ABSTRACT  Artificial intelligence (AI) provides a promising and novel direction to design future time-varying wireless networks by leading to significantly superior performances compared to conventional methods. In addition, the advanced deployment of unmanned aerial vehicles (UAVs) has boosted extensive novel research results and industrial products in terms of aerial-ground networks. However, with the rapid development of mobile networks and growing requirements for low-latency services, the conventional centralized aerial-ground network has failed to meet the time-varying expectations of mobile users in the dynamic network environment. To cope with the problems, the marriage of the aerial-ground network and innovative AI techniques, i.e., distributed artificial intelligence enabled aerial-ground network (DAIAGN), is proposed in this article, which consists of three vital components: deep reinforcement learning enabled distributed information sharing, edge intelligence enabled distributed security management, and multi-agent reinforcement learning enabled distributed decision making. The functions of the three components are elaborated, and recent related advances are surveyed in detail. A specific case study is also provided with respect to multi-agent reinforcement learning enabled distributed decision making. Furthermore, key challenges and open issues are also discussed to provide some guidances for potential future directions.

INDEX TERMS  Distributed network, aerial-ground network, artificial intelligence, reinforcement learning, edge intelligence.

I. INTRODUCTION  The past few years have witnessed the dramatic increase in the deployment of unmanned aerial vehicles (UAVs) supporting the functions of the ground network in civilian applications, such as traffic monitoring, transportation congestion control, communication enhancement, package delivery, data relaying, etc. The emerging data-craving applications, such as augmented reality (AR) executed on mobile devices, autonomous driving in vehicle-to-vehicle (V2V) networks, and super-resolution video streaming, raises tremendous challenges for the computation capabilities of aerial-ground networks [1]. Meanwhile, the rapidly increasing mobile users and access devices pose an urgent demand for an efficient network architecture enabling low latency and high reliability. However, the conventional centralized aerial-ground network fails to satisfy such demands due to the complex dynamic network environment changes. The time-varying nature of future network environment has significant influences on the communication quality of links.
between agents and remote control centers, and results in potential network congestion or deterioration. Moreover, the heavy data processing workload imposed on the control center makes network configuration inflexible, which further impacts the network efficiency.

The distributed network (DN) is a promising framework to resolve the above problems. In the DN, each agent is enabled to cooperate with other UAVs and make local decisions. Low latency is essentially ensured due to the elimination of communication costs between agents and a remote control center. Furthermore, real-time responses of agents to time-varying environmental changes result in robustness against potential network failures. Despite the distinct advantages of the DN, efficient networking of multiple agents is a challenging issue considering the time-varying network topology and task requirements. As a feasible solution to deal with the time-varying scenarios, the application of artificial intelligence (AI) techniques has become a significant area of research interests because of its online update ability and robustness against emergency. To address the issue aforementioned and fully take advantage of the DN, conventional aerial-ground networks, and state-of-the-art AI techniques such as multi-agent reinforcement learning and blockchain, a novel network architecture, distributed AI enabled aerial-ground network (DAIAGN), is proposed as depicted in Fig. 1. The DAIAGN architecture consists of three crucial components: deep reinforcement learning (DRL)-enabled distributed information sharing, edge intelligence enabled distributed security management, and multi-agent reinforcement learning (MARL) enabled distributed decision making. Although there have been plenty of discussions on AI enabled aerial networks, the cutting-edge techniques such as MARL and federated learning have not been reviewed with respect to distributed aerial-ground networks. This gap inspires our work on DAIAGN.

The remainder of the article is outlined as follows. The mechanism of distributed information sharing, which is the basis of agent cooperation, is discussed in Section II. In Section III, The policies of distributed security management are investigated. Based on the secured shared information, approaches of distributed decision making are reviewed, and a corresponding case study is provided in Section IV. Challenges and open issues are listed in Section V, followed by the conclusion of the article in Section VI.

II. DRL ENABLED DISTRIBUTED INFORMATION SHARING
To facilitate each agent in DAIAGN to utilize information from surrounding agents and respond to time-varying network status changes in real time, distributed information sharing mechanisms are necessary. In the DN, each agent will automatically collect information about surrounding environments, store the information processing history, and analyze resource requirements according to the task. With the help of information shared, the agent can actively evaluate the rewards and practical benefits of its past control decisions, thereby optimizing its decision-making policy and producing more feasible actions to meet time-varying requirements of computing tasks. In terms of centralized information sharing machenisms, the same communication or computing resources can be requested repeatedly or wasted, which leads to inefficient resource allocation. To settle the issues induced by centralized information sharing, the component called distributed information sharing is integrated in DAIAGN, which supports efficient resource utilization and contributes to decision-making optimization. Compared with conventional information sharing that takes place at the center of the network or requires the global knowledge of the entire network, agents share information only with the neighboring ones for distributed information sharing.

