Hamedan University of Technology
Iranian Association of Electrical and Electronics Engineers
Sustainable Energy and Artificial Intelligence
Online ISSN: 3060-8015
Journal Homepage: www.enai.hut.ac.ir

Optimal Switch Reduction in Distribution Networks: A Reinforcement Learning Approach for Dynamic Reconfiguration

Mohammad Amin Sa'edi, Saleh Razini*, Mohamad Amin Ghasemi

[1] *Department of Electrical Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran*
* *Corresponding Author:* s.razini@basu.ac.ir

Article Info

Article type:
Original Article

Article history:
Received 2025-05-10;
Revised 2025-05-26;
Accepted 2025-05-30.

How to cite this article:
Saedi, M. A., Razini, S. and Ghasemi, M. A. (2025). Optimal Switch Reduction in Distribution Networks: A Reinforcement Learning Approach for Dynamic Reconfiguration. *Sustainable Energy and Artificial Intelligence*, 1(3), 145-156.
DOI: 10.61186/seai.2505-1025

Abstract

The continuous fluctuations in electricity demand across different locations in the distribution network led to variations in power flow in the lines and, consequently, changes in power losses and voltage drop. With technological advancements and restructuring of the electricity industry, these changes have become more serious. So, maintaining a single network configuration at all hours may not be optimal. To address this, researchers have proposed and implemented dynamic reconfiguration of distribution networks as an effective solution. This study investigates the dynamic reconfiguration of radial distribution systems, using reinforcement learning to minimize the number of switches. The primary objectives of dynamic reconfiguration are to minimize power losses and enhance bus voltage profiles. A case study is conducted on the IEEE 33-bus standard network. Initially, it is assumed that remotely controllable switches are installed on all lines, allowing full participation in the reconfiguration process. Subsequently, the proposed algorithm for reducing the number of switches identifies critical switches that should be eliminated at each stage, followed by another round of dynamic reconfiguration. For each stage of switch reduction, an economic analysis over 20 years is performed, considering installation and annual maintenance costs alongside power loss costs. Finally, a sensitivity analysis determines the optimal number of switches based on switch costs and electricity tariff rates. The results indicate that lower electricity tariffs and higher switch costs improve the effectiveness of the proposed switch reduction method.

Keywords: Dynamic reconfiguration, Reinforcement learning, Distribution system, Loss reduction, Remotely controllable switches

1. Introduction

Facing challenges in power systems, such as load imbalance, reliability, and power losses, necessitates automation operations like reconfiguration of the distribution system. Reconfiguration involves optimizing the operational state of tie switches installed along distribution feeders. These switches can be strategically opened or closed to achieve better performance while meeting power flow constraints, ensuring a radial network topology, and maintaining required voltage levels.

Researchers have focused on several key objectives when optimizing distribution network reconfiguration, including minimizing energy losses [1], load balancing [2], improving voltage profiles [3], and enhancing system reliability [4]. In some studies, reconfiguration is approached as a multi-objective optimization problem, aiming to satisfy multiple operational goals simultaneously [5].

Given its complexity, network reconfiguration is inherently an optimization challenge, requiring advanced optimization algorithms for effective solutions [6]. Over time, heuristic and metaheuristic techniques have gained popularity in tackling this problem. Some of the most widely used methods include colonial competitive algorithms, binary bat algorithms, genetic algorithms, particle swarm optimization, adaptive reconfiguration, and reinforcement learning-based approaches [1-7].

The increasing integration of renewable distributed generation, energy storage systems, and electric vehicles (EVs) in modern power grids has introduced new technical challenges. The time-dependent nature of EV charging patterns, along with the uncertainties in distributed generation and load demand, significantly impacts the optimal configuration of the network. As a result, a fixed network topology may not be sufficient to ensure efficiency at all times, making dynamic reconfiguration a necessity [8]. In [9] dynamic reconfiguration leverages remotely controllable switches to adjust the distribution network's configuration hour by hour, optimizing power flow based on changing conditions.

Some studies propose placing reconfiguration switches on every feeder, allowing continuous adaptations in response to load variations. However, while increasing switch numbers may enhance operational flexibility, it is often financially impractical due to installation and maintenance costs. On the other hand, minimizing the number of switches reduces expenses but may limit the available reconfiguration options, potentially increasing system losses.

Thus, a trade-off must be established between reducing switch deployment costs and ensuring network efficiency. The optimal allocation of remotely controllable switches in dynamic reconfiguration remains an open research challenge, requiring a technical and economic optimization approach.

Some studies have explored different strategies for dynamic distribution network reconfiguration, including:

- Markov Decision Process-Based Dynamic Reconfiguration [10]: Considering practical operational constraints, this approach minimizes renewable generation curtailment and load shedding.
- Reinforcement Learning for Reconfiguration [11]: Applied to reduce network line losses, optimize load distribution, and enhance distributed energy resource capacity.
- Binary Integer Programming-Based Optimization [12]: Formulates the reconfiguration problem as an optimization task, evaluating switching operations between feeders under different load scenarios.
- Critical Switch Identification for Reconfiguration [13]: Develops a method for selecting key switches based on daily load curves, focusing on minimizing switching costs and balancing loads across feeders.
- Smart Grid Optimization with Remote-Controlled Switches [14]: Examines the replacement of manual switches with automated remote-controlled switches, integrating energy storage and renewable energy sources for improved reliability.
- Optimization of Renewable Energy Integration via Reconfiguration [15]: Proposes an approach to identify critical switches that facilitate efficient integration of renewable distributed generation, while accounting for uncertainties in loads and generation capacity.
- Remote-Controlled Switch Allocation for System Restoration [16]: Uses integer programming to optimize switch placement, improving system reliability while reducing operational costs.

