2003

A Heuristic Dynamic Programming Based Power System Stabilizer for a Turbogenerator in a Single Machine Power System

Wenxin Liu

Ganesh K. Venayagamoorthy
Missouri University of Science and Technology

Donald C. Wunsch
Missouri University of Science and Technology, dwunsch@mst.edu

Follow this and additional works at: http://scholarsmine.mst.edu/faculty_work

Part of the Electrical and Computer Engineering Commons

Recommended Citation
http://scholarsmine.mst.edu/faculty_work/1677

Wenxin Liu, Student Member, IEEE, Ganesh K. Venayagamoorthy Senior Member, IEEE, and Donald C. Wunsch II, Senior Member, IEEE

Department of Electrical and Computer Engineering
University of Missouri - Rolla
Rolla, MO 65409, USA
Email: wl3g3@umr.edu, ganeshv@umr.edu, and dwunsch@umr.edu

Abstract—Power system stabilizers (PSS) are used to generate supplementary control signals for the excitation system in order to damp the low frequency power system oscillations. To overcome the drawbacks of conventional PSS (CPSS), numerous techniques have been proposed in the literature. Based on the analysis of existing techniques, a novel design of power system stabilizer (PSS) based on heuristic dynamic programming (HDP) is proposed in this paper. HDP combining the concepts of dynamic programming and reinforcement learning is used in the design of a nonlinear optimal power system stabilizer. The proposed HDP based PSS is evaluated against the conventional power system stabilizer and indirect adaptive neurocontrol based PSS under small and large disturbances in a single machine infinite bus power system setup. Results are presented to show the effectiveness of this new technique.

Keywords—Neural Networks; Neuro-identifier; Neuro-control; Power System Stabilizer; Indirect Adaptive Control; On-line Training; Adaptive Critic Design; Heuristic Dynamic Programming.

I. INTRODUCTION

Currently most of the generators are equipped with voltage regulators to automatically control the terminal voltage. It is known that the voltage regulator action had a detrimental impact upon the dynamic stability of the power system. Oscillations of small magnitude and low frequency often persist for long periods of time and in some cases even present limitations on power transfer capability [1].

In the analysis and control of power system stability, two distinct types of system oscillations are usually recognized. One type is associated with generators at a generating station swinging with respect to the rest of the power system. Such oscillations are referred to as “intra-area mode” oscillations. The second type of oscillations is associated with the swinging of many machines in the one area of the system against machines in other areas. This is referred to as “inter-area mode” oscillations. Power system stabilizers (PSS) are used to generate supplementary control signals for the excitation system in order to damp both types of oscillations.

The widely used conventional power system stabilizers (CPSS) are designed using the theory of phase compensation in the frequency domain and are introduced as a lead-lag compensator. The parameters of CPSS are determined based on the linearized model of the power system. To have the CPSS provide good damping over a wide operating range, its parameters need to be fine tuned in response to both types of oscillations. Since power systems are highly nonlinear systems, with configurations and parameters that change with time, the CPSS design based on the linearized model of the power system cannot guarantee its performance in a practical operating environment. Thus, an adaptive PSS which considers the nonlinear nature of the plant and adapts to the changes in the environment is required for the power system.

To improve the performance of CPSSs, numerous techniques have been proposed for their design, such as using intelligent optimization methods (simulated annealing, genetic algorithm, tabu search) [2]-[4], fuzzy logic [5]-[6], neural networks and many other nonlinear control techniques [7]-[9]. The intelligent optimization algorithms are used to determine the optimal parameters for CPSS by optimizing an eigenvalue based cost function in an offline mode. Since the method is based on a linearized model and the parameters are not updated online, therefore they lack satisfactory performance during practical operation. The rule-based fuzzy logic control methods are well known for the difficulty in obtaining and adjusting the parameters of the rules especially online. Recent research indicates that more emphasis has been placed on the combined usage of fuzzy systems and other technologies such as neural networks to add adaptability to the design [10]. Currently, most of the nonlinear control based methods use simplified models to decrease complexity of the algorithms. Considering the complexity of practical power systems, more realistic model with less computation time is required for effective robust control over a wide range of operating conditions.
Since neural networks have the advantages of high computation speed, generalization and learning ability, they have been successfully applied to the identification and control of nonlinear systems. The work on the application of neural networks to the PSS design so far includes online tuning of CPSS parameters [11]-[12], the implementation of inverse model control [13]-[14], direct control [15] and indirect adaptive control [16]-[21]. The online tuning of CPSS parameters and the inverse model control do not update the weights of neural networks online so their performances highly depend on the quality of offline training samples which are difficult to obtain. The indirect adaptive neurocontroller design consists of two neural networks, namely the neuro-controller and the neuro-identifier. The neuro-controller is used to generate the stabilizing supplementary control signal to the plant and the neuro-identifier is used to provide a dynamic model of the plant to evaluate and update the weights of the neuro-controller. Since the plant model is not used in the direct adaptive neural network control structure, computation time is decreased. But there is no accurate way to directly evaluate the performance of the controller, especially when the system parameters are changing over time; therefore, this is not the most effective control technique.

