

Received September 5, 2021, accepted September 20, 2021, date of publication October 4, 2021, date of current version October 18, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3117451

Machine Learning Methods for UAV Flocks Management-A Survey

RINA AZOULAY¹, YORAM HADDAD¹, (Senior Member, IEEE), AND SHULAMIT RECHES²

¹Department of Computer Science, Jerusalem College of Technology, Jerusalem 9116001, Israel

²Department of Mathematics, Jerusalem College of Technology, Jerusalem 9116001, Israel

Corresponding author: Rina Azoulay (azrina@g.jct.ac.il)

This work was supported in part by the Israel Innovations Authority under Wireless Intelligent Networks (WIN), which is a research project headed by the Israeli Consortium for WIN.

ABSTRACT The development of unmanned aerial vehicles (UAVs) has been gaining momentum in recent years owing to technological advances and a significant reduction in their cost. UAV technology can be used in a wide range of domains, including communication, agriculture, security, and transportation. It may be useful to group the UAVs into clusters/flocks in certain domains, and various challenges associated with UAV usage can be alleviated by clustering. Several computational challenges arise in UAV flock management, which can be solved by using machine learning (ML) methods. In this survey, we describe the basic terms relating to UAVS and modern ML methods, and we provide an overview of related tutorials and surveys. We subsequently consider the different challenges that appear in UAV flocks. For each issue, we survey several machine learning-based methods that have been suggested in the literature to handle the associated challenges. Thereafter, we describe various open issues in which ML can be applied to solve the different challenges of flocks, and we suggest means of using ML methods for this purpose. This comprehensive review may be useful for both researchers and developers in providing a wide view of various aspects of state-of-the-art ML technologies that are applicable to flock management.

INDEX TERMS Collaborative unmanned aerial vehicles, machine learning, multi-robot systems, reviews, unmanned aerial vehicles.

I. INTRODUCTION

The technological advances in the abilities of unmanned aerial vehicles (UAVs) and the reduction in their costs have made UAVs more common than ever, with a significant potential for future use. During the Covid-19 period, UAVs could theoretically be used to assist in supply and transportation goals. Moreover, it is expected that UAVs will be integrated into the 5G and future communication networks. UAVs offer numerous applications in agriculture, maintenance, intelligence, and diverse security needs. However, several challenges should be overcome to leverage the potential of UAVs for practical goals successfully. For example, there is a need for a smart, efficient power control mechanism owing to the energy consumption limitation of UAVs, which are constrained in their weight-carrying ability. Moreover, aerial path planning should be performed carefully, while maintaining the security and safety of UAVs.

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Asif¹.

In addition to the aforementioned challenges that occur even with a single UAV, certain tasks and issues arise when considering a multi-UAV environment, in which a small or a large group should act together or should act in the same aerial environment. The formation of UAV networks, UAV clustering and coordination, and resource allocation in the UAV network should be considered and efficiently solved. A hierarchical description of the various topics relating to UAVs is provided in Figure 1.

As part of this survey, we also provide several insights into general studies that are not directly focused on UAV environments, but that may be useful when applied to UAV environments. We believe that a complete view of the existing works, both those that have been specifically developed for UAVs and those that have not, can help UAV designers to address the challenges that arise in managing and maintaining UAV flocks more efficiently.

As demonstrated in Section III, numerous existing surveys have dealt with the challenges and opportunities of UAV technologies. However, our goal is to focus on the increasingly

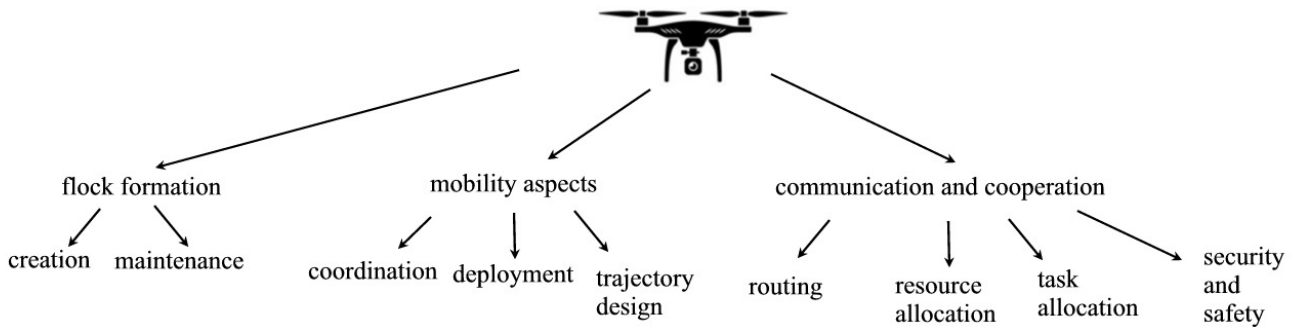


FIGURE 1. Challenges in UAV flocking.

important topic of applying machine learning (ML) methods to UAV flock domains.

The remainder of this paper is organized as follows: Section II provides basic terminology for modern ML methods that may be applicable to the improved handling of UAV flocks. Section III presents several relevant surveys that have considered UAV technologies and ML methods used for UAVs, as well as flying ad-hoc network (FANET) environments. Section IV considers relevant studies that have dealt with flock formation while Section V describes flocking and coordination issues. Studies dealing with the deployment challenge are surveyed in Section VI, while trajectories design issues are discussed in Section VII. In Section VIII, the challenges of resource allocation and power control are discussed, whereas Section IX focuses on studies that have addressed the challenging problem of task allocation in UAV flocks. Finally, Section X provides some guidelines and direction to find the best method to solve a particular challenge, and Section XI provides some insights into future works. We conclude our survey in Section XII.

II. SOME BASIC TERMS AND BACKGROUND

As the main goal of our survey is to review the application of ML methods to advance artificial aerial swarm scenarios, we first briefly present the main ML methods considered in the literature that are reviewed within the framework of this survey. Figure 2 presents a scheme with the details of several ML methods reviewed in this survey, and In Table 1, we list the main ML methods described in this survey, their main features, and their applicability to various UAV flock challenges studied in this survey.

Various UAV types were used in the reviewed papers. An overview of UAV types can be found in [44]. The authors provided a detailed description of the available simulation tools for UAV system performance, including the requirements, goals, and strengths and weaknesses of each tool. They divided the UAVs into three categories: micro UAVs (the weight of which can reach 2 kg), mini UAVs (2 to 20 kg), and small UAVs (20 to 150 kg). Micro UAVs have an endurance of several hours, and can be used for reconnaissance and inspection, whereas mini and small UAVs can last for up to

two days and can be used for data gathering. All UAV types can be used for various surveillance goals.

Another UAV classification, based on UAVs' functionalities, is suggested in the review of Yan *et al.* [125]. According to their classification, the UAVs are classified according to their communication capability, which is measured according to the radius range of the UAV data link. According to this classification, there are long-endurance UAVs, mid-range UAVs with radius range of about 650 km, short-range small UAVs with radius range less than 350 km, close-range UAVs with coverage of at least 30 km, very low cost closer-range UAVs, and commercial and consumer UAVs/drones, with very limited communication range, that may be controlled by smartphone or tablet applications.

Regarding UAV flocks, Coppola *et al.* [31] distinguished among different flock types as follows: **Teams** are typically small groups, where each agent optimizes individual objectives in a cooperative or competitive manner. **Formation** refers to a situation where each agent is typically assigned a specific sub-task, role or placement in a cooperative manner, whereas **swarms** typically refer to large groups of dispensable agents. Nevertheless, in this survey, we consider all types of UAV formations as **flocks**.

III. RECENT TUTORIALS AND SURVEYS

The aim of this study is to review recent works that have considered the concept of applying state-of-the-art ML methods to handle the challenges of UAV flocks. In recent years, several tutorials and surveys have been presented on related topics, such as the application of ML methods to solve challenges in UAV design and management, or surveys dealing with flocking challenges. However, no previous surveys have focused on the use of ML for solving UAV flocking challenges.

Therefore, in this section, we first review existing surveys and tutorials that are closely related to our goal. We categorize these various studies into two main categories. Section III-A focuses on the challenges in flock and swarm creation and management, including several flocking issues such as efficient deployment and security enhancement, whereas Section III-B focuses on surveys that have considered using ML techniques for enhancing UAV technologies and the

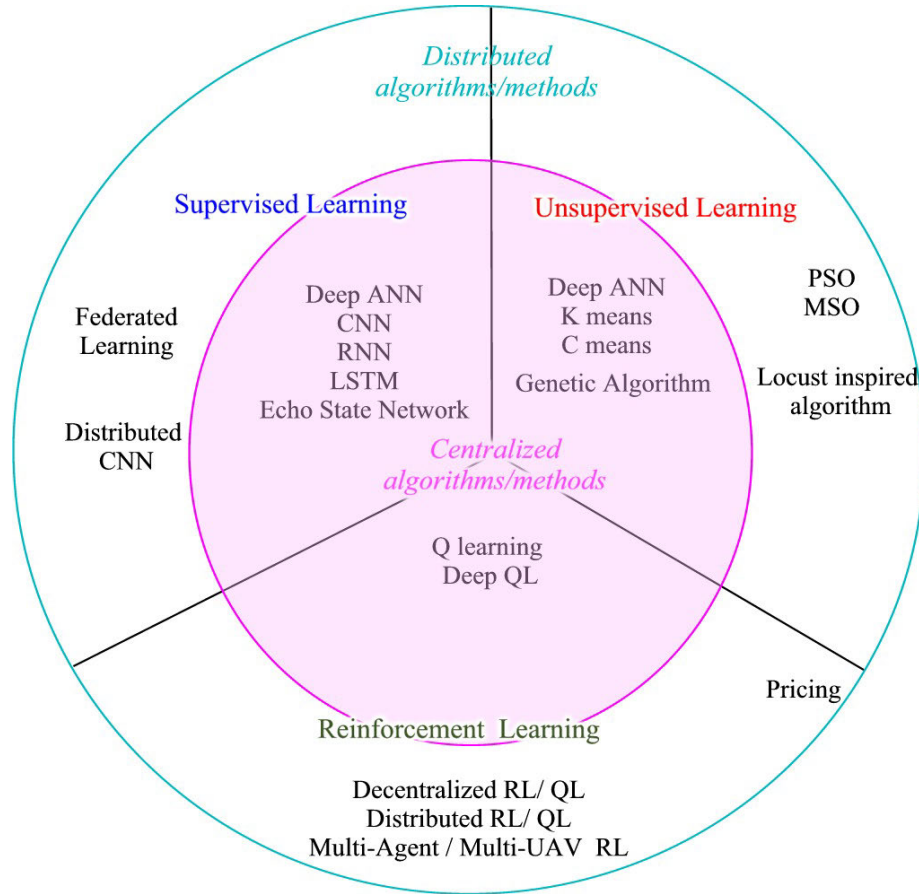


FIGURE 2. Main ML methods.

communication between UAVs. Furthermore, Section III-C discusses studies that have focused on using ML methods in this area, and finally, several comparisons and a generalization of the above.

In Table 2, we provide a list of acronyms used throughout this paper.

A. SURVEYS ON UAV FLOCKS CREATION, MANAGEMENT AND COORDINATION

The creation of groups of UAVs is important in several applications, such as wireless communication applications, agriculture applications, and security purposes. In the following section, we review several surveys that have dealt with challenges considering flock creation and management, where the flocks can be used for several application types. The significant advancements in UAV technology have resulted in applications whereby groups of UAVS (which are known as flocks or swarms) can cooperate to perform a global task. This may be the case when considering UAVs that act as cellular base stations (BSs), which should decide among themselves how to offer deployment to provide good coverage of an area for ground users. Other scenarios encompass different domains where UAVs can be used, such as photography tasks, control tasks, transportation applications, security

purposes, and agriculture domains. The common factor of all these applications is that they gave rise to several challenges whereby decisions should be made, such as UAV deployment, task allocation, and maintenance issues. In the following, we describe several surveys relating to the aforementioned issues, some of which also discuss relevant ML solutions to the problems reviewed.

The significant potential of the general use of UAVs, and more specifically, UAV flocks, was mentioned by Tahir *et al.* [110]. The authors analyzed the core characteristics of swarming drones, including the UAV mechanics, functionality, organization, modeling, and applications, as well as the autonomous aspects of such drones and drone swarms. Moreover, they conducted an online survey to determine the public awareness regarding the potential of using drone technology. The survey results demonstrated that, although most people were concerned about UAV usage, the majority did not know about the concept of a swarm of drones.

Although there is significant potential, several challenges also need to be overcome. Chung *et al.* [29] surveyed algorithms that allow the individual members of the swarm to communicate and to allocate tasks among themselves, plan their trajectories, and coordinate their flight so that the overall objectives of the swarm can be achieved efficiently.

TABLE 1. Main ML methods (C = centralized, D = distributed).

ML Method	Learning type	Approach	Main UAV Applications
DNN	Supervised/unsupervised	C/D	Resource allocation
CNN	Supervised	C	Clustering, resource allocation
RNN	Unsupervised	C	Task allocation
LSTM	Supervised	C	Deployment
Federated learning	Supervised	C/D	Resource allocation
Q-learning Reinforcement learning	Reinforcement learning	C/D	Clustering, flocking, Deployment, positioning, Resource allocation
Deep Q-learning, Deep Reinforcement learning	Reinforcement learning	C/D	Flocking, deployment, positioning, resource allocation, task allocation
Particle swarm optimization	Bio-inspired	D	Clustering, routing positioning, task allocation
Genetic algorithm	Bio-inspired	C	Clustering, deployment, task allocation
K-means	Unsupervised	C	Clustering, deployment, resource allocation, task allocation
Pricing	Unsupervised	D	Resource allocation

The authors reviewed several theoretical tools, with a specific focus on how they have been developed for and applied to aerial swarms, including theoretical algorithms, ML methods, and other advanced planning and decision-making strategies. They emphasized that an important area of further study is the development of learning and decision-making architectures that will endow swarms of aerial robots with high levels of autonomy and flexibility.

Coppola *et al.* [31] focused on micro air vehicles (MAVs), which are miniature UAVs with a size as small as 5 cm, and reviewed the challenges that must be addressed to develop MAVs for real-world operations successfully. The challenges begin at the lowest level, in terms of how the MAV design will impact the swarm behavior, and end at the highest level, where collective behaviors must be designed that can best exploit the lower-level designs, controllers, and sensors. Although the work of Coppola *et al.* did not focus on ML, they discussed several ML methods for behavior design and optimization. In particular, they investigated evolutionary robotics and reinforcement learning (RL).

Returning to the general problem of UAV flock formation, Beaver and Malikopoulos [15] presented an overview of optimization approaches for flocking behavior that provide strong safety guarantees. They summarized the results of existing optimal flocking works across engineering disciplines and presented the frontier of flocking and optimization research. They also aimed to provide a new paradigm for understanding flocking as an emergent phenomenon to be controlled, rather than desirable group behavior for agents to mimic.

Sargolzaei *et al.* [8] explored several applications of cooperative UAV missions in various fields, and reviewed emerging topics in the field of cooperative UAV control and their associated practical approaches. They surveyed the applications, algorithms, and challenges. They categorized the applications into surveillance, search and rescue, mapping, and

military applications. Similarly, they categorized the algorithms into three main classes: consensus control, flocking control, and guidance-based cooperative control. Furthermore, they investigated the challenges relating to the cooperative control and applications of cooperative algorithms and provided the related mathematics of the cooperative control algorithms.

Several studies have drawn inspiration from the behavior of flocks in nature and used bio-inspired algorithms for the formation and management of UAV flocks/swarms. Long *et al.* [71]

surveyed the literature on swarm shepherding, which is a method for controlling multiple coordinated robotic agents. The shepherding problem can be defined as the guidance of a swarm of agents from an initial location to a target location. The shepherd is responsible for the high-level path planning of the swarm and task allocation, whereas the swarm members themselves function solely on single-agent dynamics. Long *et al.* mentioned the use of ML and computational intelligence techniques for swarm control, and for shepherding in particular.

All the above surveys focused on defining general flock types and their management. Several also mentioned ML methods, but none concentrated on surveying ML methods for the creation and management of UAV flocks.

B. SURVEYS ON UAV-BASED WIRELESS COMMUNICATION NETWORKS

The incredible technological advancements in the capabilities for organizing groups of drones have considerably expanded their use as part of the 5G and beyond systems. A UAV-based wireless network is often referred to as a FANET. This term was formally defined by Bekmezci and Temel [50], who suggested several applications for FANET and discussed various design challenges. FANET design considerations were

TABLE 2. List of acronyms.

Acronym	Text
ABS	Aerial base station
ACO	Ant colony optimization
AI	Artificial intelligence
ANN	Artificial neural network
APC	Affinity propagation clustering
AQL	Adaptive Q-learning
BS	Base station
CH	Cluster head
CMTAP	Cooperative multiple task assignment problem
CNN	Convolutional neural network
CPS	Cyber-physical system
DNN	Deep neural network
DRL	Deep RL
ESN	Echo state network
FANET	Flying ad-hoc Network
FL	Federated learning
FTA	Fast task allocation
GA	Genetic algorithm
GPI	Geographical position information
HAPS	High-altitude platform station
HTER	Half total error rate
IoD	Internet of drones
LEACH	Low-energy adaptive clustering hierarchy
LEACH-C	LEACH-centralized
LP	Linear programming
LSTM	Long short-term memory
MA	Multi-agent
MADDPG	Multi-agent deep deterministic policy gradient
MARL	Multi-agent reinforcement learning
MAV	Micro air vehicle
MDP	Markov decision process
MEC	Mobile edge computing
MFG	Mean-field game
ML	Machine learning
NDP	Neuro-dynamic programming
NFV	Network function virtualization
NLP	Nonlinear programming
PSO	Particle swarm optimization
QoE	Quality of experience
RAN	Radio access network
RL	Reinforcement learning
RPW	Red palm weevil
SDN	Software-defined network
SI	Swarm intelligence
SPNE	Subgame perfect Nash equilibrium
STAPP	Simultaneous target assignment and path planning
UAV	Unmanned aerial vehicle
WSN	Wireless sensor network

investigated, including the adaptability, scalability, latency, UAV platform constraints, and bandwidth.