The related studies in the field of dynamic reconfiguration are summarized and compared in Table 1.

Table 1. Review of the related papers

Ref	Objective(s)	Method	Dynamic reconfiguration	Switch reduction
[1]	Loss reduction	Ant colony	No	No
[8]	Reliability	Reinforcement learning	No	Yes
[11]	Loss reduction	Reinforcement learning	No	Yes
[10]	Minimizing renewable generation curtailment and load shedding	Markov Decision Process	No	No
[15]	Maximizing integration of renewable distributed generation	Mixed integer programming	Yes	Yes
[16]	Reliability	Mixed integer programming	Yes	Yes

The review of existing literature reveals that most studies focus primarily on minimizing power losses, underscoring the critical role of energy efficiency and cost reduction. However, these studies often assume that all switches in the distribution network are remotely controllable, which is unrealistic due to high installation and maintenance expenses. Few studies have directly addressed the optimal number of remotely controllable switches. This paper aims to fill that gap by proposing an optimization framework for dynamic reconfiguration, leveraging reinforcement learning to optimize switch deployment and minimize distribution losses. Q-learning, being a widely studied and well-established off-policy algorithm, was chosen for its ability to handle discrete action spaces efficiently. Since the network reconfiguration problem inherently involves discrete switch operations.

The key objectives of this research include:

- Developing a reinforcement learning algorithm to optimize dynamic distribution network reconfiguration, achieving minimum system losses and voltage profile enhancement.
- Identifying the optimal number of reconfiguration switches and the best locations.
- Establishing a balanced trade-off between minimizing switch numbers and reducing system losses.

Achieving the above objectives, the main contributions of the proposed method are:

1) Three improvements in the reward function of reinforcement learning, leading to more efficient network reconfiguration results.
2) A novel switch reduction algorithm, combined with economic sensitivity analysis, to determine the optimal number of switches under varying operational conditions.

The rest of this paper is organized as follows. Section 2 presents reinforcement learning as an optimization method for switch allocation. The proposed switch reduction strategy is introduced in section 3, detailing its implementation in dynamic reconfiguration alongside an economic cost analysis. Section 4 provides a case study on the standard 33-bus distribution network, discussing results across different operating conditions. Section 5 concludes the paper, summarizing the key findings and implications of this research.

2. Reinforcement Learning

Reinforcement Learning (RL) is a subset of machine learning in which the problem is divided into two key components: the agent and the environment. The agent interacts with the environment by selecting actions based on its current state. In response, the environment provides feedback in the form of a reward and transitions to a new state, guiding the agent in learning the optimal decision-making process.

RL algorithms are generally classified into On-policy and Off-policy methods. One of the most widely used Off-policy algorithms is Q-learning, which plays a crucial role in reinforcement learning frameworks. In Q-learning, the agent first observes the environment and selects an action. The environment then returns the next state along with the reward associated with the chosen action. The agent processes this feedback to refine its decision-making strategy, continuously iterating until it converges on an optimal policy.

The optimal Q-value function in Q-learning is defined using the Bellman equation, as shown in Equation (1) [17]:

$$Q^*(s,a) = E(r(s,a) + \gamma \max_{a'} Q^*(s',a')) \tag{1}$$

This equation represents the expected cumulative reward, starting from state (s) and taking action (a), while subsequently selecting the optimal actions (a′) in future steps until reaching the maximum desired performance $Q^*(s,a)$. The discount factor γ, which falls within the range $0 \leq \gamma \leq 1$, is used to adjust the importance of future rewards, ensuring the learning process prioritizes immediate and long-term benefits effectively [18]. Once the optimal Q-value ($Q^*(s,a)$) is determined, selecting the best possible action a^* is straightforward. A greedy policy is often employed for this purpose, where the algorithm evaluates all available actions for a given state and chooses the one yielding the highest Q-value, as expressed in Equation (2):

$$a^* = arg \max_{a} Q^*(s,a) \tag{2}$$

The Q-values, which store learned values, are typically structured in the form of a matrix, indexed by both states and actions. Initially, arbitrary values are assigned, and as the agent continues to interact with the environment, the function gradually approximates the optimal Q-values based on reinforcement learning principles. The Q-value of the current state and action are updated iteratively using Equation (3):

$$\mathrm{Q(s,a)} = (1-\alpha).Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a')) \tag{3}$$

It is important to note that the updated Q-value for Q(s,a) depends not only on its previous value but also on the expected future rewards from subsequent actions. The learning rate parameter (α, where $0 \leq \alpha \leq 1$) determines the extent to which new information overrides past stored values. The term $(1-\alpha)$ signifies the proportion of prior Q-values retained as memory in the function, ensuring a balanced learning process throughout multiple iterations [18].

3. Proposed Method of Switches Reduction

This study presents a reinforcement learning-based approach to determine both the optimal number of switches and the most efficient network configuration under dynamic load conditions. The method starts by analyzing network data and providing hourly load values for each bus. Initially, it is assumed that every feeder contains remotely controllable switches, with all switches closed at the outset. The first objective of this approach is to identify the best network configuration that optimizes objectives without violating system constraints. This is accomplished using reinforcement learning algorithms.

Since deploying remotely controlled switches across all feeders is highly cost-intensive, the second objective is to minimize the number of necessary switches in the network. To achieve this, the proposed switch reduction algorithm systematically eliminates switches that undergo minimal state changes over 24 hours. Once a switch is removed, the network is reconfigured using the remaining switches, and this process is repeated until a violation of system constraints occurs. At each elimination step, an economic analysis is conducted to evaluate the trade-off between installation costs and energy losses. Ultimately, the optimal number of switches is determined by examining various electricity pricing scenarios.

