# Multi-UAVs Collaboration System based on Machine Learning for Throughput Maximization

Yu Min Park, Minkyung Lee, Choong Seon Hong
Department of Computer Science and Engineering
Kyung Hee University, 446-701
Republic of Korea
{yumin0906, minkyung0110, cshong}@khu.ac.kr

***Abstract*— Due to commercialization of the 5G network, many base stations need to enhance a reliable communication quality. Thus, many studies have still worked to provide mobility and economic benefits to the UAVs-Base Station (UAVs-BS) on behalf of ground base stations. In this paper, we propose a system to find a location where multiple users can have an optimal service throughput by considering users' requirements in Multi-UAVs communication. Based on the Air-To-Ground Path Loss Model, the virtual communication environment is established and Airtime Fairness is applied for equitable channel usage time distribution according to user requirements. Thus, we apply a collaborative algorithm with modified K-means that can distribute users to each UAV and solve communication overload problems. In addition, the Proximal Policy Optimization (PPO) algorithm is applied to set an optimal location with the maximum throughput. As a result, the proposed systems allow the Multi-UAVs to be in the locations with high service throughput for users with different demands.**

***Keywords—5G, UAVs-BS, Throughput, Machine Learning, K-Means Clustering, Reinforcement Learning***

## I. Introduction

The 5G has emerged with a number of mobile devices which have various network requirements. Due to the commercialization of 5G, each cells should be built in close-range to prevent damage from the linearity of the millimeter wave. The 5G requires numerous base stations to fulfil these requirements, however, an install of new base stations on the ground creates economic inefficiency [1]. Thus, a UAVs-Base Station (UAVs-BS) has emerged as an alternative to ground base stations [1].

An unmanned aerial vehicle (UAV) is an aircraft without a human pilot on board. And UAVs are a component of an unmanned aircraft system (UAS) [2]. The first UAVs were started for military use, but now they are used in a variety of fields, including the private and public sectors, other than the military. In addition, UAVs have a lot of potential ability in the field of communication such as cheap costs, an inherent ability for line-of-sight(LoS) communications, and a flexible movement depending on the situation. However, a wireless network on UAVs has a various limitations and problems, including optimal deployment issues, limited energy, security, etc.

The paper[3] presented an algorithm for managing distributed UAV ports to minimize energy consumption and ensure cooperation between UAVs. In the paper [4], for future smart cities, the authors proposed a study on the deployment of UAV swarms for 5G. Also, a way of UAV deployment and efficient swarm was presented by using artificial intelligence algorithm.

In this paper, we propose a way to solve the problem of Multi-UAVs optimal deployment to provide high throughput for users with different requirements in the UAVs-BS environment. There are two main differences in the content of the proposed approach from other papers. First, when UAVs move around to find the best position, we solve the problem by approaching it through reinforcement learning. This has an advantage in adapting to changing environments when comparing to conventional algorithm methods. The second involves an algorithm for how to collaborate with other UAVs when they are in a communication overload situation. For example, when there are crowds of people, such as actual performances or athletic competitions, then efficient communication through UAVs will be possible in the proposed way.

The rest of this paper is organized as follows. Section 2 looks at machine learning and the studies for maximizing throughput in UAV communication, and Section 3 describes the models used to establish the hardening learning environment of this paper and introduces the integrated model of the proposed algorithms. The performance evaluation shows the service throughput of the system model proposed in Section 4, and the last Section 5 presents the conclusion of the thesis and the direction of future research.

## II. Related work

### *A. K-Means Clustering*

$$\operatorname{argmin}_S \sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2 \tag{1}$$

K-means clustering is a part of the partitioning method of clustering. Partitioning method is a method of dividing a given data into multiple partitions. The K-means algorithm determines the sum of the squares at the center of each group and the distance from the data objects in the group as a function of cost. Also, clustering is performed by updating the groups belonging to each data object to minimize the value of this function. equation (1) is the target equation to find a set S that minimizes the sum of squares of distances for each group. Using K-means to distribute users to each UAVs has limitations that users cannot take into account their different service requirements. Therefore, in this paper, when calculating the average value, the weight for

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea goverment(MSIT) (No.2019-0-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing). *Dr. CS Hong is the corresponding author.

the service requirement includes to make the optimum distribution.

