r = \sqrt{r'^2 - H_{uav}^2} = 400\ m$, $\alpha = \frac{r}{R} = \frac{400}{600} = \frac{2}{3} < \frac{\sqrt{2}}{2}$. Following Eq. (1) in Section II-D, $h(\alpha) = 3$. Therefore, the maximum connected coverage problem for deploying a UAV network is an $h(\alpha)$-hop independently submodular maximization problem by Lemma 6 in Section V, and the proposed algorithm now delivers a $\frac{1-1/e}{2h(\alpha)+3}$ $(= \frac{1-1/e}{9})$ approximate solution.

To evaluate the performance of the proposed algorithm `ApproAlg` for the maximum connected coverage problem, we adopt the following three benchmarks. (i) Algorithm `MotionCtrl` [43] finds a distributed motion control solution for deploying K UAVs to cover as many as users while maintaining the connectivity of the UAVs. (ii) Algorithm `MCS` [35] delivers a $\frac{1-1/e}{\sqrt{K}}$-approximate solution to the problem of deploying K UAVs in a disaster area, such that a submodular function of the deployed UAVs is maximized, subject to the connectivity constraint that the subnetwork induced by the K UAVs is connected. (iii) Algorithm `GreedyLabel` [18] first assigns profits for deploying a UAV at different hovering locations in a greedy way, followed by identifying a connected subgraph with no more than K nodes such that the profit sum of nodes in the subgraph is maximized. All experiments were performed on a server with an Intel(R) Core(TM) i9-9900K CPU (3.6 GHz) and 32 GB RAM.



(a) The number of users served
(b) Algorithm running time

Fig. 5. The performance of different algorithms by varying the number m of to-be-served users from 500 to 3,000, when there are $K = 30$ UAVs.

B. Algorithm Performance

We first study the algorithm performance by varying the number m of users from 500 to 3,000, when there are $K(= 30)$ UAVs. Fig. 5(a) shows that the number of users served by algorithm `ApproAlg` is about from 8.5% to 12.5% higher than those by algorithms `MotionCtrl`, `MCS`, and `GreedyLabel`. For example, the numbers of users served by the four algorithms `ApproAlg`, `MotionCtrl`, `MCS`, and `GreedyLabel` are 2,600, 1,670, 2,395, and 1,800, respectively when there are 3,000 users in the disaster area. Fig. 5 demonstrates that more users are served by each of the four algorithms, with the increase on the number m of users. On the other hand, Fig. 5(b) plots the running times of the mentioned four algorithms, from which it can be seen that the running time of algorithm `ApproAlg` is about five seconds, much longer than those of other three algorithms. It must be mentioned such a short delay about a few seconds by algorithm `ApproAlg` is acceptable in a real UAV network, as up to

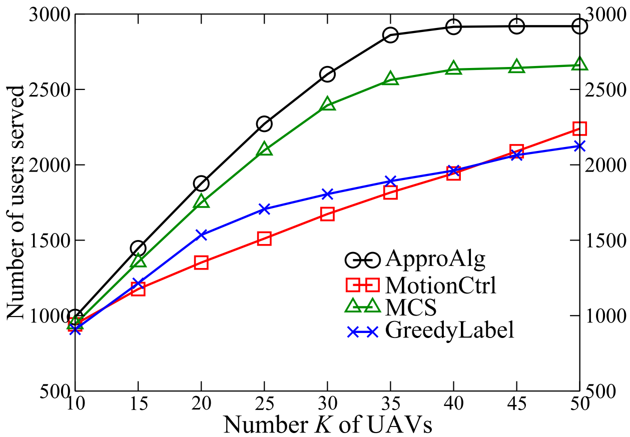


Fig. 6. The performance of different algorithms by increasing the number K of UAVs from 10 to 50, when there are $m = 3,000$ to-be-served users.

12.5% more users are served in the solution delivered by the algorithm. In the following, we do not compare the running times of the four mentioned algorithms, since their algorithm performance curves are similar to the one in Fig. 5(b) and the running time of algorithm `ApproAlg` is no greater than 15 seconds in the following groups of experiments.

