

Received 26 October 2022, accepted 9 November 2022, date of publication 16 November 2022, date of current version 23 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3222566

RESEARCH ARTICLE

HVDC Transmission Line Fault Identification: A Learning Based UAV Control Strategy

WAHEED A. KHAN¹, SHAHZAIB HAMID², M. REHAN USMAN¹, (Member, IEEE), ALI RAZA³, M. ARSLAN USMAN⁴, (Member, IEEE), CHRISTOS POLITIS⁴, (Senior Member, IEEE), AND GANDEVA BAYU SATRYA⁵, (Senior Member, IEEE)

¹Department of Electrical Engineering, Superior University, Lahore 54660, Pakistan

²Department of Research and Development, AI Studio, Johar Town, Lahore 54000, Pakistan

³Department of Electrical Engineering, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan

⁴Faculty of Engineering, Computing and Environment (ECE), Kingston University London, KT1 2EE London, U.K.

⁵School of Applied Sciences, Telkom University, Bandung 40257, Indonesia

Corresponding author: M. Rehan Usman (rehan.usman@superior.edu.pk)

This work was supported and funded by wireless world research forum (WWRF).

ABSTRACT Electricity Transmission plays an imperative role in smooth provision of power to the consumers. High voltage direct current (HVDC) system has a lead over high voltage alternate current (HVAC) system in various aspects. As DC transmissions lines transmit electricity over long distance, it is crucial to find the malfunctioning part of the line in case of faults occurrence. In this work, a decision-making based unmanned aerial vehicle (UAV) control strategy is presented for identifying fault location in HVDC transmission lines. The technique is developed on two layered control systems, i.e., command station (Leader Agent) and UAV agents (Local Agents) control. The Markov decision process (MDP) based reward policy for both agents is defined mathematically and has been implemented in MATLAB to depict their behavior. The resulting policy is optimized through the value iteration algorithm based on reward functions and transition probabilities.

INDEX TERMS Fault identification, high voltage DC (HVDC), multiagent system (MAS), markov decision process (MDP), unmanned aerial vehicle (UAV).

I. INTRODUCTION

Reward based strategy is the key to policy-based decision-making process. Such strategies are applicable to any industrial based solution where decision making is of utmost importance. When it comes to fault identification in power transmission systems, locating the precise area of fault occurrence plays a vital role in providing timely corrective actions. In the current transmission systems, especially in developing countries, the available solutions for locating the fault are either very expensive, time taking or does not exist at all. Using unmanned ariel vehicles (UAVs) combined with reward-based movement strategies can act as an alternative approach for the given problem, especially in remote areas or long-distance transmission lines. Such strategies provide input states for defining the policy function based on which

UAV movement decision is performed. In high voltage direct current (HVDC) system, fault identification in transmission lines becomes even more sensitive due to high voltage nature and bulk power transmission capacity. So, an unmanned approach is much more feasible for such systems [1], [2].

HVDC system is more reliable because it ensures safe operation as compared to high voltage alternate current (HVAC) system [3]. The HVDC grid is divided into different zones and only that zone is isolated from the network which is faulty or malfunction [4]. The DC grids are normally protected by the use of converters with the capability of fault blocking which interrupts the DC fault current. However, this limits its applicability to small systems [5]. For efficient fault protection, the second step is to identify the precise fault location so that a timely remedial action can be taken to stabilize power transmission [6]. The deployment of HVDC system strongly demands a reliable protection system and timely fault location identification ensures it. Till date, the HVDC

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna Dulizia^{id}.

system relies on technical fault identification approaches like travelling wave approach [7]. However, general drawbacks of such approaches are poor reliability and low accuracy, as they can identify an area of fault but not the location of the fault. Due to this problem, manned mission is required to recon the identified faulty area for exact fault location, which is time taking, risky and costly depending on the terrain. Keeping all this in view, an unmanned autonomous learning-based fault location identification strategy is proposed in this work. The strategy is developed considering autonomous control of UAVs using policy decision making process defined by MDP technique. The proposed work acts as basis for providing learning-based decision-making process for efficient movement of UAV to locate the line faults. The proposed approach does not involve manned missions which in case of dangerous terrains are not only cost effective but also less risky.

The paper has been divided into following sections, sections II include literature review, section III describes the proposed fault identification model for HVDC system, section IV gives results and discussions using different scenarios of the MDP implementation strategy based on multi-agent UAV system in HVDC line. Finally, section V provides conclusion and future extensions of this work.

II. LITERATURE REVIEW

To date several fault identifications schemes have been presented for power transmission systems. The directional protection scheme has been discussed in [8] for the realization of internal and external faults. The basic principle lies in the integral of reactive power. Theoretical analysis of directional characteristics is implemented on HVDC test system which is modeled in PSCAD. Sneath and Rajapakse [9] discussed a protection scheme which was implemented on an earthed bi-pole HVDC grid with the use of rate of change of voltage (ROCOV) and hybrid direct current circuit breakers. The effect of different sizes of inductors with the rise of current was studied and necessary protection thresholds on HVDC grid were also analyzed. The authors in [10] developed a transient energy-based protection scheme for HVDC transmission line by considering fault resistance and transmission distance as two main factors. In [11], the authors have proposed a unique fault location algorithm based on distributed parameter line model in which fault position and fault resistance are not required. Authors have made a comparison of their algorithm with travelling wave-based method and concluded that this algorithm is a better fault location method because it requires low sampling frequency and can use any segment of post fault data to identify faults.

Liu et al. [12] discussed travelling wave-based fault identification method using Hilbert Haung Transform and Ensemble empirical mode decomposition (EEMD) time frequency graph. Haleem et al. [13] proposed robustness technique of low to high resistance fault detection schemes at different grid and operating configurations. Fault estimation has been done in [14] using discrete wavelet transform and extreme learning machine in an HVDC transmission lines.

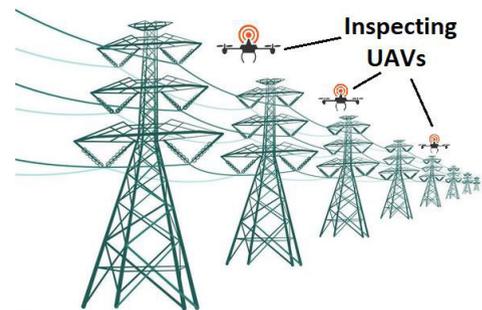


FIGURE 1. Power transmission line inspection using UAVs.

The discrete wavelet transform-coefficients have been used to find the energy of the signal and Shannon's Entropy in [15]. Yusuff et al. [16] proposed a fault location technique on a 400 KV and 361.297 km long transmission line using stationary wavelet transform and determinant function feature. In [17], a fault identification method has been presented based on single ended travelling wave theory, discrete wavelet transforms and support vector machine. The transmission system is segmented into overhead and underground cable. A robust internal faults identification system for double circuit HVDC transmission line was introduced by Yanjun et al. [18]. Currently it is used to increase the power transmission capacity in China. Niaki et al. [19] proposed wavelet-transform and cable sheath's transient voltage-based techniques to detect faults on the DC zone including HVDC cable. In [20], multi-terminal meshed network fault detection technique has been presented by taking voltage amplitudes of two DC reactors connected to the same converter terminal and voltage polarities. Epameinondas et al. [21] has discussed various HVDC topologies and compared them in terms of faults on MTDC network. Irnawan et al. [22] proposed fault protection scheme for an HVDC transmission line that connects wind turbines to grid. Kerf et al. [23] presented wavelet-based fault protection scheme for a four terminal mesh DC system. Li et al. [24] presented that the external and internal faults can be identified via measuring the amplitudes of back propagation of the travelling wave.

Most of the works presented for fault identification rely on internal parameters of the system and are unable to precisely provide the fault location. Several studies, relying on external technologies like robotics, have also been presented for the monitoring of power transmission lines. In [25], the authors have provided a design of a robot having identical arms for power lines inspection. In [26], a wheel-based climbing robot is proposed for the inspection of power lines which requires a separate line attached to the power lines. Similarly, in [27], a novel design and analysis of a robot is provided for the inspection of power lines using jumper cable attached to the power lines. The authors of [28] have provided analysis of a low weight transmission line robot for easy movement on the transmission lines. Another explored field of robotics, for transmission line fault identification, is the use of UAVs as

TABLE 1. Comparison of existing works.

Research works	Techniques	Comparison attributes						
		UAV based technique	Imaging/vision-based fault detection capability	Reward based control input	Learning/reward based optimized control	Intelligent path traversing	Transmission line inspection	Synchronized UAV Group based traversing
[25]	Inspection robot	X	X	X	X	X	✓	X
[26]	Climbing robot design	X	X	X	X	X	✓	X
[27]	Arm-Robot based inspection	X	X	X	X	X	✓	X
[28]	POLIBOT design	X	X	X	X	X	✓	X
[29]	UAV quadrotor	✓	X	X	X	✓	✓	X
[30]	Co-operative UAVs	✓	✓	X	X	X	✓	✓
[31]	Rotation invariant insulator	X	✓	X	X	X	✓	X
[32]	UAV imaging based	✓	✓	X	✓	✓	✓	X
[33]	UAV imaging vision based	X	✓	X	✓	X	✓	X
[34]	UAV intelligent monitoring	✓	✓	X	✓	✓	✓	X
[35]	UAV imaging based	✓	✓	X	✓	X	✓	X
[43]	UAV bio inspired based	✓	X	X	X	✓	X	✓
[44]	UAV mission planner based	✓	X	X	X	✓	X	X
[50]	UAV reinforcement learning	✓	X	✓	✓	✓	X	✓
This work	UAV reward-based control	✓	X	✓	✓	✓	✓	✓

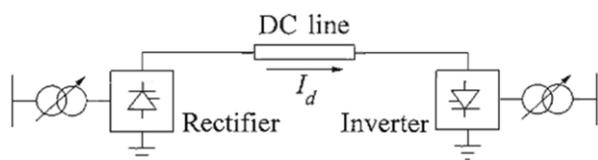


FIGURE 2. HVDC system basic structure.

shown in [29]. In [30], a cooperative communication mechanism is presented for multiple UAVs using imaging technology for power line inspection. In [31], [32], [34], and [35], multiple works have been presented for different kinds of fault identification using learning based approaches applied on aerial images taken through UAVs. In [36], the automatic UAV based inspection solution is presented where the inspection system is designed for large solar system to visualize the defects on PV module. The authors in [37], developed a target detection algorithm for UAVs used in inspection process for transmission lines. A real time fault detection model based on acceleration engine was designed for UAVs in [38]. In [39], the authors proposed a single and multi-fault detection for insulators used in power transmission lines. This method was implemented using UAVs for aerial images. In [40], a novel end-to-end network was designed for UAVs to inspect the railway system. In [41], the BSRT approach is proposed for transmission line inspection based on UAVs. The author in [42] proposed a modified model based on YOLO for detecting insulator faults used in transmission lines. The UAVs were used for taking aerial images. Most of the works presented to date using UAVs, concentrate on developing intelligent techniques for fault identification, like AI based imaging, but limited works have explored UAV based intelligent traversing strategies for fault location identification in

HVDC transmission lines, especially, using learning-based approaches. The works that involve intelligent movement of UAVs have used different application environments like reconnaissance of disaster areas [43], residential areas [44] and smart cities [45] etc. Similarly, the works involving MDP based learning strategies for autonomous movement of UAVs, presented in [45], [46], [47], [48], [49], [50], and [51], also focused on application areas other than the exploration of HVDC transmission lines especially for fault location identification. In this work, we focus on autonomous control strategy of UAVs working together to traverse the HVDC transmission lines for fault location identification. The concept of UAV based pole-to-pole transmission line inspection is shown in Figure 1. A comparison of the work presented in this paper with multiple existing techniques is given in Table 1.

As aforementioned, several works have been performed on fault detection using internal parameters like travelling wave method etc. and almost none of them can provide location of the fault. Further, use of the method depends on the specifications of the network or part of the network. This can impact on the cost of the network as multiple solutions will be required across the whole network. However, such solutions can be used to minimize the fault location identification time together with our strategy. Furthermore, the work can be enhanced by equipping UAVs with technologies like thermal and HD imaging. So, there are many potentials of this proposed strategy, however, these are not in the scope of this paper as its sole focus is to provide control strategy.

