

Research on Real-Time Operation Optimization of the Urban Rail Traction Power Supply System Integrated Storage and Feedback

Haotian Deng, Wei Liu, Jun Dai

School of Electrical Engineering, Southwest Jiaotong University, 611756, China

*Correspondence Author, 2805459734@qq.com

Abstract: Urban rail transit plays an important role in the rapid economic and social development of China. The integrated storage and feedback urban rail traction power supply system is one of the strategies for the green and low-carbon development of urban rail transit. This paper establishes a real-time operation optimization model based on the Markov decision process in the context of the integrated storage and feedback urban rail traction power supply system. It uses offline training of a deep reinforcement learning agent to optimize the control parameters of the energy feedback systems and energy storage systems in real-time to reduce the traction energy consumption of the power supply system, providing a reference for energy-saving operation of subways.

Keywords: Urban rail transit, Power supply system, Energy feedback system, Energy storage system, Real-time operation optimization.

1. Introduction

Urban rail transit plays a fundamental, pioneering, strategic, and service-oriented role in the rapid economic and social development [1]. Due to the short distances between stations and the frequent starting and stopping of trains, a significant amount of regenerative braking energy is generated [2]. The integrated storage and feedback urban rail traction power supply system can absorb part of this regenerative braking energy through the energy storage system, and the energy feedback system, which is relatively inexpensive, can reduce construction costs. This system is one of the strategies for the green and low-carbon development of urban rail transit.

However, traditional integrated storage and feedback urban rail traction power supply system face two main issues: (1) When the energy feedback by the energy feedback system cannot be fully absorbed by the station's step-down load, will cause the main substation power back-feeding, which impacts energy-saving effectiveness and may cause disturbances to the external grid; (2) The low efficiency of energy storage system and the competition for regenerative braking energy from neighboring trains hinder the system's energy-saving operation. Therefore, optimization is required.

Currently, most research by experts and scholars on the optimization of urban rail traction power supply system focuses on long-term scales. For instance, literature [3] proposes the control optimization of the energy feedback system's droop rate, which can reduce total cost by an additional 5.4% by optimizing the operational characteristics of the energy feedback system. Literature [4] presents a bi-level optimization model aimed at minimizing annual costs, optimizing the traction substations and energy feedback systems, and solving the problem with an improved particle swarm algorithm, resulting in a 3.78% reduction in project annual costs. Literature [5] establishes an energy management system for ground energy storage systems, which adapts the device's charging and discharging voltage thresholds based on train schedules and traction network no-load voltage, with field tests demonstrating significant improvements in the

system's energy-saving performance. However, these systems operate independently based on preset parameters, leading to weak coordination control. When train schedules change or neighboring devices fail, dynamic adjustments cannot be made.

Due to the stringent computational time requirements for real-time operation optimization, traditional artificial intelligence algorithms struggle to find optimal solutions at real-time scales. Therefore, this paper employs the TD3 algorithm [6] in deep reinforcement learning to train an agent, allowing the agent to process information and issue control commands in order to achieve real-time operation optimization of the integrated energy feedback urban rail traction power supply system.

2. Integrated Storage and Feedback Urban Rail Traction Power Supply System

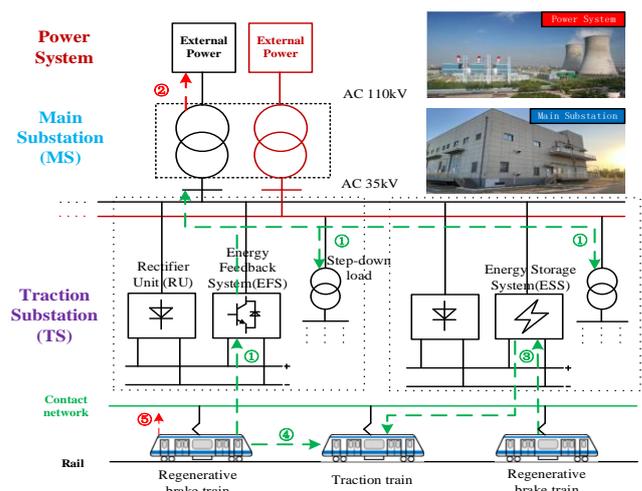


Figure 1: Integrated storage and Feedback Urban Rail Traction Power Supply System