The method follows four main steps:

1) Developing a reinforcement learning model to identify the optimal network configuration for static load.
2) Expanding the model for dynamic network reconfiguration without restrictions on switch numbers.
3) Gradually reducing the number of switches based on optimization results.
4) Conducting an economic evaluation to determine the optimal switch count.

3-1. Reinforcement Learning Model for Static Load

This section outlines a reinforcement learning-based strategy for determining the most efficient network configuration for static load. In reinforcement learning, the problem is represented using states, actions, and rewards. Since the primary goal is minimizing losses while maintaining voltage constraints, the states are defined based on voltage violations across network buses. If the voltage exceeds allowable limits due to load variations, the algorithm must adjust the network configuration to restore compliance. This is achieved by modifying switch states, either opening or closing them.

- **State Representation**

If the system has n voltage variables, the total number of possible states, each representing either constraint violation or compliance, will be 2^n. However, defining states at the bus level leads to an excessively large state space (e.g., in a 33-bus system, over 8.5 billion possible states exist). To simplify the model, voltage constraint violations are instead evaluated at the loop level rather than at individual buses.

Network loops are identified using the graph of the network, taking into account that all possible switches are closed. If voltage values across an entire loop remain within permissible limits, that loop is marked as constraint-compliant (binary value=0). Conversely, if a voltage constraint violation occurs within a loop, its corresponding binary variable is set to 1. Thus, in a network with m loops, the state is represented as an m-bit binary number, where each bit reflects the voltage compliance for a specific loop. The optimal state corresponds to a fully zero binary representation, indicating no constraint violations.

- **Action Selection**

Whenever voltage violations occur in the network, an action must be selected from the available action set (A) to restore voltage compliance. This is achieved through switch reconfiguration, ensuring that voltage constraints across all loops are satisfied.

Each switch modification represents an individual action. However, maintaining a radial network topology is a key requirement. To preserve radial connectivity, only one switch per loop may be opened at any given time. The total number of possible actions is determined by Equation (4):

$$A = \prod_{k=1}^{m} d_k \tag{4}$$

where m represents the total number of loops, and d_k denotes the number of branches in the *k*th loop.

- **Reward Function**

Since loss minimization is a critical objective, the reward function in base reference [19] is simply the negative of total system losses. After each action, the agent receives a higher reward if the loss is reduced.

- **Enhanced Reward Function**

To improve the method for dynamic load, this study introduces three key improvements to the baseline reward function:

1) In standard approaches, rewards are negatively

correlated with total energy losses. However, under dynamic load conditions, a reduction in total losses may simply result from lower system demand, rather than an improved switch configuration. To address this, we redefine the reward function using relative loss, calculated as the inverse of the ratio of losses to load variation, as shown in Equation (5):

$$r = 1/P_L \tag{5}$$

where P_L, the relative loss, is computed using Equation (6):

$$P_L = \frac{p_{loss}}{p_t} \tag{6}$$

Here, p_{loss} represents total system losses, while p_t denotes total load demand.

2) To further enhance voltage stability, an additional voltage deviation term is incorporated into the reward function. If an action reduces average voltage deviation, an extra reward is assigned, as defined in Equation (7):

$$r = (1/P_L) + 1/(\sum |V_i - 1|)/n_b) \tag{7}$$

where n_b is the total number of buses, and V_i represents the per-unit voltage at each bus. The refined reward function prioritizes configurations that simultaneously minimize losses and voltage deviations, leading to better switch selection.

3) In Q-learning, the learning rate α determines the balance between new learning and stored knowledge. During early training iterations, past experiences have limited significance. However, as training progresses, previously learned information becomes more valuable, necessitating a gradually decreasing α. The adaptive learning rate proposed in this study is defined in Equation (8):

$$\alpha = (0.9\mathrm{K} + \mathrm{k})/(\mathrm{K} + 10 \times \mathrm{k}) \tag{8}$$

where k is the current iteration, and K represents the total number of training iterations.

- **Problem-Solving Process**

The learning process begins with initializing input data and constructing the Q-matrix. The number of Q-matrix rows equals the total states, while the columns correspond to the available action set. first, all Q-values are set to zero, indicating no prior system knowledge.

The algorithm then enters an iterative training cycle, where each step involves:

1) Random selection of switch configurations.
2) Power flow analysis if the system is radial.
3) Calculating rewards and updating Q-values.
4) The state transition to the next state
5) Repeating the process until optimal Q-values are learned.

3-2. Dynamic Reconfiguration

In power distribution networks where bus loads fluctuate over time, a single fixed network configuration may not consistently deliver optimal performance. To adapt to these variations, an hourly reconfiguration process is required to ensure the system is always operating in its most efficient state. This approach involves training a Q-learning algorithm and providing it with hourly load data for all buses within the distribution network. At each time interval, power flow analysis determines the system's current state. The trained Q-matrix then suggests the optimal reconfiguration action based on the state. After implementing the action, another power flow calculation refines the network structure and evaluates the corresponding energy losses. This process repeats for every hour in a full daily cycle, maintaining an adaptive and efficient system.

3-3. Optimizing the Number of Remote-Controlled Switches

In the previous section, it was assumed that every feeder had remote-controlled switches, allowing full participation in dynamic reconfiguration. However, while this approach offers flexibility, it significantly increases system costs, making it economically impractical. To reduce expenses while still preserving the benefits of dynamic reconfiguration, this study proposes a switch optimization strategy that eliminates switches with minimal impact on network efficiency.