As mentioned earlier, a much faster and precise fault identification technique is needed that is not only able to identify line faults but can also provide accurate and efficient fault location. For such reasons, an efficient learning-based strategy is devised using UAVs to detect and locate the fault

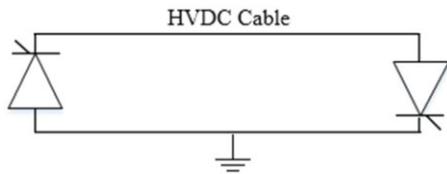


FIGURE 3. Back-to-back scheme.

efficiently, especially in areas where manned missions are not possible. In this context, limited work has been performed using a learning-based UAV approach for fault identification in HVDC transmission lines. The major contributing factors of this work are summarized as follows:

- An unmanned and efficient HVDC Transmission line inspection strategy using UAVs is presented for faulty portion identification in the transmission line.
- Provides basis for the AI/Machine-Learning based approaches for autonomous control of the UAVs to be used in Power sector applications.
- Provides Reward based UAV control strategy for learning networks,
 - where the reward equations are derived based on the control strategy and
 - Then using MATLAB, the reward tables are derived for each possible scenario based on the states of the Leader agent (Command Center) and the local agents (UAV agents).
- Intuition is provided using the application of HVDC power system transmission lines to explain and support the control strategy.
- Validation of the presented control strategy is provided through MATLAB based simulation results.

III. SYSTEM MODEL

This section provides details of the system model for fault location identification in HVDC transmission lines using a reward based multiagent system. The section starts with a brief description of the general HVDC transmission system with its link classifications. The latter subsections concentrate on the multiagent reward-based strategy.

A. HIGH VOLTAGE DC TRANSMISSION SYSTEM

HVDC transmission system is also termed as electrical super-highway because it can transmit a large electrical power over long distances with less electrical losses. These transmission lines can carry 100kV to 1500kV respectively [1]. In HVDC transmission system there is no skin effect hence, the conductor area is fully utilized. Also, the inductive and capacitive parameters do not apply any limit on the transmission capacity as well as length of the DC cable [2].

The basic structure of HVDC transmission system is shown in Figure 2. This system consists of AC side rectifier, transformers, DC side inverter and DC transmission lines or cables. A brief description of HVDC link classification is provided as follows.

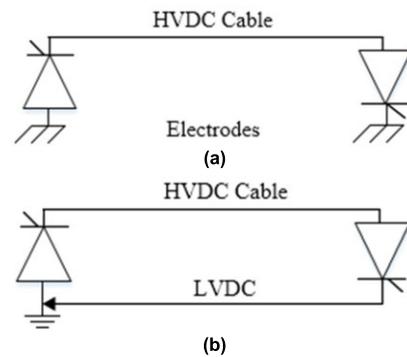


FIGURE 4. HVDC mono-polar transmission system: (a) with ground return path (b) with metallic return pat.

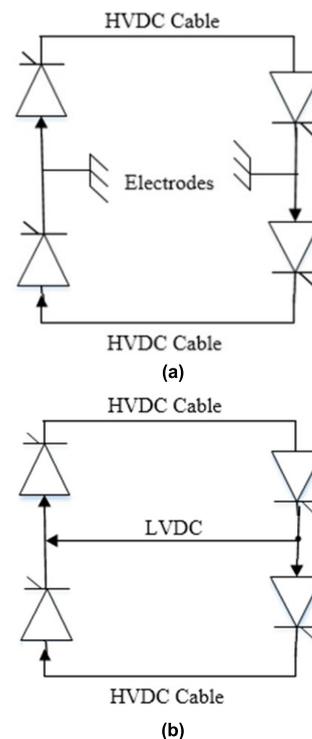


FIGURE 5. HVDC bi-polar transmission system (balanced operation): (a) with electrodes (b) with low voltage DC (LVDC).

Back-to-back: If the rectifier and inverter are placed on the same station, then this scheme is termed as back-to-back as shown in Figure 3. It is used in contiguous AC grids that are not synchronized and in meshed grids [52].

Mono-polar: Mono-polar HVDC transmission system is used for very long power transmission especially undersea cable transmission with its return path to ground or sea electrodes [52]. The system is shown in Figure 4(a). If the environmental constraints or existing infrastructure limit the use of electrodes, then in that case a metallic return path can be applied as shown in Figure 4(b).

Bipolar: This type of scheme is required if the single pole is not able to handle the desired transmission capacity. It can also be used when the rejection power of the load is

low or there is need of high energy availability [52]. If the transmission distance is short or infrastructure restrictions of using electrodes, then a metallic return path is used instead of electrodes. Both configurations, i.e., with electrodes and low voltage DC are shown in Figure 5(a) and (b) respectively.

Homo-polar: The Homo-polar system has conductors with negative polarity. Its advantage lies in its less installation cost [52]. The system is shown in Figure 6.

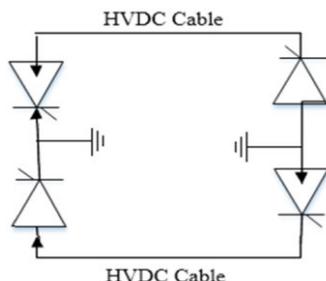


FIGURE 6. HVDC homo-polar transmission system.

As the research focuses on policy-based input for Learning approaches to derive autonomous UAV control for fault location identification in HVDC transmission lines, so it is important to mention that the proposed work is primarily for the faults related to the transmission lines, mainly:

- Line to Line Faults
- Line to Ground Faults
- Double line to ground Faults

After the brief description of HVDC transmission system with its link classifications, the description of multi-agent system (MAS), policy derivation and control strategy are provided in the next subsections.

B. MULTI-AGENT SYSTEM (MAS)

As aforementioned, multiple methods are used for fault detection in HVDC transmission system, however, using UAVs with implementation of reward policy-based movement makes the detection process more efficient. The process is safer due to its autonomous and unmanned nature. The technique implemented in this paper for UAV movement is hierarchical multi-agent system (MAS) where agents are considered as UAVs, that will monitor the HVDC transmission line. The distribution possibilities of UAVs in different areas of the HVDC transmission line through multiagent system is explained taking a simple example for better understanding shown in Figure 7. The example scenario in Figure 7 can be interpreted as a chunk of HVDC transmission line divided into five partition areas considering eight UAV agents controlled by a command center called the Leader agent. It is considered that there can be a maximum of two UAVs in one partition, if the number of UAVs is more than that; the leader agent must divide the UAV agents appropriately. To explain it further a total of four cases have been taken to check the efficiency of our system shown in Figure 7 and explained as follows:

- **Case 1:** First case starts with all the agents in the area 3. To make sure that the presented system works smoothly according to the defined conditions, the network will divide these UAV agents into all areas for achieving better efficiency such that Area 1 and 5 have 1 agent and other areas have 2 agents each shown in Figure 7(a).
- **Case 2:** Second possibility starts with 3 UAV agents in the area 2. Other areas have a satisfactory number of UAV agents; thus, our network takes 1 agent and moves it to area 1 which is close by and has 1 agent shown in Figure 7(b).
- **Case 3:** Third possible case starts with 3 UAV agents in the area 4. However, both close by areas 3 and 5 have satisfactory number of agents, thus our network takes 1 agent and moves it to area 1 but in doing so complete topology is changed as shown in Figure 7(c).
- **Case 4:** Forth case starts with 3 UAV agents in area 4 and 5. However, area 5 is the last one and area 3 has only one agent. This topology is complex and to get satisfactory results 1 agent is moved to area 1 and one agent is moved to area 2 as shown in Figure 7(d).

Based on the cases in above example, the general design and derivation of MAS for fault location identification in HVDC transmission lines is provided in the next sections.

C. IMPLEMENTATION OF MAS ON HVDC

Normally the HVDC transmission system is divided into different zones as shown in Figure 8. It is recommended to divide the system into adjacent overlapping zones for better protection as when a fault occurs in the overlapping area, the circuit breakers of both the zones will open. After a fault occurrence, efficient fault location identification is required in the transmission lines, right after the breaker operation, which is zone 2 shown in Figure 8 and is the focus of this work. Further, the proposed work offers independence of 24/7 monitoring or monitoring once the fault has occurred, depending on what is required, and which one is more convenient. Like the example in Figure 7, for deriving the policy and reward functions, the zone 2 in Figure 8 can be divided into different sections called partition areas, each having 'm' UAV agents, controlled by a command center called the leader agent. The UAV agents must be distributed efficiently, as shown in the example given in Figure 7, to perform searches with specified goals of locating Faults. Further details of how the transmission line can be divided into subsections and partitions for better coverage by UAV agents are discussed later in the results and discussion section.

The approach proposed in this work has been considered through modifications in the work performed in [53] implementing multi-agent systems. The system includes an implementation of leader (command and control center) and local agents (UAVs) approach. For this purpose, an effective implementation of Markov decision process (MDP) model is important for both leader and local agents given as follows:

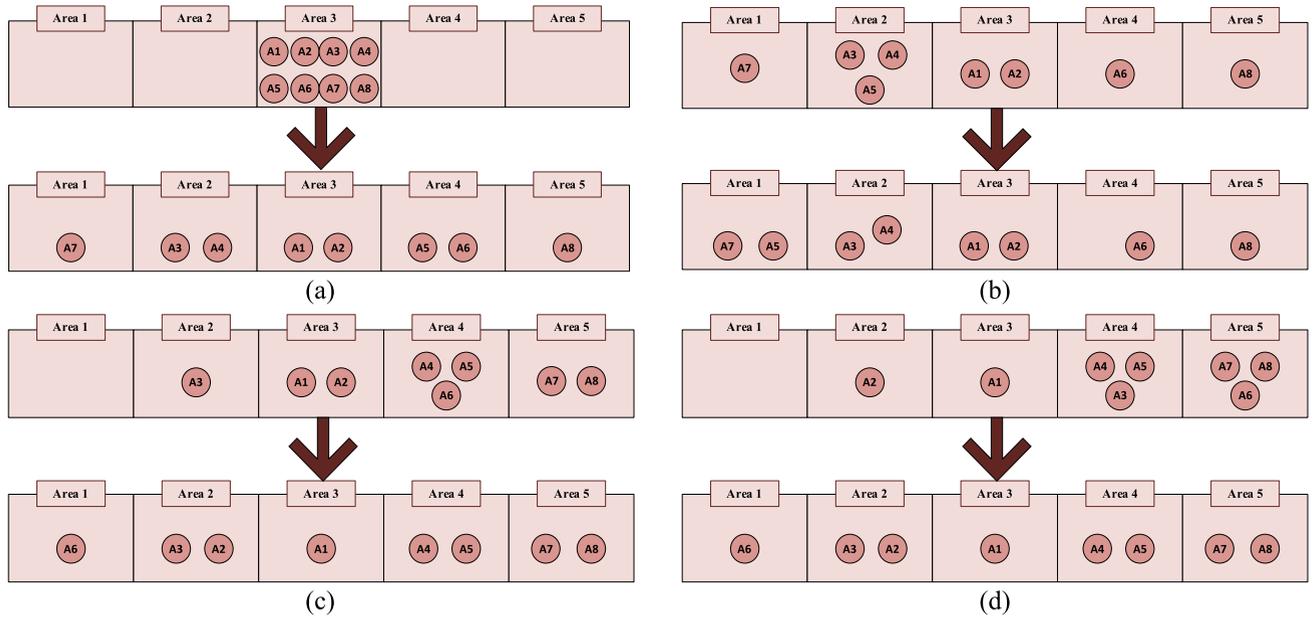


FIGURE 7. UAV multiagent system distribution possibilities, an example.

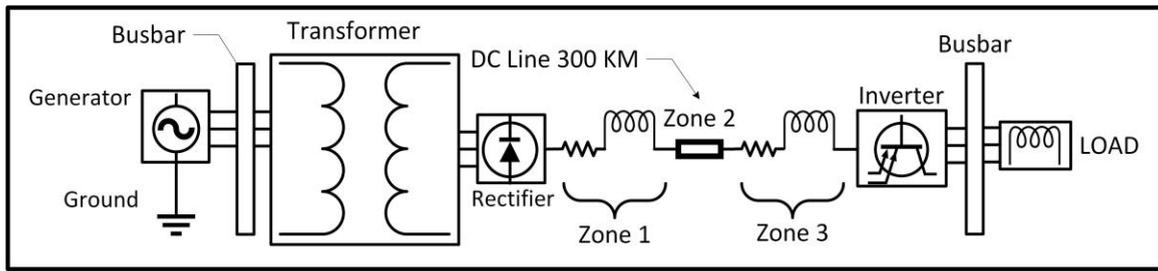


FIGURE 8. HVDC transmission system.