Hayat *et al.* [42] discussed the characteristics and requirements of UAV networks from a communications and networking perspective. They surveyed and quantified the quality-of-service requirements, network-relevant mission parameters, data requirements, and minimum data to be transmitted over the network. They reported experimental results from many projects and investigated the suitability of existing communication technologies for supporting reliable aerial networking.

Recent implementations and projects using multi-UAV systems were surveyed by Hentati and Fourati [45]. They presented a review of the recent UAV communication protocols and mechanisms and summarized current research in the area of monitoring via UAVs. They discussed well-known design issues relating to UAV communication protocol, including mobility, adaptability, reliability, scalability, latency, UAV platform constraints, bandwidth allocation, power consumption, computational power, localization, security, and privacy. Another recent study by Boccadoro and Grieco [85] categorized the multifaceted aspects of the Internet of drones (IoD), including several fields in which drones and swarms of drones may be useful. Their goal was to provide an overview of the current state of research activities regarding the IoD.

The survey of Fotouhi *et al.* [38] focused on regulation issues that arise when integrating UAVs into cellular systems. They reviewed the available UAV types and potential solutions to interference issues, as well as challenges and opportunities for assisting cellular communications with UAV-based flying relays and BSs. They also considered testing, regulation, and security challenges.

Oubbati *et al.* [99] summarized the main characteristics of software-defined network (SDN) and network function virtualization (NFV) technologies that are used for the efficient management of UAV-assisted mobile networks. They presented an in-depth discussion relating to both the design challenges of UAV networks and their principal use cases.

The authors emphasized that AI techniques and ML methods are expected to play a crucial role in optimizing UAV-assisted networks in various aspects, such as optimizing the resource allocation and scheduling, enhancing the network prediction and boosting the network performance. However, multiple challenges require further investigation, such as the high computational processing, high energy consumption, and high latency.

A special focus on the algorithmic challenges involved in UAV-based networks was provided in the study of Ullah *et al.* [114]. They investigated the joint optimization problem to enhance the system efficiency of UAV-assisted next-generation communication systems. They focused on the following challenges: flight trajectory, time scheduling, altitude optimization, aerial and relay BSs, energy harvesting, power transfer, optimal power consumption, and resource allocation. They presented an in-depth study regarding various successive optimization algorithms and methods, such as mobile edge computing (MEC) techniques and SDNs, as well as ML and deep learning applications that play a vital role in 5G and B5G communication.

The potential of using ML methods for UAV-based communication networks has been the highlight of several recent surveys. Mozaffari *et al.* [79] surveyed the applications and challenges of UAV usage in communication networks, as aerial BSs (ABSs), or as flying mobile terminals. Specifically, they also noted issues to which ML can be applied, namely channel modeling and trajectory optimization.

Saad *et al.* [98] surveyed the countless applications of UAVs and UAV swarms, and the challenges that should be addressed when considering these technologies. They reviewed research problems pertaining to UAV network performance analysis and optimization, including the physical layer design, trajectory path planning, resource management, multiple access, cooperative communications, standardization, control, deployment strategies, and security issues. Sharma *et al.* [103] provided insights into the latest UAV communication technologies through the investigation of suitable task modules, antennas, resource handling platforms, and network architectures. They considered the applicability of ML, encryption, and optimization techniques to ensure long-lasting and secure UAV communications.

Moreover, several surveys have mainly focused on describing the major ML methods and their application to UAV wireless communication networks, which we review in Section III-C.

Owing to the common use of skimmers in security areas, one of the most challenging issues in the implementation of a wireless UAV-based network is the need for protection and security.

Shakeri *et al.* [101] surveyed several design challenges of multi-UAV systems for cyber-physical system (CPS) applications. They considered the modern generation of systems with synergic cooperation between computational and physical potentials that can interact with humans through several new mechanisms. UAVs can act as part of such systems, and the potential of ML to improve the efficiency of CPSs was discussed.

Wang *et al.* [116] presented a comprehensive survey on UAV networks from a CPS perspective. They attempted to provide novel insight into the state of the art in UAV networks. The basics of the three cyber components, namely communication, computation, and control, were discussed, including the requirements, challenges, solutions, and advances. The three UAV network hierarchies, namely the cell level, system level, and system of the system level, were investigated according to the CPS scale, with the architecture, key techniques, and applications of the UAV network explicitly demonstrated under each hierarchy. Blockchain applications in UAV networks have even been considered in [111]. Such applications include network security, decentralized storage, inventory management, and surveillance. It worth noting that they also considered ML as one of the blockchain applications in other emerging areas.

As the focus of our survey is on the use of UAV flocks/swarms, we present the following surveys relating to the use of UAV swarms in wireless communication networks. Li *et al.* [66] presented space-air-ground integrated networks and discussed the research challenges faced by the emerging integrated network architecture. They also provided a review of various 5G techniques based on UAV platforms that could be categorized into different domains, such as the physical layer and network layer, as well as joint communication, computing, and caching. They defined a UAV-based swarm

network as a network based on a swarm of UAVs that offers ubiquitous connectivity to ground users. The benefits of such a network are that it is highly flexible with rapid provision features, and it is a feasible solution to enable fast, effective recovery and communication expending.

Another survey relating to UAVs for wireless networks was conducted by Zhang *et al.* [135]. They discussed several aspects of using air-ground integrated mobile edge networks, in which UAVs are employed to assist the MEC network. Among the advantages of this approach, they provided the following examples: the easy deployment of such networks, the high mobility of UAVs, and the ability to provide MEC services ubiquitously and reliably. They also considered the advantages as well as the challenges of using UAV swarms, which can complete tasks more efficiently and can improve the chance of successful missions. Furthermore, they can improve the scalability and reduce the probability of system failure. However, they noted that the means of hiring UAVs requires further discussion.

The above survey focused on the use of UAVs as part of communication networks. However, as mentioned previously, UAVs have several purposes, and somehow, different challenges arise when considering a communication network from when aiming to connect the UAVs among themselves. XiJun *et al.* [28] focused on UAV swarm communication architectures and routing protocols. They detailed four communication architectures for UAV swarms. They also discussed applicable scenarios and introduced a systematic overview of as well as feasibility research on routing protocols. Finally, they investigated the open research issues of UAV swarm communication architectures and routing protocols. However, they hardly addressed ML methods.

In the following section, we present surveys that have focused on ML methods used for solving different challenges arising in UAV technologies and UAV swarms.

C. SURVEYS ON APPLICATION OF ML TO UAV-RELATED CHALLENGES

In the following, we mainly focus on ML methods used for UAV control and UAV flock control. Considering a group of UAVs in terms of a communication network offers several advantages as well as faces various challenges. In fact, general-purpose UAVs can form a D2D wireless network among themselves in order to be coordinated. In other applications, the mesh network of the UAVs can be used as a cloud that provides a wireless network for ground users.

Bithas *et al.* [16] presented a survey on the use of ML techniques for UAV-based communication and the improvement of various design and functional aspects. Their survey focused on UAVs that are used for communication tasks, as part of 5G and beyond communication systems, in which the following aspects were detailed: channel modeling, interference management, parameter configuration, security and safety, resource management, position detection, and localization. However, our survey differs from the study of Bithas *et al.* in terms of the following aspects: (a) we concentrate

on the challenges relating to the planning, organization, and maintenance of UAVs flocks, rather than dealing with the challenges of a single UAV; (b) we provide a brief description of classical and innovative ML methods and their use in the construction, management, and utilization of UAV flocks for different goals; and (c) we provide an overall survey of related studies that may be applicable for these challenges, even if they did not directly consider the UAV flocking goal.

Several studies have reviewed the capability of artificial neural networks (ANNs) and advanced deep learning methods to address some of the above challenges relating to UAV applications. Carrio *et al.* [19] reviewed several applications of deep learning for UAVs, including the most relevant developments as well as their performances and limitations.

Mao *et al.* [89] performed a comprehensive survey of the applications of deep learning algorithms for different network layers, including physical layer modulation/coding, data link layer access control, resource allocation, and routing layer path searching and traffic balancing. They also discussed the use of deep learning to enhance other network functions, such as network security and sensing data compression.

Chen *et al.* [25] considered the potential of using ANNs for solving various wireless networking problems. They presented a detailed overview of several recent ANN types, including recurrent, spiking, and deep neural networks, which are pertinent to wireless networking applications. For each ANN type, they presented the basic architecture as well as specific examples that are particularly important for wireless network design. Moreover, they provided an in-depth overview of the variety of wireless communication problems that can be addressed using ANNs. For each individual application, they presented the main motivation for using ANNs, along with the associated challenges and use cases.

Another survey considering deep learning for wireless networking was presented by Zhang *et al.* [23]. They introduced and compared state-of-the-art deep learning models, and provided guidelines for model selection towards solving networking problems, whereby the various scenarios and applications in wireless networking were grouped, ranging from mobile traffic analytics to security and emerging applications.

The recent study of [57] examined how ML can be used in the area of radio access networks (RANs) with UAVs. The paper investigated which types of ML methods are useful for designing UAV-based RANs (U-RANs), with a particular focus on supervised RL strategies. They discussed the advantages and potential of using ML algorithms in U-RANs in numerous network design scenarios that could not be easily solved by conventional model-based methods. They investigated the different types of ML methods that are useful for designing U-RANs by focusing on supervised and RL strategies in particular.

Al-Turjman *et al.* [4] presented an overview of the use of artificial intelligence (AI) in UAV communications. They reviewed the communication protocols and technologies that have been applied in UAV communications and discussed

several AI-based classification and image-based techniques that are used for UAV communication. Another recent study on UAV-based communication with ML method is the study of Wu *et al.* [119], which focused on the exploitation of advanced techniques, such as intelligent reflecting surfaces, short packet transmission, energy harvesting, joint communication and radar sensing, and edge intelligence, to meet the diversified service requirements of future wireless systems. They also reviewed ML methods specifically for UAV trajectory and communication design. Interestingly, Wu *et al.* considered the three research dimensions that are the focus of our survey, namely UAV technology, swarm control, and ML utilization. However, Wu *et al.* focused on advanced technologies that are used for UAV-based communication networks, whereas our study focuses on ML methods that are used for multipurpose UAV flocks and swarms.

The majority of the above surveys focused on the utilization of ML methods for solving FANET challenges, but none of them considered the challenges that are specific to UAV clustering, which is an important issue in several practical areas. Table 3 summarizes the afore-mentioned reviews, with a special mention of the main focus of each one. In particular, for each survey, we note whether it focused on UAV/aerial vehicles, implicitly considered flocks and flock formation, and whether ML methods and applications were also surveyed. In the table, the notation “V” indicates a main concern, whereas “+” denotes that several relevant ideas were mentioned, but with no main focus on this issue, and finally, × means that the study (almost) did not consider the mentioned issue.

To summarize, Table 3 presents a list of surveys related to the topic of UAVs flocks, and the challenges relating to their design and management.

The contribution and novelty of our survey compared to the surveys mentioned previously and summarized in Table 3 are based on the fact that we concentrate on using ML methods for UAV flock management. Our scope is in the intersection among ML, swarm intelligence (SI), and studies relating to UAVs, which none of the previous surveys have encompassed. Figure 3 illustrates this point.

IV. FLOCKS FORMATION AND MAINTENANCE

The possible co-existence of large swarms of aerial vehicles has logically led to the necessity of grouping them into flocks. The grouping of UAVs into flocks may be useful for several applications, such as UAV-based communication networks, wireless sensor networks (WSNs), information gathering, image collection, and military applications. The clustering of UAVs implies many tasks, such as dividing them into clusters, where each cluster has a cluster head (CH) that is responsible for data aggregation and transmission to the BS.

The first problem that arises when dealing with UAV organization into flocks is that of flock formation; that is, how to cluster the UAVs into flocks. This includes (a) the affiliation of each UAV to a certain flock; (b) the internal topology within the flock; (c) the internal routing protocol

TABLE 3. Summary of scope of surveys.

Reference	Scope	UAVs	Flocks	ML
Tahir et al. [110]	Future usage of UAVs and flocks	V	V	×
Chung et al. [29]	Aerial swarm robotics	V	V	+
Coppola et al. [31]	Challenges of Micro Air Vehicles	V	×	+
Beaver et al. [15]	Optimal robotic flocking	+	V	+
Sargolzaei et al. [8]	Flock cooperation control	V	V	+
Long et al. [71]	Swarm control	×	V	V
Bekmezci et al. [50]	Flying ad-hoc networks	V	V	×
Hayat et al. [42]	Civil applications	V	V	×
Hentati et al. [45]	Design issues	V	V	+
Boccardo et al. [85]	Internet of Drones	V	+	×
Fotouhi et al. [38]	Standards of UAV based communications	V	×	+
Ullah et al. [114]	Optimizing communication efficiency	V	+	V
Mozaffari et al. [79]	UAV based communication	V	×	+
Saad et al. [98]	UAV based networks	V	+	V
Sharma et al. [103]	Networking technologies	V	×	+
Shakeri et al. [101]	Cyber-physical applications	V	V	V
Wang et al. [116]	Cyber UAV perspective	V	V	V
Alladi et al. [111]	Blockchain applications	V	×	+
Li et al. [66]	UAV communication	V	+	V
Oubhati et al. [82]	SDN and NFV for UAV assisted networks	V	V	+
Zhang et al. [135]	Air-Ground Integrated Networks	V	+	+
XiJun et al. [28]	UAV swarm communication	V	V	+
Bithas et al. [16]	ML for UAV communication systems	V	×	V
Carrio et al. [19]	Deep Learning for UAVs	V	×	V
Mao et al. [89]	Deep Learning for Wireless networks	+	+	V
Chen et al. [25]	ANNs for wireless networks	V	×	V
Zhang et al. [23]	Deep Learning for wireless networks	×	+	V
Kouhdaragh et al. [57]	ML for UAV-Based Networks Design	V	×	V
Alsami et al. [5]	AI for robotic communication	V	V	V
Al-Turjman et al. [4]	AI-based UAV communication system	V	×	V
Wu et al. [119]	Sensing and Intelligence	V	V	V

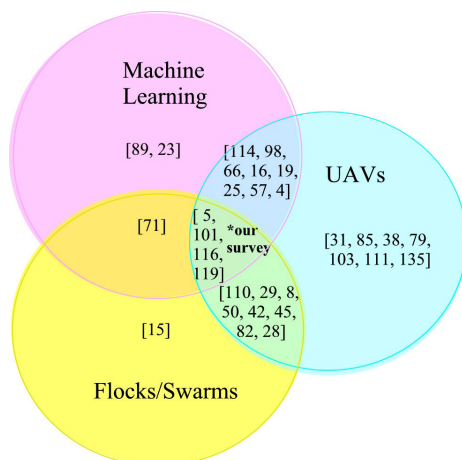


FIGURE 3. Main focus of related surveys.

used in the flock; (d) the required updates of the internal and external structure of the flock given external changes; (e) task allocation within the flock; and (f) task allocation updates given external changes.

The aim of this section is to survey ML methods for UAV flocking. Clustering methods based on bio-inspired algorithms and particle swarm optimization (PSO) are presented in Section IV-A, whereas the issue of how ML methods can be applied to clustering formation is investigated in Section IV-B.

A. BIO-INSPIRED ALGORITHMS

Based on the fact that flocking behavior occurs widely in nature, several studies have suggested the use of bio-inspired algorithms to achieve efficient flocking behavior. Owing to their popularity, we describe these in a separate section.

The PSO algorithm is an evolutionary computing technique that is modeled according to the social behavior of a flock of birds. In the context of PSO, a swarm refers to numerous potential solutions to the optimization problem, where each potential solution is known as a particle. The aim of the PSO is to determine the particle position that results in the best evaluation of a given fitness function. Regarding the clustering challenge, Latiff *et al.* [64] were the first to present energy-aware clustering for WSNs using a PSO algorithm. They defined a cost function that considers the intra-cluster distance, available energy of each node, and each CH candidate. The simulation results indicated that the proposed protocol using the PSO algorithm provided a higher network lifetime, delivered more data, and produced better clustering stations compared to LEACH and LEACH-C.

Following Latiff *et al.*, several variations of the basic PSO algorithm were suggested. Kuila and Jana [87] presented linear programming/nonlinear programming (LP/NLP) formulations of the clustering and routing problems in WSNs, followed by two proposed algorithms based on PSO. The clustering algorithm considered the energy conservation of

the nodes through load balancing, whereas the routing algorithm considered a tradeoff between the transmission distance and number of hop counts. In the clustering phase, the routing overheads of the CHs were considered for balancing their energy consumption.

Suganthi and Rajagopalan [108] considered the problem of the construction of energy-effective multi-swarms in dynamic environments and presented variations of PSO that are specifically formulated for excellent functioning in dynamic settings. The primary notion is the extension of single-population PSO as well as charged PSO techniques through the construction of interactive multi-swarms. They also proposed updating as well as recalculating algorithms for connected dominating sets given changes in the ad-hoc wireless network topology.

Collotta *et al.* [30] proposed a fuzzy logic-based mechanism for wireless sensor networks, which defines the sleeping time of the sensor devices. A PSO algorithm was introduced to obtain the optimal values and parameters of the proposed fuzzy logic controller so as to achieve the best results regarding the battery life of the sensor nodes.

A hybrid approach for the clustering scheme was proposed by Aftab *et al.* [1] for drone-based cognitive IoT, which uses a hybrid mechanism of glowworm swarm optimization [117] and the dragonfly algorithm [78]. Whereas cluster formation and CH selection were performed using glowworm swarm optimization, the swarm joining and cluster management mechanisms were inspired by the dragonfly algorithm. Tracking methodology was also proposed using the dragonfly algorithm rules, which aided in more effective management of the cluster topology. The cluster maintenance was performed by a mechanism to identify dead cluster members. Finally, an efficient route selection function was proposed.

Arafat and Moh [7] suggested clustering and CH selection methods based on PSO for reducing the energy consumption and balancing the energy utilization, thereby increasing the network lifetime. According to the PSO method, after computing the search region, a particle is assigned to a subswarm, which is known as a local swarm. The local swarm overlapping, overcrowding, and convergence are verified before the next iteration starts. After obtaining the node location information, the node is assigned into a cluster based on the nearest location point. In the second phase, the cluster sizes are balanced by moving the UAVs from an excessively large cluster to their nearest cluster. The fitness values for the CH selection in the PSO algorithm are based on the weighted sum of three parameters: the intra-cluster distance, UAV energy level, and inter-cluster distance.