We conduct an extra group of ablation experiment. Notice that there are two major differences between the proposed algorithm `ApproAlg` and algorithm `GreedyLabel` in [18]. The first one is that, unlike the algorithm in [18] that assigns profits to nodes in one way only, we assign profits to nodes in multiple ways. We use `GreedyLabel+multiAssign` to represent the algorithm that combines the algorithm in [18] and the multiple ways of node profit assignments in this paper. The second difference is that we propose a new tree decomposition technique. We use `GreedyLabel+newTreeDecomp` to represent the algorithm that combines the algorithm in [18] and the new tree decomposition technique in this paper. Fig. 5(a) shows that the number of users served by algorithm `GreedyLabel+newTreeDecomp` is much larger than that by algorithm `GreedyLabel+multiAssign`, which indicates that the performance improvement of the proposed algorithm `ApproAlg` is mainly contributed by the proposed new tree decomposition technique. On the other hand, the running time of the proposed algorithm `ApproAlg` is mainly prolonged by the multiple ways of node profit assignments, see Fig. 5(b).

We then investigate the performance of different algorithms by increasing the number K of UAVs from 10 to 50, when there are $m = 3,000$ users. Fig. 6 plots that the number of users served by each algorithm increases with more UAVs. In addition, the deployed UAVs by algorithm `ApproAlg` serve 97% ($\approx \frac{2{,}915}{3{,}000}$) of users when there are $K = 40$ UAVs, while the deployed UAVs by the other three algorithms serve no more than 88% ($\approx \frac{2{,}633}{3{,}000}$) of users.

We also study the performance of different algorithms by varying the communication range R between two UAVs from $500\ m$ to $1,000\ m$ while fixing the communication range r' of a user at $500\ m$, when $m = 3,000$, $K = 30$. Recall that $r = \sqrt{r'^2 - H_{uav}^2}$ and $\alpha = \frac{r}{R}$, where $H_{uav} = 300\ m$. It can be seen that the value of α decreases from 0.8 to 0.4 when R grows from 500 m to 1,000 m. Then, following Eq. (1), the value of $h(\alpha)$ decreases from 4 to 2. Fig. 7 illustrates that the number of users served by each of the four algorithms `ApproAlg`, `MotionCtrl`, `MCS`, and `GreedyLabel` increases with the growth of the communication range R between two UAVs. The rationale behind the phenomenon is that less numbers of relaying UAVs are needed when the communication range R is larger, and more UAVs thus can be used to serve the users. Fig. 7 also plots the difference between the numbers of users served by algorithms `ApproAlg`, `MotionCtrl`, `MCS`, and `GreedyLabel`. For example, the number of users served by algorithm `ApproAlg` is about 20% larger than the one by algorithm `MCS` when the communication range R between two UAVs is $500\ m$, while the number by algorithm `ApproAlg` is only about 2.2% larger than the one by algorithm `MCS` when $R = 1,000\ m$.

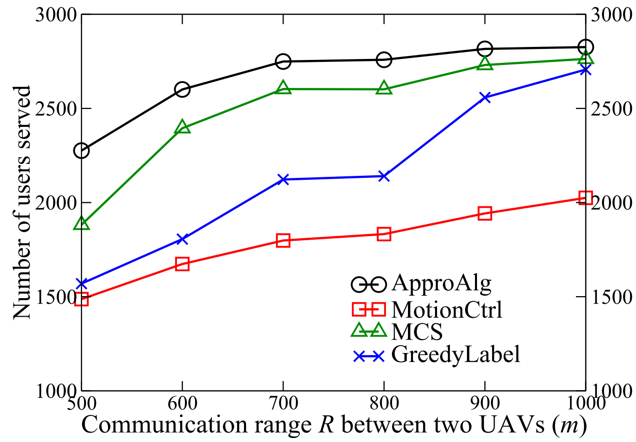


Fig. 7. The performance of different algorithms by varying the communication range R between two UAVs from 500 m to 1,000 m while fixing $r' = 500$ m, when $m = 3,000$ users and $K = 30$ UAVs.

We finally evaluate the performance of different algorithms by varying the communication range r' between a user and a UAV from 400 m to 600 m while fixing the communication range R between two UAVs at 600 m, when $m = 3,000$ users and $K = 30$ UAVs. Recall that $r = \sqrt{r'^2 - H_{uav}^2}$ and $\alpha = \frac{r}{R}$, where $H_{uav} = 300\ m$. It can be seen that the value of α increases from 0.44 to 0.87 when r' grows from 400 m to 600 m. Then, following Eq. (1) in Section II-D, the value of $h(\alpha)$ increases from 2 to 4. Fig. 8 shows that the number of users served by each of the four algorithms `ApproAlg`, `MotionCtrl`, `MCS`, and `GreedyLabel` increases with the growth of the communication range r' between a UAV and a user, since more users will be served by deployed UAVs. In addition, Fig. 8 indicates that the number of users served by algorithm `ApproAlg` is about 8% larger than those by the other three algorithms.