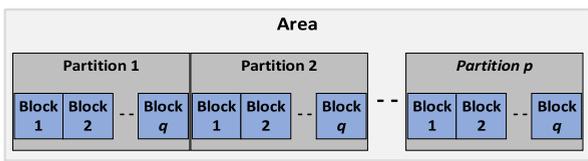


FIGURE 9. Partition and block distribution of HVDC transmission line for UAV agent movement.

1) MDP MODEL FOR LEADER AGENT

For leader agent (command and control), the MDP model includes the states, actions, and transition probabilities, which are defined as follows:

a: STATES

Four state variables are used in this work with each having its own function, i.e., a is the UAV agent state which depicts that either UAV agent is in on or off state, g depicts the goal completion, c gives the UAV agent partition (location) and b is the status of partition (faulty, not faulty, and dangerous).

The state equation comprising of these variables is given as:

$$\begin{aligned}
 S &= \{s_1, s_2, \dots, s_n\} \\
 s_i &= \{a_1, a_2, \dots, a_m, g, c_1, c_2, \dots, c_m, b_1, b_2, \dots, b_p\} \\
 i &\in \{1, 2, \dots, n\}, \\
 a_j &\in \{0, 1\}, \quad j \in \{1, 2, \dots, m\} \\
 b_k &\in \{0, 1, 2\}, \quad k \in \{1, 2, \dots, p\}, \\
 c_j &\in \{1, 2, \dots, p\}, \quad j \in \{1, 2, \dots, m\}, \\
 g_w &\in \{0, 1\}, \quad w \in \{1, 2, \dots, m\}, \quad (1)
 \end{aligned}$$

where, S is the total number of states, s_i is the state space, m is the number of UAV agents and p is the number of partitions shown in the Figure 9.

b: ACTIONS

There is a total of m agents and p partitions thus, the maximum number of actions is given by $m \times p = M$. An agent a can move to any partition k . The set of actions M is given as:

$$M = \{\mu_{1,1}, \mu_{1,2}, \dots, \mu_{2,1}, \mu_{2,2}, \dots, \mu_{m,p}, NOOP\} \quad (2)$$

where, $NOOP$ is the abbreviation for no operation.

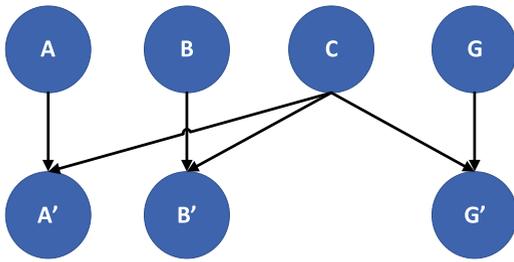


FIGURE 10. State transition conditions for leader agent.

c: REWARD FUNCTION

The reward function is the most important part in any MDP process. The states and reward function are defined according to the application, which in our case is exploration of HVDC transmission lines for fault location identification. The reward function shows how important any task is from the number of choices. In this model the reward function includes the following:

1. The completion of goals g (fault detected).
2. Avoiding congestion of agents in one partition $\alpha_{j,k}$.
3. Avoiding dangerous partitions $\beta_{j,k}$.

The equation for the reward function is given as

$$R(s) = (\alpha_{1,1} + \alpha_{1,2} + \dots + \alpha_{2,1} + \alpha_{2,2} + \dots + \alpha_{m,p}) \times [\lambda_1 e^{g_w} + \lambda_2 e^{\alpha_{1,1} + \alpha_{1,2} + \dots + \alpha_{2,1} + \alpha_{2,2} + \dots + \alpha_{m,p}} + \lambda_3 e^{-(\beta_{1,1} + \beta_{1,2} + \dots + \beta_{2,1} + \beta_{2,2} + \dots + \beta_{m,p})}]$$

$$\alpha_{j,k} = \begin{cases} 1 & \text{agents}_{j,k} \leq 2 \\ 0 & \text{otherwise,} \end{cases} \quad j \in \{1, 2, m\}$$

$$\beta_{j,k} = \begin{cases} 1 & \text{agents}_{j,k} > 1 \text{ and } b_k == 2 \\ 0 & \text{otherwise,} \end{cases} \quad i \in \{1, 2, 3\}$$

(3)

After the definition of the reward function, the second most important part is the state transition probability.

d: STATE TRANSITION PROBABILITY

Transition probability typically shows the probability of moving from one state to another state. It may vary according to any given state and thus it is very important to have probabilities that agree with the application. These probabilities are typically calculated from an external source and are utilized here for the purpose of research. The state transition conditions are shown in Figure 10, where the next states of the agent A' , partition status B' and goal G' are dependent on their respective previous states and the partition status state C .

Practically, there are multiple ways of defining the conditional probabilities involved in the problem. These can be available from statistical data or can be learned through real time environment. The methods of obtaining the probabilities do not lie in the scope of this work, instead it shows how to use them in a MDP model.

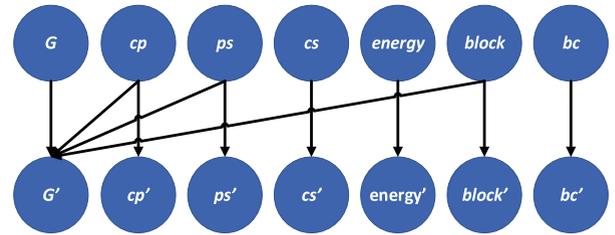


FIGURE 11. State transition conditions for member UAV agents.

2) MDP MODEL FOR LOCAL AGENT

a: STATES

For the local individual UAV agent, in the proposed model, seven variables are considered, i.e., partition status ps which depicts that a partition is normal (0), unknown (1) or faulty (2), companion present cp which depicts that an agent's current partition has a companion (0) or not (1), goal g (fault detected (1) or not (0)), battery status $energy$ which depicts three levels; empty (0), half (1) and full (2), companion location status cs which depicts the location of companion agents depending on the partition number, block location in a partition is given by $block$ and the block status is given by bc_q which gives that a block is explored (1), not explored (2) or dangerous (3) and q gives the total number of blocks in a partition as shown in Figure 9. Using all above variables, the generalized member state equation is given as:

$$S = \{s_1, s_2, \dots, s_n\}$$

$$s_i = \{ps, cp, g_w, energy, cs, block, bc_1, bc_2, \dots, bc_q\}$$

$$ps \in \{0, 1, 2\},$$

$$cs \in \{1, 2, 3 \dots p\},$$

$$energy \in \{0, 1, 2\},$$

$$cp \in \{1, 2, 3 \dots p\},$$

$$block \in \{1, 2, \dots, q\}$$

$$bc \in \{1, 2, 3\}$$

$$g_w \in \{0, 1\}, w \in \{1, 2, \dots, m\}$$

(4)

b: ACTIONS

Each UAV agent can take any of the five actions, given below, which are given to it by the leader commanding agent. These actions simply are related to the movement. Thus, the actions can be movement in any four directions or no movement at all. To sum it all, the following are the movements which are given by the leader agent.

$$M = \{MovLeft, MovRight, MovStraight, MovBack, stay\}$$

(5)

c: REWARD FUNCTION DEFINITION

Reward function is a main part of the MDP model and should be defined with caution. A problem with a single variable definition can alter the whole policy.

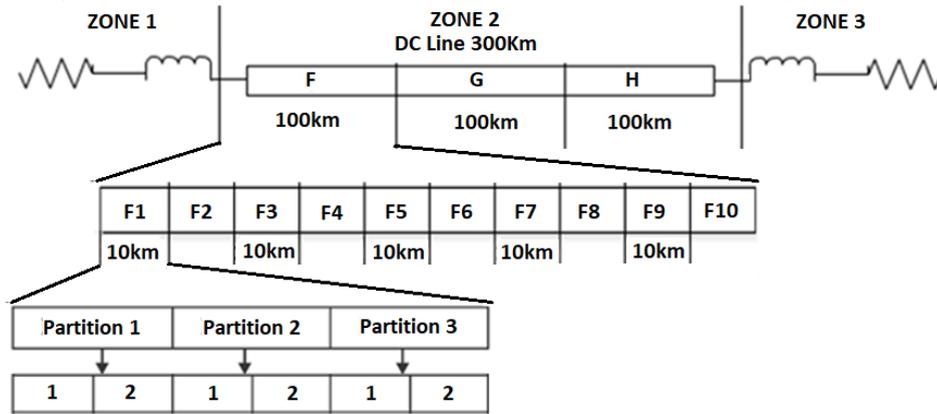


FIGURE 12. HVDC transmission line zone distribution experimentation scenario.

The reward function definition for member UAV agent is dependent upon the following parameters:

- 1) Maximum area coverage.
- 2) Congestion avoidance (at most two UAV agents allowed the same partition).
- 3) Status Check (UAV agent must send its active/inactive status report to the Leader agent).
- 4) Hazardous block avoidance.
- 5) Maximize battery utility.

$$\begin{aligned}
 R(s) &= \sum_{i=1}^5 \gamma_i r_i, \gamma_i \in [0, \infty) \\
 r_1 &= \sum_{l=1}^q (bc_l == 2) + \delta_1 (ps == 0) \\
 &\quad + \frac{\delta_1}{2} (ps == 1), \delta_1 \in (0, \infty) \\
 r_2 &= (cp \oplus cs) \\
 r_3 &= energy \\
 r_4 &= \sum_{l=1}^q (block == l) * (bc_l \neq 3) \\
 r_5 &= g_1 + g_2
 \end{aligned} \tag{6}$$

where, r_1 prompts the UAV agent for area exploration, as the block search area increases, the reward function also increases with it, γ_1 is the value of coefficient, for this scenario, and is taken to be a large positive number for increasing the exploration area. r_2 is the reward for congestion avoidance, γ_2 is also taken to be large positive number, as its purpose is also to maximize the reward function, the 3rd reward r_3 gives the battery utility status, defined by three levels, i.e., low, medium or full, γ_3 is taken about ten times larger than that of γ_2 for the missions where level of battery is most important, on the other hand, lower value of γ_3 may be used if exploration is more important (full value of energy will take reward function to maximum value), r_4 is for hazardous block avoidance to avoid UAV agent failure, i.e., the block may have high temperature, or unsustainable terrain (leader agent must make sure to not assign agents to such blocks), in the last fifth variable r_5 equation gives goal achievement.

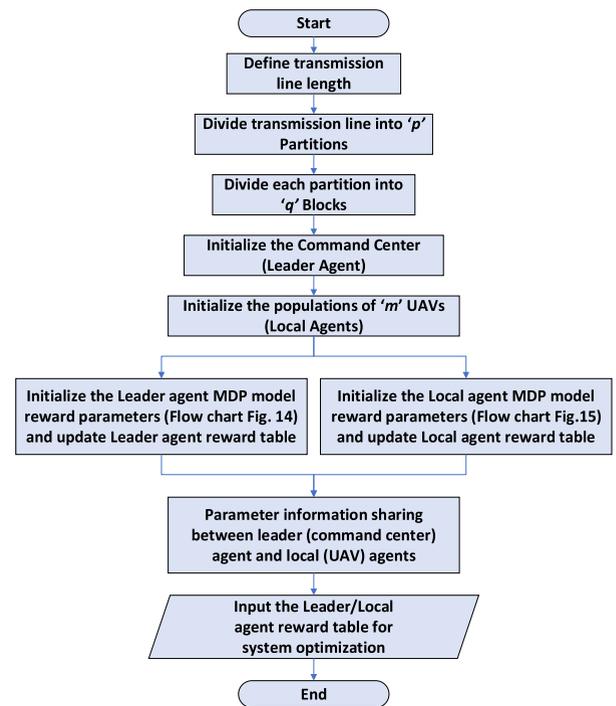


FIGURE 13. Flowchart for the presented reward-based strategy.

d: STATE TRANSITION PROBABILITY

As aforementioned, transition probability gives the change from one state to another. The state transition conditions of member agent state are shown in Figure 11. The state transition depends on seven variables of the member UAV agent, where the next state of goal G' depends upon the previous states of goal G , companion present cp , partition status ps , and block. The remaining all other states, in Figure 11, depend on their respective previous states.

IV. RESULTS AND DISCUSSION

To validate the proposed strategy, this section provides simulation-based outcomes and discussion on the results.