Ganesan *et al.* [39] aimed to provide an effective solution to elect the CH among multiple drones at different periods based on the various physical constraints of the drones. The elected CH acts as the decision maker and assigns tasks to the other drones. In the case where the CH fails, the next eligible drone is re-elected as the leader. The CH is elected dynamically based on the following three parameters: its current position (closer to the BS and all other drones), residual

energy, and velocity, using a method based on PSO. This elected leader CH alone communicates with the BS and other drones, thereby decreasing the communication energy, which in turn extends the network lifetime. Clusters are formed using spider monkey optimization based on the proximity of the drones to one another, their connectivity to other nodes (using RSSI), and the residual energy of the drones. The clusters that are formed will have equal average residual energy for increasing the network lifetime further. In the above study, the simulation demonstrated that the results of the proposed distributed method were more efficient than the results of the PSO-C algorithm, in which the fitness value is based only on the residual energy, in terms of extending the network lifetime by reducing the communication energy consumption.

Azoulay and Reches [12], [13] developed a multi-agent (MA) blackboard-based protocol to form clusters of self-interested UAVs with close routes. According to the proposed protocol, the formed clusters are saved in a centralized blackboard, and any new UAV can either join an existing flock or create a new one and publish it in the centralized blackboard. The simulation results of this approach demonstrated a load reduction in crowded environments.

Another bio-inspired clustering algorithm for highly dynamic large-scale ad-hoc UAV networks was suggested by Yu *et al.* [124], which transplants the foraging model of *Physarum polycephalum* to the field of highly dynamic large-scale ad-hoc UAV networks to adapt to these networks, and conducts cluster formation and maintenance effectively. They demonstrated through simulations that their algorithm outperformed the classical clustering algorithm in terms of the average link connection lifetime and average CH lifetime, which could make the cluster structure more stable.

B. ML METHODS FOR FLOCK FORMATION

As highlighted in the literature, the problem of determining optimal clusters, even in static environments, is known to be NP-hard [2], [73]; thus, the different methods suggested for relatively static or highly dynamic networks are heuristics. Given that ML exhibits good performance in various NP-hard problems [53], several studies have suggested the use of advanced ML methods to handle the clustering challenge. The basic clustering problem is a variation of the dominating set problem [100]. A dominated set of graph G is a subset S of the vertex set of G , in which each vertex in G is either in S or adjacent to a vertex in S . A minimum dominating set is a dominating set of the smallest size in a given graph. The size of the minimum dominating set is known as the domination number of the graph. In wireless sensor applications, each sensor is considered as a node and an edge exists between two nodes if and only if they have a direct connection to one another. Moreover, the set of the CHs is the dominated set of the WSN graph. He *et al.* [43] suggested the use of a Hopfield neural network to minimize the weakly connected dominating set for the self-configuration of WSNs. They explored the convergence procedure by simulating the subsequent model and tested the model on complete graphs to observe the

robustness of the parameters in the model. Furthermore, they presented an improved hill climbing algorithm with directed convergence, heuristic adjustment, and threshold compensation. They demonstrated the efficiency of their method by means of simulation and investigated the effect of the transmission radius on the cluster sizes. The results indicated that it is important to select a suitable transmission radius to ensure that the network has good stability and a long lifespan.

When viewing the clustering and head selection problems along the timeline, the decision in one period may influence the future environmental state and future decision making; thus, RL can be applied to solve the clustering problem over time. Pan *et al.* [88] proposed an energy-aware CH selection method based on binary PSO and K-means to extend the network lifetime. The selection criteria of the objective cost function were based on minimizing the intra-cluster distance as well as the distance between the CHs and BS, and optimizing the energy consumption of the entire network. Moreover, the sensor nodes were divided into several clusters based on the K-means algorithm at the beginning, which could reduce the complexity of the overall algorithm. They demonstrated by simulation that their technique could outperform LEACH-C and PSO-C in terms of the network lifetime.

Yuan *et al.* [129] proposed a genetic algorithm (GA)-based network clustering method that provided a framework for dynamically optimizing wireless sensor clusters. The residual energy, expected energy expenditure, distance to the BS, and number of nodes in the vicinity were used as the arguments for the fitness function, where each GA chromosome represented a designation map of CHs. Given the CHs, the cluster members were formed following the nearest neighbors rule. In each transmission round, the network structure was updated dynamically to achieve a balance of the residual energy of the sensor nodes, and hence, to extend the network longevity.

Tolba and Alarifi [112] proposed an RL-based technique known as adaptive Q-learning (AQL) for clustering in a distributed sensor network to improve the network performance with a minimum energy-overhead tradeoff. AQL operates in two distinct phases, namely those for CH selection and forwarder selection. The decision-making system is used to qualify nodes based on their past behavior over transmission. AQL improves both the inter- and intra-cluster communication. The experimental results demonstrated the effectiveness of the proposed learning technique by improving the network lifetime with a high request-response rate, thereby minimizing the delay, overhead, and request failures.

Finally, Govindara and Deepa [40] considered a WSN for the IoT and proposed a capsule neural network architectural model to achieve better performance by minimizing the network energy overhead for the WSN-based IoT. The clustering process in sensor networks contributes to improving the network quality by controlling the energy consumption rate and improving the data accuracy rate. The above authors aimed to develop a new capsule neural network model that can maintain the network energy at an optimal level, leading to

Clustering

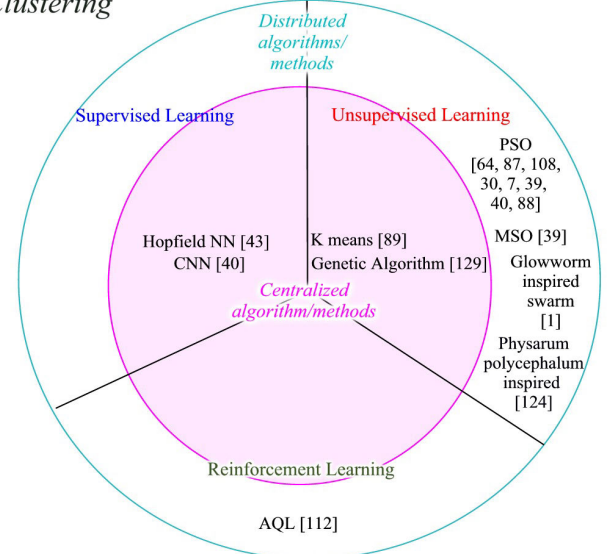


FIGURE 4. ML methods used for the clustering challenge.

improved throughput and higher accuracy with a reasonable network overhead. The capsule neural network architecture is a plausible neural model and has been proven to be effective for routing and optimization operations wherein the activation of the capsule is calculated at the forward pass time.

The idea is to add structures known as “capsules” to a convolutional neural network (CNN) and to reuse output from several of those capsules to form more stable representations (with respect to various perturbations) for higher capsules. The simulation results that were obtained established the reliability and effectiveness of the proposed capsule neural network learning for the energy optimization of the IoT in sensor networks compared to existing methods in the literature.

To clarify the various algorithms used for clustering, Table 4 summarizes the studies that have used bio-inspired and ML methods to solve the clustering challenge, while figure 4 presents the different types of ML methods used for the clustering challenge.

V. FLOCKING STRATEGIES AND UAV COORDINATION

Several challenges in flock management involve the physical location of the flocks. This includes determining the optimal deployment of UAVs, physical coordination of UAVs, and trajectory optimization. In the following section, we survey several recent studies that have suggested the use of ML methods to address the location-related challenges. We focus on certain geographical challenges of flock formation, which are mainly determined by their localization and mobility, to avoid collisions and to achieve energy efficiency. We consider the following related problems: flocking strategies to avoid collisions between flock members, deployment challenges in UAV-based wireless networks, and multi-UAV path planning challenges. Each related problem is discussed in a

TABLE 4. Creation and maintenance of flocks and clusters (C = centralized, D = distributed).

Publication	Challenge	Optimization Criteria	ML Method	Approach
Latiff et al. [64]	Clustering	Intra-cluster distance, Energy stability	PSO	D
Kuila and Jana [87]	Clustering, Routing	Energy conservation, and load balancing	LP/NLP, PSO	D
Suganthi and Rajagopalan [108]	Multi-swarm Construction	energy effectiveness	PSO	D
Collotta et al. [30]	WSN clustering	Battery life	Fuzzy logic PSO	D
Aftab et al. [1]	Drone cluster Management and routing	Clustering construction time and lifetime, energy consumption	Glowworm swarm optimization	D
Arafat and Moh [7]	Localization and clustering	Network lifetime	PSO	D
Ganesan et al. [39]	CH election, Clustering	Link lifetime, cluster lifetime	PSO+MSO	D
Yu et al. [124]	UAV clustering	Connection lifetime, CH lifetime	foraging model of Physarum polycephalum	D
He et al. [43]	WSN clustering	Stability, lifetime	Hopfield NN	C
Pan et al. [88]	Clustering	Energy consumption	PSO+K-means	C/D
Yuan et al. [129]	Clustering	Network lifetime	GA	C
Tolba and Alarifi [112]	WSN clustering	Network lifetime	AQL	D
Govindaraj and Deepa [40]	WSN clustering	Optimal energy Maintenance	CNN	C

separate section and relevant ML methods are surveyed, with each of these challenges discussed and solved.

Given a flying flock of objects, it is important for the flock members to coordinate to avoid cohesion. The basic flocking challenges of any flock of flying objects working independently were defined as follows by Reynolds [94]: the birds attempt to stick together, and to avoid collisions with one another and with other objects in their environment. There are three basic rules for maintaining the flocking behavior, as follows: separation: avoid collisions with nearby flockmates; alignment: attempt to match velocity with nearby flockmates; and cohesion: attempt to stay close to nearby flockmates. Similar to the swarming behaviors that are observed in animals and insects, autonomous coordination among UAVs should be guaranteed. In this section, we consider several studies that have used ML methods to maintain stable flock behavior.

A distributed approach for the flocking problem was suggested by Olfati-Saber [81] in the area of MA dynamic systems, where the agents were autonomous and no centralized manager existed. Olfati-Saber considered this problem and suggested several distributed flocking algorithms for this purpose. Two cases of flocking in free space and the presence of multiple obstacles were considered. Moreover, three flocking algorithms were proposed: two for free flocking and one for constrained flocking. A systematic method was provided for the construction of cost functions (or collective potentials) for the flocking, in which the collective potentials penalized deviation from a class of lattice-shaped objects. Several simulation results were presented to demonstrate the performance of the 2D and 3D flocking, split/rejoin

maneuver, and squeezing maneuver for hundreds of agents using the proposed algorithms.

Maza *et al.* [76], [77] considered civil domains in which multiple aerial robots with sensing and actuation capabilities were available and presented the AWARE project for the autonomous coordination and cooperation of UAVs. The different components of the AWARE platform and scenario in which the multi-UAV missions were carried out were described, including surveillance with multiple UAVs, sensor deployment, and fire threat confirmation. Key issues in multi-UAV systems, such as distributed task allocation, conflict resolution, and plan refinement, were solved in the execution of the missions. Constraint-based temporal planning and scheduling were used for solving the planning problems, and the contract net protocol was used to manage the distributed task allocation process among the different UAVs. They demonstrated by means of experiments with real UAVs that the developed architecture allowed a broad spectrum of missions to be covered: surveillance, sensor deployment, fire confirmation, and extinguishing.

Quintero [92] developed a novel flocking algorithm that enabled multiple UAVs to locate themselves in the flock to distribute a given sensing task among the group members, assuming a leader-follower network topology. They focused on the control policy of the followers, and developed a cost function that is a function of the distance and heading with respect to the leader, as well as a stochastic kinematic model that facilitates flocking. Thereafter, they used dynamic programming to minimize the expected cost of each follower. Quintero *et al.* assumed that there is a predefined flock leader that is known to the entire flock, and their aim was to optimize

the distance and heading of the followers with regard to the leader.

Xu *et al.* [122] proposed a distributed neuro-dynamic flocking design for ensuring that the UAVs follow three heuristic flocking rules, namely cohesion, separation, and alignment, in an optimal manner, using the neuro-dynamic programming (NDP) technique. First, an innovative cost function was developed by combining the system cohesion, separation, and alignment performance. Subsequently, a novel neural network was proposed to approximate the minimized cost function value by using the Hamilton–Jacobi–Bellman equation in an online and forward-in-time manner. Thereafter, near-optimal flocking could be achieved by minimizing the estimated cost function.

Flocking control can be performed by the learning-based model predictive control (LBMPC), developed by Aswani *et al.* [10]. Model predictive control (MPC) is a process control method which aims to satisfy a set of constraints, where LBMPC aims to handle system constraints, and to optimize performance with respect to a cost function, by using statistical identification tools to learn model uncertainties. Aswani *et al.* [9] apply LBMPC to UAV control, and demonstrate by their experiments that the LBMPC yields to less overshoot on the UAV and to better robustness to mis-learning, w.r.t. the referred linear MPC method.

Hung and Givigi [49] proposed the use of a model-free RL method to enhance the autonomous coordination among UAVs in a swarm. They used a leader–follower policy, whereby Peng’s $Q(\lambda)$ [84] with a variable learning rate was employed by the followers to learn a control policy that facilitates flocking. The problem was structured as a Markov decision process (MDP), in which the agents were modeled as UAVs that experience stochasticity owing to disturbances such as winds and control noises, as well as weight and balance issues. The learned policies were compared to dynamic programming methods that were solved using stochastic optimal control. The simulation results demonstrated the feasibility of the proposed learning approach for enabling agents to learn how to flock in a leader–follower topology while operating in a non-stationary stochastic environment.

Tsai [113] focused on vision-based collision avoidance for UAVs. First, images from the UAV cameras were fused based on deep CNNs. Thereafter, a recurrent neural network (RNN) was constructed to obtain high-level image features for object tracking and to extract low-level image features for noise reduction. The system distributed the calculation among the multiple UAVs to perform object detection, tracking, and collision avoidance efficiently. To achieve high overall system performance, the study adopted the half total error rate (HTER) accuracy measurement, which is based on the false rejection rate and false acceptance rate.

Jafrai *et al.* [52] proposed a flocking design framework for MA systems using an RL technique, which was appropriate for real-time implementation. The controller that they developed was based on a bio-inspired RL controller, relying on a computational model of emotional learning in the

mammalian limbic system. They used this model for practical MA systems in the presence of uncertainty and dynamics.

In the work of Venturini *et al.* [115], the authors considered a general MARL framework for the initial exploration and surveillance of a swarm of independent UAVs. Their scheme followed the framework in which observations of other agents are used to make decisions and to avoid collision, thereby encouraging cooperation. They defined a deep Q network algorithm. They subsequently demonstrated its efficiency with limited training, compared it to the look-ahead search heuristic, and showed that the MARL scheme can explore the environment and reach the targets more rapidly. They also performed a transfer learning experiment, demonstrating that agents trained on a different map can learn to adapt to a completely new scenario much faster than when restarting the training from scratch.

Anicho *et al.* [6] compared the performance of RL and SI methods for solving the problem of coordinating multiple high-altitude platform stations (HAPSs) for communication area coverage. Swarm coordination techniques are essential for developing autonomous capabilities for the control and management of multiple HAPSs. The authors observed that the RL approach exhibited superior overall peak user coverage with unpredictable coverage dips, whereas the SI-based approach exhibited lower coverage peaks but better coverage stability and faster convergence rates.

Sharma and Ghose [104] developed several basic swarming laws for UAVs. They demonstrated that when the cohesion rule is applied, an equilibrium condition is reached, in which all of the UAVs settle at the same altitude on a circle with a constant radius. They proposed a decentralized autonomous decision-making approach that could achieve collision avoidance, and developed algorithms with the aid of these swarming laws for two types of collision avoidance, namely group-wise and individual, in the 2D plane and 3D space. Their experimental results demonstrated the effectiveness of their approach, in which self-organized flight cluster collision avoidance was successfully achieved by the UAV swarms.

Wu *et al.* [120] studied the problem of path conflicts for UAV clusters and established a method for calculating the collision probabilities of UAVs under the constraints of the mission space and number of UAVs. In the cluster flight mode, the automatic tracking and prediction of UAV cluster tracks were implemented to avoid path conflicts in the clusters. To address the inconsistency problem owing to noise, a state estimation method based on the Kalman algorithm was proposed. The cluster state prediction and collision probability were calculated to prevent the clusters of formation UAVs conflicting on paths during flight. Finally, the simulation results verified the validity and effectiveness of the proposed method in multi-UAV formation flight planning.

Table 5 and Figure 5 summarize the studies relating to flocking and collision avoidance among UAVs using ML methods.

TABLE 5. ML applied to flocking and coordination.

Publication	Challenge	Optimization Criteria	ML Method
Olfati-Saber [81]	Flocking challenges	connectivity, energy	Distributed flocking algorithms
Maza et al. [76], [77]	Architecture	execution cost	Temporal planning, contract net
Quintero [92]	Localization	distance and heading	DP
Xu et al. [122]	Ensuring flocking rules	cohesion, separation, and alignment	NDP
Hung and Givigi [49]	Coordination	flocking cost function	Q-learning
Tsai [113]	Vision-based collision avoidance	HTER	RNN
Jafrai et al. [52]	Flocking design	multi-objective properties	Bio-inspired RL
Venturini et al. [115]	Exploration and surveillance	target reaching efficiency	Deep Q-learning
Anicho et al. [6]	Coordinating	coverage	RL and SI
Sharma and Ghose [104]	Collision avoidance	swarm size and stability	Swarm laws Decentralized alg.
Wu et al. [120]	object tracking	noise, collision prediction	Kalman algorithm

VI. UAV DEPLOYMENT

An important challenge in UAV based wireless communication network is the challenge of UAVs deployment. In this UAV based network, the UAVs function as BSs, and their deployment affects the network efficiency and coverage. In fact, the problem of determining the optimal placement of ABSs to maximize the coverage is known to be a NP-hard problem [74], and in this section, we survey several ML-based methods for addressing this challenge.