VII. Related Work

The use of UAVs as aerial base stations (BS) recently has gained lots of attentions in public communications. For example, Zhao *et al.* [43] presented a motion control algorithm for deploying a given number K of UAVs to cover as many

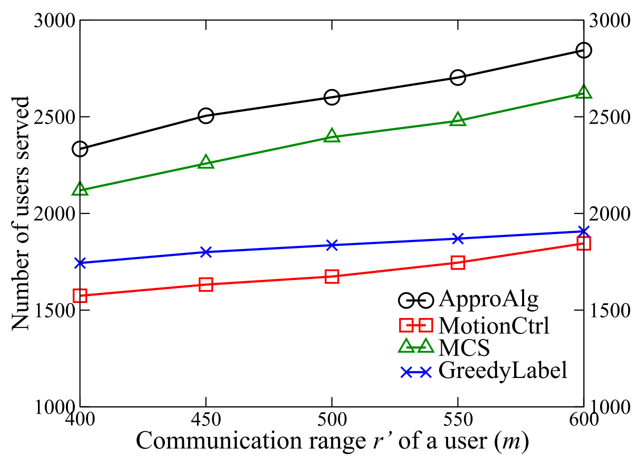


Fig. 8. The performance of different algorithms by varying the communication range r' of a user from 400 m to 600 m while fixing $R = 600$ m, when $m = 3,000$ users and $K = 30$ UAVs.

as users while maintaining the connectivity among UAVs. Liu *et al.* [23] considered the similar problem and proposed a deep reinforcement learning (DRL) based algorithm. Yang *et al.* [38] investigated the problem of scheduling the movement of multiple UAVs to fairly provide communication services to mobile ground users for a given period, by using the DRL method, too. Shi *et al.* [27] studied the problem of planning the flying trajectories of multiple UAVs for a period such that the average UAV-to-user pathloss in the network is minimized, assuming that a user can be served by only a single UAV during the period. They decoupled the problem into multiple subproblems, and solved the subproblems separately.

There are several studies on the special cases of the h-hop independently submodular maximization problem, subject to the connectivity constraint that the induced subgraph of G by a subset of nodes in V is connected. For example, Khuller *et al.* [17], [18] devised a $\frac{1-1/e}{12}$-approximation algorithm for the *budgeted connected dominating set (BCDS) problem*, which is to find a set S of K nodes in a graph G such that the number of nodes dominated by the nodes in S is maximized, subject to the constraint that the induced subgraph $G[S]$ is connected. Notice that the objective function of the BCDS problem is 3-hop independently submodular with $h = 3$. Lamprou *et al.* [21] recently proposed an improved algorithm. They adopted the similar profit assignment procedure and tree decomposition technique as those in [17] and [18], and obtained an improved solution by decomposing a tree with a different size. They improved the approximation ratio from $\frac{1-1/e}{12}(\approx 0.05267)$ to $\frac{1-e^{-\frac{7}{8}}}{11}(\approx 0.05301)$, which is still much smaller than the approximation ratio $\frac{1-1/e}{2h+3}$ $(= \frac{1-1/e}{9} \approx 0.0702)$ of the proposed algorithm in this paper when $h = 3$. Huang *et al.* [15] proposed a $\frac{1-1/e}{8(\lceil 2\sqrt{2}\alpha\rceil+1)^2}$-approximation algorithm for the maximum connected coverage problem with $h = 4$. where $\alpha = \frac{r}{R}$, r and R are the sensing range and communication range of a sensor respectively, and $0 < r \leq R$. It can be seen that $\frac{1-1/e}{128} \leq \frac{1-1/e}{8(\lceil 2\sqrt{2}\alpha\rceil+1)^2} \leq \frac{1-1/e}{32}$, as $0 < \alpha \leq 1$. Yu *et al.* [39], [40] recently improved the approximation ratio to $\frac{1-1/e}{8(\lceil \frac{4}{\sqrt{3}}\alpha\rceil+1)^2}$, where $\frac{1-1/e}{128} \leq \frac{1-1/e}{8(\lceil \frac{4}{\sqrt{3}}\alpha\rceil+1)^2} \leq \frac{1-1/e}{32}$. It can be seen that both approximation ratios in [15], [39], and [40] are no greater than $\frac{1-1/e}{32}$.

