An Intelligent Assistant for Obtaining the Optimal Policy during Operation Transients in a HRSG

A Reyes B., M. A. Degadillo V., P. H. Ibargüengoytia

Instituto de Investigaciones Eléctricas
Av. Reforma 113 Col. Palmira
C.P. 62490 Temixco, Morelos, México
{areyes, madv, pibar}@iie.org.mx

Abstract. This paper shows how to use Markov Decision Processes to guide the actions of the operation personnel during a transient in a heat recovery steam generator (HRSG). The critical events shown correspond to situations where a strong participation of human operators is required. The aim is to help the steam generation control system to set the plant on a safe operation state. The proposed system also takes into account that either the actuators or control valves in the loop can fail. The possible recommended actions are of parallel nature so that different people can command different control devices to control the situation. A reward function specify the desired and non-desired operation states.

Introduction

The electric power industry is continuously searching for ways to improve the efficiency and reliability with which it supplies energy. Although the fundamental technologies of power generation, transmission and distribution change quite slowly, the power industry has been quick to explore new technologies that might assist its search and eventually to adopt those that show benefits.

This general tendency has held true to form from the various artificial intelligence technologies such as general planners, expert systems, artificial neural networks, inductive learning, fuzzy logic, genetic algorithms, and more. Some researchers have applied every form of AI tool in at least prototype form to one or more problem areas in the power industry, and new practical applications of AI appear with increasing frequency. In some cases, AI tools augment or replace existing technologies. In others AI tools enable solutions to problems previously addressed only by natural intelligence, creating new applications for computers.

A recent example of these applications is COMMAS (Condition Monitoring Multi-Agent System) [Mangina, McArthur, McDonald, Moyes, 2001], a system developed to make easy the task of monitoring the efficiency in an industrial gas turbine. The aim of this work is to improve the accuracy of the current monitoring systems by building small software components which contain partial information about the state
of the plant to monitor. Another application case is a system based on genetic algorithms and Monte Carlo simulation for design purposes. The idea of the system was to optimize the design reliability of an electric power system in their facets of generation, transmission and distribution [Su, Lii; 1999].

There are some other problems in the electric power industry whose complexity also requires uncertainty management and optimization approaches. The family of techniques that allows to deal with this kind of problems is known as Markov Decision Processes (MDP). [Littman, 1996, Boutilier, Dean, Hanks 1995], which bases its strength in the computation of an optimal policy, in an accessible and stochastic environment, and in a state transition model. [Agueda and Ibarguengoytia 2002] shows a planning architecture based on MDPs to deal with the uncertainty that a classical level controller observes in the drum of a conventional steam generator by means of the actions of a single agent.

In this work we extend the problem to deal with uncertain critical disturbances in the whole feedwater and steam generation operation by assuming that more than one action at one time is required to take on different control devices. The main equipment involved such as gas turbines, heat recovery steam generators and the steam turbine must operate simultaneously and in a coordinated way so that if an action is taken on one part of the process other parts can be affected too. The general idea is to find a fast and optimal way to solve discordances when an unexpected fault occurs.

The present work primary explains the problem, some of its instances, and specifies a possible solution. It briefly shows the MDP formalization, the details of the technique implementation to solve the problem, some alternatives to deal with the computational complexity, results found, and advantages and disadvantages of this kind of methods.

1.0 Problem Statement

A Combined Cycle Power Plant (CCPP) allows to recover energy from the Gas Turbine (GT) outlet gases by means of a Heat Recovery Steam Generator (HRSG) which produces steam for a Steam Turbine (ST). In this way electricity is generated in two forms, one with gas turbine generator(s) and the other with a steam turbine generator. A CCPP may include one or more trains of GT-HRSG with usually one steam turbine.

The HRSG (figure 1) is a mayor equipment consisting of several piping banks, each one designed to: (1) rise the feedwater temperature to the drum water temperature (Economizer), (2) get the water enthalpy to produce the amount of steam required (High Pressure Evaporator), superheat the saturated steam to the conditions required by the steam turbine (Superheater). The HRSG also includes a water-steam separator (Steam Drum). This drum is
associated with the feedwater system that includes a deareator, a surge tank and a feedwater pump.

In normal conditions, with the GT at base load, the steam production, and consequently the electric power production, may be increased by about 5% above the base generation. This may be done with additional burners (Afterburners) located in the duct of the GT outlet gases to HRSG.

