Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Load Frequency Control of Re Integrated Smart Deregulated Power System Using Reinforcement Learning Based Controller

Nausheen Bano[1], T. Anil Kumar[2]
[1]*Research Scholar, Anurag University, Hyderabad.*
[2]*HOD, EEE Anurag University, Hyderabad.*

ARTICLE INFO

Received: 29 Dec 2024

Revised: 15 Feb 2025

Accepted: 24 Feb 2025

ABSTRACT

When there is a lag between generation and load in a power system, frequency deviation happens. The frequency deviation in the deregulated power system is caused by consumers switching to various DISCOs, which causes the load on DISCOs to fluctuate the frequency which causes the undesirable operation in the the power system. The EV aggregators are introduced in each control area to supply power to the DISCOs in the event of a contract violation. This paper presents a reinforcement learning controller for load frequency control of a Smart Deregulated Power System (SDPS) that consists of two control areas, each of which contains thermal, solar PV plants, and hydro, wind plants respectively. The superiority of reinforcement learning controller over model predictive and robust controllers is that the neural network is trained from the control and system parameters of the open access power system under various operating scenarios. The Reinforcement learning controller is called Actor-Critic agent based Deep Deterministic Policy Gradient (DDPG) controller tested on two area model of deregulated power system under different possible contract scenarios and under various operating conditions. The Actor-Cretic reinforcement learning approach for LFC compared to FOPI and PI controllers under different possible contract scenarios in smart deregulated environment.

Keywords: Control Area (CA); Load Frequency Control (LFC), Smart Deregulated Power System (SDPS), EV Aggregators, Reinforcement learning.

INTRODUCTION

The electric power system evolved over many decades, in this period, traditional networks evolved into sophisticated, reliable, efficient, and sustainable smart grids thanks to financial, technological, environmental, and political incentives [1-3]. Smart grids allow customers to actively participate in the electrical markets. The rapid advancements in vehicle-to-grid technology mean that plug-in electric vehicles, one type of distributed energy storage, are set to play a significant part in emergency reliability services [4-6]. Electric vehicles (EVs) are growing in popularity because these vehicles consume less gasoline and release fewer greenhouse emissions into the atmosphere [7]. The state of the environment affects the output power of renewable resources. The regulation of frequency in power systems is significantly impacted by large-scale variable generation. In order to balance generation and demand, the power system must account for the unpredictability introduced by renewable energy sources, which calls for the addition of spinning reserve [8].

Shifting from conventional to deregulated power systems in the early decades created additional uncertainty to power systems. A typical power system's generating, distribution as well as transmission components are held by a utility that is vertically integrated (VIU), which provides customers with regulated electricity rates. In a market of open energy, GENCO's may choose to take part in the LFC duty or not. Distribution companies (Discos) may enter into agreements with generators (GENCO's) or independent power producers (IPPs) to supply electricity in certain regions [9]. In a deregulated power system, frequency regulation becomes more complex. A fixed controller may not suffice the demand of deregulated system. In such cases, a robust controller significance increases for such a system.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The review of literature on LFC of deregulated power systems describes a number of different control strategies. The operating framework required for the deregulated power system's can be found in [10].In the deregulated electricity system, a controller based on ramp following has been studied [11]. For a deregulated power system or electrical grid, a decentralized neural network controller is available [12]. In the case of a deregulated power system, traditional integral controller improvements have been modified through genetic algorithm optimization [13]. It was developed to use a genetic algorithm to optimize the PID controller architecture in the context of deregulated power [14]. The LFC design of the deregulated electricity system makes use of both reduced order observer techniques and optimal output feedback control [15]. For two levels of independence, the electrical system that is deregulated is designed using the LFC method and the Internal Model Control technique [16]. The LFC of the deregulated electrical system's robustness analysis was conducted using the structured singular value approach [17]. The LFC consisting of multiple area thermal system that have a PID controller of fractional order has been described by the researcher [18].

Different trial and error methods and traditional tuning techniques like Ziegler-Nicholas are used to gradually modify PID controller gain values based on experience; nevertheless, these are approaches not perform better in a variety of operating circumstances and random load variations [19]. Over the years, different control strategies that are smart and optimized are utilized for LFC have been presented as research work by researchers. To adjust the PID parameters using fuzzy logic, the system that can be used is considered is an adaptive neural fuzzy inference system (ANFIS) [20-21]. However, due to its lack of adaptability, a fuzzy system necessitates ground experience to modify the membership functions and acquiring the necessary specialized knowledge can be difficult [22].

The fractional order (FO) controller has attracted the attention due to its enhanced capability to modify the dynamics of the system [23-26]. Because of its intrinsic flexibility, a number of researchers have recommended the FO controller for the power industry over the conventional PI and PID controller [27],[28].It has recently been found that reinforcement learning (RL) based control strategies are promising for the current grid. In [29], a critical review of the literature on reinforcement learning-based control of electric power systems is presented. Because reinforcement learning can learn on its own through interactive trial and error using data it gathers from the ever-changing environment, it performs better than standard control approaches. Therefore, reinforcement learning performs better while making decisions and solving control issues in the actual world. Numerous studies have been conducted in the literature that use reinforcement learning systems to control the frequency of an interconnected area. In the uninterrupted action domain, a new technique for controlling frequency via DRL is presented in [30]. However, due of the agents' parallel learning behavior, this kind of method is devoid of a continuous gradient signal [31].

The Deep Deterministic Policy Gradient (DDPG) does not call for the action division and states, which Lilliclip et al. [32] presented as a solution to continuous control problems. Overestimating Q-values is a drawback of DDPG, which is comparable to deep Q-networks (DQN), which update the Q-value and can lead to incremental bias and subpar policy [33-34]. Moreover, a given dataset may result in less-than-ideal convergence when load generation is continuously variable, since the authors [35] initially used to construct the PID controller to enable collect data for the agent's initialization.

In this present paper, the Deep Deterministic Policy Gradient (DDPG) optimization technique is used to train the Actor Critic agent as an LFC controller for frequency management in SDPS in different transaction Scenarios. The paper is drafted as follows: model of the deregulated power system discussed in detail in section 2. Whereas section 3. covers the Deep Deterministic Policy gradient technique for LFC of a SDPS. In section 4. discussed about simulation result and concussions are given in section-5.