There are other investigations on maximizing the values of other submodular functions, not h-hop independently submodular functions, subject to connectivity constraints. For example, Kuo *et al.* [20] considered the problem of deploying K wireless routers in a wireless network such that a submodular function of the deployed K routers is maximized, subject to the constraint that the subnetwork induced by the K routers is connected, for which they proposed a $\frac{1-1/e}{5(\sqrt{K}+1)}$-approximation algorithm, where e is the base of the natural logarithm.

There are flourishing studies on maximizing the value of a submodular function *without* connectivity constraints. For monotone submodular functions, Nemhauser *et al.* [26] considered a problem of choosing K elements from a set such that a submodular function of the chosen K elements is maximized. They devised a $(1-1/e)$-approximation algorithm for the problem and showed that the result is tight. They also extended their result to the submodular function maximization problem under the constraint of the intersection of M matroids, and proposed a $\frac{1}{M+1}$-approximation algorithm [11], and this approximation ratio later is further improved to $\frac{1}{M+\epsilon}$ by Lee *et al.* [22] when $M \geq 2$, where ϵ is a given constant with $0 < \epsilon \leq 1$. Calinescu *et al.* [7] and Filmus *et al.* [10] proposed a randomized $(1-1/e)$-approximation algorithm for maximizing a submodular problem under a matroid constraint, respectively, while Buchbinder devised a deterministic $(1/2+\epsilon)$-approximation algorithm [4], [5]. Sviridenko [29] proposed a $(1-1/e)$-approximation algorithm for maximizing a submodular function subject to a linear constraint, while Kulik *et al.* [19] extended to the solution to multiple linear constraints by giving a $(1-1/e-\epsilon)$-approximation algorithm.

A. Technical Novelties

We are motivated by the study in [18]. There are two essentially technical differences between our work and the work in [18]. The first one is that, unlike the algorithm in [18] that assigns profits to nodes in one way only, we assign profits to nodes in multiple ways. We show that there is a tree T in G with the profit sum of nodes in T being no less than $(1-1/e)OPT$ among one of the multiple profit assignments, and the number of edges in T is no greater than $(K-1)h$, which is less than the number $(Kh-1)$ in [18] when $h \geq 2$.

The second one is that the traditional tree decomposition technique adopted in [18] decomposes a tree T with $(K-1)h$ edges into $4h$ subtrees so that the number of nodes in each subtree is no more than $\frac{K}{2}$. We here propose a novel tree decomposition technique that decomposes a tree T with $(K-1)h$ edges into $2h+3$ subtrees, such that the number of nodes in each subtree is no more than $\frac{K}{2}$, by exploring important structure properties of the tree T. Note that $2h+3 < 4h$ for any integer h if $h \geq 2$. By utilizing the proposed tree decomposition technique, we devise a novel approximation algorithm for the h-hop independently submodular maximization problem, and its approximation ratio is $\frac{1-1/e}{2h+3}$ when $h \geq 2$.

VIII. Conclusion

In this paper, we studied the novel h-hop independently submodular maximization problem, which generalizes many optimization problems arisen in different domains, such as the MCCP problem of deploying a connected UAV network to serve as many users as possible. We then devised a $\frac{1-1/e}{2h+3}$-approximation algorithm for the problem, where e is the base of the natural logarithm. The proposed algorithm has many potential applications, and one direct corollary from this result is a $\frac{1-1/e}{11}$-approximate solution to the MCCP problem when $h = 4$, which significantly improves its currently best $\frac{1-1/e}{32}$-approximate solution [40].

References

[1] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[2] K. Avrachenkov, P. Basu, G. Neglia, B. Ribeiro, and D. Towsley, "Online myopic network covering," 2012, *arXiv:1212.5035*.

[3] I. Bor-Yaliniz and H. Yanikomeroglu, "The new frontier in RAN heterogeneity: Multi-tier drone-cells," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 48–55, Nov. 2016.

[4] N. Buchbinder and M. Feldman, "Deterministic algorithms for submodular maximization problems," *ACM Trans. Algorithms*, vol. 14, no. 3, p. 32, 2018.