Park *et al.* [83] considered the problem of the optimal deployment of multi-UAVs to provide high throughput for users with different requirements in the UAV-BS environment. Based on the air-to-ground path loss model, the virtual communication environment was established, in which air-time fairness was applied for equitable time distribution of the channel usage according to user requirements. RL was applied to determine the best UAV positions, and a collaborative algorithm with modified K-means was employed to distribute users to each UAV to solve communication overload problems.

In the study of Klaine *et al.* [56], the aim was to maximize the number of users covered by the system in an emergency scenario, where the drones were limited by both backhaul and RAN constraints. For this scenario, Klaine *et al.* proposed the use of RL to determine the optimal position of the UAVs. The proposed solution was compared to different positioning strategies, such as deploying the drones in fixed random positions, fixing the drones around a circle centered in the middle of the area at evenly spread angles, or deploying the drones in the locations of hotspots of a previously destroyed network. The results demonstrated that the Q-learning solution outperformed all other methods in all considered metrics. Liu *et al.* [67] proposed a deep RL (DRL) variation known as the deep deterministic policy gradient method for energy-efficient UAV control in the context of providing communication coverage for ground users.

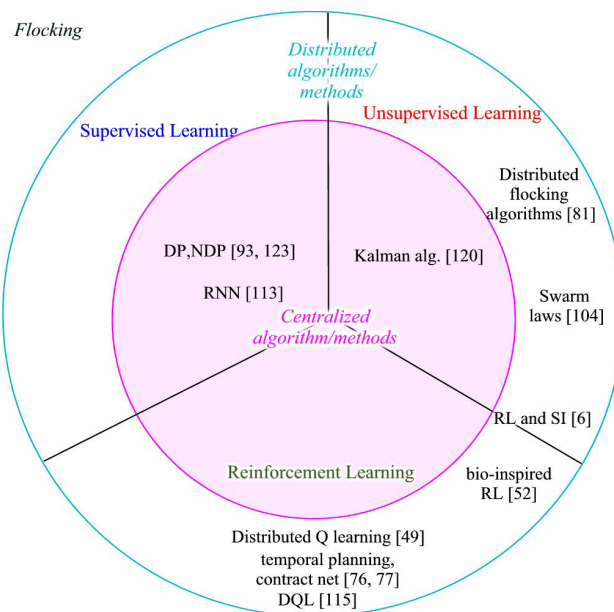


FIGURE 5. ML methods used for the flocking challenges.

The control policy considers the UAV movements in each time slot and the aim is to optimize the communication coverage, fairness, energy consumption, and connectivity.

The deployment decision becomes more complicated when the environment itself, the coverage map in particular, is initially unknown, and the effect of each UAV position on the coverage is not known in advance. Zhang *et al.* [134] suggested an ML framework based on a Gaussian mixture model and a weighted expectation maximization algorithm to predict the potential network congestion. Based on the predicted congestion and traffic, the optimal deployment of the UAVs could be achieved in a manner that minimized the transmit power required to satisfy the communication demand of users in the downlink, while minimizing the power required for the

UAV mobility. Qiu *et al.* [91] considered the NP-hard problem of maximizing the coverage rate of N ground users by the simultaneous placement of multiple ABSs with a limited coverage range. In their study, Qiu *et al.* applied the DRL method for representing the state by a coverage bitmap to capture the spatial correlation between ground users and ABSs, and for effectively learning the action and reward function given dynamic interactions with the complicated propagation environment.

In the study of Liu *et al.* [70], the maximization criteria consisted of the ground user opinions. They formulated the problem of the joint non-convex 3D deployment and dynamic movement of the UAVs, where the goal was to maximize the sum of the mean opinion score (MOS) of the ground users. They demonstrated that the problem was NP-hard and proposed a Q-learning-based solution to handle the problem. An algorithm based on a combination of the GA and K-means was used to obtain the cell partition of the users. Subsequently, a Q-learning-based deployment algorithm was proposed to achieve 3D placement of the UAVs when the users were static. Finally, a Q-learning-based movement algorithm was presented to obtain the 3D dynamic movement of the UAVs. Several recent studies have suggested different types of learning to accelerate the deployment process. Liu *et al.* [68] developed a fast positioning algorithm for the deployment of UAVs serving as BSs, which achieved the goal of maximizing the sum of the downlink rates in the multi-UAV communication network. They designed a geographical position information (GPI) learning algorithm to learn the GPI relationship between the users and UAVs and demonstrated that the GPI learning method enabled rapid suboptimal deployment in the case of multiple UAVs and users.

Another aspect that can be considered in optimal UAV deployment is the expectation of user requests.

Liu *et al.* [18] considered the problem of navigation control for a group of UAVs that served as mobile BSs. The UAVs were supposed to fly around a target area to provide long-term communication coverage for ground mobile users. Liu *et al.* designed a decentralized DRL-based framework to control each UAV in a distributed manner. Their goals were to maximize the temporal average coverage score achieved by all UAVs in a task, maximize the geographical fairness of all considered points of interest (PoIs), and minimize the total energy consumption, while maintaining the UAVs connected and located inside the area borders. They designed the state, observation, action space, and reward in an explicit manner and modeled each UAV using deep neural networks (DNNs).

In the remainder of this section, we consider studies that aimed to deal with a joint optimization problem, namely the deployment challenge that considers user association, power control, and other optimization problems that are involved in implementing a UAV-based wireless network.

Dai *et al.* [37] investigated the problem of the efficient deployment of UAVs while guaranteeing the quality-of-service requirements. The UAV played the role of a

coordinator to provide a high-quality communication service for ground users as well as to maximize the benefits of caching. They proposed an RL-based approach to solve the multi-objective deployment problem, while maintaining an optimal tradeoff between the power consumption and backhaul saving. Owing to the interdependent relationships among different UAVs, they adopted RL based on the local search approach to determine the 3D placement, minimum transmit power, and cache strategy of each UAV. Finally, the minimum number of UAVs was provided together with the efficient deployment scheme.

In another study, Dai *et al.* [36] leveraged game theory to model the problem of imperfect channel state information and proposed a robust and distributed learning algorithm to identify the multi-UAV plan flight paths while simultaneously gathering local information until the optimal deployment location was determined. They demonstrated via simulation that the suggested distributed probabilistic learning algorithm converged to a stochastic stable state, which maximized the optimization objective; that is, the sum of the alpha-fairness of all ground terminals.

Chen *et al.* [26] studied the problem of the proactive deployment of cache-enabled UAVs for optimizing the quality-of-experience (QoE) of wireless devices in a cloud RAN. In the considered model, the network could leverage human-centric information such as the visited locations, requested contents, gender, job, and device type of users to predict the content request distribution and mobility pattern of each user. Subsequently, given these behavior predictions, the proposed approach sought to determine the user-UAV associations, optimal UAV locations, and contents to cache at the UAVs. The problem was formulated as an optimization problem, the goal of which was to maximize the QoE of the users while minimizing the transmit power used by the UAVs. To solve this problem, they proposed the use of conceptor-based echo state networks (ESNs). The ESNs could effectively predict the content request distribution and mobility pattern of each user when limited information on the states of the users and network was available. Based on the predictions of the content request distributions and mobility patterns of the users, the user-UAV association, UAV locations, and contents to cache at the UAVs could be optimally determined.

Several studies considering the deployment challenge have suggested the use of clustering algorithms to cluster the ground user, and subsequently to determine the UAV deployment based on the clustered users. Several studies proposed running this scheme iteratively until convergence is reached. Kang *et al.* [97] proposed a cluster-based UAV deployment scheme to reduce the data traffic as well as service delays and to improve the coverage of BSs. User groups were formed using the K-means clustering algorithm. Thereafter, the optimal UAV locations were determined given the user clusters. Furthermore, a long short-term memory (LSTM)-based caching scheme was proposed to cache the popular contents on UAVs. Yang *et al.* [127] considered the sum power

minimization problem via jointly optimizing the user association, power control, computation capacity allocation, and location planning in a UAV-based network. They proposed a low-complexity algorithm for solving these subproblems iteratively. The compressive sensing-based algorithm was proposed for the user association subproblem. For the computation capacity allocation subproblem, the optimal solution was obtained in a closed form. For the location planning subproblem, the optimal solution was effectively obtained via a 1D search method. Finally, to obtain a feasible solution for this iterative algorithm, a fuzzy C-means clustering-based algorithm was proposed.

Koushik *et al.* [58] studied an aerial network in which certain UAVs were used as gateway nodes to aggregate the data from other UAVs and send these to a nearby aircraft that acted as a control node for the UAVs in the swarm. To handle the dynamic swarm topology and time-varying link conditions, they designed the following solution: A deep Q-learning algorithm was used to determine the optimal links between nodes, and an optimization algorithm was applied to fine-tune the position of the UAV node locally to optimize the overall network performance.

Finally, Nguyen *et al.* [62] considered both the deployment and resource allocation challenges for a distributed UAV wireless network in disaster situations. They proposed a rapid user clustering model based on the K-means procedure for the UAV network. Furthermore, they proposed distributed real-time power allocation to maximize the end-to-end sum rate and embedded programming into the UAV devices for rapid recovery of the network to support a large number of users in disaster communication.

It can be concluded that the various studies have considered different aspects of the UAV deployment challenge. They differ in terms of the maximization criteria, data availability, and additional problems that are considered jointly with the deployment decision. Table 6 and Figure 6 summarize the studies dealing with deployment, path planning, and collision avoidance using ML methods.

VII. TRAJECTORY OPTIMIZATION AND NAVIGATION

In this section, we extend our survey by reviewing the geographical challenges of flock management. We deal with the trajectory and path planning challenges and review recent studies that have aimed to address these challenges using ML methods. Trajectory planning refers to moving from point A to point B without collisions, which can be computed using both discrete and continuous methods. Trajectory planning is considered as a major topic in robotics as it paves the way for autonomous vehicles.

Several recent studies [93], [95], [96] considered the issue of path planning and collision avoidance in UAV flocks using the MDP. Bayerlein *et al.* [14] leveraged the use of RL, where the UAV acted as an autonomous agent in the environment, to learn the trajectory that maximized the sum rate of the transmission during the flying time. Movement decisions were made directly by a neural network. The algorithm did

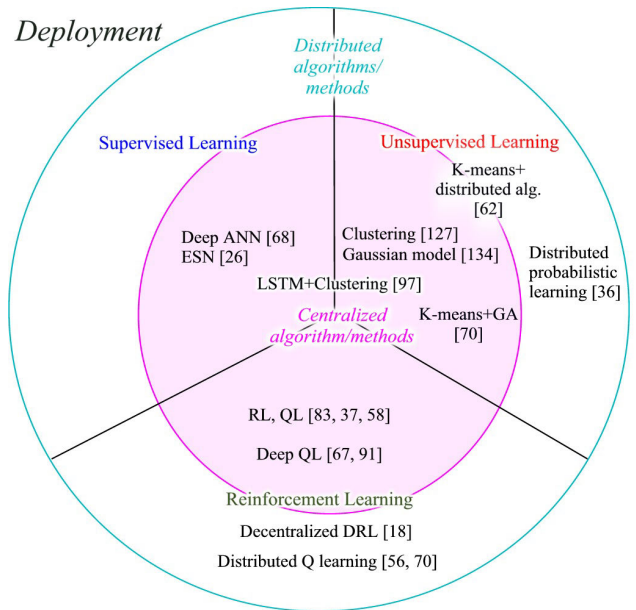


FIGURE 6. ML methods used for the deployment.

not require explicit information regarding the environment and could learn the network topology to improve the system-wide performance.

Zeng *et al.* [131] formulated the UAV trajectory optimization problem as the minimization of the weighted sum of its mission completion time and the expected communication outage duration. They demonstrated that the formulated problem could be transformed into an equivalent MDP. Thereafter, it was proposed to use an RL technique, such as the classical Q-learning approach, to learn the UAV action policy, namely the flying direction in this context. As the output of the MDP involves a continuous state space that essentially has an infinite number of state-action pairs, they applied a DNN to approximate the Q function. They used the dueling network architecture with multi-step learning to train the DNN. The proposed DRL-based trajectory design did not require any prior knowledge regarding the channel model or propagation environment. It only utilized the raw signal measurement at each UAV as the input to improve its radio environmental awareness.

In the study of Hu *et al.* [51], a decentralized RL method was adopted to solve the UAV trajectory design problem. To coordinate multiple UAVs performing real-time sensing tasks, they first proposed a sense-and-send protocol, and analyzed the probability of successful valid data transmission using nested Markov chains. Thereafter, they formulated the decentralized trajectory design problem and propose an enhanced multi-UAV Q-learning algorithm to solve this problem. The simulation results demonstrated that the proposed enhanced multi-UAV Q-learning algorithm converged faster and achieved higher utilities for UAVs in real-time task-sensing scenarios.

Multi-ant colony optimization (ACO) was suggested by Cekmez *et al.* [20] for the obstacle avoidance path planning

TABLE 6. ML application for deployment.

Publication	Challenge	Optimization Criteria	ML Method
Park <i>et al.</i> [83]	Deployment	throughput	RL, K-means
Klaine <i>et al.</i> [56]	Deployment	coverage	Distributed Q-learning
Liu <i>et al.</i> [67]	UAV movements	coverage, fairness, energy consumption, and connectivity	DRL
Zhang <i>et al.</i> [134]	Congestion prediction	power minimization	Gaussian model
Qiu <i>et al.</i> [91]	Action and reward function	coverage rate	DRL
Liu <i>et al.</i> [70]	Deployment and movement	mean opinion score	Q-learning, GA, K-means
Liu <i>et al.</i> [68]	Deployment	sum rate	DNN
Liu <i>et al.</i> [18]	Navigation	coverage, fairness, energy minimization	Decentralized DRL
Chen <i>et al.</i> [26]	User requests and mobility prediction	QoE	Echo state network
Dai <i>et al.</i> [37]	Deployment	quality of service	RL + local search
Dai <i>et al.</i> [36]	Path planning, deployment	alpha-fairness	Probabilistic distributed learning
Kang <i>et al.</i> [97]	Deployment, caching,	reducing delays improved coverage	Clustering, LSTM
Yang <i>et al.</i> [127]	User association, power control, computation capacity allocation, and location planning	power minimization	Iterative algorithm, clustering
Koushik <i>et al.</i> [58]	Routing and deployment	throughput	Q-learning, Optimization
Nguyen <i>et al.</i> [62]	Deployment and resource allocation	throughput	K-means, Distributed algorithm

challenge. According to the multi-ACO method, numerous ant colonies attempt to determine an optimal solution cooperatively by exchanging their valuable information with one another. Cekmez *et al.* aimed to implement obstacle avoidance UAV path planning using this algorithm, and they experimentally demonstrated that this approach achieved effective path planning for UAVs compared to a single-colony ACO approach.

A. TRAJECTORY DESIGN WITH ADDITIONAL CHALLENGES

In recent studies, the UAV trajectory challenge in wireless communication networks has also been discussed in combination with the related resource allocation challenges.

Challita *et al.* [21], [22] considered a UAV cellular network and aimed to minimize the interference caused by the ground network, as well as the wireless transmission latency. They modeled the problem as a distributed multi-UAV game, in which the objective of each UAV was to learn its path, transmit power level, and association vector autonomously and jointly. For the proposed game, the cell association vector, trajectory optimization, and transmit power level of the UAVs were closely coupled with one another and their optimal values varied according to the network dynamics. Challita *et al.* suggested the use of DRL based on ESN cells for optimizing the trajectories of multiple cellular connected UAVs in an online manner to achieve convergence to equilibrium. The deep ESN architecture was trained to allow each UAV to map each observation of the network state to an action, with the aim of minimizing a sequence of

time-dependent utility functions. Each UAV used the ESN to learn its optimal path, transmit power level, and cell association vector at different locations along its path. The proposed algorithm was demonstrated to reach subgame perfect Nash equilibrium (SPNE) upon convergence.

Liu *et al.* [121] addressed the challenge of the joint trajectory design and power control of UAVs that served as BSs. Their goal was to improve the user throughput, while satisfying the user rate requirement. Their solution was based on an MA Q-learning based placement algorithm to determine the initial deployment of the UAVs. Furthermore, an ESN-based prediction algorithm was used to predict the mobility of users and an MA Q-learning-based method was considered for the trajectory acquisition and power control algorithm for the UAVs.

The objective in the study of Khamidehi *et al.* [54] was to determine the trajectory of multiple ABSs, such that the sum rate of the users served by each ABS was maximized. For this purpose, along with the optimal trajectory design, they also considered the optimal power and subchannel allocation, which are of great importance for supporting users with the highest possible data rates. Thus, they divided the problem into two subproblems: ABS trajectory optimization, and joint power and subchannel assignment. Subsequently, they developed a distributed Q-learning-based algorithm to solve these subproblems efficiently, which did not require a significant amount of information exchange between the ABSs and core network. The simulation results demonstrated that although Q-learning is a model-free RL technique, it has a remarkable capability to train ABSs to optimize their trajectories based

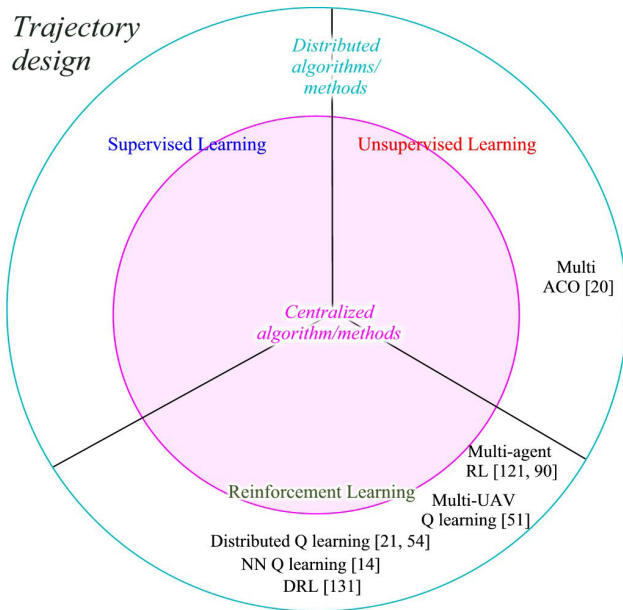


FIGURE 7. ML methods used for trajectory design.

on the received reward signals, which carry information from the network topology.

