JOURNAL OF NANO- AND ELECTRONIC PHYSICS Vol. **15** No 3, 03004(6pp) (2023)
ЖУРНАЛ НАНО- ТА ЕЛЕКТРОННОЇ ФІЗИКИ Том **15** № 3, 03004(6cc) (2023)

Energy Management in Electric Vehicles Using Improved Swarm Optimized Deep Reinforcement Learning Algorithm

M.A. Jawale[1,*], A.B. Pawar[2], Sachin K. Korde[3], Dhananjay S. Rakshe[4], P. William[3], Neeta Deshpande[5]

[1] *Department of Information Technology, Sanjivani College of Engineering, SPPU, Pune, India*
[2] *Department of Computer Engineering, Sanjivani College of Engineering, SPPU, Pune, India*
[3] *Department of Information Technology, Pravara Rural Engineering College, SPPU, Pune, India*
[4] *Department of Computer Engineering, Pravara Rural Engineering College, SPPU, Pune, India*
[5] *Department of Computer Engineering, R H Sapat College of Engineering, Management Studies and Research, SPPU, Pune, India*

(Received 20 May 2023; revised manuscript received 16 June 2023; published online 30 June 2023)

The internal combustion engine-based transportation system is causing severe problems such as rising levels of pollution, rising petroleum prices, and the depletion of natural resources. To divide power between the engine and the battery in an effective manner, a sophisticated energy management system is required to be put into place. A power split strategy that is efficient may result in higher fuel economy and performance of Electric Vehicles (EVs). In this paper, we propose the reinforcement learning method using Deep Q learning (DQL), which is a novel Improved Swarm optimized Deep Reinforcement Learning Algorithm (IS-DRLA) designed for energy management control. To perform an update on the weights of the neural network, this method computes the use of a modified version of the swarm optimization technique. After that, the suggested IS-DRLA system goes through training and verification using high-precision realistic driving conditions, after which it is contrasted with the standard approach. The performance indices such as State of Charge (SOC) and fuel consumption and loss function are analyzed for the efficiency of the proposed method (IS-DRLA). According to the findings, the newly proposed IS-DRLA is capable of achieving a higher training pace with a lower overall fuel consumption than the conventional policy, and its fuel economy comes very close to matching that of the worldwide optimal. In addition to this, the adaptability of the suggested strategy is demonstrated by utilizing a different driving schedule.

Keywords: Energy management, Electric vehicles, Improved Swarm optimized Deep Reinforcement Learning Algorithm (IS-DRLA).

DOI: 10.21272/jnep.15(3).03004 PACS number: 88.85.Hj

1. INTRODUCTION

Electric vehicles are environmentally friendly, green, sustainable, clean, and efficient. HEVs employ batteries to store electrical energy that is used to power the motor and small engine. It may provide higher fuel economy while emitting less harmful pollutants [1]. The HEV's two energy sources, “the engine and battery,” must be regulated in such a way that they operate within their efficient ranges while meeting the driver's power requirement [2]. As a result, an intelligent control approach that can efficiently split power between the engine and battery is necessary. There are several architectures available, including series, parallel, and power split. There is no mechanical connection between an electric motor and an engine in series architecture. The engine charges the battery via the generator, and the battery then powers the motor, which drives the wheels. For the car to operate properly, series hybrids require two independent energy conversion mechanisms. Parallel hybrids enable both power sources to operate concurrently for optimal performance [3]. The transmission and drive train, on the other hand, are more complex and expensive components. The parallel arrangement is more difficult to implement than the series set up, but it offers numerous benefits not available with the series configuration. The power split arrangement was designed [4] as a means of compensating for the shortcomings that are inherent in both the series and parallel systems. It makes use of both electrical and mechanical power couplers in its configurations. In this configuration, the engine and battery can power the vehicle either independently or together, and the engine also can charge the battery at the same time.

The remaining part of this research is organized into 5 sections, section 2: literature survey, section 3: Problem Formulation of Energy Management, section 4: Proposed approach, section 5: Results from simulations and discussions on the proposed approach, and section 6: conclusion.