[5] N. Buchbinder, M. Feldman, and M. Garg, "Deterministic $(1/2 + \epsilon)$-approximation for submodular maximization over a matroid," in *Proc. 30th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, 2019, pp. 241–254.

[6] N. Buchbinder, M. Feldman, J. Seffi, and R. Schwartz, "A tight linear time (1/2)-approximation for unconstrained submodular maximization," *SIAM J. Comput.*, vol. 44, no. 5, pp. 1384–1402, Jan. 2015.

[7] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, "Maximizing a submodular set function subject to a matroid constraint (extended abstract)," in *Proc. Int. Conf. Integer Program. Combinat. Optim. (IPCO)*, 2007, pp. 182–196.

[8] S. Chandrasekharan *et al.*, "Designing and implementing future aerial communication networks," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 26–34, May 2016.

[9] L. Deng *et al.*, "Approximation algorithms for min-max cycle cover problems with neighborhoods," *IEEE/ACM Trans. Netw.*, vol. 28, no. 4, pp. 1845–1858, Aug. 2020.

[10] Y. Filmus and J. Ward, "A tight combinatorial algorithm for submodular maximization subject to a matroid constraint," in *Proc. IEEE 53rd Annu. Symp. Found. Comput. Sci.*, Oct. 2012, pp. 659–668.

[11] M. L. Fisher, G. L. Nemhausert, and L. A. Wolsey, "An analysis of the approximations for maximizing submodular set functions—II," *Math. Program. Study*, vol. 8, pp. 73–87, Jan. 1978.

[12] N. Garg, "Saving an epsilon: A 2-approximation for the k-MST problem in graphs," in *Proc. 37th Anuu. ACM Symp. Theory Comput. (STOC)*, 2005, pp. 396–402.

[13] J. Gong, T.-H. Chang, C. Shen, and X. Chen, "Flight time minimization of UAV for data collection over wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1942–1954, Sep. 2018.

[14] Q. Guo *et al.*, "Minimizing the longest tour time among a fleet of UAVs for disaster area surveillance," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 2451–2465, Jul. 2022.

[15] L. Huang, J. Li, and Q. Shi, "Approximation algorithms for the connected sensor cover problem," in *Proc. 21st Int. Conf. Comput. Combinatorics (COCOON)*, 2015, pp. 183–196.

[16] D. S. Johnson, M. Minkoff, and S. Phillips, "The prize collecting Steiner tree problem: Theory and practice," in *Proc. 11th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, 2000, pp. 760–769.

[17] S. Khuller, M. Purohit, and K. K. Sarpatwar, "Analyzing the optimal neighborhood: Algorithms for budgeted and partial connected dominating set problems," in *Proc. 25th Annu. ACM-SIAM Symp. Discrete Algorithms*, Jan. 2014, pp. 1702–1713.

[18] S. Khuller, M. Purohit, and K. K. Sarpatwar, "Analyzing the optimal neighborhood: Algorithms for partial and budgeted connected dominating set problems," *SIAM J. Discrete Math.*, vol. 34, no. 1, pp. 251–270, Jan. 2020.

[19] A. Kulik, H. Shachnai, and T. Tamir, "Maximizing submodular set functions subject to multiple linear constraints," in *Proc. 20th Annu. ACM-SIAM Symp. Discrete Algorithms*, Jan. 2009, pp. 545–554.

[20] T.-W. Kuo, K. C.-J. Lin, and M.-J. Tsai, "Maximizing submodular set function with connectivity constraint: Theory and application to networks," *IEEE/ACM Trans. Netw.*, vol. 23, no. 2, pp. 533–546, Apr. 2015.

[21] I. Lamprou, I. Sigalas, and V. Zissimopoulos, "Improved budgeted connected domination and budgeted edge-vertex domination," *Theor. Comput. Sci.*, vol. 858, pp. 1–12, Feb. 2021.

[22] J. Lee, M. Sviridenko, and J. Vondrák, "Submodular maximization over multiple matroids via generalized exchange properties," in *Proc. 12th Int. Workshop Approximation Algorithms Combinat. Optim., 13th Int. Workshop Randomization Approximation Tech. Comput. Sci. (APPROX-RANDOM)*, 2009, pp. 244–257.

[23] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, Sep. 2018.

[24] Y. Liu and W. Liang, "Approximate coverage in wireless sensor networks," in *Proc. 30th Annu. IEEE Conf. Local Comput. Netw. (LCN)*, Nov. 2005, pp. 68–75.