The schematic of the integrated storage and feedback urban rail traction power supply system is shown in Figure 1. Compared to traditional power supply systems, the energy flow is more complex. The distribution of regenerative

braking energy from the trains is as follows: 1) It is feedback to the AC network through the energy feedback system and is used by step-down loads such as escalators, lighting, and air conditioning; 2) Some of the feedback energy is sent back to the external power supply through the main substation, and this energy cannot be utilized by the traction power supply system; 3) The energy storage system can store regenerative braking energy for use by traction trains, achieving a "peak shaving and valley filling" effect; 4) Neighboring traction trains directly use the regenerative braking energy from the braking trains; 5) Onboard braking resistors consume regenerative braking energy.

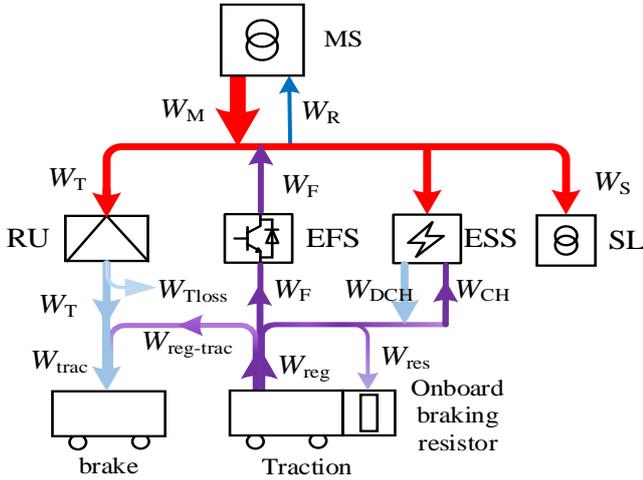


Figure 2: Energy Flow Diagram

The energy consumption evaluation indicators are shown in Figure 2. system energy consumption is generally evaluated from two perspectives: The energy consumption of the main substation, denoted as W_M ; The energy consumption of the main substation excluding the step-down load energy consumption, denoted as $W_M - W_S$ [7]. We select the latter and define the system's traction energy consumption W_{TR} as the energy consumption evaluation indicator. The derivation is as follows:

$$W_{TR} = W_T - W_F + W_R \quad (1)$$

3. Markov Decision Process-Based Model Construction

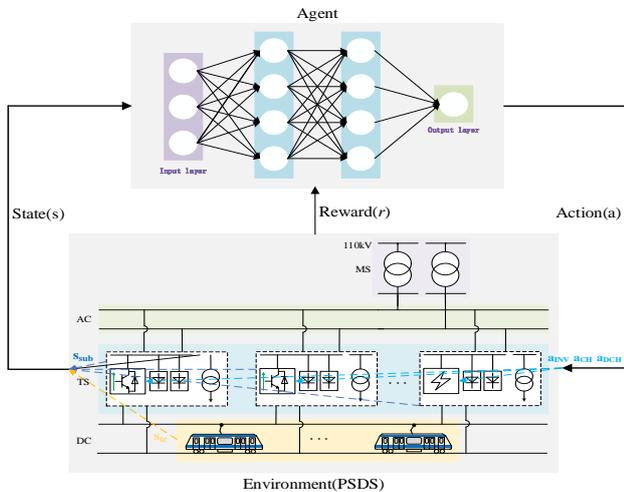


Figure 3: MDP Model Schematic

A schematic of the MDP model is shown in Figure 3.

3.1 Environment

The real-world environment of urban rail transit involves many uncontrollable factors. Considering the safety of rail transit operation, it is extremely difficult for the agent to perform dynamic interactions and online training in the real environment. Therefore, a simulation platform (PSDS), corrected with real-world measured data, is used as the simulation environment for the agent's dynamic interaction and offline training.