The risk with the indirect adaptive neurocontroller scheme is that the training of the controller is carried out all the time and this can lead to instability under uncertainties and large disturbances. In this paper, a novel heuristic dynamic programming (HDP) based optimal power system stabilizer is proposed. HDP is a class of adaptive critic designs which provides optimal control. With adaptive critic designs, neural networks with fixed weights are used as tools for implementing optimal controllers which is a potential benefit in overcoming stability issues. The proposed HDP based PSS is evaluated on a single machine infinite bus power system against those of CPSS and indirect adaptive neurocontrol designs. Simulation results are provided to show the performances of the different controllers.

The power system model is described in section II. The introduction to HDP and the design of the HDP based PSS are described in section III. The training process of the HDP-PSS is described in section IV. Some simulation results are provided in section V.

II. POWER SYSTEM MODEL

The single machine infinite bus power system (SMIB) model used to evaluate the IDNC is shown in Fig. 1. The SMIB called the plant in this paper consists of a synchronous generator, a turbine, a governor, an excitation system and a transmission line connected to an infinite bus. The model is built in MATLAB/SIMULINK environment using the Power System Blockset [22]. In Fig. 1, \( P_{REF} \) is the mechanical power reference, \( P_{SF} \) is the feedback through the governor, \( T_M \) is the turbine output torque, \( V_{INF} \) is the infinite bus voltage, \( V_{REF} \) is terminal voltage reference, \( V_f \) is terminal voltage, \( V_A \) is the voltage regulator output, \( V_f \) is field voltage, \( V_E \) is the excitation system stabilizing signal, \( \Delta \omega \) is the speed deviation, \( V_{PSS} \) is the PSS output signal, \( P \) is the active power and \( Q \) is the reactive power at the generator terminal.

In Fig. 1, the switch \( S_j \) is used to carry out tests on the power system with HDP based controller (HDPC), indirect neural network control based controller (IDNC) and conventional PSS (CPSS) and without PSS (with switch \( S_j \) at position 1, 2, 3 and 4 respectively). Switch \( S_j \) is used to select between normal operation and training phase (position 1 and 2 respectively).

The synchronous generator is described by a seventh order d-q axis set of equations with the machine current, speed and rotor angle as the state variables.

The turbine is used to drive the generator and the governor is used to control the speed and the real power. The block diagram of a separately excited turbine and a conventional governor are shown in Fig. 2.

The excitation system for the generator is modeled according to IEEE Std. 421.5 [23]. The block diagram of the excitation system is shown in Fig. 3.

The CPSS consists of two phase-lead compensation blocks, a signal washout block, and a gain block. The input signal is the rotor speed deviation \( \Delta \omega \) [24]. The block diagram of the CPSS is shown in Fig. 4.

![Fig. 1 System model configuration.](image)

![Fig. 2 Block diagram of the turbine and the governor.](image)

![Fig. 3 Block diagram of the excitation system.](image)

![Fig. 4 Block diagram of the CPSS.](image)
The parameters for the generator, AVR, excitation system, turbine and governor are given in Appendix A [23]-[25].

III. HDP BASED PSS DESIGN

A. Background

Adaptive critic designs (ACDs) are neural network designs capable of optimization over time under conditions of noise and uncertainty. A family of ACDs was proposed by Werbos [26] as new optimization technique combining the concepts of reinforcement learning and approximates dynamic programming. For a given series of control actions that must be taken sequentially, and not knowing the effect of these actions until the end of the sequence, it is possible to design an optimal controller using the traditional supervised learning neural network.

