Abstract - This paper proposes a reinforcement learning based SVC controller to improve the damping of power systems in the presence of load model parameters uncertainty. The proposed method is trained over a wide range of typical load parameters in order to adapt the gains of the SVC stabilizer. The simulation results show that the tuned gains of the SVC stabilizer using Reinforcement Learning can provide better damping than the conventional fixed-gains SVC stabilizer. To evaluate the usefulness of the proposed method, we compare the response of this method with PD controller. The simulation results show that our method has the better control performance than PD controller.

Keywords: Load Model, SVC Model, Reinforcement Learning, critic, PD Controller.

1 Introduction

Static Var Compensator (SVC) is one of the FACTS devices that are widely used by several utilities to support the voltage of power transmission systems [1]. The SVCs with supplementary signal in their voltage control loops can effectively enhance the damping of power system oscillations and improve power systems stability [2]. In last few years, many researchers have proposed techniques for tuning SVC stabilizers to enhance the damping of electromechanical oscillations of power systems [3-6]. The basic limitation of these techniques is that the influence of load model parameters on tuning SVC stabilizers has not been taken account. Almost all of SVC stabilizers have been tuned based on fixed load parameters. The parameters of typical loads seasonally vary, and in case change over day, consequently, the SVC stabilizers tuned under fixed load parameters may become unacceptable under other load parameters. Damping out the electromechanical oscillations associated with synchronous generators can enhance the damping characteristic of a power system. In this paper, a reinforcement learning based intelligent controller for SVC stabilizer is proposed to damp out these electromechanical oscillations. Simulations performed on the power system show that the proposed method improves the system damping over a wide range of typical load model parameters.

The outline of this paper is as follows. In Section 2 and 3, we introduce the structure of the system model. The structure of the reinforcement learning is then developed in Section 4. In Section 5, we apply the proposed architecture to the system model. Section 6 summarizes the conclusions.

2 System Model

Figure 1, shows the power system under study. The system consists of a synchronous generator connected to a large system through a transmission line. The generator is equipped with automatic voltage regulator and governor-turbine control systems. The load and the SVC are connected to the generator bus. The overall system model is obtained using the component connection model technique.

2.1 Load Model

This paper follows the recommendation of the IEEE working group and utilities [7,8] in utilizing the voltage-dependent load model for composite load representation. Utilities normally perform field tests, or in some cases perform regression analysis to establish system load models to be used for power-flow and stability studies, these models are in the form of:

\[
P_L = P_{L0} V_t^{n_p} \\
Q_L = Q_{L0} V_t^{n_q}
\]  

(1)

Where \(P_L\) and \(Q_L\) are the load active and reactive power. \(V_t\) is the load bus voltage. \(n_p\) and \(n_q\) are the load parameters. \(P_{L0}\) and \(Q_{L0}\), and \(V_0\) are the nominal value of load active power, load reactive power, and bus voltage prior to a disturbance. The load representation given in equation (1) makes possible the modeling of all typical voltage dependent load models by selecting appropriate values of load parameters \((n_p, n_q)\). With load parameters equal 0,
1, or 2, the load model represents constant power, constant current, or constant impedance characteristics. The values of \(n_p\) and \(n_q\) depend on the nature of the load and can vary between 0 to 3.0 for \(n_p\) and 0 to 4.0 for \(n_q\). The measurement of typical values of \(n_p\) and \(n_q\) of various kinds of typical power system composite loads are reported in [8]. These measurement values are required for control parameter adaptation.

2.2 SVC Model

The model of thyristor-controlled reactor SVC type considered in this study is shown in Fig. 2. A single time constant \(T_a\) and gain \(K_a\) represent the thyristors firing control system. The dynamic equation of SVC is given as:

\[
\frac{d\Delta B_{svc}}{dt} = \frac{1}{T_a} \left[ -\Delta B_{svc} + k_v (V_{ref} - V_i + \Delta V_i) \right]
\]

(2)

The variable inductive susceptance \(B_L\) is given by

\[
B_{svc} = j \left( \frac{2\pi - 2\alpha + \sin 2\alpha}{\pi x_s} \right) + \frac{1}{X_C}, \quad \pi \leq \alpha \leq \pi
\]

(3)

Where \(x_s\) is the reactance of the fixed inductor of the SVC and \(\alpha\) is the thyristor firing angle.

\[\text{Fig. 1: Single – line diagram for power system}\]

