

Reinforcement Learning applied to Power System Oscillations Damping

D. Ernst[†]

Department of Electrical and Computer Engineering - Institut Montéfiore
University of Liège - Sart-Tilman B28 - B4000 Liège - Belgium

[†] Research Fellow F.N.R.S.

Phone: +32-4-3669637 - Fax: +32-4-3662984
ernst@montefiore.ulg.ac.be

Abstract

This paper investigates the use of reinforcement learning in electrical power system oscillations damping. The approach consists in using temporal-difference learning algorithms to control a FACTS (Flexible Alternative Current Transmission System) so as to damp power system oscillations. The proposed approach is based only on local measurements and frees itself from the knowledge of power system dynamics. An illustration is carried out on a one machine infinite bus system.

1 Introduction

Reinforcement learning (RL) takes its origin in optimal control theory and dynamic programming ([1]). It aims at approximating by experience solutions to problems of unknown dynamics. Year after year, the techniques evolved leading to a panel of more and more efficient algorithms. From a theoretical point of view, many breakthroughs have been realized notably concerning the convergence of the algorithms and their applications to nonlinear systems ([7], [6]). Also the steady increase in computer capacities makes RL methods more and more feasible. Therefore, the power system community started getting interested in such techniques, but surprisingly more in market modelling ([4]) rather than in nonlinear system control. However new needs appeared recently in power system dynamics control, especially with the introduction of new devices based on power electronics, like Flexible Alternating Current Transmission systems (FACTS).

In this paper we focus on how to control by means of RL algorithms a FACTS device in order to damp power system oscillations. These latter phenomena become even more important with the growth of extensive power systems and especially with the interconnection of these systems with ties of limited capacity.

Basically, the RL approach proposed in this paper to

control the FACTS consists of an adaptive closed-loop control that tends to maximize a function, image of the quality of the oscillations damping. The only signal used by the controller is the electrical power transferred in the line measured at fixed intervals. This is roughly a discrete time optimal control approach in which the state dynamics (and observation equations) are unknown.

The main advantages of such a controller are that it frees oneself from any knowledge of the power system dynamics equations, adapts itself to changing conditions and is able to act in a stochastic environment.

The methodology used is briefly explained in the next section and an illustration is carried out on a one machine infinite bus system.

2 Methodology

Reinforcement learning algorithms adjust a closed-loop control rule which is a mapping from the system states to the control action. The RL methods we use in this paper belong to the *temporal-difference* type of methods that suppose the existence of a *reward* r_{t+1} associated to the transition from the state at instant t to the state at instant $t + 1$. Then the mapping is done such that at each instant t , the algorithms tend to maximize the *expected return* defined as the discounted sum of rewards over future time steps, that is $\sum_{k=0}^{\infty} \gamma^k r_{t+(k+1)}$ where γ , $0 \leq \gamma < 1$ is a discount factor.

Many difficulties appear when we want to apply successfully these algorithms to the control of a FACTS device, among them, the absence of knowledge about the power system state. Only local measurements are available. To overcome this difficulty, a technique commonly used in partially observable environments is applied to define a suitable state for the RL algorithms ([2]). The state at instant t is not only given by the observation done at time t , but also by a succession of past observations and control actions of the power system. Moreover, observations being relative to continuous variables, they have to be discretized in order

to deal with a finite set of states. Discretization of the control space is also necessary in order to have a finite set of control actions. The finer the discretization, the better the control. Unfortunately this increases the time to converge to a satisfactory solution.

Another problem concerns an appropriate reward definition. A good reward must satisfy two main properties. The first one is to insure that once the control law has converged, the quality of the obtained control is sufficient. The second one is to avoid a too slow convergence of the reinforcement learning algorithm. The reward chosen in this work is based on the notion of *signal error*. If P_{e_t} is the electrical power transmitted in the line at instant t then the reward r_t is $-|P_{e_t} - \overline{P_e}|$ where $\overline{P_e}$ is the estimated average value of the electrical power transferred in the line. With such a reward, the reinforcement learning algorithm will attempt to minimize the distance between $\overline{P_e}$ and P_{e_t} , and so to damp the electrical power oscillations.

