
A Multi-objective Optimization Planning Framework for Active Distribution System Via Reinforcement Learning

Hongtao Li¹, Cunping Wang¹, Hao Tian^{2,3,*}, Zhigang Ren¹,
Ergang Zhao² and Lina Xu²

¹State Grid Beijing Electric Power Research Institute, Beijing, 100075, China

²Wuxi Research Institute of Applied Technologies Tsinghua University, Wuxi, 214072, China

³La Consolacion University Philippines, 3000 Bulacan, Catmon Rd, Malolos
E-mail: tianhao@cw Xu.edu.cn

*Corresponding Author

Received 20 March 2023; Accepted 11 May 2023;
Publication 26 August 2023

Abstract

The effective planning of active distribution networks is crucial for utility companies to make informed decisions regarding investments in distributed generation, reliability assessment, reactive power planning, substation revisions, and feeder repositioning. However, the dynamic nature of the solution space makes it challenging for model-based optimization methods to ensure computational performance in active distribution network planning. To address this issue, this study proposes a planning method that focuses on improving computational performance through the continuous updating of the planning model's solution space during the reinforcement learning training process. Based on simulations conducted on the IEEE 33-bus test

Distributed Generation & Alternative Energy Journal, Vol. 38.6, 1741–1762.

doi: 10.13052/dgaej2156-3306.3862

© 2023 River Publishers

system, the proposed planning strategy successfully enhances computational performance while minimizing investment costs compared to other strategies. With the proposed method, the investment cost and the operation cost are reduced by 32.42% and 23.91%, respectively.

Keywords: Planning, active distribution network planning, reinforcement learning, renewable energy source.

1 Introduction

The distribution network is a vital component of the power system responsible for distributing the electricity generated by power plants to end-users. With the growing use of distributed generation, which offers the benefits of low-carbon environmental protection and flexible installation in the distribution network, the term “active distribution networks” refers to networks with control systems for dispersed energy resources [1]. The introduction of large-scale distributed generation transforms a traditional passive distribution network with a single flow into an active distribution network with multiple power sources, increasing the complexity and unpredictability of system operation and planning [2]. Active management and planning are required to address issues arising from distributed generation access, optimize distributed generation placement, and maximize the benefits of distributed generation in distribution networks. An active distribution network utilizes advanced communication and information technologies to actively manage the network, thereby scaling up access to distributed power sources, energy storage, and demand-side management, among others. The present management strategy encourages the adoption of distributed power sources and increases environmental stress by enabling them to actively participate in system regulation based on operational requirements [3]. The standard approach to planning distributed generation assumes that “installation is forgotten.” Therefore, it is crucial to investigate the optimal planning problems for distributed power sources in active distribution networks based on active management models.

In order to achieve a clean energy supply and net zero carbon emissions, it is crucial to incorporate renewable energy sources through active distribution network planning. In [4], SVC devices were introduced to increase the hosting capacity of photovoltaic generation. By upgrading conventional feeders and substations, and strategically investing in distributed generations [6, 7], energy storage systems [8], and reactive power sources [4, 5], a utility company can achieve certain performance targets, such as reducing carbon

emissions and minimizing unserved load. Current methodologies predefine a mix of active distribution network management technologies to find the best solutions within a fixed space. System reliability was improved in [9] by properly configuring energy storage systems [10], distribution automated systems, and distribution switches. In [11], the ideal mix, location, and size of wind turbines, photovoltaic panels, and energy storage systems were determined to increase revenue in the active distribution network. References [12, 13] focused on enhancing distribution network resilience, while [14] optimized active distribution network facilities to meet the growing demand for electric vehicle charging. Reference [15] employed renewable energy sources, energy storage systems, and demand response to reduce carbon emissions. If the solution areas are dynamically updated, these performance metrics can be further improved.

The presence of distributed generation has an impact on various aspects of traditional distribution networks, including line energy flow, node voltages, and power network topology. While research has mainly focused on placement and capacity of distributed generation, grid planning, and multi-stage planning, some studies have examined conventional distribution networks. Researchers have developed models using particle swarm optimization algorithms to address load uncertainty and identify the ideal solution. Other studies have created computational formulas and optimization models to determine the appropriate size and location of distributed generation units, considering factors such as node voltages and line current. Additionally, some researchers have categorized distributed generation into power controllable and power stochastic types and developed operation optimization models. Reference [21] provides a simple planning method that evaluates the status of multistage or multi-scenario modelling for distribution systems. However, many studies still rely on passive management techniques, which do not accurately represent the ability to adapt to distributed energy. To address this issue, recent research has focused on hierarchical modelling techniques, distributed generation capabilities, and optimization techniques for active distribution system planning.