MODELLING OF VARIOUS COMPONENTS OF SMART DEREGULATED POWER SYSTEM (SDPS)

This section presents the modelling of SDPS with integration of hydro, wind in CA-1and thermal, solar PV plants in CA-2 and EV Aggregators are integrated in each control area to supply power to the DISCOs under contract violation scenario. Any GENCOs can sell power to any DISCO at competitive pricing by utilizing "DISCO Participation matrix" (DPM), it's easy to comprehend the contract. The values in the rows (horizontal direction) and columns (vertical direction) of DISCO Participation matrix represent generating units and DISCOs respectively. The total number of the elements in the column is one. The entry in the column is zero, if the DISCO is not drawn any power [36-39].

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The DISCO Participation matrix (DPM) is given by

$$DPM = \begin{bmatrix} cpf11 & cpf12 & cpf13 & cpf14 \\ cpf21 & cpf22 & cpf23 & cpf24 \\ cpf31 & cpf32 & cpf33 & cpf34 \\ cpf41 & cpf42 & cpf43 & cpf44 \end{bmatrix}$$

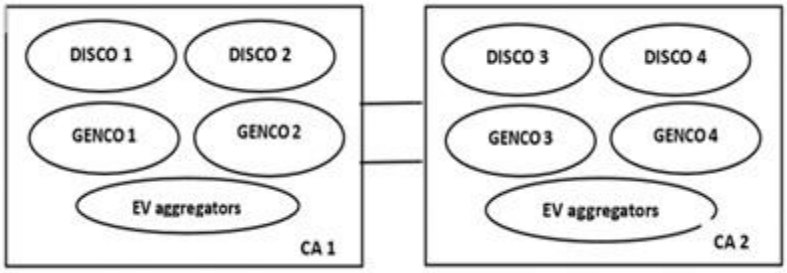


Figure 1. Two area system in deregulated power system

$$cpf_{ij} = \frac{\text{j}^{\text{th}}\ \text{DISCO'Spower demand out of i}^{\text{th}}\ \text{GENCO[p. u M. W]}}{\text{j}^{\text{th}}\ \text{DISCO'S total power demand[p. u. M. W]}}$$

$$\sum cpfij = 1$$

$cpfij$ = for j$^{\text{th}}$ DISCOM load Contract Participation Matrix in terms of i$^{\text{th}}$ GENCO. If there are more GENCOs in a given area in accordance with their involvement in AGC, ACE has to be divided among them; this is indicated by the "ACE participation factor" (apf).

The area control error in the multi area power system

ACE=B$\Delta f + \Delta P_{Tie}$ (1)

The Area Contract Error Participation facto

$\sum_{j=1}^{n_i} apf_{ji}$=1 (2)

The total Load Demand

$\Delta P_{Total} = \Delta P_{loc,i} + \Delta P_{di}$ (3)

The local contracted load demand given as

$\Delta P_{loc,i} = \sum_{j=1}^{m_i} \Delta P_{L\ j-i}$ (4)

The un contracted load demand given as

$\Delta P_{di} = \sum_{j=1}^{m_i} \Delta P_{UL\ j-i}$ (5)

The deviation in scheduled tie line power flow

$\zeta_1 = \sum_{\substack{k=1 \\ i \neq k}}^{N} \Delta P_{Tie,ik,scheduled}$ (6)

$\Delta P_{Tie,ik\ scheduled} = \sum_{j=1}^{n_i} \sum_{t=1}^{m_k} apf_{(s_i+j)(s_k+i)} \Delta P_{L\ t-k} - \sum_{t=1}^{n_k} \sum_{j=1}^{m_i} apf_{(s_k+i)(s_i+j)} \Delta P_{L\ j-i}$ (7)

$\Delta P_{tie\ ,ik\ error} = \Delta P_{tie,i,actual} - \zeta_1$ (8)

Generation of power of each GENCO

$\Delta P_{m,k-i} = \rho_{ki} + apf_{ki} \sum_{j=1}^{m_i} \Delta P_{UL\ j-i}$ k=1, 2....n_i (9)

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

ρ_{ki} is contracted load demand $GENCO_{k-i}$

$$\rho_{ki} = \sum_{j=1}^{N} \sum_{I=1}^{m_i} cpf_{(s_i+k)(s_j+t)}\, \Delta P_{L\ j-i} \quad (10)$$

$$\int A\dot{C}E_1 dt = \left\{B_1 \Delta f_1 + \Delta P_{tie12} - \left(\sum_{i=1}^{2} \sum_{j=3}^{4} cpf_{ij} \Delta P_{Lj} - \sum_{i=3}^{4} \sum_{j=1}^{2} cpf_{ij} \Delta P_{Lj}\right)\right\} \quad (11)$$

Similarly for the other area it is given by

$$\int A\dot{C}E_2 dt = \left\{B_2 \Delta f_2 + a_{12} \Delta P_{tie12} - a_{12}\left(\sum_{i=1}^{2} \sum_{j=3}^{4} cpf_{ij} \Delta P_{Lj} - \sum_{i=3}^{4} \sum_{j=1}^{2} cpf_{ij} \Delta P_{Lj}\right)\right\} \quad (12)$$

A. MODEL OF THERMAL PLANT

The thermal power plant [36] can be described in terms of various models as follows:

Governor Model, $G_g = \frac{\mathrm{K_g}}{\mathrm{T_g S+1}}$

Reheater Model, $G_r = \frac{\mathrm{K}_r T_r S+1}{T_r S+1}$

Steam turbine Mode, $G_t = \frac{K_t}{T_t S+1}$

Generator Model = $\frac{\mathrm{K_p}}{T_P\, S+1}$

With *Tp, Tr, Tg* and Tt representing the generator, reheater governor time and turbine constants, and Kp, Kr, Kg and Kt representing the generator, reheater governor time and turbine gains.

B. MODEL OF HYDRO PLANT

The hydal power plant is described as follows [36]:

Governor Model, $G_g = \frac{\mathrm{K_{gh}}}{T_{gh}\ S+1}$

Transient droop compensator = $\frac{T_{RS}\, S+1}{T_{Rh}\, S+1}$

Turbine Model, $G_t = -\frac{\mathrm{T_W\ S+1}}{0.5\mathrm{T_W\ S+1}}$

Where in the penstock, Tw is the nominal start time of water, $_{\mathrm{Tgh}}$, $_{\mathrm{TRs}}$, and $_{\mathrm{TRh}}$ are the governor's time constants, transient droop, and reset time constants of the hydro governor, and $_{\mathrm{Kgh}}$ is the governor's gain.

