

Article

A Reinforcement Learning Routing Protocol for UAV Aided Public Safety Networks

Hassan Ishtiaq Minhas¹, Rizwan Ahmad^{1,*} , Waqas Ahmed², Maham Waheed¹, Muhammad Mahtab Alam³  and Sufi Tabassum Gul²

¹ School of Electrical Engineering and Computer Science, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan; hminhas.msee17seecs@seecs.edu.pk (H.I.M.); maham.waheed@seecs.edu.pk (M.W.)

² Department of Electrical Engineering, Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad 45650, Pakistan; waqas@pieas.edu.pk (W.A.); tabassumgul@gmail.com (S.T.G.)

³ Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, 19086 Tallinn, Estonia; muhammad.alam@taltech.ee

* Correspondence: rizwan.ahmad@seecs.edu.pk

Abstract: In Public Safety Networks (PSNs), the conservation of on-scene device energy is critical to ensure long term connectivity to first responders. Due to the limited transmit power, this connectivity can be ensured by enabling continuous cooperation among on-scene devices through multipath routing. In this paper, we present a Reinforcement Learning (RL) and Unmanned Aerial Vehicle (UAV) aided multipath routing scheme for PSNs. The aim is to increase network lifetime by improving the Energy Efficiency (EE) of the PSN. First, network configurations are generated by using different clustering schemes. The RL is then applied to configure the routing topology that considers both the immediate energy cost and the total distance cost of the transmission path. The performance of these schemes are analyzed in terms of throughput, energy consumption, number of dead nodes, delay, packet delivery ratio, number of cluster head changes, number of control packets, and EE. The results showed an improvement of approximately 42% in EE of the clustering scheme when compared with non-clustering schemes. Furthermore, the impact of UAV trajectory and the number of UAVs are jointly analyzed by considering various trajectory scenarios around the disaster area. The EE can be further improved by 27% using Two UAVs on Opposite Axis of the building and moving in the Opposite directions (TUOAO) when compared to a single UAV scheme. The result showed that although the number of control packets in both the single and two UAV scenarios are comparable, the total number of CH changes are significantly different.

Keywords: clustering; d2d communication; disasters; energy conservation; network lifetime; public safety networks; reinforcement learning



Citation: Minhas, H.I.; Ahmad, R.; Ahmed, W.; Waheed, M.; Alam, M.M.; Gul, S.T. A Reinforcement Learning Routing Protocol for UAV Aided Public Safety Networks. *Sensors* **2021**, *21*, 4121. <https://doi.org/10.3390/s21124121>

Academic Editor: Sergio Toral Marín

Received: 31 December 2020

Accepted: 29 March 2021

Published: 15 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Man-made disasters such as terrorism can result in both the loss of life and critical infrastructure. It is estimated that an underdeveloped country like Pakistan incurred direct losses of 127 billion dollars in the last 17 years or so due to terrorism [1]. In addition, the attacks like Army Public School (APS) Peshawar, in which 150 students lost their lives, left a huge social and psychological impact on society [2]. This event has resulted in the complete overhaul of the security infrastructure and caused indirect losses to the economy. Numerous other terrorist incidents such as on the Pakistan Navy Ship (PNS) Mehran, and General Headquarter (GHQ) are examples in which terrorists attacked a building, took hostages, and/or destroyed critical infrastructure. In these scenarios, to disrupt coordination, communication/cellular infrastructure is often taken out either by the authorities or terrorists. The on-scene devices carried by the trapped victims are unable to communicate to first responders or law enforcement agencies. The information (location and number of victims, audio, video, and images)

captured by on-scene available devices can provide timely information to first responders for carrying out a coordinated rescue operation [3,4].

In these situations, a Device to Device (D2D) network provides an alternative method of communication and connectivity among devices [5,6]. The D2D network along with the presence of an Unmanned Aerial Vehicle (UAV) can ensure the information collected by the devices to reach the Command Center (CC). This situation is depicted in Figure 1. Since the transmission power of the devices are limited, it may be possible that some of the devices are unable to reach the UAV. It therefore becomes imperative to cluster the devices. Clustering is a process in which the network is divided into small substructures called clusters, based on node degree, mobility, weights, etc., [7]. This small substructure consists of Cluster Members (CMs) and a Cluster Head (CH). In the network, the CMs communicate with their respective CHs to forward their data to the UAV. The CHs can also rely on each other in forwarding data to the UAV and subsequently to the CC.

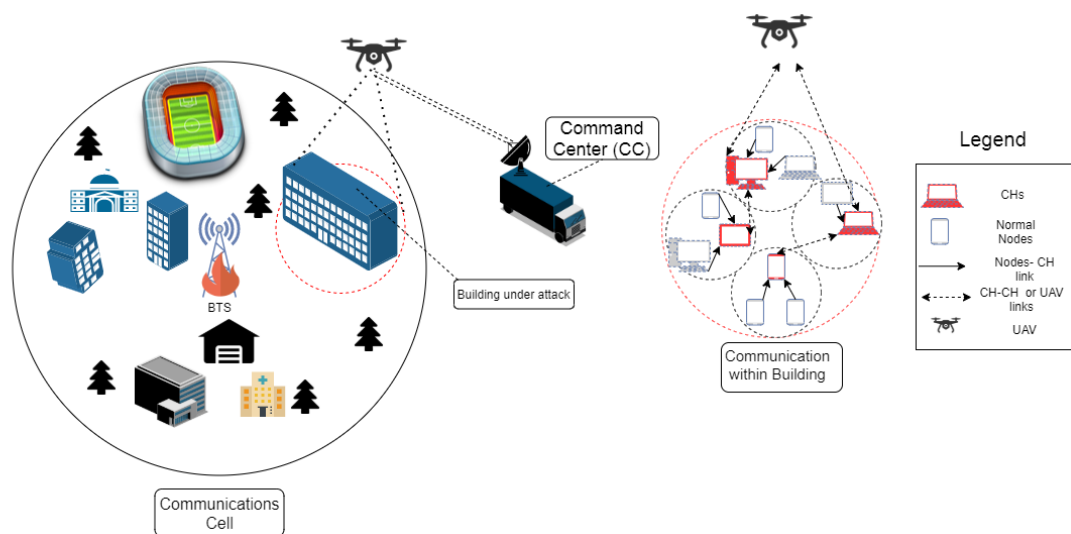


Figure 1. Disaster-affected region.

Multiple clustering techniques are studied in the literature based on different applications [8–20]. For example, the Link Cluster Algorithm (LCA) introduces the concept of a Gateway (GW) node to provide better connectivity among neighboring CHs [8]. However, this scheme is unstable due to numerous ID exchanges between nodes. The Least Cluster Change (LCC) [9] reduces the cost of re-clustering, which gives stability to the clusters. In [10], the authors proposed a new underlay clustering-based D2D network for the partial and disconnected network. It forms dynamic clusters using adhoc base stations or mobile devices. In [11], low frame-sized beacon signals are broadcasted by CHs to decrease signaling overhead. CMs are declared based on the SINR values. The α -Stability Structure Clustering algorithm (α -SSCA) is proposed in [12] in which CH is selected based on a score function calculated by the exchange of hello messages between neighboring nodes. The concept of quasi clusters, a special cluster within a cluster, are introduced in [13] to help in reducing transmission power resulting in a longer network lifetime.

In [15], the authors proposed a new clustering scheme in which first clusters are formed by dividing the area into multiple small partitions and then selecting the CH in those partitions. These will be selected based on energy. To prolong network lifetime and to avoid a blind spot issue in the scheme, re-clustering is performed when the energy of the existing CH reaches a certain threshold. This scheme suffers from scalability issues. In our earlier work [16], we compared different basic clustering schemes such as Clustering without GW (CG), Clustering with GW (CWG), and No Clustering (NC) in terms of throughput and energy. Simulation results show that the performance of CWG is best in terms of throughput and energy.

Another important clustering technique is called the K mean clustering, which is mostly used in Wireless Sensor Networks (WSNs). In [17], K means and LEACH-C are combined to prolong the lifetime of the network. At the start, the K mean algorithm is used to form clusters, then LEACH-C is applied to each cluster. The method reduces the overhead and increases the packet success rate. In [18], the authors have used the K means L layer algorithm, which results in a decreased number of clusters enhancing network lifetime. In [19], an energy-efficient K mean clustering protocol is proposed to optimize packet size based on the channel conditions. This approach reduces energy consumption and increases the overall network lifetime. In [20], an optimum value of K is obtained using the elbow method and afterwards, clustering will be applied using the K mean algorithm. Simulations show that after running the elbow method on different points, there comes a point where gain gained by increasing K will drop. This is the optimum value of K.

In the absence of a cellular network, relaying critical data from the devices to the CC is another important concern. Several routing protocols/algorithms exist in the literature for multi-hop/multi-user networks. An emergency routing technique based on body-to-body networks, known as an Optimized Routing Approach for Critical and Emergency Networks (ORACE-NET) is proposed in [21]. It establishes temporary network connectivity for relief works in disaster-affected areas. The results show that ORACE-NET performs better in terms of energy consumption and throughput. Considering network size, different routing schemes perform differently. Interference Aware Routing (IAR), Shortest Path Routing (SPR), and Broadcast Routing (BR) are tested in [22] to find the shortest emergency route in a disaster scenario. The simulation results show that the BR has the highest packet success ratio for small networks, while IAR performs better for large networks. In [23], the authors applied Simultaneous Wireless Information and Power Transfer (SWIPT) on CHs to achieve better performance in a disaster-affected region.