2. LITERATURE SURVEY

The scope of research [6] was to propose a hierarchical “Deep Reinforcement Learning (DRL)” approach for scheduling the energy consumption of “Distributed Energy

* jawale.madhu@gmail.com

Resources (DERs)", such as "Energy Storage Systems (ESS) and electric vehicles (EV)". Study [7], conditional probability and Markov chains are used to simulate the PEV departure time and necessary energy consumption while driving. Additionally, the PVWatt model and the "Adaptive Neuro-Fuzzy Inference System (ANFIS)" each model the performance of the PV and the need for the home load. Then, by avoiding needless charging/discharging schemes, a "Model Predictive Control (MPC)" is made to reduce the cost of energy as well as make the PEV battery last longer. To get over battery constraints, this research [8] suggests a "hybrid EMS for the series-parallel PHEV" that makes use of a rule-based control method and a "Genetic Algorithm (GA)"-based optimization technique. The goal of the research presented is to suggest a new technique for recognizing driving behavior for a "P1-P2 series parallel electric vehicle utilizing a Long Short-Term Memory Recurrent Neural Network (LSTM RNN) in conjunction with an Energy Consumption Minimization Strategy (ECMS)". By concentrating on effective driving mode transition rather than single mode optimization, novelty is attained. Computational and experimental verification of the proposed approach to assess power-sharing superiority based on battery SOC is described, demonstrating the results. The proposed method has been discussed and compared to the standard DTC method, and now it has been experimentally validated.

3. PROBLEM FORMULATION OF ENERGY MANAGEMENT

Enhancing the electric vehicle's performance and efficiency through energy management is the primary focus of this strategy. "As described below, the cost function, a trade-off between fuel economy and sustainable electric quantity, is employed as the evaluation indicator to measure the efficacy of the energy management technique":

$$J = \int_{t_0}^{t} \left\{ \alpha \cdot \dot{f}_{rate}(t) + \beta \cdot [SOC(t) - SOC(t_0)]^2 \right\} dt \qquad (1)$$

The instantaneous fuel consumption rate across the period $[t_0, t]$ is represented by the term "f_{rate}," where "α" and "β" are both positive weighting factors. Additionally, the inequality constraints should be satisfied in the manner described below to guarantee dependability and safety:

$$SOC_{min} \le SOV(t) \le SOC_{max} \qquad (2)$$

$$P_{b,min} \le P_b(t) \le P_{b,max} \qquad (3)$$

$$n_{g,min} \le n_g(t) \le n_{g,max} \qquad (4)$$

$$I_{g,min} \le I_g(t) \le I_{g,max} \qquad (5)$$

$$T_{eng,min} \le T_{eng}(t) \le T_{eng,max} \qquad (6)$$

4. PROPOSED APPROACH

In this study we propose the reinforcement learning method using Deep Q learning (DQL), a novel Improved Swarm optimized Deep Reinforcement Learning Algorithm (IS-DRLA) intended for energy management control, and it employs an improved swarm optimization algorithm to update the neural network weights.

4.1 Vehicle Model

A moving vehicle experiences a variety of forces, such as rolling resistance, slope resistance, and aerodynamic drag, which all work against the movement of the vehicle and lower its speed. Theorems acting on the car are usually denoted by equation (7).

$$F_r = \frac{1}{2}\rho A_f C_D (V - V_W)^2 + Pf_r + Mgsin\alpha \qquad (7)$$

where P is the force acting on the center of a stationary tyre, f_r is the rolling resistance, α an is the road angle, A_f is the frontal area of the vehicle, the C_D is the aerodynamic drag of the vehicle's body, p is the density of the air, V is the speed of the vehicle, and V_w is the wind speed component in the direction in which the vehicle is moving. The wheels of a hybrid electric vehicle (HEV) are driven by a motor, charged by the generator, which is part of the hybrid electric vehicle's (HEV) planetary gear system (PGS). The PGS has a carrier, a sun, a ring gear, and several pinion gears. The generator is linked to the sun, the engine to the carrier, and the final drive to the ring gear. The governing equations between the various gear ratios and circle diameters are written as equation (8)

$$\omega_r * r_r = -\omega_s * r_s + \omega_c (r_s + r_r) \qquad (8)$$

m_r, m_s, m_c are the radii of the ring and the sun, and the angular velocity of the carrier, respectively. Torques exerted on the sun, r_r, r_s ring, and carrier can be approximated using the equation (9)

$$T_c = -k_{ys} T_s = -k_{yr} T_r \qquad (9)$$

The torque acting on the carrier, sun, and ring gears are denoted by $T_c, Ts, and\ T_r$ respectively; $k_{yr} = (1 + i_g)/i_g$ and $k_{ys} = (1 + i_g)$ ang i_g is the gear ratio. At a given velocity, the relationship between the engine speed m_e, the motor speed m_m, and the generator speed m_g is given by equation (10).

$$\frac{N_r}{N_s + N_r} * \omega_m + \frac{N_s}{N_s + N_r} * \omega_g = \omega_e \qquad (10)$$

In a Toyota Prius, the tooth numbers N_r and N_s refer to the ring and sun gears, respectively. Prius's $N_r = 78$, and $N_s = 30$.