Qie *et al.* [90] considered a multi-UAV target assignment scenario, where a flock of UAVs was supposed to fly to targets that were distributed at different locations with the shortest total flight distance, with certain fixed threat areas that the UAVs could not enter. Of course, collision avoidance between UAVs was required, and it was assumed that there was only one type of target and all UAVs were identical. This combinatorial optimization problem included two main subproblems: target assignment and path planning. They proposed a simultaneous target assignment and path planning (STAPP) method based on an MA deep deterministic policy gradient (MADDPG) algorithm, which is a type of MARL algorithm. In STAPP, the multi-UAV target assignment problem was first constructed as an MA system. Thereafter, the MADDPG framework was used to train the system to solve the target assignment and path planning simultaneously, according to a corresponding reward structure. The proposed system could deal with dynamic environments effectively, as its execution required only the locations of the UAVs, targets, and threat areas.

Table 7 and Figure 7 summarize the part of our survey related to the trajectory design of UAVs using ML methods.

VIII. RESOURCE ALLOCATION IN AERIAL FLOCKS

A challenging issue when considering flock management is the allocation of communication resources among the flock members, given that the flock members organize an ad-hoc mesh network for their internal and external communication tasks. Note that D2D network challenges arise in FANETs as well as in other ad-hoc networks, such as VANET, fixed sensor networks, and even any wireless network. Thus, in this section, we survey not only studies that have considered

FANETs, but also those that can easily be applied to FANETs. We review these studies for the sake of completeness, as we ought to provide designers of UAV flocks with all of the elements that can be used to develop solutions to their specific scenarios.

Many types of resources can be allocated among flock members, as described in the following:

- **Transmission time:** This can be managed via time-division multiple access (TDMA), which allows several users to share the same frequency channel by transmission at different time slots.
- **Frequency channel:** This can be managed via frequency-division multiple access (FDMA), which consists of dividing the available bandwidth into several non-overlapping frequency channels, where each channel is allocated to a different user.
- **Power control:** This is important to achieve the optimal transmission power for each D2D transmission so as to reduce the interference on other concurrent transmissions, while maintaining a sufficiently strong received signal at the target receiver. It is also important for avoiding the near-far problem (mainly in the CDMA technique) and saving battery life.
- **Multiple input, multiple output (MIMO) antennas:** These enable the control of a new dimension, namely the possibility of focusing on a specific target via beam-forming, and also improve the received signal via signal processing.

Three different situations arise when considering resource allocation problems in UAV flocks or other D2D mesh networks:

- The cluster is completely covered by a ground BS.
- The cluster is partially covered by a ground BS.
- The cluster does not detect any ground BS in the entire area.

Clearly, the most difficult situations are those in which no BS is detected, as well as those that may occur in uninhabited areas, areas following a disaster, or for flying flocks. In such situations, all of the decisions should be made by the flock members, with no involvement from external management.

Another level of distinction is in the goal function that one is attempting to maximize. Cui *et al.* [34] specified four fundamental energy-efficient metrics that have received the most research attention in the literature:

- GEE: the system global energy efficiency;
- WSEE: the weighted sum energy efficiency;
- WPEE: the weighted product energy efficiency; and
- WMEE: the weighted minimum energy efficiency.

Additional optimization criteria include::

- maximizing the system throughput;
- maximizing the quality of service, in a manner that is compatible with the request types;
- minimizing the average latency; and
- minimizing the maximal latency of any device.

TABLE 7. ML application for trajectory design.

Publication	Challenge	Optimization Criteria	ML Method
Bayerlein <i>et al.</i> [14]	Trajectory	sum rate	NN Q-learning
Zeng <i>et al.</i> [131]	Trajectory optimization	Weighted sum of mission completion	DRL
Hu <i>et al.</i> [51]	Trajectory design	successful Transmissions	Multi-UAV Q-learning
Cekmez <i>et al.</i> [20]	Path planning	fitness	Multi-ACO
Challita <i>et al.</i> [21]	Subgame perfect equilibrium	routing, power	Distributed DRL
Liu <i>et al.</i> [121]	Trajectory design and power control	Throughput	MA Q-learning
Khamidehi <i>et al.</i> [54]	Trajectory	sum rate	Distributed Q-learning
Qie <i>et al.</i> [90]	Target assignment	Shortest flight distance	MA RL

In general, the common resource allocation problems are known to be NP-hard [2], [73]. Thus, over the years, various suboptimal resource allocation algorithms have been developed to deal with such problems [35], [106].

Another difference among resource allocation policies is how decisions are made regarding the distribution of resources, for which three decision-making styles are used:

- **Centralized resource management:** One of the UAVs, known as the CH, is responsible for the decision-making process. Centralized management can also be implemented by the ground BS in situations where the cluster is covered or partially covered by such a BS.
- **Decentralized decision-making process:** The decision-making process is handled in a distributed manner by the network/flock members. Coordination among the decisions of the members is achieved using spectrum sensing and message passing.
- **Hybrid approach:** Decisions are made in a manner that combines the distributed decisions of the flock members with the central management of the manager.

In the remainder of this section, we describe several studies that have leveraged ML methods to address the challenge of resource allocation in ad-hoc networks. For each considered study, we mention the challenge and the main details of the ML method or methods used to address the challenge. We structure the section according to the ML method used to solve the resource allocation problem.

A. DEEP LEARNING-BASED METHODS

Several recent studies in communication resource management have suggested the use of deep learning models to determine efficient optimal solutions to the scheduling and control decisions made in a UAV wireless network. Using deep learning schemes, a multilayer neural network is trained, where the network inputs are the network state (in a given representation), and the output that should be trained is the resource allocation decision. A DNN can be trained either with a supervised training scheme using resource allocation solutions that are calculated using any optimization method, or with an unsupervised scheme by calculating the value of

the neural network output and optimizing this value by changing the neural network weights. In the following, we describe several works of both types.

Sun *et al.* [109] proposed a deep learning-based scheme for real-time resource management over interference-limited wireless networks. In particular, they suggested the use of deep learning to approximate the behavior of a certain optimization algorithm. Their theoretical results indicated that it is possible to learn a well-defined optimization algorithm highly effectively by using finite-sized DNNs. To demonstrate their claim, they constructed such a DNN for the power control problems over either the IC or the IMAC channel and trained the DNN to approximate the behavior of the WMMSE algorithm effectively [106].

Cui *et al.* [34] used a deep-learning approach for the scheduling of interfering links in a dense wireless network with full frequency reuse. They employed a DNN architecture that takes the geographic spatial convolutions of the interfering or interfered with neighboring nodes as the input over multiple feedback stages to learn the optimal solution. To ensure fairness, they proposed a scheduling approach that uses the sum rate optimal scheduling algorithm over subsets of links for maximizing the proportional fairness objective over the network. In their study, Cui *et al.* suggested two methodologies for the neural network training: the first is a supervised learning process, in which the network is trained by suboptimal solutions to the resource allocation problem, as computed by FPLinQ [105], which is an algorithm based on the fractional programming approach, and the second is an unsupervised training process, in which the sum rate is supposed to be maximized. According to their comparison, the unsupervised sum rate scheme outperformed the supervised learning scheme for layouts containing links with similar distances.

Matthiesen *et al.* [75] developed a deep learning system for energy-efficient power control in wireless networks. They used an optimal reduced-complexity branch-and-bound procedure to determine the globally optimal power policy for a large set of situations, and subsequently used the solved set as the training set for a DNN. The numerical results demonstrated that the proposed method achieved near-optimal

performance in suggesting the optimal power control policy. Azoulay *et al.* [11] used a DNN to learn the optimal power control and to request scheduling in a MANET cluster, where the optimization problem was mathematically defined and the DNN was trained by examples that were solved by an optimal solver. The simulation results revealed that the DNN could serve as a computationally inexpensive component of expensive optimization algorithms in real-time tasks, providing a very good approximation solution for these problems with a very low average run time. Ahmed *et al.* [3] developed a deep learning-based resource allocation model with the objective of maximizing the total network throughput by performing joint resource allocation for both the power and channel. They used a supervised learning approach to train the model and obtained the training data by solving the non-convex optimization problem using a GA.

However, despite the similarities between WSNs and UAV networks, certain differences should be considered, such as the high mobility and large distances of the UAVs in an ad-hoc UAV wireless network. Thus, in the remainder of this section, we focus on ML methods that have been used for resource allocation in UAV wireless networks.

Chen *et al.* [27] considered the problem of joint caching for a network of cache-enabled UAVs that served wireless ground users over the long-term evolution (LTE) licensed and unlicensed (LTE-U) bands. They proposed a distributed algorithm based on the ML framework of the liquid state machine (LSM) to predict the user content request distribution and to enable the UAVs to select the optimal resource allocation strategies that maximized the number of users with stable queues depending on the network states autonomously. The optimization problem that was solved was a combined user association, spectrum allocation, and content caching problem.

CNN methods were used by Lee *et al.* [65] to learn the efficient transmit power control strategy for maximizing either the spectral efficiency or energy efficiency. Moreover, they proposed a form of deep power control that could be implemented in a distributed manner with local channel state information, allowing the signaling overhead to be greatly reduced. Through simulations, they demonstrated that the deep power control could achieve almost the same or even higher spectral efficiency and energy efficiency compared to a conventional power control scheme, with a much lower computation time.

In summary, when considering the use of deep learning for any resource allocation management, the above related studies have suggested: (a) how to present the input, (b) the value to be optimized, and (c) how to train the network. Given all of the above, the recent studies succeeded in solving different scheduling and resource allocation problems by using deep learning with algorithms based on steps (a) to (c). Previous works have demonstrated the benefit of the deep learning approach in solving complex optimization problems, and in particular, complex resource allocation problems.

B. RL AND DRL

Owing to the high dynamics of UAV networks, several studies have considered RL-based methods to manage the resource allocation in these environments. RL methods can consider the effect of each episodic decision on the future network behavior. This is particularly important when the goal is to maximize an objective function such as the quality of service or latency, where the aim is to avoid long queues and to fulfil the user requirements as soon as possible, because in this situation, each decision may be crucial to avoid long queues in the future system behavior as a result of current decisions. RL and DRL methods can be applied to enable resource management decisions that consider the system behavior over time.

Zhang *et al.* [132] proposed an efficient energy and radio resource management framework based on intelligent power cognition for solar-powered UAVs. In particular, they suggested the use of RL for adjusting the energy harvesting, information transmission, and flight trajectory to improve the utilization of solar energy, with two primary goals: staying aloft over a long period and achieving high communication performance.

Hu *et al.* [46] suggested the use of RL for the following three essential parts in the cellular UAV network: protocol design, trajectory control, and resource management. They presented a distributed sense-and-send protocol to coordinate the UAVs, proposed the use of a Q-learning algorithm for the trajectory control and resource management, and introduced different types of RL approaches and their applications, including the user association, power management, and subchannel allocation.

Cao and Liang [123] proposed a DRL framework for channel and power allocation in a UAV communication system, in which the UAVs were used as BSs. With the proposed framework, the UAV stations could allocate both the channels and transmit power for the uplink transmission of the IoT nodes, with the aim of maximizing the minimum energy efficiency of the IoT nodes.

Ciu *et al.* [32], [33] addressed the challenge of the autonomous resource allocation of multiple UAV communication networks with the goal of maximizing the long-term rewards. To model the uncertainty of the environments, they formulated the resource allocation problem as a stochastic game, in which each UAV becomes a learning agent and each resource allocation solution corresponds to an action taken by the UAVs. They developed an MARL-based resource allocation algorithm to solve the formulated stochastic game of multi-UAV networks. Specifically, each UAV, as an independent learning agent, discovered its best strategy according to its local observations using the Q-learning algorithm. All agents implemented a decision algorithm independently but shared a common structure based on Q-learning. They provided convergence proof of the proposed MARL-based resource allocation algorithm, and demonstrated by simulations that the proposed MARL-based resource allocation

algorithm for the multi-UAV networks could achieve a trade-off between the information exchange overhead and system performance.

C. DISTRIBUTED AND MA-BASED METHODS

We proceed by surveying ML distributed methods that can be applied for resource allocation in UAV wireless networks.

Zeng *et al.* [130] used a distributed swarm approach for the power control and scheduling of UAV flocks. In particular, they used distributed federated learning (FL) algorithms within a UAV swarm that consisted of a leading UAV and several following UAVs. In their study, the UAVs trained their ability to share their local FL with the flock leader.

Chen *et al.* [24] considered the problem of spectrum access in multi-UAV networks using a game theoretic perspective. They formulated the demand-aware joint channel slot selection problem as a weighted interference mitigation game, and subsequently designed a utility function considering the features of the multi-UAV network; for example, certain rewards based on the channel and slot selection. They proved that the formulated game was an exact potential game with at least one pure-strategy Nash equilibrium.

Thereafter, they applied the distributed log-linear algorithm to achieve the Nash equilibrium, and proposed a low-complexity and realistic channel and slot initialization scheme for the UAVs to accelerate the convergence. They demonstrated by simulation that the network utility and aggregate interference level that were achieved from the learning solution were very close to the best Nash equilibrium.

Liu *et al.* [69] studied the problem of downlink power allocation in UAV-based wireless networks with ABSs. They proposed a price-based optimal power allocation scheme and modeled the interaction between the UAVs and ground users as a Stackelberg game. As the leaders of the game, the UAVs selected the optimal power price to maximize their own revenue. Each ground user that belonged to the networks selected the optimal power strategy to maximize its own utility. To solve the equilibrium program with the equilibrium constraint problem of the Stackelberg game, they investigated the lower equilibrium of the ground users and proposed a distributed iterative algorithm to obtain the equilibrium of the lower game.

A distributed power control problem in a dense UAV network was formulated as a mean-field game (MFG) by Zhang *et al.* [136]. They proposed a two-stage scheme to alleviate the congestion of data traffic and interference effects. In the proposed scheme, a high-altitude platform (HAP) performs semi-persistent scheduling and allocates time frequency resources in a non-orthogonal manner, while the UAVs autonomously perform distributed power control. They formulated the centralized resource allocation problem as a roommate matching problem and developed a novel time slot allocation algorithm to solve it. Thereafter, the distributed power control of massive UAVs was formulated as an MFG, which was solved based on the finite

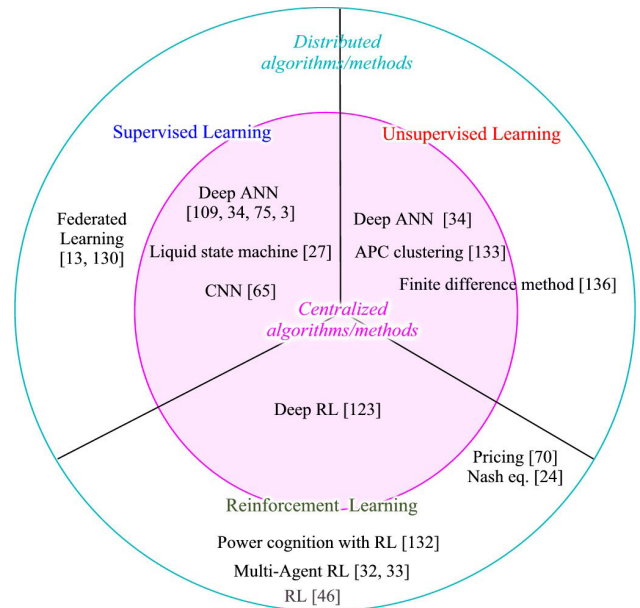


FIGURE 8. ML methods used for resource allocation.

difference method. The simulation results demonstrated that the proposed scheme could significantly improve the reliability of the communication in dense UAV networks.

In a recent study, Zhang *et al.* [133] proposed a potential game-based power control algorithm to optimize the power allocation of the UAVs. The power control problem was proven to be an exact potential game in which the Nash equilibrium was reached. First, the rule of determining neighbors was defined by fixing a threshold. Thereafter, the power strategies were updated continuously until convergence was reached and the final optimal power allocation result was obtained. They proposed a clustering algorithm by means of affinity propagation clustering (APC). The UAVs and their associated users were clustered into groups according to the mutual interference, whereby UAVs who potentially received severe interference from one another had a stronger tendency to join the same cluster. The APC also provided an unsupervised learning technique to determine the number of clusters and CHs to reduce interference automatically.

Table 8 and Figure 8 summarizes the different studies, with their main goal, ML method that was used to achieve the main goal, how the method was tested and the validation thereof, and whether the project code can be downloaded.

IX. TASK ALLOCATION IN UAV FLOCKS

The problem of multi-type UAV task allocation involves determining the best combination of UAVs to cover certain targets.

Huang *et al.* [63] considered the multi-type UAV task allocation problem, which dealt with the question of how to determine the best combination of UAVs to cover certain targets. They assumed that certain amounts of resources were required by the targets and that the UAV formations had to meet the resource demands of the targets. They presented

a multi-type UAV coordinated task allocation method based on cross-entropy. Their study focused on the task allocation situation in which different types of UAVs were assigned to several tasks and the tasks required certain resources. The cross-entropy method obtained random samples from the candidate solutions and then used them to update the allocation probability matrix.

In [86], the authors presented an algorithm for the routing of multiple UAVs and UAV swarms to a set of locations while meeting the constraints of the time on target, total mission time, enemy radar avoidance, and total path cost optimization. They aimed to develop a massive array of autonomous UAVs that were capable of working together towards a common goal. They used a combinatorics problem known as the vehicle routing problem with time windows (VRPTW) to model the routing situation of the UAVs.

In [80], a planning approach for a platform composed of multiple heterogeneous unmanned aerial systems (UASs) was presented. The research activities were focused on the interoperability, task allocation, and task planning problems within the system. A consolidated planning technique was applied to generate low-level plans from a mission described in C-BML automatically, thereby filling the gap between the interoperability and automatic plan generation for missions in which multiple heterogeneous aerial platforms, both simulated and real, were involved. The architecture integrated symbolic, geometric, and task allocation planners. The approach was tested in missions involving multiple surveillance and 3D map generation tasks with a team of simulated and real UASs.

Hu *et al.* [47] presented a new architecture for UAV clustering to enable efficient multi-modal, multi-task offloading. They considered the problem in which the user distribution changes dynamically in real time. Thus, the deployment of static/fixed edge computing nodes in state-of-the-art networks failed to meet their computing demands. They constructed a model known as UAV-M3T for vehicular VR/AR gaming. The UAV-M3T architecture used AI-based decision making for the collaborative optimization of the UAV team and network resources to improve the task performance and resource efficiency of the UAVs. They proposed an AI-based decision-making framework to facilitate the UAV cooperation and joint optimization of the computing, caching, and communication resources. The deployment of UAV clusters is performed in real time, using RNN based on real-time perception.