C. MODEL OF WIND POWER PLANT

The doubly-fed induction generator (DFIG) of a wind turbine (WT) is the subject of this study. The following output power parameters are achieved by wind turbines when they transform wind energy into electricity [34].

$$P_W = 0.5\rho\, AC_p v_w^3$$

In this case, ρ, *A, Cp*, and V stand for, respectively, air density, blade swept area, power coefficient, and wind speed. The power coefficient of a wind turbine is

$$(0.44 - 0.0167\beta) sin\left(\frac{\pi(\lambda - 2)}{15 - 0.3\beta}\right) - 0.00184(\lambda - 3)$$

where

λ -tip speed ratio

β - Blade pitch angle.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The transfer function can be obtained using the equation [40]

$$G_W = \frac{K_{PW1}\,(1+S\,T_{PW1})}{1+S}\,\frac{K_{Pw2}}{1+S\,T_{PW2}}\,\frac{K_{PW3}}{1+S}$$

where $_{TPW1}$ and $_{TPW2}$ are the wind plant system constants and $_{KPW1,}$ $_{KPW2}$, and $_{KPW3}$ are the wind power plant's gains.

D. PHOTO VOLTAIC PLANT

Both series and parallel connections are made between the photovoltaic cells. The nonlinear relationship between voltage and current results from variations in solar energy throughout the day. A maximum power point tracker (MPPT) is used to boost the power output from a solar PV panel. The solar photo voltaic plant's transfer function is explained as follows [40]:

$$G_{pv} = \frac{K_{PV1}\,S + K_{PV2}}{S^2 + T_{PV1}\,S + T_{PV2}}$$

Where $_{TPV1}$ and $_{TPV2}$ are the PV system's time constants and $_{KPV1}$ and $_{KPV2}$ are the gains of the Photo voltaic system with MPPT. Under some circumstances, a PV system's MPPT can be recovered using the incremental conductance (IC) technique.

At the right $\frac{dP_{pv}}{dV_{pv}} > 0$

$At\ MPP\ \frac{\mathrm{dP_{pv}}}{\mathrm{dV_{pv}}} = 0$

$At\ the\ left\ \frac{\mathrm{dP_{pv}}}{\mathrm{dV_{pv}}} < 0$

E. MODEL OF EV AGGREGATORS

In Figure 2. is a block diagram model of electrical vehicle fleets as a whole, including battery chargers, PFCs, and LFCs. The battery chargers are utilized to regulate the power transaction between the grid and batteries. If all the EVs suddenly disconnect from the grid, it will lead to undesirable frequency response. Usually, every electric vehicle may have droop characteristics that are defined by a dead band function to solve this problem. The dead bandies' minimum limit (Δf_{LL}) and maximum limit (Δf_{UL}) are measured at 10 MHz and -10 MHz, respectively. TEVi is the battery's time constant, and KEVi is the EV gain. The SOC of EVs reveals the value of KEVi. ΔP_{AG}^{max} and ΔP_{AG}^{min} indicate the maximum limit and minimum limit of the power output of electric vehicle fleets [41].

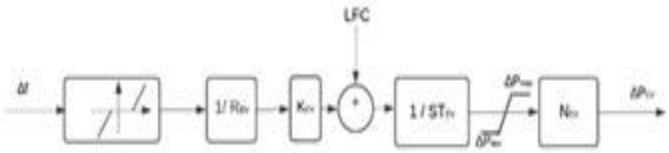


Figure 2. Electric Vehicle model

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

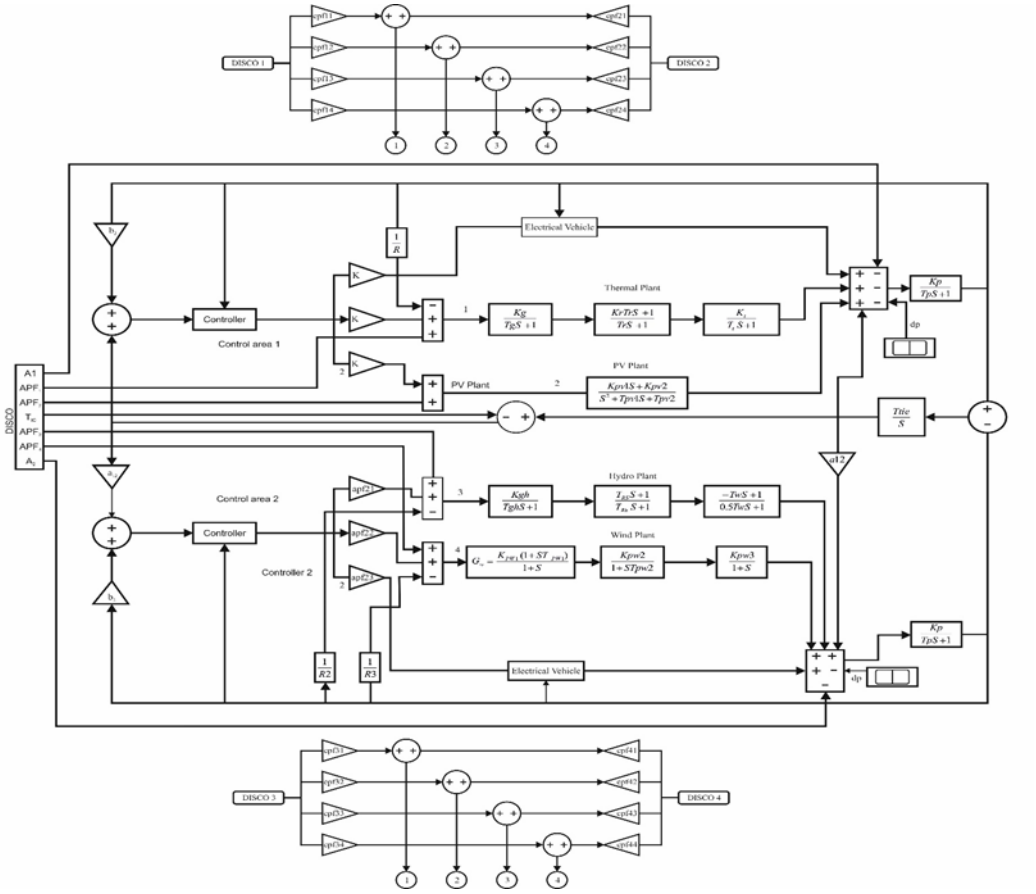


Figure 3. Two Area Interconnected Deregulated Power System (Environment)