Once the devices are clustered, the UAV position and trajectory plays a crucial role in determining the Energy Efficiency (EE) of the network. Mainly, in the disaster situation the UAV acts as a relay node for on-scene devices [24–33]. For UAV deployment in PSN, the authors in [27] discovered the optimal altitude for a UAV that maximizes coverage. In [29], authors proposed a UAV-assisted solution to establish energy-efficient connectivity in a disaster-affected region in the presence of Critical Nodes (CNs). In [30], authors proposed a D2D-based solution using UAVs that can reduce the response time significantly. In [31], authors used Reinforcement Learning (RL) technique to deploy UAVs in a disaster scenario to maximize total user coverage. In [32], the intelligent placement of UAVs as temporary aerial base stations is discussed for public safety communications. In [34], authors proposed a UAV-assisted vehicular communication framework using Software Defined Networking (SDN) to reduce the processing cost of vehicles. UAV will act as a flying relay and helps in forwarding data to a Mobile Edge Computing (MEC) server. This algorithm reduces the average system cost by half. In [35], the authors proposed a new cellular network for UAVs to support a high data rate. Three transmission modes of UAV with a network, UAVs, and devices i.e., U2N, U2U, and U2D are studied. Authors in [36] proposed a multi agent deep RL-based UAV framework assisted by MEC in which UAVs will assist users on the ground. Results showed considerable gain in terms of fairness and energy consumption. In this paper [37], the authors used UAVs in a disaster environment to extract information from its one-hop devices using wireless power transfer technique. The graph traversal method is used to reduce the energy cost of the UAV to one third of the total energy.

In [38], authors reviewed and discussed different UAV-Aided Wireless Sensor Networks (UAWSNs). The advantage of these networks is increased coverage and maximum energy consumption at the cost of variable paths and mobility issues resulting in coverage problems in these networks. In [39–46], different UAV WSN-structured routing (flat, cluster-based, tree-based, and location-based) protocols are proposed. Authors in [47] used UAV communication to provide rescue operations in disaster-affected areas. UAVs are spread over the entire area to provide network coverage. Gateway UAVs are further used to deliver the information to the main network. In this work, the aim is to maximize the data

rate while considering battery consumption. To address the problem of gateway UAV selection, authors in [48] proposed a gateway UAV selection algorithm named Battery-Aware Multi Arm Bandit (BA-MAB). They have also explored the use of machine learning. Two kinds of UAVs are present in this work: Access UAVs and gateway UAVs. The objective is to maximize the data rate while consuming minimum energy.

In [49], authors used drones for surveillance and data collection in buildings. Sensors are used by the drones to navigate the buildings to identify and pinpoint the problems. Deep RL is applied along with curriculum learning and neural networks.

Existing work on UAV and clustering mainly focuses on optimizing data collection in WSNs. However, as demonstrated from the above literature, PSN is another important use of UAVs and clustering simultaneously. The energy dynamics of on-scene devices in a disaster scenario is highly dynamic compared to other uses, which amplifies the complexity of ensuring end-to-end connectivity. Therefore, in this paper:

- We first analyze the impact of different clustering schemes and a UAV presence on the performance of multihop routing in a disaster scenario. We then present a RL approach to ensure end-to-end connectivity and improve Energy Efficiency (EE) of PSNs.
- We consider the mobility of UAVs around the disaster area. Multiple UAV trajectories are devised in order to improve the coverage of clusters in the disaster area while ensuring EE.

This paper is organized as follows: In Section 2, the detailed system model is presented. Section 2.1 presents network throughput and delay, Section 2.2 presents the energy model, and the problem formulation is presented in Section 2.3. In Section 3, routing methods are discussed in detail with clustering, route discovery, routing, and control overhead in Sections 3.1–3.4 respectively. Section 4 provides a comparison of clustering schemes with respect to energy and throughput. Sections 4.1 and 4.2 discuss reinforcement-based routing and combined RL and UAVs trajectory optimization. In the end, Section 5 gives the conclusion and future directions.

2. System Model

We consider a man-made disaster scenario in which terrorists attack a large building. It is assumed that the normal cellular/wireless infrastructure is either blocked by security forces or destroyed by the terrorists. We further assume the presence of some on-scene devices (from here on called nodes) held by the trapped victims. If provided with an adequate emergency communication network, the information carried by these nodes can yield great insight for the law enforcement agencies. However, it may not be possible for the emergency communication network to provide coverage to all the nodes simultaneously. In this situation, the nodes can cooperatively communicate with each other to form a D2D multi-hop network. To simplify communication between multiple nodes, the devices can form clusters assisted through the D2D network. The UAV and CC are deployed outside the building perimeter to collect the information from these clusters. The clusters within range of CC communicate directly with the CC and the clusters outside the range of the CC communicate with the CC through a UAV. This situation is depicted in Figure 2, in which the UAV (acting as a relay) is placed outside the building to connect the nodes with the CC used by law enforcement agencies. To protect the CC from an ambush by the terrorists, the CC is deployed slightly away from the building. The symbols are mentioned in Table 1.

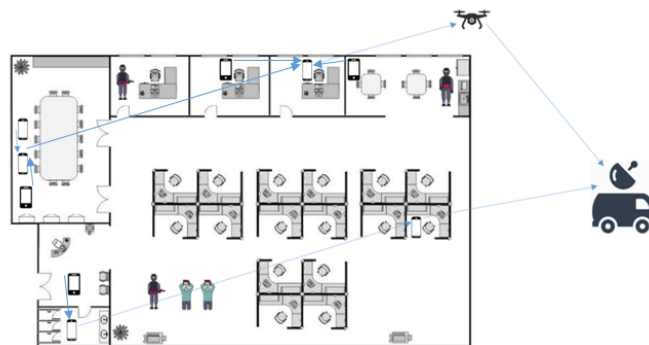
Table 1. Table of symbols.

Name	Symbol	Name	Symbol
Number of nodes	N	Number of clusters	i
Cluster radius	R_{max}	CH Tx radius	T_{max}
K value	K	Average distance of each K	W_k
Distance	d	Transmissions	T
Reward function	Rw	Learning rate	V
Discount factor	γ	Path loss	PL
Carrier frequency	f_c	Number of walls	n_w
Number of floors	n_f	Floor loss	F_l
Throughput	r	Path	ρ
Routing matrix	H	Packet size	L
Energy Efficiency	EE	Delay	δ
Transmitting power	P_t	Bandwidth	B
Noise power	N_o	Packet size	L
Beta	β	Alpha	α

Figure 3 shows a two-tier network with N nodes that are randomly distributed in $100\text{ m} \times 100\text{ m}$. The nodes are divided into i clusters and the set of nodes in each cluster is denoted by N_i , where $i = \{1, \dots, I\}$. Considering a limitation on the transmit power of on-scene devices, the maximum distance between CH and its CM is restricted to R_{max} . Whereas, a CH in this network can communicate with the other CHs over a maximum distance denoted by T_{max} . The pathloss [50] between links separated by distance d is calculated from:

$$PL(d) = 20 \times \log_{10}(d) + 46.4 + 20 \times \log_{10}(f_c/5) + 12 \times n_w + 17 + 4(n_f - 1) \quad (1)$$

where the path loss exponent is assumed to be 2, f_c is the carrier frequency, n_w is the number of walls (taken as 1), and n_f is the number of floors (taken as 0) as we assume that the devices are on the same floor.

**Figure 2.** Disaster hit area with Unmanned Aerial Vehicle (UAV) and Command Center (CC).

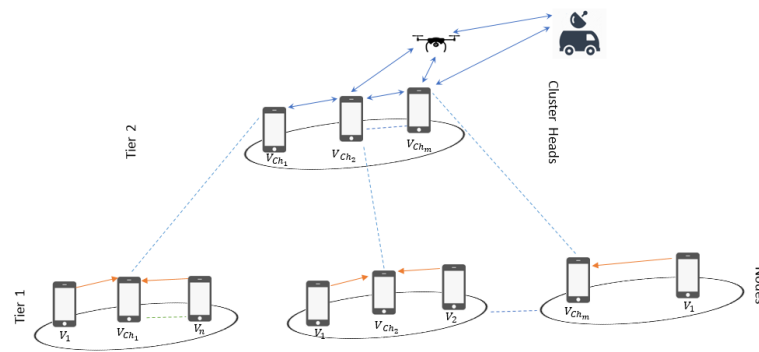


Figure 3. Two-tier graph.

2.1. Network Throughput and Delay

In this scenario, the total network throughput (r) is the sum of throughput from each CM to the CC. Given d_{mn} represents the link distance between m and n , then the throughput is defined as:

$$r_{mn} = B \log_2 \left(1 + \frac{P_t}{N_o P L(d_{mn})} \right). \quad (2)$$

The CHs in the network can cooperate with each other to forward their data to the UAV or CC. Since a CM in cluster i ($s_i \in N_i$) takes multiple hops to reach CC, we define a routing matrix for each CM, denoted by H_{i,s_i} . It is assumed that the routing matrix remains the same for some time horizon and also remains the same for all the nodes in a cluster. Therefore, the routing matrix is defined as $H_{i,s_i} = (h_{mn}) \in \mathbb{R}^{(K+1) \times (K+2)}$, where h_{mn} denotes the status of the connection between m and n , $m \in \{s_i, CH_1, \dots, CH_I, UAV, CC\}$ and $n \in \{CH_1, \dots, CH_I, UAV, CC\}$, and $h_{mn} = 1$ ($h_{mn} = 0$) indicates the presence (absence) of path between the nodes m and n . For example, if the 2nd node in the 1st cluster forward their data to CC through CH3 and UAV, then the routing matrix for such a configuration can be written as:

$$H_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

The first entry $h_{11} = 1$ indicates the link between the 2nd node and 1st CH, $h_{23} = 1$ indicates the link condition between CH1 and CH3, $h_{34} = 1$ connects CH3 to the UAV, and $h_{45} = 1$ successfully terminates the information at CC.