A. REINFORCEMENT LEARNING METHODS

In [107], a review that focused on RL techniques for dynamic task scheduling was presented. The results of the study were addressed by means of the state of the art in RL used in dynamic task scheduling and a comparative review of those techniques. Although it did not focus on UAVs, it was connected to several studies in that area.

Shamsoshoara *et al.* [102] studied the problem of the limited spectrum in UAV networks and suggested a relay-based

cooperative spectrum leasing scenario in which a group of UAVs in the network cooperatively forwarded data packets to a ground primary user for spectrum access. Each UAV either joined a relaying group to provide a relaying service for the or performed data transmission to the UAV fusion center. MARL was used for the task allocation among UAVs and simulation results were presented to verify the convergence to the optimal solution. The proposed method is a general form of the Q-learning algorithm for a distributed MA environment.

Kim and Morrison [55] jointly determined the system resources (system design), task allocation, and waypoint selection in a stochastic context. They formulated the problem as an MDP and employed DRL to obtain state-based decisions. The study of [137] also used the MDP model. These authors presented a Q-learning-based fast task allocation (FTA) algorithm through neural network approximation and prioritized experience replay, which effectively offloaded the online computation to an offline learning procedure. Specifically, in the proposed approach, a Q network was developed that encoded the allocation rules. The Q network could handle totally different tasks by learning the allocation rules. The task allocation problem for heterogeneous UAVs in the presence of environmental uncertainty was formulated as an MDP. Subsequently, a Q-learning-based FTA algorithm using neural network approximation and prioritized experience replay was presented.

The study of [126] proposed a task-scheduling algorithm based on RL, which enabled the UAV to adjust its task strategy automatically and dynamically using its calculation of the task performance efficiency. A decentralized networking protocol was presented to coordinate the movement of UAVs and to achieve real-time networking of UAV clusters. The expansion strategy was applied to solve the problem of UAV networking in the initial state and DRL was employed to solve the dynamic allocation problem of the wireless channel, so as to optimize the time delay of the UAV data transmission. Finally, an example was provided to introduce the application of the above methods to solve the problem of UAV cluster task scheduling.

In [17], a spatiotemporal scheduling framework for autonomous UAVs using RL was presented. The framework enabled UAVs to determine their schedules autonomously to cover the maximum pre-scheduled events that were spatially and temporally distributed in a given geographical area and over a pre-determined time horizon. The designed framework had the ability to update the planned schedules in the case of unexpected emergency events. The UAVs were trained using the Q-learning algorithm to determine an effective scheduling plan.

B. BIO INSPIRED ALGORITHMS

In [61], a bio-inspired approach for efficient task allocation was presented. The objective was to provide each UAV with the capability to make independent task allocation decisions to ensure autonomous control, an efficient running time, and the mean time to search and rescue a survivor. They developed

TABLE 8. Summary of ML-based resource allocation methods (C = centralized, D = distributed).

Publication	Resources	Model Objective	Learning Method	Approach	UAV
Sun et al. [109]	Power control	Throughput	Deep learning	C	×
Cui et al. [34]	Link scheduling	Sum rate	Deep learning	C	V
Zeng et al. [130]	Power allocation and scheduling	Convergence rate	Distributed distributed swarm approach	D	V
Matthiesen et al. [75]	Power control	energy efficiency	Deep learning	C	×
Azoulay et al. [11]	Power control	Throughput	DNN	C	V
Ahmed et al. [3]	Resource allocation	throughput	Deep learning	C	×
Chen et al. [27]	User association, spectrum allocation	max stable queues and content caching	Liquid state machine	D	V
Lee et al. [65]	Power control	Spectral efficiency/energy efficiency	CNN	C/D	×
Zhang et al. [132]	Power control	throughput	RL	D	V
Hu et al. [46]	User association, power management, subchannel alloc.	maximize successful transmissions	RL	D	V
Cao et al. [123]	Channel and power	Energy efficiency	DRL	C	V
Ciu et al. [32], [33]	Dynamic resources	Long-term rewards	MA RL	D	V
Zeng et al. [130]	Power control, scheduling	coverage	federated learning	D	V
Chen et al. [24]	Spectrum access	stability	Log-linear algorithm to achieve NE	D	V
Liu et al. [69]	Power allocation	Maximize own values	pricing	D	V
Zhang et al. [136]	Power control	reliability	Finite difference method	C+D	V
Zhang et al. [133]	Power control	Data traffic	APC clustering	C	V

a new locust-inspired heuristic with explicit mapping of the algorithm components to the existing biological locust colony, and the proposed algorithm was customized to the task allocation problem. The proposed approach was inspired by the nature of the locust species, and their autonomous and elastic behavior in response to inside and outside stimuli. The experimental results demonstrated the superiority of the performance of LIAM compared to the well-established MUTA heuristics, including the auction-based, max-sum, ACO, and opportunistic task allocation (OTA) algorithms, where it successfully maintained a significantly higher throughput as well as lower task completion and execution times.

In [118], a multi-UAV reconnaissance task allocation model was presented that considered the heterogeneity of targets as well as the constraints of the UAVs and targets. The optimization objective was to minimize the weighted sum of the task execution time and total UAV consumption. A modified GA was developed to solve the extended multiple Dubins traveling salesman problem (MDTSP). Double-chromosome encoding was used to describe the allocation results of targets to UAVs. Opposition-based learning and multiple mutation operators were employed to improve the optimality and convergence efficiency. The effectiveness of the algorithm was validated by numerical experiments on scenarios of different scales.

In [48], a hierarchical method was presented to solve the task assignment problem for multiple UAV teams. Tasks involving multiple UAVs formed into several teams that cooperatively attacked numerous ground targets were considered. In this scenario, the targets had to be pre-assigned to the UAVs. A hierarchical solution framework was developed to divide the task assignment problem into three subproblems: target clustering, cluster allocation, and target assignment. First, the targets were clustered based on their geographic positions. Thereafter, the clusters were assigned to different UAV teams, with one cluster per team, using the Hungarian algorithm. Finally, the targets were assigned to different team members for each team using an improved ACO (IACO) algorithm, the ant windows of which were adaptive to accelerate the convergence. Through this process, all UAVs were assigned targets along with the visit sequence. The problem was solved by means of clustering algorithms and integer linear programming, using a mixed integer linear programming model and the IACO. The ACO algorithm is a combinatorial optimization algorithm that simulates the foraging process of ants, which is performed using chemical signals known as pheromones.

In [60], the author introduced a bio-inspired algorithm for task allocation in multi-UAV search and rescue missions based on locust behavior. The locust species was selected as it represents an extreme example of adaptive control and high

plasticity to respond to internal and external stimuli. To evaluate the proposed algorithm, a strictly controlled experimental framework was designed to demonstrate the potential role of the proposed algorithm in such missions by providing a high net throughput and an efficient communication scheme.

In [128], the cooperative multiple task assignment problem (CMTAP) of heterogeneous fixed-wing UAVs performing the Suppression of Enemy Air Defense (SEAD) mission against multiple ground stationary targets was discussed. The study presented a modified GA with a multi-type gene chromosome encoding strategy. First, the multi-type gene encoding scheme was introduced to generate feasible chromosomes that satisfied the UAV capability, task coupling, and task precedence constraints. Thereafter, the Dubins car model was adopted to calculate the mission execution time (the objective function of the CMTAP model) of each chromosome and to make each chromosome conform to the UAV maneuverability constraint.

In [72], the process of allocating pesticide spraying tasks by multiple UAVs was investigated. A model was proposed to produce a satisfactory allocation scenario for pesticide spraying and to provide the flight route for each UAV. The model extended the distance between two points to the Dubins path distance and dynamically determined the time windows of the pesticide spraying according to the temperature conditions at the time that the UAVs performed the tasks. The study proposed a GA to solve the aforementioned problem and provided the encoding, crossover, and mutation methods in the algorithm.

The study of [59] proposed a bacteria-inspired heuristic for the efficient distribution of tasks among deployed UAVs. The use of multi-UAVs is a promising concept for combating the spread of the red palm weevil (RPW) in palm plantations. The proposed bacteria-inspired heuristic was used to resolve the multi-UAV task allocation problem when combating RPW infestation. The performance of the proposed algorithm was benchmarked in simulated detect-and-treat missions against three long-standing multi-UAV task allocation strategies, namely opportunistic task allocation, the auction-based scheme, and the max-sum algorithm, as well as a recently introduced locust-inspired algorithm for the allocation of multi-UAVs.

Table 9 and Figure 9 summarize the task allocation studies surveyed in this section.

X. GUIDELINES FOR ML METHOD SELECTION

In this study we surveyed several challenges related to UAV flocks, and described some recent ML based solutions for the different challenges. Given the information about most successful approaches, we would like to provide, in the following section, some guidelines to researchers attempting to select an appropriate ML method to solve a particular problem at hand. First of all we define some criteria to classify a given problem instance, and then we describe decision tree for each challenge.

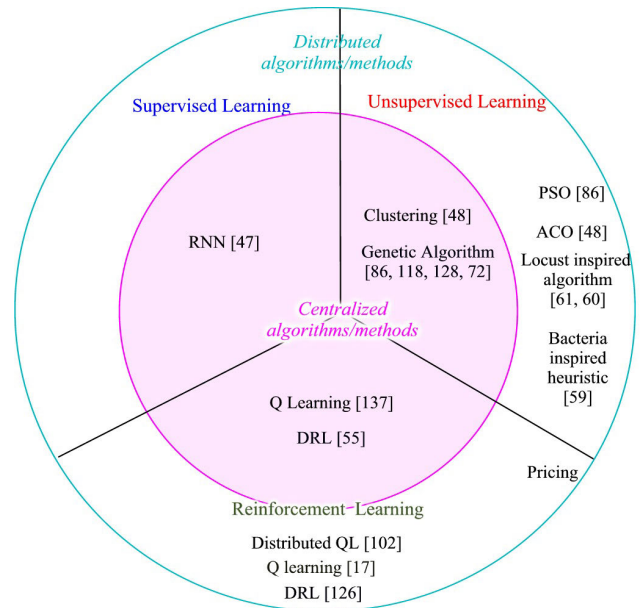


FIGURE 9. ML methods used for the task allocation challenge.

- **Flock challenge:** we considered the following challenges: clustering (flock formation) challenge; coordination, deployment and trajectory design; resource and task allocation. In addition, sometimes two or more challenges are combined, for example, if resource and task allocation decision are taken together.
- **Distributed, centralized or hybrid solution:** In general, centralized solution assumes a central controller that collect all UAV's information and manages the learning process. There are challenges where such situation is not realistic. For example, if online rapid power control is required, while the overhead of transmission the information to the central controller is too high. In such a case, distributed solution is used where each unit collects the relevant information to it and gets its own decision. A third possible approach is the hybrid approach: for example, small groups (cluster) of UAVs may be managed by a local manager (cluster head) for each group. This approach combines the advantages of the centralized solution: efficiency, available information, etc., with the advantages of the distributed solution: time and space lower complexity, even though the solution is mainly sub-optimal. Note that in domains which require distributed management, methods such as deep learning, which require large amount of information about the environment, will not be applicable, given the lack of the availability of this data at each node. However, deep reinforcement learning, where each UAV learns from its historical moves and from observable information in its own environment, can be used for addressing several challenges.
- **Cooperative UAVs or Self interested Multi UAVs:** When considering different flock challenges, there may be two types of situations. In most of the surveyed

TABLE 9. Summary of ML-based task allocation methods.

Publication	Challenge	Optimization Criteria	ML Method
Huang et al. [63]	Coordinated task allocation-	team rewards	Cross-entropy based algorithm
Pohl and Lamont [86]	UAV mission planning	path length	PSO, GA
Munoz-Morera et al. [80]	Task allocation	cost	planning
Hu et al. [47]	Multi-modal, multi-task task offloading	min delay	RNN
Shamsoshoara et al. [102]	UAV partition and task allocation	communication rate	Distributed Q-learning
Kim and Morrison [55]	Waypoint selection	reward	DRL
Zhao et al. [137]	Task allocation for heterogeneous UAVs in uncertain environment	Reward and coverage	DRL
Yang et al. [126]	Minimizing sum power for UAV-enabled MEC network	Transmission delay	DRL
Bouhamed et al. [17]	Spatiotemporal scheduling	coverage	Q-learning
Kurdi et al. [61]	UAV allocation to survivors	Throughput	Locust inspired
Wang et al. [118]	Task allocation for heterogeneous targets	path length	GA
Hu et al. [48]	Task allocation for multiple teams	Overall reward	Clustering + Hungarian alg. + ACO
Kurdi et al. [60]	Coordinated task allocation	Throughput	Locust inspired
Ye et al. [128]	Cooperative task assignment against ground stationary targets	Mission execution time	GA
Luo et al. [72]	Pesticide spraying task allocation	Total profit	GA
Kurdi et al. [59]	Task allocation among deployed UAVs	Net throughput	Bacteria-inspired heuristic

situation the UAVs are cooperative. This means that they belong to the same owner and attempt to reach a common goal or perform a common task. On the other side, there are also Multi UAV situations, in which each UAV has its own motivation, interests and goals to perform. In this situation, each suggested solution protocol should ensure that each UAV will be motivated to follow the protocol solution. Given self interested multi UAV situations, methods such as reinforcement learning, deep reinforcement learning, pricing mechanisms or other game theory based protocol should be applicable, rather than standard ML methods that do not consider the interests and motivation of each selfish UAV.

- **Optimization criteria:** In the literature of optimization challenges in flock control and management, there are several types of optimization criteria, that are used. Some of them are relevant to most or all challenges, while most of them are relevant only in part of the domains. In particular, the criteria of energy efficiency, and battery lifetime, can be considered for optimization of different aspect of UAV's flock, thus it may be relevant in most of the considered challenges. Coverage, throughput maximization, and latency minimization, are relevant in most challenges that are related to aerial communication networks. Issues of fairness are more relevant in multi-UAV domains, where each UAV has its own goals and incentives. Finally, in task allocation

and deployment challenges, task completion rate and coverage success should be considered too. In fact, different optimization criteria can be used by the same ML method, by changing the objective function which the ML method aims to optimize.

- **Does the Required big-data exist or can it be produced:** Most of the new machine learning methods rely on a large amount of information to perform their training. Thus, in domains where big data exists, or can be created, such machine learning methods, and especially deep learning methods, can be successfully utilized. In particular, big data is required in deep learning methods, including the different popular DNNs, CNNs, RNNs, etc., which all rely on training phase done using large amount of available training data.
- **Training phase can be performed:** Given the fact that all challenges considered in our survey should find solutions in real time, the requirement in run time is critical. However, if the main characteristics of the environment are known in advance, a training phase can be performed before getting to real time situations. In some domains, the main challenge situations is indeed already known, thus training phase can be performed in order to train the machine learning method. However, if we consider dynamic environment that frequently changes and where no key characteristics can be exhibited ahead of time, then training phase may

results in learning a model that is not applicable to the real time situations. Therefore, when a training phase is not possible, it would be better to use algorithmic solutions, or methods like Particle Swarm Optimization, which do not require pre-training phase.

- **Computational available resources:** Some of the machine learning methods required a large amount of time and space to be performed. Therefore, when there are limited resources, one may prefer using machine learning methods with lower resources requirement, as well as using non-optimal greedy algorithms to handle the challenge. Deep learning methods, or deep reinforcement methods, will not be considered in such cases. However, in situations where large computational resources (time and space) are available for the training phase, but not in real time, still deep learning methods may still be useful, if training is done off-line. Then in real time the UAV behaves according to the process learned in the training phase.

In Figure 10 we show for each UAV flock challenge, the main recent ML that were recommended for solving it. The description is given as a decision tree, where the leaves of the decision tree describes the main ML methods that have been successfully used for this type of problem.

Figure 10 can be used as a guideline to choose the most appropriate machine learning method, given the problem attributes. Surely, there are challenges which should be solved in distributed manner, such as clustering, flocking, and UAV coordination, where no central unit can solve them. Thus methods such as PSO, DP, RL or DRL can be used in this case, by each decision maker unit in the flock. Other challenges that can be solved also in centralized manner, are UAV deployment and resource and task allocation. When a centralized unit solves problem, it may use all the available information, and a DNN, CNNs or RNNs, can be applied to solve it. In addition, if the problem can be defined as a MDP, and the effect of the current decision of the future system status should be considered, then the RL or DRL can be used to solve the challenge.

Note that of course there are additional approaches that are not described in our discussion, as this survey focuses on ML methods. In addition, we recommend the readers to explore other effective methods, that can of course be effective for solving various problems and challenges.

XI. OPEN ISSUES

Several open issues can be considered in the different challenges involving UAV flocks. In the following section, we describe some of the interesting open issues that can be considered.

A. COMPARISON OF DIFFERENT ML METHODS

First, most studies have suggested one or two ML-based methods to solve the flock challenge, and subsequently compared the performance of their methods to classical

algorithmic solutions presented in the literature. However, a comparison of different ML methods has rarely been conducted, while this may be important to enable the correct mapping of an ML method to each specific challenge. For example, when considering flocking and coordination challenges, where different algorithmic and ML methods were suggested, a comprehensive study comparing the different approaches for the same problem and their efficiency can be profitable for the system designers.

B. DEFINING UNIFORM OPTIMIZATION CRITERIA

When describing the different studies on controlling and managing aerial flocks, we observed that different optimization criteria exists, which may even conflict with each other. A uniform set of optimization criteria may be helpful in order to be able to compare the efficiency of various methods and solution. For example, when considering the resource allocation challenge, we can suggest the criteria of average throughput and average latency, but other criteria, that take into consideration fairness and starvation avoidance, may be also important.

C. ML METHODS COMBINATIONS

The combination of more than one DRL method may also be efficient for different real problems. For example, a CNN and LSTM can be combined to achieve superior solutions when considering sequences of several prediction problems over a time series given coverage maps. Moreover, DRL and CNN or LSTM can be combined to achieve better decision making in complex spatial or sequential models.