DEEP DETERMINISTIC POLICY GRADIENT

The Reinforcement Learning (RL) is a subset of machine learning that focuses on making decisions to maximize cumulative rewards in a given scenario. Unlike supervised learning, which uses a training dataset with predefined answers, RL learns from experience. In RL, an agent learns to achieve a goal in an uncertain and potentially complicated environment by taking actions and receiving feedback in the form of rewards or penalties. Optimal action selection is a difficult challenge in continuous control problems, as the agent must select from a much larger and frequently infinite array of actions. Policy gradients like Soft Actor-Critic, Proximal Policy Gradient (PPO), Deep Deterministic Policy Gradient (DDPG), Trust Region Policy Optimization (TRPO), are commonly employed in reinforcement learning (RL) applications for continuous control problems. Under these circumstances, the agent picks up a policy that associates states with acts directly. The purpose of this study is to look at DDPG.

Inspired by Deep Q-Network, Deep Deterministic Policy Gradient (DDPG) is an off-policy deep reinforcement technique that does not require a model. Actor-Critic Policy Gradient serves as its basis. The Deterministic policy ensures that the RL environment always performs the same action for a given state. In model-free RL, the agent learns by exploring its surroundings and acquiring experience through trial and error. The agent tests numerous actions to see which ones give better results and then updates its policy accordingly. The Actor-Critic RL seeks to determine the best policy for the agent in a given environment by combining two components: actor and critic. It combines value-based and policy-based techniques, with the Actor controlling agent behaviour using the Policy gradient and the Critic evaluating the Agent's action based on the value-function.

A INTRODUCTION TO ACTOR-CRITIC AGENT BASED DEEP DETERMINISTIC POLICY GRADIENT

Actor: The Actor learns the ideal policy by investigating the surroundings.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Critic: The Critic evaluates the value of each action done by the Actor to decide whether it will result in a greater reward, advising the Actor on the best course of action to take [42]. The Actor then uses the Critic's remarks to alter its policies and make informed decisions, resulting in better overall performance [42].

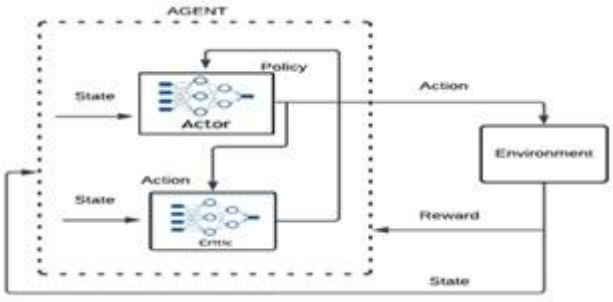


Figure 4. Actor critic Control training

Based on DQN techniques, DDPG, or Deep Deterministic Policy Gradient, is a reinforcement learning algorithm. DDPG specifically makes use of two DQN approaches, Replay buffer and Target network. In addition to these techniques, DDPG uses two sets of Actor-Critic neural networks for function approximation. Both sets are made up of an Actor network and a Critic network, which have the same structure and parameters [43].

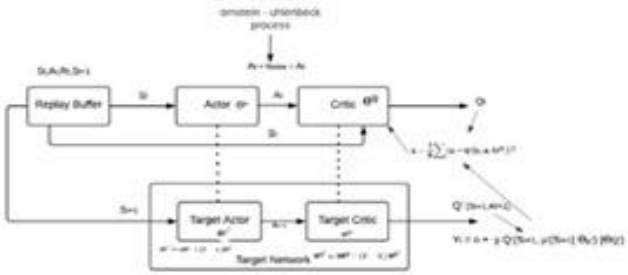


Figure 5. Deep Deterministic Policy gradient

Soft Target Updates

Soft target updates are used in DDPG to gradually update the target network weights rather than copying them immediately from the Actor-Critic network [43].The Soft Target approach significantly enhances the stability of learning, and soft target updates are implemented as follows:

$$\theta^{Q'} = \tau\theta^{Q} + (1-\tau)\theta^{Q'}$$

$$\theta^{\mu'} = \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}$$

At each time step, just a small portion of the weights from the Actor-Critic network are passed on to the target network. This proportion is set by a hyperparameter called the target update rate (τ).

The DDPG optimization approach is used to train the Actor Critic-based agent as an LFC controller. The following elements comprise the proposed interconnected deregulated system, with each area having its own frequency controller (agent) trained in it:

ENVIRONMENT

The term "environment" refers to everything in a linked system except the agent. Based on observations from the environment, an agent selects an appropriate action at each time step; the environment reacts by providing a reward and by gathering new observations opposing the chosen action.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

STATE

In the proposed deregulated system, the agent's input and state is called the ACE.

REWARD

The environment gives response to assess whether the system is accomplishing its objectives by comparing the agent's behavior to each observation. To fulfill objective functions, the agent is thereby guided by the reward function to engage in activities that maximize values.

$R_t = -\sum_{i=1}^{T} B\Delta f + \Delta P_{tie,error}$

Where negative value minimizes the error, thus maximizes the reward.

ACTION

It is the output of the agent to the power system which occurs as a control signal (ΔP_c) and the policy that maximizes reward at a given state determines its value.

In the proposed deregulated interrelated system, the agent's input/state are represented by the area control error (ACE). To compute the agent's action in both areas, the states are given as proportionate integrals of the ACE.

B.DESIGN OF ACTOR CRITIC CONTROLLER FOR LFC OF SMART DEREGULATED POWER SYSTEM WITH EV AGGRAGATORS

The primary goal is to integrate renewable energy sources (RES) and choose random step disruptions in the power system to direct actor critic agents, thereby minimizing frequency as well as the tie-line power under uncertain circumstances. Agent outputs the signal to the environment after accepting the frequency response will be in the form of an ACE from the environment. The networks of actor and critic comprise the agent. The state is given as input to the actor which gives control signal as the action which is given as the input to the critic. The critic compares the action of the actor with the new state value and gives value based on the reward and tells the actor how good the action was. The critic is an estimated value function, whereas actor is a policy framework that decides what to do. We improved the actor critic agent by adding a new layer made up of the function y = abs(weights) $*$ x in place of the actor-network's fully connected layer. While gradient descent optimization may provide negative results, this new layer guarantees positive weights. Table-I lists the actor critic controller's parameters.