Let $\hat{\zeta}_{i,s_i}$ defines the throughput for all the possible paths from the node in s_i , then the throughput on the paths defined by the routing matrix H_{i,s_i} can be written as $\hat{\zeta}_{i,s_i} = (r_{mn} h_{mn}) \in \mathbb{R}^{(K+1) \times (K+2)}$. Since a node takes multiple hops to reach CC, we assume that the minimum throughput in all the multi-hop links is the throughput of the node. Let λ_{i,s_i} be the minimum value of $\hat{\zeta}_{i,s_i}$, then the sum throughput of all the nodes can be written as:

$$\zeta = \sum_{i \in I} \sum_{s_i} \lambda_{i,s_i}. \quad (4)$$

The number of hops can be directly computed from the rank of H_{i,s_i} denoted by ρ_{i,s_i} . Using λ_{i,s_i} , packet size L , and transmission rate, the end-to-end delay (ignoring propagation delay) of a node indexed by s_i in cluster i can be calculated as:

$$\delta_{i,s_i} = \sum_m \sum_n \frac{L h_{mn}}{r_{mn}}, \quad (5)$$

whereas, the sum end-to-end delay of all nodes in the network are given as:

$$\Delta = \sum_{i \in I} \sum_{s_i} \delta_{i,s_i}. \quad (6)$$

2.2. Energy Model

Assuming that all the nodes transmit with constant power P_t , the transmission energy $E_{tx}(i, s_i)$ of CM (indexed by s_i of cluster i) can be written as:

$$E_{tx}(i, s_i) = \frac{LP_t}{r_{s_i i}}. \quad (7)$$

Assuming, a CM has an initial energy $E_{initial}(i, s_i)$, then the residual energy of the node after T transmissions becomes:

$$E_{res}(i, s_i) = E_{initial}(i, s_i) - \sum_T E_{tx}(i, s_i). \quad (8)$$

Since a CH forwards data of all the CMs, the transmission energy of i th CH becomes:

$$E_{res}(i) = E_{initial}(i) - \sum_T \sum_{s_i} E_{tx}(i), \quad (9)$$

where, $E_{initial}(i)$ is the initial energy of CH, and

$$E_{tx}(i) = \sum_n \frac{LP_t h_{in}}{r_{in}}. \quad (10)$$

The total energy consumption in the networks after T transmissions can now be written as:

$$E_{tot} = \sum_{i \in I} \sum_{s_i} E_{tx}(i, s_i) + \sum_{i \in I} E_{tx}(i). \quad (11)$$

2.3. Problem Formulation

The objective in this work is to maximize throughput, however, maximizing throughput can lead to higher energy consumption. Since the devices have limited energy, the higher energy consumption can lead to dead nodes in the network and subsequently network singularities. Therefore in this paper, we aim to increase EE. Based on the sum throughput and energy calculations (4) and (11) in the previous subsections, we can define the EE as:

$$EE = \frac{\zeta}{E_{tot}} = \frac{\sum_{i \in I} \sum_{s_i} \lambda_{i,s_i}}{\sum_{i \in I} \sum_{s_i} E_{tx}(i, s_i) + \sum_{i \in I} E_{tx}(i)}. \quad (12)$$

The objective function can now be written as:

$$\begin{aligned} \max_{H_{i,s_i}} \quad & EE \\ \text{s.t.} \quad & E_{res}(i) > 0 \\ & R_{max} \leq 30m. \end{aligned} \quad (13)$$

In the following we present an intelligent routing technique based on RL to maximize the EE.

3. Reinforcement Learning-Based Routing

To achieve the above objective function, we define a routing methodology which involves three steps. In the first step, the devices form clusters to decrease the transmission energy. In the second step, route discovery takes place and end-to-end paths are found. Note that routing discovery is dependent on the type of underlying clustering schemes as different clustering schemes generate different network configurations or end-to-end paths. In the last step, based on routing discovery we apply RL to determine transmission paths to improve the EE of the network.

As discussed earlier, clustering can impact route discovery, therefore, to optimize the route discovery we compare the performance of different clustering schemes.

3.1. Clustering

In this paper, we consider two types of clustering schemes: (1) Clustering-energy which is a distributed clustering scheme and (2) clustering-K mean which is a centralized clustering scheme.

3.1.1. Clustering-Energy

In this scheme, we assume that the nodes have no prior information about the energy and distance of other nodes. Initially, a node will check its E_{res} , and based on its level decides about its current status as either CH or CM. If the node energy is 100 times above the energy threshold, then it will immediately declare itself as a CH. Otherwise, it will wait for a small interval equivalent to a single transmission round (for other nodes to declare themselves as CH) and then declares itself as a CH. If the node declares itself as a CH then it will broadcast its CH Identification (ID) and E_{res} information. This CH ID will be selected randomly between 0 and I_{max} . This broadcast is limited by the $Tx_{range} = 30$ m. The nodes which receive this broadcast will compare their own E_{res} with the received information. If the E_{res} of the node is higher compared to that of the current CH, it will declare itself as a CH and broadcast its own E_{res} and CH ID. On receiving this broadcast, the node that has previously declared itself as a CH will change its status to CM and send an association request to the new CH. All the other nodes except CH which receive this association request will change their CH. On the other hand, if the E_{res} of the receiving node is lower than the CH, then they will send an association request to the CH and change their status as member nodes.

Afterwards, the CH will broadcast its cluster ID at $Range = 1.5 \times Tx_{range}$. All the CHs that receive this broadcast will forward it until all the CHs receive this ID. If some CH have selected the same ID before, then this will resolve the issue because this broadcast and each CH will now have a unique CH-ID. The process of re-clustering will start when the E_{res} of the CH reaches the energy threshold value. At that point the CH will broadcast a CH dead message. Nodes that receive this message will start the CH selection process.

We observe the nodes falling in the overlapping zones of two clusters can act as a GW node. This will provide an added degree of freedom in route discovery. Therefore, based on the above procedure, we derive another scheme called clustering-energy-GW. The only difference is that the GW nodes are formed if a member node receives the broadcast of two or more CHs. The node will set and broadcast its status as a GW node. These schemes are easy to implement and need no central authority, however, this is achieved at a higher cost of cluster formation and CH selection.

3.1.2. Clustering-Kmean

K mean clustering is an unsupervised machine learning algorithm to cluster nodes in the network. From a network perspective, it is a centralized scheme and requires distance information of all the nodes. This algorithm comprises of two key steps. In the first step, K centres are placed randomly in the given geographic area and all nodes must associate themselves with the closest centre. In the next step, the mean of each centre with the nodes is calculated and these new means then become new centres. These same steps are

repeated until the criterion function (C_r) becomes minimum. We can find this function by the formula:

$$C_r = \sum_{i=1}^I \sum_{d_i} |x_{d_i} - x_i|. \quad (14)$$

Here $(x_{d_i} - x_i)$ gives the average distance of nodes with point x_i in cluster i . In the end, the node which is closest to the respective centre will then become a CH and all other nodes will become member nodes of the cluster. Performance of K mean clustering mostly depends on the ideal K value. An elbow [20] algorithm is used to identify the optimum value of K.

The above discussed scheme can be modified by incorporating the GW nodes. The chances of GW nodes in this clustering are very rare and the new scheme is called the clustering-K mean-GW. These schemes have a low maintenance cost although a central authority is needed in this case. The comparison of all these schemes are shown in Table 2. All the schemes which are discussed in Table 2 are modified by us.

Table 2. Comparison between clustering schemes. GW: Gateway.

Scheme Name	Clustering Type	Clustering Overlapping	Gate Way	Location Awareness
Clustering-energy	Energy based	High	No	Not required
Clustering-energy-GW	Energy based	High	Yes	Not required
Clustering-k mean	K mean	Low	No	Required
Clustering-k mean-GW	K mean	Low	Yes	Required

3.2. Route Discovery

After clustering, CHs will have complete information about its clusters. In this special scenario, nodes only need to communicate with the temporary CC, so CHs can keep the routing table for CC. To find this routing table, the CH will start the route discovery process with the destination always set as a CC in which the UAV can also act as a relay. Initially, the CH will broadcast the Route Request (RREQ), and the CHs that receives this RREQ will reply with a Route Reply (RREP). The routing tables are maintained at the CHs. If a packet is transmitted successfully, then acknowledgment will be received. If this is not the case, then it will re-transmit the packet. With three consecutive packet re-transmissions, the respective CH will start the route discovery again and if the problem persists, then the CH will declare the destination inaccessible. The route discovery process is again initiated after every 250 transmissions and the same steps are repeated.

3.3. Routing

For routing, the node that needs to transmit will forward its packet to its respective CH. The CH will then forward this packet to the destination based on its routing table H_{i,s_i} .

In this paper, we apply RL to update the entries of the routing table to improve EE. Similar to [51], we use a linear functional approximation for the cost function. The proposed cost function is defined as:

$$Cost = \beta \left(\alpha \left(\frac{d_{mn}}{\max(d_{mn})} \right) + (1 - \alpha) \left(\frac{d_{n(\text{UAV/CC})}}{\max(d_{n(\text{UAV/CC})})} \right) \right) + (1 - \beta) \left(\frac{E_{mn}}{\max(E_{mn})} \right) \quad (15)$$

where, α is the weighting factor for shortest distance to the UAV and the next hop and β is the weighting factor to provide a balance between the distance and energy, and m and n are sender and receiver respectively. d_{mn} is the distance between m and n while $\max(d_{mn})$ is the maximum distance between m and n . Similarly E_{mn} is the transmission energy between m and n while $\max(E_{mn})$ is the maximum energy cost for the next hop. The above cost function consists of two sections, the distance and energy. This cost is only checked at a CH. The distance is balanced by α which considers the weight of the next

hop and the corresponding distance of the next hop to UAV/CC. Choosing a minimum distance path to the UAV/CC can cause significant load balancing issues. The energy consumption of forwarding CHs in such a path will also increase significantly. On the contrary, only choosing the minimum distance path for the next hop may not always be feasible as it can increase the number of hops or the distance to the UAV. The value of α finds the tradeoff in terms of distance which is subsequently used by the parameter β to manage the tradeoff among immediate energy expenditure in the next hop and distance. For example, if a CH has connectivity to multiple CHs through which it can forward its data. The scaling factor β is proportionate to the immediate energy cost of the next hop and distance cost of different routes. The terms $\max(E_{mn})$, $\max(d_{mn})$ and $\max(d_{nUAV/CC})$ normalize the energy and distance of all the possible routing paths in the next hop.