Another type of challenge, which may benefit, from utilizing ML methods combinations, is reflecting the effects of the flock decision on the long term flock management, or the effect of the flock member on its own long term status. These effects can be incorporated by using RL and DRL methods, probably combined with other ML methods. For example, the DRQN method [41], which is a combination of RNN and deep Q learning method, developed for challenging partially observed decision processes, can be successfully applied for complex design and control challenges in the aerial domain.

D. DISTRIBUTED ML METHODS

Distributed algorithms have been widely suggested for flock management. However, when considering a UAV that performs its tasks alone without interacting in a flow (hereafter referred to as a standalone UAV), the use of distributed algorithms is rare. A standalone UAV behaves in a manner that is best for itself, and it will cooperate with other UAVs only if this will improve its own situation. Thus, any mechanism that is proposed for such constellations should be proven to be stable. DRL methods can be used in such situations, considering self-learning by each UAV or by each UAV group, but it is not clear whether centralized learning can be implemented by any standalone CH or cluster presenter, given the fact that it may learn a model that is biased for its own profit rather than for the benefit of the entire flock or system.

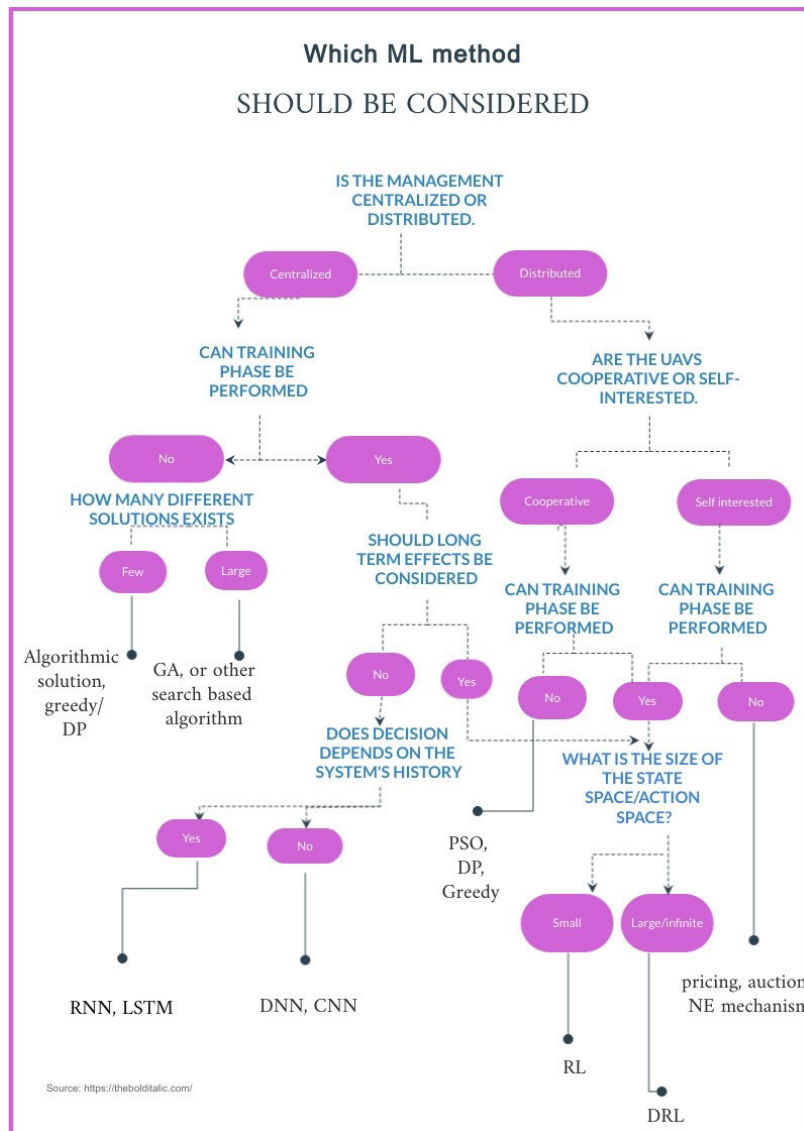


FIGURE 10. Which ML method should be considered: a decision tree view.

E. REAL WORLD DATA-SETS

Another challenge that will affect future developments in this area is the need to obtain real-world-based standard flock databases, with data regarding their motions and actions, for use as comparative databases for ML training when solving algorithm challenges relating to flock formation and maintenance. Databases for standard UAV motions and actions may also be useful to compare different ML methods, which is already the case in other domains in which ML is widely used, such as NLP and machine vision. Such datasets may include several data about UAV flocks, including the geographical aspects of UAV motion, and relevant information about the strength of the wireless coverage of the flying ad-hoc network. The geographical data included in the datasets should include the three dimensional locations of the UAVs, which may be important to the development of efficient methods for solving the flocking challenges in the aerial domain.

F. EMBEDDING REPRESENTATION

Another important issue that can be studied, when considering DL methods to solve flocking challenges, is how to represent the aerial domain in a way that will enable easy training of the DL model, as well as enabling scalability of the learning process. We should enable the learning process to be done on small aerial domains, and also to manage flocks working in larger and complex domains. Some recent studies suggested embedding representation for graph information to solve classical optimization problems on graphs using ML. Therefore some of these ideas can be helpful to the embedding challenge of the aerial technical properties.

G. CONSIDERING THE HYBRID PROBLEM

Finally, the abilities of different deep learning schemes make it possible to take a comprehensive planning approach, and to consider all aspects of flock management and use.

This enables the provision of an overall ML solution that solves all of the involved challenges, including flock formation, flock maintenance, resource allocation, task allocation, and UAV deployment, simultaneously. Such a harmonic view may be efficient and can provide a real-time, rapid solution to several applications of UAV flocks, which are expected to become increasingly popular in the near future, as the cost of UAV-based technology becomes lower and their use becomes more crucial.

XII. CONCLUSION

This paper has surveyed several complex issues relating to UAV flock formation, maintenance, and related challenges, which can be efficiently handled using ML methods. We have provided a review that is as current as possible, and in each section we also discussed relevant open issues. We hope that our study will be helpful for researchers and developers who wish to apply ML methods for solving challenges relating to UAV flocks, and in particular, regarding FANETs.

The technological advancements in the production of UAVs, together with the need to manage them by grouping them into clusters, offers enormous potential in various fields, including commercial, industrial, agricultural, and security applications. Moreover, several challenges arise in determining the correct manner in which to manage UAV clusters. ML methods can enable rapid and efficient solutions to these challenges in real time for problems that can be considered in advance, so that the learning system can be trained in a manner that will make it possible to handle the various challenges in real time effectively. Such challenges appear in several main areas: flock formation and management; flock deployment, coordination, and navigation; inter-flock and intra-flock communication resource allocation; and finally, task allocation in flocks that are responsible for handling predefined tasks. To the best of our knowledge, this survey is the first attempt to address the range of issues associated with the creation, management, and maintenance of UAV clusters comprehensively, with particular emphasis on machine-based solutions, to provide an effective and rapid response to these important issues.

REFERENCES

- [1] F. Aftab, A. Khan, and Z. Zhang, "Hybrid self-organized clustering scheme for drone based cognitive Internet of Things," *IEEE Access*, vol. 7, pp. 56217–56227, 2019.
- [2] P. Agarwal and C. Procopiuc, "Exact and approximation algorithms for clustering," in *Proc. 9th Annu. ACM-SIAM Symp. Discrete Algorithms*, 1999, pp. 658–667.
- [3] K. I. Ahmed, H. Tabassum, and E. Hossain, "Deep learning for radio resource allocation in multi-cell networks," *IEEE Netw.*, vol. 33, no. 6, pp. 188–195, Nov./Dec. 2019.
- [4] F. Al-Turjman and H. Zahmatkesh, "A comprehensive review on the use of AI in UAV communications: Enabling technologies, applications, and challenges," in *Unmanned Aerial Vehicles in Smart Cities*. Cham, Switzerland: Springer, 2020, pp. 1–26.
- [5] S. H. Alsamhi, O. Ma, and M. S. Ansari, "Survey on artificial intelligence based techniques for emerging robotic communication," *Telecommun. Syst.*, vol. 72, pp. 483–503, Mar. 2019.
- [6] O. Anicho, P. B. Charlesworth, G. S. Baicher, A. Nagar, and N. Buckley, "Comparative study for coordinating multiple unmanned HAPS for communications area coverage," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2019, pp. 467–474.
- [7] M. Y. Arafat and S. Moh, "Localization and clustering based on swarm intelligence in UAV networks for emergency communications," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8958–8976, Oct. 2019.
- [8] A. Sargolzaei, A. Abbaspour, and C. D. Crane, "Control of cooperative unmanned aerial vehicles: Review of applications, challenges, and algorithms," in *Optimization, Learning, and Control for Interdependent Complex Networks*, vol. 1123. Atlanta, GA, USA: Springer, 2020, pp. 229–255.
- [9] A. Aswani, P. Bouffard, and C. Tomlin, "Extensions of learning-based model predictive control for real-time application to a quadrotor helicopter," in *Proc. Amer. Control Conf. (ACC)*, Jun. 2012, pp. 4661–4666.
- [10] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.
- [11] R. Azoulay, K. Danilchenko, Y. Haddad, and S. Reches, "Transmission power control using deep neural networks in TDMA-based ad-hoc network clusters," in *Proc. IWCMC*, Jun. 2021, pp. 406–411.
- [12] R. Azoulay and S. Reches, "UAV flocks forming for crowded flight environments," in *Proc. JCAART*, 2019, pp. 154–163.
- [13] R. Azoulay and S. Reches, "Flocks formation model for self-interested UAVs," *Intell. Service Robot.*, vol. 14, no. 2, pp. 157–174, Apr. 2021.
- [14] H. Bayerlein, P. De Kerret, and D. Gesbert, "Trajectory optimization for autonomous flying base station via reinforcement learning," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
- [15] L. E. Beaver and A. A. Malikopoulos, "An overview on optimal flocking," *Annu. Rev. Control*, vol. 51, pp. 88–99, 2021.
- [16] P. S. Bithas, E. T. Michailidis, N. Nomikos, D. Vouyioukas, and A. G. Kanatas, "A survey on machine-learning techniques for UAV-based communications," *Sensors*, vol. 19, no. 23, p. 5170, Nov. 2019.
- [17] O. Bouhamed, H. Ghazzai, H. Besbes, and Y. Massoud, "A generic spatiotemporal scheduling for autonomous UAVs: A reinforcement learning-based approach," *IEEE Open J. Veh. Technol.*, vol. 1, pp. 93–106, 2020.
- [18] C. H. Liu, X. Ma, X. Gao, and J. Tang, "Distributed energy-efficient multi-UAV navigation for long-term communication coverage by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 19, no. 6, pp. 1274–1285, Jun. 2020.
- [19] A. Carrio, C. S. Pérez, A. Rodriguez-Ramos, and P. Campoy, "A review of deep learning methods and applications for unmanned aerial vehicles," *J. Sensors*, vol. 2017, pp. 1–13, Aug. 2017.
- [20] U. Cekmez, M. Ozsiginan, and O. K. Sahingoz, "Multi colony ant optimization for UAV path planning with obstacle avoidance," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2016, pp. 47–52.
- [21] U. Challita, W. Saad, and C. Bettstetter, "Deep reinforcement learning for interference-aware path planning of cellular-connected UAVs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–7.
- [22] U. Challita, W. Saad, and C. Bettstetter, "Cellular-connected UAVs over 5G: Deep reinforcement learning for interference management," 2018, *arXiv:1801.05500*. [Online]. Available: <http://arxiv.org/abs/1801.05500>
- [23] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.
- [24] J. Chen, Q. Wu, Y. Xu, Y. Zhang, and Y. Yang, "Distributed demand-aware channel-slot selection for multi-UAV networks: A game-theoretic learning approach," *IEEE Access*, vol. 6, pp. 14799–14811, 2018.
- [25] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.
- [26] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017.
- [27] M. Chen, W. Saad, and C. Yin, "Liquid state machine learning for resource and cache management in LTE-U unmanned aerial vehicle (UAV) networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1504–1517, Mar. 2019.
- [28] X. Chen, J. Tang, and S. Lao, "Review of unmanned aerial vehicle swarm communication architectures and routing protocols," *Appl. Sci.*, vol. 10, no. 10, p. 3661, May 2020.

- [29] S.-J. Chung, A. A. Paranjape, P. Dames, S. Shen, and V. Kumar, "A survey on aerial swarm robotics," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 837–855, Aug. 2018.
- [30] M. Collotta, G. Pau, and V. Maniscalco, "A fuzzy logic approach by using particle swarm optimization for effective energy management in IWSNs," *IEEE Trans. Ind. Electron.*, vol. 64, no. 12, pp. 9496–9506, Dec. 2017.
- [31] M. Coppola, K. N. McGuire, C. De Wagter, and G. C. H. E. de Croon, "A survey on swarming with micro air vehicles: Fundamental challenges and constraints," *Frontiers Robot. AI*, vol. 7, p. 18, Feb. 2020.
- [32] J. Cui, Y. Liu, and A. Nallanathan, "The application of multi-agent reinforcement learning in UAV networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.
- [33] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 729–743, Feb. 2020.
- [34] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1248–1261, Jun. 2019.
- [35] C. Wu and D. P. Bertsekas, "Distributed power control algorithms for wireless networks," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 504–514, Mar. 2001.
- [36] H. Dai, H. Zhang, M. Hua, C. Li, Y. Huang, and B. Wang, "How to deploy multiple UAVs for providing communication service in an unknown region?" *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1276–1279, Aug. 2019.
- [37] H. Dai, H. Zhang, B. Wang, and L. Yang, "The multi-objective deployment optimization of UAV-mounted cache-enabled base stations," *Phys. Commun.*, vol. 34, pp. 114–120, Jun. 2019.
- [38] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 4th Quart., 2019.
- [39] R. Ganesan, X. M. Raajini, A. Nayyar, P. Sanjeevikumar, E. Hossain, and A. H. Ertas, "BOLD: Bio-inspired optimized leader election for multiple drones," *Sensors*, vol. 20, no. 11, p. 3134, Jun. 2020.
- [40] S. Govindaraj and S. N. Deepa, "Network energy optimization of IOTs in wireless sensor networks using capsule neural network learning model," *Wireless Pers. Commun.*, vol. 115, no. 3, pp. 2415–2436, Dec. 2020.
- [41] M. Hausknecht and P. Stone, "Deep recurrent Q-learning for partially observable MDPs," in *Proc. AAAI Fall Symp. Ser.*, 2015, pp. 1–9.
- [42] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2624–2661, 4th Quart., 2016.
- [43] H. He, Z. Zhu, and E. Mäkinen, "A neural network model to minimize the connected dominating set for self-configuration of wireless sensor networks," *IEEE Trans. Neural Netw.*, vol. 20, no. 6, pp. 973–982, Jun. 2009.
- [44] A. I. Hentati, L. Krichen, M. Fourati, and L. C. Fourati, "Simulation tools, environments and frameworks for UAV systems performance analysis," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 1495–1500.
- [45] A. I. Hentati and L. C. Fourati, "Comprehensive survey of UAVs communication networks," *Comput. Standards Interfaces*, vol. 72, Oct. 2020, Art. no. 103451.
- [46] J. Hu, H. Zhang, L. Song, Z. Han, and H. V. Poor, "Reinforcement learning for a cellular internet of UAVs: Protocol design, trajectory control, and resource management," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 116–123, Feb. 2020.
- [47] L. Hu, Y. Tian, J. Yang, T. Taleb, L. Xiang, and Y. Hao, "Ready player one: UAV-clustering-based multi-task offloading for vehicular VR/AR gaming," *IEEE Netw.*, vol. 33, no. 3, pp. 42–48, May 2019.
- [48] X. Hu, H. Ma, Q. Ye, and H. Luo, "Hierarchical method of task assignment for multiple cooperating UAV teams," *J. Syst. Eng. Electron.*, vol. 26, no. 5, pp. 1000–1009, 2015.
- [49] S.-M. Hung and S. N. Givigi, "A Q-learning approach to flocking with UAVs in a stochastic environment," *IEEE Trans. Cybern.*, vol. 47, no. 1, pp. 186–197 Jan. 2017.
- [50] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (FANETs): A survey," *Ad Hoc Netw.*, vol. 11, no. 3, pp. 1254–1270, 2013.
- [51] J. Hu, H. Zhang, and L. Song, "Reinforcement learning for decentralized trajectory design in cellular UAV networks with sense-and-send protocol," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6177–6189, Aug. 2019.
- [52] M. Jafari, H. Xu, and L. R. G. Carrillo, "A biologically-inspired reinforcement learning based intelligent distributed flocking control for multi-agent systems in presence of uncertain system and dynamic environment," *IFAC J. Syst. Control*, vol. 13, Sep. 2020, Art. no. 100096.
- [53] E. B. Khalil, H. Dai, Y. Zhang, B. Dilkina, and L. Song, "Learning combinatorial optimization algorithms over graphs," in *Proc. Neural Inf. Process. Syst.*, 2017, pp. 1–11.
- [54] B. Khamidehi and E. S. Sousa, "Reinforcement learning-based trajectory design for the aerial base stations," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–6.
- [55] I. Kim and J. R. Morrison, "Learning based framework for joint task allocation and system design in stochastic multi-UAV systems," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2018, pp. 324–334.
- [56] P. V. Klaine, J. P. B. Nadas, R. D. Souza, and M. A. Imran, "Distributed drone base station positioning for emergency cellular networks using reinforcement learning," *Cogn. Comput.*, vol. 10, no. 5, pp. 790–804, Oct. 2018.
- [57] V. Kouhdaragh, F. Verde, G. Gelli, and J. Abouei, "On the application of machine learning to the design of UAV-based 5G radio access networks," *Electronics*, vol. 9, no. 4, p. 689, Apr. 2020.
- [58] A. M. Koushik, F. Hu, and S. Kumar, "Deep Q-learning-based node positioning for throughput-optimal communications in dynamic UAV swarm network," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 3, pp. 554–566, Sep. 2019.
- [59] H. Kurdi, M. F. AlDaoud, S. Al-Megren, E. Aloboud, A. S. Aldawood, and K. Youcef-Toumi, "Adaptive task allocation for multi-UAV systems based on bacteria foraging behaviour," *Appl. Soft Comput.*, vol. 83, Oct. 2019, Art. no. 105643.
- [60] H. Kurdi, J. How, and G. Bautista, "Bio-inspired algorithm for task allocation in multi-UAV search and rescue missions," in *Proc. AIAA*, Jan. 2016, p. 1377.
- [61] H. A. Kurdi, E. Aloboud, M. Alalwan, S. Alhassan, E. Alotaibi, G. Bautista, and J. P. How, "Autonomous task allocation for multi-UAV systems based on the locust elastic behavior," *Appl. Soft Comput.*, vol. 71, pp. 110–126, Oct. 2018.
- [62] L. D. Nguyen, K. K. Nguyen, A. Kortun, and T. Q. Duong, "Real-time deployment and resource allocation for distributed UAV systems in disaster relief," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [63] L. Huang, H. Qu, and L. Zuo, "Multi-type UAVs cooperative task allocation under resource constraints," *IEEE Access*, vol. 6, pp. 17841–17850, 2018.
- [64] N. M. A. Latiff, C. C. Tsimenidis, and B. S. Sharif, "Energy-aware clustering for wireless sensor networks using particle swarm optimization," in *Proc. IEEE 18th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2007, pp. 1–5.
- [65] W. Lee, M. Kim, and D.-H. Cho, "Deep power control: Transmit power control scheme based on convolutional neural network," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1276–1279, Jun. 2018.
- [66] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, Apr. 2019.
- [67] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, Sep. 2018.
- [68] J. Liu, Q. Wang, X. Li, and W. Zhang, "A fast deployment strategy for UAV enabled network based on deep learning," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Aug. 2020, pp. 1–6.
- [69] X. Liu, L. Li, F. Yang, X. Li, W. Chen, and W. Xu, "Price-based power allocation for multi-UAV enabled wireless networks," in *Proc. 28th Wireless Opt. Commun. Conf. (WOCC)*, May 2019, pp. 1–5.
- [70] X. Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-UAV networks: Deployment and movement design," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8036–8049, Aug. 2019.
- [71] N. K. Long, K. Sammut, D. Sgarioto, M. Garratt, and H. A. Abbass, "A comprehensive review of shepherding as a bio-inspired swarm-robotics guidance approach," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 21, no. 4, pp. 3039–3071, Aug. 2020.
- [72] H. Luo, Y. Niu, M. Zhu, X. Hu, and H. Ma, "Optimization of pesticide spraying tasks via multi-UAVs using genetic algorithm," *Math. Problems Eng.*, vol. 2017, pp. 1–16, Nov. 2017.