C.STEPS FOR THE IMPLEMENTATION OF ACTOR CRITIC CONTROLLER

1. **Initialize**:

1. Initialize the **actor network** (policy network) with random weights (ϕ).
2. Initialize the **critic network** (value function) with random weights(θ).
3. Initialize the **environment** and start from an initial state st(ACE)
4. Set the **discount factor** γ and learning rates αa(for the actor) and αc (for the critic).

2. **For each episode**:

1. Set the initial state st(ACE)
2. **Repeat** for each time step until the episode ends:
3. **Actor**: Choose an action at (ΔPc) according to the policy π(ΔPc |ACE).
4. Take action at in the environment, observe the reward rt, and transition to the next state st+1.
5. **Critic**: Compute the temporal difference (TD) error: δt= R_t+γV(st+1)−V(ACE)

Where: $R_t = -\sum_{i=1}^{T} B\Delta f + \Delta P_{tie,error}$

3.Update the Critic: Update the critic by minimizing the TD error: θ←θ+αcδt∇θV(ACE)

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4.Update the Actor: Update the actor's policy using the TD error ϕ ← ϕ +αaδt∇θπlogπ(ΔPc |ACE)

5. Repeat until convergence (i.e., until the agent's performance stops improving or reaches the desired number of episodes).

Figure 6. Training plot for actor critic

Table -I Hyper parameters of proposed Actor Critic Control

Parameters	Value
No of Observations	2
No of Actions	2
Actor Network	
Hidden Layers	3
Number of neurons in hidden layer 1	96
Number of neurons in hidden layer 2	16
Number of neurons in hidden layer 3	6
Learning rate	1e-5
Activation function	Relu
Critic Network	
Hidden Layers	3
Number of neurons in hidden layer 1	64
Number of neurons in hidden layer 2	32
Number of neurons in hidden layer 3	32
Learning rate	1e-5
Training	
Sample Time	0.1 sec
Max Episode	50

RESULTS AND DISCUSSION

Scenario-I: Poolco based transaction

In this scenario, GENCOs of each area contribute equally in AGC, with ACE participation factors being represented in the DPM. It is presumed that the load change affects area 1 only. The load is requested by DISCO 1 and DISCO 2, for an assumption that a load demand of 0.1 MW per unit is there. The Disco participation matrix (DPM) is given by

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$DPM = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The total demand of DISCOs and the Contract Participation Factor (cpf) can be used to characterize the intended production of a GENCO in pu MW.

$\Delta PG_1 = 0.5(0.1) + 0.5(0.1) = 0.1$ puMW

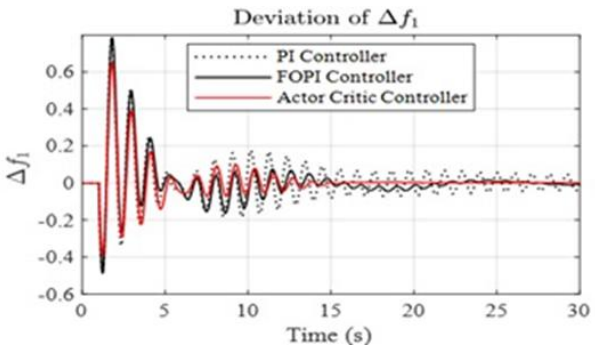
$\Delta PG_2 = 0.5(0.1) + 0.5(0.1) = 0.1$ puMW

$\Delta PG_3 = 0$ puMW

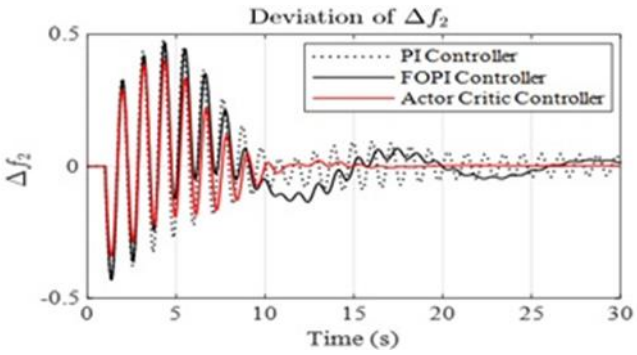
$\Delta PG_4 = 0$ puMW

$\Delta Ptie_{1-2} = 0$ puMW

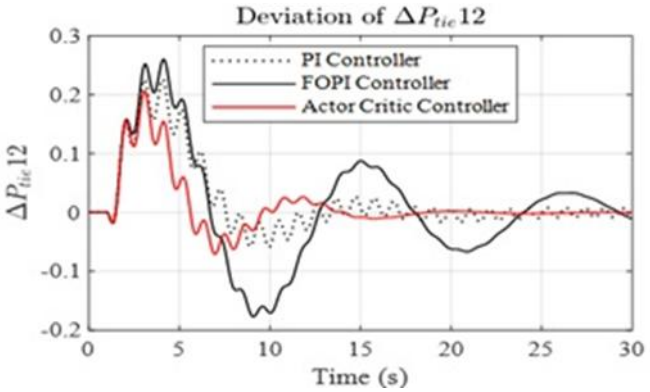
Poolco transactions shown from (a) to (c) with Pl, FOPI and Actor Critic controller.



(a) Change in frequency in Area 1



(b) Change in frequency in Area 2



(c) Tie line power

Table 2 The dynamic response for LFC under Poolco Transaction given below.

	PI Controller	FOPI Controller	Actor Critic Controller
Δf1	-0.0013	-0.0006	0.0000
Δf2	-0.0011	-0.0008	0.0000
ΔP_tie12	0.0057	0.0012	0.0000
Settling time (sec)	25	22	15
% overshoot	74%	70%	62%

Scenario II: Bilateral transaction

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

There is a contract in place for this kind of transaction between DISCOs and GENCOs from the same area as well as from the other locations. GENCOs are required to supply 0.1p.u. of power to each DISCO. The matrix's components are

$$apf_1 = 0.75; apf_2 = 0.25; apf_3 = 0.5; apf_4 = 0.5$$

$$DPM = \begin{bmatrix} 0.5 & 0.25 & 0 & 0.3 \\ 0.2 & 0.25 & 0 & 0 \\ 0 & 0.25 & 1 & 0.7 \\ 0.3 & 0.25 & 0 & 0 \end{bmatrix}$$