### B. Reinforcement Learning

Reinforcement learning takes benefits to get correct parameters in the direction of high reward, repeating the actions established in a given environment. Figure 1 shows a scenario plot of the process in which an agent learns to get rewards for its behavior in a given environment. In situations where AI needs to be learned in an environment that is not already in the past, Reinforcement learning is appropriate even among machine learning because there are no data sets to be learned.

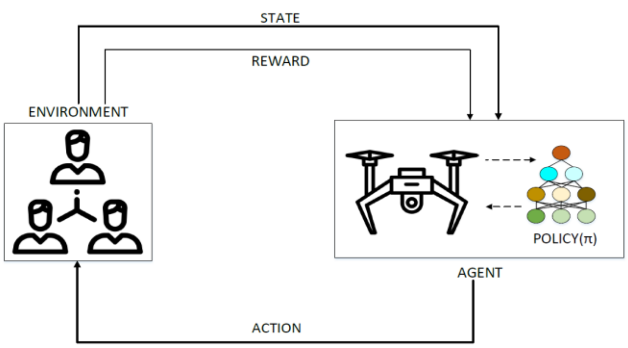


**Figure 1. The perception-action-learning loop**

There are also several methods in Reinforcement Learning and can be broadly classified as Value-based and Policy-based. Unlike Value-based, Policy-based is suitable for learning real robots or drones with continuous space. In Policy-based , Trust Region Policy Optimization (TRPO) and Proximal Policy Optimization (PPO) are the typical methods. PPO is a simpler and more general-purpose method than TRPO, a technique that has evolved TRPO [5]. Therefore, in this paper, the method of finding the location of UAV for maximizing service throughput is studied by using PPO, among reinforcement learning.

### C. Optimization Deployment in UAVs Communications

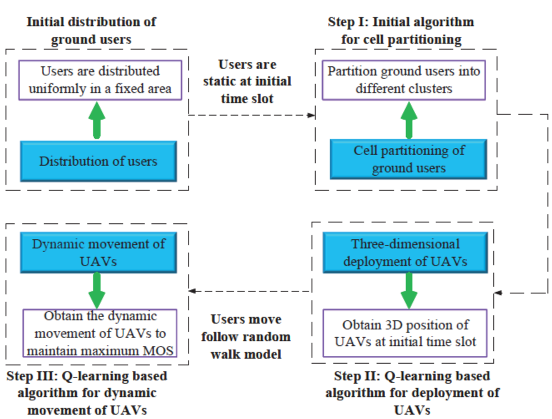


**Figure 2. The procedure for solving the problem of deploying and movement design of multiple UAVs in [7]**

The paper[6] proposes a global optimization algorithm based on Tabu-Search as a way to maximize throughput. When the range of UAVs to move is a constant grid, the optimized position was found in the range that guarantees communication with the user. In this way, the average throughput increases in around 26% compared to the initial batch. However, the UAV movement to the grid has the limitation to make errors in real environments. Therefore, this paper presents a way to find the optimum position, with the range of UAVs moving in a continuous space.

In the paper[7], the authors found a position to maximize Quality of User communication Experience(QoE) in communication with UAVs. As a way to find location, they applied a reinforcement learning based on Q-learning. Also, through K-means clustering, UAVs were placed in each cluster after users were classified. In this paper, the objective is to find a position in UAVs communication that maximizes throughput instead of QoE. Also, it proposes a collaborative algorithm in a communication overload situation not considered in this paper

## III. System Model

This paper apply a machine learning as a way to deploy UAVs so that they can produce maximum throughput, considering the service requirements of each user, when multiple users communicate using Multi-UAVs. First, we use Modified K-means clustering to cluster considering users' location and service requirements, then each UAV uses a reinforcement learning model as a way to find a location to maximize throughput. For the simulation of this study, the network model in Section 3.1, the collaboration algorithm in Section 3.2, and the reinforcement learning model in Section 3.3 are proposed.