The process of distributed information sharing in DAIAGN generally can be divided into two procedures: determining which UAVs and ground users (GUs) share information via dynamic access control and supporting information sharing via efficient computational resources utilization. To be specific, dynamic access control aims at providing a time-dependent access policy according to the time-varying UAV network topology. Moreover, the computational resources utilization policy needs to be adjusted based on the time-varying resource requirements and the access policy.

The problems of dynamic access control and computational resources utilization essentially involve tremendous decision variables and are hard to solve since the corresponding optimization problems are non-convex. Conventional convex optimization based methods reduce the original non-convex problem to several tractable convex subproblems at the expense of accuracy, and the subproblems are enabled to be solved in an iterative manner. Despite the fact that such conventional methods have achieved remarkable success for typical one-shot resource allocation problems in static scenarios, the enforced recalculation of the entire optimization process restricts the real-time application of these methods in dynamic scenarios. To solve the two sequential optimization problems in dynamic scenarios, Markov decision process (MDP) together with its extensions such as partially observable Markov decision process (POMDP) and semi-MDP (SMDP) has been widely adopted to provide a feasible mathematical model. In settings of MDP, the optimization objective of the sequential problem is to maximize an accumulative reward with respect to discrete steps of the assigned task execution period, where the reward of each step is determined according to practical engineering requirements, i.e., reward engineering. DRL is a powerful technique to solve an MDP owing to its robustness against unexpected events in dynamic environments brought by reinforcement learning and feature extraction capability brought by deep neural networks. Although DRL approaches are time-consuming at the training stage, the requirement of timeliness can be satisfied with a trained DRL model at the execution stage in practical applications. Given the end-to-end architecture of DRL, recalculation of the entire optimization process is unnecessary, and time for execution keeps a constant level.
In this section, recent works will be reviewed on DRL enabled dynamic access control and efficient computational resources utilization as listed in Table 1. As can be concluded from Table 1, most recent works focus on the scenario where multiple GUs are served by multiple UAVs, and deep Q learning is recognized as an effective method for distributed information sharing.

A. DRL ENABLED DYNAMIC ACCESS CONTROL

In cases of conventional access control, nodes of the network remain spatially static or temporally steady, and the corresponding optimization problem is essentially time-invariant. Compared to conventional access control, the problem of dynamic access control for the UAV network puts forward a complicated sequential optimization problem coupling the UAV trajectory design and the time-dependent access control policy. In this subsection, recent state-of-the-art works will be discussed on DRL based dynamic access control with various optimization objectives.

1) THROUGHPUT MAXIMIZATION

To reduce the information transmission latency between UAVs and GUs, it is significant to increase the total throughput. The biggest challenge of DRL based throughput maximization lies in the hybrid action space issue, i.e., the action space of each UAV is continuous while that of the GU is discrete. This issue can be addressed by transforming the original discrete actions into samples from a continuous distribution [2]. In consideration of the access limitation and the unified objective of throughput maximization, the authors modeled the joint optimization of UAV trajectory and UAV-GU access control as a mixed cooperative-competitive game, which provides a game-theoretical view of the access control problem.

2) COMMUNICATION EFFICIENCY MAXIMIZATION

In large-scale practical applications, the interaction between numerous GUs and multiple UAVs is complicated in consideration of large-scale time-varying network topology,
complex channel interference and a large amount of data transmission, etc. These factors could lead to a huge state space, e.g., the joint position space of agents grows exponentially with respect to the number of the agents. Simple extensions of communication efficiency maximization algorithms applied in small-scale cases can fail in such large-scale problems. According to mean field theory specially developed for dynamic complex systems, a Fokker-Planck-Kolmogorov equation was constructed to represent the dynamics of each UAV, and the corresponding communication efficiency optimization problem was formulated into an N-player mean field game in [3]. To solve the mean field game, the authors combined trust region policy optimization with neural network feature embedding methods, and proposed a scalable mean field DRL algorithm.

3) NETWORK CAPACITY MAXIMIZATION
In a massive wireless Internet of Things (IoT) network where UAVs relay data from IoT devices to remote servers, probabilistic mutual interference of IoT devices can severely impact the network capacity, and complicates the access control analysis. Moreover, the issue of energy consumption further constrains the network life. To settle the aforementioned problems, the authors in [4] proposed a constrained DRL algorithm based on Lagrangian primal-dual policy optimization. In order to attain the maximal long-term network capacity while ensuring energy sustainability of UAVs, the algorithm dynamically controlled the altitude of UAVs and channel access probability of IoT devices.

B. DRL ENABLED EFFICIENT COMPUTATIONAL RESOURCES UTILIZATION
Considering the sensitive delay requirements of intelligent applications such as autonomous driving, it is inefficient and infeasible to rely on a network center to collect information from all the agents. To support a feasible UAV-GU or UAV-UAV distributed information sharing mechanism, it is of significance to utilize computational resources efficiently. In this subsection, inspiring works on DRL based efficient computational resources utilization will be reviewed.