The process begins by executing the dynamic reconfiguration algorithm and identifying switches that remain unchanged throughout the day-these switches are the first candidates for removal. Next, the method systematically eliminates switches that undergo the fewest open-close operations within the daily cycle. After each switch removal, the network reconfiguration process is repeated, ensuring the remaining switches maintain optimal system performance. This iterative elimination process continues until all remaining switches are deemed essential for operational efficiency.

3-4. Economic Analysis

To determine the optimal number of switches, this study evaluates their impact on total system costs over the operational lifespan. These costs include investment and operational expenses for remote-controlled switches and energy loss costs incurred throughout the project's duration.

During the switch reduction process, system performance is simulated at each stage, calculating the combined cost of switch deployment and energy losses over the project's lifetime. The configuration that minimizes costs is selected as the optimal solution.

The total system cost is expressed as the Net Present Cost (NPC), defined in Equation (9):

$$NPC_{total} = NPC_{loss} + NPC_{cap,sw} + NPC_{m,sw} \quad (9)$$

Where NPC_{loss} is the net present cost of energy losses, $NPC_{cap,sw}$ is the net present cost of switch investment, and $NPC_{m,sw}$Net present cost of switch maintenance.

The annual cost of energy losses, $C_{loss,}$ is determined by multiplying the total annual energy loss, $E_{loss,}$ by the electricity tariff, T_{elec}, as shown in Equation (10). Annual energy losses are derived from daily loss values, using Equation (11), where $loss_{24}$ represents daily network losses.

$$C_{loss} = E_{loss} \times T_{elec} \quad (10)$$

$$E_{loss} = loss_{24} * 365 \quad (11)$$

On average, annual maintenance costs for switches amount to 3% of the total switch investment, $C_{cap,sw}$. The total annual maintenance cost, $C_{m,sw}$ is given in Equation (12):

$$C_{m,sw} = 0.03 \times C_{cap,sw} \times N_{sw} \quad (12)$$

where N_{sw} represents the number of remote-controlled switches utilized in dynamic reconfiguration. The total investment cost for switches depends on the unit price per switch, $P_{switch,}$ and the required switch count, N_{sw}, calculated using Equation (13):

$$C_{Cap,sw} = N_{sw} \times P_{switch} \quad (13)$$

The net present cost (NPC) of all expenses, based on annual costs (A), is determined using Equation (14):

$$NPC = \left[\frac{(1+i)^n - 1}{i(1+i)^n}\right] \times A \quad (14)$$

where n represents the operational period (years) and *i* denotes the annual interest rate.

4. Case Study

4-1. Overview of the Network

The proposed method has been tested on the standard IEEE 33-bus radial distribution network under various operating conditions. Figure 1 illustrates the network's basic configuration. This network consists of 37 branches, each equipped with a switch that can be opened or closed. When all switches remain closed, the network forms five loops, which disrupt the radial topology [20].

Table 2 lists these loops and their corresponding branches. To preserve the radial configuration, one switch per loop must be opened while ensuring that no loop has multiple open branches, as this would compromise network continuity. The goal of dynamic reconfiguration is to select five optimal switches for each hour to improve network performance in response to fluctuating generation and demand.

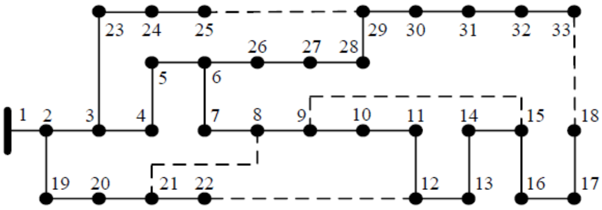

Fig. 1. IEEE 33 bus system

Table 2. Loops and the relative branches in the case study

loop	Branches [20]
1	2-3-4-5-6-7-18-19-20-33
2	9-10-11-12-13-14-34
3	8-9-10-11-21-33-35
4	6-7-8-15-16-17-25-26-27-28-29-30- 31-32-34-36
5	3-4-5-22-23-24-25-26-27-28-37

All actions for reconfiguration are the possible choices of selecting one switch in each loop. So the total number of possible reconfiguration actions will be 10 × 7 × 7 × 16 × 11 = 86,240, although not all are feasible. Similarly, the number of possible system states, based on voltage constraints across network loops, is $2^5 = 32$.

To implement dynamic reconfiguration, various hourly load profiles were simulated across the 33 buses of the distribution system. Figure 2 shows the hourly load profile for bus 15, which follows a typical residential demand pattern.

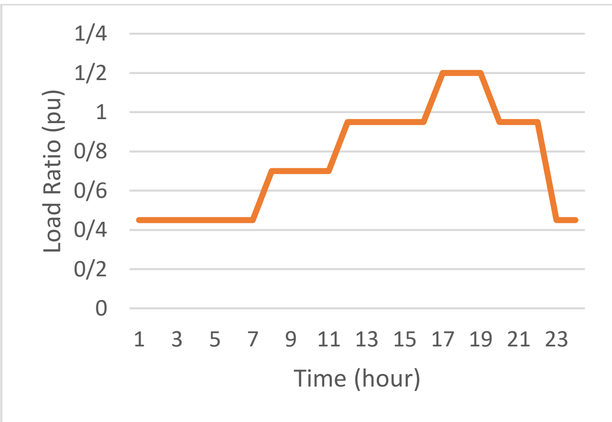

Fig. 2. A residential load curve of in the system

4-2. Validating the Model for Static Load

To verify the reliability of the proposed method, a test was conducted using static load conditions, similar to those in reference [19], with the primary objective of minimizing system losses. The results were compared with findings from previous studies. The optimal configuration obtained in this study involves opening switches on lines 7, 11, 14, 31, and 37, while keeping all other switches closed. Under this setup, total system losses will be 112.95 kW, and the average voltage deviation across buses will be 0.165p.u.