Various planning methods are being explored for active distribution network planning, including distributed generation planning, strategy implementation models, demand side response, and optimization methods. Researchers have investigated heuristic algorithms such as particle swarm optimization [29], robust optimization [30, 31], and genetic algorithm [32, 33] for optimizing active distribution network architecture [28]. However, very few studies have considered both distributed generation and demand side response

in these planning methodologies. This paper proposes a solution to the distribution network planning problem by utilizing both the non-dominated sorting genetic algorithm and tabu search method. These algorithms are frequently used to address the workshop scheduling problem, and their combination provides complementary benefits in calculation and search, resulting in faster algorithm convergence.

Compared to standard planning techniques, hierarchical multi-objective planning that considers multiple planning objectives can enhance the comprehensiveness and effectiveness of distribution network planning. Reference [22] proposes a hierarchical planning paradigm for implementing active distribution network planning control. The lower level planning involves building a distributed power optimization model for an operational distribution network, while the higher level planning focuses on generating distributed energy-related challenges from the system network loss sensitivity perspective. In addition, Reference [22] examines crucial assessment metrics for active distribution networks, management capability planning, and constructs a hierarchical model approach. The study also builds a distribution system planning model that takes into account the association between balancing network loss, line investments, power price, carbon emission expenditure, and policy subsidies using an active management mode. To develop hierarchical multi-objective procedural programming, this study suggests taking into account the distributed power grid connection perspective.

According to literature, active distribution network planning places a strong emphasis on distributed generation planning, and several studies have explored this topic [23]. The study found that demand-side management and network load can affect the method for mapping distributed generation and the detailed yearly cost of the planning scheme when energy production is scheduled. Reference [24] examines methods for scheduling distributed stored energy and distributed generation. Reference [25] outlines a method for creating a distributed energy forecast for an active distribution system. Furthermore, the use of renewable energy can increase the share of clean energy in an area, improve the environment, and promote efficient and safe operation of the power system [26, 27].

Reinforcement learning has gained popularity in the field of active distribution networks due to its ability to adapt to the environment and make optimal decisions without requiring a model. Researchers have developed a collaborative multiple voltage control using smart inverters [35], and a batch-constrained reinforcement learning approach for dynamic distribution network reconfiguration using limited historical operational data [37]. Literature

extensively documents the use of reinforcement learning for managing active distribution systems, including a multi-agent deep reinforcement approach for voltage control using on-load tapping changers and solar inverters [34], addressing overvoltage problems in distribution systems with significant solar penetration [38], and creating a safe volt/var control using a restricted soft actor-critic method [36]. Furthermore, research has used reinforcement learning to plan the stochastic operation of active distribution systems [39]. There is still great potential for using reinforcement learning to dynamically explore solution spaces for active distribution system planning issues.

A new method for active distribution network planning is proposed in this article. The method utilizes reinforcement learning to dynamically update solution spaces, resulting in an active distribution network planning scheme that guarantees desired performance. The following is a summary of this paper's significant contributions:

- Instead of addressing the model-based optimization problem with fixed solution spaces, this study models the dynamic update of the active distribution network solution spaces as a Markov decision-making process to achieve the essential performances.
- Reinforcement learning is used to train the deep neural networks in the active distribution network to make sure it performs as predicted.
- The proposed approach is tested on the industry-recognized IEEE 33-bus system, and the numerical outcomes demonstrate that it is possible to guarantee the active distribution network's predicted performance at a lower investment cost than in other circumstances.

2 Problem Description

Passive distribution networks are currently in use, but they differ from active distribution networks in their inability to actively monitor and regulate dispersed generation and load. While traditional distribution systems utilize a single network topology to manage load variation, active distribution networks employ stochastic and dynamic planning techniques and consider source-grid-load coordinated scheduling. Additionally, active distribution networks can actively regulate multiple controllable resources to prevent network failures, while traditional distribution networks can only control their systems after a failure has occurred. Despite its use of passive control technology to manage one-way power flow, traditional distribution networks can still implement distributed hierarchical management and incorporate demand-side management components.

The planning of the active distribution network faces three main challenges. Firstly, the selection of optimal locations for energy storage systems, units, and SVCs introduces binary variables and creates linkages between current and new nodes. Secondly, traditional distribution network planning encounters increasing complexity due to the unpredictability of distributed generation and management challenges. Finally, the planning is a multi-objective uncertain nonlinear mixed integer problem that requires balancing conflicting objectives and restrictions within a targeted time frame. As distributed generation increases, load growth becomes more difficult to predict, leading to power loss, under or overvoltage, and modifications to fault events. Moreover, conflicts may arise between independent investors installing distributed generation and distribution network businesses maintaining grid security and power quality. Effective voltage management, reactive power balance, and relay protection are crucial for the system's dependable and secure operation. To transition from passive to active management, the distribution network must update its current automation system.