The adaptive critic method determines optimal control laws for a system by successively adapting two ANNs, namely, an action neural network (which dispenses the control signals) and a critic network (which learns the desired performance index for some function associated with the performance index). These two neural networks approximate the Hamilton-Jacobi-Bellman equation associated with optimal control theory. The adaptation process starts with a non-optimal, arbitrarily chosen control by the action network; the critic network then guides the action network toward the optimal solution at each successive adaptation. During the adaptations, neither of the networks needs any “information” of an optimal trajectory, only the desired cost needs to be known. Furthermore, this method determines optimal control policy for the entire range of initial conditions and needs no external training, unlike other neuro-controllers [27].

The design ladder of ACDs includes three basic implementations: Heuristic Dynamic Programming (HDP), Dual Heuristic Programming (DHP) and Globalized Dual Heuristic Programming (GDHP), in the order of increasing power and complexity. The interrelationships between members of the ACD family have been generalized and explained in [28]. In this paper, HDP approach is adopted for the design of a power system stabilizer.

B. General Control Structure

The HDP-PSS consists of three neural networks, which are the action, identifier and critic networks. The action network is used to generate the stabilizing control signals; the identifier network is used to model the plant and estimate its output; the critic network is used to estimate cost-to-go function $J$ given by the Bellman’s equation. The general structure of the HDPC is shown in Fig. 5.

To simply the description of the training process, it is necessary to clarify the time step definitions: Both $V_{ps}(k)$ and $\Delta \omega(k)$ signals are sampled at time step $k$, but $\Delta \omega(k)$ is not the response for the control signal $V_{ps}(k)$. Due to the time lag property of the plant, the impact of the control signal $V_{ps}(k)$ is reflected in the next time sample of the output signal $\Delta \omega(k+1)$. The following sections describe the designs of the three neural networks.

C. Identifier Neural Network Design

The identifier neural network is developed using the series-parallel Nonlinear Auto Regressive Moving Average (NARMA) model [29]. The model output $\hat{y}$ at time $k+1$ depends on both past $n$ values of output and $m$ past values of input. The neuro-identifier output equation takes the form given by (1).

$$\hat{y}(k+1) = f \left[ y(k), y(k-1), \ldots, y(k-n+1), u(k), u(k-1), \ldots, u(k-m+1) \right]$$

(1)

Where $y(k)$ and $u(k)$ represent the output and input of the plant to be controlled at time $k$. For this particular system, $y$, $u$ and $\hat{y}$ are the speed deviation $\Delta \omega$ of the plant, the output of the action network $V_{ps}$ and the estimated plant output $\Delta \hat{\omega}(k)$ by the identifier network respectively. Here both $m$ and $n$ are chosen to be 2. One reason for choosing three time step values is because a third order model of the system is sufficient for the study of transient stability. The other reason is that more time delays means more computation and one author's previous work verified that three time delays is enough for this kind of problem [25].

The identifier network is a multi-layer feedforward network trained with backpropagation (BP) algorithm. The numbers of neurons in the input, hidden and output layers are six, ten and one respectively. Considering the ranges of $\Delta \omega$ and $V_{ps}$ to speed up the training process, the scaling factors for $\Delta \omega$ and $V_{ps}$ are chosen to be 400 and 2 respectively.

The training process of the identifier network is shown in Fig. 6. The inputs to the identifier network are $\{\Delta \omega(k-1), \Delta \omega(k-2), \Delta \omega(k-3), \Delta \omega(k-4), V_{ps}(k-1), V_{ps}(k-2), V_{ps}(k-3)\}$ and its output is $\Delta \hat{\omega}(k)$. The desired output is the output of the plant $\Delta \omega(k)$. The cost function for training the identifier network is given by (2).
\[ J_f(k) = \frac{1}{2} \varepsilon_f(k)^2 = \frac{1}{2} (\Delta \omega(k) - \Delta \hat{\omega}(k))^2 \] (2)

During pre-training of the identifier, the switch \( S_2 \) is at position 2 so that small magnitude Pseudo Random Binary Signal (PRBS) is used to replace the actual network to excite all possible dynamics of plant [21]. During the post-training, the switch \( S_2 \) is at position 1 so that the actual control signal calculated by the action network can be fed to both the plant and the identifier [25].

