Journal of Applied Automation Technologies

Review Paper

Reinforcement Learning Models for Optimizing Emergency Resource Dispatch Strategies: A Review

Bartosz Dabrowski[1, *] and Rafa Walczak[1]

[1] Faculty of Electronics and Information Technology, Gdansk University of Technology, 80-233 Gdansk, Poland
*Corresponding author: b.dabrowski@pg.edu.pl

Abstract. With socioeconomic development, emergency resource dispatch plays a critical role in responding to sudden incidents. As an emerging machine learning technique, reinforcement learning offers new avenues for optimizing emergency resource dispatch strategies. This paper systematically reviews reinforcement learning models for optimizing emergency resource dispatch strategies. It first delves into the theoretical foundations of reinforcement learning, including Markov decision processes and the core concepts of reinforcement learning algorithms. Subsequently, relevant models are categorized and reviewed based on two dimensions: reinforcement learning algorithm types and characteristics of emergency resource dispatch problems. Furthermore, key influencing factors of reinforcement learning models in emergency resource dispatch are analyzed, such as state space design, action space definition, reward function design, and environment modeling. Current challenges and future development directions are identified, aiming to provide references for further research.

Keywords: *reinforcement learning; emergency resource dispatch; model review; optimization strategy*

Received on 4 Jan 2023, Accepted on 10 March 2023, Published on 18 March 2023

Introduction

In today's society, various emergencies frequently occur, such as natural disasters, accidents, and public health incidents, posing severe threats to people's lives and property safety as well as the stable operation of society [1]. As a critical component of emergency management, emergency resource dispatch aims to swiftly and efficiently allocate and deploy limited emergency resources to the areas of greatest need following an incident. This maximizes the reduction of disaster losses while safeguarding people's lives and ensuring basic living needs [2]. However, emergency resource dispatch faces numerous challenges, including the sudden and unpredictable nature of incidents [2], the limited and uneven distribution of resources [3], time constraints, and complex and dynamic field environments [4]. Traditional dispatch methods often fail to meet practical demands, making research on optimizing emergency resource dispatch strategies critically important [5].

Reinforcement learning, as a machine learning technique based on trial-and-error learning, has achieved remarkable results in numerous fields in recent years, providing new perspectives and methods for solving complex decision-making problems [6]. The core principle of reinforcement learning involves an agent interacting with its environment and learning optimal behavioral strategies based on reward signals from the environment to maximize cumulative rewards [7]. Unlike traditional model-based optimization methods, reinforcement learning does not require precise prior knowledge of the environment's mathematical model. Instead, it adapts to the environment through continuous exploration and learning. This characteristic gives it unique advantages when addressing complex, dynamic, and uncertain emergency resource dispatch scenarios [8]. In emergency material distribution, road conditions may change in real time due to disasters, and demand points may fluctuate over time. Reinforcement learning can dynamically adjust delivery routes and material

DOI: 10.64972/jaat.2023v1.31

allocation plans based on real-time information [9]. In emergency personnel dispatch, it can optimize personnel deployment and task assignment in real time according to on-site rescue progress and personnel status, thereby improving rescue efficiency [10].

Currently, reinforcement learning research in emergency resource scheduling is gaining increasing attention. Scholars have proposed various reinforcement learning-based models from different perspectives [11]. However, these studies remain in a phase of ongoing development and refinement, with several pressing issues requiring resolution [12]. First, significant performance variations exist among different algorithms in terms of convergence speed, stability, and adaptability, necessitating in-depth analysis and comparison [13]. Second, the emergency resource dispatch problem itself exhibits multidimensional complexity. Effectively integrating reinforcement learning models with these problem characteristics to construct targeted and practical dispatch models represents a key research direction [14]. Furthermore, the performance of reinforcement learning models in practical applications is influenced by multiple factors, including state space design, action space definition, reward function design, and environment modeling. The rationality and effectiveness of these factors directly impact the scheduling outcomes and decision quality of the model. However, systematic research and optimization of these critical influencing factors remain relatively insufficient [15-17].

This paper aims to provide a systematic review of reinforcement learning models for optimizing emergency resource dispatch strategies. By thoroughly analyzing the theoretical foundations of reinforcement learning, comprehensively categorizing existing models, and exploring their key influencing factors and challenges, it seeks to offer references for future research. First, this paper will elaborate on relevant reinforcement learning theories, including the core concepts of Markov decision processes and reinforcement learning algorithms, laying the groundwork for subsequent model analysis. Subsequently, existing models are categorized and reviewed from two dimensions: reinforcement learning algorithm types and characteristics of emergency resource dispatch problems. The features, advantages, and limitations of each model category are analyzed. Next, key influencing factors of reinforcement learning models in emergency resource dispatch are explored in depth. Detailed analysis is conducted using practical examples, proposing corresponding optimization strategies and methods. Finally, the paper summarizes major challenges in current research and outlines future development directions, aiming to advance the application and development of reinforcement learning in emergency resource scheduling.

Theoretical Analysis of Reinforcement Learning

Markov Decision Process

The Markov Decision Process (MDP) forms the mathematical foundation of reinforcement learning, providing a framework for modeling sequential decision problems [18]. An MDP consists of a state space (S), action space (A), transition probability (P), reward function (R), and discount factor (γ), as detailed in Table 1. The state space S encompasses all possible environmental states, while the action space A represents the set of actions an agent can execute [19]. Transition probabilities denote the likelihood of moving from states to states after executing action a [20]. The reward function indicates the reward obtained when executing action, a in state s or transitioning to states. The discount factor γ balances the importance of future rewards versus immediate rewards.

Table 1. Components of an MDP

No	Element	Description
1	State Space (S)	Describes the features and conditions of the environment at a certain time
2	Action Space (A)	Actions that the agent can take in a given state
3	Transition Probability (P)	Probability of transitioning to another state after taking an action in a certain state
4	Reward Function (R)	Measures the immediate reward gained by the agent after taking an action in a state
5	Discount Factor (γ)	Balances the weight between immediate and future rewards

The elements outlined in Table 1 collectively establish a comprehensive modeling paradigm that is highly adaptable to a variety of sequential decision-making problems. By abstracting the environment into well-defined states and actions, the MDP framework enables researchers to systematically analyze complex scenarios where uncertainty and change are inherent. In emergency resource dispatch, this abstraction is particularly valuable,

DOI: 10.64972/jaat.2023v1.31

as it allows for the rigorous equationtion of dispatch policies that can adapt dynamically to evolving field conditions and unforeseen events, creating a solid theoretical underpinning for subsequent algorithmic design.

To facilitate a clearer understanding of how the Markov Decision Process (MDP) framework operates within reinforcement learning, it is beneficial to present its structure in a graphical form. This visual representation helps illustrate the relationships between the essential elements—states, actions, transition probabilities, rewards, and discount factors—and how they interact dynamically over time. Figure 1 provides a conceptual overview of the MDP, highlighting the agent's interaction with the environment and the flow of information throughout the process of determining decisions. This figure serves as a useful reference for readers to grasp the foundational mechanics of reinforcement learning models.

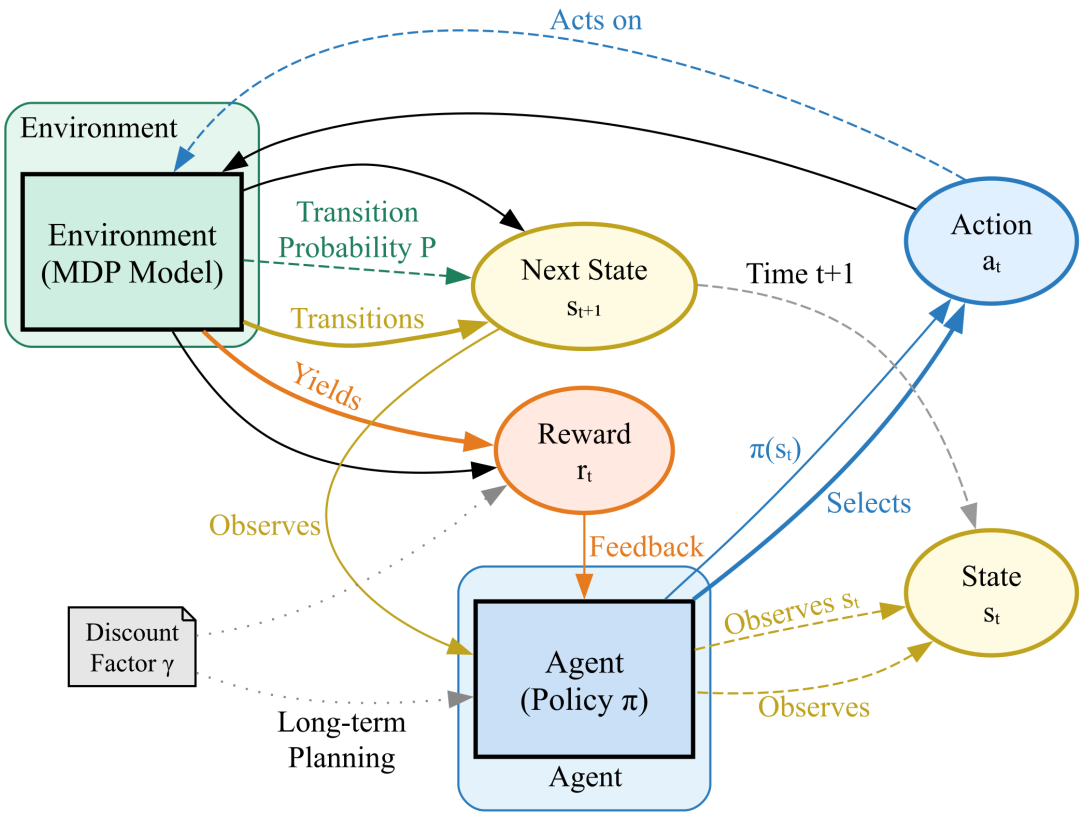

Figure 1. Structure of the Markov Decision Process.