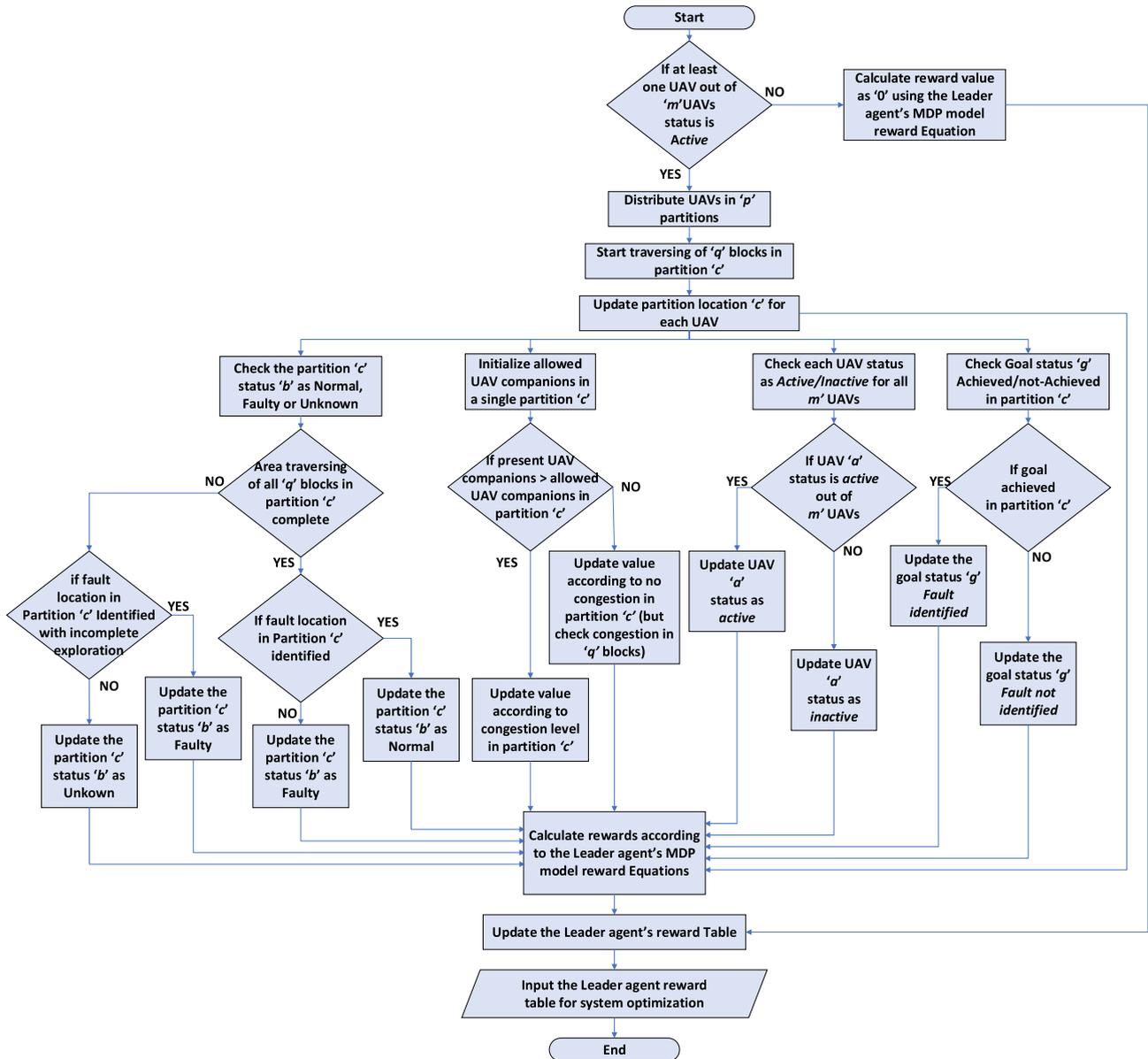


FIGURE 14. Reward based optimization flowchart for leader agent.

The simulation scenario is defined as follows: the strategy is applied on zone 2 (transmission lines) provided in Figure 8. The zone is divided into three sections F, G, and H, each considering 100 km transmission line as shown in Figure 12. Zone F of the transmission line is divided into ten small zones with each zone covering 10 kilometers. If we take one of the sub zones into consideration, it is again divided into three specific partitions, each of these partitions is divided into two equal blocks as shown in Figure 12. The purpose of these blocks is to foresee which part of the partition is hazardous and which partition is friendly enough for the agent to move there. As aforementioned, leader agent is considered as command-and-control center and local agents are UAVs.

The resulting policy is optimized through the value iteration algorithm [53] using the transition probabilities and reward functions defined in previous section. The value iteration equation which gives the optimal solution is given as:

$$V^* = \max_A \sum_s T(s, a, s') [R(s, a, s') + \gamma V^*(s')] \quad (7)$$

where, R is the reward function, T is the transition probability, s is the current state of MDP, s' is the previous state of MDP, and a is the action taken by the UAV agent. The iterative algorithm will keep running till it converges at an optimized result. This section is divided into two parts, i.e., the first part provides results and discussion on MDP based reward table derivations using MATLAB for understanding the behavior

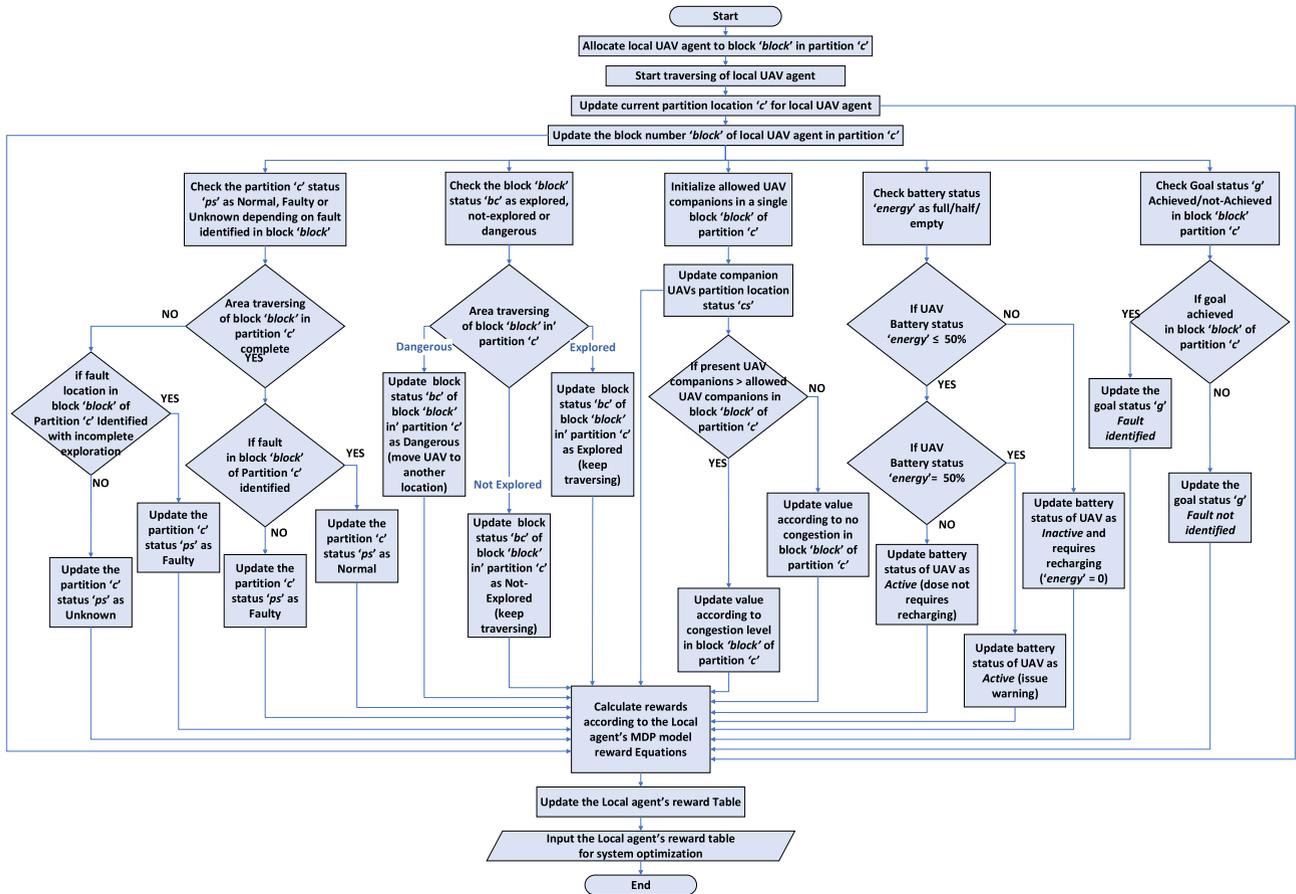


FIGURE 15. Reward based optimization flowchart for Local UAV agents.

of both leader agent and local UAV agents, the second part provides simulation results, also derived using MATLAB, showing the behavior of UAVs for fault location identification in transmission lines based on the presented reward-based strategy where the time taken to reach the fault location is also calculated depending on distance and velocity of UAVs. The overall system flowchart is shown in Figure 13 and the flow charts for reward-based system optimization for both leader and local UAV agents are shown in Figures 14 and 15.

A. MDP BASED LEADER AND LOCAL AGENT REWARD RESULTS AND DISCUSSION

This section provides the leader and local UAV agent reward tables and their behavior understanding based on cases taken from these tables. For both leader and local UAV agent cases, let’s assume that UAVs are trying to explore through an HVDC transmission line during normal condition with a goal to search for faults. Unknown constraints are if any transmission line has fault it is declared as faulty. If the transmission line does not contain fault, it is considered as normal condition and if the part of transmission line is not explored yet it is considered unknown.

TABLE 2. Initial conditions of leader agent.

Status of Partition 1	Status of Partition 2	Status of Partition 3	Status of Goal
UNKNOWN	UNKNOWN	UNKNOWN	UNACHIEVED

1) LEADER AGENT BEHAVIOR

Two cases are discussed for better understanding of Leader agent’s reward function behavior, i.e., 1) with uneven distribution of UAVs and 2) with even distribution of UAVs. MDP is used to derive the states of leader agent and the reward for each state of leader agent is calculated to formulate a complete policy. The values of leader agent reward are dependent on the variables discussed in Section III. The initial conditions for the variables associated with leader agent are depicted in Table 2. The MDP policy of the controlling leader agent is given as follows:

a: MDP POLICY

The MDP policy of leader agent is dependent on multiple state variables defined in section III of this paper. The targeted UAV movements are dependent on these state variables and

TABLE 3. MDP policy: states definition.

Leader Agent Variables	State Definition		
UAV Agent Status	Inactive		Active
	0		1
Partition Status	Normal	Unknown	Faulty
	0	1	2
Goal	Achieved		Not Achieved
	1		0

TABLE 4. Behavior of leader agent reward function with odd no. of UAV agents.

Cases	UAV Agent Status	Goal Status	Partition to which agent is assigned	Partition Status	Reward
1	00000	00	11111	111	0
2	00001	00	11111	111	840.01
3	00011	00	11111	111	1.6803x10 ³
4	00111	00	11111	111	2.5205x10 ³
5	01111	00	11111	111	3.3606x10 ³
6	11111	00	11111	111	4.2008x10 ³
7	11111	01	11111	111	4.6303x10 ³
8	11111	10	11111	111	4.6303x10 ³
9	11111	11	11111	111	5.7980x10 ³
10	11111	11	22222	111	5.7980x10 ³
11	11111	11	33333	111	5.7980x10 ³
12	11111	11	12312	222	1.2049x10 ⁴
13	11111	11	12312	111	1.2146x10 ⁴
14	11111	11	12312	101	1.2157x10 ⁴
15	11111	11	12312	000	1.2265x10 ⁴

TABLE 5. Behavior of leader agent reward function with even no. of UAV agents.