![Fig. 6 Training of the neuro-identifier during pre-control (the dashed line shows backpropagation path).](image)

D. Critic Neural Network Design

The critic network is also a multi-layer feedforward network trained with BP algorithm. The number of neurons in the input, hidden and output layers are chosen to be three, six and one respectively. The inputs to the neuro-controller are the speed deviation \( \Delta \omega \) and its two previous values and the output of the critic network is the estimated cost-to-go function \( J \), which is defined as:

\[ J(k) = \sum_{i=0}^{\infty} \gamma^i U(k+i) \] (3)

Where \( \gamma \) is the discount factor for finite horizon problems with the range of [0, 1] and is chosen to be 0.5 in this design. \( U(k) \) is the utility function or the local cost. Due to the inertia of the plant, the local/immediate cost \( U(k) \) at every time step is dependent on the present and past speed deviations [25] and is given by:

\[ U(k) = [0.4 \Delta \hat{\omega}(k) + 0.4 \Delta \omega(k-1) + 0.16 \Delta \omega(k-2)]^2 \] (4)

The training process of the critic network can be clearly seen from Fig. 7. During training, first the critic network is fed with the three time delayed outputs of the identifier \( (\Delta \hat{\omega}(k), \Delta \omega(k-1), \Delta \omega(k-2)) \), to calculate the estimated cost-to-go function \( J(k) \). Then critic network is fed with \( [\Delta \omega(k+1), \Delta \omega(k), \Delta \omega(k-1)] \) to calculate the estimated cost-to-go function \( J(k+1) \). According to the Bellman’s definition of \( J(k) \), \( J(k) = \gamma J(k+1) + U(k) \). Therefore, \( \gamma J(k+1) + U(k) \) is the desired target output for \( J(k) \) during the critic network training.

![Fig. 7 Training process of the critic network (the dashed line shows backpropagation path).](image)

E. Action Neural Network Design

The action network is a multi-layer feedforward network trained with BP algorithm. The number of neurons in input, hidden and output layers is three, six and one respectively. The inputs to the action network are the speed deviation \( \Delta \omega \) and its two previous values and its output is the control signal \( V \).

The training process of the action network is illustrated in Fig. 8. The purpose of action network training is to minimize the estimated cost-to-go function by the critic network with effective control signals. In HDP, \( \partial J / \partial \hat{\omega} \) is backpropagated through the critic and identifier networks in order to evaluate the performance of the action network and update its weights accordingly.

![Fig. 8 Training process of the action network (the dashed lines show backpropagation paths).](image)

F. Training Procedure

The general training procedure and more details on ACD is described in [27]. It consists of three separate training cycles: training of the critic network, training of the identifier network and training of the action network. The training frequency for each training cycle may be different. To decrease the computation burden of the training process, the training times for each training sample is set to 1 and the learning rate is set.

273
to 0.1 with a sampling frequency of 20 Hz. The three training cycles are alternated until an acceptable plant performance is achieved.

IV. SIMULATION RESULTS

To evaluate the performance of the HDPC, the system response of the HDPC is compared with the cases where there is no PSS, with a CPSS and with an indirect adaptive neurocontrol based PSS (IDNC) [21] in the system. The comparison is carried out under different kinds of operating conditions and disturbances. These disturbances are namely: a three phase short circuit at the infinite bus, step changes in the terminal voltage reference and change transmission line impedance. All these disturbances are carried out under two operating points, \( P=0.5 \text{ pu}, Q=0.02 \text{ pu} \) and \( P=0.6 \text{ pu}, Q=0.05 \text{ pu} \).

A. Simulation Results at \( P=0.5\text{pu}, Q=0.02\text{pu} \)

1) 200ms three phase short circuit: Figs. 9 and 10 are the comparisons of the system responses under a 200ms three phase short circuit fault occurring at 1 second. It can be seen that CPSS has better damping of the speed deviation than when there is no CPSS in the system; IDNC has better damping than CPSS while HDP has the best damping. From Fig. 10, it can be seen that the terminal voltage responses are comparable for this particular fault.

2) 10% stepchange in the terminal voltage reference: Figs. 11 and 12 are the comparisons of the system response to a 10% step change in \( V_{\text{ref}} \) (1.1 pu to 1.21 pu) at 1 second and 10% decrease (1.21 pu to 1.1 pu) at 8 second. Again, the HDP provides the best damping to the speed deviation for this kind of disturbance and the terminal voltage responses are similar.

3) Change in transmission line impedance: Fig. 13 is the comparison of the system responses to a change in transmission line impedance. During this case, the impedance of the transmission line is changed from \( Z_1=0.025 + j0.7559 \) pu to \( Z_2=0.05 + j1.5 \) pu at 1 second. Again, the HDP provides the best damping to the speed deviation of the four controllers.