1) OFFLOADING TASK NUMBER MAXIMIZATION
In DAIAGN, the optimization of offloading task management can be extremely complicated since it involves multi-dimensional resource utilization with respect to UAVs, GUs, multi-access edge computing (MEC) servers and base stations (BSs). This issue was investigated in [7], where a distributive optimization problem was formulated to maximize the number of offloaded tasks while satisfying heterogeneous quality-of-service (QoS) requirements. To solve the optimization problem formulated, the authors proposed a multi-agent deep deterministic policy gradient (MADDPG) based method, which achieved higher delay/QoS satisfaction ratios than the single-agent version and random schemes.

2) COMPUTATION DELAY MINIMIZATION
The computation delay is a crucial issue for AI-enabled applications in distributed aerial-ground networks, where IoT devices generate a large amount of data. Considering the limitation of computation capacity and battery life of IoT devices, UAVs are used to assist computation. Since the computational node capacity, tasks, and channel quality are time-varying, an adaptive computational resource policy is necessary. A DRL based collaborative computation offloading and resource allocation scheme was proposed in [8], where the optimization objective was minimizing computation delay and energy consumption of both IoT devices and UAVs by adaptive learning from the dynamic aerial-ground network.

III. EDGE INTELLIGENCE ENABLED DISTRIBUTED SECURITY MANAGEMENT
With a growing need for government agencies to establish wireless communications with agents for combating crimes and terrorism, security management and privacy preserving in DAIAGN are significant. The demands, which keep the agents free of potential security issues and privacy concerns while maintaining decision making efficiency in dynamic environments, bring forth distributed security management mechanisms. Specifically, in dynamic environments where there exist potential moving eavesdroppers and time-varying jamming attacks, the inefficient conventional centralized security management strategies may fail. Taking U2G communications for example, the U2G channel link is dominated by the line-of-sight (LoS) fading because of the high flying altitude of UAVs, which can be easily wiretapped by the ground or aerial eavesdroppers. In this case, a single UAV cannot obtain the desired secure transmission rate in the presence of multiple potential eavesdroppers. Therefore, it is necessary to adopt a multi-UAV cooperative data transmission scheme to achieve security management in such scenario.

There are mainly two directions for a DN to maintain secure, local data privacy, which intends to protect local data and eliminate security threats at the source end, and information transmission security, which is aimed at protecting communication links and enhancing security in the transmission procedure. In context of edge computing, UAVs and ground IoT devices which possess local computing capability can be treated as edge devices and computing resources. Motivated by this point of view, we apply the emerging edge intelligence (EI), the fruit of the marriage of edge computing and AI, to enable distributed security management following the two aforementioned directions.

As a crucial technique in EI, federated learning (FL), which leaves the private raw data on distributed edge devices instead of sharing the entire local dataset among all devices in common deep learning practice, has been extensively investigated by researchers to address the issue of local data privacy. In terms of information transmission security, EI presents an advanced mechanism, blockchain, which
replaces the conventional centralized security management mechanism. In Section III, recent representative works on FL enabled local data privacy preservation and blockchain enabled transmission security maintenance will be reviewed as listed in Table 2, most of which aim at heterogeneous network settings.

A. FL ENABLED LOCAL DATA PRIVACY PRESERVATION

As a classical problem in privacy and distributed optimization, local data privacy preservation has been extensively investigated by researchers in the optimization and data processing communities. However, conventional methods are unable to tackle the scale of heterogeneous federated networks where there are numerous ground static edge devices such as IoT devices and aerial moving devices (e.g., UAVs). Although FL provides an advanced framework to overcome the problem, system constraints such as energy consumption and communication efficiency can make FL challenging to implement in practice. In this subsection, recent works on FL enabled local data privacy preservation will be reviewed especially involving the aforementioned two problems.

1) ENERGY EFFICIENCY

In UAV-assisted applications in smart cities, data from civilians are required. According to general data protection regulation (GDPR), there are privacy concerns for agencies that prohibit direct sharing of user data. In this case, joint consideration of user data privacy and energy efficiency for UAVs is significant. A feasible example for such joint design is provided in [11], where an FL-based aerial-ground air quality sensing framework was proposed for fine-grained 3D air quality index (AQI) monitoring and forecasting combining light-weight Dense-MobileNet. The authors also introduced a light-weight Dense-MobileNet model in the FL framework to achieve end-to-end learning from haze images to AQI scale distribution.

2) COMMUNICATION EFFICIENCY

For secure control problems involving a massive population of UAVs, local control model of each UAV cannot be directly shared and the complex interaction is difficult to analytically resolve. The straightforward method by considering the interactions among UAVs to make a flock requires an impractically huge inter-UAV communication capability of the UAV network. To settle the problem, a mean field game framework was introduced to establish a theoretical tool for analysis of the control of the UAV flock in [12]. FL was applied to share only neural network parameters among UAVs, and the combination of FL and mean field game further mitigates the communication costs and stability concerns of the UAV network.