Case	Agent State	Goal Status	Partition to which agent is assigned	Partition Status	Reward
1	111111	11	112233	000	1.4718x10 ⁴
2	111111	11	112233	100	1.4623x10 ⁴
3	111111	11	112233	110	1.4588x10 ⁴
4	111111	11	112233	111	1.4576x10 ⁴
5	111111	11	112233	222	1.4433x10 ⁴
6	111111	11	111233	222	6.8408x10 ³
7	111111	11	112223	222	6.8408x10 ³
8	111111	11	333333	222	6.9105x10 ³
9	111111	00	112233	000	1.2801x10 ⁴
10	111111	00	112222	000	5.1834x10 ³

b: CASE A: UNEVEN DISTRIBUTION OF UAVs AND LEADER AGENT BEHAVIOR

To simulate and discuss this case, five UAV agents, two goals for each agent and three partitions are considered as shown in Figure 16. The behavior of the leader agent’s reward function in this case is shown in Table 4. This table is derived by simulating the leader agent’s reward function in Section III using MATLAB. The table shows reward changes according to the changing value of each state (values defined in Table 3). The simulation table of leader agents reward behavior is too large to be presented here so, fifteen different states, given in Table 4, are chosen from the actual table, derived using MATLAB, to showcase the behavior of leader agent. The reward is directly dependent on the agent’s state. If the state of any agent is inactive, then those agents cannot participate in computation of reward values. In case 1, all agents are inactive, so the reward function returns 0, as our reward function is UAV agent dependent, and it becomes inactive if all agents are inactive. The reward increases with the changing state of each UAV agent from inactive to active. This can be seen from cases 2 to 6 where one agent is activated in the second case, two agents in case three, until the activation of all agents in case 6 shows the increasing value of reward in the first six cases. The next two cases, i.e., case 7 and 8, depict the behavior according to goal completion, where it can be seen that goal completion also increases the reward values. In the ninth state both goals are completed therefore the reward value increases further. In all the cases from 1 to 11 all the UAV agents are in the same partition which is why from case 9 to 11 the behavior of reward function does not change when the number of partitions is changed as all UAV agents are still in same partition. The same partition case in this work is not recommended to achieve better search results and this kind of problem is considered as congestion. Which is why in the next states due to different partitions assigned to UAV agents the reward increases further showing that congestion is reduced. In case 12 the congestion is removed with two UAVs in the first partition, two UAVs in second partition and one in third.

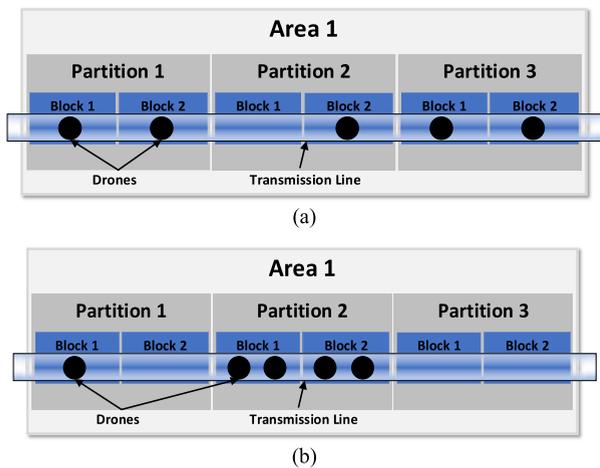


FIGURE 16. UAV distribution example with odd no. of UAV agents.

their definition is provided in Table 3. The behavior analysis of leader agent’s reward function is shown in Tables 4 and 5 with respect to odd and even distribution of UAV agents respectively. The even and odd cases are discussed as follows.

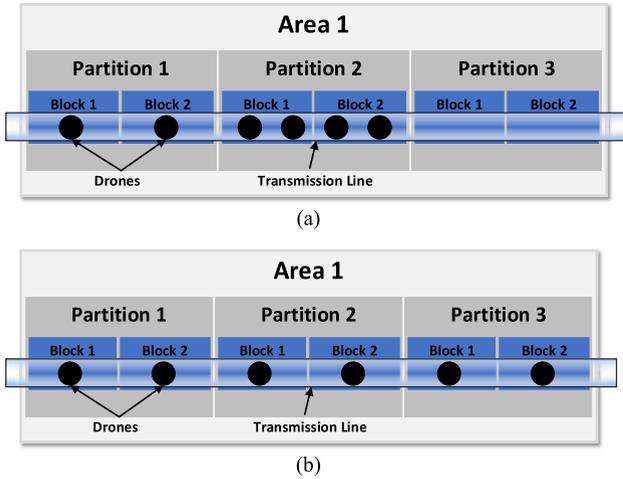


FIGURE 17. UAV distribution example with even no. of UAV agents.

In this case the state of all partitions is assigned dangerous value. In the preceding case 13, the state of all three partitions is changed to unknown instead of dangerous which further increases the reward. For the last chosen case 15, the partition status of all partitions is changed to known increasing the reward value further. Moreover, in the same case all agents are active, goals are completed, and no congestion is found in this case, therefore it depicts the maximum reward. Next, we will discuss the advantage of having an even number of drone agents.

c: CASE B: EVEN DISTRIBUTION OF UAVs AND LEADER AGENT BEHAVIOR

Table 5 depicts the representation of Leader agent’s reward function behavior with six agents with two goals each and three partitions. Each of these partitions includes two blocks which will be further explained in the next section for member UAV agent reward function. In this table the major difference occurs when we have Case 1 of Table 5 showing all agents in active state, both goals completed, no congestion, i.e., having two UAVs in each partition, one for each block, shown in Figure 17 and all partition status are known. Therefore, the reward is maximum for this case as the agent assignment is equal for all partitions. The major significance of having an even number of UAVs is the maximization of Leader agent’s reward function which is visible if we compare the results of case 15 in Table 4 and case 1 in Table 5 showing that case 1 in Table 5 has higher value. The remaining Table 5 shows similar behavior of reward as in Table 4. As mentioned earlier, it is not possible to show the complete reward tables, so their graphical results are shown in Figures 18 and 19 for Table 4 and Table 5 respectively. A similar trend can be seen in these results, i.e., the reward values in terms of even UAV agents, shown in Figure 19, is higher as compared to the values for odd number of UAV agents, shown in Figure 18.

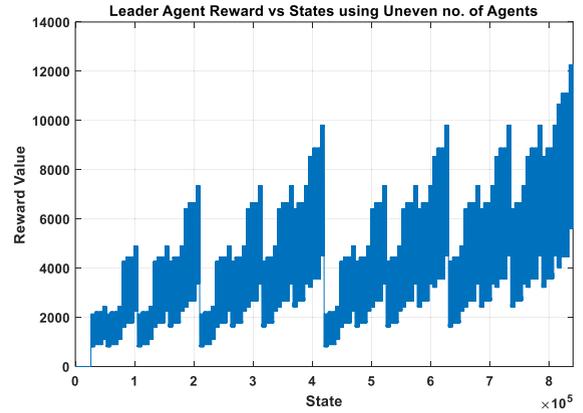


FIGURE 18. Reward value vs leader agent states with odd distribution of UAV agents.

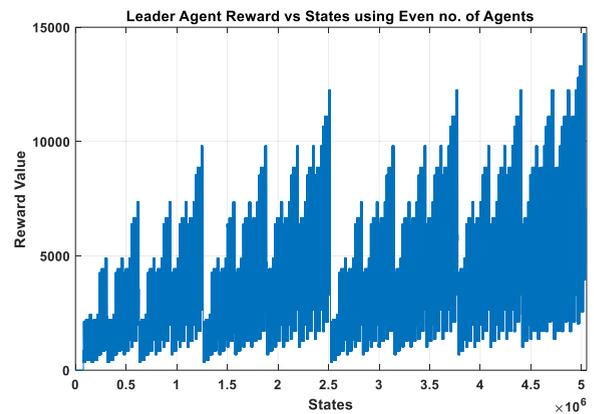


FIGURE 19. Reward value vs leader agent states with even distribution of UAV agents.

TABLE 6. Initial status for member UAV agent.

Partition Status (PS)	Current Partition (CP)	Companion Location Status (CS)	Status of Goal
UNEXPLORED	PARTITION 1	PARTITION 1	UNACHIEVED

d: MOVEMENT FOR CONGESTION

In the cases where congestion is present in Table 4 and V, the movement of UAV, based on the reward, is such that it moves to maximize the reward and coverage. One such case is depicted in Figure 17(a), where partition 2 has four UAVs and the reward is minimum in this case. The reward-based policy takes two UAVs to partition 3 to complete uniform UAV assignment across all partitions and maximizes the reward by assigning one UAV to each block (Figure 17(b)).

2) MEMBER AGENT BEHAVIOR

As in Leader agent, MDP is used to detect the states of UAV member agent so the reward for each state of UAV is calculated to formulate a complete policy. The member

TABLE 7. MDP policy: states definition.

Local Agent Variables	State Definition		
Partition Status	Normal	Unknown	Faulty
	0	1	2
Current Partition	Partition 1	Partition 2	Partition 3
	1	2	3
Goal	Achieved		Not Achieved
	1		0
Battery Energy	Empty	Half	Full
	0	1	2
Companion Location status	Partition 1	Partition 2	Partition 3
	1	2	3
Block Number	Block 1		Block 2
	1		2
Block Status	Unexplored	Explored	Dangerous
	1	2	3

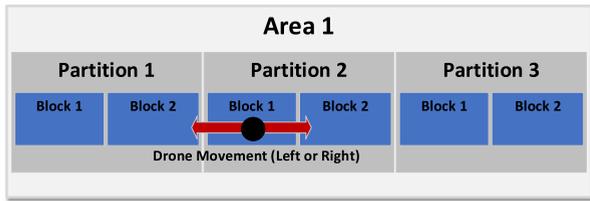


FIGURE 20. UAV agent movement.

UAV agent’s reward values are dependent on the variables discussed in the local agent’s MPD policy given in section III. The initial conditions for the variables associated with UAVs are provided in Table 6. The policy-based block to block movement of a member UAV agent is shown in Figure 20 (left or right).

a: MDP POLICY

The MDP policy of member agent is dependent on several factors with their specific state variable values. The member UAV movements are dependent on these state values. The state variables alongside their state definition are shown in Table 7. The behavior of member UAV agent’s reward function, based on state definition in Table 7, is shown in Table 8.

In Table 8, 23 cases are considered from the original Table simulated in MATLAB, but like the leader agents reward tables, this table also cannot be shown here due to being too large. However, the complete table values are shown graphically in the result shown in Figure 21. The cases 1 and 2 for local agent depicts partition status to be normal, current partition of the drone is taken as 1, goals are completed, energy is full and block status is dangerous. The difference between both cases is companion status. In case 1 where companion is in partition 3 there is no congestion therefore the reward is bigger than case 2 where companion is present in the same partition as the member UAV itself. Cases 3 to 5

TABLE 8. Reward behavior of local UAV agent.

Case	Parti tion Statu s	Curr ent Parti tion	Goal Statu s	Ener gy	Com pani on Statu s	Bloc k	Bloc k Statu s	Rew ard Func tion
1	0	1	11	2	3	2	3	9300
2	0	1	11	2	1	2	3	9200
3	0	1	00	2	3	2	3	8100
4	0	1	00	2	3	2	2	8600
5	0	1	00	2	3	2	1	8500
6	0	3	00	2	3	2	1	8400
7	0	2	00	2	2	2	1	8400
8	0	2	01	2	2	1	1	9000
9	0	2	11	2	2	1	1	9600
10	0	2	11	1	2	2	1	8600
11	0	2	11	0	2	2	1	7600
12	1	2	11	0	2	2	1	4600
13	1	2	11	2	2	2	1	6600
14	1	2	00	0	2	2	1	3400
15	2	2	11	0	2	2	1	1600
16	2	2	11	0	2	2	2	1700
17	2	2	11	1	1	2	3	2300
18	2	2	11	2	1	2	3	3300
19	2	2	11	0	1	2	3	1300
20	2	2	11	0	2	2	3	1200
21	2	2	00	0	2	2	1	400
22	2	2	00	0	2	2	2	500
23	2	2	00	0	2	2	3	100

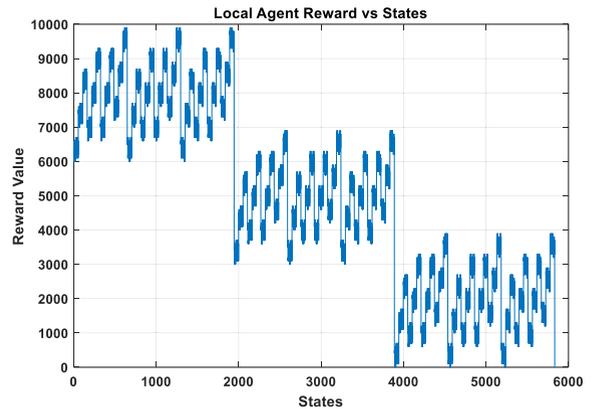


FIGURE 21. Reward vs state for local agent (UAV).