The Markov Decision Process (MDP) diagram clearly illustrates how agents interact with their environment through a closed feedback loop of state observation, action selection, and reward acquisition. This framework is foundational for reinforcement learning, as it systematically models the sequential nature of decision-making under uncertainty. By capturing the dynamic interplay between the agent and its surroundings, the MDP enables researchers to formalize complex dispatch problems and develop tailored optimization strategies. Understanding these interactions is crucial for designing algorithms that can adapt to the rapidly changing conditions typical of emergency resource allocation scenarios.

The core property of an MDP is its Markovian nature, meaning the probability distribution of the next state depends solely on the current state and the selected action, independent of past states [21]. This enables MDPs to simplify complex decision-making processes by transforming them into manageable mathematical forms. The objective of an MDP is to find the optimal policy π that maximizes the cumulative reward obtained by the agent during interactions with the environment [22]. The policy π can be deterministic, selecting a specific action for a given state, or stochastic, choosing different actions with certain probabilities for a given state.

DOI: 10.64972/jaat.2023v1.31

The value function and Q-function are crucial for evaluating policy quality in MDPs. The value function represents the expected cumulative reward starting from state s and following policy π; the Q-function estimates the quality of action a taken at states. The Bellman equation forms a crucial theoretical foundation for reinforcement learning, tightly linking the state value function, action-state value function, and policy to provide mathematical tools for optimizing agent behavior [23]. The calculation methods for the value function and Q-function differ slightly depending on the policy. Markov Decision Processes provide the mathematical foundation for learning reinforcement by formalizing sequential decision-making problems under uncertainty. An MDP is typically defined as:

$$\mathrm{MDP} = (\mathrm{S}, \mathrm{A}, \mathrm{P}, \mathrm{R}, \gamma) \qquad (1)$$

In this definition, S is the set of all possible states, A is the set of available actions, P represents the transition probability function, R is the reward function, and γ is the discount factor. This compact equationtion allows agents to systematically model their environment and make optimal decisions over time.

Core Concepts of Reinforcement Learning Algorithms

The core of reinforcement learning algorithms lies in enabling agents to learn optimal policies through interaction with the environment, thereby maximizing long-term cumulative rewards [24]. These algorithms fall into two major categories: model-based [25] and model-free [26]. Model-based algorithms require constructing an environment model and optimize policies through planning and simulation [25]; model-free algorithms directly learn policies through trial-and-error interactions with the environment [26].

Among these two categories, model-free methods are particularly favored in scenarios where Building a realistic environmental model is challenging or impractical. Popular model-free approaches, such as Q-learning and policy gradient methods, have demonstrated remarkable effectiveness across a variety of complex tasks. These methods enable agents to adapt to dynamic and high-dimensional environments by leveraging value functions or directly optimizing policy parameters. On the other hand, model-based algorithms offer the advantage of improved sample efficiency by utilizing learned or given models for planning, but often face challenges related to model accuracy and computational complexity. Model-based versus model-free options approaches is largely guided by the specific characteristics of the task and the availability of environment information, making reinforcement learning a versatile framework applicable to a broad spectrum of real-world problems.

Policy constitutes one of reinforcement learning's core elements, subdivided into deterministic policies [27] and stochastic policies [28]. Under deterministic policies, the agent always selects the same action for a given state [27]; stochastic policies permit the agent to choose different actions with certain probabilities, enhancing exploratory flexibility [28]. Common reinforcement learning algorithms include Q-learning [29], SARSA [30], policy gradient methods [31], and actor-critic methods [32]. A comparison of reinforcement learning algorithm characteristics is shown in Table 2.

Table 2. Comparison of Reinforcement Learning Algorithm Characteristics

Algorithm Type	Core Features	Applicable Scenarios
Q-learning	Model-free Value-based Tabular or combined with Deep Neural Networks DQN	Small state and action spaces or handling high-dimensional state spaces with deep learning
SARSA	Model-free Value-based Updates considering actually executed actions	Scenarios requiring real-time interaction and online learning
Policy Gradient	Directly optimizes policies Can handle continuous action spaces	High-dimensional action spaces requiring flexible policy representation
Actor-Critic	Combines policy gradient and value estimation Uses two neural networks for collaborative learning	High-dimensional state and action spaces Complex decision problems

As shown in Table 2, each reinforcement learning algorithm type possesses distinct characteristics that make it more suitable for specific types of problem domains. The ability to select or design an appropriate algorithm based on the size of the state-action space and the nature of the dispatch scenario is critical for achieving effective learning and robust decision-making. Additionally, the strengths and weaknesses highlighted in this comparison underscore the importance of tailoring algorithmic approaches to the operational requirements and constraints of emergency resource dispatch, ensuring both practical feasibility and theoretical soundness.

Q-learning is a typical model-free reinforcement learning algorithm that selects optimal actions by learning Q-values for state-action pairs [29]. Its core principle involves iteratively updating the Q-table by adjusting current Q-values based on the current reward and the maximum Q-value for the next state. Q-learning is simple to implement, exhibits good convergence, and is suitable for problems with small state and action spaces. Q-learning is one of the most popular model-free reinforcement learning algorithms. It updates the action-value function iteratively based on agent experience using the following rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max\_\{a'\}Q(s',a') - Q(s,a)] \tag{2}$$

In this equation, $\mathrm{Q(s,a)}$ is the current value estimate for state s and action a, α is the learning rate, r is the received reward, s ' is the next state, and γ is the discount factor. This update mechanism allows the agent to learn optimal behavior through repeated interactions with the environment.

Policy gradient methods directly optimize the parameters of the policy function, maximizing cumulative rewards through gradient ascent [31]. Unlike value-based methods such as Q-learning, policy gradient approaches directly refine the policy itself, enabling handling of continuous action spaces and making them more suitable for complex decision problems. Their core principle involves updating policy parameters by estimating policy gradients, progressively enhancing policy performance. Estimating policy gradients typically requires Monte Carlo sampling or other approximation methods.

The Actor-Critic method combines the strengths of policy gradient and value function estimation by training two neural networks to learn the policy and value function respectively, collaborating to optimize the policy [32]. The actor network generates actions based on the current policy, while the critic network evaluates the quality of these actions and adjusts the actor network's parameters accordingly. This architecture enhances the efficiency and stability of policy updates, making it suitable for complex problems with high-dimensional state and action spaces.

Classification and Review of Emergency Resource Dispatch Models Based on Reinforcement Learning

Classification by Reinforcement Learning Algorithm Type

Q-learning-based Models

Q-learning is a typical model-free, value-based reinforcement learning algorithm that selects optimal actions by learning Q-values for state-action pairs [33]. Its core principle involves iteratively updating the Q-table by adjusting current Q-values based on immediate rewards and the maximum Q-value of the next state. In emergency resource scheduling, [34] developed a Q-learning-based model to address resource allocation and dispatch challenges. In urban emergency relief material dispatch, [35] defines different rescue tasks and resource allocation scenarios as states, while assigning actions to allocate resources to various tasks. Through continuous learning and updating of the Q-table, the model learns optimal resource allocation actions across diverse states, thereby achieving efficient emergency relief material dispatch. Table 3 illustrates the application characteristics of the Q-learning-based emergency resource dispatch model across different scenarios.

Table 3. Application Characteristics of the Q-learning-Based Emergency Resource Dispatch Model

Application Scenario	State Definition	Action Definition	Advantages	Disadvantages
Urban Emergency Relief Material Dispatch	Different rescue tasks and resource allocation situations	Actions allocating resources to different tasks	Can learn optimal resource allocation strategies Improves rescue efficiency	Low learning efficiency when state and action spaces are large
Disaster Area Emergency Communication Resource Dispatch	Status of communication base stations and communication demands	Actions allocating communication resources to different base stations	Optimizes communication resource allocation Enhances communication reliability	Difficulty handling continuous action spaces

Table 3 demonstrates how Q-learning-based models can be flexibly adapted to various emergency resource dispatch scenarios by appropriately defining state and action spaces. These models facilitate the identification of optimal strategies through iterative learning, gradually improving allocation policies as more experience is gathered. Despite their limitations in handling large-scale or continuous spaces, Q-learning models remain a foundational approach, especially in structured environments or when computational resources are limited. Their simplicity and clear interpretability also make them attractive for initial deployments and pilot studies in emergency management.

Policy Gradient-Based Models

In recent years, the complexity of emergency resource dispatch scenarios has underscored the need for more flexible and adaptive decision-making frameworks. Traditional optimization methods, which often rely on discrete and static representations of resources and actions, may fall short in capturing the nuanced dynamics and uncertainties inherent in real-world emergencies. Reinforcement learning, by contrast, offers a dynamic and data-driven approach capable of continuously improving dispatch strategies through direct interaction with the environment. This adaptability is particularly valuable when resource demands fluctuate and environmental conditions evolve rapidly. Among the various reinforcement learning paradigms, approaches that can effectively manage continuous variables—such as resource quantities, vehicle trajectories, or supply levels—are especially promising. The ability to model and optimize in continuous spaces allows for more granular and precise control, which is vital for maximizing the efficiency and responsiveness of emergency operations.