4.2 Objective Purpose Formula and Restrictions

Here, the objective function provided in equation (11) is applied to optimize HEV performance.

$$J = m_{ft} \quad (11)$$

Where m_{ft} is the overall fuel consumption rate for a certain time interval. The fuel efficiency is expressed by the equation, where P_b is the power drawn from the batteries equation (12)

$$P_b = V_{OC} * I_b - I_b^2 * R_b \quad (12)$$

V_{oc} – open-circuit voltage, R_b – battery resistance, and I_b – battery current. The time rate of SOC can be represented by the equation (13): SOC. (14)

$$SOC = -\frac{I_b}{Q_b} \quad (13)$$

$$SOC = \frac{V_{oc}\sqrt{{V_{oc}}^2 - 4P_b R_b}}{2P_b Q_b} \quad (14)$$

Quantity of batteries denoted Q_b. The equation below describes the relationship between the MG1/MG2 ratio, the engine, and the desired torques and speeds equation (15).

$$\left.\begin{array}{c} T_{M/G1} = -\frac{1}{1+R}[T_e] \\ \omega_{M/G1} = -R\zeta\omega_{req} + (1+R)\omega_e \\ T_{M/G2} = -\frac{1}{(1+R)}\left[\frac{(1+R)T_{req}}{\zeta} + RT_e\right] \\ \omega_{M/G2} = \zeta\omega_{req} \end{array}\right\} \quad (15)$$

It is important to take into account the constraints listed in eq. while numerically calculating the objective function for the EV equation (16).

$$\left\{\begin{array}{c} \omega_{e,min} \le \omega_e \le \omega_{e,max} \\ \omega_{mg1,min} \le \omega_{mg1} \le \omega_{mg1,max} \\ \omega_{mg2,min} \le \omega_{mg2} \le \omega_{mg2,max} \\ T_{2,min \le T_e \le T_{e,max}} \\ \omega_{mg1,min} \le \omega_{mg1} \le \omega_{mg1,max} \\ \omega_{mg2,min} \le \omega_{mg2} \le \omega_{mg2,max} \\ SOC_{min} \le SOC \le SOC_{max} \end{array}\right. \quad (16)$$

where $\omega_{e,min}$, $\omega_{e,max}$, $\omega_{mg_1,min}$, $\omega_{mg_1,max}$, $\omega_{mg_2,min}$, $\omega_{mg_2,max}$, $T_{e,min}$, $T_{e,max}$, $T_{mg_1,min}$, $T_{mg_1,max}$, $T_{mg_2,min}$, $T_{mg_2,\alpha x}$, $SOCmin$, and $SOCmax$ are the lowest and highest numbers for the engine's speed, MG_1, and MG_2's torque, and the SOC's limiting range, correspondingly.

4.3 Energy Management Strategy Based on Improved Swarm Optimized Deep Reinforcement Learning Algorithm (IS-DRLA)

To create a very effective method, optimization using particle swarm only requires a few elementary steps. To arrive at an optimal solution, PSO first selects a set of parameters that are statistically likely to yield success and then multiplies those numbers by a standard random term. The algorithm avoids the problem of impulsive convergence, which plagues most search procedures. Particles created at the outset of PSO continue to exist until a solution is located. Two primary aspects control the particle motion: 1) a global best result between particles and 2) an iteration-to-iteration best solution between particles. After gathering data from successive iterations, the "pbest" is the best answer the algorithm has considered so far. The best results visited by any particle are stored in a variable named gbest, which is passed from particle to particle. The cognitive and social elements, denoted by "pbest" and "gbest," are responsible for optimal performance. If a new, superior solution is identified at each iteration, the "pbest" and "gbest" values for each particle are updated. With each iteration, the process of identifying "pbest" and "gbest" continues until either the required outcomes are achieved or it is determined that no further viable solutions can be discovered within the application search area. Both the speed and the distance a particle travels are determined by its velocity. Each particle in the D-dimensional space is described by the PSO as $X_i = (x_{i1}, x_{i2},\ldots, x_{iD})$, where i denote the particle number and x_{iD} is the number of parameters used to define the solution. A velocity in each dimension is calculated separately using the formula $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$, whereas the position in space is given by $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$. At the end of each cycle, the particle's best possible location (pbest) is compared to the best possible position (gbest) throughout the whole simulation, and the particle is moved in a random direction toward pbest or gbest. Velocity is updated as given by equation (17).