B. BLOCKCHAIN ENABLED TRANSMISSION SECURITY MAINTENANCE

The decentralized nature of blockchain makes it a superior choice for transmission security maintenance in the DN. As a result of the distributed consensus mechanism, any transaction can occur and be validated at the edge device end without involvement of third-party devices in context of EI, which essentially eliminates potential threats brought by the transmission between edge devices and third-party devices. The transaction automatically takes place when a certain condition is satisfied, namely smart contracts. All smart contracts are immutable since each device keeps a local record of the committed transactions, which assures the traceability of transaction data. In this subsection, works on blockchain enabled transmission security maintenance will be reviewed from the perspective of transaction traceability.

1) ENERGY TRADING

Considering the limited battery power, the persistence of long-term task execution has become a significant issue in the expansion of UAV-based applications, which gives birth to the area of energy trading between UAVs and charging stations. Conventional energy trading methods face the issues of inefficiency and security during trading. To address the issues, an advanced blockchain framework, based on the tangle data structure, is used to create a distributed network of UAVs and charging stations in [18]. The method allows the UAVs to buy energy from the charging station in exchange for tokens, and a game-theoretic model is used for deciding the buying strategy of energy for UAVs.

2) EDGE COMPUTING RESOURCE ALLOCATION

Due to the open characteristics of UAV communication and mobile edge computing paradigm, there are serious security and privacy issues in edge computing resources allocation between edge computing servers and UAVs. In [19], a resource pricing and trading scheme based on Stackelberg dynamic game was proposed to optimally allocate edge computing resources between servers and UAVs, and blockchain technology was applied to record the entire resources trading process to protect the security and privacy. In the established mechanism, the resource price is controlled by servers, and the UAVs follow the price announced to make optimal decisions on the edge computing resources demands.

IV. MARL ENABLED DISTRIBUTED DECISION MAKING

In dynamic environments, the agents need to work cooperatively and efficiently to meet the time-varying requirements of tasks. The centralized working manner fails to satisfy such demand because of the infeasible communication delay from agents to the remote control center. Instead, the working manner of distributed decision making is employed in DAIAGN. Following the distributed decision making manner, agents (UAVs, IoT devices, ground vehicles, etc.) can autonomously adjust behaviors according to the time-varying network status and environmental changes. Moreover, with the shared information from other agents and collected information from surrounding environments, each agent is enabled to make more rewarding decisions based on its current states. What information should be collected from environments depends
TABLE 2. Edge intelligence based methods for distributed security management.

| Purpose                      | Application            | Proposed   | Method       | Heterogeneous |
|------------------------------|------------------------|------------|--------------|---------------|
| Local Data Privacy Preservation | Data Collection       | [11]        | GCN          | ✓             |
| Local Data Privacy Preservation | Robust Control       | [12]        | MPG Learning | ✓             |
| Local Data Privacy Preservation | Crowdsensing         | [13]        | Q Learning   | ✓             |
| Local Data Privacy Preservation | Ground Object Classification | [14] | CNN          | ✓             |
| Local Data Privacy Preservation | Intelligent Transportation | [15] | Federated Learning | ✓             |
| Local Data Privacy Preservation | UAV Classification    | [16]        | CNN          | ✓             |
| Local Data Privacy Preservation | Data Aggregation      | [17]        | Auction       | ✓             |
| Transmission Security Maintenance | Energy Trading    | [18]        | Stackelberg Game | ✓             |
| Transmission Security Maintenance | Resource Allocation | [19]        | Stackelberg Game | ✓             |
| Transmission Security Maintenance | Resource Cooperation | [20]        | Contract      | ✓             |
| Transmission Security Maintenance | Data Collection      | [21]        | ALP (Blockchain) | ✓             |
| Transmission Security Maintenance | Spectrum Trading    | [22]        | Stackelberg Game | ✓             |
| Transmission Security Maintenance | Intrusion Detection | [23]        | Q Learning    | ✓             |

Abbreviations: GCN: graph convolutional network, CNN: convolutional neural network, MPG: mean field game, ALP: adaptive linear prediction.

on the requirements of the assigned tasks. In terms of moving target tracking tasks with a swarm of UAVs, position information about potential obstacles close to each UAV is important because a UAV must adjust its heading direction and speed to avoid collision.

In the DN, the most significant issue for agents is to form a self-organizing network and cooperatively perform a complex task with time-varying requirements. As an innovative branch of reinforcement learning, MARL has been attracting substantial research interest in the past few years. Many recent research results have shown that MARL is powerful and successful for distributed decision making tasks such as distributed network congestion control, collision avoidance for moving vehicles, distributed UAV-GU access control, etc. Adopting MARL based methods, agents can learn to make the best decisions according to the multi-agent rewards in various problem settings, e.g., average rewards in cooperative settings, conflicting rewards in competitive settings, and heterogeneous rewards in mixed settings.