[25] M. Moradi, K. Sundaresan, E. Chai, S. Rangarajan, and Z. M. Mao, "SkyCore: Moving core to the edge for untethered and reliable UAV-based LTE networks," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2018, pp. 35–49.

[26] G. Nemhauser and L. A. Wolsey, "An analysis of the approximations for maximizing submodular set functions—I," *Math. Program.*, vol. 14, no. 1, pp. 265–294, 1978.

[27] W. Shi *et al.*, "Multi-drone 3-D trajectory planning and scheduling in drone-assisted radio access networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8145–8158, Aug. 2019.

[28] C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modelling the scaling properties of human mobility," *Nature Phys.*, vol. 6, no. 10, pp. 818–823, Sep. 2010.

[29] M. Sviridenko, "A note on maximizing a submodular set function subject to a knapsack constraint," *Oper. Res. Lett.*, vol. 32, no. 1, pp. 41–43, 2004.

[30] W. Xu, T. Li, W. Liang, J. X. Yu, N. Yang, and S. Gao, "Identifying structural hole spanners to maximally block information propagation," *Inf. Sci.*, vol. 505, pp. 100–126, Dec. 2019.

[31] W. Xu, W. Liang, and X. Lin, "Approximation algorithms for min-max cycle cover problems," *IEEE Trans. Comput.*, vol. 64, no. 3, pp. 600–613, Mar. 2015.

[32] W. Xu, W. Liang, X. Lin, and J. X. Yu, "Finding top-k influential users in social networks under the structural diversity model," *Inf. Sci.*, vols. 355–356, pp. 110–126, Aug. 2016.

[33] W. Xu *et al.*, "Approximation algorithms for the generalized team orienteering problem and its applications," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 176–189, Feb. 2021.

[34] W. Xu, M. Rezvani, W. Liang, J. X. Yu, and C. Liu, "Efficient algorithms for the identification of top-k structural hole spanners in large social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 5, pp. 1017–1030, May 2017.

[35] W. Xu *et al.*, "Throughput maximization of UAV networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 2, pp. 881–895, Apr. 2022.

[36] W. Xu *et al.*, "Minimizing the deployment cost of UAVs for delay-sensitive data collection in IoT networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 2, pp. 812–825, Apr. 2022.

[37] W. Xu *et al.*, "Approximation algorithms for the team orienteering problem," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, Jul. 2020, pp. 1389–1398.

[38] P. Yang, X. Cao, X. Xi, W. Du, Z. Xiao, and D. Wu, "Three-dimensional continuous movement control of drone cells for energy-efficient communication coverage," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6535–6546, Jul. 2019.

[39] N. Yu, H. Dai, A. X. Liu, and B. Tian, "Placement of connected wireless chargers," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, Apr. 2018, pp. 387–395.

[40] N. Yu, H. Dai, G. Chen, A. X. Liu, B. Tian, and T. He, "Connectivity-constrained placement of wireless chargers," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 909–927, Mar. 2021.

[41] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3747–3760, Jun. 2017.

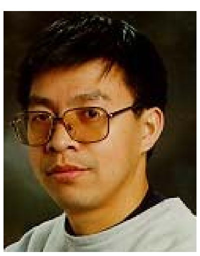
[42] J. Zhang *et al.*, "Minimizing the number of deployed UAVs for delay-bounded data collection of IoT devices," in *Proc. 40th IEEE Int. Conf. Comput. Commun. (INFOCOM)*, May 2021, pp. 1–10.

[43] H. Zhao, H. Wang, W. Wu, and J. Wei, "Deployment algorithms for UAV airborne networks toward on-demand coverage," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2015–2031, Sep. 2018.

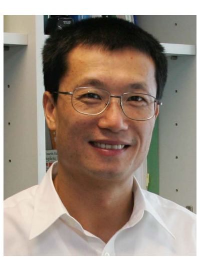
Wenzheng Xu (Member, IEEE) received the B.Sc., M.E., and Ph.D. degrees in computer science from Sun Yat-sen University, Guangzhou, China, in 2008, 2010, and 2015, respectively. He is currently an Associate Professor at Sichuan University. Also, he was a Visitor at The Australian National University and The Chinese University of Hong Kong. His research interests include wireless *ad-hoc* and sensor networks, UAV networks, mobile computing, approximation algorithms, combinatorial optimization, online social networks, and graph theory.