TABLE I. DEFINITION AND VALUE OF PARAMETER

| ***Parameter*** | ***Description*** | ***value*** |
|---|---|---|
| f | Transmitter frequency | 1.5Ghz |
| $P_t$ | UAVs-BS transmitting power | 10dBw |
| $G_t$ | Antenna gains for the base station | 8dB |
| $G_r$ | Antenna gains for the user. | 0dB |
| L | Total System Losses | 8dB |
| B | Bandwidth | 10MHz |
| $N$ | Noise power | 107dBw |

### A. Network Model

*1) The Air-To-Ground Path Loss Model:* Model for calculating signal-to-noise ratio (SNR) due to signal attenuation effect at distance($d$) during UAVs communication [6].

$$\mathrm{PL(dB)} = 20\log_{10}\left(\frac{4\pi df}{c}\right) + P(LoS)\eta_{LoS} + P(NLoS)\eta_{NLoS} \tag{2}$$

$$P_r(d) = P_t + G_t - PL + G_r - L \tag{3}$$

$$\mathrm{SNR} = \frac{singnal\ power}{noise\ power} = \frac{P_r(d)}{N} \tag{4}$$

With SNR and equation 5, the data capacity($C_i$) of the user(i) can be obtained.

$$C_i = B\log_2(1 + SNR) \tag{5}$$

*2) Airtime Fairness:* It is a model for equitable channel usage time$(t_i)$ distribution according to users' different service needs [8]. The model provides the correct Throughput Capacity$(C_{t,j})$ according to the data capacity$(C_j)$ of the user$(j)$.

$$\sum_{j \in U} t_j \leq 1 \tag{6}$$

$$C_{T,j} = t_j C_j \quad for \ \forall j \in S \tag{7}$$

The goal of this paper is to increase the final throughput by adding throughput capacity to each user. As a result, the problem that needs to be solved in this paper is to find the location of the UAV that satisfies equation 8.

$$max \sum_{j \in S} C_{T,i} \tag{8}$$

### B. Collaboration Algorithm

*1) Modified K-Means Clustering based on User Demand:* The normal K-means are clustered considering only the location $(u_i)$ of the users. However, because the user's requirements are not taken into account, they are not clustered to have the maximum throughput. Thus, as shown in Equation 10, the user's demand$(d_i)$ in UAV was weighted and constant $\alpha$ multiplied by a constant to modify the mean expression.

$$\mathrm{D} = \sum_{i \in M_j} d_i \tag{9}$$

$$(x_j^*, z_j^*) = \frac{\sum_{i \in M_j}(x_i, z_i) * \alpha * d_i}{D} \tag{10}$$

*2) Collaborate with fellow UAVs in the event of overload:* Clustered results did not consider to the extent that UAVs can handle. In other words, either UAVs can cause communication overload. Therefore, steps should be taken to get out of it. In this paper, we select the nearest UAV to escape from the network overload. Then, we deploy the UAV $(j)$ in a position$(k_j)$ where the throughput can be maximized through the reinforcement learning model. Part 2 of Algorithm 1 describes the algorithm in detail.

### C. Reinforcement Learning Model

Simulation of reinforcement learning was done with Unity and learned through the ML-Agents components and Tensorflow. When the episode begins, 10 users and one UAV are randomly placed within a radius of 250 meters. At this time, users have different service requirements and calculate the overall throughput to users according to the expressions of the network models used. Action is to move the UAV to any points and then calculates the throughput and compares it with the previous throughput. If the throughput has increased, it gives a reward an increased return, otherwise the episode and update the parameters in the learning model will end. Then, it repeat the episode to learn. Figure 3 shows the overall structure of the learning scenario.