Existing MARL based works on distributed decision making for aerial-ground network generally fall into two topics: trajectory design and target tracking. The fundamental difference between trajectory design and target tracking lies in the knowledge of final destination. As for trajectory design problems, the final destination is known and UAVs undertake computation tasks while ensuring to reach the destination. Compared with trajectory design, the final destination in target tracking problems is not predictable and UAVs are expected to track the target according to its past trajectory and current state information. In Section IV, we focus on the two representative areas of research interest on distributed decision making and discuss corresponding state-of-the-art techniques based on MARL, as listed in Table 3. As can be deduced from Table 3, MADQL and MADDPG have been widely adopted in distributed decision making and exhibit good performances in both trajectory design and target tracking tasks.

A. MARL ENABLED TRAJECTORY DESIGN

The primary challenge of trajectory design is the coupling of UAV flight actions and resource management in a UAV-assisted communication network. To be specific, the achievable data transmission rates from UAVs to BSs or GUUs are affected by the relative distances between them, which are further concerned with the UAV trajectory and the geographical distribution of BSs or GUUs. In this subsection, MARL based state-of-the-art works on trajectory design will be reviewed and discussed.

1) TRANSMISSION SUCCESS PROBABILITY MAXIMIZATION

As facilities for sensing devices in aerial-ground network, UAVs can transmit sensory data to terrestrial devices. In principle, there are two different transmission modes in terms of distances between UAVs and devices: 1) direct transmission for short distances, namely UAV-to-device (U2D) transmission, and 2) indirect transmission for long distances, which takes ground BSs as relays and relies on BSs to forward the data to devices, namely cellular-enabled transmission. To ensure QoS of multi-UAV-to-multi-device transmission, an MARL based trajectory design algorithm combining deep Q-network (DQN) was proposed in [24] which designed a joint sensing and transmission protocol to coordinate multiple UAVs performing sensing tasks, and meanwhile maximizes the transmission success probability.

2) AGE OF INFORMATION MINIMIZATION

To evaluate the freshness of the collected data, age of information (AoI) has been recognized as the metric, which is defined as the elapsed time after the latest successful transmission of collected data. For brevity of writing, we use AoI of UAVs to represent the AoI of the collected data of UAVs.
Since a high AoI means that the collected data may be temporally inconsistent with current states of the data sources, it is of necessity to minimize AoI during the execution of data collection/sensing tasks. Based on deep deterministic policy gradient (DDPG), a multi-UAV trajectory design algorithm was developed in [25], which minimized the AoI of UAVs given a sensing and transmission protocol.

### 3) COMMUNICATION COVERAGE MAXIMIZATION

To fully leverage maneuverability and flexibility of UAVs, a swarm of UAVs can be organized to cooperatively provide effective communication coverage for a long period of time. In [26], the authors developed a distributed DDPG method to maximize the temporal average coverage score achieved by all UAVs in a task in consideration of the geographical fairness of point-of-interests and total energy consumptions. Different from the original version of DDPG, where UAVs share the same actor and critic network, the distributed DDPG proposed in [26] considered that each UAV has a separate actor and critic network.

### B. MARL ENABLED TARGET TRACKING

Since the trajectory of the target is not known to UAV trackers in target tracking problems, the UAVs are expected to learn from the target’s previous trajectory and/or the images/signals of the target to track the target [31], [32]. For a radio frequency (RF) moving target, authors in [31] proposed a constrained action-based multi-agent Q-learning (MQL) tracking method for UAVs to autonomously make flight decisions that takes into account the uncertainty of reference point location. In [32], an MARL based UAV control method was proposed to track multiple first responders in 3D environments in the presence of obstacles and occlusions. The method selects the optimal joint control actions according to the Cramer-Rao lower bound (CRLB) of the joint measurement likelihood function to achieve high tracking performance.

Different from the aforementioned works, an end-to-end cooperative MARL approach is provided [33], termed as multi-UAV soft actor-critic (MUSAC), which takes only position information of UAVs and the target as input, and outputs tracking strategies of UAV trackers. The advantage of this method is that the single requirement of position information breaks the physical limitation of photographing equipment and image processing ability of UAVs in practical applications. The two-dimensional target tracking with a swarm of rotary-wing UAVs where UAVs and the target move at a fixed altitude is depicted in Fig. 2. We consider that the location of the target at each moment is available to the UAVs, while the motion states of the target, i.e. the flight speed and elevation angle, are unpredictable.

The tracking trajectory is presented in Fig. 3. The basic definitions about the types of tracking failures are demonstrated as follows. Type-1 failure is the essential tracking failure caused by the case where the target is beyond the sensing range of both UAVs, Type-2, Type-3 and Type-4 failures are caused by violation of UAV-target safe distance constraint, UAV network connectivity constraint, and UAV-UAV safe distance constraint, respectively. As can intuitively conclude from the figure, the designed mechanism and the MUSAC algorithm are effective in terms of the formulated target tracking problems. Both UAVs have exhibited the expected ability of intelligent target tracking and autonomous collision avoidance. We continue to demonstrate the reward comparison results of the proposed MUSAC algorithm with the three popular RL baselines. As shown in Fig. 3, the proposed MUSAC outperforms the three baselines in terms of the final convergent mean episode reward significantly.