$$V_{id}^{(t+1)} = \omega * V_{id}^{(t+1)} + U[0,1] * \psi_1 * \left(p_{id}^{(t)} - x_{id}^{(t)}\right) + U[0,1] * \psi_2 * \left(p_{gd}^{(t)} - x_{id}^{(t)}\right) \quad (17)$$

The position is updated using this velocity (t + 1) and given as equation (18),

$$X_{id}^{(t+1)} = X_{id}^{(t)} + V_{id}^{(t+1)} \quad (18)$$

In this equation, $U[0, 1]$ represents a uniform random distribution, t is a time index, and ψ_1 and ψ_2 represent the respective weights of the local best and global best solutions' trade-offs and implications on the particle's total velocity.

4.4 Characteristics of Deep Reinforcement Learning

Proposed techniques for reinforcement learning that make use of discontinuous system parameters are one-step Q-learning and enhanced swarm. That once concentration of the separate points approaches the cutoff, there is a chance that the "curse of dimensionality" issue will manifest. The value function in the suggested algorithm is built using a neural network that performs a comparable purpose. A Q network receives state variables as inputs and directly generates control actions. The only difference between the proposed algorithm's settings for state variables, control actions, and rewards is that the DQL approach's states are constant.

There are connecting different layers that make up the neural net: the input layer, the hidden layers, and the output layer. Each deep learning model has 32 neurons, and the number of layers for input and output can be adjusted by system parameters and other control methods. Using an evaluation network and a target network, both of which are neural networks, the optimal control method can be determined. The Q-value used for evaluation is the result of Input 1, whereas the Q-value needed for success is the result of Input 2. The loss function can be written in terms of the mean-variance expectation of Q-values as follows:

$$L(w) = \left[E(r + \gamma \, {}^{max}_{a'} Q(s', a', w') - Q(S, A, W))^2\right] \quad (19)$$

Where $r + \gamma \, {}^{max}_{a'}$ the goal Q-value is represented by $Q(s', a', w')$, and the evaluation Q-value is represented by $Q(S, A, W)$. The weights of the evaluation network and the target network, respectively, are denoted by the letters w andw', and optimizing weights results in the optimization of the neural network. In the emergence, we indicated that earlier research in deep RL for EM generally used the gradient descent technique to change weight values using the minimization of the error function as the optimization problem. Given the strong inter-sample correlations, the training efficiency of the evaluation network is increased by randomly selecting batches of data from the replay buffer to be used for a fixed number of timesteps in the training procedure. The foregoing description of the training process's optimization strategy is in-depth. In addition, Table 1 introduces the DQL algorithm's primary parameters.

Table 1 – DQL algorithm's primary parameters

Main Parameters	Value
Sample batches size	64
Learning rate	0.0001
Replay buffer capacity	1200
Amount to be discounted	0.95
Aspect e of Initial Exploration	1.0

Algorithm 1 contains the DQL algorithm's pseudo-code.

Algorithm 1: Deep Q-Learning Algorithm
Randomly weight Q evaluation network
Initialize $\hat{Q}$ *with w = w*
Start replay buffer B with N
For occurrence $= 1$ *to M*
Reorganize early state $s_1 = (SOC^1 n_g^1, P_{dem}^1)$
For $t = 1$ *to T*
$a_t \leftarrow \varepsilon - greedy\ (s_t, Q)$
Execute a_t*; observe* s_{t+1} *and* r_t
Store the vector $(s_t,\ a_t, r_t, s_{t+1})$ *in reply buffer B*
Sample random batch of $(s_t,\ a_t, r_t, s_{t+1})$ *from B*
If terminal S_{j+1}
Set $Y_j = r_j$
Else
Set $Y_j = r_j + \gamma max_{\ a_{j+1}} \hat{Q}(S_{j+1}, a_{j+1}, w')$
End if
Calculate loss function $L(w) = E[(y_J - Q(s_j, a_j, w))^2]$
Performance optimization method improved swarm base on $L(w)$*.*
Rest $\hat{Q}$ *with weight w′ = w*
End for

First, initializing sets up two networks and the replay buffer. Then a nested loop is used, with the "inside loop designating the time-steps inside a training episode and the outside loop designating the training episodes". The inner loop is where control actions are carried out and neural networks are updated.