Once the costs are calculated in the current epoch, the routing is carried out. All the steps will remain the same except after the cost function RL is applied by using the equation given below:

$$RL_{mn} = (1 - \gamma)RL'_{mn} + \gamma(Rw_{mn} + V(\max(RL'_{m'n'}))). \quad (16)$$

Here RL'_{mn} is the previous value of RL_{mn} , Rw_{mn} is the reward obtained from communication which in this case is 0 for an unsuccessful transmission and 1 for a successful transmission, γ is the discount factor varying between $0 \leq \gamma \leq 1$ which tells us how much importance we want to give to the current rewards, and V is the learning rate varying between $0 \leq V \leq 1$ that tells us to what extent these RL values are updated after each iteration. $\max(RL'_{m'n'})$ is the maximum calculated cost for the next hop. $(1 - \gamma)RL'_{mn}$ takes a weight of the old RL'_{mn} value and then by adding the learned value which is the combination of Rw_{mn} and current $\max(RL'_{m'n'})$. This means an action is taken after looking at the old, current, and future rewards as shown in Algorithm 1.

Algorithm 1: Reinforcement Learning Algorithm.

Input: $N = 1, 2, \dots, i$ // Number of nodes

m & n // sender & receiver

Rw_{mn} // Reward function

$\gamma \in [0, 1]$ // Discount factor

$V \in [0, 1]$ // Learning rate

Output: RL_{mn}

Initialize $Rw_{mn}=1$

while RL is not converged **do**

 Start in state $s \in N$

while s is CH **do**

$\max(RL'_{m'n'})$ // Maximum calculated cost for the next hop

Rw_{mn} // Current reward

RL'_{mn} // Previous RL value

$RL(mn) \leftarrow (1 - V).RL'_{mn} + \gamma.(Rw_{mn} + V(\max(RL'_{m'n'})))$ $RL'_{mn} \leftarrow RL_{mn}$

end

end

3.4. Control Overhead

This subsection summarizes the cost of control overhead in clustering and non-clustering schemes. The control overhead in the above schemes can be categorized as a beacon message overhead, clustering overhead, and routing overhead.

- **Beacon messages overhead:** These beacon messages are sent by the nodes to find the information about their respective neighbors. The number of beacon messages sent are dependent on the number of nodes, N , in the environment. The nodes which are in its vicinity will reply. These beacon messages are resent after every 10 s to renew

the neighbor's information. Since there is no mobility in our scenario the neighbor change is only possible due to DNs.

- **Clustering overhead:** To calculate the clustering overhead, we categorize the clustering schemes as centralized and decentralized schemes. The schemes that employ K mean clustering are centralized schemes because they require the location information of all the nodes. We assume that after the exchange of beacon messages, location information of all the nodes is forwarded/relayed to the CC. The CC performs K mean clustering and inform the nodes about their clusters and CHs. Here, we do not consider the overhead of passing location information to CC and K mean information back to the nodes from CC. This approximate overhead can be readily found from the achievable capacity [52]. The schemes that employ clustering energy are distributed schemes. The clustering overhead is calculated when a node broadcasts control packets to declare itself as the CH. The receiving nodes will reply accordingly as discussed in Section 3.1.1. These control packets depend on the number of clusters i and the number of CMs in each cluster.
- **Routing overhead:** Once the CHs are formed, this step includes the amount of control overhead involved in the discovery of neighboring CHs and the routing path. It is assumed that the routing tables are only maintained at the CHs, which will reduce the overall routing overhead. For route discovery, the CHs will send the control packets to the neighboring CHs which will then forward the control packets all the way to the CC. The CC will confirm the routing path for each CH through a reverse response as discussed in Section 3.2.

4. Performance Analysis

For simulations, we consider an area of $100 \times 100 \text{ m}^2$ with $N = 100$ nodes is considered as shown in Figure 4 with the UAV placed at the edge and CC placed further away for safety reasons. The UAV and CC are static and their placement allows a limited number of nodes in their transmission range. These nodes (GW and CH) are consequently used to reach the UAV and CC in a multi-hop manner resulting in rapid depletion of their battery.

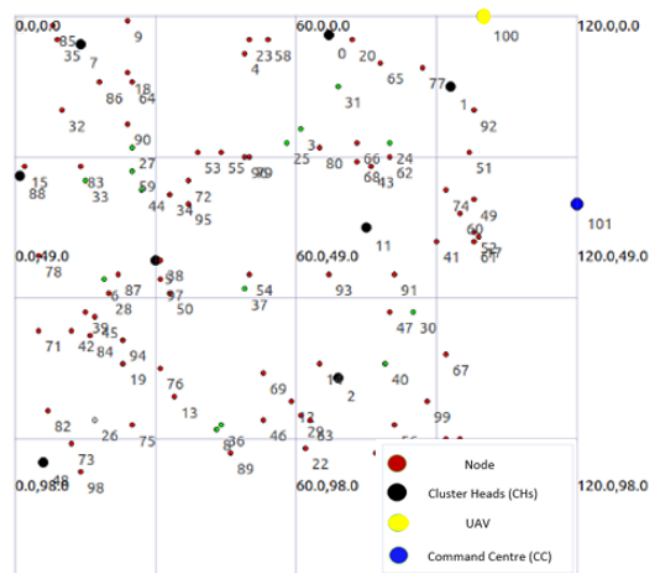


Figure 4. Topology of the network in NS-3.

In Figure 4, the blue color is used for CC, the yellow color is used for UAV, the black color is used for CHs, the green color is used for GWs, the red is used for CMs, and the white color is used for Dead Nodes (DNs). The white nodes are not active and unable to communicate due to low or no $E_{initial}$. This topology is obtained after running NS-3-based simulations for several rounds. The simulation parameters are shown in Table 3.

Table 3. Simulation parameters.

Parameter	Values
Number of Devices (N)	100
Network Grid	100 m × 100 m
CC Placement	(120, 35) m
UAV Placement (Initial)	(100, 0, 10) m
Size of Data Packet (L)	1024 bytes
Header Size	40 bytes
Initial Power Level	0 to 1 J
E_{Tx}	50 nJ/bit
E_{Rx}	50 nJ/bit
Threshold	4.35 mJ
Cluster Range, R_{max}	30 m
CH Tx Range, T_{max}	45 m
Distance b/w UAV and CC	60 m
Max Transmissions in a Round (N_{Tx})	5
I_{max}	1024
α	0.5
β	0.5
Discount Factor (γ)	0.8
Learning Rate (V)	0.4

For a fair comparison, we include the results of non-clustering schemes based on Dijkstra [53] along with schemes discussed in Section 3.1. To get a better insight into the schemes performance, we assigned random energy varying between 0 and 1 J to all the nodes. Figure 5 shows the number of DNs where the total number of DNs is very high because nodes have random $E_{initial}$ and they die quickly. The figure shows that both Dijkstra-based schemes had the highest number of DNs. This is approximately 57% higher compared to clustering-energy at 250 s, respectively.

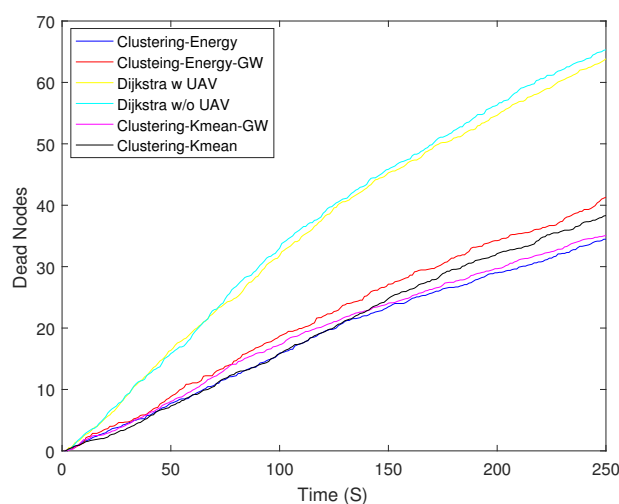
**Figure 5.** Number of dead nodes.

Figure 6 shows the residual energy of different schemes. As expected, the Dijkstra-based scheme had the lowest residual energy at 250 s. The Dijkstra-based scheme consumed 61% higher energy as compared with clustering-K mean-GW at 250 s. Figure 7 shows the throughput of all the schemes. Until 90 s, Dijkstra with UAV shows the highest throughput, but after 90 s its curve starts to saturate due to the increasing number of DNs. Clustering-energy-GW has the highest throughput after 250 s i.e., 38 and 40% higher as compared to Dijkstra with UAV and clustering-energy, respectively. Figure 8 shows the EE of different

schemes. The clustering-K mean-GW has highest throughput per unit of energy until 35 s. Beyond 35 s, clustering-K mean shows highest throughput starting from 75 s to 180 s. Both the Dijkstra-based schemes have the lowest EE and decrease further with time. The clustering-K mean has a 29 and 45% higher EE as compared to clustering-energy and Dijkstra without UAV. Figure 9 shows end-to-end delay. The delay is very high in all the schemes because of the high hop count. In comparison, Dijkstra without UAV shows the highest delay followed by Dijkstra with UAV. Clustering-K mean-GW offers a 75% lower delay when compared with Dijkstra without the UAV.