In recent years, active distribution network planning has become crucial to the growth of sustainable energy sources and the power industry. Its primary objective is to determine the optimal location, size, and type of producing facilities to meet future load demand while maintaining the security and dependability of the power system and establishing the ideal generation mix. Active distribution system planning models may be short-term, mid-term, or long-term, with a planning horizon ranging from a few years to several decades. The active distribution system planning model comprises two distinct modules: the operational evaluation module and the investment decision module. The former selects the production mix that can meet the growth in yearly peak load and overall energy consumption based on predicted utilization hours, while the latter evaluates the feasibility of the proposed generation mix using probabilistic or deterministic production simulation methods. However, such compartmentalized models only offer possible development plans rather than optimal ones. Recent advances in computing power have enabled the integration of the operational evaluation and investment decision modules into a detailed active distribution network planning model. This integrated approach allows for more accurate simulation of power systems operating at various load levels, resulting in a more reasonable and reliable generation expansion strategy. Techniques have been proposed to incorporate precise operational limitations into active distribution network planning, including the use of load blocks generated from the daily load curve. While load blocks require less processing time than accounting

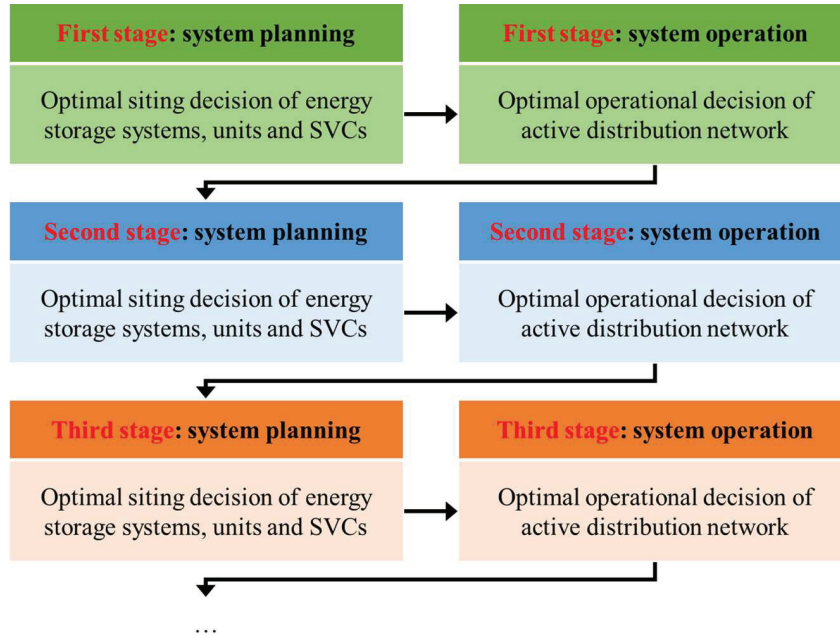


Figure 1 Optimization planning framework for the active distribution system.

for all operating limitations across several days, they may not effectively depict the ongoing fluctuation and volatility of renewable energy sources.

Demand Side Management (DSM) refers to strategies implemented on the consumer side of electricity usage to optimize electricity consumption, ensure rational usage, and support effective distribution of available resources [34]. With increasing power demand, passive distribution networks have evolved into active distribution networks that incorporate demand response technology. Active distribution network management encompasses various factors, including demand-side response to peak and valley load characteristics that drive up operational costs. By reducing peak load demand through improved consumption patterns, user guidance, and peak load shifting, demand-side management can balance the load and lessen the pressure on distribution system cables and equipment. Pricing and incentive-based measures such as interruptible capacity planning and time-of-use power pricing can also be employed to support demand response. Prior to implementing demand response, power workers must ensure the effective functioning of the power system by taking into account user characteristics to reduce distribution system loss, power purchase costs, and equipment wear and tear.

Effective demand side response can enhance equipment voltage compliance, extend equipment useful life, and improve component operating conditions, which leads to improved system safety and higher control power over the load. By matching the user's electrical demand, demand side response can reduce the risk of overload operations and greatly enhance power supply security.