Policy gradient methods directly optimize the parameters of the policy function by maximizing cumulative rewards through gradient ascent [36]. Unlike value-based Q-learning algorithms, policy gradient approaches directly optimize the policy itself, enabling them to handle continuous action spaces and making them more suitable for complex decision-making problems [36-37]. In emergency resource dispatch, policy gradient-based models have been applied to address continuous resource allocation and complex dynamic scheduling problems [38]. For path planning and resource allocation of emergency rescue vehicles, policy gradient models effectively handle continuous state and action information such as vehicle speed and fuel levels. By continuously adjusting policy parameters, they optimize rescue vehicle routes and resource allocation schemes, enhancing rescue efficiency [39]. Table 4 compares the characteristics of policy gradient models versus other models when addressing continuous action space problems.

Table 4. Comparison of Characteristics Between Policy Gradient Models and Other Models for Continuous Action Space Problems

Model Type	Handling Ability	Policy Optimization Method	Convergence Speed	Stability
Policy Gradient	Strong	Directly optimize policy parameters	Slower	Better
Q-learning	Weak	Indirect via Q-value table	Faster	Worse
DQN	Moderate	Approximate Q-value with deep neural networks	Faster	Moderate

The comparative analysis in Table 4 highlights the practical considerations when selecting reinforcement learning algorithms for continuous action space problems in emergency resource dispatch. Policy gradient methods, by directly optimizing the policy, offer greater stability and flexibility in complex, real-world scenarios where discrete representations may fall short. The trade-offs between convergence speed and stability must be carefully balanced according to the urgency and complexity of the dispatch tasks. These insights provide valuable guidance for practitioners seeking to match algorithmic choices to the unique demands of their operational environments.

Deep Reinforcement Learning-Based Models

As emergency resource dispatch scenarios become increasingly complex, traditional reinforcement learning algorithms often encounter limitations when dealing with high-dimensional state and action spaces or highly nonlinear relationships among variables. The need to process vast amounts of heterogeneous data—such as geographic information, resource inventories, traffic flows, and environmental factors—demands more powerful and scalable modeling techniques. To address these challenges, researchers have explored the integration of deep learning with reinforcement learning, giving rise to a new class of algorithms capable of capturing intricate patterns and dependencies within the data. By employing deep neural networks, these methods can automatically extract latent features from raw input, reducing the need for manual feature engineering and enabling more flexible adaptation to dynamic environments. This synergy between deep

DOI: 10.64972/jaat.2023v1.31

learning and reinforcement learning significantly broadens the range of problems that can be effectively tackled, particularly in domains where traditional methods struggle to scale or generalize.

Deep reinforcement learning combines the strengths of deep neural networks and reinforcement learning, enabling automatic extraction of state features and learning of complex decision functions [40]. Common deep reinforcement learning algorithms include Deep Q-Network (DQN) [41] and Deep Deterministic Policy Gradient (DDPG) [42]. In emergency resource dispatch, deep reinforcement learning models are widely applied to address challenges involving large-scale, high-dimensional state spaces and complex dynamic environments [43]. Reference [43] demonstrates that DQN models can resolve high-dimensional state spaces and intricate decision-making problems in large-scale emergency resource dispatch. By leveraging deep neural networks for state feature extraction and value function approximation, these models enhance the accuracy and efficiency of dispatch decisions. DDPG excels in emergency resource scheduling problems involving continuous action spaces. For instance, in multi-agent emergency rescue tasks, it coordinates actions among agents to achieve efficient resource allocation and task assignment [44]. Table 5 lists several common deep reinforcement learning algorithms along with their application scenarios and characteristics in emergency resource scheduling.

Table 5. Applications of Deep Reinforcement Learning Algorithms in Emergency Resource Scheduling

Algorithm Type	Application Scenario	Advantages	Disadvantages
DQN	Large-scale emergency resource dispatch	Can handle high-dimensional state spaces Improves decision accuracy	Parameter sensitive Prone to overfitting
DDPG	Emergency resource dispatch with continuous action spaces	Suitable for continuous action spaces Coordinates multi-agent actions	Requires large sample size Long training time
Dueling	Emergency resource dispatch requiring accurate state value estimation	More accurate state value estimation Improves learning efficiency	Complex structure Difficult to implement

The information in Table 5 illustrates the versatility and power of deep reinforcement learning in addressing diverse emergency resource scheduling challenges. By leveraging the feature extraction capabilities of neural networks, these algorithms significantly enhance the agent’s ability to process high-dimensional and complex environmental data. However, the increased computational requirements and training complexity necessitate careful consideration of resource constraints and deployment environments. This underscores the importance of algorithm selection and parameter tuning in achieving effective and scalable solutions for real-world emergency scenarios.

In deep reinforcement learning, Deep Q-Networks (DQN) use neural networks to approximate the Q-function for high-dimensional state spaces. The loss function minimized during DQN training is defined as:

$$L(\theta) = E\_\{(s,a,r',s') \sim D\}[(r + \gamma \max\_\{a'\}Q(s',a';\theta^\wedge -) - Q(s,a;\theta))^\wedge 2] \tag{3}$$

Here, θ denotes the parameters of the current Q -network, $\theta^\wedge$ - represents the parameters of the target network, and D is the experience replay buffer. This loss encourages the network to predict Q-values that are consistent with the Bellman optimality equation, enabling learning in large-scale problems.

Classification of Emergency Resource Dispatch Problems by Characteristics

Dynamic Emergency Resource Dispatch Models

In the field of emergency management, the ability to swiftly and efficiently allocate resources under uncertain and fluctuating conditions is paramount. Emergency resource dispatch scenarios are inherently characterized by unpredictability, stemming from factors such as sudden-onset disasters, shifting demands, and rapidly changing on-ground realities. Traditional static resource allocation methods, which often rely on pre-defined plans and rigid distribution rules, frequently fall short in addressing the unpredictable evolution of emergencies. As a result, there has been increasing interest in the development of intelligent dispatch models capable of perceiving environmental changes and autonomously adjusting allocation decisions. These models leverage advances in machine learning and artificial intelligence to enable a higher degree of adaptability and

DOI: 10.64972/jaat.2023v1.31

responsiveness. By continuously monitoring key environmental indicators and feedback from ongoing operations, intelligent dispatch systems can identify emerging needs, reprioritize resource distribution, and coordinate actions across multiple agents, ultimately supporting more effective and resilient emergency response strategies.

Emergency resource dispatch typically occurs within dynamically changing environments, such as evolving disasters, resource depletion, and replenishment [45]. Dynamic emergency resource dispatch models can adjust resource allocation strategies in real time based on environmental changes. Dynamic emergency resource allocation models based on deep reinforcement learning learn online and adapt to environmental dynamics, promptly allocating resources to areas of greatest need to enhance the timeliness and effectiveness of emergency response [46]. These models typically account for temporal factors and state transitions, employing methods such as temporal difference learning to update value functions or policies. Table 6 illustrates the application characteristics of dynamic emergency resource allocation models across different disaster scenarios.

Table 6. Application Characteristics of Dynamic Emergency Resource Scheduling Models in Different Disaster Scenarios

Disaster Scenario	Model Features	Advantages	Disadvantages
Forest Fire Emergency Dispatch	Considers fire spread dynamics and resource consumption	Can respond timely to fire changes Reasonably allocate resources	High real-time data requirements High computational complexity
Flood Emergency Rescue Dispatch	Adjusts resource allocation based on flood water level changes and affected population dynamics	Improves rescue efficiency Reduces disaster losses	High environmental uncertainty Model robustness needs improvement

From the perspectives outlined in Table 6, it is clear that dynamic emergency resource scheduling models must be responsive and adaptable to rapidly changing disaster environments. Incorporating real-time data and environmental dynamics enhances the timeliness and appropriateness of allocation decisions, ultimately strengthening the effectiveness of emergency response efforts. Nevertheless, the increased complexity and data requirements highlight the need for robust simulation and data management tools, ensuring that models remain both accurate and computationally feasible in practice.

Given the complexity of deep reinforcement learning algorithms, especially in the context of emergency resource dispatch, a schematic diagram can enhance our comprehension of the main technical workflow. Visualizing the Deep Q-Network (DQN) architecture clarifies how state information is processed by neural networks to estimate optimal actions. As shown in Figure 2, the DQN structure encompasses input layers for environmental states, hidden layers for feature extraction, and output layers for Q-value prediction, thus supporting efficient learning in high-dimensional spaces. This illustration makes the underlying algorithmic process more accessible and tangible to researchers and practitioners.

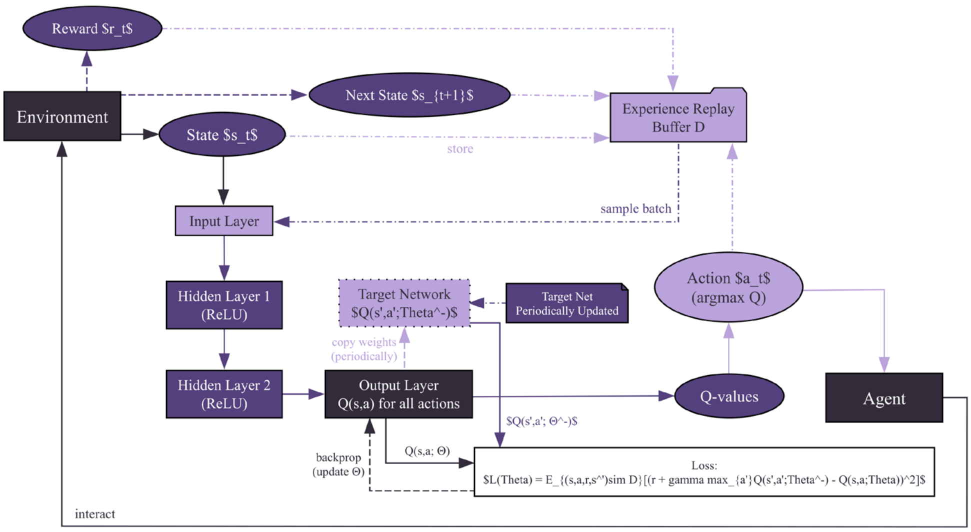

Figure 2. Deep Q-Network architecture.