3 Problem Formulation

Power systems experience poorly damped electromechanical oscillations due to small disturbances. These oscillations may sustain and grow if no adequate damping is available. Sustained oscillations in power systems are undesirable because they can lead to fatigue of machine shafts and to system separation. Therefore, it is desired that these oscillations be well damped. Power system stabilizer is the most widely used to enhance the damping of these oscillations. In recent years, there are a number of FACTS devices that have been extensively used in providing additional damping to power system oscillations. Static Var Compensator (SVC) is one of these FACTS devices. The SVC with additional damping loop can effectively damp out these electromechanical oscillations. The auxiliary (stabilizing) signal \(\Delta V_s\) in Figure 2 is given as:

\[
\Delta V_s = \left( \frac{T_w}{1 + sT_w} \right) \left( k_p + \frac{k_I}{s} \right) \Delta \omega
\]

(4)

Where \(k_p\) and \(k_I\) are the gain-settings of the SVC stabilizer and \(T_w\) is the washout time constant. The stabilizer uses the generator speed deviation or bus frequency deviation \((\Delta \omega)\) as a feedback signal to generate the auxiliary stabilizing signal \(\Delta V_s\) (stabilizer output signal) to the SVC. The signal \(\Delta V_s\) causes fluctuations in the SVC susceptance \(B_L\) and hence in the bus voltage. If the SVC stabilizer is tuned correctly the voltage fluctuations act to modulate the power transfer to damp out the electromechanical oscillations mode.

The proposed approach is to use a reinforcement learning to continuously re-tune the SVC stabilizer gains \((k_p, k_I)\) based on real-time measurements of load parameters by training reinforcement learning over a wide range of typical load parameters.

4 Reinforcement Learning

Reinforcement learning methods embody a general Monte Carlo approach to dynamic programming for solving optimal control problems [9-12]. Q-learning procedures converge on value functions for state-action pairs that estimate the expected sum of future reinforcements, which reflect behavior goals that might involve costs, errors, or profits. To define the Q-learning algorithm, we start by representing a system to be controlled as consisting of a discrete state space, \(S\), and a finite set of actions, \(A\), that can be taken in all states [13,14]. A policy is defined by the probability, \(\pi(S, a)\), that action \(a\) will be taken in state \(S\). Let the reinforcement resulting from applying action \(a_t\) while the system is in state \(S_t\) be \(R(S_t, a_t)\). \(Q_\pi(S_t, a_t)\) is the value function given
state $s_t$ and action $a_t$, assuming policy $\pi$ governs action selection from then on. Thus, the desired value $Q_\pi(s_t,a_t)$ is:

$$Q_\pi (s_t,a_t) = E_\pi \left \{ \sum_{k=0}^{T} \gamma^k R(s_{t+k},a_{t+k}) \right \}$$

where $\gamma$ is a discount factor between 0 and 1 that weights reinforcement received sooner more heavily than reinforcement received later. This expression can be rewritten as an immediate reinforcement plus a sum of future reinforcement:

$$Q_\pi (s_t,a_t) = E_\pi \left \{ R(s_t,a_t) + \gamma \sum_{k=1}^{T} \gamma^{k-1} R(s_{t+k},a_{t+k}) \right \}$$

In dynamic programming, policy evaluation is conducted by iteratively updating the value function until it converges on the desired sum. By substituting the estimated value function for the sum in the above equation, the iterative policy evaluation method from dynamic programming results in the following update to equation, the iterative policy evaluation method from estimated value function for the sum in the above convergence on the desired sum. By substituting the by iteratively updating the value function until it converges on the desired sum. By substituting the by iteratively updating the value function until it converges on the desired sum. By substituting the

$$\Delta Q_\pi (s_t,a_t) = \alpha_t \left [ R(s_t,a_t) + \gamma \max_{a_{t+1}} Q_\pi (s_{t+1},a_{t+1}) \right ] - Q_\pi (s_t,a_t)$$

$$\Delta Q_\pi (s_t,a_t) = \alpha_t \left [ R(s_t,a_t) + \gamma \sum_{a'_{t+1}} P_{s_{t+1}|s_t,a_t} \max_{a_{t+1}} Q_\pi (s_{t+1},a') \right ] - Q_\pi (s_t,a_t)$$

where $0 \leq \alpha_t \leq 1$. To improve the action-selection policy and achieve optimal control, the dynamic programming method called value iteration can be applied. This method combines steps of policy evaluation with policy improvement. Assuming we want to maximize total reinforcement, as would be the case if reinforcements are profits or proximity to a destination, the Monte Carlo version of value iteration for the $Q$-function is:

$$\Delta Q_\pi (s_t,a_t) = \alpha_t \left [ R(s_t,a_t) + \gamma \max_{a_{t+1}} Q_\pi (s_{t+1},a') \right ] - Q_\pi (s_t,a_t)$$

This is what has become known as the Q-learning algorithm. A general structure of reinforcement learning is depicted in Figure 3.