$\Delta PG1 = cpf_{11} \times \Delta PL1 + cpf_{12} \times \Delta PL2 + cpf_{13} \times \Delta PL3 + cpf_{14} \times \Delta PL4$

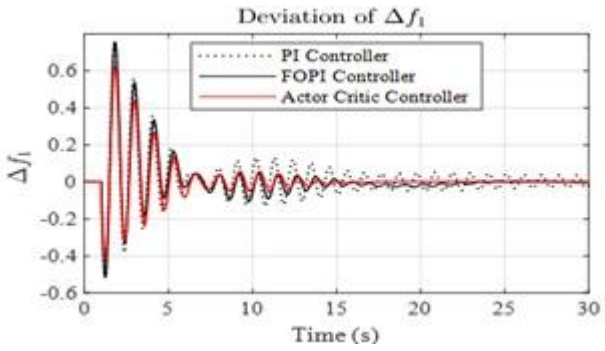
$= 0.5*0.1 + 0.25*0.1 + 0*0.1 + 0.3*0.1 = 0.105puM$

$\Delta P_{G2} = 0.045puMW$

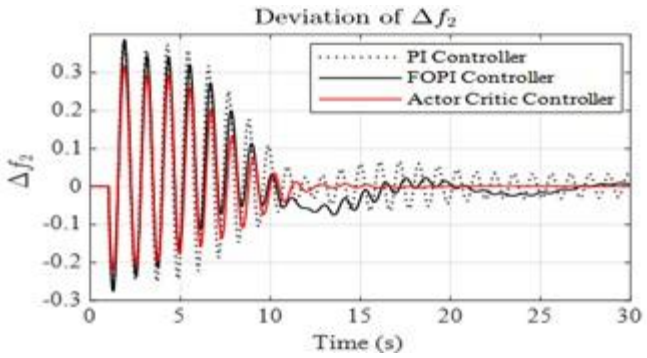
$\Delta P_{G3} = 0.195puMW$

$\Delta P_{G4} = 0.105puMW$

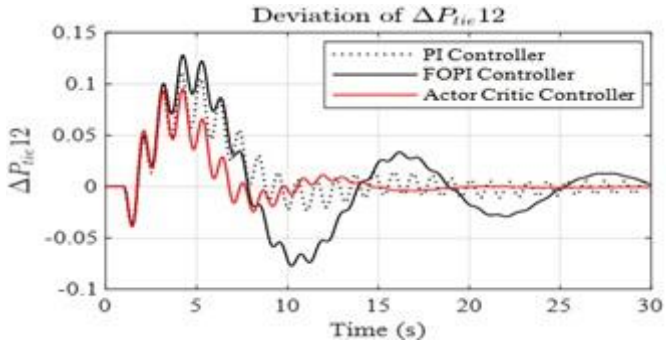
The responses for the bilateral transactions shown from (a) to (c) with Pl, FOPI and Actor- Critic controller.



(a) Change in frequency in Area 1



(b) Change in frequency in Area 2



(c) Tie line power

Table 3 The dynamic response for LFC under Bilateral transaction given below.

	PI Controller	**FOPI Controller**	**Actor Critic Controller**
Δf1	0.0090	-0.0008	0.0000

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Δf2	-0.0092	-0.0015	0.0000
ΔP_tie12	0.0043	0.0001	0.0000
Settling time (sec)	28	25	18
% overshoot	72%	68%	60%

Scenario III: With contract violation Bilateral transaction

The DISCO may sometime violate the bilateral transaction and may demand more power. Case two is taken into consideration for the cause of visualizing a contract violation, with DISCO's demand for 0.1puMW of extra power. This extra power is supplied by EVs.

The total local load in area-1.

(ΔP_{L1LOC}) = Load of DISCO 1 + load od DISCO2 = $(0.1 + 0.1) + 0.1 = 0.3$puMW.

Similarly , total load in area II = load of DISCO3 + load of DISCO4 = 0.2pu MW

$$\Delta P_{tie12schedule} = \sum_{j=1}^{2}\sum_{j=3}^{4} cpfij\Delta pLj - \sum_{j=3}^{4}\sum_{j=1}^{2} cpfij\Delta pLj$$

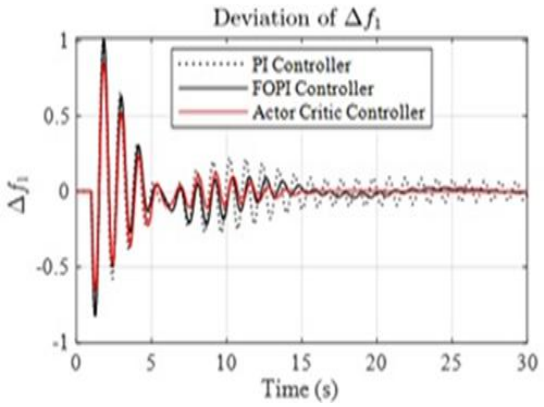
$\Delta P_{tie12,schedule} = -0.08$pu MW

$\Delta P_{G1} = cpf_{11} * \Delta P_{L1} + cpf_{12} * \Delta P_{L2} + cpf_{13} * \Delta P_{L3} + cpf_{14} * \Delta P_{L4}$

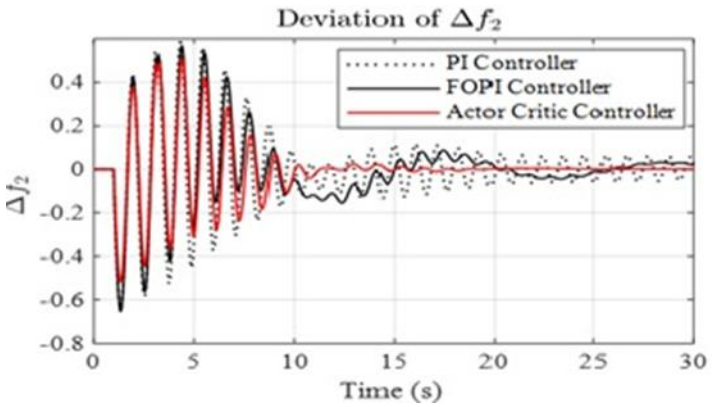
$= 0.5(0.1+0.1)+(0.25*0.1)+(0*0.1)+(0.3*0.1) = 0.155$puMW

$\Delta P_{G2} = 0.065$puMW, $\Delta P_{G3} = 0.195$puMW, $\Delta P_{G4} = 0.085$puMW