### V. CHALLENGES AND OPEN ISSUES

The AI enabled aerial-ground distributed network is believed to be promising structure for current and future time-varying dynamic network. We are still facing critical challenges and open issues in order to facilitate its wide range of applications.

### A. FEDERATED REINFORCEMENT LEARNING FOR AERIAL-GROUND NETWORKS

In Section III, we have reviewed some works on FL based security management while all of them focus on supervised learning, which limits the applications. In self-organized aerial-ground networks, a combination of RL and FL essentially breaks such limitation and extend the applicability of FL to more general learning scenarios.
combination, namely federated reinforcement learning, can highly boost the network efficiency and keep local data privacy, which is a promising technique for future aerial-ground networks.

**B. SCALABLE MARL FOR SWARM INTELLIGENCE**

Although there have been some inspiring works on MARL based swarm intelligence as reviewed in Section IV, these works generally suffer from the issue of scalability of a medium-scale problem. The success of mean field methods as reported in [3] results from a statistical modeling of complex interaction between agents for large-scale problems. However, this modeling can fail in medium-scale problems where interactions need more specific consideration and the state/action space is much more complicated. To address this issue, a scalable MARL method for swarm intelligence is necessary.

**C. DECENTRALIZED-TRAINING RL FOR HIGHLY DYNAMIC UAV NETWORKS**

In the research community of RL, the most popular RL methods all follow the centralized-training-decentralized-execution (CTDE) manner, which facilitates training in laboratory. In future highly dynamic UAV networks, the CTDE manner may fail due to requirements of real-time update of neural network parameters. This issue calls for more feasible decentralized-training-decentralized-execution (DTDE) methods to support applications in highly dynamic environments.

**D. HIERARCHICAL BLOCKCHAIN FOR SECURITY MANAGEMENT**

Most existing works on blockchain enabled security management uses a unified blockchain, which will encounter performance deterioration in heterogeneous aerial-ground networks. By contrast, a hierarchical blockchain framework can adapt to various task requirements of different layers of the network. Nevertheless, the application of the hierarchical blockchain is not direct since how to manage the hierarchical blockchain in terms of time-varying network topology is another complex problem.

**E. AI-ENABLED MULTI-AGENT COMBAT IN AERIAL-GROUND NETWORKS**

AI has provided future directions for fighting crime in sense of adoption of multiple autonomous vehicles. Since the criminals/evaders can take initiative action to prevent themselves from being caught, autonomous vehicles/pursuers need to adjust their strategies according to the actions of the criminals. This case is called multi-agent combat in context of game theory. To achieve smart monitoring in cities, AI-enabled multi-agent combat is worth investigation in aerial-ground networks.

**F. RELIABLE COMMUNICATION SUPPORTING DENSE INTELLIGENT AERIAL-GROUND NETWORKS**

Communication supporting common dense networks (e.g., dense IoT networks) has been widely discussed in the community of communication in the past ten years. The dense intelligent aerial-ground network raises up much more complicated issues involving efficient local model parameters sharing, environmental information synchronization, AI-enabled multi-UAV consensus and time series information interaction. These requirements call for more dedicated AI-driven design for reliable communication techniques to support applications of dense intelligent aerial-ground network.
VI. CONCLUSION

Emerging intelligent applications require large amount of data to cope with time-varying requirements in dynamic environments. Substantial efforts from industry and academia are necessary since traditional network architectures are not feasible for such applications. In this article, a novel framework of distributed network is introduced to replace the conventional centralized network to perform complex tasks. The fundamental components of the architecture generally cover the entire runtime of an application, i.e., information sharing, security management, and decision making. As a result, the proposed architecture is flexible enough to be adopted and adapted to different aerial-ground intelligent applications and scenarios. The advantages of MARL enabled distributed decision are assessed by implementing a case study with respect to multi-UAV assisted moving target tracking. In addition, key challenges and open issues were discussed in order to provide an enlightening guidance for future research directions.

To summarize, the research on aerial-ground distributed network is promising, and further research efforts are needed to bring the cutting-edge technology to maturity in the future.

REFERENCES

[1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, “6G wireless networks: Vision, requirements, architecture, and key technologies,” IEEE Veh. Technol. Mag., vol. 14, no. 3, pp. 28–41, Sep. 2019.

[2] R. Ding, Y. Xu, F. Gao, and X. Shen, “Trajectory design and access control for air–ground coordinated communications system with multi-agent deep reinforcement learning,” IEEE Internet Things J., vol. 9, no. 8, pp. 5785–5798, Apr. 2022.