Figure 10 shows Packet Delivery Ratio (PDR) of different schemes. In the beginning, both the non-clustering schemes (Dijkstra) have a higher PDR upto 75 s. Afterwards their performance degrades and Dijkstra w/o UAV has the lowest PDR. Compared to (Dijkstra), the PDR of clustering schemes decays slowly. The clustering-energy-GW have the highest PDR after 150 s and it is 9.2% higher than that of clustering-K mean at 250 s.

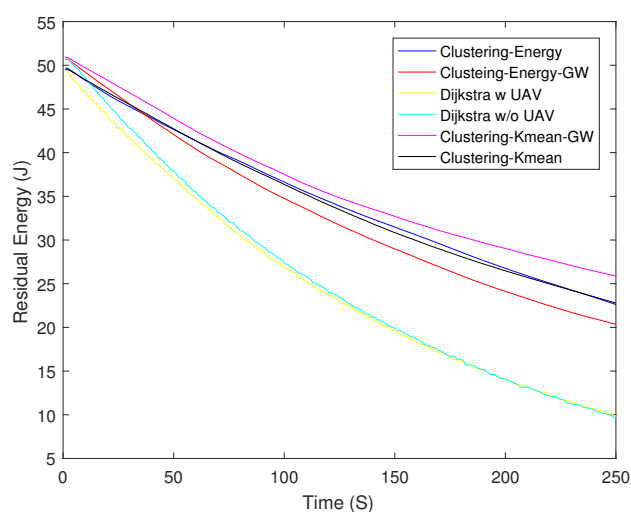


Figure 6. Residual energy.

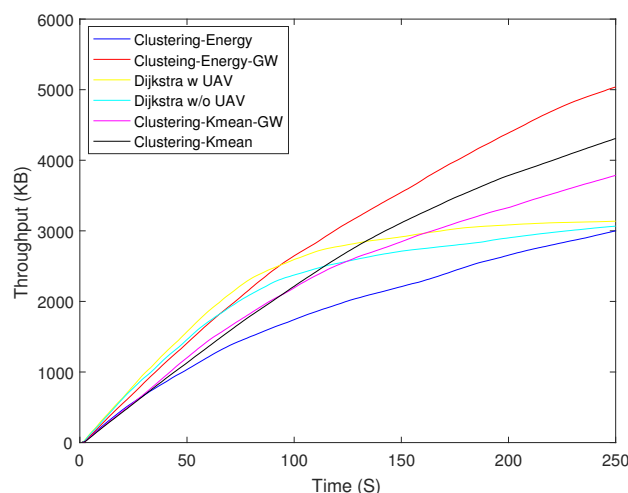


Figure 7. Throughput.

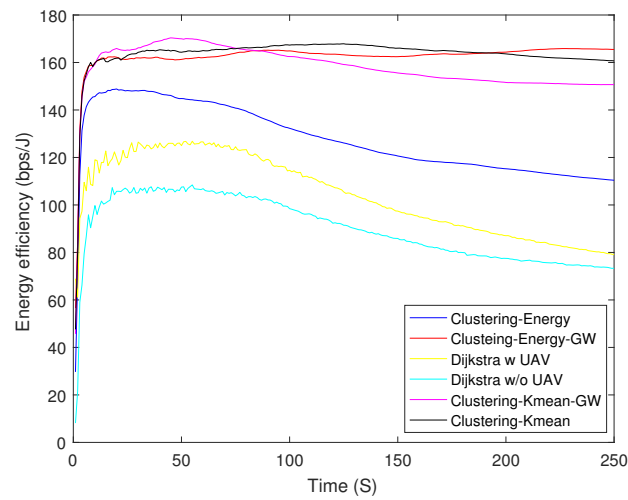


Figure 8. Energy efficiency.

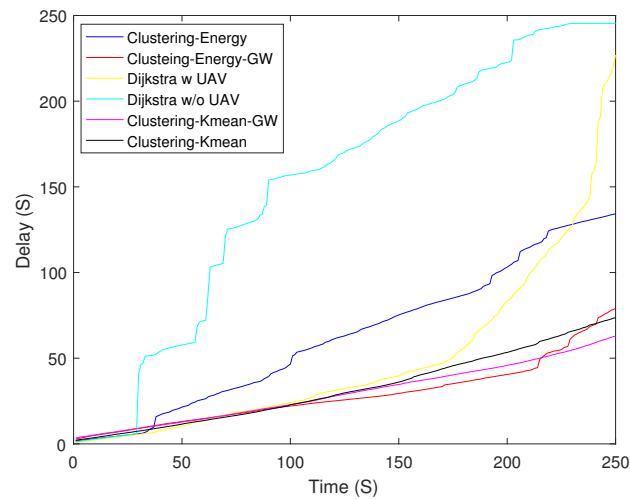


Figure 9. End-to-end delay.

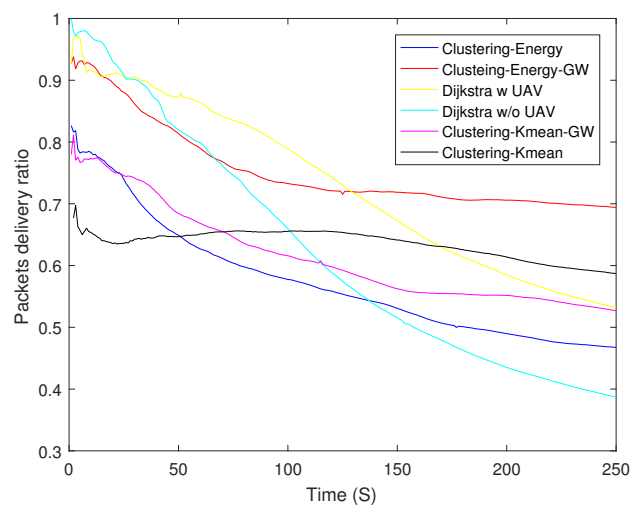


Figure 10. Packet delivery ratio.

4.1. Reinforcement-Based Routing

From earlier results we observed that clustering-K mean performs the best in EE, therefore, we apply RL on the best performing scheme. Figure 11 shows the EE between the best perform-

ing scheme clustering-K mean and clustering-K mean-RL. The results show an improvement of up to 15% at 10 s when compared with the underlying scheme. In addition, we have also compared RL-based distance and energy only variant of Equation (15). Its worth mentioning that RL-based distance only variant performed much better compared to both conventional Dijkstra and Dijkstra with UAV. Figure 12 shows PDR between the clustering-K mean and its RL variant. RL-based K mean variant shows at least a 4% improvement.

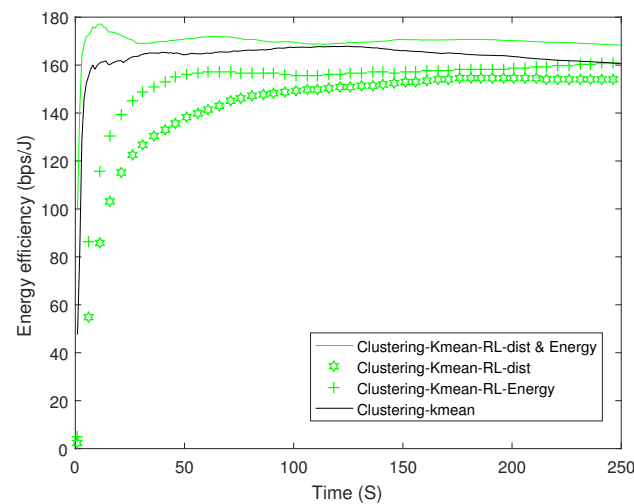


Figure 11. Energy efficiency of the reinforcement learning scheme. RL: Reinforcement Learning.

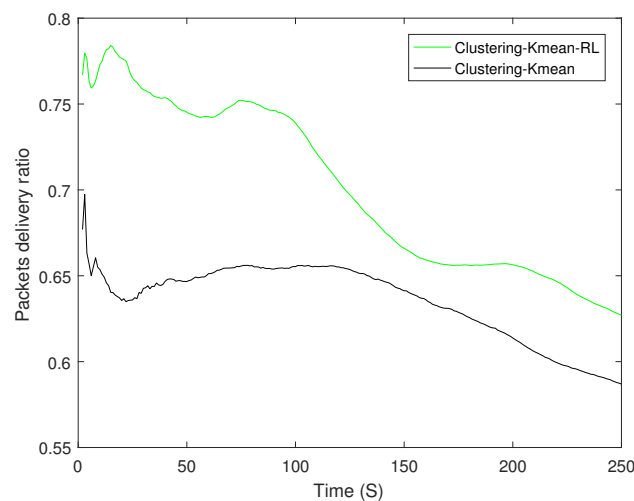


Figure 12. Packet delivery ratio of the reinforcement learning scheme.

Figure 13 shows a comparison of the total number of control packets sent by all the schemes including clustering and Dijkstra-based schemes. Both Dijkstra-based schemes transmit almost the same number of control packets. It is interesting to observe that the control overhead of Dijkstra-based schemes is approximately 90% higher than the clustering-energy-GW. Figure 14 shows the same comparison between different clustering schemes. The control overhead of clustering-energy-GW is 32% higher when compared to clustering-energy, clustering-K mean-GW, and clustering-K mean-RL.

Figure 15 shows the number of CH changes in different clustering schemes. The energy-based schemes show the highest number of CH changes when compared with the K mean clustering schemes because of their distributed nature. Clustering-energy-GW shows 53% more CH changes when compared with clustering-K mean-GW.