Responses for Contract Violation transactions shown from (a) to (c) with Pl, FOPI and Actor critic controller



(a) Change in frequency in Area 1

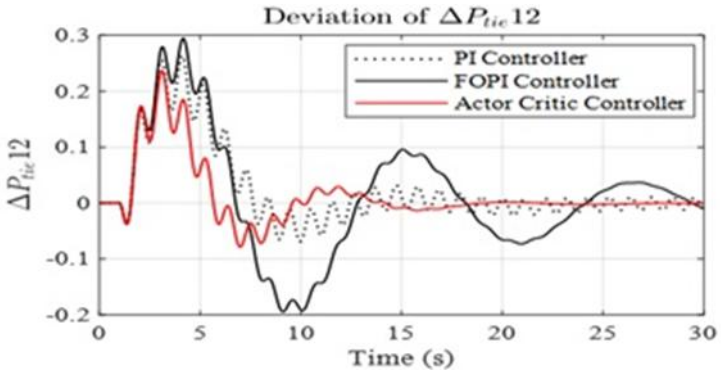


(b) Change in frequency in Area 2

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



(c)Tie line power

Table 4 The dynamic response for LFC under on contract violation.

	PI Controller	FOPI Controller	Actor Critic Controller
Δf1	0.00313	0.00005	0.00000
Δf2	-0.00578	-0.00042	0.00000
ΔP_tie12	0.00823	0.00129	0.00000
Settling time (sec)	30	28	16
% overshoot	82%	80%	68%

The frequency response and tie line power and output power of various power plant are compared with PI, FOPI and Actor Critic controller for poolco, bilateral and contract violation Scenario of deregulated power system are shown in waveforms. The Actor Critic controller based on DDPG algorithm shows considerable good performance compared to PI and FOPI controllers.

5.CONCLUSION

In paper presents two area deregulated power system is taken which comprises of thermal and PV plants in CA-1 and CA-2 consist of wind, hydal and EV aggregators are integrated in both control areas to supply power to the DISCOs in the contract violation scenario. The Actor Critic-based agent is trained as an LFC controller for frequency regulation of SDPS using the DDPG optimization technique. The superiority of the R-L control methodology over other controlled strategies is its uniqueness to train the neural network for the system and control environment of SDPS. The performance of Actor Critic controllers is compared to that of PI and FOPI controllers. The results demonstrate that, in contrast to PI and FOPI, the Deep Deterministic Policy Gradient (DDPG) reinforcement learning controller based on actor-critic agent controls frequency effectively under various transactions of the deregulated electricity system.

Appendix A

Rated power= 2000 MW, f=50Hz, H = 5, D = .00833, B_i =. 425p.u.MW /Hz, 2∏T_{12} = 0.545, Kp= 120, T_p = 20sec, R_i=2.4 Hz/ p.u.MW, a_{12} = -1

Power Plant	Parameters	Value
Thermal	K_r	0.33Hz / pu MW
	K_g	1
	K_t	1sec
	T_g	.08sec
	T_r	10sec
	T_t	0.18sec

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Photovoltaic	K_{PV1}	-18
	K_{PV2}	900
	T_{PV1}	100sec
	T_{PV2}	50sec
Hydro	T_{gh}	10sec
	K_{gh}	1
	T_{Rs}	2sec
	T_{Rh}	1sec
	T_{w}	1sec
Wind	K_{pw1}	1.25
	K_{pw2}	1.4
	T_{pw1}	1sec
	T_{pw2}	.041sec
Electric Vehicle	K_{EV}	1
	T_{EV}	1sec

Appendix B

Parameters	PI Controller		FOPI Controller	
	Area 1	Area 2	Area 1	Area 2
K_p	0.390	0.250	0.35	0.19
K_i	0.490	0.260	0.520	0.29
λ	---	---	0.8	0.8

REFRENCES

[1] D. X. Zhang, X. Q. Han, and C. Y. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids", CSEE Journal of Power and Energy Systems, vol. 4, no. 3, pp. 362–370, Sep. 2018.
[2] M. L. Tuballa and M. L. Abundo, "A review of the development of Smart Grid technologies", Renewable and Sustainable Energy Reviews, vol. 59, pp. 710–725, Jun. 2016.
[3] R. C. Qiu and P. Antonik, Smart Grid and Big Data: Theory and Practice, New York: Wiley Publishing, 2017.
[4] J. R. Pillai and B. Bak-Jensen, "Integration of vehicle-to-grid in the western Danish power system," IEEE Trans. Sustain. Energy, vol. 2, no. 1, pp. 12–19, Jan. 2011.
[5] H. Yang, C. Y. Chung, and J. Zhao, "Application of plug-in electric vehicles to frequency regulation based on distributed signal acquisition via limited communication", IEEE Trans. Power Syst., vol. 28, no. 2, pp. 1017–1026, May 2013.
[6] H. Liu, Z. Hu, Y. Song, J. Wang, and X. Xie, "Vehicle-to-grid control for supplementary frequency regulation considering charging demands", IEEE Trans. Power Syst., vol. 30, no. 6, pp. 3110–3119, Nov. 2015
[7] Liang Chen, Rui Ma," Clean energy synergy with electric vehicles: Insights into carbon footprint",Elsevier,Energy strategy Reviews,volume 53,May 2024.
[8] Erick F. Alves; Louis Polleux; Gilles Guerassimoff; Magnus Korpås; Elisabetta Tedeschi,"Allocation of Spinning Reserves in Autonomous Grids Considering Frequency Stability Constraints and Short-Term Solar Power Variations", IEEE Access ,volume 11,March 2023
[9] H. Bevrani, "Robust Power System Frequency Control", 1st ed. New York: Springer, 2009
[10] Richard D Christie and Anjan Bose, "Load Frequency control Issues in Power System Operations After Deregulation", IEEE Trans. on Power Syst., vol.11, no.3, pp.1191-1200, Sept 1996.
[11] Bjorn H Bakken and Ove S Grande, "Automatic Generation Control in a Power System", IEEE Trans. On Power Systems., vol.13, no.4, pp.1401- 1406, Nov 1998.
[12] Shayeghi H and Shayanfar H A and O P Malik, "Robust decentralized neural networks based LFC in a