3.2 State Space

The state space must accurately describe the environmental characteristics of the time-varying traction power supply system:

(1) Traction Substation State:

The state of each traction substation includes operational statuses such as voltage, power, and the operation of the regenerative braking energy utilization device installed at the substation. The operational state of each traction substation is represented by the traction network voltage and DC-side output power.

The energy feedback system and rectifier unit do not start simultaneously, so the operational state of the energy feedback system does not need to be considered separately. When the output power of the traction substation is negative, it indicates that the rectifier unit is shut down, and the energy feedback system is providing power. The traction substations with energy feedback systems have two state variables.

When the energy storage system is charging, the rectifier unit does not start, and these two operations are mutually exclusive. However, when the energy storage system is discharging, the rectifier unit will activate the traction train. The charge state of the energy storage system directly impacts its operational state. Therefore, the traction substations with energy storage systems, in addition to the traction network voltage and DC-side output power, also need to provide the power and charge state of the energy storage system. The traction substations with energy storage systems have four state variables.

The traction substation state s_{sub} can be expressed as:

$$s_{sub} = [s_{sub,EFS} \ s_{sub,ESS}] \quad (2)$$

$$s_{sub,EFS} = [U_{TS,1} P_{TS,1} U_{TS,2} P_{TS,2} \cdots U_{TS,i} P_{TS,i} \cdots U_{TS,N_{EFS}} P_{TS,N_{EFS}}] \quad (3)$$

$$s_{sub,ESS} = [U_{TS,1} P_{TS,1} P_{ESS,1} SOC_1 \cdots U_{TS,i} P_{TS,i} P_{ESS,i} SOC_i \cdots U_{TS,N_{EFS}} P_{TS,N_{EFS}} P_{ESS,N_{ESS}} SOC_{N_{ESS}}] \quad (4)$$

Where, $U_{TS,i}$ represents the traction network voltage of the i th traction substation; $P_{TS,i}$ represents the DC-side output power of the i th traction substation; N_{EFS} represents the number of traction substations equipped with energy feedback systems; $P_{ESS,i}$ represents the power of the energy storage system at the i th traction substation; SOC_i represents the charge state of the energy storage system at the i th traction substation; and N_{ESS}

is the number of traction substations equipped with energy storage systems.

(2) Train State:

At different operational times, the number of trains in operation on the line may vary based on the departure interval. If the operation information of a single train is used as the state input, it would result in an uncertain number of state nodes in the neural network input, which is not conducive to stable training of the DRL (Deep Reinforcement Learning) model. Therefore, the total traction power and total regenerative braking power of the trains in the traction substation intervals of the line are used as the state input for the trains, as shown in the equation (5). At each time, the number of traction substation intervals is a fixed value, ensuring a fixed state space dimension.

$$\begin{aligned} P_{trac,i,j} &= \sum_{k=0}^{N_{tr,i,j}} P_{tr,k} P_{tr,k} > 0 \\ P_{RBE,i,j} &= \sum_{k=0}^{N_{tr,i,j}} P_{tr,k} P_{tr,k} < 0 \end{aligned} \quad (5)$$

Where, $P_{trac,i,j}$ represents the total traction power of the train between traction substation i and traction substation j . $P_{RBE,i,j}$ represents the total regenerative braking power of the train between traction substation i and traction substation j . $P_{tr,k}$ represents the power of the k th train. $N_{tr,i,j}$ represents the number of trains between traction substation i and traction substation j .