B. Simulation Results at \( P=0.6\text{pu}, Q=0.05\text{pu} \)

1) 200ms three phase short circuit: Figs. 14 and 15 are comparisons of the system responses under a 200ms three phase short circuit fault occurring at the infinite bus. The findings of the simulation results are similar to those conclusion in A. 1 above for the first operating point.

\[274\]
3) **Change in transmission line impedance:** Fig. 18 is the comparison of the system responses to a simulated transmission line fault. The impedance of the transmission line changes from $Z_1 = 0.025 + j0.7559$ pu to $Z_2 = 0.0125 + j0.378$ pu at 1 second. For these tests, the HDP still has the best performance.

VI. CONCLUSION

To overcome the drawbacks of conventional power system stabilizers, a HDP based power system stabilizer (PSS) design is presented in this paper. The proposed method is evaluated on a single machine infinite bus power system. The design of the HDP is based on only the speed deviation signals of the synchronous generator. Therefore, the computations involved in the neural network design are minimal. This is desirable for practical hardware implementation on the power station platforms. In addition, the online training computational demand is reduced once the action network is trained for optimal performance over a number of operating points. Simulation results for different kinds of disturbances and operating conditions demonstrate the effectiveness and robustness of the HDP. Such a nonlinear adaptive PSS will yield better and fast damping under small and large disturbances even with changes in system operating conditions. Better and fast damping means that generators can operate more close to their maximum generation capacity. Thus, ensuring that generators remain stable under sever faults such as three phase short circuits. This means that more power generated per invested dollar.
APPENDIX A

Table I: Parameters of the Single Machine Infinite Bus Power System in Fig. 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{\omega} )</td>
<td>0.69s</td>
</tr>
<tr>
<td>( T_{\phi} )</td>
<td>0.25s</td>
</tr>
<tr>
<td>( X_{\phi} )</td>
<td>0.205pu</td>
</tr>
<tr>
<td>( T_{\phi} )</td>
<td>0.66s</td>
</tr>
<tr>
<td>( T_{x} )</td>
<td>27ms</td>
</tr>
<tr>
<td>( X_{x} )</td>
<td>0.164pu</td>
</tr>
<tr>
<td>( T_{w} )</td>
<td>33ms</td>
</tr>
<tr>
<td>( T_{w} )</td>
<td>38ms</td>
</tr>
<tr>
<td>( T_{d} )</td>
<td>1.98pu</td>
</tr>
<tr>
<td>( T_{q} )</td>
<td>26.4ms</td>
</tr>
<tr>
<td>( X_{d} )</td>
<td>2.09pu</td>
</tr>
<tr>
<td>( X_{q} )</td>
<td>0.213pu</td>
</tr>
<tr>
<td>( T_{d0} )</td>
<td>0.15</td>
</tr>
<tr>
<td>( T_{q0} )</td>
<td>0.594</td>
</tr>
<tr>
<td>( T_{d0} )</td>
<td>0.884</td>
</tr>
<tr>
<td>( T_{q0} )</td>
<td>2.662</td>
</tr>
<tr>
<td>( T_{\text{MIN}} )</td>
<td>0</td>
</tr>
<tr>
<td>( T_{\text{MAX}} )</td>
<td>1.2</td>
</tr>
<tr>
<td>( K_{d} )</td>
<td>20</td>
</tr>
<tr>
<td>( T_{d} )</td>
<td>0.264</td>
</tr>
<tr>
<td>( K_{q} )</td>
<td>50</td>
</tr>
<tr>
<td>( T_{q} )</td>
<td>0.01</td>
</tr>
<tr>
<td>( V_{\text{MIN}} )</td>
<td>10</td>
</tr>
<tr>
<td>( V_{\text{MAX}} )</td>
<td>10</td>
</tr>
<tr>
<td>( V_{\text{MIN}} )</td>
<td>10</td>
</tr>
<tr>
<td>( T_{F} )</td>
<td>0.46</td>
</tr>
<tr>
<td>( K_{\text{STAB}} )</td>
<td>25</td>
</tr>
<tr>
<td>( T_{F} )</td>
<td>10</td>
</tr>
<tr>
<td>( T_{d} )</td>
<td>0.76</td>
</tr>
<tr>
<td>( T_{q} )</td>
<td>0.1</td>
</tr>
<tr>
<td>( T_{d} )</td>
<td>0.76</td>
</tr>
<tr>
<td>( T_{q} )</td>
<td>0.1</td>
</tr>
<tr>
<td>( V_{\text{FSS_MIN}} )</td>
<td>0.1</td>
</tr>
<tr>
<td>( V_{\text{FSS_MAX}} )</td>
<td>0.1</td>
</tr>
</tbody>
</table>