DOI: 10.64972/jaat.2023v1.31

The Deep Q-Network (DQN) architecture depicted here demonstrates the integration of deep learning with reinforcement learning principles. By utilizing layered neural networks, DQN is capable of extracting rich features from high-dimensional state spaces, making it particularly suitable for complex environments such as emergency resource scheduling. This approach allows the agent to approximate value functions more accurately, facilitating more effective decision-making. The visualization highlights the core workflow where state inputs are transformed into Q-value predictions, thereby supporting robust policy learning. Such architectures have opened new possibilities for applying reinforcement learning to real-world operational challenges in emergency management.

Multi-Objective Emergency Resource Dispatch Model

In the realm of emergency resource dispatch, real-world operations are rarely driven by a single objective. Instead, decision-makers must often contend with a wide array of sometimes conflicting goals, such as responding swiftly to emergencies, utilizing limited resources in the most effective manner, and ensuring that operational costs remain manageable. The inherently complex and uncertain nature of emergency scenarios further amplifies these challenges. Traditional single-objective optimization methods, while useful in straightforward contexts, frequently fall short in capturing the full scope of operational priorities present in emergency management. As such, the development of models capable of addressing multiple objectives simultaneously has become a focal point for both researchers and practitioners. By integrating various performance indicators into the decision-making process, these multi-objective models provide a more holistic perspective, enabling dispatch strategies that are not only efficient but also equitable and sustainable. Leveraging advances in reinforcement learning, these approaches are able to dynamically balance trade-offs among competing objectives, adapting to evolving conditions and stakeholder requirements throughout the resource allocation process.

In practical emergency resource dispatch, multiple objectives often need to be considered, such as minimizing rescue time, maximizing resource utilization efficiency, and minimizing rescue costs [47-49]. The multi-objective emergency resource dispatch model constructs a multi-objective optimization function and employs reinforcement learning algorithms to balance and optimize among these objectives [50]. Reference [51] established a reinforcement learning-based multi-objective emergency resource scheduling model. It uses rescue time and resource demand satisfaction as objective functions, converting the multi-objective function into a single-objective function for optimization by assigning different weights. These models achieve balance among different objectives, providing more comprehensive and reasonable scheduling solutions. Table 7 lists common objective types in multi-objective emergency resource scheduling models and their weighting methods.

Table 7. Objective Types and Weighting Methods in Multi-Objective Emergency Resource Scheduling Models

Objective Type	Weight Setting Method	Advantages	Disadvantages
Minimize Rescue Time	Assign higher weights based on expert experience and historical data	Clearly emphasizes key points Improves rescue efficiency	Subjective weight determination
Maximize Resource Demand Satisfaction	Combine expert opinions and actual demand	Considers resource guarantee Enhances rescue effect	Requires accurate demand forecasting
Minimize Rescue Cost	Assign weights considering resource value and budget constraints	Reduces rescue cost Improves resource utilization	May conflict with other objectives such as rescue time

Table 7 emphasizes the necessity of balancing multiple objectives in emergency resource scheduling. Assigning appropriate weights to each goal enables the equationtion of composite optimization objectives that reflect the real-world priorities of decision-makers. This approach allows for the reconciliation of potentially conflicting demands—such as minimizing time versus cost—leading to more holistic and practical scheduling solutions. The appropriate calibration of these weights remains a challenging but essential aspect of multi-objective model design, warranting further research and methodological innovation.

In emergency resource dispatch, multiple objectives such as minimizing time and cost while maximizing efficiency must be balanced. A widely used approach is to assign weights to each objective and combine them into a single optimization function:

$$\min F = w\_1 T + w\_2 C - w\_3 S \qquad (4)$$

In this function, F is the composite objective, T denotes total rescue time, C is total cost, S is satisfaction or resource utilization, and w_1, w_2, w_3 are the respective weights. This weighted sum allows decision-makers to tune the trade-off among different goals according to operational priorities.

Multi-Agent Emergency Resource Dispatch Models

In recent years, the increasing complexity and scale of emergency scenarios—such as natural disasters, industrial accidents, and large-scale public health emergencies—has highlighted the limitations of single-agent resource dispatch approaches. Traditional centralized decision-making frameworks often struggle to cope with rapidly evolving environments, distributed information, and the need for real-time adaptability. As emergency operations frequently involve multiple teams, vehicles, or autonomous systems operating in parallel, there is a growing demand for decentralized models that can support flexible and coordinated resource allocation. Multi-agent systems have emerged as a promising solution by distributing intelligence and autonomy across multiple entities, allowing for concurrent decision-making and local adaptation. By leveraging reinforcement learning techniques, these systems can dynamically learn optimal collaboration strategies, balance workload distribution, and mitigate conflicts among agents. The integration of multi-agent reinforcement learning not only enhances scalability and fault tolerance but also paves the way for more resilient and responsive emergency management frameworks.

In complex emergency rescue scenarios, collaboration and coordination among multiple agents are typically required [52]. Multi-agent emergency resource dispatch models treat each agent as an independent decision-making unit, utilizing reinforcement learning algorithms to enable communication and cooperation among agents, thereby jointly accomplishing emergency resource dispatch tasks [53]. A dynamic task allocation model for aerial flood rescue based on multi-agent reinforcement learning leverages the advantages of multi-agent reinforcement learning in handling distributed dynamic problems, achieving efficient task allocation [54]. These models fully utilize the distributed computing power and collaborative capabilities of multiple agents to enhance the efficiency and robustness of emergency resource scheduling. Table 8 compares the characteristics of multi-agent and single-agent emergency resource scheduling models.

Table 8. Comparison of Characteristics Between Multi-Agent and Single-Agent Emergency Resource Dispatch Models

Model Type	Decision Unit	Collaboration Ability	Complexity	Applicable Scenarios
Multi-Agent Model	Multiple agents	Strong	High	Complex emergency rescue scenarios involving multiple participants
Single-Agent Model	Single agent	Weak	Low	Simple emergency resource dispatch problems Centralized decision-making

As illustrated in Table 8, the distinction between multi-agent and single-agent models has significant implications for the complexity and effectiveness of emergency resource dispatch. Multi-agent approaches offer improved collaboration and scalability, which are crucial in complex, distributed rescue operations. However, they also introduce additional challenges in communication and coordination, demanding sophisticated algorithmic solutions. By understanding these trade-offs, researchers and practitioners can better design dispatch systems that align with the organizational structure and operational demands of their emergency response frameworks.

Analysis of Key Influencing Factors for Reinforcement Learning Models in Emergency Resource Dispatch

State Space Design and Feature Extraction

In the context of reinforcement learning, the design of the state space is a foundational step that directly influences the model's ability to learn effective policies. A well-constructed state space encapsulates the essential information required for the agent to accurately interpret the environment and make informed decisions. In emergency resource dispatch scenarios, this involves identifying and incorporating environmental variables, operational constraints, and key metrics that capture the evolving needs and situational complexities inherent to disaster response. State representation must strike a balance between comprehensiveness and tractability, ensuring that relevant features are neither omitted nor excessively detailed to the point of impeding learning efficiency.

The state space serves as the input interface for reinforcement learning models, reflecting the model's ability to perceive the environment. In emergency resource dispatch, its design must encompass all critical factors affecting decision-making [55]. As demonstrated in [56], the state space for post-disaster emergency material dispatch may include regional disaster severity, demand volume, material storage point locations and inventory levels, transportation vehicle positions, and load status.

Beyond the core physical and logistical elements, advanced state representations may also integrate various contextual and operational factors. For instance, agent-specific attributes such as resource allocation history, priority levels of affected zones, and recent changes in environmental conditions can be encoded to enhance situational awareness. Additionally, the inclusion of dynamic factors—such as evolving road accessibility, communication network status, and resource consumption rates—enables the model to adapt more flexibly to real-world uncertainties. Careful feature engineering and dimensionality reduction techniques, such as principal component analysis or autoencoders, can further optimize the state space, reducing redundancy and improving computational efficiency. Ultimately, the effectiveness of reinforcement learning in emergency dispatch depends largely on how well the state space captures the multifaceted realities of the operational environment.

Feature extraction is a critical step in state space design. Through dimensionality reduction and feature selection, redundant information is eliminated while retaining the most valuable features for decision-making, thereby enhancing model efficiency and performance [57]. Principal Component Analysis (PCA) and Deep Autoencoders are commonly used feature extraction methods [58]. State space design prioritizes different aspects across scenarios: urban emergency response emphasizes resource distribution and traffic conditions, while forest firefighting focuses on fire dynamics and geographic information such as fire location, spread rate, wind direction, and wind speed [59-60]. Feature extraction methods are compared in Table 9.

Table 9. Feature Extraction Methods Comparison

Method Name	Applicable Scenario	Advantages and Disadvantages
Principal Component Analysis	Pure numerical linear correlated high-dimensional data	Advantages High computational efficiency Disadvantages May lose nonlinear features
Autoencoder	Pure complex nonlinear correlated high-dimensional data	Advantages Strong feature extraction ability Disadvantages High computational resource requirements

Table 9 highlights the importance of selecting appropriate feature extraction methods in the construction of effective state spaces. The choice between linear and nonlinear techniques depends on the nature of the input data and the complexity of the scenario. Proper feature extraction not only improves model performance but also reduces computational burdens, which is particularly beneficial in time-sensitive emergency response situations. Continued refinement of these methods will further enhance the adaptability and precision of reinforcement learning-based dispatch models.