Fig. 3. presents the bus voltage profile, confirming that voltage constraints are satisfied at all buses. Table 3 provides a comparison of the proposed method with other approaches, demonstrating its enhanced efficiency.

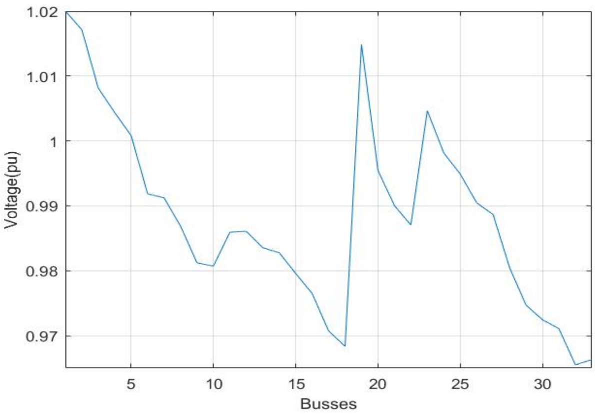

Fig. 3. Voltage profile after reconfiguration for static load

Table 3. Results of different methods for static load

minimum bus voltage	Power loss(kW)	Opened switches	Reconfiguration method
0.963	124.65	7-10-13-25-31	Q learning [19]
0.962	116.98	6-9-14-32-37	Fuzzy Adaptive Evolutionary Programming [17]
0.965	112.95	7-11-14-31-37	Proposed method

4-3. Results of Dynamic Reconfiguration

4-3-1. System Performance Before Reconfiguration

To establish a baseline, the system was initially set up with predefined open branches, represented by dashed lines in Fig. 1. Without implementing reconfiguration, power flow analysis under dynamic load conditions revealed total energy losses of 2109.18kWh over 24 hours. Table 4 details hourly losses and voltage deviations.

4-3-2. System Performance After Dynamic Reconfiguration with All Switches

Applying dynamic reconfiguration with switches installed on all branches, the system generated six distinct configurations, each tailored to specific hours of operation. Table 4 presents the open branches, system losses, and voltage deviations for each hour, along with total system losses.

4-3-3. Performance After Reconfiguration Using Critical Switches

Following full-system reconfiguration, 13 switches were identified as critical, meaning they frequently change state and significantly influence network efficiency. These critical switches include: 6, 7, 9, 10, 13, 14, 16, 27, 28, 30, 31, 35, and 37.

Switches that remained inactive during the dynamic reconfiguration process were excluded from further optimization, and the Q-learning algorithm was retrained exclusively using critical switches to allow more efficient configuration selection.

The updated dynamic reconfiguration results, including open branches, system losses, and voltage deviation profiles, are presented in Table 4. Under this revised setup, the number of actively changing switches decreased from 13 to 8, and total energy losses were further reduced, demonstrating improved algorithm convergence.

The enhancement stems from reducing the number of possible reconfiguration actions, which simplifies the Q-matrix structure, allowing for faster and more efficient training. While this setup yields lower losses and better voltage stability, retaining numerous remote-controlled switches remains costly. Therefore, the next phase involves applying the switch reduction algorithm.

4-3-4. Results After Implementing the Switch Reduction Algorithm

To further optimize switch deployment, the algorithm identifies switches with minimal activity and shortest operational duration over a full day, progressively removing them while maintaining system functionality. A key condition in this process is ensuring at least one active switch per network loop. The process continues until the power flow analysis converges, confirming robust system operation.

Upon execution, switch 6 was identified as the first candidate for removal. After eliminating switch 6, a new dynamic reconfiguration was performed, with updated open branches, system losses, and voltage deviations recorded in Table 5. Following switch 6's removal, total system losses increased by only 9.3kW, indicating minimal efficiency degradation while simultaneously reducing installation costs. However, determining the ideal switch count requires a full economic assessment, covered in the next section.

Further iterations identified switches 6, 28, 14, and 10 as the next removal candidates. The results after eliminating switches 6, 28, and 14 are shown in Table 5. However, removing switch 10 disrupted the algorithm's ability to generate optimal hourly configurations, signaling the termination of the switch reduction process.

The final configuration reveals that switches 7, 13, 31, and 37 remain open throughout the day, while switches 9 and 10 change state only once within 24 hours. In this setup, only two switches are required for dynamic reconfiguration, which leads to the lowest capital cost, but total energy losses increase slightly compared to previous configurations.

4-4. Economic Evaluation and Discussion

To analyze the financial impact of the proposed dynamic reconfiguration method and find the optimum switches, the cost models outlined in Section 3 are applied. The economic assessment of the system using critical switches is as follows:
Before reducing the number of switches, the system operates with eight remote-controlled switches, leading to an annual energy loss of:

$$E_{loss} = 1782.08 * 365 = 650459.2 \quad (15)$$

Considering an electricity tariff, T_{elec}, and an interest rate of 4% over 20 years, the Net Present Cost (NPC) of energy losses is calculated as:

$$NPC_{loss} = \left[\frac{(1+0.04)^{20}-1}{0.04\times(1+0.04)^{20}}\right]\times 650459 \times T_{elec} = 8839737 \times T_{elec} \quad (16)$$

The capital investment cost required for switch installation is:

$$C_{cap,sw} = 8 \times P_{switch} \quad (17)$$

where P_{switch} represents the unit cost per switch. The maintenance cost per switch, taking into account periodic servicing expenses, is given by:

$$NPC_{m,sw} = \left[\frac{(1+0.04)^{20}-1}{0.04\times(1+0.04)^{20}}\right] \times 0.03 \times P_{switch} = 0.407 \times P_{switch} \quad (18)$$

Thus, the total initial financial commitment for the system, combining energy losses, investment, and maintenance, is:

$$NPC_{total} = 8839737 \times T_{elec} + 11.25 \times P_{switch} \quad (19)$$

Similar evaluations were conducted for each stage of switch elimination, and the results are summarized in Table 6, showing how the optimal number of switches and total system costs vary with switch price and electricity tariff.