4. RESULTS AND DISCUSSIONS

In this research, we proposed the IS-DRLA is intended for energy management control, and it employs an improved swarm optimization algorithm to adjust the neural network values. The existing methods such as "Federated Reinforcement Learning (FRL)" [16] and "Hybrid Convolutional Neural Network- Long- and Short-Term Memory (HCNN-LSTM)" [17] are used to analyze the performance indices such as State of Charge (SOC) and fuel consumption and loss function.

Fig. 1 displays the fuel consumption curves over episodes for the proposed and existing approaches. Naturally, the proposed strategy's rate of convergence is significantly higher than that of existing techniques, and its fuel consumption is likewise lower than that of the latter. The outcome demonstrates that by removing unproductive control, the heuristic planning method can enhance the speed and effectiveness of the energy management strategy. Furthermore, using the algorithm for real-time control is made simpler by the method's need for fewer training episodes.

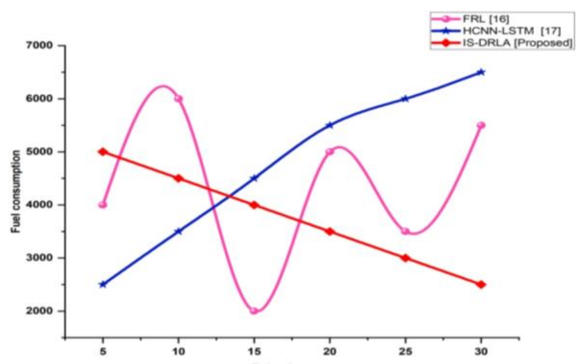

Fig. 1 – Fuel consumption

Fig. 2 shows the SOC trajectories for the three distinct control strategies. The SOC curves for each strategy show that they are all capable of keeping SOC levels relatively constant within the range of 0.61 to 0.72. SOC version of IS-DRLA is more similar to existing approaches, which suggests that "the proposed algorithm is closer to the global optimal solution". Heuristic planning, which omits undesirable control actions during training, is said to have had an impact on the trajectory differences between traditional methods.

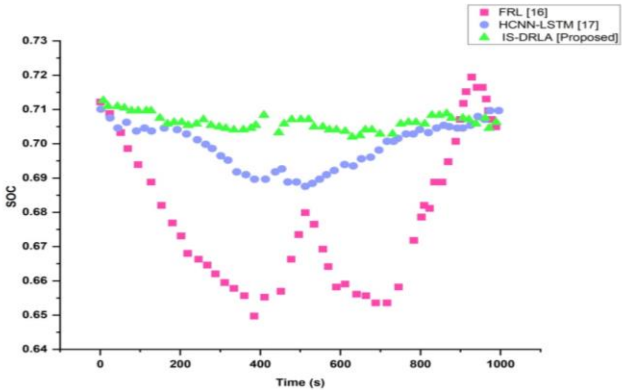

Fig. 2 – SOC trajectories for three different control techniques

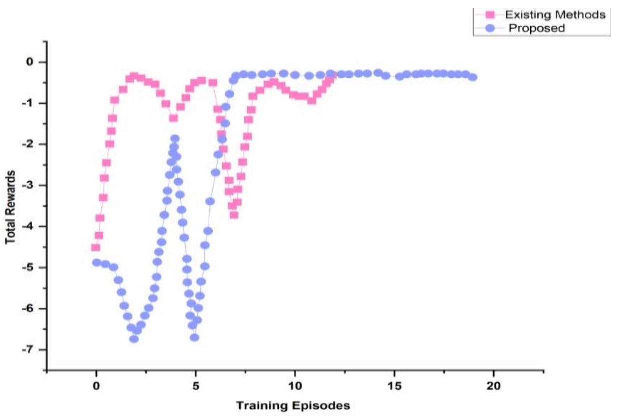

Fig. 3 – Overall reward trajectories for the proposed and existing approaches

In addition Fig. 3 depicts the total rewards' trajectories for the two methods, and "we can see that the reward for the IS-DRLA optimizer reaches its maximum value sooner", supporting its speed. Locally magnified views of the finest reward levels imply that the new approach saves energy and improves optimization.