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

deregulated power system", Electric Power systems Research, vol.77, pp.241-251, 2007.
[13] Demiroren A and Zeynelgil H L, "GA application to optimization of AGC in three-area power system after deregulation", Electrical Power and Energy Systems, vol.29, pp.230-240, 2007.
[14] Praghnesh Bhatt, Ranjit Roy and S P Ghoshal, "Optimized multi area AGC simulation in restructured Power systems", Electrical Power and Energy Systems, vol.32, pp.311-322, 2010.
[15] Elyas Rakhshani and JavadSadeh, "Practical viewpoints on load frequency control problem in a deregulated power system", Energy Conver- sion and Management, vol.51, pp.1148-1156, 2010.
[16] Wen Tan, Hongxia Zhang and Mei Yu, "Decentralized load frequency control in deregulated environments", Electrical Power and Energy Sys- tems, vol.41, pp.16-26, 2012.
[17] Wen Tan and Hong Zhou, "Robust analysis of decentralised load frequency control for multi-area power systems", Electrical Power and Energy Systems, vol.43, pp. 996-1005, 2012.
[18] Sanjoy Debbarma, Lalit Chandra Saikia and NidulSinha, "AGC of a multi-area thermal system under deregulated environment using a non- integer controller", Electric Power Systems Research, vol.95, pp.175-183, 2013.
[19] H. A. Yousef, Power System Load Frequency Control: Classical and Adaptive Fuzzy Approaches, 1st ed. Boca Raton, FL, USA: CRC Press, 2017.
[20] S. Aziz, H. Wang, Y. Liu, J. Peng, and H. Jiang, "Variable universe fuzzy logic-based hybrid LFC control with real-time implementation", IEEE Access, vol. 7, pp. 25535–25546, 2019.
[21] H. A. Yousef, K. AL-Kharusi, M. H. Albadi, and N. Hosseinzadeh, "Load frequency control of a multi-area power system: An adaptive fuzzy logic approach", IEEE Trans. Power Syst., vol. 29, no. 4, pp. 1822–1830, Jul. 2014
[22] H. M. Hasanien and S. M. Muyeen, "A Taguchi approach for optimum design of proportional-integral controllers in cascaded control scheme", IEEE Trans. Power Syst., vol. 28, no. 2, pp. 1636–1644, May 2013.
[23] I. Pan and S. Das, "Kriging based surrogate modelling for fractional order control of microgrids", IEEE Trans. Smart Grid, vol. 6, no. 1, pp. 36–44, Jan. 2015.
[24] M. I. Alomoush, "Load frequency control and automatic generation control using fractional-order controllers", Elect. Eng., vol. 91, no. 7, pp. 357–368, Mar. 2010.
[25] I. Podlubny, I. Petráš, B. M. Vinagre, P. O'Leary, and L. Dorcák, "Analogue realizations of fractional-order controllers", Nonlin. Dyn., vol. 29, nos. 1–4, pp. 281–296, Jul. 2002.
[26] C. Ismayil, R. S. Kumar, and T. K. Sindhu, "Optimal fractional order PID controller for automatic generation control of two-area power systems", Int. Trans. Elect. Energy Syst., vol. 25, no. 12, pp. 3329–3348, Dec. 2015.
[27] S. Debbarma, L. C. Saikia, and N. Sinha, "AGC of a multi-area thermal system under deregulated environment using a non-integer controller", Elect. Power Syst. Res., vol. 95, pp. 175–183, Feb. 20
[28] I. Pan and S. Das, "Fractional order AGC for distributed energy resources using robust optimization", IEEE Trans. Smart Grid, doi: 10.1109/TSG.2015.2459766.
[29] M. Glavic, "(Deep) reinforcement learning for electric power system control and related problems: A short review and perspectives", Annu. Rev. Control, vol. 48, pp. 22–35, Jan. 2019.
[30] Z. Yan and Y. Xu, "Data-driven load frequency control for stochastic power systems: A deep reinforcement learning method with continuous action search", IEEE Trans. Power Syst., vol. 34, no. 2, pp. 1653–1656, Mar. 2019.
[31] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments", in Proc. Adv. Neural Inf. Process Syst., 2017, pp. 6382–6393
[32] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning", 2015, arXiv:1509.02971
[33] H. Dong and H. Dong, Deep Reinforcement Learning. Singapore: Springer, 2020
[34] Z. Yan and Y. Xu, "A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system", IEEE Trans. Power Syst., vol. 35, no. 6, pp. 4599–4608, Jun. 2020
[35] Junaid Khalid, Makbul A.M Ramli, Muhammad saud khan and taufal hidayat, "Efficient load frequency control of renewable integrated power system: A Twin delayed DDPG- based reinforcement learning approach", IEEE Access., volume 10 ,12 May 2022
[36] Hossam Hassan Ali, Ahmed M. Kassem, Mujahed Al-Dhaifallah, (Member, IEEE), And Ahmed Fathy, "Multi-Verse Optimizer for Model Predictive Load Frequency Control of Hybrid Multi-Interconnected Plants Comprising Renewable Energy", IEEE Access., Volume 8, June 22, 2020,

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[37] Kumar, K. V., Kumar, T. A., & Ganesh, V. (2016, November). Chattering free sliding mode controller for load frequency control of multi area power system in deregulated environment. In 2016 IEEE 7th power India international conference (PIICON) (pp. 1-6). IEEE.

[38] Reddy, Palle Jayabharath, and Thalluru Anil Kumar. "AGC of three-area hydro-thermal system in deregulated environment using FOPI and IPFC." 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS). IEEE, 2017.

[39] Anil Kumar, T., Ramana, N. V. “Tuning of Sliding mode observer optimal parameters for load frequency control in coordination with frequency controllable HVDC link in multi area deregulated power system”, International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO),2015.

[40] Y. Arya, “AGC of PV-thermal and hydro-thermal power systems using CES and a new multi-stage FPIDF-(1+PI) controller”, Renew. Energy, vol. 134, pp. 796–806, Apr. 2019.

[41] Sanjoy Debbarma, “Utilizing Electric Vehicles for LFC in Restructured Power Systems Using Fractional Order Controller”, IEEE Trans. On Smart Grid, volume: 8,pp. 2554 – 2564, Nov. 2017.

[42] Grondman, I.; Busoniu, L.; Lopes, G.A.; Babuska, R. A survey of actor–critic reinforcement learning: Standard and natural policy gradients. IEEE Trans. Syst. Man Cybern. Part C 2012, 42, 1291–1307.

[43] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, “Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications,” IEEE Trans. Cybern., vol. 50, no. 9, pp. 3826–3839, Sep. 2020.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.