Action Space Definition and Decision Granularity

A well-designed action space is fundamental to the success of reinforcement learning models in emergency resource dispatch. The definition of the action space directly determines the agent’s ability to interact with and influence its environment, thus shaping the learning process and the ultimate effectiveness of dispatch decisions. In practical applications, the action space should closely reflect the operational realities and constraints of emergency response, incorporating factors such as resource availability, logistical limitations, and regulatory requirements. Moreover, it is essential to ensure that the action space remains both expressive and computationally manageable. Overly simplistic action spaces may fail to capture the complexity of real-world decision-making, whereas excessively complex or high-dimensional action spaces can impede learning efficiency and hinder model performance. By carefully balancing expressiveness and tractability, researchers can construct action spaces that facilitate effective learning while supporting the nuanced decision-making required in dynamic and uncertain emergency scenarios.

The action space defines the set of operations an agent can execute in a given state, while decision granularity impacts model complexity and decision accuracy [61]. In emergency resource dispatch, action space design must consider practical feasibility and constraints. As described in [62], actions may include allocating a specific quantity of supplies to a target location, dispatching vehicles carrying designated cargo to a destination, or activating/deactivating power generation equipment. Regarding decision granularity, finer granularity enables more precise resource scheduling control but increases model complexity and computational load [63]. Coarser granularity simplifies the model but may reduce decision accuracy, requiring trade-offs based on specific problem contexts [64]. The impact of different decision granularities on the action space is compared in Table 10.

Table 10. Comparison of the Impact of Different Decision Granularities on the Action Space

Decision Granularity	Action Space Complexity	Decision Accuracy	Computational Resource Requirement
Fine	High	High	High
Medium	Medium	Medium	Medium
Coarse	Low	Low	Low

The comparison presented in Table 10 underscores the critical impact of decision granularity on both model complexity and scheduling accuracy. Fine-grained actions enable more nuanced and precise resource deployment, which can be vital in high-stakes emergencies. However, this increased precision comes at a cost in terms of computational demand and model complexity. Striking the right balance between granularity and feasibility is essential for ensuring that reinforcement learning models remain practical and effective in real-world dispatch operations.

The choice of decision granularity significantly influences the efficiency and effectiveness of reinforcement learning models in emergency resource scheduling. To better demonstrate this impact, a comparative visualization is provided. Figure 3 depicts the relationship between different levels of decision granularity and model performance, including computational complexity and scheduling accuracy. By examining this figure, readers can more intuitively appreciate the trade-offs involved in designing the action space for practical applications.

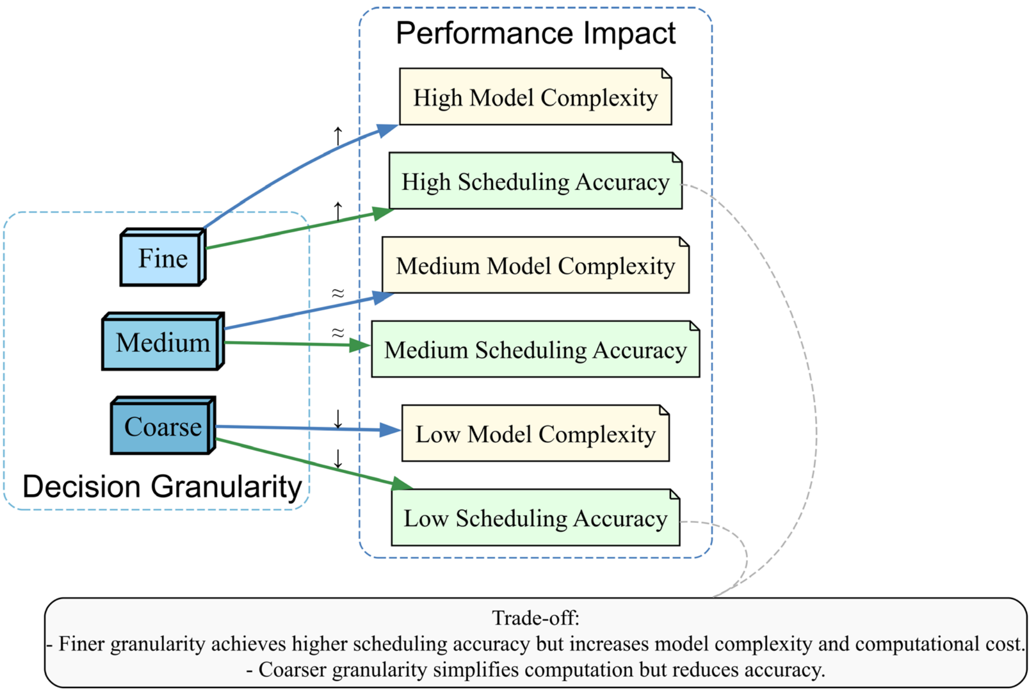

Figure 3. Effect of decision granularity on performance.

This figure provides a comparative view of how decision granularity influences the performance of reinforcement learning models in emergency resource dispatch. By visualizing the trade-offs between computational complexity and scheduling accuracy, it becomes evident that finer-grained decision-making allows for more precise resource control, while also demanding greater computational resources. Conversely, coarser granularity reduces model complexity at the possible expense of solution quality. Understanding this balance is essential for practitioners aiming to deploy scalable, efficient solutions that remain practical under operational constraints. The insights gained from such analysis can guide the customization of action spaces to best fit the needs of diverse emergency scenarios.

Reward Function Design and Optimization Objective Implementation

Designing an effective reward function is a fundamental challenge in reinforcement learning applications for emergency resource dispatch. The reward function serves as the primary channel through which operational priorities and constraints are translated into actionable learning signals for the agent. In practice, reward structures must be capable of capturing the often complex and competing objectives present in real-world emergency scenarios, such as efficiency, fairness, and risk minimization. A thoughtfully designed reward function should reflect both immediate and long-term goals, incentivizing the agent not only to complete urgent deliveries but also to allocate resources judiciously and adapt responsively to changing circumstances. Moreover, the function must be carefully calibrated to prevent unintended behaviors, such as resource hoarding or neglecting less accessible regions. In addition to direct performance metrics, auxiliary factors—such as system resilience, equity among affected areas, and compliance with operational guidelines—can be integrated into the reward framework to guide more holistic decision-making. By establishing a nuanced and context-aware reward system, reinforcement learning models can better align with the multifaceted demands of emergency management and drive continual improvement in dispatch strategies.

The reward function is pivotal in reinforcement learning, as it determines the learning direction of the agent. It must be meticulously designed based on the actual application scenario to ensure the dispatch system achieves its intended objectives [65]. In emergency resource dispatch, positive feedback can be established—for instance, awarding higher scores for successfully delivering urgently needed supplies to remote areas. Conversely, negative feedback can be implemented—such as deducting corresponding points for exceeding the budget. Design must balance multiple objectives to avoid model bias caused by single-metric focus [66-68]. For forest firefighting dispatch, considerations include extinguishing efficiency alongside rescuer safety and resource consumption costs [69].

Designing an effective reward function is essential for guiding the learning process in emergency dispatch scenarios. A typical composite reward function may be equationted as:

$$R = \lambda\_1 \times \text{ Delivered } - \lambda\_2 \times \text{ Delay } - \lambda\_3 \times \text{ Cost } \tag{5}$$

Here, Delivered refers to the quantity of resources delivered, Delay is the time taken, Cost is the total expenditure, and $\lambda\_1, \lambda\_2, \lambda\_3$ are the coefficients balancing these objectives. This structure provides flexibility to shape the agent's behavior in line with operational requirements.

Environmental Modeling and Simulation

Environmental modeling is fundamental to the advancement of reinforcement learning research in emergency resource dispatch. Real-world emergencies are characterized by high complexity, unpredictability, and a multitude of interacting factors, making direct experimentation both risky and impractical. By constructing representative models of disaster scenarios, researchers can systematically explore the effects of different variables—including resource availability, infrastructure integrity, and agent decision-making—on overall system performance. Such simulations not only facilitate the evaluation of novel algorithms under controlled conditions but also support the identification of potential bottlenecks and vulnerabilities in existing response strategies. Moreover, the use of synthetic environments enables the replication of rare yet critical events, which are otherwise difficult to observe or reproduce in real life, thereby supporting the development of robust and resilient dispatch policies.

Given the difficulty of directly using real-world emergency scenarios for experimentation and data collection, environmental modeling and simulation serve as crucial tools for researching reinforcement learning applications in emergency resource dispatch [70]. By establishing accurate environmental models, disaster scenarios, resource dynamics, and agent behaviors can be simulated, providing training and testing platforms for reinforcement learning algorithms. Common simulation tools include MATLAB [71], Simulink [72], and

AnyLogic [73], which enable visualization and dynamic simulation of emergency resource dispatch processes, helping researchers better understand and optimize dispatch strategies [74].

In addition, environmental simulation offers a flexible and cost-effective means for benchmarking different reinforcement learning approaches across a variety of operational contexts. Scenarios such as multi-agent coordination, logistics under road disruptions, and allocation under uncertain demand can all be emulated, allowing for comparative analysis of algorithm performance and generalizability. These virtual testbeds also support iterative model refinement, as researchers can rapidly adjust parameters, introduce new constraints, or simulate policy interventions without the risks or costs associated with real-world deployment. Through continuous simulation and validation, the insights gained can ultimately inform the design of more effective, data-driven emergency resource dispatch systems, bridging the gap between theoretical research and practical application.