3 Model Formulation

Reinforcement learning's sole objective is to learn from interaction and accomplish a goal. The so-called Markov decision-making process is a particular type of stochastic process that frequently describe this process. The agent and environmental concepts are essential. Components of the Markov decision-making process [40]. The agent is in charge of selecting a solution after solving the puzzle. The agent primarily engages with the environment to gather data. The ongoing relationship between the agent and the environment is described by the technique below: The agent receives feedback in the form of incentives as a result of the actions they take and their environment. The agent aims to maximize the sum of its acquired rewards over a predetermined (or indefinite) time scale. The relationship between the agent and the environment happens at certain discrete intervals of time $t = 0, 1, 2, 3 \dots$. After getting an accurate description of the state of the environment, the agent selects a course of action for each time step t . The environment reacts to the agent by sending a numerical signal relating to the action that was chosen by the agent after it has completed it. The signals are considered as a reward in this instance. The agent's objective is to maximize the rewards' discounted total as determined by $\sum_p \gamma Reward_{t+p}$. The task of translating states to actions at each time step is then left to the agent. The sign π_t stands for the agent's policy, which refers to this mapping. Once again, this procedure is carried out until convergence is reached before the system enters a new state [40]. Finally, the reward of acting in the initial state and following the policy is clearly expressed as $q(s, a) = E[\sum_p \gamma Reward_{t+p} | S_t = s, A_t = a]$. Using this function, the optimum policies for the current situation can be determined. In summary, the first subsection's long-term dynamic storage investment problem is reformulated using the Markov decision-making process framework, which serves as the foundation of the solution approach. The reward and transit functions are described in depth, as well as the state and action sets. This paper tackles the overestimation bias issue with Q-learning by using a novel approach termed

metamodeling in combination with artificial datasets. As a result, the second paragraph goes into the specifics and presumptions of the above technique. The final step-by-step methodology is described before a test case and its results.

In order to construct optimal policies using appropriate algorithms, it is essential to provide definitions for all the actions and state sets, transition and reward functions, and discount factor within every Markov decision-making process.

State:

$$S = \{s_{i,n}, \forall i \in I, \forall n \in Node\} \tag{1}$$

$$s_{i,n} \in \{0, 1\}$$

$s_{i,b}$ is a binary value of the state space, indicating whether to invest equipment i in node n or not.

Action: The set of potential actions for the problem is defined by the agent’s available actions. In the second scenario, the agent is expected to respond at one of the various defined levels based on a discrete-time Markov structure. Therefore, the following vector can be used to define the agent’s action:

$$A = \{a_i, \forall i \in I\} \tag{2}$$

$$a_i \in \{1, 2, 3, \dots, N\}$$

a_i is an integer value of the action space, an indication of which node the equipment i is invested at.

Environment: This study evaluates the effectiveness of different planning strategies, based on RES serving capability and the number of unserved loads in the active distribution system. To evaluate this performance at each stage, convex optimization models are used for each state, as shown below:

$$network\ constraints : \left\{ \begin{array}{l} \sum_j (p_{ji} - r_{ji}|I_{ji}|^2) + \tilde{p}_i^g - S_i \cos \varphi_i = \sum_k p_{ik}, \forall i \\ \sum_j (q_{ji} - x_{ji}|I_{ji}|^2) + \tilde{q}_i^g - S_i \sin \varphi_i = \sum_k q_{ik}, \forall i \\ v_i - v_j - 2(p_{ij}r_{ij} + q_{ij}x_{ij}) \\ \quad + |I_{ij}|^2(r_{ij}^2 + x_{ij}^2) = 0, \forall ij \\ v_i |I_{ij}|^2 = p_{ij}^2 + q_{ij}^2, \forall ij \end{array} \right. \tag{3}$$

$$\text{constraints of cutting power : } \begin{cases} \Delta p_{Cwind} \leq P_{Cwind.max} \\ \Delta p_{Csolar} \leq P_{Csolar.max} \\ \Delta p_{Claod} \leq P_{Cload.max} \end{cases} \quad (4)$$

$$\text{constraints of equipments : } \begin{cases} \begin{cases} E_{es,t+1} = E_{es,t} - \eta p_{es.c} + \eta p_{es.d}, \\ 0 \leq p_{es.d} \leq a_{es} u_{es.d} p_{es.max}, \\ 0 \leq p_{es.c} \leq a_{es} u_{es.c} p_{es.max}, \\ a_{es} E_{es.min} \leq E_{es,t} \leq a_{es} E_{es.max} \end{cases} \\ a_{unit} P_{unit.min} \leq p_{unit} \leq a_{unit} P_{unit.max} \\ a_{SVC} P_{SVC.min} \leq p_{SVC} \leq a_{SVC} P_{SVC.max} \end{cases} \quad (5)$$

Constraint (3) consists of convexified power flow restrictions, which include constraints on voltage, network power flow, branch current, and relaxed second-order cone constraints. Constraint (4) imposes limitations on renewable energy sources and loads that need to be respected. Constraint (5) outlines the operating limitations for the energy storage system, unit, and SVC. \tilde{p}_i^g is the energy output of generations in the distribution network, considering the uncertain RES generations.