**Algorithm 1:** Proposed Collaboration Algorithm

**Input:** $u_i, d_i, k_j, \forall i \in U, \forall j \in D$
**Output:** $M_j^*, k_j, \forall j \in D$

**Part 1:** Weighted K-means Clustering based on User Demand
**Input:** $u_i, d_i, k_j, \forall i \in U, \forall j \in D$
**Output:** $M_j, k_j, \forall j \in D$
**do**
  $M_j^p \leftarrow M_j \ for \ each \ j \in D$
  **for** i ∈ U
    Choose j ∈ D s.t. minDistance(i, j)
    Add i to $M_j$
  **for** j ∈ D
    $k_j \leftarrow (\sum_{i \in M_j} u_i * \alpha * d_i)/(\sum_{i \in M_j} d_i)$
**while** $M_j \neq M_j^p \ for \ each \ j \in D$

**Part 2:** Cooperation Algorithm in the overload
**Input:** $M_j, u_i, d_i, k_j, \forall i \in U, \forall j \in D$
**Output:** $M_j^*, k_j, \forall j \in D$
**for** j ∈ D
  **if** $C_j's \ number > 10$
    **for** k ∈ D (k ≠ j)
      **if** $M_j's \ number < 10$
        Choose i ∈ $M_j$ s.t. minDistance(i, k)
        Delete i to $M_j$ $And \ Add \ i \ to \ M_k$
        **break**
    **continue**
$M_j^* \leftarrow M_j \ for \ each \ j \in D$

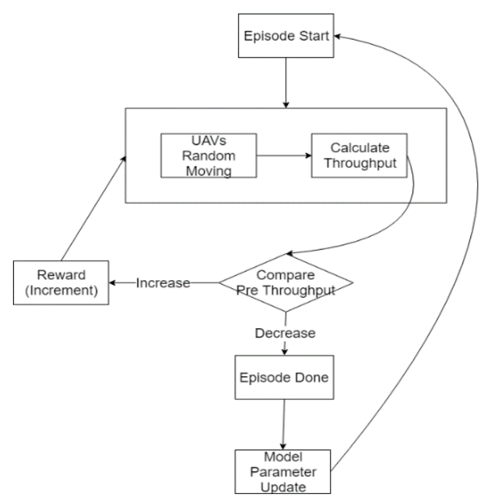


**Figure 3. Learning Scenario Sequence Diagram**

## IV. Simulation Result

A simulation of this paper processes with 30 users and 3 UAVs. Users have different needs and are randomly located in [250m, 250m] zones. The initial location of the UAVs is arbitrarily located due to the nature of the K-means algorithm, and when performing the collaboration algorithm, the resting UAVs are located at the origin [0m, 0m, 0m]. The situation begins with the central ground base station assigning users to UAVs to facilitate communication. Also, the UAV is assumed to be in the optimum position for the height and is fixed at

200M. At the end of the clustering, the reinforcement learning model in the UAV ensures that users continue to find the best position for changing locations and demands.

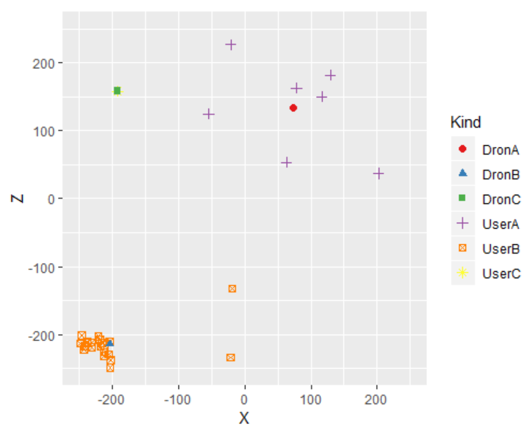


**Figure 4. Deploying by Normal K-means Clustering**

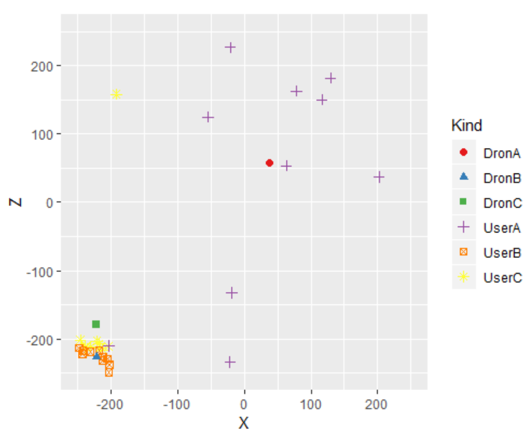


**Figure 5. Deploying by Proposed Algorithm**

The throughput of the proposed algorithm is compared with the normal K-means clustering for performance analysis. Figure 4 and 5 show deployments through the normal K-means clustering and through proposed algorithms, respectively, when users are concentrated in certain areas. Figure 4 shows that Drone B is in an overload situation and that Table 2 shows that it is processing approximately 111 Mbps. On the other hand, the proposed algorithms allows Drone A and Drone C to reduce the overload of Drone B, increasing in the throughput, around 153 Mbps.