3 Illustrations

The simplified power system model used to illustrate the FACTS controlled by reinforcement learning algorithm consists of a single synchronous generator connected to an infinite inertia machine through a transmission line. The FACTS is modelled like a variable capacitance installed in series with the transmission line. The set of possible values for the capacitance is reduced to two elements, allowing the RL algorithms to use only a bang-bang control to damp the electrical power oscillations. The state at time t used in the algorithm is composed of P_{e_t} , $P_{e_{t-1}}$ and the value of the control action decided at time $t - 1$.

The temporal-difference reinforcement learning algorithm used in our experiments is known as Sarsa(λ) with replacing traces (see [7]).

The graphics of figure 1 represent the evolution of the electrical power transmitted in the line when the power system is subject to the same perturbation for different stages of the learning process. At the beginning the damping produced is poor but as the control process proceeds, it becomes better and better.

4 Conclusions

This paper aimed to draw the main lines of applying reinforcement learning techniques to control FACTS devices in order to damp power system oscillations. The main difficulties were an appropriate definition of the state and the reward function used in the algorithms, two major issues for a good use of RL techniques. The same procedure has been successfully applied to the control of a FACTS device on a much larger power system with a more accurate modelling ([3]). Other types of RL algorithms like prioritized sweeping ([5]) have also been experimented and have shown even better results than the Sarsa(λ) algorithm used here.

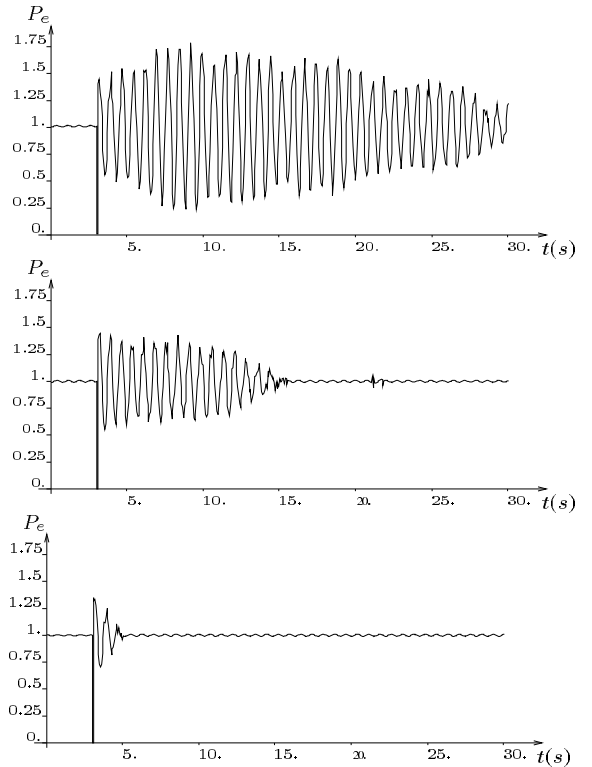


Figure 1: Variation of the electrical power transmitted in the line for different stages of the learning while the power system is subjected to a 50 ms duration short-circuit

References

- [1] Richard Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [2] Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L. Littman. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 1994.
- [3] Ernst D. and Wehenkel L. FACTS devices controlled by means of reinforcement learning algorithms. Submitted to Power System Computation Conference. 2002.
- [4] Steven A. Harp, Sergio Brignone, Bruce F. Woltenberg, and Tariq Samad. SEPIA. A Simulator for Electric Power Industry Agents. *IEEE Control Systems Magazine*, 20(4):53–59, August 2000.
- [5] A.W. Moore and Atkeson C.G. Prioritized Sweeping: Reinforcement Learning with Less Data and Less Real Time. *Machine Learning*, 13:103–130, 1993.
- [6] R. Munos. A study of reinforcement learning in the continuous case by the means of viscosity solutions. *Machine Learning*, 40:265–299, 2000.
- [7] R.S. Sutton and A.G. Barto. *Reinforcement Learning, an introduction*. The MIT Press, 1998.