Reward: The reward function is the final Markov decision-making process component to be defined. This is a key component of the definition since it influences the signals that the agent receives from the environment or rewards. The main forces behind the agent's derivation of the best possible policies are these signals.

$$\begin{aligned} \text{Reward} = & - \left[\sum_i \widetilde{Cost}_{inv,i} a_i + Cost_{loss} \sum_{line,t} I_{line,t} r_{line,t} \right. \\ & + \sum_t (\widetilde{Cost}_{wind} \Delta p_{Cwind} + \widetilde{Cost}_{solar} \Delta p_{Csolar} \\ & \left. + \widetilde{Cost}_{load} \Delta p_{Cload}) \right] \quad (6) \end{aligned}$$

The reward function of the initial investment cost and yearly operating costs is described by function (6), where The first term is the equipment's original investment cost, the second is network loss, the third is the cost of decreasing wind and photovoltaic power, and the fourth is load curtailment cost. $\widetilde{Cost}_{inv,i}$, \widetilde{Cost}_{wind} , \widetilde{Cost}_{solar} , \widetilde{Cost}_{load} are the cost of investment,

wind curtailment, solar curtailment and load curtailment considering the uncertainty from the government tax policies, environment, etc.

4 Solving Method

What results in the Q-learning method's overestimation bias under highly stochastic circumstances has been demonstrated. Since the action space and state space are small, the reinforcement learning method based on Q learning is the most efficient. In practical terms, this implies that the agent might try to transition to a particular state, even if transitioning to other states would be the best course of action, if it believes there is a chance of receiving an exceedingly "excellent" reward. This section of the essay discusses how this phenomenon pertains to the scenario under consideration and proposes a potential solution to mitigate its impact. It is crucial to note that the operational-level assumptions enumerated will be utilized solely in this study's investigation. The goal is to devise a computationally efficient and unbiased method for computing the loss of the load cost component in the reward function.

Assuming that the issue arises when the agent receives inaccurate signals about the optimal approach, these signals serve as incentives for the agent during each control period of the problem. The reward function contains two negative components, namely capital costs and outage penalties. Increasing storage capacity has a direct effect on investment costs, while outage penalties are largely dependent on the stochastic occurrence of outages. If there are few or infrequent outages during a decision period, the agent may choose to continue operating the system as it is and accept the outages instead of taking measures to prevent them, such as making investments. This outcome generates the previously mentioned misleading signals. In the best-case scenario, this occurrence would result in a slower convergence rate of the solution algorithm. However, in the worst-case scenario, it could result in the development of suboptimal strategies. To reduce the impact of this phenomenon on the outage part of the reward function, a unique approach must be developed using synthetic data and function approximation.

In contrast to real data collection and investigations, simulated datasets are created dynamically using methods. In this case, a simulated dataset with multiple input variables and a single output feature, the interruption cost, can be generated using current simulation techniques. A function approximation technique can then be used to map the desired and approximate output from the given inputs. Since this cost component is dependent on the storage

Table 1 The procedure of reinforcement learning for active distribution network planning

The procedure of reinforcement learning	
1	initialization: Q table
2	for every episode do :
3	Observe current state s
4	Select action a based on the Q table according to ε -greedy policy
5	Obtain a new state and reward from the environment
6	Update the Q table based on the Bellman function
7	Update the current state
8	End for

unit in the system, a vector can be created using the necessary attributes to estimate the cost of an outage.

In the initial stage of this process, it became evident that a systematic approach was necessary for generating data to be included in the data collection. Each of these findings is based on n distinct and independent simulations of the system, with the results subsequently being averaged. To ensure independent observations, each input feature of a data set (including all time characteristics and electricity production for all storage units) is randomly chosen from the corresponding ranges. We generate n network model realizations after selecting the input data for a specific observation and inducing outages for each test. The corresponding outputs (outage cost) are then computed by evaluating the outcomes of these simulations.

Prior to presenting the numerical case studies and results, a comprehensive and schematic outline of the proposed methodology is provided. While the categorizing Q-learning approach remains the foundation of the algorithm, a pre-processing stage involving the generation of synthetic datasets and function approximation has been incorporated. Table 1 illustrates the process, which comprises the standard phases of the Q-learning algorithm modified to suit the specific situation.

5 Results and Discussion

This section presents case studies that were conducted to validate the proposed approach in various technology combinations. The proposed algorithms were implemented using TensorFlow Python backend and executed on an Intel Core 1.6 GHz system with 8 GB of RAM. The optimization model was solved using GUROBI. To avoid computational complexity, a k-means clustering method was employed to group daily operating vectors and select

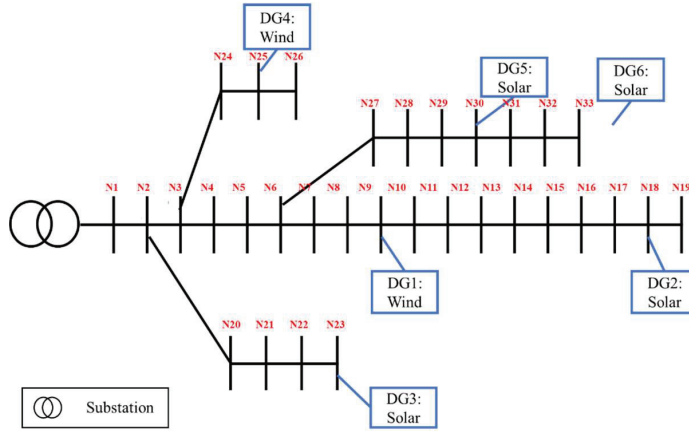


Figure 2 Illustration of the original IEEE 33 system before planning.