TABLE II. THROUGHPUT TABLE BY METHOD

| | ***Proposed Algorithm*** | ***Normal K-means Algorithm*** |
|---|---|---|
| *Total Throughput (Mbps)* | 153.618 | 111.779 |

Figure 6 shows a graph of the throughput of the proposed algorithm and the normal K-means clustering in a random case. When users are distributed evenly, such as case C, case D, case F, case G, the proposed algorithm does not perform very well. However, under overload conditions such as case H, case I, case J, it can be seen that the normal K-means differs by 50 Mbps in throughput.

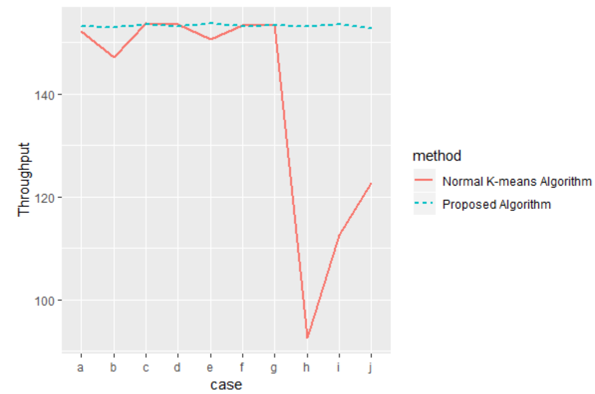


**Figure 6. Throughput graph by Methods**

## V. CONCLUSION AND FUTURE WORK

In this paper, we used techniques called PPO with reinforcement learning and modified K-means clustering to deploy Multi-UAVs in the service throughput maximization as considering a location with different service requirements. The proposed method is able to increase the throughput by about 1.5 times in an overload situation compared to the normal method. Also, this paper shows the potential ability to deliver reliable 5G services with UAVs-BS systems. In addition, UAVs communication to overloaded areas shows to provide the stable service environment. In the future, limitations such as energy efficiency and security did not consider in this paper. Therefore, we will do it to address energy efficiency problem in a future work.

## REFERENCES

[1] Li, Bin, Zesong Fei, and Yan Zhang. "UAV communications for 5G and beyond: Recent advances and future trends." *IEEE Internet of Things Journal* (2018).

[2] "ICAO's circular 328 AN/190 : Unmanned Aircraft Systems". "*https://www.icao.int/Meetings/UAS/Documents/Circular%20328_en.pdf*" ICAO. Retrieved 22 May 2019.

[3] Lynskey, Jared, Kyi Thar, Thant Zin Oo, and Choong Seon Hong. "Facility Location Problem Approach for Distributed Drones." *Symmetry* 11, no. 1 (2019): 118.

[4] Qi, Fei, Xuetian Zhu, Ge Mang, Michel Kadoch, and Wei Li. "UAV Network and IoT in the Sky for Future Smart Cities." *IEEE Network* 33, no. 2 (2019): 96-101.

[5] Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal policy optimization algorithms." *arXiv preprint arXiv:1707.06347* (2017).

[6] ur Rahman, Shams, Geon-Hwan Kim, You-Ze Cho, and Ajmal Khan. "Positioning of UAVs for throughput maximization in software-defined disaster area UAV communication networks." *Journal of Communications and Networks* 20, no. 5 (2018): 452-463.

[7] Liu, Xiao, Yuanwei Liu, and Yue Chen. "Reinforcement Learning in Multiple-UAV Networks: Deployment and Movement Design." *arXiv preprint arXiv:1904.05242* (2019).

[8] Danna, Emilie, Subhasree Mandal, and Arjun Singh. "A practical algorithm for balancing the max-min fairness and throughput objectives in traffic engineering." In *2012 Proceedings IEEE INFOCOM*, pp. 846-854. IEEE, 2012