The particular driving cycle that was employed to train the neural network validates the speed and optimality of the IS-DRLA. However, driving cycles in the actual world are unpredictable and changeable.

5. CONCLUSION

To better regulate energy consumption, we present the IS-DRLA, a neural network whose weights are updated via an enhanced swarm optimization technique. These strategies are designed to maximize efficiency. It is possible to conclude that the new variant of the technique, which goes by the proposed, performs significantly better than the traditional method when it comes to the quickness and optimality of solving the problem of energy management. The recommended DQL methodology with the upgraded swarm optimizer achieves faster learning speed and lower fuel usage than the normal DQL strategy and is near the optimum solution. Deep reinforcement learning may be adapted to various operating modes. The goal of the ongoing research is to develop a novel approach that can be applied to the problem of energy management, allowing for better sample selection and training. This is being done as a reaction to the foregoing analysis of the possible problem of insufficient samples collected during the planning portion of the suggested technique. In addition, future studies will concentrate on the architectural examination of neural network models in deep reinforcement training. The system's efficacy will be verified using both an equipment test setup and actual road-going automobiles.

REFERENCE

1. L. Hu, Q. Tian, C. Zou, J. Huang, Y. Ye, X. Wu, *Renew. Sustain. Energy Rev.* **162**, 112416 (2022).
2. U. Rehman, K. Yaqoob, M.A. Khan, *Int. J. Electrical Power Energy Syst.* **134**, 107358 (2022).
3. P. Xie, L. Jin, G. Qiao, C. Lin, C. Barreneche, Y. Ding, *Renew. Sustain. Energy Rev.* **160**, 112263 (2022).
4. Z. Dimitrova, W.B. Nader, *Energy* **239**, 121933 (2022).
5. A.G. Abo-Khalil, M.A. Abdelkareem, E.T. Sayed, H.M. Maghrabie, A. Radwan, H. Rezk, A.G. Olabi, *Int. J. Thermofluid.* **13**, 100134 (2022).
6. S. Lee, D.H. Choi, *Sensors* **20** No 7, 2157 (2020).
7. M. Yousefi, N. Kianpoor, A. Hajizadeh, M. Soltani, *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, 2201 (2019).
8. N. Ding, K. Prasad, T.T. Lie, *Int. J. Energy Res.* **45** No 2, 1627 (2021).

Управління енергією в електричних транспортних засобах з використанням вдосконаленого алгоритму глибокого підсилення

M.A. Jawale[1], A.B. Pawar[2], Sachin K. Korde[3], Dhananjay S. Rakshe[4], P. William[3], Neeta Deshpande[5]

[1] *Department of Information Technology, Sanjivani College of Engineering, SPPU, Pune, India*
[2] *Department of Computer Engineering, Sanjivani College of Engineering, SPPU, Pune, India*
[3] *Department of Information Technology, Pravara Rural Engineering College, SPPU, Pune, India*
[4] *Department of Computer Engineering, Pravara Rural Engineering College, SPPU, Pune, India*
[5] *Department of Computer Engineering, R H Sapat College of Engineering, Management Studies and Research, SPPU, Pune, India*

Транспортна система на основі двигуна внутрішнього згоряння створює серйозні проблеми, такі як зростання рівня забруднення атмосфери та виснаження природних ресурсів. Для ефективного розподілу енергії між двигуном і батареєю необхідна складна система управління енергією. Ефективна стратегія розподілу потужності може призвести до кращої економії палива та продуктивності електромобілів (EV). У статті ми пропонуємо метод навчання з підкріпленням з використанням глибокого навчання Q (DQL), який є новим алгоритмом з підкріпленням (IS-DRLA), оптимізованим для групи Improved Swarm, розробленим для контролю управління енергією. Щоб виконати оновлення вагових коефіцієнтів нейронної мережі, цей метод обчислює використання модифікованої версії методу оптимізації роя. Після цього запропонована система IS-DRLA проходить навчання та перевірку з використанням високоточних реалістичних умов водіння, після чого вона порівнюється зі стандартним підходом. Індекси продуктивності, такі як стан заряду (SOC) і функція витрат і втрат палива, аналізуються на ефективність запропонованого методу (IS-DRLA). Відповідно до висновків, нещодавно запропонований IS-DRLA здатний досягати вищого темпу навчання з нижчим загальним споживанням палива.

Ключові слова: Енергоменеджмент, Електричні транспортні засоби, Удосконалений алгоритм глибокого підсилення навчання (IS-DRLA).