Table 2 Reinforcement learning algorithm parameters

	Symbols	Value
Learning rate	–	10^{-3}
Batch size	N_B	64
Discount factor	γ	0.9
Episode size	T	96

a subset of representative days that adequately capture various operating conditions, such as load, wind, and solar, across seasons. The second module was designed to select multiple sample days and streamline the model before conducting the proposed tasks. Load, wind, and solar yearly profiles were normalized and separated into daily profiles, which were then combined to generate a high-dimensional vector. This process was repeated to establish all representative days of each year across the planned horizon. The proposed technique was tested on the IEEE 33-bus distribution network, as shown in Figure 2. Table 2 lists all the parameters for the reinforcement learning algorithm, while Table 3 displays the investment at each node. Finally, the desired performances were achieved.

To highlight the advantages of the proposed approach, Table 4 compares the proposed example with other scenarios. Case A has the same investment as the baseline example at all stages. In Case B, a mathematical optimization method is utilized.

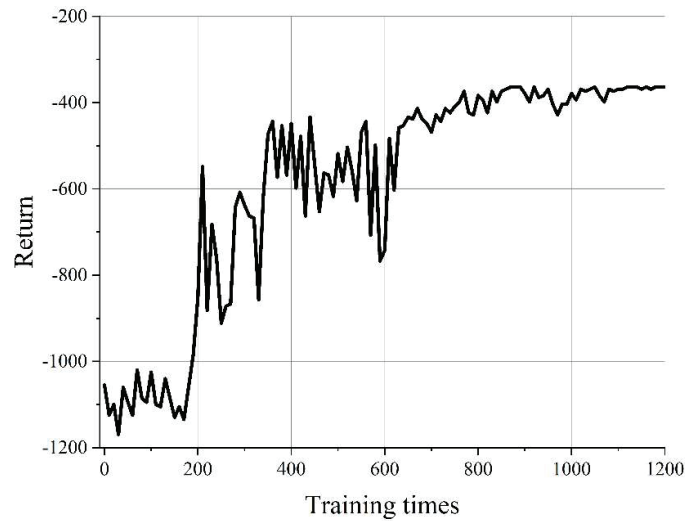
In Table 4, the numbers in parentheses represent the possible nodes for installing equipment, and the feasible space of the decision variables is

Table 3 Investment result using the proposed method

	Energy Storage System	Units	SVC
Step 1	Invest in Node 14	Invest in Node 14	No
Step 2	Invest at Node 29	No	Invest in Node 6
Step 3	Invest in Node 21	Invest in Node 5	Invest in Node 3
Step 4	No	Invest in Node 32	No

Table 4 Comparison with other cases

Case	Investment Cost (k\$)	Operation Cost (k\$)	Total Cost (k\$)
Base case	231.32	132.42	363.74
Case A	321.43	132.42	453.85
Case B	352.19	174.03	526.22

**Figure 3** Convergence result of reinforcement learning.

only partially explored in this case. As a result, despite the relatively low investment cost, there is minimal improvement in performance. Considering the high cost of energy storage system installation, we expand the potential nodes for SVC and units. The results show that while the load curtailment is significantly reduced, the photovoltaic curtailment remains below the target level.

As shown in Figure 3, the return of the proposed reinforcement learning can be converged to -363.74 , meaning the total planning cost of the distribution network is 363.74 k\$ within 1000 training times.

6 Conclusion

This research presents a novel performance-oriented active distribution network planning method using reinforcement learning. Unlike traditional methods that optimize an objective function over a known solution space, reinforcement learning dynamically updates the active distribution system until the desired performance is achieved. The proposed solution effectively achieves the target performance at a low investment cost by minimizing load, wind, and photovoltaic curtailment. As the element set expands to include feeders, switches, and charging stations, and more complex performance indices such as dependability, social conflict, and EV hosting capacity are introduced, the approach can become even more sophisticated over time. With the proposed method, the investment cost and the operation cost are reduced by 32.42% and 23.91%, respectively.

Acknowledgements

This research was Supported by the Research on Economic Capacity Increase and Flexible Network Structure Construction Technology of Distribution Network for High Proportion Distributed Energy Access (Project No. 52022322000U).