The final control elements associated to the HRSG are: (1) a control valve to regulate the fuel to the afterburners, (2) the feedwater control valve, and (3) the steam pressure control valve to the ST.

Some CCPP are designed to operate with a coordinated control that embraces the automatic operation of the complete system. In this mode of operation all main equipments may be taken from startup to the base load of the plant. However there are many requirements, and the fact that several equipments may operate simultaneously makes it difficult to maintain the coordinated control in cases of severe disturbances or in cases of fail of any equipment or control instrumentation.

Figure 1. Feedwater and main steam systems simplified diagram.

During normal operation, the conventional feedwater control system commands the feedwater control valve to regulate the steam drum level. However, when severe disturbances take place, this traditional control loop is not longer able to stabilize the drum level. One of this disturbances is a partial or total electric load rejection. Another case is presented when a sudden high steam demand occurs. In this case , the steam-water equilibrium point, which depends on the steam pressure, moves causing an enthalpy change of both fluids (steam and water). Consequently, the enthalpy
change causes an increment in the water level because of a strong water displacement to the steam drum. The three-element control system reacts closing the feedwater control valve. However, a water increase is needed instead of a feedwater decrease. Under these circumstances, the participation of the operation staff is more necessary to help the control system to decide the actions that should be taken in order to overcome the transient.

These operational problems could be overcome by means of an intelligent system that shows recommendations to operators about how to make the best action on the process and correct the problem. The system should be able to find an optimal path according to the crisis dimension, take into account that actuators are not perfect and they could fail to produce non-desired effects. Finally, the system should be conscious that the recommended actions could have parallel nature and be taken by more than one person, either from the control room or locally from the process itself.

**Markov Decision Processes**

An MDP $M = \langle S, A, \Phi, R \rangle$, where $S$ is a finite set of states of the system, $A$ is a finite set of actions, $\Phi: A \times S \to \Pi(S)$ is the state transition function, mapping an action and a state to a probability distribution over $S$ for the possible resulting state. The probability of reaching state $s'$ by performing action $a$ in state $s$ is written $\Phi(a, s, s')$. $R: S \times A \to \mathbb{R}$ is the reward function. $R(s, a)$ is the reward the system receives if it takes action $a$ in state $s$.

A policy for an MDP is a mapping $\pi: S \to A$ that selects an action for each state. Given a policy, we can define its finite-horizon value function $V^\pi_n: S \to \mathbb{R}$, where $V^\pi_0(s)$ is the expected value of applying the policy $\pi$ for $n$ steps starting in state $s$. The value function is defined inductively with $V^\pi_0(s) = R(s, \pi(s))$ and $V^\pi_{n+1}(s) = R(s, \pi(s)) + \sum_{u \in S} \Phi(\pi(s), s, u) V^\pi_n(u)$. Over an infinite horizon, a discounted model is frequently used to ensure policies have a bounded expected value. For some $\beta$ chosen so that $\beta < 1$, the value of any reward from the transition after the next is discounted by a factor of $\beta$, and the one after that $\beta^2$, and so on. Thus, if $V^\pi(s)$ is the discounted expected value in state $s$ following policy $\pi$ forever, we must have $V^\pi(s) = R(s, \pi(s)) + \beta \sum_{u \in S} \Phi(\pi(s), s, u) V^\pi(u)$, which yields a set of linear equations in the values of $V^\pi$.

A solution to an MDP is a policy that maximizes its expected value. For the discounted infinite-horizon case with any given discount factor $\beta$, there is a policy $V^*\pi$ that is optimal regardless of the starting state (Howard 1960) that satisfies the following equation: $V^*(s) = \max_a \{ R(s, a) + \beta \sum_{u \in S} \Phi(a, s, u) V^*(u) \}$.

Two popular methods for solving this equation and finding an optimal policy for an MDP are: (1) value iteration and (2) policy iteration (Puterman 1994).
In policy iteration, the current policy is repeatedly improved by finding some action in each state that has a higher value than the action chosen by the current policy for the state. The policy is initially chosen at random, and the process terminates when no improvement can be found. This process converges to an optimal policy (Puterman 1994).