Challenges and Future Development Directions

Challenges

Despite the significant progress made in applying reinforcement learning to emergency resource dispatch, the practical deployment of these models continues to encounter a range of substantial obstacles. The inherently dynamic and uncertain nature of emergency scenarios often requires reinforcement learning systems to operate under unpredictable and rapidly evolving conditions. Moreover, the variability of disaster scales, resource availability, and communication constraints poses considerable challenges for both algorithm design and real-world implementation. Many existing studies rely on idealized simulations or simplified assumptions that do not fully capture the complexity found in actual emergencies. As a result, there is often a gap between academic research and operational application, with issues such as data sparsity, incomplete information, and system robustness requiring further attention. Addressing these practical concerns is crucial for enhancing the reliability and adaptability of reinforcement learning-based dispatch strategies in real-world emergency management.

Reinforcement learning optimization of emergency resource dispatch strategies faces challenges including complex environment modeling, high real-time requirements, multi-agent coordination, and data privacy and security [75].

Complex Environment Modeling

Real-world emergency scenarios involve highly dynamic and uncertain environments, including random disaster occurrences, dynamic resource availability, and real-time traffic changes, making precise modeling challenging [76]. Developing environment models that accurately capture these complexities remains a critical hurdle for applying reinforcement learning in emergency resource allocation. In natural disaster relief, factors like aftershocks, flood expansion, and sudden weather shifts can impact rescue operations and resource demands, yet these elements are difficult to capture with precise mathematical models [77-78].

High Real-Time Requirements

In emergency response, time is critical, demanding extremely high real-time performance for resource allocation [79]. However, the training and decision-making processes of reinforcement learning models may require significant time, particularly in complex models and large-scale state spaces, making it difficult to meet real-time scheduling demands [80]. Deep reinforcement learning models typically require vast amounts of training data and computational resources, with training processes potentially lasting hours or even days—an approach impractical in actual emergency scenarios [81].

Multi-Agent Coordination

Achieving efficient coordination and cooperation among agents to avoid conflicts and resource wastage poses a significant challenge in multi-agent emergency resource allocation [82]. Multiple agents may possess differing objectives and interests, making the design of effective communication mechanisms and coordination strategies

DOI: 10.64972/jaat.2023v1.31

for collaborative operation an urgent issue. In multi-UAV emergency rescue missions, drones must avoid collisions while rationally dividing tasks to collectively accomplish rescue operations [83].

Data Privacy and Security

When training models using real-world emergency data, data privacy and security are paramount. Emergency data may contain sensitive information, such as personal details of affected individuals or detailed rescue operation plans, necessitating robust technical safeguards. During public health events, patient personal information must be strictly protected to prevent data breaches [84].

Future Development Directions

Addressing the aforementioned challenges, this section outlines future development directions for reinforcing learning-based optimization of emergency resource allocation strategies: 1) Enhance interdisciplinary integration with operations research, systems engineering, computer science, and other fields. Leverage theories and methodologies from these disciplines to further refine reinforcement learning models, thereby improving the scientific rigor and effectiveness of emergency resource allocation. 2) Continuously refine existing reinforcement learning algorithms to enhance their adaptability, convergence speed, and decision accuracy in complex environments. 3) Leverage advanced hardware technologies such as cloud computing, edge computing, and the Internet of Things (IoT) to provide robust computational support and data transmission guarantees for training and deploying reinforcement learning models, thereby improving model real-time performance and response speed. 4) Strengthen the integration of theoretical research with practical applications. Conduct more case studies and demonstration projects to validate and refine reinforcement learning models for optimizing emergency resource allocation, driving their widespread adoption in emergency management.

Conclusion

This paper provides a systematic review of reinforcement learning models for optimizing emergency resource dispatch strategies. Beginning with the theoretical foundations of reinforcement learning, it categorizes and analyzes existing models while exploring key influencing factors and challenges. Research indicates that reinforcement learning holds broad application prospects in the field of emergency resource dispatch, effectively enhancing the scientific rigor and precision of resource allocation to provide critical support for emergency rescue operations. However, the application of reinforcement learning in this domain remains in its developmental stage, necessitating further research and exploration in model refinement, algorithm optimization, and multidisciplinary integration. Moving forward, we should strengthen the integration of theoretical innovation and practical application, fully leverage advanced computational technologies and data analysis methods, and further enhance the efficiency and effectiveness of emergency resource dispatch. This will enable faster and more efficient responses to complex and dynamic emergencies, thereby playing a greater role in safeguarding social stability and protecting people's lives and property.

In addition to the aforementioned points, it is crucial to recognize the importance of realistic scenario modeling and data quality in the development of reinforcement learning-based dispatch systems. Effective simulation environments that accurately reflect the uncertainties and complexities of emergency situations are fundamental for training and evaluating algorithms. Moreover, the integration of domain knowledge with reinforcement learning frameworks can facilitate more robust decision-making under various constraints, such as limited resources, evolving situational information, and diverse stakeholder needs. Collaboration between emergency management experts and technical researchers is essential to ensure that algorithmic strategies align with real-world operational requirements. By fostering interdisciplinary cooperation and emphasizing continuous feedback from field applications, reinforcement learning approaches can be progressively tailored to address the nuanced demands of emergency resource management, ultimately contributing to more resilient and adaptive response systems.

References

[1] Chouba, I., Amodeo, L., Arbaoui, T., Yalaoui, F., Laplanche, D., & Sanchez, S. (2022). Optimization of French Hospital Department Activities Using an Efficient Discrete Event Simulation Model. Applied Sciences (Switzerland), 12(19), 21. DOI:10.3390/app12199498

[2] Sujatha, K., Bhavani, N. P. G., Hari, R., Kumar, K. S., Jayachitra, N., Krishnaveni, S., Thivya, K. S., Kalaivani, A., & Sankari, B. R. (2021). Power management technique for energy-efficient communication systems in telemedicine. Green Technological Innovation for Sustainable Smart Societies: Post Pandemic Era, 115–143. DOI:10.1007/978-3-030-73295-0_6

[3] Han, C. S., & Thorup, K. R. (2021). Integrated equilibrium planning for emergency logistics warehouse allocation based on internet plus mode. International Journal of Information Technology and Management, 20(3), 316–334. DOI:10.1504/IJITM.2021.116054

[4] Razavilar, N., & Taleshi, J. M. (2021). Cost-Effectiveness Analysis of Transcatheter Arterial Embolization Techniques for the Treatment of Gastrointestinal Bleeding in the United States. Value in Health, 24(4), 477–485. DOI:10.1016/j.jval.2020.10.027

[5] LI, H., LIU, C., ZENG, Q., HE, H., REN, C., WANG, L., & CHENG, F. (2021). Mining emergency event logs to support resource allocation. IEICE Transactions on Information and Systems, E104D(10), 1651–1660. DOI:10.1587/transinf.2021EDP7029

[6] Davodabadi, A., Daneshian, B., Saati, S., & Razavyan, S. (2021). Prioritization of patients in ICU: Composite approach of multiple-criteria decision-making and discrete event simulation. Brazilian Journal of Operations and Production Management, 18(1). DOI:10.14488/BJOPM.2021.008

[7] Pei, Z., Yuan, Y., Yu, T., & Li, N. (2022). Dynamic Allocation of Medical Resources During the Outbreak of Epidemics. IEEE Transactions on Automation Science and Engineering, 19(2), 663–676. DOI:10.1109/TASE.2021.3102491

[8] Yin, M., Li, W., Feng, L., Yu, P., & Qiu, X. (2020). Emergency communications based on throughput-aware D2D multicasting in 5G public safety networks. Sensors (Switzerland), 20(7), 1901. DOI:10.3390/s20071901

[9] Majumder, R., & Ghose, D. (2023). Three-player game-theoretic allocation of indivisible resources during natural disasters. Engineering Optimization, 55(9), 1577–1596. DOI:10.1080/0305215X.2022.2101047

[10] Bahari, A., & Asadi, F. (2020). A simulation optimization approach for resource allocation in an emergency department healthcare unit. Global Heart, 15(1). DOI:10.5334/GH.528

[11] Anantha Kumar, R., & Kartheeban, K. (2021). Resource allocation using Dynamic Pricing Auction Mechanism for supporting emergency demands in Cloud Computing. Journal of Parallel and Distributed Computing, 158, 213–226. DOI:10.1016/j.jpdc.2021.07.016

[12] Martin, D. A., Guillen, M., Farro, A., Condori, M., Dreyfuss, A., & Nagdev, A. (2022). Role of Tele-ultrasound for Teaching Ultrasound-guided Nerve Blocks in the Emergency Department: A Case Series from Peru. Clinical Practice and Cases in Emergency Medicine, 6(3), 204–207. DOI:10.5811/cpcem.2022.2.55417

[13] Ramirez-Nafarrate, A., Araz, O. M., & Fowler, J. W. (2021). Decision Assessment Algorithms for Location and Capacity Optimization under Resource Shortages. Decision Sciences, 52(1), 142–181. DOI:10.1111/deci.12418

[14] Becker, A. S., Erinjeri, J. P., Chaim, J., Kastango, N., Elnajjar, P., Hricak, H., & Vargas, H. A. (2022). Automatic Forecasting of Radiology Examination Volume Trends for Optimal Resource Planning and Allocation. Journal of Digital Imaging, 35(1), 1–8. DOI:10.1007/s10278-021-00532-4

[15] Nyimbili, P. H., & Erden, T. (2021). Comparative evaluation of GIS-based best–worst method (BWM) for emergency facility planning: perspectives from two decision-maker groups. Natural Hazards, 105(1), 1031–1067. DOI:10.1007/s11069-020-04348-3