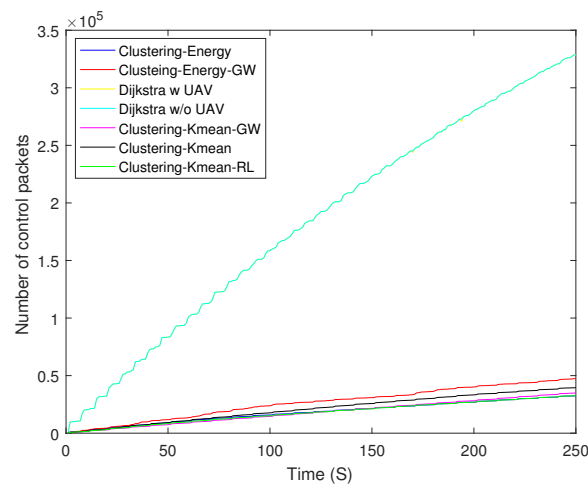


Figure 13. Control packets sent by all the schemes.

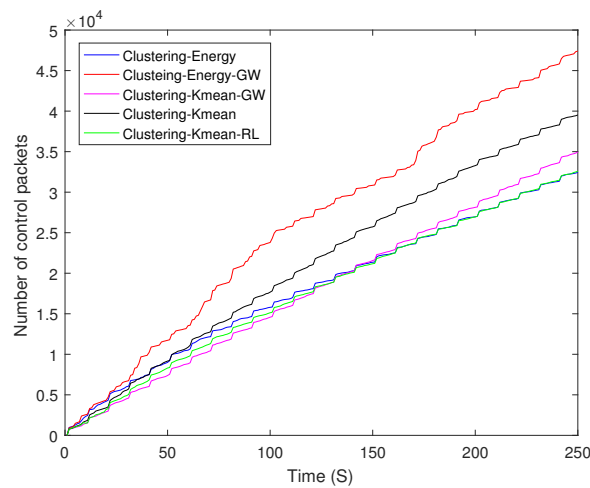


Figure 14. Control packets sent by different clustering schemes.

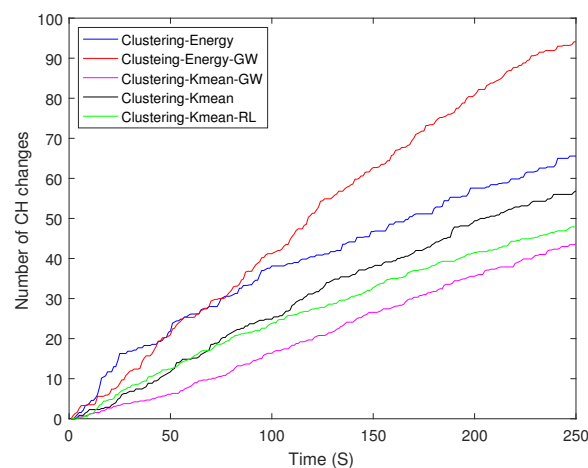


Figure 15. Number of CH changes in different clustering schemes.

4.2. Combined RL and UAV(s) Trajectory Optimization

In the previous sections, the UAV was considered to be a static node. However, the UAVs act as flying relays and with an adequately designed flight trajectory they can provide uniform coverage to all the CHs, thus increasing EE. In this paper we consider the placement of single and multiple UAVs and analyze the impact of their trajectory. The

purpose of multiple UAVs is to decrease the number of hops and provide an improved EE. In the disaster scenarios, the energy of the devices is very critical so to further improve EE and connectivity, a multiple-UAV scenario is applied. The speed of UAVs in motion is 2.70 m/s. The transmission range of a UAV is 60 m. If the CHs falls in the transmission range of two UAVs, the CH will choose the UAV closest to the CC. The RL is used to find the best routing paths in the scenario given in Figure 16. Figure 16a provides three dimensional view of the disaster scenario and Figure 16b shows the trajectory paths of UAV's around the disaster scenario. We consider pre-defined UAV flight paths around the area which are discussed below.

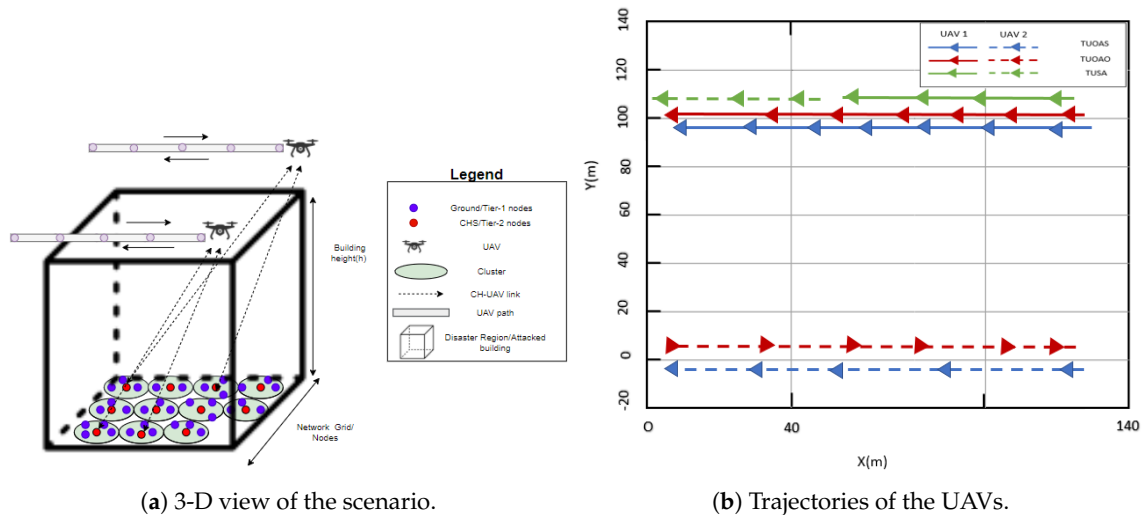


Figure 16. Trajectories of UAVs around the disaster area.

- Two UAVs on Opposite Axis and Same direction (TUDOAS):** In this scheme, the two UAVs are placed at the edge of the building. Both the UAVs are placed on the opposite corners of the building. They start moving from the same side of the building, as shown in Figure 16. Both these UAVs are moving in parallel to each other but the direction they are following is the same. The trajectory they are following is along the straight line alongside the building. When they reach the opposite corner of the building, they will follow the same path backwards. The CHs that are in the range of any of these UAVs will send their packets through the respective UAV. Thus, the Equation (15) for the two UAVs will be modified as:

$$Cost_1 = \beta \left(\alpha \left(\frac{d_{m_1 n_1}}{\max(d_{m_1 n_1})} \right) + (1 - \alpha) \left(\frac{d_{n_1(UAV_1)}}{\max(d_{n_1(UAV_1)})} \right) \right) + (1 - \beta) \left(\frac{E_{m_1 n_1}}{\max(E_{m_1 n_1})} \right). \quad (17)$$

Here $Cost_1$ is the cost associated with UAV₁ in which m_1 and n_1 are the sender and receiver nodes associated to the UAV₁:

$$Cost_2 = \beta \left(\alpha \left(\frac{d_{m_2 n_2}}{\max(d_{m_2 n_2})} \right) + (1 - \alpha) \left(\frac{d_{n_2(UAV_2)}}{\max(d_{n_2(UAV_2)})} \right) \right) + (1 - \beta) \left(\frac{E_{m_2 n_2}}{\max(E_{m_2 n_2})} \right). \quad (18)$$

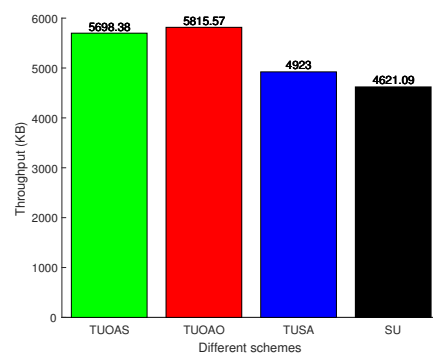
Similarly $Cost_2$ is the associated cost with the UAV₂. While m_2 and n_2 are the sender and receiver nodes associated to the UAV₂. In the case, sender m is in the direct range of CC, the cost will be calculated using Equation (15). The above equations can also be used for all other two UAV schemes presented below.

- Two UAVs on Opposite Axis and Opposite Direction (TUOAO):** In this scheme two UAVs are placed at the opposite corner of the building as shown in Figure 16. Both UAVs move along a straightline alongside the building towards their respective direction. By moving in this way they will help in maximize the coverage area of the building affected by the disaster. When a UAV reaches the edge of the building

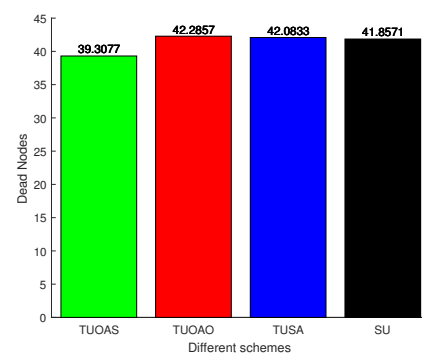
it will follow the same path backwards and it keeps on doing this till the end of the simulation.

- **Two UAVs moving on Same Axis (TUSA):** In this scheme, both the UAVs were placed on the same axis separated by 60 m, as shown in Figure 16. Both the UAVs move in the same direction and on reaching their respective endpoint they follow the same path backwards. The maximum separation between them remains the same.
- **Single UAV in motion (SU):** In this scheme, we placed a single UAV at the corner of the building. The UAV moves alongside the building and traverses the same path on its way back from the end of the building.