The train state s_{tr} can be expressed as:

$$s_{tr} = [P_{trac,1,2} P_{RBE,1,2} P_{trac,2,3} P_{RBE,2,3} \cdots P_{trac,i,j} P_{RBE,i,j} \cdots P_{trac,N_{TS}-1,N_{TS}} P_{RBE,N_{TS}-1,N_{TS}}] \quad (6)$$

3.3 Action Space

The action space consists of three types of encoding: The startup voltage of the energy feedback system. The charging voltage of the energy storage system. The discharging voltage of the energy storage system. Since the number of traction substations equipped with energy feedback systems differs from the number of traction substations equipped with energy storage systems, the action space is represented as a one-dimensional vector. These three action types can be expressed as a_{INV} , a_{CH} , a_{DCH} :

$$\begin{aligned} a_{INV} &= [k_{inv,1} k_{inv,2} \cdots k_{inv,i} \cdots k_{inv,N_{EFS}}] \\ a_{CH} &= [k_{ch,1} k_{ch,2} \cdots k_{ch,i} \cdots k_{ch,N_{ESS}}] \\ a_{DCH} &= [k_{dch,1} k_{dch,2} \cdots k_{dch,i} \cdots k_{dch,N_{ESS}}] \end{aligned} \quad (7)$$

Where, $k_{inv,i}$ represents the startup voltage coefficient of the i th energy feedback system; $k_{ch,i}$ and $k_{dch,i}$ represent the charging and discharging voltage coefficients of the i th energy storage system, respectively. $k_{inv,i}$, $k_{ch,i}$, $k_{dch,i} \in (-1,1)$.

The action space is re-normalized as shown in the equation:

$$\begin{aligned} U_{inv,i} &= (U_{br} + U_{d0})/2 + k_{inv,i} \times [(U_{br} - U_{d0})/2] \\ U_{ch,i} &= (U_{br} + U_{d0})/2 + k_{ch,i} \times [(U_{br} - U_{d0})/2] \\ U_{dch,i} &= (U_{d0} + U_{Tmin})/2 + k_{dch,i} [(U_{d0} - U_{Tmin})/2]_{Tmin} \end{aligned} \quad (8)$$

Where, U_{br} represents the onboard braking resistor's startup

voltage; U_{d0} represents the no-load voltage of the rectifier unit.

3.4 Reward Function

The agent's goal is formally represented as a special signal called a reward, which is transmitted from the environment to the agent. In the urban rail traction power supply system, the states of the traction substations at different times are independent. The reward function r in this paper is defined based on the energy consumption evaluation indicator, as shown in the equation (9). R_p is used for constant optimization effects and R_n is used to apply penalties when constraints are not satisfied.

$$r = R_p - R_n \quad (9)$$

$$R_p = (P_{TR}^{ref} - P_{TR})/1000 \quad (10)$$

Where, represents the reference system's current step traction power, and P_{TR} represents the agent's current step traction power.

When the control parameters issued by the agent at this step cause the traction network voltage or rail potential of the power supply system to exceed the limit, the system will feedback a large negative reward as a penalty to the agent, as expressed in the equation (11):

$$R_n = \begin{cases} 0 & \text{satisfy constraints} \\ R_{fail} & \text{dissatisfy constraints} \end{cases} \quad (11)$$

This design provides clear and comprehensive guidance for the agent to learn a reasonable strategy in the real-time operation optimization of the integrated energy feedback urban rail traction power supply system.

4. Case Study

4.1 Project Overview

To verify the effectiveness of the real-time operation optimization model for the integrated storage and feedback traction power supply system proposed in this paper, a flexible traction power supply simulation system was built using a real-world subway line as an example, and a deep reinforcement learning agent was trained offline.

The subway project's line topology is shown in Figure 4. The traction power supply system uses a centralized DC 1500V contact rail system. The total length of the line is 59.55 km, with 2 main substations (MS1 and MS2), 18 traction substations (TS1-TS18), and 7 step-down substations (SS1-SS7). Among the traction substations, there are two types of equipment installations: rectifier units (RU) with energy feedback systems (EFS) and rectifier units with energy storage systems (ESS). Except for substations TS7 and TS8, all other substations are equipped with stations. The power supply system parameters are shown in Table 1, and the operational line uses the "4B" train formation, with train parameters shown in Table 2. The full-day operating schedule is shown in Table 3.