[3] D. Chen, Q. Qi, Z. Zhuang, J. Wang, J. Liao, and Z. Han, “Mean field deep reinforcement learning for fair and efficient UAV control,” IEEE Internet Things J., vol. 8, no. 2, pp. 813–828, Jan. 2021.

[4] S. Khairy, P. Balaprakash, L. X. Cai, and Y. Cheng, “Constrained deep reinforcement learning for energy sustainable multi-UAV based random access IoT networks with NOMA,” IEEE J. Sel. Areas Commun., vol. 39, no. 4, pp. 1101–1115, Apr. 2021.

[5] J. Chen, L. Guo, J. Jia, J. Shang, and X. Wang, “Resource allocation for IRS assisted SFG NOMA transmission: A MADRL approach,” IEEE J. Sel. Areas Commun., vol. 40, no. 4, pp. 1302–1316, Apr. 2022.
[6] Y. Cao, S.-Y. Lien, Y.-C. Liang, K.-C. Chen, and X. Shen, “User access control in open radio access networks: A federated deep reinforcement learning approach,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3721–3736, Jun. 2022.

[7] H. Peng and X. Shen, “Multi-agent reinforcement learning based resource management in MEC- and UAV-assisted vehicular networks,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 131–141, Jan. 2021.

[8] A. M. Seid, G. O. Boateng, S. Anokye, T. Kwantwi, G. Sun, and G. Liu, “Collaborative computation offloading and resource allocation in multi-UAV-assisted IoT networks: A deep reinforcement learning approach,” *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12203–12218, Aug. 2021.

[9] Y.-C. Wu, C. Lin, and T. Q. S. Quek, “A robust distributed hierarchical online learning approach for dynamic MEC networks,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 2, pp. 641–656, Feb. 2022.

[10] X. Xu, B. Shen, S. Ding, G. Srivastava, M. Bilal, M. R. Khoosravi, V. G. Menon, M. A. Jan, and M. Wang, “Service offloading with deep Q-network for digital twinning-empowered internet of vehicles in edge computing,” *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1414–1423, Feb. 2022.

[11] Y. Liu, Y. Ni, X. Li, S. H. Ahmed, W. Y. R. Lim, and C. Miao, “Federated learning in the sky: Aerial-ground air quality sensing framework with UAV swarms,” *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9827–9837, Jun. 2021.

[12] H. Shiri, J. Park, and M. Bennis, “Communication-efficient massive UAV online path control: Federated learning meets mean-field game theory,” *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6840–6857, Nov. 2020.

[13] S. Liu, C. Zheng, Y. Huang, and T. Q. S. Quek, “Distributed reinforcement learning for privacy-preserving dynamic caching,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 3, pp. 749–769, Mar. 2022.

[14] Y. Liu, Z. Ma, X. Liu, S. Ma, and K. Ren, “Privacy-preserving object detection for medical images with faster R-CNN,” *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 69–84, 2022.

[15] W. Y. R. Lim, J. Huang, Z. Xiong, J. Kang, D. Niyato, X.-S. Hua, C. Leung, and C. Miao, “Towards federated learning in UAV-enabled internet of vehicles: A multi-dimensional contract-matching approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5140–5154, Aug. 2021.

[16] K. Huang, X. Liu, S. Fu, D. Guo, and M. Xu, “A lightweight privacy-preserving CNN feature extraction framework for mobile sensing,” *IEEE Trans. Depend. Sec. Comput.*, vol. 18, no. 3, pp. 1441–1455, Jun. 2021.

[17] M. Zhang, L. Yang, S. He, M. Li, and J. Zhang, “Privacy-preserving data aggregation for mobile crowdsensing with externality: An auction approach,” *IEEE Access*, vol. 29, no. 3, pp. 1046–1059, Jun. 2021.

[18] V. Hassija, V. Chamola, D. G. N. Krishna, and M. Guizani, “A distributed framework for energy trading between UAVs and charging stations for critical applications,” *IEEE Access*, vol. 69, no. 5, pp. 5391–5402, May 2020.

[19] H. Xu, W. Huang, Y. Zhou, D. Yang, M. Li, and Z. Han, “Edge computing resource allocation for unmanned aerial vehicle assisted mobile network with blockchain applications,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3107–3121, May 2021.

[20] Y. M. Saputra, D. N. Nguyen, D. T. Hoang, T. X. Vu, E. Dutkiewicz, and S. Chatzinotas, “Federated learning meets contract theory: Economic-efficiency framework for electric vehicle networks,” *IEEE Trans. Mobile Comput.*, vol. 21, no. 8, pp. 2803–2817, Aug. 2022.

[21] X. Xu, H. Zhao, H. Yao, and S. Wang, “A blockchain-enabled energy-efficient data collection system for UAV-assisted IoT,” *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2431–2443, Feb. 2022.