References

- [1] Rupolo D, Pereira Junior BR, Contreras J, et al. A new parallel and decomposition approach to solve the medium-and low-voltage planning of large-scale power distribution systems. *Int J Electr Power Energy Syst* 2021;132:107191.
- [2] Nick M, Cherkaoui R, Paolone M. Optimal planning of distributed energy storage systems in active distribution networks embedding grid reconfiguration. *IEEE Trans Power Syst* 2018;33(2):1577–90.
- [3] Alobaidi AH, Khodayar M, Vafamehr A, et al. Stochastic expansion planning of battery energy storage for the interconnected distribution and data networks. *Int J Electr Power Energy Syst* 2021;133:107231.
- [4] Xu X, Li JY, Xu Z, et al. Enhancing photovoltaic hosting capacity-A stochastic approach to optimal planning of static var compensator devices in distribution networks. *Appl Energy* 2019;238:952–62.

- [5] Amrane Y, Boudour M, Belazzoug M. A new Optimal reactive power planning based on Differential Search Algorithm. *Int J Electr Power Energy Syst* 2015;64: 551–61.
- [6] Shen X, Shahidehpour M, Zhu S, et al. Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. *IEEE Trans Smart Grid* 2018;9(2):1425–33.
- [7] Ugranlı F. Analysis of renewable generation's integration using multi-objective fashion for multistage distribution network expansion planning. *Int J Electr Power Energy Syst* 2019;106:301–10.
- [8] Gao HJ, Wang LF, Liu JY, et al. Integrated day-ahead scheduling considering active management in the future smart distribution system. *IEEE Trans Power Syst* 2018;33 (6):6049–61.
- [9] Xie S, Hu Z, Yang L, et al. Expansion planning of active distribution system considering multiple active network managements and the optimal load-shedding direction. *Int J Electr Power Energy Syst* 2020;115:105451.
- [10] Narimani A, Nourbakhsh G, Arefi A, Ledwich GF, Walker GR. SAIDI constrained economic planning and utilization of central storage in rural distribution networks. *IEEE Syst J.* 2019;13(1):842–53.
- [11] Ehsan A, Yang Q. Coordinated investment planning of distributed multi-type stochastic generation and battery storage in active distribution networks. *IEEE Trans Sustain Energy* 2019;10(4):1813–22.
- [12] Ghasemi M, Kazemi A, Bompard E, et al. A two-stage resilience improvement planning for power distribution systems against hurricanes. *Int J Electr Power Energy Syst* 2021;132:107214.
- [13] Najafi Tari A, Sebastian MS, Tourandaz Kenari M. Resilience assessment and improvement of distribution networks against extreme weather events. *Int J Electr Power Energy Syst* 2021;125:106414.
- [14] Mozaffari M, Abyaneh HA, Jooshaki M, et al. Joint expansion planning studies of EV parking lots placement and distribution network. *IEEE Trans Ind Inform* 2020; 16(10):6455–65.
- [15] Melgar-Dominguez OD, Pourakbari-Kasmaei M, Lehtonen M, Sanches Mantovani JR. An economic-environmental asset planning in electric distribution networks considering carbon emission trading and demand response. *Electr Power Syst*
- [16] Xiang Y, Wang Y, Su YC, Sun W, Huang Y, Liu JY. Reliability correlated with optimal planning of the distribution network with a distributed generation [J]. *Electr Power Syst Res* 2020;186:106391.

- [17] Yu T, Feng B, Wei DN, Liu SK, Zhang BB, Ji L. Source-network-load-storage coordinated optimal scheduling for active distribution network with a distributed generation [J]. *Water Resource Hydropower Eng* 2021;52(6):215–22.
- [18] Kong T, Cheng HZ, Li G, Xie H. Review of power distribution network planning [J]. *Power Syst Technol* 2009;33(19):92–9.
- [19] Zhang J.Z. Research on the Locating and Sizing of Multi-type Distributed Generations and the Optimal Operation [D]. North China Electric Power University, 2015.
- [20] Xue HB. Research on distribution network planning with a distributed generation [D]. Xi'an University of Technology; 2018.
- [21] Nie ML, Wang F, Chen C, Wang LX, Dong XZ. Multi-objective distribution network planning considering reliability [J]. *Proc CSU-EPSCA* 2016;28(1):10–6.
- [22] Cai Y, Lin J, Wan C, Song YH. A Bi-level stochastic programming approach for strategic active distribution network operators in the electricity market [J]. *Proceedings of the CSEE*, 2016; 36(20): 5391–5402+5715.
- [23] Koutsoukis NC, Georgilakis PS, Hatzargyriou ND. Multistage coordinated planning of active distribution networks[J]. *IEEE Trans Power Syst* 2018;33(1):32–44.
- [24] Zhang SX, Yuan JY, Cheng HZ, Li K. Optimal distributed generation planning in active distribution network considering demand side management and network reconfiguration [J]. *Proc CSEE* 2016;36(S1): 1–9.
- [25] Wu XM, Dang J, Ren F, Wang SK. Research on Optimal dispatch of active distribution network with distributed energy storage [J]. *J Phys: Conf Ser* 2020; 1634(1): 012121 (6pp).
- [26] Li X, Shan WL, Du DJ, Fei MR. Bilevel Planning of active distribution networks considering demand-side management and distributed generation penetration [J]. *Sci Sin Inform* 2018;48:1333–47.
- [27] Tian LL. Research on the energy management strategy of an active distribution network for improving new renewable energy harvesting [D]. Beijing Jiaotong University; 2018.
- [28] Ebrahimi H, Marjani SR, Talavat V. Optimal planning in active distribution networks considering nonlinear loads using the MOPSO algorithm in the TOPSIS framework [J]. *Int Trans Electric Energy Syst* 2019;30(3):17.