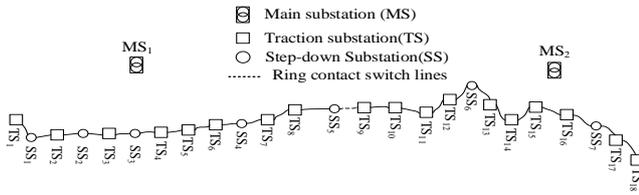


Figure 4: Subway Line Power Supply

Table 1: Power Supply System Parameters

Parameters	Value
Contact rail resistance (Ω/km)	0.0083
Rail resistance (Ω/km)	0.02
Rail-to-Earth Transition Resistance (Ω/km)	15
Rated power of rectifier unit ($\text{kV} \cdot \text{A}$)	1950
Capacity of energy storage system(MJ)	160
Maximum charging and discharging power /kW	3000
Charge/discharge efficiency/(%)	93
Maximum of SOC	1
Minimum of SOC	0.25
Rated power of energy feedback system/kW	3000
Rated capacity of step-down substation/($\text{kV} \cdot \text{A}$)	1250
Average load of step-down transformer /%	10(Spring)
Simulation step /s	1

Table 2: Train Parameters

Parameters	Value
Train formation	3MIT
Load	AW2
Weight/t	167.88
Rated voltage /V	1500
Maximum braking voltage /V	1800
Pollination method	contact rail
Auxiliary power /kW	46(Spring)
Structural speed /(km/h)	120

Table 3: Full-Day Operation Schedule

Departure interval (s)	Operating hours	Running time (h)
600	5:00-6:00, 21:00-23:00	3
240	6:00-7:00, 9:00-13:00, 16:00-17:00, 19:00-21:00	8
120	7:00-9:00	2
300	13:00-16:00	3

4.2 Operation Optimization Results Analysis

Among the 18 traction substations on the line, 12 substations are equipped with energy feedback systems, and 6 substations are equipped with energy storage systems. In the reference system, the charging voltage of energy storage system is 1750V, the discharging voltage is 1600V, and the startup voltage of energy feedback system is 1750V.

Using the full-day operating schedule from Table 3, the agent was trained with varying train departure intervals as input conditions. The agent's iterative process is shown in Figure 5, where the light blue curve represents the cumulative reward at each iteration, and the dark blue curve represents the average reward.

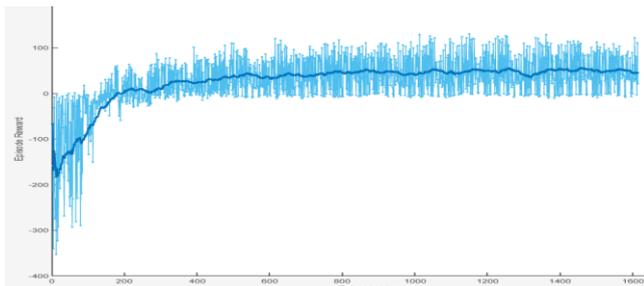


Figure 5: Iterative Process

From Figure 5, it can be seen that the agent stabilizes after 800 iterations, with an average reward of 50.6057. The reward varies significantly with different departure intervals, indicating that the arrangement of trains differs under various departure intervals, leading to substantial differences in energy utilization between trains and different energy-saving potential. The agent increases the regenerative braking energy utilization rate and reduces the traction energy consumption of the power supply system by changing the startup voltage of the regenerative braking energy utilization device, thereby influencing the energy flow on the line every second.

Taking a 240-second departure interval as an example, the hourly energy averages in the reference system and the optimized system by the agent are shown in Table 4.

Table 4: Hourly energy averages

Case	W_{TR} /kWh	W_{res} /kWh	W_F /kWh	W_R /kWh	W_{CH} /kWh	W_{DCH} /kWh
Reference	9206	354	370	144	578	500
Optimization	8845	29	671	47	1022	884

Compared to the reference system, WTR in the optimized system is reduced by 361 kWh, achieving an energy-saving rate of 3.92%. The energy that was originally consumed by the onboard braking resistors is now utilized by the energy feedback system and energy storage systems after optimization. Additionally, the real-time optimization system changes the energy flow on the line every second. As the amount of feedback energy increases, the power sent back by the main substation actually decreases, significantly improving the utilization of regenerative braking energy and reducing the traction energy consumption of the power supply system.