depict the importance of block status keeping all other variables constant. In these cases, block is dangerous in case 3 thus the reward is minimum in this case. For case 5 the block is unexplored therefore the reward function performs better than case 3. Case 4 performs best in this scenario because block status is explored in this case. Cases 6 and 7 depict the fact that the reward function remains the same in case of congestion. In case 6 congestion is in partition 3 and in case 7 congestion is in partition 2. Cases 8 and 9 depict the variance caused due to goal completion with all other variables being constant. In case 9, where both goals are completed, it performs better than case 8, where one goal is completed. Cases 10 and 11 depict the difference in reward created due to battery level of the drone. Reward is more for case 10 when battery level is half as compared to case 11 when battery level is empty. Cases 12 and 13 also depict the battery level but in this scenario battery level

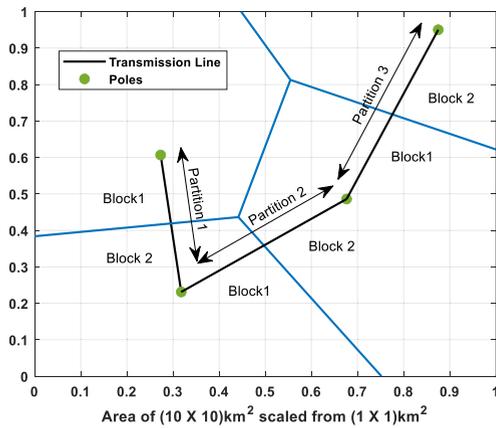


FIGURE 22. Simulation environment: Area (10 × 10) km² scaled from (1 × 1) km², 3 Pole-to-pole partitions (4 poles), 6 blocks (two blocks in each partition).

shifts from completely empty to full, therefore, the reward change is more than the change between cases 10 and 11. Cases 14 and 15 depict the difference created due to goal status change from unachieved to achieved. Case 15, where a goal is achieved, has a better reward value than 14, where both goals are unachieved. Case 16 takes the same values of all variables as in case 15 other than the block status which is changed from explored to unexplored resulting in decrease of reward. For cases 17 to 19 the response of battery is checked by keeping the block status to dangerous. The reward is maximum for the case 18 where battery is at maximum level and minimum for the case 19 where battery is empty. Case 20 takes the same variable values as in case 19 with the presence of congestion when both drones are present in partition 2. In this scenario the reward function decreases. Cases 21, 22 and 23 depict the minimum values of reward function with congestion present, goals are unachieved, partition is dangerous, energy is drained to completely empty. The only difference between these cases is the value of block status. Reward is minimum for test case 23 where the block status is dangerous and maximum for case 22 where block status is explored. The behavior of the complete reward table of a member UAV agent is shown in result given in Figure 21. The policy based optimized movement of UAV member agent will be from block to block based on all the factors explained above and is detailed in the second part of this section provided next.

B. SIMULATION RESULTS SHOWING UAV MOVEMENT AND FAULT LOCATION IDENTIFICATION

This section provides the MATLAB based simulation results for showing the movement of UAVs and fault location identification in transmission lines after the system optimization through the reward tables derived in the previous section. For showing the behavior of UAV agents, four poles are considered connecting three transmission lines. A total area of 1 × 1 km² is considered scaled over 10 × 10 km².

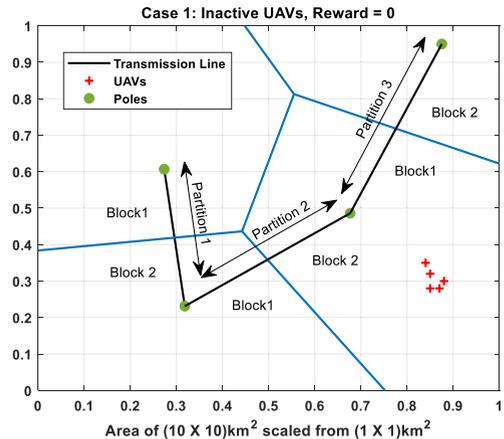


FIGURE 23. Case 1 of Table 4: All UAVs at initial position with inactive status.

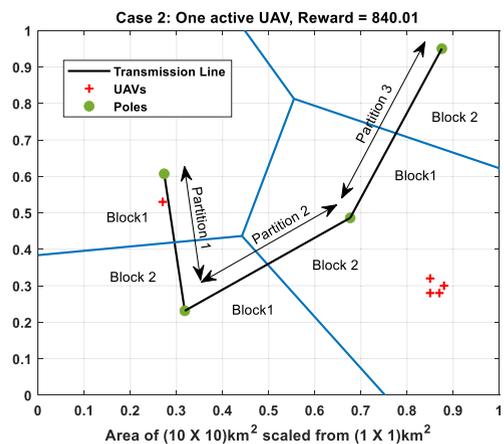


FIGURE 24. Case 2 of Table 4: One UAV active and assigned to partition 1's block 1.

Three pole-to-pole transmission lines are considered. To create more realistic terrain, Voronoi topology is considered for uniform distribution of partitions and blocks in the transmission line. Each pole-to-pole transmission line presents a single partition, and each partition is divided into two blocks using the Voronoi area division line. The environment is shown in Figure 22. To keep uniformity in explanation of the results related to previous sections, 5 UAV agents are considered and for congestion avoidance two UAVs are assigned to one partition with each block having one UAV. The results for this section are also provided in two parts: to represent different scenarios from the reward Table 4 and to calculate UAV's time to reach the fault location by considering a single partition with two blocks using distance equation given as:

$$s = v \times t$$

$$t = \frac{s}{v} \tag{8}$$

where, v is the velocity of the UAV and s is the UAV's distance to fault.

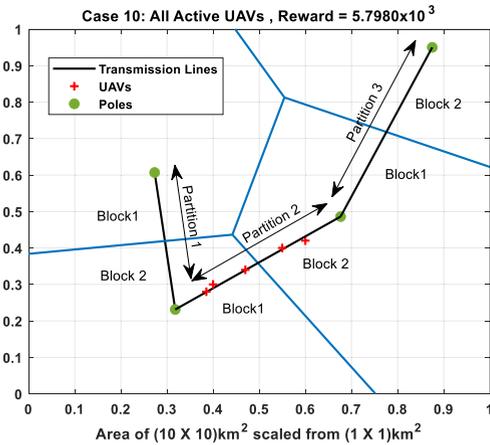


FIGURE 25. Case 10 of Table 4: All UAV active and assigned to partition 2 (congestion behavior).

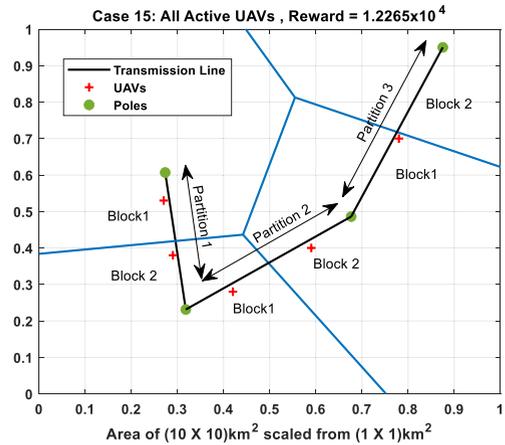


FIGURE 27. Case 15 of Table 4: All UAV active with no congestion and no faults.

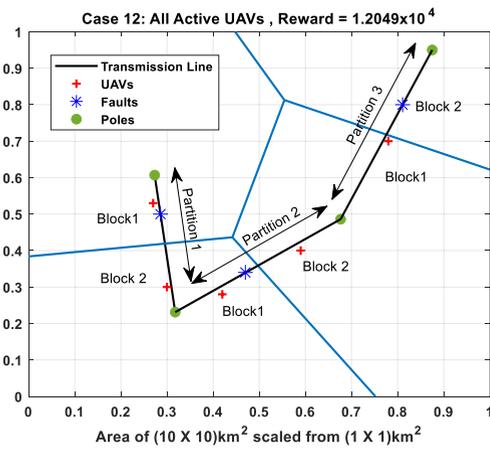


FIGURE 26. Case 12 of Table 4: All UAV active with no congestion and one fault introduced in each partition.

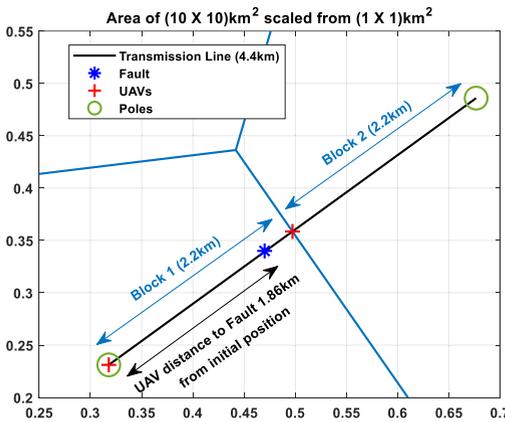


FIGURE 28. UAV transmission line traversing: fault location = block 1, UAV distance to fault = 1.86km, UAV velocity $v = 120\text{km/h}$ and 60km/h .

1) REWARD TABLE BEHAVIOUR DEPICTION

The results shown in Figures 23 to 26 provide the reward table behavior depiction by choosing different scenarios from Table 4. Figure 23 presents Case 1 from Table 4. It shows that when all UAVs are at an initial position and are not assigned to any partition, they are considered inactive and the reward in this case is calculated to be 0. The result in Figure 24 shows Case 2 of Table 4. It is shown that one UAV is assigned to block 1 of partition 1 and that the UAV has reached the assigned partition, with the status of two other partitions as unknown. Here, one thing to understand is that all partitions are healthy, and no fault has yet been introduced. The reward in this case is 840. Figure 25 shows Case 10 of the reward Table 4. It is shown that all the UAVs are now assigned to partition 2 with three UAVs in block 1 and two UAVs in block 2. This case depicts the congestion behavior of the reward-based control strategy and in this scenario the reward is 5.7980×10^3 . Yet no fault has been introduced to any partition. The result in Figure 26 shows Case 12 from the reward Table 4. In this case, two depictions are shown, firstly it is

shown that once the congestion avoidance becomes active, the UAVs, that were in partition 2 in previous result, readjust themselves to maximize the reward value. It is visible in this result that 2 UAVs have moved to partition 1, 2 remained in partition 2 and 1 has moved to partition 3 in a way that each block in each partition has one UAV. In partition 3, there is only one UAV in block 1 because of the odd number of UAVs. Secondly, in this case a single fault is introduced in each partition and the UAVs have completed the goal of identifying the fault locations, due to which the reward value is now 1.2049×10^4 . The result in Figure 27 is shown for Case 15 in the reward Table 4. After successful identification of fault location in previous result, this result is produced by removing the faults from each partition. In this result, each UAV reports to the leader agent that the partition status is now normal, hence achieving its goal. In this case, it can be seen that the reward value has now increased to 1.2265×10^4 . The significance of these results is to show that as soon as the situation improves, the system maximizes the reward value for better optimization.

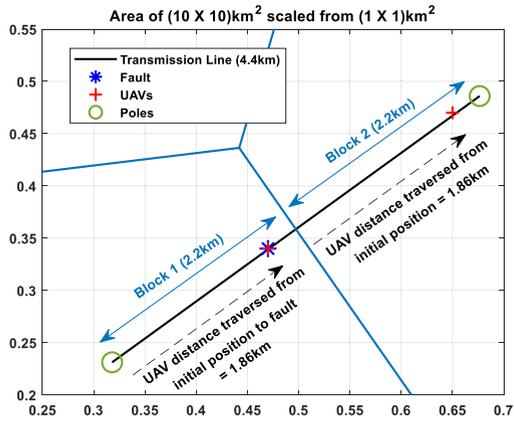


FIGURE 29. UAV transmission line traversing: fault location = block 1, UAV distance to fault = 0km (fault location reached), UAV velocity $v = 120\text{km/h}$ and 60km/h .