[22] L. Zhai, Y. Zou, J. Zhu, and B. Li, “Improving physical layer security in IRS-aided WPCN multicast systems via Stackelberg game,” *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1957–1970, Mar. 2022.

[23] X. Lu, L. Xiao, G. Niu, X. Ji, and Q. Wang, “Safe exploration in wireless security: A safe reinforcement learning algorithm with hierarchical structure,” *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 732–743, 2022.

[24] F. Wu, H. Zhang, J. Wu, and L. Song, “Cellular UAV-to-device communications: Trajectory design and mode selection by multi-agent deep reinforcement learning,” *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4175–4189, Jul. 2020.

[25] F. Wu, H. Zhang, J. Wu, Z. Han, H. V. Poor, and L. Song, “UAV-to-device underlay communications: Age of information minimization by multi-agent deep reinforcement learning,” *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4461–4475, Jul. 2021.

[26] 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.

[27] L. Wang, K. Wang, C. Pan, W. Xu, N. Aslam, and L. Hanzo, “Multi-agent deep reinforcement learning-based trajectory planning for multi-UAV assisted mobile edge computing,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 73–84, Mar. 2021.

[28] K. K. Nguyen, T. Q. Duong, T. Do-Duy, H. Claussen, and L. Hanzo, “3D UAV trajectory and data collection optimisation via deep reinforcement learning,” *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2358–2371, Apr. 2022.

[29] S. Yin and F. R. Yu, “Resource allocation and trajectory design in UAV-aided cellular networks based on multiagent reinforcement learning,” *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2933–2943, Feb. 2022.

[30] Y.-J. Chen, K.-M. Liao, M.-L. Ku, F. P. Tso, and G.-Y. Chen, “Multi-agent reinforcement learning based 3D trajectory design in aerial-terrestrial wireless caching networks,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8201–8215, Aug. 2021.

[31] Y.-J. Chen, D.-K. Chang, and C. Zhang, “Autonomous tracking using a swarm of UAVs: A constrained multi-agent reinforcement learning approach,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13702–13717, Nov. 2020.

[32] J. Moon, S. Papaioannou, C. Laoudias, P. Kolios, and S. Kim, “Deep reinforcement learning multi-UAV trajectory control for target tracking,” *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15441–15455, Oct. 2021, doi: 10.1109/JIOT.2021.3073973.

[33] Z. Xia, J. Du, J. Wang, C. Jiang, Y. Ren, G. Li, and Z. Han, “Multi-agent reinforcement learning aided intelligent UAV swarm for target tracking,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 931–945, Jan. 2022.

---

**ZHAOYUE XIA** (Member, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree in information and communication engineering. His research interests include resource allocation, optimization, and the applications of machine learning in Internet of Things networks.

---

**JUN DU** (Senior Member, IEEE) received the B.S. degree in information and communication engineering from the Beijing Institute of Technology, in 2009, and the M.S. and Ph.D. degrees in information and communication engineering from Tsinghua University, Beijing, in 2014 and 2018, respectively. From October 2016 to September 2017, she was a Sponsored Researcher and she visited the Imperial College London. Currently, she is an Assistant Professor with the Department of Electrical Engineering, Tsinghua University. Her research interests include communication, networking, resource allocation, and system security problems of heterogeneous networks and space-based information networks. She was a recipient of the Best Student Paper Award from IEEE GlobalSIP, in 2015, the Best Paper Award from IEEE ICC 2019, and the Best Paper Award from IWCMC, in 2020.
YONG REN (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electronic engineering from the Harbin Institute of Technology, China, in 1984, 1987, and 1994, respectively. He worked as a Postdoctoral Researcher at the Department of Electronics Engineering, Tsinghua University, China, from 1995 to 1997, where he is currently a Professor with the Department of Electronics Engineering and the Director of the Complexity Engineered Systems Laboratory (CESL). He holds 60 patents and has authored or coauthored more than 300 technical papers in the behavior of computer network, P2P network, and cognitive networks. He has served as a Reviewer of *IEICE Transactions on Communications*, *Digital Signal Processing*, *Chinese Physics Letters*, *Chinese Journal of Electronics*, *Chinese Journal of Computer Science & Technology*, and *Chinese Journal of Aeronautics*. His current research interests include complex systems theory and its applications to the optimization and information sharing of the internet, the Internet of Things and ubiquitous networks, cognitive networks, and cyber-physical systems.

ZHAN HAN received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was a Research and Development Engineer at JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an Assistant Professor at Boise State University, ID, USA. Currently, he is a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received the NSF Career Award, in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society, in 2011, the EURASIP Best Paper Award for the EURASIP Journal on Advances in Signal Processing, in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS), in 2016, and several best paper awards in IEEE conferences. He was an IEEE Communications Society Distinguished Lecturer, from 2015 to 2018, he has been an AAAS Fellow, since 2019, and he has been an ACM Distinguished Member, since 2019. He is 1% highly cited researcher, since 2017, according to Web of Science.

***