[16] Hejja, K., & Hesselbach, X. (2019). Evaluating impacts of traffic migration and virtual network functions consolidation on power aware resource allocation algorithms. Future Generation Computer Systems, 101, 83–98. DOI:10.1016/j.future.2019.06.015

[17] Schütte, M., & Hartart, F. (2019). Questionnaire-based evaluation of mass casuality incident (MCI) exercises: A tool for learning-oriented feedback. Notfall Und Rettungsmedizin, 22(6), 522–527. DOI:10.1007/s10049-019-0584-2

[18] Wang, B., Sun, Y., Zhao, N., & Gui, G. (2020). Learn to Coloring: Fast Response to Perturbation in UAV-Assisted Disaster Relief Networks. IEEE Transactions on Vehicular Technology, 69(3), 3505–3509. DOI:10.1109/TVT.2020.2967124

[19] Kao, J. H., Hsiao, P. H., Oyang, Y. J., Sun, W. Z., Sun, C. H., Liaw, H. T., & Chang, S. W. (2020). The study for dispatch decision of medical emergency resources with real-time spatial analysis. Journal of Supercomputing, 76(3), 1604–1627. DOI:10.1007/s11227-018-2358-x

[20] Peixoto, J. P. J., Costa, D. G., Portugal, P., & Vasques, F. (2023). A geospatial dataset of urban infrastructure for emergency response in Portugal. Data in Brief, 50. DOI:10.1016/j.dib.2023.109593

[21] Yousefi, M., & Yousefi, M. (2020). Human resource allocation in an emergency department: A metamodel-based simulation optimization. Kybernetes, 49(3), 779–796. DOI:10.1108/K-12-2018-0675

[22] El-Bouri, R., Eyre, D. W., Watkinson, P., Zhu, T., & Clifton, D. A. (2021). Hospital Admission Location Prediction via Deep Interpretable Networks for the Year-Round Improvement of Emergency Patient Care. IEEE Journal of Biomedical and Health Informatics, 25(1), 289–300. DOI:10.1109/JBHI.2020.2990309

[23] Guan, X., & Xue, J. (2023). Energy-Efficient Computing Offloading Based on Multi-UAV Dispatch via NOMA in Emergency Communication Networks. Wireless Personal Communications, 133(1), 199–226. DOI:10.1007/s11277-023-10764-y

[24] Rosen, T., Ferrante, L. E., Liu, S. W., Benton, E. A., Mulcare, M. R., Stern, M. E., Biese, K., Hwang, U., & Sanon, M. (2020). Managing Older Adults with Presumed COVID-19 in the Emergency Department: A Rational Approach to Rationing. Journal of the American Geriatrics Society, 68(8), 1631–1635. DOI:10.1111/jgs.16651

[25] Yang, L. (2021). An intelligent method to select emergency allocation path of groundwater resource based on game theory. Arabian Journal of Geosciences, 14(12), 1–8. DOI:10.1007/s12517-021-07464-x

[26] Wang, Y., Su, Z., Xu, Q., Li, R., Luan, T. H., & Wang, P. (2023). A Secure and Intelligent Data Sharing Scheme for UAV-Assisted Disaster Rescue. IEEE/ACM Transactions on Networking, 31(6), 2422–2438. DOI:10.1109/TNET.2022.3226458

[27] Bu, B. (2022). Network coding data distribution technology between streams of emergency system based-wireless multi-hop network. Computer Communications, 186, 22–32. DOI:10.1016/j.comcom.2022.01.008

[28] Wu, L., & Wang, W. (2019). Resource allocation optimization of UAVs-enabled air-ground collaborative emergency network in disaster area. International Journal of Performability Engineering, 15(8), 2133–2144. DOI:10.23940/ijpe.19.08.p13.21332144

[29] Ghannad, P., Lee, Y. C., Friedland, C., & Yang, E. (2019). Optimizing the Socioeconomic Benefit of Post-Disaster Strategies by Prioritizing Reconstruction of Damaged Facilities. Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019, 180–187. DOI:10.1061/9780784482445.023

[30] Khelladi, F., Boudali, M., Orjuela, R., Cassaro, M., Basset, M., & Roos, C. (2022). An Emergency Hierarchical Guidance Control Strategy for Autonomous Vehicles. IEEE Transactions on Intelligent Transportation Systems, 23(5), 4319–4330. DOI:10.1109/TITS.2020.3043485

[31] Guo, Z., Hu, B., & Chen, S. (2023). A Decision for Throughput Optimization in UAV-Enabled Emergency Outdoor–Indoor Fairness Communication. Drones, 7(7). DOI:10.3390/drones7070460

[32] Yao, L., Xu, Z., & Chen, X. (2019). Sustainable water allocation strategies under various climate scenarios: A case study in China. Journal of Hydrology, 574(000), 529–543. DOI:10.1016/j.jhydrol.2019.04.055

[33] Cao, S., Li, D., & Cao, S. (2020). Optimization research of emergency resource on rail transit allocation location considering multiple transport modes. Journal of Physics: Conference Series, 1486(7), 072015 (8pp). DOI:10.1088/1742-6596/1486/7/072015

[34] Padda, S. S., Glass, J. R., & Stahlschmidt, Z. R. (2021). When it's hot and dry: life-history strategy influences the effects of heat waves and water limitation. Journal of Experimental Biology, 224(7), jeb.236398. DOI:10.1242/jeb.236398

[35] Gao, Y., Liu, C., Zeng, Q., & Duan, H. (2021). Two Effective Strategies to Support Cross-Organization Emergency Resource Allocation Optimization. Mobile Information Systems, 2021(2), 1–15. DOI:10.1155/2021/7965935

[36] Kain, T., & Fowler, R. (2019). Preparing intensive care for the next pandemic influenza. Critical Care, 23(1). DOI:10.1186/s13054-019-2616-1

[37] Ranabhat, C. L., Acharya, S. P., Adhikari, C., & Kim, C. B. (2023). Universal health coverage evolution, ongoing trend, and future challenge: A conceptual and historical policy review. Frontiers in Public Health, 11. DOI:10.3389/fpubh.2023.1041459

[38] Jia, L., Wang, L., & Qin, Y. (2022). Resource Allocation on High-Speed Railway Emergency Management. High-Speed Railway Operation Under Emergent Conditions, 265–286. DOI:10.1007/978-3-662-63033-4_7

[39] Yang, Z., Marti, A., Chen, Y., & Marti, J. R. (2023). Optimal Resource Allocation to Enhance Power Grid Resilience Against Hurricanes. IEEE Transactions on Power Systems, 38(3), 2621–2629. DOI:10.1109/TPWRS.2022.3193133

[40] Tran, D. H., Nguyen, V. D., Chatzinotas, S., Vu, T. X., & Ottersten, B. (2022). UAV Relay-Assisted Emergency Communications in IoT Networks: Resource Allocation and Trajectory Optimization. IEEE Transactions on Wireless Communications, 21(3), 1621–1637. DOI:10.1109/TWC.2021.3105821

[41] Rezapour, S., Baghaian, A., Naderi, N., & Farzaneh, M. A. (2022). Dynamic on-site treatment strategy for large-scale mass casualty incidents with rescue operation. Computers and Industrial Engineering, 163(163-), 163. DOI:10.1016/j.cie.2021.107796

[42] Austin, E. E., Blakely, B., Tufanaru, C., Selwood, A., Braithwaite, J., & Clay-Williams, R. (2020). Strategies to measure and improve emergency department performance: A scoping review. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 28(1). DOI:10.1186/s13049-020-00749-2

[43] Xu, L., Wang, Z., Liu, Y., & Xing, L. (2021). Energy allocation strategy based on fuzzy control considering optimal decision boundaries of standalone hybrid energy systems. Journal of Cleaner Production, 279(6), 123810. DOI:10.1016/j.jclepro.2020.123810

[44] Sun, Y., Ling, J., Chen, X., Kong, F., Hu, Q., & Biancardo, S. A. (2022). Exploring Maritime Search and Rescue Resource Allocation via an Enhanced Particle Swarm Optimization Method. Journal of Marine Science and Engineering, 10(7). DOI:10.3390/jmse10070906

[45] Yao, Z., Cheng, W., Zhang, W., & Zhang, H. (2021). Resource Allocation for 5G-UAV-Based Emergency Wireless Communications. IEEE Journal on Selected Areas in Communications, 39(11), 3395–3410. DOI:10.1109/JSAC.2021.3088684

[46] Campbell, K. H., Pettker, C. M., & Goffman, D. (2020). Consolidation of obstetric services in a public health emergency. Seminars in Perinatology, 44(7). DOI:10.1016/j.semperi.2020.151281

[47] Melman, G. J., Parlikad, A. K., & Cameron, E. A. B. (2021). Balancing scarce hospital resources during the COVID-19 pandemic using discrete-event simulation. Health Care Management Science, 24(2), 356–374. DOI:10.1007/s10729-021-09548-2

[48] Santhiapillai, F. P., & Ratnayake, R. M. C. (2022). Risk-based prioritization method for planning and allocation of resources in public sector. TQM Journal, 34(4), 829–844. DOI:10.1108/TQM-02-2021-0039

[49] Arumuganathan, P., Adcock, A. K., Espinosa, C., & Findley, S. (2023). Analysis of Telestroke Usage in Rural Critical Access Emergency Departments. Telemedicine and E-Health, 29(12), 1828–1833. DOI:10.1089/tmj.2022.0408