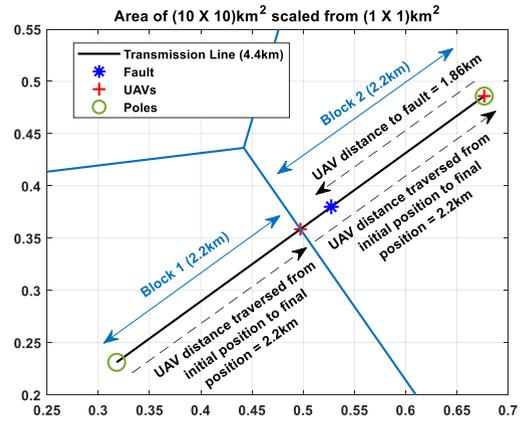


FIGURE 31. UAV transmission line traversing: fault location = block 2, UAV distance to fault = 1.86km , UAV velocity $v = 120\text{km/h}$ and 60km/h .

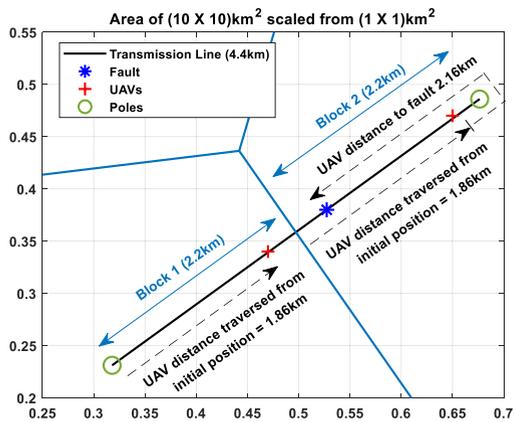


FIGURE 30. UAV transmission line traversing: fault location = block 2, UAV distance to fault = 2.16km , UAV velocity $v = 120\text{km/h}$ and 60km/h .

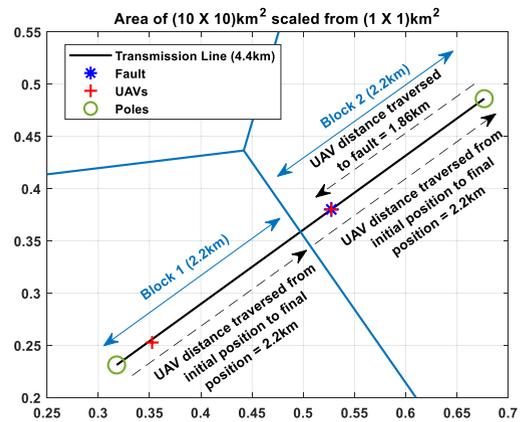


FIGURE 32. UAV transmission line traversing: fault location = block 2, UAV distance to fault = 0km (fault location reached), UAV velocity $v = 120\text{km/h}$ and 60km/h .

2) FAULT LOCATION IDENTIFICATION TIME CALCULATION

In this subsection, the results for UAV's time to locate the fault, while traversing the transmission line, are shown. The results are derived by choosing partition 2 from the environment shown in Figure 22. Two UAVs are considered, one assigned to each block. To depict the time to reach the fault, equation (8) is used. Distance s in (8) is calculated by substituting x and y coordinates in the line equation given as:

$$s = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (9)$$

In Figure 28, it is shown that each block is of 2.2km with UAVs at the initial position. A fault is introduced in block 1 of the partition and the distance to fault from the UAV, assigned to block 1, is 1.86km . In Figure 29, it is shown that the UAV from block 1 has reached the fault location by traversing the distance of 1.86km . The calculation of time taken to reach the fault is dependent on the velocity of the UAV which varies depending on the type of UAV used. To show as an example, in this result, it is assumed that the velocity v of the UAVs in each block is 120km/h , resulting in calculated time to reach the fault = 0.93mins given from equation (8).

If we change the value of velocity v to 60km/h , then the time to reach the fault changes to 1.86mins . In the result in Figure 30, now the fault is introduced in block 2 at an instant when UAV has traversed the distance of 1.86km from its initial position. In this case it is shown that now the distance to reach the fault is 2.16km because the UAV must traverse to the final position, as shown in Figure 31, and then it will start traversing backwards towards the initial position. While traversing backwards it is shown in Figure 32 that the UAV in block 2 reaches the fault after covering the distance of 2.16km calculated from the position when the fault was introduced. In this case, considering $v = 120\text{km/h}$ the time to reach the fault, calculated from (8) is 1.08mins and in the case of $v = 60\text{km/h}$ the calculated time to reach the fault increases to 2.16mins . These results are deduced for depiction of time calculation considering ideal scenario whereas in real-time environment, factors like wind and height etc. may affect the results. However, real-time testing is not in the scope of this work and will be considered as part of future extensions, because the focus of this paper is to present and validate the reward-based control strategy for the fault location identification in HVDC transmission lines.

V. CONCLUSION

An intelligent control strategy is presented in this paper using UAVs for fault location identification in HVDC transmission lines. The technique is developed for autonomous movement of UAV agents to explore HVDC transmission lines based on MDP Reward based policy. Reward equations are derived for both Leader (command center) and Local (UAVs) agents. Complete reward tables are derived, through MATLAB, using the reward functions for every possible scenario. The graphical results, derived using MATLAB, provide the reward behavior for both commanding leader agent and the member UAV agents. The resulting policy is optimized through the value iteration algorithm using the transition probabilities and defined reward functions behaviors. The validation of system optimization through the MDP based reward tables is provided using MATLAB based simulation results. The results are provided by creating different scenarios in MATLAB using the reward Table 4. The results show the movement of UAVs and reward behavior for scenarios like congestion avoidance and fault occurrence etc. Furthermore, based on simulation results, calculation for time to reach the fault location is also derived using the fault distance and UAV velocity by considering two velocities, i.e., 120km/h and 60km/h. The results are derived considering the ideal scenario whereas in real-time environment, factors like wind and height etc. may affect the results. As the work focuses on presenting the reward-based control strategy for the fault location identification in HVDC transmission lines, real-time testing is not in the scope of this work and will be considered as part of future extensions.

The presented work provides a baseline for the future work using learning techniques, such as machine learning/deep learning, for autonomous control and movement of UAVs in the power field applications. Such techniques are capable of learning and evaluating completely on their own, so enhanced traversing strategies can be developed by integrating vision-based or sensor-based technologies etc.

ACKNOWLEDGMENT

This work was supported and funded by wireless world research forum (WWRF). The authors would like to thank their colleagues at WWRF for their valuable contributions.

REFERENCES

- [1] W. Long and S. Nilsson, "HVDC transmission: Yesterday and today," *IEEE Power Energy Mag.*, vol. 5, no. 2, pp. 22–31, Mar. 2007.
- [2] P. Bresesti, W. L. Kling, R. L. Hendriks, and R. Vailati, "HVDC connection of offshore wind farms to the transmission system," *IEEE Trans. Energy Convers.*, vol. 22, no. 1, pp. 37–43, Mar. 2007.
- [3] V. Akhmatov, M. Callavik, C. Franck, S. Rye, T. Ahndorf, M. Bucher, H. Müller, F. Schettler, and R. Wiget, "Technical guidelines and pre-standardization work for first HVDC grids," *IEEE Trans. Power Del.*, vol. 9, no. 1, pp. 327–335, Feb. 2013.
- [4] M. A. Dumitrescu, *Network Protection and Automation Guide*. Saint-Ouen-sur-Seine, France: Alstom Grid, 2011. Accessed: Jan. 5, 2022. [Online]. Available: <https://rpa.energy.mn/wp-content/uploads/2016/07/network-protection-and-automation-guide-book.pdf>
- [5] L. Tang and B. T. Ooi, "Locating and isolating DC faults in multi-terminal DC systems," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1877–1884, Jul. 2007.
- [6] R. Dantas, J. Liang, C. E. Ugalde-Loo, A. Adamczyk, C. Barker, and R. Whitehouse, "Progressive fault isolation and grid restoration strategy for MTDC networks," *IEEE Trans. Power Del.*, vol. 33, no. 2, Apr. 2018, Art. no. 909918.
- [7] L. van der Sluis, *Transients in Power Systems*. Hoboken, NJ, USA: Wiley, 2001.
- [8] S. Luo, X. Dong, S. Shi, and B. Wang, "A directional protection scheme for HVDC transmission lines based on reactive energy," *IEEE Trans. Power Del.*, vol. 31, no. 2, pp. 559–567, Apr. 2016.
- [9] J. Sneath and A. D. Rajapakse, "Fault detection and interruption in an earthed HVDC grid using ROCOV and hybrid DC breakers," *IEEE Trans. Power Del.*, vol. 31, no. 3, pp. 973–981, Jun. 2016.
- [10] X. Zheng, T. Nengling, Y. Guangliang, and D. Haoyin, "A transient protection scheme for HVDC transmission line," *IEEE Trans. Power Del.*, vol. 27, no. 2, pp. 718–724, Apr. 2012.
- [11] J. Suonan, S. Gao, G. Song, Z. Jiao, and X. Kang, "A novel fault-location method for HVDC transmission lines," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 1203–1209, Apr. 2010.
- [12] J. Liu, J. Duan, H. Lu, and Y. Sun, "Fault location method based on EEMD and traveling-wave speed characteristics for HVDC transmission lines," *J. Power Energy Eng.*, vol. 3, no. 4, pp. 106–113, 2015.
- [13] N. M. Haleem and A. D. Rajapakse, "Local measurement based ultra-fast directional ROCOV scheme for protecting bi-pole HVDC grids with a metallic return conductor," *Int. J. Electr. Power Energy Syst.*, vol. 98, pp. 323–330, Jun. 2018.
- [14] F. Ünal and S. Ekici, "A fault location technique for HVDC transmission lines using extreme learning machines," in *Proc. 5th Int. Istanbul Smart Grid Cities Congr. Fair (ICSG)*, Apr. 2017, pp. 125–129.
- [15] S.-P. Gao, X. Chu, Q.-Y. Shen, X.-F. Jin, J. Luo, Y.-Y. Yun, and G.-B. Song, "A novel whole-line quick-action protection principle for HVDC transmission lines using one-end voltage," *Int. J. Electr. Power Energy Syst.*, vol. 65, pp. 262–270, Feb. 2015.
- [16] A. A. Yusuff, A. A. Jimoh, and J. L. Munda, "Fault location in transmission lines based on stationary wavelet transform, determinant function feature and support vector regression," *Electr. Power Syst. Res.*, vol. 110, pp. 73–83, May 2014.
- [17] H. Livani and C. Y. Evrenosoglu, "A single-ended fault location method for segmented HVDC transmission line," *Electr. Power Syst. Res.*, vol. 107, pp. 190–198, Feb. 2014.
- [18] Y. Ma, H. Li, G. Wang, and J. Wu, "Fault analysis and traveling-wave-based protection scheme for double-circuit LCC-HVDC transmission lines with shared towers," *IEEE Trans. Power Del.*, vol. 33, no. 3, pp. 1479–1488, Jun. 2018.
- [19] S. H. A. Niaki, H. K. Karegar, and M. G. Monfared, "A novel fault detection method for VSC-HVDC transmission system of offshore wind farm," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 475–483, Jun. 2015.
- [20] R. Li, L. Xu, and L. Yao, "DC fault detection and location in meshed multiterminal HVDC systems based on DC reactor voltage change rate," *IEEE Trans. Power Del.*, vol. 32, no. 3, pp. 1516–1526, Jun. 2017.
- [21] E. Kontos, R. T. Pinto, S. Rodrigues, and P. Bauer, "Impact of HVDC transmission system topology on multiterminal DC network faults," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 844–852, Apr. 2015.
- [22] R. Imawan, K. Srivastava, and M. Reza, "Fault detection in HVDC-connected wind farm with full converter generator," *Int. J. Electr. Power Energy Syst.*, vol. 64, pp. 833–838, Jan. 2015.
- [23] K. De Kerf, K. Srivastava, M. Reza, D. Bekaert, S. Cole, D. Van Hertem, and R. Belmans, "Wavelet-based protection strategy for DC faults in multi-terminal VSC HVDC systems," *IET Gener., Transmiss., Distrib.*, vol. 5, no. 4, pp. 496–503, Apr. 2011.
- [24] X. Li, S. Lin, and S. Jiang, "Identification method for internal and external faults based on amplitude feature of backward traveling wave for HVDC transmission line," in *Proc. 5th Int. Conf. Electr. Utility Deregulation Restructuring Power Technol. (DRPT)*, Nov. 2015, pp. 1166–1169.
- [25] Z. Qing, Z. Xiao-Long, L. Xin-ping, X. Jie, Z. Ting, and W. Cheng-Jiang, "Mechanical design and research of a novel power lines inspection robot," in *Proc. Int. Conf. Integr. Circuits Microsyst. (ICICM)*, Nov. 2016, pp. 363–366.
- [26] X. Yue, H. Wang, and Y. Jiang, "A novel 110 kV power line inspection robot and its climbing ability analysis," *Int. J. Adv. Robot. Syst.*, vol. 14, no. 3, May 2017, Art. no. 1729881417710461.
- [27] C. M. Shruthi, A. P. Sudheer, and M. L. Joy, "Dual arm electrical transmission line robot: Motion through straight and jumper cable," *Automatika*, vol. 60, no. 2, pp. 207–226, Apr. 2019.