Table 4. Result of dynamic reconfiguration before switch reduction

hour	Before Reconf.		Dynamic Reconfiguration with all switches			Dynamic Reconfiguration with critical switches		
	Losses (kW)	Avg. Voltage Deviation (p.u)	Open Branches	Losses (kW)	Avg. Voltage Deviation (p.u)	Open Branches	Losses (kW)	Avg. Voltage Deviation (p.u)
0	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
1	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
2	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
3	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
4	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
5	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
6	133.76	0.017	31-27-14-9-7	98.14	0.01	37-31-14-9-6	91.08	0.012
7	76.82	0.012	37-31-13-9-6	76.82	0.012	37-31-13-9-7	75.85	0.011
8	76.82	0.012	37-31-13-9-6	76.82	0.012	37-31-13-9-7	75.85	0.011
9	76.82	0.012	37-31-13-9-6	76.82	0.012	37-31-13-9-7	75.85	0.011
10	76.82	0.012	37-31-13-9-6	76.82	0.012	37-31-13-9-7	75.85	0.011
11	56.99	0.01	30-28-14-9-7	57.74	0.01	37-31-14-9-7	55.93	0.008
12	56.99	0.01	30-28-14-9-7	57.74	0.01	37-31-14-9-7	55.93	0.008
13	56.99	0.01	30-28-14-9-7	57.74	0.01	37-31-14-9-7	55.93	0.008
14	56.99	0.01	30-28-14-9-7	57.74	0.01	37-31-14-9-7	55.93	0.008
15	56.99	0.01	30-28-14-9-7	57.74	0.01	37-31-14-9-7	55.93	0.008
16	82.39	0.015	37-35-16-10-7	90.31	0.014	37-31-13-9-7	78.12	0.012
17	82.39	0.015	37-35-16-10-7	90.31	0.014	37-31-13-9-7	78.12	0.012
18	82.39	0.015	37-35-16-10-7	90.31	0.014	37-31-13-9-7	78.12	0.012
19	79.28	0.013	37-31-14-10-6	79.08	0.013	31-28-13-10-7	77.25	0.011
20	79.28	0.013	37-31-14-10-6	79.08	0.013	31-28-13-10-7	77.25	0.011
21	79.28	0.013	37-31-14-10-6	79.08	0.013	31-28-13-10-7	77.25	0.011
22	47.74	0.009	37-31-13-9-7	47.12	0.008	31-28-13-10-7	47.62	0.009
23	47.74	0.009	37-31-13-9-7	47.12	0.008	31-28-13-10-7	47.62	0.009
Daily Average	87.88	0.013	--------	78.15	0.011	-------	74.25	0.01
Energy Loss in 24 Hours	2109.18 (kWh)		1885.42 (kWh)			1782.07 (kWh)		

Table 5. Result of the switch reduction algorithm

hour	Dynamic Reconfig. with Removal of Switch 6			Dynamic Reconfig. with Removal of Switches 6 and 28			Dynamic Reconfig. with Removal of Switches 6, 28, and 14		
	Open Branches	Losses (kW)	Avg. Voltage Deviation (p.u)	Open Branches	Losses (kW)	Avg. Voltage Deviation (p.u)	Open Branches	Losses (kW)	Avg. Voltage Deviation (p.u)
0	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
1	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
2	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
3	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
4	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
5	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
6	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01	37-31-13-9-7	92.95	0.01
7	37-31-14-9-7	75.31	0.01	37-31-13-9-7	75.86	0.011	37-31-13-9-7	75.86	0.011
8	37-31-14-9-7	75.31	0.01	37-31-13-9-7	75.86	0.011	37-31-13-9-7	75.86	0.011
9	37-31-14-9-7	75.31	0.01	37-31-13-9-7	75.86	0.011	37-31-13-9-7	75.86	0.011
10	37-31-14-9-7	75.31	0.01	37-31-13-9-7	75.86	0.011	37-31-13-9-7	75.86	0.011
11	37-31-14-9-7	55.93	0.008	37-31-13-10-7	56.97	0.008	37-31-13-9-7	56.37	0.008
12	37-31-14-9-7	55.93	0.008	37-31-13-10-7	56.97	0.008	37-31-13-9-7	56.37	0.008
13	37-31-14-9-7	55.93	0.008	37-31-13-10-7	56.97	0.008	37-31-13-9-7	56.37	0.008
14	37-31-14-9-7	55.93	0.008	37-31-13-10-7	56.97	0.008	37-31-13-9-7	56.37	0.008
15	37-31-14-9-7	55.93	0.008	37-31-13-10-7	56.97	0.008	37-31-13-9-7	56.37	0.008
16	37-31-13-10-7	78.95	0.013	37-31-14-9-7	77.08	0.012	37-31-13-9-7	78.12	0.012
17	37-31-13-10-7	78.95	0.013	37-31-14-9-7	77.08	0.012	37-31-13-9-7	78.12	0.012
18	37-31-13-10-7	78.95	0.013	37-31-14-9-7	77.08	0.012	37-31-13-9-7	78.12	0.012
19	31-28-14-9-7	76.41	0.011	37-31-14-9-7	77.54	0.011	37-31-13-10-7	78.38	0.011
20	31-28-14-9-7	76.41	0.011	37-31-14-9-7	77.54	0.011	37-31-13-10-7	78.38	0.011
21	31-28-14-9-7	76.41	0.011	37-31-14-9-7	77.54	0.011	37-31-13-10-7	78.38	0.011
22	37-31-14-9-7	46.83	0.008	37-31-13-9-7	47.12	0.008	37-31-13-9-7	47.12	0.008
23	37-31-14-9-7	46.83	0.008	37-31-13-9-7	47.12	0.008	37-31-13-9-7	47.12	0.008
Daily Average		74.64	0.01	--------	74.87	0.01	-------	74.98	0.01
Energy Loss in 24Hours	1791.37 (kWh)			1797.08 (kWh)			1799.7 (kWh)		