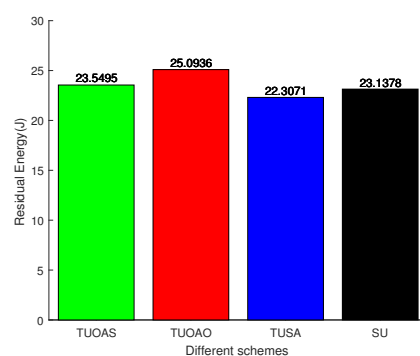
We have considered variable energy for the nodes varying between 0 and 1 J. The bars in Figure 17a show that all the schemes achieve their maximum throughput after 250 s. The TUOAO records the highest throughput and provide 2%, 15%, and 20% gain when compared with TUOAS, TUSA, and SU, respectively. Figure 17b presents the number of DNs for each scheme. The difference between the schemes TUOAO, TUSA, and SU was minor and only TUOAS showed 7% lesser DNs compared to the other three schemes. Figure 17c shows the residual energy after 250 s. TUOAO has the most residual energy whereas TUOAS was second with 6% lower residual energy, and SU and TUSA are third and fourth, respectively. Figure 17d shows EE. All the schemes present a higher performance compared to the SU case. TUOAO presents the highest gain in EE, whereas the EE of TUOAS is 10% lower. The EE of TUSA and SU is 23% and 27% lower when compared to TUOAO. Figure 18 compares the number of control packets sent for different trajectory schemes. The scheme with a single UAV sends the least amount of control packets, 14% lower than that of TUOAS. Figure 19 shows the number of CH changes against time. The schemes with two UAVs have a higher number of CH changes. This is mainly due to the routing overhead induced by the movement of the UAVs, which result from frequent changes in routing paths. Intuitively, the SU scheme in comparison has 50% fewer CH changes than the TUOAS.



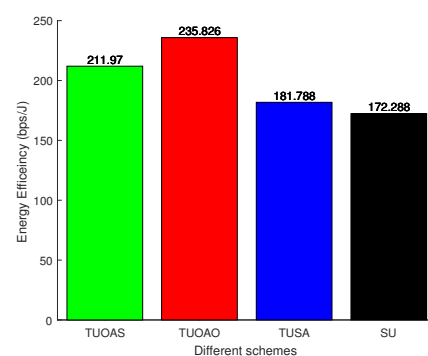
(a) Throughput.



(b) Number of dead nodes.



(c) Residual energy.



(d) Energy efficiency.

Figure 17. Results after 250 s.

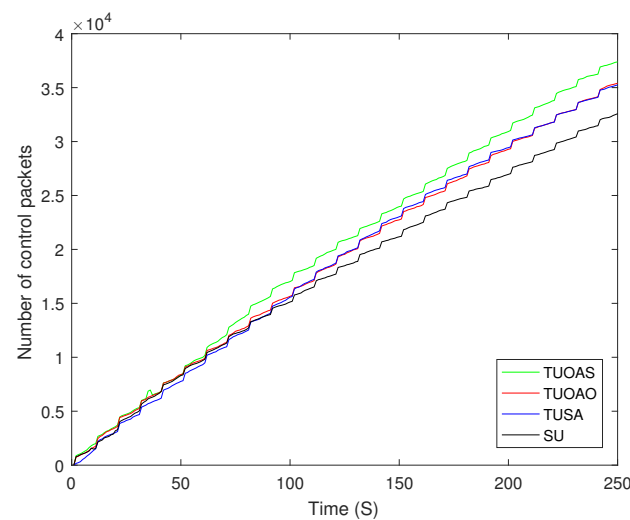


Figure 18. Number of control packets.

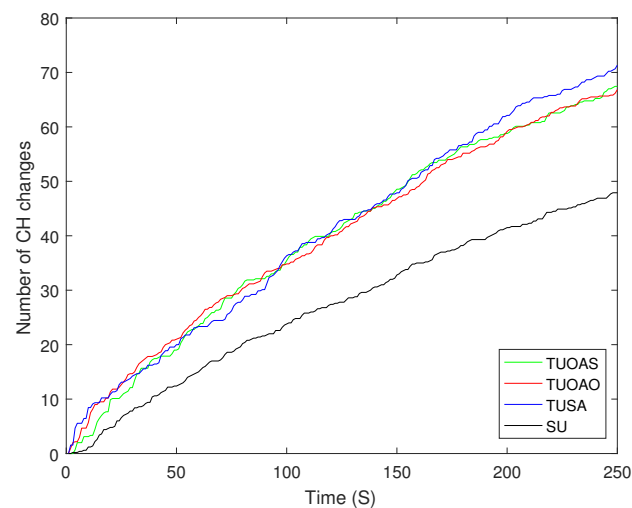


Figure 19. Number of CH changes.

5. Conclusions

In this work, multiple routing schemes were evaluated for a man-made disaster scenario. The main concern in these scenarios is the network lifetime achieved through EE. As expected, the simulation results showed clustering schemes had a longer network lifetime when compared with non-clustering schemes. In terms of throughput, clustering-energy-GW had 40% higher throughput than clustering-energy. However, in terms of EE, clustering-K mean showed the best performance. We then applied RL on the clustering-K mean which further improved the EE by 15% at 10 s. Inclusion of multiple UAVs further improved EE, however, the amount of improvement was highly dependent on the trajectory of these UAVs. In a rectangular disaster area, maximum EE was achieved when the UAVs started scanning linearly from the opposite ends of the building while maximizing the coverage. However, this came at the cost of increased routing overhead and cluster changes. In future, we plan to get a holistic picture by incorporating the impact of control packets. It is also possible to explore the use of localization techniques to find the exact location of the nodes, which is useful in planning multi-UAV deployment.

Author Contributions: H.I.M., R.A. and W.A. contributed the key idea and defined the problem statement. R.A. and W.A. helped with the system model, simulation framework, basic clustering schemes, and UAV trajectories. H.I.M. performed the implementation, analysis, and simulations. M.W. contributed to the design of the RL based approach. M.M.A. and S.T.G. provided statistical

input and integration of results. H.I.M., R.A., W.A. and M.W. were involved in the preparation of the original draft of the paper. M.M.A. and S.T.G. helped review the paper. The improvements in the write up were contributed by all. All authors have read and agreed to the published version of the manuscript.

Funding: This research has received funding from the NATO-SPS funding grant agreement no. G5482.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Zakaria, M.; Jun, W.; Ahmed, H. Effect of terrorism on economic growth in Pakistan: an empirical analysis. *Econ. Res. Ekonomska Istraživanja* **2019**, *32*, 1794–1812. [[CrossRef](#)]
- Qureshi, R.; Gulraiz, A.; Shahzad, Z. An Analysis of Media's Role: Case Study of Army Public School (APS) Peshawar Attack. *Soc. Commun.* **2016**, *2*, 20–30. [[CrossRef](#)]
- Masood, A.; Scazzoli, D.; Sharma, N.; Moullec, Y.L.; Ahmad, R.; Reggiani, L.; Magarini, M.; Alam, M.M. Surveying pervasive public safety communication technologies in the context of terrorist attacks. *Phys. Commun.* **2020**, *41*, 101109. [[CrossRef](#)]
- Ali, K.; Nguyen, H.X.; Shah, P.; Vien, Q.T.; Bhuvanansundaram, N. Architecture for public safety network using D2D communication. In Proceedings of the 2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Doha, Qatar, 3–6 April 2016; pp. 206–211.
- Shaikh, F.S.; Wismüller, R. Routing in multi-hop cellular device-to-device (D2D) networks: A survey. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2622–2657. [[CrossRef](#)]
- Muraoka, K.; Shikida, J.; Sugahara, H. Feasibility of capacity enhancement of public safety LTE using device-to-device communication. In Proceedings of the 2015 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, 28–30 October 2015; pp. 350–355.
- Anupama, M.; Sathyanarayana, B. Survey of cluster based routing protocols in mobile adhoc networks. *Int. J. Comput. Theory Eng.* **2011**, *3*, 806. [[CrossRef](#)]
- Ephremides, A.; Wieselthier, J.E.; Baker, D.J. A design concept for reliable mobile radio networks with frequency hopping signaling. *Proc. IEEE* **1987**, *75*, 56–73. [[CrossRef](#)]
- Chiang, C.C.; Wu, H.K.; Liu, W.; Gerla, M. Routing in clustered multihop, mobile wireless networks with fading channel. In Proceedings of the IEEE SICON, Singapore, 14–17 April 1997; Volume 97, pp. 197–211.
- Fodor, G.; Parkvall, S.; Sorrentino, S.; Wallentin, P.; Lu, Q.; Brahma, N. Device-to-device communications for national security and public safety. *IEEE Access* **2014**, *2*, 1510–1520. [[CrossRef](#)]
- Lu, Q.; Miao, Q.; Fodor, G.; Brahma, N. Clustering schemes for D2D communications under partial/no network coverage. In Proceedings of the 2014 IEEE 79th Vehicular Technology Conference (VTC Spring), Seoul, Korea, 18–21 May 2014; pp. 1–5.
- Guizani, B.; Ayeb, B.; Koukam, A. Hierarchical cluster-based link state routing protocol for large self-organizing networks. In Proceedings of the 2011 IEEE 12th International Conference on High Performance Switching and Routing, Cartagena, Spain, 4–6 July 2011; pp. 203–208.
- Laha, A.; Cao, X.; Shen, W.; Tian, X.; Cheng, Y. An energy efficient routing protocol for device-to-device based multihop smartphone networks. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 5448–5453.
- Chang, T.C.; Wei, C.; Hsu, M.; Lin, C.; Su, Y.T. Distributed clustering and spectrum-based proximity device discovery in a wireless network. In Proceedings of the 2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Nara, Japan, 1–3 June 2016; pp. 1–4.
- Islam, N.; Dey, S.; Sampalli, S. Energy-Balancing Unequal Clustering Approach to Reduce the Blind Spot Problem in Wireless Sensor Networks (WSNs). *Sensors* **2018**, *18*, 4258. [[CrossRef](#)]
- Minhas, H.I.; Ahmad, R.; Ahmed, W.; Alam, M.M.; Magarani, M. On the impact of clustering for Energy critical Public Safety Networks. In Proceedings of the 2019 International Symposium on Recent Advances in Electrical Engineering (RAEE), Islamabad, Pakistan, 28–29 August 2019; Volume 4, pp. 1–5.
- Echoukairi, H.; Kada, A.; Bouragba, K.; Ouzzif, M. A novel centralized clustering approach based on K-means algorithm for wireless sensor network. In Proceedings of the 2017 Computing Conference, London, UK, 18–20 July 2017; pp. 1259–1262.
- Gupta, A.; Shekhar, N. A novel K-means L-layer algorithm for uneven clustering in WSN. In Proceedings of the 2017 International Conference on Computer, Communication and Signal Processing (ICCCSP), Chennai, India, 10–11 January 2017; pp. 1–6.