At a 240-second departure interval, the energy distribution at the 217th second of the line is shown in Table 5. During this second, multiple trains on the line are braking. In the reference system, the regenerative braking energy is primarily feedback to the AC side by the energy feedback system, but the step-down load cannot absorb the energy, resulting in a large amount of power being sent back to the main substation. This not only wastes regenerative braking energy but also impacts the external grid. In the optimized system, the agent lowers the charging voltage of the energy storage system and increases the EFS's startup voltage, allowing most of the regenerative braking energy to be absorbed by the energy storage system for use by subsequent traction trains. This significantly reduces the feedback power, and as a result, the power sent back to the main substation is also reduced.

Table 5: 217th second energy distribution

Case	P_{TR} /kW	P_{res} /kW	P_F /kW	P_R /kW	P_{CH} /kW	P_{DCH} /kW
Reference	-1543	0	5754	4211	454	0
Optimization	-1065	0	1708	643	4479	0

Using the full-day operating schedule from Table 3, the agent optimized the power supply system's control parameters for the entire day's operation. The hourly traction energy consumption of the power supply system under different departure intervals is shown in Table 6. The total traction energy consumption of the reference system for the entire day is 171 MWh, while the total traction energy consumption of the optimized system is 166 MWh. The system's traction

energy consumption is reduced by 4994 kWh, achieving an energy-saving rate of 2.92%.

Table 6: Comparison of full-day energy consumption

W_{TR}/kWh	120	150s	240s	300s	600s
Reference	17295	14307	9206	7416	3965
Optimization	17215	13984	8845	7172	3776

5. Conclusion

This paper presents an integrated energy feedback urban rail traction power supply system as the operating environment, with traction substation and train operation data serving as states, and the control parameters of regenerative braking energy utilization devices as actions. The traction energy consumption of the power supply system is used to define the reward function. A real-time operation optimization model based on the Markov decision process is developed, and a deep reinforcement learning (DRL) agent is employed to optimize the control parameters of the regenerative braking energy utilization devices in real-time. The proposed system achieves a 2.92% reduction in traction energy consumption over the course of a full day's operation.

References

- [1] Shang M, Zhou Y, Mei Y, et al. Energy-Saving Train Operation Synergy Based on Multi-Agent Deep Reinforcement Learning on Spark Cloud [J]. IEEE Transactions on Vehicular Technology, 2022, PP: 1-13.
- [2] Wei L, Jian Z, Hui W, Tuojian W, Ying L, Xiaowen Y, et al. Modified AC/DC Unified Power Flow and Energy-Saving Evaluation for Urban Rail Power Supply System With Energy Feedback Systems [J], IEEE Transactions on Vehicular Technology, 2021, 70(10): 9898-9909.
- [3] Zhang G, Tian Z, Tricoli P, et al. Inverter Operating Characteristics Optimization for DC Traction Power Supply Systems [J]. IEEE Transactions on Vehicular Technology, 2019, 68(4): 3400-10.
- [4] Qian X, Wei L, Qianfeng Y, Xiaodong Z, Haohao G, Haotian D, Peng X, et al. Bi-level optimal design for DC traction power supply system with reversible substations [J], IEEE Transactions on Transportation Electrification, 2023, PP(99): 1-1.
- [5] Liu Y, Yang Z, Wu X, et al. Adaptive Threshold Adjustment Strategy Based on Fuzzy Logic Control for Ground Energy Storage System in Urban Rail Transit [J]. IEEE Transactions on Vehicular Technology, 2021, 70(10): 9945-56.
- [6] Fujimoto S, Hoof H v, Meger D. Addressing Function Approximation Error in Actor-Critic Methods; proceedings of the International Conference on Machine Learning, F, 2018 [C].
- [7] Zhang J, Liu W, Tian Z, et al. Urban Rail Substation Parameter Optimization by Energy Audit and Modified Salp Swarm Algorithm [J]. IEEE Transactions on Power Delivery, 2022, 37(6): 4968-78.