### 5 Simulation Results

The reinforcement learning under consideration has two inputs $n_p$ and $n_q$ (load parameters) and produces the values of proportional and integral gains $K_P$ and $K_I$ (SVC stabilizer gains) separately. A number of simulations have been performed with the proposed method. So that comparison, the fixed-gains SVC stabilizer is tuned first to yield the best damping characteristic to the system at load model parameters. The reinforcement learning SVC stabilizer is a robust controller as far as the variation in load parameters while the system is stable with chosen load model parameters is concerned. From the figures it can be observed that the reinforcement learning based SVC stabilizer provides better damping than the fixed-gain SVC stabilizer. It can be observed from Figs. 5-7, that the system with fixed-gain SVC stabilizer will become unstable under typical load parameters while the system is stable with reinforcement learning SVC stabilizer. From this figures can be concluded that the reinforcement learning the SVC stabilizer is a robust controller as far as the variation in load model parameters is concerned.

### 6 Conclusion

A novel technique based on reinforcement learning is proposed to adapt the PI stabilizer gains of SVC to improve the damping characteristic of a power system over a wide range of typical load model parameters. The reinforcement learning was developed to adapt the gains of the SVC stabilizer. The proposed reinforcement learning was trained based on real-time measurements of SVC stabilizer. The proposed reinforcement learning was trained based on real-time measurements of SVC stabilizer.
load model parameters. The SVC stabilizer gains can be
determined by the reinforcement learning, which makes
the proposed SVC stabilizer relatively simple and suitable
for practical implementation for on-line implementation.
Simulation results show that the proposed method is more
effective means for improving the dynamic performance
of the power system than the fixed-gains SVC stabilizer.

Reference

[1] CIGRE and IEEE Working Groups on FACTS,
“FACTS Overview,” IEEE Power Engineering Society,
95TP108, Apr. 1995

Enhancement by Robust Tuning of SVC Controllers In
the Presence of Load Parameters Uncertainty", IEEE
Trans. On Power Systems Vol. 13, No.4, pp. 1298-1303,
Nov. 1998.

[3] Chang, C.S.; Yu, Q.Z.; Liew, A.C.; Elangovan, S.,
"Genetic algorithm tuning of fuzzy SVC for damping
power system inter-area oscillations", Fourth International

"Fuzzy Adaptive Model Reference Approach-Based
Power System Static VAR Stabilizer", Electric Power

Var Compensator Controller Over a Wide Range of Load
Models Using an Artificial Neural Network”, Electric

compensate the influence of voltage dependent loads”,
Proceedings of the 36th IEEE Conference on , Volume: 3

on Power Systems, Vol. PWRS-1, No. 4, pp. 112-118,

[8] W.Xu, E.Vaahedi, Y. Mansour and J. Tamby,
"Voltage Stability Load parameter Determination from
Field Tests On B.C. Hydro’s System", IEEE Trans. On

of reinforcement learning”, IEEE, 1996.

control through fuzzy Q-learning algorithms”, ICIPS ’97.

time: advantage updating”, IEEE World Cong. on
Computational Intelligence, Vol. 4, pp. 2448-2453, 1994

[12] Min-Soeng Kim; Sun-Gi Hong; Ju-Jang Lee, “Self-
organizing fuzzy inference system by Q-learning”, FUZZ-

controller”, IEEE World Congress on Computational

Multiobjective Satisfactory Fuzzy Logic Controller using
Reinforcement Learning”, International Journal of Fuzzy

Fig. 4: The system time response of at np=nq=2 for 5%
change in input torque Tm
Upper: Reinforcement Learning SVC Stabilizer, Lower:
fixed-gains SVC Stabilizer
Fig. 5: The system time response of at $np=nq=1$ for 5% change in input torque $T_m$
Upper: Reinforcement Learning SVC Stabilizer, Lower: fixed-gains SVC Stabilizer

Fig. 6: The system time response of at $np=1.3$, $nq=3$ for 5% change in input torque $T_m$
Upper: Reinforcement Learning SVC Stabilizer, Lower: fixed-gains SVC Stabilizer

Fig. 7: The system time response of at $np=2.7$, $nq=1.6$ for 5% change in input torque $T_m$
Upper: Reinforcement Learning SVC Stabilizer, Lower: fixed-gains SVC Stabilizer