- [28] E. J. Lima, M. H. S. Bomfim, and M. A. D. M. Mourão, "POLIBOT—Power lines inspection RoBOT," *Ind. Robot, Int. J.*, vol. 45, no. 1, pp. 98–109, Jan. 2018.
- [29] C. Deng, J.-Y. Liu, Y.-B. Liu, and Y.-Y. Tan, "Real time autonomous transmission line following system for quadrotor helicopters," in *Proc. Int. Conf. Smart Grid Clean Energy Technol. (ICSGCE)*, Oct. 2016, pp. 61–64.
- [30] C. Deng, S. Wang, Z. Huang, Z. Tan, and J. Liu, "Unmanned aerial vehicles for power line inspection: A cooperative way in platforms and communications," *J. Commun.*, vol. 9, no. 9, pp. 687–692, 2014.
- [31] T. Jabid and M. Z. Uddin, "Rotation invariant power line insulator detection using local directional pattern and support vector machine," in *Proc. Int. Conf. Innov. Sci., Eng. Technol. (ICISSET)*, Oct. 2016, pp. 1–4.
- [32] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, "Insulator detection in aerial images for transmission line inspection using single shot multibox detector," *IEEE Access*, vol. 7, pp. 9945–9956, 2019.
- [33] R. M. Prates, R. Cruz, A. P. Marotta, R. P. Ramos, E. F. S. Filho, and J. S. Cardoso, "Insulator visual non-conformity detection in overhead power distribution lines using deep learning," *Comput. Electr. Eng.*, vol. 78, pp. 343–355, Sep. 2019.
- [34] V. N. Nguyen, R. Jossen, and D. Roverso, "Intelligent monitoring and inspection of power line components powered by UAVs and deep learning," *IEEE Power Energy Technol. Syst. J.*, vol. 6, no. 1, pp. 11–21, Mar. 2019.
- [35] D. Pernebayeva, A. Irmanova, D. Sadykova, M. Bagheri, and A. James, "High voltage outdoor insulator surface condition evaluation using aerial insulator images," *High Voltage*, vol. 4, no. 3, pp. 178–185, Sep. 2019.
- [36] X. Li, Q. Yang, Z. Chen, X. Luo, and W. Yan, "Visible defects detection based on UAV-based inspection in large-scale photovoltaic systems," *IET Renew. Power Gener.*, vol. 11, no. 10, pp. 1234–1244, Aug. 2017.
- [37] J. Liu, R. Jia, W. Li, F. Ma, H. M. Abdullah, H. Ma, and M. A. Mohamed, "High precision detection algorithm based on improved RetinaNet for defect recognition of transmission lines," *Energy Rep.*, vol. 6, pp. 2430–2440, Nov. 2020.
- [38] B. K. Wang, X. Y. Peng, M. Jiang, and D. T. Liu, "Real-time fault detection for UAV based on model acceleration engine," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9505–9516, Dec. 2020.
- [39] J. Han, Z. Yang, Q. Zhang, C. Chen, H. Li, S. Lai, G. Hu, C. Xu, H. Xu, D. Wang, and R. Chen, "A method of insulator faults detection in aerial images for high-voltage transmission lines inspection," *Appl. Sci.*, vol. 9, no. 10, p. 2009, May 2019.
- [40] Y. Wu, Y. Qin, Z. Wang, X. Ma, and Z. Cao, "Densely pyramidal residual network for UAV-based railway images dehazing," *Neurocomputing*, vol. 371, pp. 124–136, Jan. 2020.
- [41] W. Cao, L. Zhu, J. Han, T. Wang, and Y. Du, "High voltage transmission line detection for UAV based routing inspection," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics*, Jul. 2013, pp. 554–558.
- [42] C. Liu, Y. Wu, J. Liu, Z. Sun, and H. Xu, "Insulator faults detection in aerial images from high-voltage transmission lines based on deep learning model," *Appl. Sci.*, vol. 11, no. 10, p. 4647, May 2021.
- [43] M. R. Usman, M. A. Usman, M. A. Yaq, and S. Y. Shin, "UAV reconnaissance using bio-inspired algorithms: Joint PSO and penguin search optimization algorithm (PeSOA) attributes," in *Proc. 16th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2019, pp. 1–6.
- [44] F. Hassan, M. R. Usman, S. Hamid, M. A. Usman, C. Politis, and G. B. Satrya, "Solar powered autonomous hex-copter for surveillance, security and monitoring," in *Proc. IEEE Asia Pacific Conf. Wireless Mobile (APWiMob)*, Apr. 2021, pp. 188–194.
- [45] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of artificial intelligence and machine learning in smart cities," *Comput. Commun.*, vol. 154, pp. 313–323, Mar. 2020.
- [46] J. Sandino, F. Vanegas, F. Gonzalez, and F. Maire, "Autonomous UAV navigation for active perception of targets in uncertain and cluttered environments," in *Proc. IEEE Aerosp. Conf.*, Mar. 2020, pp. 1–12.
- [47] L. Xu, J. Liu, L. Xie, and X. He, "Multi-UAV navigation and recharging for fair and sustainable coverage in wireless networks," in *Proc. 3rd Int. Conf. Adv. Inf. Sci. Syst. (AISS)*, Nov. 2021, pp. 1–6.
- [48] T. Guo, N. Jiang, B. Li, X. Zhu, Y. Wang, and W. Du, "UAV navigation in high dynamic environments: A deep reinforcement learning approach," *Chin. J. Aeronaut.*, vol. 34, no. 2, pp. 479–489, Feb. 2021.
- [49] O. Walker, F. Vanegas, F. Gonzalez, and S. Koenig, "A deep reinforcement learning framework for UAV navigation in indoor environments," in *Proc. IEEE Aerosp. Conf.*, Mar. 2019, pp. 1–14.
- [50] C. Wang, J. Wang, J. Wang, and X. Zhang, "Deep-reinforcement-learning-based autonomous UAV navigation with sparse rewards," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6180–6190, Jul. 2020.
- [51] Y. Yang, K. Zhang, D. Liu, and H. Song, "Autonomous UAV navigation in dynamic environments with double deep Q-networks," in *Proc. AIAA/IEEE 39th Digit. Avionics Syst. Conf. (DASC)*, Oct. 2020, pp. 1–7.
- [52] High Voltage Direct Current Transmission. *Proven Technology for Power Exchange*. Siemens. Accessed: Aug. 10, 2022. [Online]. Available: https://www.brown.edu/Departments/Engineering/Courses/ENGN1931F/HVDC_Proven_TechnologySiemens.pdf
- [53] S. Hamid, A. Nasir, and Y. Saleem, "Reinforcement learning based hierarchical multi-agent robotic search team in uncertain environment," *Mehran Univ. Res. J. Eng. Technol.*, vol. 40, no. 3, pp. 645–662, Jul. 2021.



WAHEED A. KHAN received the B.Sc. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2010, and the M.Sc. degree in electrical engineering from The University of Lahore, Lahore, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electrical Engineering, The Superior University, Lahore. He is currently working as a Lecturer with The Superior University. His research interests include smart HVDC systems, smart energy management systems, and power electronics.



SHAHZAIB HAMID received the B.S. degree in electrical engineering from the University of Central Punjab, Lahore, Pakistan, in 2015, and the M.S. degree in computer engineering with a specialization in artificial intelligence from the University of Engineering and Technology, Lahore, in 2017. He is currently pursuing the Ph.D. degree with COMSATS University Lahore, Pakistan. He is currently working as the Head Data Scientists with the AI Studio Lahore, Pakistan. His research interests include data science, AI, machine learning, computer vision, and NLP.



M. REHAN USMAN (Member, IEEE) received the B.S. degree in electrical engineering from COMSATS University Islamabad (CUI), Lahore Campus, Pakistan, in 2008, the M.S. degree in electrical engineering (telecommunications) from BTH, Sweden, in 2010, the M.S. degree in project management and operational development from KTH, Sweden, in 2013, and the Ph.D. degree in IT convergence engineering from the Kumoh National Institute of Technology (KIT), South Korea, in 2018. He is currently working as an Associate Professor with the Department of Electrical Engineering, Superior University, Lahore, and enrolled as a Guest Lecturer/Researcher at Kingston University London, U.K. His research interests include resource and interference management in Next Gen. HetNets. (5G and beyond), massive MIMO, NOMA, cyber security, smart solutions for energy/power systems, telehealthcare, ML/DL based medical imaging, and video quality assessment.



ALI RAZA was born in Multan, Pakistan. He received the B.S. and M.S. degrees in electrical engineering, in 2012 and 2017, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan. He has served as an Electrical Engineer at Masood Spinning Mills, Kabir Wala, Pakistan. During his job, he has supervised many mega projects of solar system up to 7MWatts. His main research interest includes HVDC systems. His research interests include smart grid, demand side response, HVDC systems, and power electronics.



CHRISTOS POLITIS (Senior Member, IEEE) is currently a Professor (Chair) of wireless communications with the School of Computer Science and Mathematics (CSM), Kingston University. He is the Co-Director of the Digital Information Research Centre (DIRC) with a staff of about 25 academics, 20 postdoctoral researchers, and over 60 Ph.D. students, making it one of the largest in the field in the U.K. Upon joining Kingston University in 2007. He has Co-Founded and led the Research Group on Wireless Multimedia and Networking (WMN). He teaches modules on wireless systems and networks. Prior to this post, he worked for Ofcom, the U.K. Regulator and Competition Authority, as a Senior Research Manager. While at the University of Surrey, he was a Postdoctoral researcher working on virtual distributed testbeds at the Centre for Communication Systems Research (now the 5G Innovation Centre). This was preceded by placements with Intracom-Telecom SA and Maroussi 2004 SA in Athens, Greece. He has managed to raise funding from the EU and U.K. Research and Technology Frameworks under the ICT and Security Program. He holds two patents. He has published over 170 papers in international journals and conferences and chapters in ten books. He has advised several governmental and commercial organizations on their research program/agendas and portfolios. He is an U.K. Chartered Engineer and a member of the Technical Chamber of Greece.



M. ARSLAN USMAN (Member, IEEE) was born in 1988. He received the B.S. degree in electrical engineering from the COMSATS Institute of Information Technology, Lahore, Pakistan, in 2010, the M.S. degree in electrical engineering with specialization in signal processing from Blekinge Tekniska Hogkola (BTH), Karlskrona, Sweden, in 2013, and the Ph.D. degree in IT convergence engineering from the Kumoh National Institute of Technology (KIT), South Korea, in 2018. After completion of the Ph.D., he continued working with the WENS Laboratory as a Postdoctoral Fellow at the Department of IT Convergence Engineering, Kumoh National Institute of Technology, till August 2018. He is currently a member of Wireless Multimedia and Networking (WMN) Research Group and also a Lecturer in computer science with the Faculty of Engineering, Computing and Environment (ECE), Kingston University London, U.K. His research interests include video quality assessment, image classification and detection, and next generation wireless networks, (MIMO, NOMA, and small cells in next generation networks).



GANDEVA BAYU SATRYA (Senior Member, IEEE) received the bachelor's degree in informatics engineering from STT Telkom, in March 2008, the master's degree in informatics engineering and media informatics from IT Telkom, Bandung, Indonesia, in June 2012, and the Ph.D. degree in security communication in next-generation networks from the Department of IT Convergence Engineering, Kumoh National Institute of Technology, South Korea, in February 2019. He has been a member of the Research Center of Internet of Things (RC IoT), Telkom University, since June 2019. He is currently a Lecturer and a Researcher with the Faculty of Informatics, Telkom University, Bandung. His research interests include routing protocol, packet scheduling, and security communication in next-generation networks. He is appointed as a Technical Activities Coordinator of the IEEE Communication Society (ComSoc) Indonesian Section.

• • •