Table 6. Economic Analysis

Reconfiguration method	Daily configuration states (No.)	Daily loss (kWh)	Total cost
Basic configuration (before)	-----	2109.18	-----
Reconfiguration using 8 switches	4	1782.07	$8839737 \times T_{elec} + 11.25 \times P_{switch}$
Reconfiguration using 6 switches	4	1791.37	$8886034 \times T_{elec} + 8.44 \times P_{switch}$
Reconfiguration using 4 switches	3	1797.08	$8914143 \times T_{elec} + 5.62 \times P_{switch}$
Reconfiguration using 2 switches	2	1799.7	$8927135 \times T_{elec} + 2.81 \times P_{switch}$

The findings highlight a clear trade-off: increasing the number of remote-controlled switches reduces energy losses, but comes with higher installation costs.

To determine the optimal switch count, a sensitivity analysis was adopted for the switch price under two scenarios for the electricity price:

1) Global average electricity tariff: 8¢/kWh
2) Subsidized tariff (reflecting production cost of electricity in Iran): 1¢/kWh

When evaluating total system costs using an electricity tariff of 8¢/kWh, the cost functions yield:

$$NPC_{tot,8} = 707178 + 11.25 \times P_{sw} \quad (20)$$

$$NPC_{tot,6} = 710882 + 8.44 \times P_{sw} \quad (21)$$

$$NPC_{tot,4} = 713131 + 5.62 \times P_{sw} \quad (22)$$

$$NPC_{tot,2} = 714170 + 2.81 \times P_{sw} \quad (23)$$

Comparing these cost functions, the economic threshold price for switches is \$828. This means if the switch price is below \$828, deploying eight switches is the most cost-effective approach. But in the case where of switch price exceeds \$828, utilizing the switch reduction algorithm is a more economical solution.

Repeating the same cost evaluations under a subsidized tariff (1¢/kWh) results in:

$$NPC_{\text{tot},8} = 88397 + 11.25 \times C_{\text{sw}} \quad (24)$$

$$NPC_{\text{tot},6} = 88860 + 8.44 \times C_{\text{sw}} \quad (25)$$

$$NPC_{\text{tot},4} = 89141 + 5.62 \times C_{\text{sw}} \quad (26)$$

$$NPC_{\text{tot},2} = 89271 + 2.81 \times C_{\text{sw}} \quad (27)$$

Under subsidized pricing, the economic threshold price for switches is only \$104. This means, deploying eight switches is economically viable only if the switch price is below \$104.

If the switch price exceeds \$104, reducing the number of switches leads to lower total system costs.

Given that this study examines a medium-voltage (MV) distribution network, tie switches for reconfiguration at this voltage level are relatively expensive. While a higher switch count improves voltage stability and reduces system losses, it also significantly increases infrastructure costs. The findings suggest that expanding switch deployment is cost-effective only when electricity tariffs are high, allowing the reduction in losses to offset switch costs. In low-tariff scenarios, the switch reduction algorithm becomes essential, as the economic threshold (\$104) is far lower than actual MV switch prices.

5. Conclusion

This study explored optimal switch deployment for dynamic reconfiguration in power distribution systems. Conventional approaches assume that each feeder line is equipped with remote-controlled switches, resulting in high installation costs. The proposed method, leveraging reinforcement learning (Q-learning), optimizes network reconfiguration to minimize annual energy losses and enhance voltage profiles. Several improvements were made to the reward function, leading to better algorithm convergence. Dynamic load was modeled as six distinct hourly demand profiles, producing six optimized network configurations per day. These configurations were determined through sequential switching operations. To reduce infrastructure costs, a gradual switch reduction process was introduced, systematically eliminating switches with minimal contribution to reconfiguration efficiency. The algorithm was recalibrated at each step, determining the optimal number of switches based on economic comparisons.

Findings confirm that increasing switch count lowers system losses but requires a higher initial investment. The optimal switch configuration was determined by evaluating total investment vs. operational costs, revealing that economic feasibility depends on electricity tariffs and switch prices. Two tariff scenarios were analyzed, showing that for medium-voltage networks, the switch reduction algorithm significantly reduces total costs. Future research could explore alternative state representations within reinforcement learning to enhance algorithm adaptability for varying grid conditions.

Despite the successful performance of the proposed method in improving system operation, the algorithm has certain limitations in identifying the optimal system conditions under specific scenarios. This is because the states are defined based on voltage violations in the distribution network loops. Therefore, two cases can be envisioned where voltage violations occur at two buses within the same loop. However, since the algorithm is loop-sensitive, it cannot differentiate between these two cases and will provide the same response for both. Future studies can work on refining the state definition within the reinforcement learning algorithm to address this limitation.

References

[1] Uikey, P., & Kori, A. (2021, July). In Distribution System Minimum Loss Reconfiguration using Ant Colony Optimization Algorithm. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1196-1199). IEEE.