- [29] Jordehi AR. Particle swarm optimisation with opposition learning-based strategy: an efficient optimisation algorithm for day-ahead scheduling and reconfiguration in active distribution systems [J]. *Soft Comput* 2020;24(24):18573–90.
- [30] Malee RK, Chundawat AS, Maliwar N, Sharma AK. distributed generation integrated distribution system expansion planning with uncertainties [J]. *J Intell Fuzzy Syst* 2018;35(5): 4997–5006.
- [31] Babaei S, Jiang RW, Zhao CY. Distributionally robust distribution network configuration under random contingency [J]. *IEEE Trans Power Syst* 2020;35(5): 3332–41.
- [32] Koutsoukis N, Georgilakis P. A chance-constrained multistage planning method for active distribution networks [J]. *Energies* 2020;12(21):4154.
- [33] Gao HJ, Liu JY. Coordinated planning considering different types of distributed generation and load in active distribution network [J]. *Proc CSEE* 2016; 36(18): 4911–4922+5115.
- [34] Ebrahimi H, Marjani SR, Talavat V. Optimal planning in active distribution networks considering nonlinear loads using the MOPSO algorithm in the TOPSIS framework [J]. *Int Trans Electric Energy Syst* 2019;30(3):17.
- [35] Jordehi AR. Particle swarm optimisation with opposition learning-based strategy: an efficient optimisation algorithm for day-ahead scheduling and reconfiguration in active distribution systems [J]. *Soft Comput* 2020;24(24):18573–90.
- [36] Malee RK, Chundawat AS, Maliwar N, Sharma AK. distributed generation integrated distribution system expansion planning with uncertainties [J]. *J Intell Fuzzy Syst* 2018;35(5): 4997–5006.
- [37] Babaei S, Jiang RW, Zhao CY. Distributionally robust distribution network configuration under random contingency [J]. *IEEE Trans Power Syst* 2020;35(5): 3332–41.
- [38] Koutsoukis N, Georgilakis P. A chance-constrained multistage planning method for active distribution networks [J]. *Energies* 2020;12(21):4154.
- [39] Gao HJ, Liu JY. Coordinated planning considering different types of distributed generation and load in active distribution network [J]. *Proc CSEE* 2016; 36(18): 4911–4922+5115.
- [40] Sutton R, Barto A. Reinforcement learning: an introduction. The MIT Press; 2015.

Biographies



Hongtao Li received his bachelor's degree from Tianjin University in 1996 and his master's degree in power systems and automation from North China Electric Power University in 2005. He is currently working at State Grid Beijing Electric Power Company. His research area is distribution network operation and control technology.



Cunping Wang received his bachelor's degree in electrical engineering from Tianjin University in 2008 and Doctor's degree in electrical engineering from Huazhong University of Science and Technology in 2013. He is currently working at State Grid Beijing Electric Power Company. His research areas include active distribution network operation and control, and high reliable power supply for important users.



Hao Tian received his master's degree in electrical engineering from North China University of Technology. He is currently working as a teacher at Wuxi University. He is mainly engaged in energy Internet planning, power system reliability assessment, power system analysis, power grid planning and other research work.



Zhigang Ren received the master's degree in high voltage and insulation technology from Xi'an Jiaotong University in 2009. He is currently working at State Grid Beijing Electric Power Company. His research area is power cable operation and condition monitoring technology.



Erang Zhao received the bachelor's degree in electrical engineering from Hebei of University Technology in 2014 and the master's degree in electrical engineering from Tsinghua University in 2021, respectively. He is currently working as a researcher at Wuxi Research Institute of Applied Technologies, Tsinghua University. His research areas include power system operation and planning, microgrid operation.



Lina Xu received the bachelor's degree in Electronic Information Science and Technology from Xuzhou University of Engineering in 2018. She is currently working at Wuxi Research Institute of Applied Technologies, Tsinghua University. She is mainly engaged in engineering software development work.