[50] Mishra, S., Biswas, S., & Bhatnagar, S. (2020). Palliative care delivery in cancer patients in the era of Covid-19 outbreak: Unique needs, barriers, and tools for solutions. Indian Journal of Palliative Care, 26(5), S130–S141. DOI:10.4103/IJPC.IJPC_194_20

[51] Xu, Y. H., Yu, G., & Yong, Y. T. (2021). Deep Reinforcement Learning-Based Resource Scheduling Strategy for Reliability-Oriented Wireless Body Area Networks. IEEE Sensors Letters, 5(1), 1. DOI:10.1109/LSENS.2020.3044337

[52] Liu, S., Tian, J., Zhai, C., & Li, T. (2023). Joint computation offloading and resource allocation in vehicular edge computing networks. Digital Communications and Networks, 9(6), 1399–1410. DOI:10.1016/j.dcan.2022.12.002

[53] Zhang, H., Xi, S., Jiang, H., Shen, Q., Shang, B., & Wang, J. (2023). Resource Allocation and Offloading Strategy for UAV-Assisted LEO Satellite Edge Computing. Drones, 7(6). DOI:10.3390/drones7060383

[54] Tsai, J. C. H., Weng, S. J., Liu, S. C., Tsai, Y. Te, Gotcher, D. F., Chen, C. H., Chou, C. A., & Kim, S. H. (2020). Adjusting daily inpatient bed allocation to smooth emergency department occupancy variation. Healthcare (Switzerland), 8(2). DOI:10.3390/healthcare8020078

[55] Wu, J., Zhang, J., Ma, J., Zhao, M., Wang, X., Gao, X., & Zhang, H. (2021). Energy Efficient 5G LoRa Ad-Hoc Network for Smart Grid Communication. ICEIEC 2021 - Proceedings of 2021 IEEE 11th International Conference on Electronics Information and Emergency Communication, 17–20. DOI:10.1109/ICEIEC51955.2021.9463809

[56] Smith, T. G., Ahmad, I., Takhar, A., Surda, P., & El-Boghdadly, K. (2020). Unconventional multidisciplinary team strategy for tracheostomy in COVID-19. Anaesthesia Reports, 8(2), 178–182. DOI:10.1002/anr3.12074

[57] Fang, J., Hu, J., Wei, J., Liu, T., & Wang, B. (2020). An efficient resource allocation strategy for edge-computing based environmental monitoring system. Sensors (Switzerland), 20(21), 1–16. DOI:10.3390/s20216125

[58] Yu, X., Hu, Z., Wang, X., & Luo, W. (2023). Ranking teaching–learning-based optimization algorithm to estimate the parameters of solar models. Engineering Applications of Artificial Intelligence, 123(Pt.A), 106217–106225. DOI:10.1016/j.engappai.2023.106225

[59] Wang, P., Wang, F., & Zhu, Z. (2022). The Evaluation Method of Community Emergency Resource Allocation Based on Coordination Sensor Information Collection. Journal of Sensors, 2022. DOI:10.1155/2022/8715028

[60] Zhou, H., Zhang, H., Chen, B., Huo, J., & Lin, H. (2023). Optimum Vehicle Evacuation Plan for Nuclear Emergency Using Fuzzy Credibility Theory and Improved Genetic Algorithm. Arabian Journal for Science and Engineering, 48(8), 10517–10536. DOI:10.1007/s13369-023-07663-6

[61] Wang, W., Ye, C., & Tian, J. (2023). SGGTSO: A Spherical Vector-Based Optimization Algorithm for 3D UAV Path Planning. Drones, 7(7), 27. DOI:10.3390/drones7070452

[62] Lu, J., Huang, M., Wu, W., Wei, Y., & Liu, C. (2023). Application and Improvement of the Particle Swarm Optimization Algorithm in Source-Term Estimations for Hazardous Release. Atmosphere, 14(7), 18. DOI:10.3390/atmos14071168

[63] Niu, C., & Yun, L. (2023). Optimization study of emergency logistics distribution vehicle scheduling based on saving algorithm. Proceedings of SPIE, 10. DOI:10.1117/12.2678841

[64] Davarian, F., & Behnamian, J. (2022). Robust finite-horizon scheduling/rescheduling of operating rooms with elective and emergency surgeries under resource constraints. Journal of Scheduling, 25(6), 625–641. DOI:10.1007/s10951-022-00741-x

[65] Chen, F., Chen, J., & Liu, J. (2022). Forecast of flood disaster emergency material demand based on IACO-BP algorithm. Neural Computing and Applications, 34(5), 3537–3549. DOI:10.1007/s00521-021-05883-1

[66] Zhang, J., Ren, G., & Song, J. (2023). Resilience-based optimization model for emergency bus bridging and dispatching in response to metro operational disruptions. PLoS ONE, 18(3 March). DOI:10.1371/journal.pone.0277577

[67] Yu, Y., Liu, H., Tian, Y., Li, X., & Shang, Y. (2023). Hierarchical Optimization of Scheduling Scheme for Mixed Helicopter Fleet in Disaster Relief. Arabian Journal for Science and Engineering, 48(5), 7195–7206. DOI:10.1007/s13369-023-07649-4

[68] Park, Y., Lee, S., Sung, I., Nielsen, P., & Moon, I. (2023). Facility Location-Allocation Problem for Emergency Medical Service with Unmanned Aerial Vehicle. IEEE Transactions on Intelligent Transportation Systems, 24(2), 1465–1479. DOI:10.1109/TITS.2022.3223509

[69] Lyng, J. W., Baldino, K. T., Braude, D., Fritz, C., March, J. A., Peterson, T. D., & Yee, A. (2022). Prehospital Supraglottic Airways: An NAEMSP Position Statement and Resource Document. Prehospital Emergency Care, 26(S1), 32–41. DOI:10.1080/10903127.2021.1983680

[70] Raj, R., & Jain, V. (2023). Resource and traffic control optimization in MMAP[c]/PH[c]/S queueing system with PH retrial times and catastrophe phenomenon. Telecommunication Systems, 84(3), 341–362. DOI:10.1007/s11235-023-01053-x

[71] Zai, W., Wang, J., & Li, G. (2023). A Drone Scheduling Method for Emergency Power Material Transportation Based on Deep Reinforcement Learning Optimized PSO Algorithm. Sustainability (Switzerland), 15(17). DOI:10.3390/su151713127

[72] Wang, X., Tang, Y., Qi, Q., Wang, G., & Bi, B. (2022). Optimization of Emergency Transportation Organization of Holiday Tourism Traffic. Wireless Communications and Mobile Computing, 2022. DOI:10.1155/2022/2264876

[73] Ding, Z., Xu, X., Jiang, S., Yan, J., & Han, Y. (2022). Emergency logistics scheduling with multiple supply-demand points based on grey interval. Journal of Safety Science and Resilience, 3(2), 179–188. DOI:10.1016/j.jnlssr.2022.01.001

[74] Wang, X., Chen, G., & Xu, S. (2022). Bi-objective green supply chain network design under disruption risk through an extended NSGA-II algorithm. Cleaner Logistics and Supply Chain, 3. DOI:10.1016/j.clscn.2021.100025

[75] He, L., & Xie, Z. (2022). Optimization of Urban Shelter Locations Using Bi-Level Multi-Objective Location-Allocation Model. International Journal of Environmental Research and Public Health, 19(7). DOI:10.3390/ijerph19074401

[76] Kim, D., Kan, Y., Aum, Y. J., Lee, W., & Yi, G. (2023). Hotspots-based patrol route optimization algorithm for smart policing. Heliyon, 9(10), 16. DOI:10.1016/j.heliyon.2023.e20931

[77] Zhang, T., Jiang, P., Sheng, K., & Wang, X. (2023). Distribution vehicle +drone joint distribution path optimization based on K-means optimization genetic algorithm. Proceedings of SPIE, 12591(000), 74. DOI:10.1117/12.2668580

[78] Wang, K., Yuan, W., & Yao, Y. (2023). Path optimization for mass emergency evacuation based on an integrated model. Journal of Building Engineering, 68. DOI:10.1016/j.jobe.2023.106112

[79] Chen, S., Yang, J., & Lang, P. (2023). Research On Emergency Parcel Transfer and Structure Optimization of E-Commerce Logistics Network Based on Multi-Objective Planning. BCP Business & Management, 48, 1–11. DOI:10.54691/bcpbm.v48i.5243

[80] Dai, Z. (2023). Site optimization study of emergency material centre based on improved particle swarm algorithm. Proceedings of SPIE, 12799(000), 108. DOI:10.1117/12.3005969

[81] Chang, Z., Punnen, A. P., Zhou, Z., & Cheng, S. (2023). Solving dynamic satellite image data downlink scheduling problem via an adaptive bi-objective optimization algorithm. Computers and Operations Research, 160(Dec.), 1.1-1.16. DOI:10.1016/j.cor.2023.106388

[82] Alkanhel, R., Rafiq, A., Mokrov, E., Khakimov, A., Muthanna, M. S. A., & Muthanna, A. (2023). Enhanced Slime Mould Optimization with Deep-Learning-Based Resource Allocation in UAV-Enabled Wireless Networks. Sensors, 23(16). DOI:10.3390/s23167083

[83] Zhang, B., Wang, L., & Zhou, C. (2023). Emergency repair scheduling algorithm for multi-point fault of power communication considering multi team cooperation. Proceedings of SPIE, 12800(000), 192. DOI:10.1117/12.3004110

[84] Wan, M., Ye, C., & Peng, D. (2023). Multi-period dynamic multi-objective emergency material distribution model under uncertain demand. Engineering Applications of Artificial Intelligence, 117(Pt.A). DOI:10.1016/j.engappai.2022.105530