[2] Muruganantham, B., Selvam, M. M., Gnanadass, R., & Padhy, N. P. (2017, December). Energy loss reduction and load balancing through network reconfiguration in practical LV distribution feeder using GAMS. In *2017 7th International Conference on Power Systems (ICPS)* (pp. 509-513). IEEE.

[3] Shetty, V. J., & Ankaliki, S. G. (2019, August). Joint Reconfiguration of Electrical Distribution System for Power Loss Reduction and Voltage Profile Enhancement: Using BPSO. In *2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC)* (pp. 458-463). IEEE.

[4] Li, Z., Wu, W., Zhang, B., & Tai, X. (2019). Analytical reliability assessment method for complex distribution networks considering post-fault network reconfiguration. *IEEE Transactions on Power Systems*, *35*(2), 1457-1467.

[5] Sam, A., Daraskar, S. S., Gotham, H. V., Pemmada, S., & Patne, N. R. (2022, July). A Novel Multi-Objective based Optimal Reconfiguration of A

Distribution Network by A Closed Switch Heuristic Method. In *2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP)* (pp. 01-05). IEEE.

[6] Karami, S., Tofighi, A., & Kia, M. (2019). Optimal Reconfiguration in Smart Distribution Network in the Presence of DG Resources under Uncertainties. *Journal of Energy Planning and Policy Research*, *5*(3), 123-150.

[7] Rajaram, R., Kumar, K. S., & Rajasekar, N. (2015). Power system reconfiguration in a radial distribution network for reducing losses and to improve voltage profile using modified plant growth simulation algorithm with Distributed Generation (DG). *Energy Reports*, *1*, 116-122.

[8] Oh, S. H., Yoon, Y. T., & Kim, S. W. (2020). Online reconfiguration scheme of self-sufficient distribution network based on a reinforcement learning approach. *Applied energy*, *280*, 115900.

[9] Saedi, M.A., Razini, S., & Ghasemi, M.A. (2023). Dynamic reconfiguration of distribution network in presence of electric vehicle charging station using RL, 31st international conference on electrical engineering.

[10] Wang, C., Lei, S., Ju, P., Chen, C., Peng, C., & Hou, Y. (2020). MDP-based distribution network reconfiguration with renewable distributed generation: Approximate dynamic programming approach. *IEEE Transactions on Smart Grid*, *11*(4), 3620-3631.

[11] Gao, Y., Shi, J., Wang, W., & Yu, N. (2019, October). Dynamic distribution network reconfiguration using reinforcement learning. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)* (pp. 1-7). IEEE.

[12] Chen, C. S., & Cho, M. Y. (2002). Energy loss reduction by critical switches. *IEEE Transactions on Power Delivery*, *8*(3), 1246-1253.

[13] Chen, C. S., & Cho, M. Y. (2002). Determination of critical switches in distribution system. *IEEE transactions on power delivery*, *7*(3), 1443-1449.

[14] Santos, S. F., Fitiwi, D. Z., Cruz, M. R., Santos, C., & Catalão, J. P. (2019). Analysis of switch automation based on active reconfiguration considering reliability, energy storage systems, and variable renewables. *IEEE Transactions on Industry Applications*, *55*(6), 6355-6367.

[15] Lei, S., Hou, Y., Qiu, F., & Yan, J. (2017). Identification of critical switches for integrating renewable distributed generation by dynamic network reconfiguration. *IEEE Transactions on Sustainable Energy*, *9*(1), 420-432.

[16] López, J. C., Rider, M. J., Garcia, A. V., Cavalcante, P. L., Miranda, L. F., & Martins, L. L. (2017, September). Optimization approach for the allocation of remote-controlled switches in real-scale electrical distribution systems. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)* (pp. 1-6). IEEE.

[17] Gholizadeh, N., Kazemi, N., & Musilek, P. (2023). A comparative study of reinforcement learning algorithms for distribution network reconfiguration with deep Q-learning-based action sampling. *Ieee Access*, *11*, 13714-13723.

[18] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

[19] Sampath, K., Pattabiraman, S., Kannan, M., & Narayanan, K. (2019, July). Power loss minimization in radial distribution system through network reconfiguration. In *2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP)* (pp. 1-5). IEEE.

[20] Mahdavi, M., Alhelou, H. H., Hatziargyriou, N. D., & Jurado, F. (2021). Reconfiguration of electric power distribution systems: Comprehensive review and classification. *IEEE Access*, *9*, 118502-118527.

Biography

MohamadAmin Sa'edi received his B.Sc. degree in Electrical Engineering from the University of Shahid Rajaee Teacher Training University, Tehran, Iran, in 2020, and his M.Sc also in Electrical Engineering, from Bu-Ali Sina University University , hamedan, Iran, in 2023. He is currently a Phd student in the Department of Electrical Engineering at Bu-Ali Sina University. His research interests include the modeling and control of power systems and microgrids.

Saleh Razini is a lecturer in the department of electrical engineering, Bu-Ali Sina university, Hamedan, Iran. He received the B.Sc. Degree in power electrical engineering, from Bu-Ali Sina University, in 2008, the M.Sc. degree in power systems and high voltage from K.N.Toosi university of technology, Tehran, in 2011. and Ph.D at Bu-Ali Sina university in 2017. His main research interests are power system planning and management, energy policy, and renewable energies.

MohamadAmin Ghasemi received his B.Sc. degree in Electrical Engineering from the University of Tehran, Tehran, Iran, in 2008, and his M.Sc. and Ph.D. degrees, also in Electrical Engineering, from Sharif University of Technology (SUT), Tehran, Iran, in 2010 and 2016, respectively. He is currently an Assistant Professor in the Department of Electrical Engineering at Bu-Ali Sina University. His research interests include the modeling and control of power systems, microgrids, and inverter-based renewable energy systems.