19. Razzaq, M.; Devi Ningombam, D.; Shin, S. Energy efficient K-means clustering-based routing protocol for WSN using optimal packet size. In Proceedings of the 2018 International Conference on Information Networking (ICOIN), Chiang Mai, Thailand, 10–12 January 2018; pp. 632–635.
20. Bholowalia, P.; Kumar, A. Article: EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN. *Int. J. Comput. Appl.* **2014**, *105*, 17–24.
21. Arbia, D.B.; Alam, M.M.; Attia, R.; Hamida, E.B. ORACE-Net: A novel multi-hop body-to-body routing protocol for public safety networks. *Peer-Peer Netw. Appl.* **2017**, *10*, 726–749. [[CrossRef](#)]
22. Yuan, H.; Guo, W.; Wang, S. Emergency route selection for D2D cellular communications during an urban terrorist attack. In Proceedings of the 2014 IEEE International Conference on Communications Workshops (ICC), Sydney, Australia, 10–14 June 2014; pp. 237–242.
23. Hassan, A.; Ahmad, R.; Ahmed, W.; Magarini, M.; Alam, M.M. UAV and SWIPT Assisted Disaster Aware Clustering and Association. *IEEE Access* **2020**, *8*, 204791–204803. [[CrossRef](#)]
24. Li, X.; Guo, D.; Grosspietsch, J.; Yin, H.; Wei, G. Maximizing mobile coverage via optimal deployment of base stations and relays. *IEEE Trans. Veh. Technol.* **2015**, *65*, 5060–5072. [[CrossRef](#)]
25. Zeng, Y.; Zhang, R.; Lim, T.J. Wireless communications with unmanned aerial vehicles: Opportunities and challenges. *IEEE Commun. Mag.* **2016**, *54*, 36–42. [[CrossRef](#)]
26. Mozaffari, M.; Saad, W.; Bennis, M.; Debbah, M. Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 3949–3963. [[CrossRef](#)]
27. Al-Hourani, A.; Kandeepan, S.; Lardner, S. Optimal LAP altitude for maximum coverage. *IEEE Wirel. Commun. Lett.* **2014**, *3*, 569–572. [[CrossRef](#)]
28. Košmerl, J.; Vilhar, A. Base stations placement optimization in wireless networks for emergency communications. In Proceedings of the 2014 IEEE International Conference on Communications Workshops (ICC), Sydney, Australia, 10–14 June 2014; pp. 200–205.
29. Hassan, A.; Ahmad, R.; Ahmed, W.; Magarini, M.; Alam, M.M. Managing Critical Nodes in UAV assisted Disaster Networks. In Proceedings of the 2020 17th Biennial Baltic Electronics Conference (BEC), Tallinn, Estonia, 6–8 October 2020; pp. 1–5. [[CrossRef](#)]
30. Alam, M.M.; Le Moullec, Y.; Ahmad, R.; Magarini, M.; Reggiani, L. A Primer On Public Safety Communication in the Context of Terror Attacks: The NATO SPS “COUNTER-TERROR” Project. In *Advanced Technologies for Security Applications*; Palestini, C., Ed.; Springer: Dordrecht, The Netherlands, 2020; pp. 19–34.
31. Valente Klaine, P.; Nadas, J.; Souza, R.; Imran, M. Distributed Drone Base Station Positioning for Emergency Cellular Networks Using Reinforcement Learning. *Cogn. Comput.* **2018**, *10*. [[CrossRef](#)]
32. Hydher, H.; Jayakody, D.N.K.; Hemachandra, K.T.; Samarasinghe, T. Intelligent UAV deployment for a disaster-resilient wireless network. *Sensors* **2020**, *20*, 6140. [[CrossRef](#)] [[PubMed](#)]
33. Lin, N.; Fu, L.; Zhao, L.; Min, G.; Al-Dubai, A.; Gacanin, H. A Novel Multimodal Collaborative Drone-Assisted VANET Networking Model. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 4919–4933. [[CrossRef](#)]
34. Zhao, L.; Yang, K.; Tan, Z.; Li, X.; Sharma, S.; Liu, Z. A Novel Cost Optimization Strategy for SDN-Enabled UAV-Assisted Vehicular Computation Offloading. *IEEE Trans. Intell. Transp. Syst.* **2020**. [[CrossRef](#)]
35. Zhang, S.; Zhang, H.; Song, L. Beyond D2D: Full Dimension UAV-to-Everything Communications in 6G. *IEEE Trans. Veh. Technol.* **2020**, *69*, 6592–6602. [[CrossRef](#)]
36. Wang, L.; Wang, K.; Pan, C.; Xu, W.; Aslam, N.; Hanzo, L. Multi-Agent Deep Reinforcement Learning Based Trajectory Planning for Multi-UAV Assisted Mobile Edge Computing. *IEEE Trans. Cogn. Commun. Netw.* **2020**. [[CrossRef](#)]
37. Atif, M.; Ahmad, R.; Ahmad, W.; Zhao, L.; Rodrigues, J.J.P.C. UAV-Assisted Wireless Localization for Search and Rescue. *IEEE Syst. J.* **2021**, 1–12. [[CrossRef](#)]
38. Arafat, M.Y.; Habib, M.A.; Moh, S. Routing Protocols for UAV-Aided Wireless Sensor Networks. *Appl. Sci.* **2020**, *10*, 4077. [[CrossRef](#)]
39. Zhan, C.; Zeng, Y.; Zhang, R. Energy-Efficient Data Collection in UAV Enabled Wireless Sensor Network. *IEEE Wirel. Commun. Lett.* **2018**. [[CrossRef](#)]
40. Gomez, J.M.; Wiedemann, T.; Shutin, D. Unmanned Aerial Vehicles in Wireless Sensor Networks: Automated Sensor Deployment and Mobile Sink Nodes. In Proceedings of the International Conference on Intelligent Autonomous Systems, Baden-Baden, Germany, 11–15 June 2018.
41. Uddin, M.A.; Mansour, A.; Jeune, D.L.; Ayaz, M.; Aggoune, E.H.M. UAV-Assisted Dynamic Clustering of Wireless Sensor Networks for Crop Health Monitoring. *Sensors* **2018**, *18*, 555. [[CrossRef](#)] [[PubMed](#)]
42. Zema, N.R.; Mitton, N.; Ruggeri, G. Using location services to autonomously drive flying mobile sinks in wireless sensor networks. In Proceedings of the International Conference on Ad Hoc Networks, San Remo, Italy, 1–2 September 2015; pp. 180–191.
43. Villas, L.A.; Guidoni, D.L.; Maia, G.; Pazzi, R.W.; Ueyama, J.; Loureiro, A.A. An energy efficient joint localization and synchronization solution for wireless sensor networks using unmanned aerial vehicle. *Wirel. Netw.* **2015**, *21*, 485–498. [[CrossRef](#)]
44. Abu-Salih, A.T.; Seno, S.A.H. Energy-efficient data gathering framework-based clustering via multiple UAVs in deadline-based WSN applications. *IEEE Access* **2018**, *6*, 72275–72286. [[CrossRef](#)]
45. Dong, M.; Ota, K.; Lin, M.; Tang, Z.; Du, S.; Zhu, H. UAV-assisted data gathering in wireless sensor networks. *J. Supercomput.* **2014**, *70*, 1142–1155. [[CrossRef](#)]

46. Okcu, H.; Soy Turk, M. Distributed clustering approach for UAV integrated wireless sensor networks. *Int. J. Hoc Ubiquitous Comput.* **2014**, *15*, 106–120. [[CrossRef](#)]
47. Hashima, S.; Hatano, K.; Mohammed, E. Multiagent Multi-Armed Bandit Schemes for Gateway Selection in UAV Networks. In Proceedings of the 2020 IEEE Globecom Workshops (GC Wkshps), Taipei, Taiwan, 7–11 December 2020; pp. 7–11.
48. Mohamed, E.M.; Hashima, S.; Aldosary, A.; Hatano, K.; Abdelghany, M.A. Gateway Selection in Millimeter Wave UAV Wireless Networks Using Multi-Player Multi-Armed Bandit. *Sensors* **2020**, *20*, 3947. [[CrossRef](#)]
49. Hodge, V.J.; Hawkins, R.; Alexander, R. Deep reinforcement learning for drone navigation using sensor data. *Neural Comput. Appl.* **2020**, *33*, 2015–2033 [[CrossRef](#)]
50. Bultitude, Y.D.J.; Rautiainen, T. IST-4-027756 WINNER II D1. 1.2 V1. 2 WINNER II Channel Models. In *EBITG, TUI, UOULU, CU/CRC, NOKIA*; Tech. Rep; 2007. Available online: <http://www.ero.dk/93F2FC5C-0C4B-4E44-8931-00A5B05A331B?frames=no&> (accessed on 10 March 2021).
51. Kiani, F.; Amiri, E.; Zamani, M.; Khodadadi, T.; Abdul Manaf, A. Efficient intelligent energy routing protocol in wireless sensor networks. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 618072. [[CrossRef](#)]
52. Gupta, P.; Kumar, P.R. The capacity of wireless networks. *IEEE Trans. Inf. Theory* **2000**, *46*, 388–404. [[CrossRef](#)]
53. Johnson, D.B. A note on Dijkstra's shortest path algorithm. *J. ACM* **1973**, *20*, 385–388. [[CrossRef](#)]