- [73] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [74] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, Mar. 2017.
- [75] B. Matthiesen, A. Zappone, E. A. Jorswieck, and M. Debbah, "Deep learning for real-time energy-efficient power control in mobile networks," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [76] I. Maza, F. Caballero, J. Capitan, J. R. M. de Dios, and A. Ollero, "Experimental results in multi-UAV coordination for disaster management and civil security applications," *J. Intell. Robot. Syst.*, vol. 61, no. 1, pp. 563–585, 2011.
- [77] I. Maza, K. Kondak, M. Bernard, and A. Ollero, "Multi-UAV cooperation and control for load transportation and deployment," *J. Intell. Robot. Syst.*, vol. 57, pp. 417–449, Jan. 2010.
- [78] S. Mirjalili, "Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Comput. Appl.*, vol. 27, no. 4, pp. 1053–1073, 2016.
- [79] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 3rd Quart., 2019.
- [80] J. Muñoz-Morera, I. Maza, F. Caballero, and A. Ollero, "Architecture for the automatic generation of plans for multiple UAS from a generic mission description," *J. Intell. Robot. Syst.*, vol. 84, nos. 1–4, pp. 493–509, Dec. 2016.
- [81] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.
- [82] O. S. Oubbati, M. Atiquzzaman, P. Lorenz, M. H. Tareque, and M. S. Hossain, "Routing in flying ad hoc networks: Survey, constraints, and future challenge perspectives," *IEEE Access*, vol. 7, pp. 81057–81105, 2019.
- [83] Y. M. Park, M. Lee, and C. S. Hong, "Multi-UAVs collaboration system based on machine learning for throughput maximization," in *Proc. 20th Asia-Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Sep. 2019, pp. 1–4.
- [84] J. Peng and R. J. Williams, "Incremental multi-step Q-learning," *Mach. Learn.*, vol. 22, nos. 1–3, pp. 283–290, 1996.
- [85] P. Boccadoro, D. Striccoli, and L. A. Grieco, "An extensive survey on the Internet of Drones," 2020, *arXiv:2007.12611*. [Online]. Available: <http://arxiv.org/abs/2007.12611>
- [86] A. J. Pohl and G. B. Lamont, "Multi-objective UAV mission planning using evolutionary computation," in *Proc. Winter Simul. Conf.*, Dec. 2008, pp. 1268–1279.
- [87] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach," *Eng. Appl. Artif. Intell.*, vol. 33, pp. 127–140, Aug. 2014.
- [88] Q. Pan, Q. Ni, H. Du, Y. Yao, and Q. Lv, "An improved energy-aware cluster heads selection method for wireless sensor networks based on k -means and binary particle swarm optimization," in *Advances in Swarm Intelligence*, vol. 8795. Cham, Switzerland: Springer, 2014.
- [89] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2595–2621, 4th Quart., 2018.
- [90] H. Qie, D. Shi, T. Shen, X. Xu, Y. Li, and L. Wang, "Joint optimization of multi-UAV target assignment and path planning based on multi-agent reinforcement learning," *IEEE Access*, vol. 7, pp. 146264–146272, 2019.
- [91] J. Qiu, J. Lyu, and L. Fu, "Placement optimization of aerial base stations with deep reinforcement learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [92] S. A. P. Quintero, G. E. Collins, and J. P. Hespanha, "Flocking with fixed-wing UAVs for distributed sensing: A stochastic optimal control approach," in *Proc. Amer. Control Conf.*, Jun. 2013, pp. 2025–2031.
- [93] D. R. Vikranth and K. Siddhabathula, "Autonomous air traffic control—Collision avoidance for UAVs using MDP," *Int. J. Comput. Sci. Inf. Technol. Secur.*, vol. 6, no. 1, pp. 446–454, 2016.
- [94] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 25–34, Aug. 1987.
- [95] S. Ragi and E. K. P. Chong, "UAV path planning in a dynamic environment via partially observable Markov decision process," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2397–2412, Oct. 2013.
- [96] S. Temizer, M. Kochenderfer, L. Kaelbling, T. Lozano-Perez, and J. Kuchar, "Collision avoidance for unmanned aircraft using Markov decision processes," in *Proc. AIAA Guid., Navigat., Control Conf.*, 2010, p. 8040.
- [97] S. W. Kang, K. Thar, and C. S. Hong, "Unmanned aerial vehicle allocation and deep learning based content caching in wireless network," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2020, pp. 793–796.
- [98] W. Saad, M. Bennis, M. Mozaffari, and X. Lin, *Wireless Communications and Networking for Unmanned Aerial Vehicles*. Cambridge, U.K.: Cambridge Univ. Press, 2020.
- [99] O. S. Oubbati, M. Atiquzzaman, T. A. Ahanger, and A. Ibrahim, "Softwarization of UAV networks: A survey of applications and future trends," *IEEE Access*, vol. 8, pp. 98073–98125, 2020.
- [100] S. Butenko, X. Cheng, C. A. Oliveira, and P. M. Pardalos, "A new heuristic for the minimum connected dominating set problem on ad hoc wireless networks," in *Recent Developments in Cooperative Control and Optimization*. New York, NY, USA: Springer, 2004, pp. 61–73.
- [101] R. Shakeri, M. A. Al-Garadi, A. Badawy, A. Mohamed, T. Khattab, A. K. Al-Ali, K. A. Harras, and M. Guizani, "Design challenges of multi-UAV systems in cyber-physical applications: A comprehensive survey and future directions," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3340–3385, 4th Quart., 2019.
- [102] A. Shamsoshoara, M. Khaledi, F. Afghah, A. Razi, and J. Ashdown, "Distributed cooperative spectrum sharing in UAV networks using multi-agent reinforcement learning," in *Proc. 16th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2019, pp. 1–6.
- [103] A. Sharma, P. Vanjani, N. Paliwal, C. M. W. Basnayaka, D. N. K. Jayakody, H.-C. Wang, and P. Muthuchidambaramanathan, "Communication and networking technologies for UAVs: A survey," *J. Netw. Comput. Appl.*, vol. 168, no. 8, pp. 149–169, 2020.
- [104] R. K. Sharma and D. Ghose, "Collision avoidance between UAV clusters using swarm intelligence techniques," *Int. J. Syst. Sci.*, vol. 40, no. 5, pp. 521–538, May 2009.
- [105] K. Shen and W. Yu, "FPLinQ: A cooperative spectrum sharing strategy for device-to-device communications," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2323–2327.
- [106] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [107] C. Shyalika, T. Silva, and A. Karunananda, "Reinforcement learning in dynamic task scheduling: A review," *Social Netw. Comput. Sci.*, vol. 1, no. 6, p. 306, Nov. 2020.
- [108] S. Suganthi and S. P. Rajagopalan, "Multi-swarm particle swarm optimization for energy-effective clustering in wireless sensor networks," *Wireless Pers. Commun.*, vol. 94, no. 4, pp. 2487–2497, 2017.
- [109] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.
- [110] A. Tahir, J. Böling, M.-H. Haghbayan, H. T. Toivonen, and J. Plosila, "Swarms of unmanned aerial vehicles—A survey," *J. Ind. Inf. Integr.*, vol. 16, Dec. 2019, Art. no. 100106.
- [111] T. Alladi, V. Chamola, N. Sahu, and M. Guizani, "Applications of blockchain in unmanned aerial vehicles: A review," *Veh. Commun.*, vol. 23, Jun. 2020, Art. no. 100249.
- [112] A. Alarifi and A. Tolba, "Optimizing the network energy of cloud assisted Internet of Things by using the adaptive neural learning approach in wireless sensor networks," *Comput. Ind.*, vol. 106, pp. 133–141, Apr. 2019.
- [113] Y.-H. Tsai, "Vision-based collision avoidance for unmanned aerial vehicles by recurrent neural networks," *World Acad. Sci., Eng. Technol., Int. J. Comput. Inf. Eng.*, vol. 13, no. 4, pp. 196–200, 2019.
- [114] Z. Ullah, F. Al-Turjman, U. Moatasim, L. Mostarda, and R. Gagliardi, "UAVs joint optimization problems and machine learning to improve the 5G and beyond communication," *Comput. Netw.*, vol. 182, Dec. 2020, Art. no. 107478.
- [115] F. Venturini, F. Mason, F. Pase, F. Chiariotti, A. Testolin, A. Zanella, and M. Zorzi, "Distributed reinforcement learning for flexible UAV swarm control with transfer learning capabilities," in *Proc. 6th ACM Workshop Micro Aerial Vehicle Netw. Syst. Appl.*, Jun. 2020, pp. 1–6.
- [116] H. Wang, H. Zhao, J. Zhang, D. Ma, J. Li, and J. Wei, "Survey on unmanned aerial vehicle networks: A cyber physical system perspective," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1027–1070, 2nd Quart., 2020.

- [117] J. Wang, Y.-Q. Cao, B. Li, S.-Y. Lee, and J.-U. Kim, "A glowworm swarm optimization based clustering algorithm with mobile sink support for wireless sensor networks," *Internet Technol. J.*, vol. 16, no. 5, pp. 825–832, 2015.
- [118] Z. Wang, L. Liu, T. Long, and Y. Wen, "Multi-UAV reconnaissance task allocation for heterogeneous targets using an opposition-based genetic algorithm with double-chromosome encoding," *Chin. J. Aeronaut.*, vol. 31, no. 2, pp. 339–350, 2018.
- [119] Q. Wu, J. Xu, Y. Zeng, D. W. K. Ng, N. Al-Dhahir, R. Schober, and A. L. Swindlehurst, "A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence," 2020, *arXiv:2010.09317*. [Online]. Available: <http://arxiv.org/abs/2010.09317>
- [120] Z. Wu, J. Li, J. Zuo, and S. Li, "Path planning of UAVs based on collision probability and Kalman filter," *IEEE Access*, vol. 6, pp. 34237–34245, 2018.
- [121] X. Liu, Y. Liu, Y. Chen, and L. Hanzo, "Trajectory design and power control for multi-UAV assisted wireless networks: A machine learning approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7957–7969, Aug. 2019.
- [122] H. Xu and L. R. G. Carrillo, "Distributed near optimal flocking control for multiple unmanned aircraft systems," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2015, pp. 879–885.
- [123] Y. Cao, L. Zhang, and Y.-C. Liang, "Deep reinforcement learning for channel and power allocation in UAV-enabled IoT systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [124] Y. Yu, L. Ru, and K. Fang, "Bio-inspired mobility prediction clustering algorithm for ad hoc UAV networks," *Eng. Lett.*, vol. 24, no. 3, pp. 83–92, 2016.
- [125] C. Yan, L. Fu, J. Zhang, and J. Wang, "A comprehensive survey on UAV communication channel modeling," *IEEE Access*, vol. 7, pp. 107769–107792, 2019.
- [126] J. Yang, X. You, G. Wu, M. M. Hassan, A. Almogren, and J. Guna, "Application of reinforcement learning in UAV cluster task scheduling," *Future Gener. Comput. Syst.*, vol. 95, no. 1, pp. 140–148, 2019.
- [127] Z. Yang, C. Pan, K. Wang, and M. Shikh-Bahaei, "Energy efficient resource allocation in UAV-enabled mobile edge computing networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4576–4589, Sep. 2019.
- [128] F. Ye, J. Chen, Y. Tian, and T. Jiang, "Cooperative multiple task assignment of heterogeneous UAVs using a modified genetic algorithm with multi-type-gene chromosome encoding strategy," *J. Intell. Robot. Syst.*, vol. 100, no. 2, pp. 615–627, Nov. 2020.
- [129] X. Yuan, M. Elhoseny, H. K. El-Minir, and A. M. Riad, "A genetic algorithm-based, dynamic clustering method towards improved WSN longevity," *J. New. Syst. Manage.*, vol. 25, no. 1, pp. 21–46, Jan. 2017.
- [130] T. Zeng, O. Semiari, M. Mozaffari, M. Chen, W. Saad, and M. Bennis, "Federated learning in the sky: Joint power allocation and scheduling with UAV swarms," 2020, *arXiv:2002.08196*. [Online]. Available: <http://arxiv.org/abs/2002.08196>
- [131] Y. Zeng, X. Xu, S. Jin, and R. Zhang, "Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning," 2020, *arXiv:2003.07574*. [Online]. Available: <http://arxiv.org/abs/2003.07574>
- [132] J. Zhang, M. Lou, L. Xiang, and L. Hu, "Power cognition: Enabling intelligent energy harvesting and resource allocation for solar-powered UAVs," *Future Gener. Comput. Syst.*, vol. 110, pp. 658–664, Sep. 2020.
- [133] J. Zhang, G. Chuai, and W. Gao, "Power control and clustering-based interference management for UAV-assisted networks," *Sensors*, vol. 20, no. 14, p. 3864, Jul. 2020.
- [134] Q. Zhang, M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Machine learning for predictive on-demand deployment of UAVs for wireless communications," in *Proc. IEEE Global Commun. Conf.*, Dec. 2018, pp. 1–6.
- [135] W. Zhang, L. Li, N. Zhang, T. Han, and S. Wang, "Air-ground integrated mobile edge networks: A survey," *IEEE Access*, vol. 8, pp. 125998–126018, 2020.
- [136] Z. Zhang, L. Li, X. Liu, W. Liang, and Z. Han, "Matching-based resource allocation and distributed power control using mean field game in the NOMA-based UAV networks," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2018, pp. 420–426.
- [137] X. Zhao, Q. Zong, B. Tian, B. Zhang, and M. You, "Fast task allocation for heterogeneous unmanned aerial vehicles through reinforcement learning," *Aerosp. Sci. Technol.*, vol. 92, pp. 588–594, Sep. 2019.



RINA AZOULAY received the B.A. degree in economics and computer science and the M.Sc. degree in computer science from Bar-Ilan University, and the Ph.D. degree in computer science from Bar-Ilan University, in 2001. She is currently a Senior Lecturer with the Computer Science Department, Jerusalem College of Technology, Israel. Her main research interests include protocols for reaching agreements, application of machine learning methods, and intelligent tutoring systems.



YORAM HADDAD (Senior Member, IEEE) received the B.Sc. (Hons.), Engineering Diploma (Hons.), and M.Sc. (Hons.) (radiocommunications) degrees from SUPELEC (leading engineering school), Paris, France, in 2004 and 2005, the Ph.D. degree (Hons.) in computer science and networks from Telecom ParisTech, in 2010, and the Habilitations (HDR) degree from Sorbonne University, in 2019. Since 2010, he has been a Faculty Member (actually at the rank of an Associate Professor) with Jerusalem College of Technology (JCT)-Lev Academic Center, Jerusalem, Israel, and the Head of FTNet Laboratory. From 2011 to 2012, he was a Kreitman Postdoctoral Fellow with Ben-Gurion University, Israel. He published about 50 papers in international conferences, such as SODA), book chapters, and journals. His main research interests include wireless networks and algorithms for networks, wireless software defined networks (SDN) and ML for 5G cellular networks and beyond, and lightweight security for the IoT. He served on the technical program committee for major IEEE conferences. He was a recipient of Henry and Betty Rosenfelder Outstanding Researcher Award in 2013 and the Computer Science Outstanding Lecturer Award in 2018. He was the Chair of a working group at European COST action on Autonomous Control for a Reliable Internet of Services (ACROSS).



SHULAMIT RECHES received the B.A. degree in mathematics from Jerusalem College, the M.Sc. degree in mathematics and computer science from Bar-Ilan University, and the Ph.D. degree in computer science from Bar-Ilan University, in 2011. She is currently a Senior Lecturer with the Mathematics Department, Jerusalem College of Technology, Israel. Her main research interests include artificial intelligence and algebraic combinatorics.

...