Hongbin Xie received the B.Sc. degree in computational finance from Sichuan University, China, in 2020, where she is currently pursuing the master's degree with the College of Computer Science. Her current research interests include UAV scheduling and networking.

Chenxi Wang received the B.E. degree in the Internet of Things from Sichuan Agricultural University, China, in 2021,where he is currently pursuing the master's degree with the College of Computer Science. His current research interests include UAV scheduling and networking.



Weifa Liang (Senior Member, IEEE) received the B.Sc. degree in computer science from Wuhan University, China, in 1984, the M.E. degree in computer science from the University of Science and Technology of China in 1989, and the Ph.D. degree in computer science from the Australian National University in 1998. He is currently a Professor with the Department of Computer Science, City University of Hong Kong. Prior to the current position, he was a Professor at The Australian National University. His research interests include design and analysis of energy efficient routing protocols for wireless *ad-hoc* and sensor networks, the Internet of Things, edge and cloud computing, network function virtualization, and software-defined networking, design and analysis of parallel and distributed algorithms, approximation algorithms, combinatorial optimization, and graph theory. He serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.



Xiaohua Jia (Fellow, IEEE) received the B.Sc. and M.Eng. degrees from the University of Science and Technology of China in 1984 and 1987, respectively, and the D.Sc. degree in information science from The University of Tokyo in 1991. He is currently a Chair Professor with the Department of Computer Science, City University of Hong Kong. His research interests include cloud computing and distributed systems, computer networks, wireless sensor networks, and mobile wireless networks. He is an Editor of IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS (2006–2009), *Journal of World Wide Web*, *Wireless Networks*, and *Journal of Combinatorial Optimization*. He is the General Chair of ACM MobiHoc 2008, the TPC Co-Chair of IEEE MASS 2009, the Area-Chair of IEEE INFOCOM 2010, the TPC Co-Chair of IEEE GlobeCom 2010, and the Panel Co-Chair of IEEE INFOCOM 2011.

Zichuan Xu (Member, IEEE) received the B.Sc. and M.E. degrees in computer science from the Dalian University of Technology, China, in 2008 and 2011, respectively, and the Ph.D. degree in computer science from The Australian National University in 2016. From 2016 to 2017, he was a Research Associate at the Department of Electronic and Electrical Engineering, University College London, U.K. He is currently a Full Professor and a Ph.D. Advisor with the School of Software, Dalian University of Technology. He is also a 'Xinghai Scholar' with the Dalian University of Technology. His research interests include mobile edge computing, serverless computing, network function virtualization, software-defined networking, algorithmic game theory, and optimization problems.

Pan Zhou (Senior Member, IEEE) received the B.S. degree in the advanced class from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2006, and the Ph.D. degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, in 2011. He is currently an Associate Professor with the School of Electronic Information and Communications, HUST. From 2011 to 2013, he was a Senior Technical Member with Oracle America Inc., Boston, MA, USA. His current research interests include security and privacy, machine learning and big data analytics, and information networks.

Weigang Wu (Member, IEEE) received the B.Sc. and M.Sc. degrees from Xi'an Jiaotong University, China, in 1998 and 2003, respectively, and the Ph.D. degree in computer science from Hong Kong Polytechnic University in 2007. He is currently a Full Professor at the School of Data and Computer Science, Sun Yat-sen University, China. His research interests include distributed systems and wireless networks, especially cloud computing platforms, and *ad-hoc* networks. He has published more than 60 papers in major conferences and journals. He has served as a member of Editorial Board of two international journals, *Frontiers of Computer Science* and *Ad Hoc & Sensor Wireless Networks*. He is also an organizing committee member/a program committee member for many international conferences. He is a member of the ACM.

Xiang Chen (Member, IEEE) received the B.E. and Ph.D. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2002 and 2008, respectively. From July 2008 to July 2012, he was with the Wireless and Mobile Communication Technology Research and Development Center, Research Institute of Information Technology, Tsinghua University. From August 2012 to December 2014, he was with the Tsinghua Space Center, School of Aerospace, Tsinghua University. Since January 2015, he has been serving as an Associate Professor with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China. He is also with the Research Institute, Tsinghua University, Shenzhen, as a Chief Researcher (part-time). His research interests mainly focus on 5G wireless communications, wireless big data, the Internet of Things, and software radio.