In value iteration, optimal policies are produced for successively longer finite horizons until they converge. It is relatively simple to find an optimal policy over \( n \) steps \( \pi^*_n(.) \), with value function \( V^*_n(.) \) using the recurrence relation: 
\[
\pi^*_n(s) = \arg \max_a \left( R(s, a) + \beta \sum_{s' \in S} \Phi(a, s, u) V^*_{n-1}(u) \right),
\]
starting condition \( V^*_0(.) = 0 \ \forall \ s \in S \), where \( V^*_n \) is derived from the policy \( \pi^*_m \) as described earlier. The algorithm to the optimal policy for the discounted infinite case in a number of steps that is polynomial in \(|S|, |A|, \log \max_{s,a} |R(s,a)| \) and \( 1/(1-\beta) \).

**System Development**

The solution of an MDP allows to set an optimal action policy in critical situations where the HRSG control system requires the joint interaction of operation personnel in control room and in the field to take the plant to a safe operation condition. The particular case study presented next corresponds to the effects on a HRSG when a high load demand or a load reject suddenly occurs in the electric system.

According to the previous section, the problem formalization as an MDP requires a finite set of states, a finite set of joint actions to apply on the process, an state transition model, a reward function, and a discount factor.

![Recommended curve](image)

**Figure 2. HRSG operation curve in steady conditions**

The set of states is obtained from the discretization of the HRSG operation curve in steady conditions (figure 2), which relates the main steam flow \( (Fms) \) to the drum pressure \( (Pd) \). The set of actions is conformed by the combination of the possible
actions on the feedwater valve (Vfw) and the main steam valve (Vms). It is important to mention that in this general version it is supposed that regulation of the main steam flow by opening or closing the Vfw is possible. It is also supposed that the drum pressure is regulated by opening or closing the Vms, although in both cases there are more control elements involved that were not considered.

The transition model can be obtained from historic data related to the effectiveness with which the actuators and control valves really operate in that part of the process. It is also important to say that the nature of actions remains the same unless a change on the state is recorded. The reward function is assigned by a human expert in the operation of the HRSG. The states with positive reward correspond to those states matching with the normal operation curve. The selection of the discount factor can be a function of the number of discrete states and its value can be in the range of 0 to 1.

Table 1 shows the set of joint actions for this application. PosVms and PosVfw represent increments (+), decrements (-) or null actions (0) for the valves Vms and Vfw, respectively. On the other hand, table 1b shows an example of the transition matrix, where the element $\Phi_{ij}$ represent the probability of transition from state $i$ to state $j$ given the action $a_j$.

Table 1 shows that the actions are joint and require to operate simultaneously on different control devices. It is possible to see in the transition model that the higher probabilities correspond to the desired states and the rest of them decrease as the effect of the joint action locates the plant away from the new state.
Result Analysis

The intelligent system was implemented using Java2 by Sun Microsystems and tested through different cases of reward values and changes in the transition model. In this application, the state space in the HRSG was partitioned as a 3x4, 8x6 and 8x8 matrix to have a total of 12, 48 and 64 states respectively. Some tests were also performed using sets of 5 and 9 actions. In each case, the reward function awarded the states matching the operation curve and punished the others. To approximate the value function, the value iteration algorithm was used, which converged easily. Figure 3 shows the representation of the problem after obtaining the optimal policy in one of the run cases.

Figure 3 shows that the value iteration algorithm effectively found the optimal policy, which is denoted in red color. The first number of each state represents the utility value in the last iteration. The central value is the state reward value. The number in the left lower corner of each state is an identifier denoting the state itself. In the first run test (12 states), the algorithm converged in 6 and 7 iterations for sets of 9 and 5 actions respectively. In the second case the solution was reached in 10 and 15 iterations, and in the 64-state case the MDP was solved in 11 iterations for 9 actions, and 17 iterations for 5 actions. It can be concluded that, the more actions are available, the easier it is to find the solution, although in small state spaces it is not so difficult and the number of actions is not very important. In this application, the
probability of success in the operation of a control device was 70%, and the remaining 30% was the probability of failure.

With the idea of showing the utilization of the obtained policy, there is a section in the graphical user interface to enter data corresponding to the drum pressure and the main steam flow, and there are also two fields to show the resulting state and the recommended joint action for that state. In the run example, the drum pressure is rising to the upper limit due to a load reject, however, despite the effort of the control system to regulate the feedwater supply to the drum, the situation is critical and the participation of operation personnel is required. The system then suggests to open the steam valve in order to reduce the pressure in the drum, and to open the feedwater valve to reach the state 6 corresponding to the safe operation.

Conclusions

The main features of an intelligent system to assist operators in the treatment of transients in a HRSG were described in this work. The system suggests complementary or supplementary actions to the control system in order to optimally reach a normal operation through joint actions. The power of representing a system as an MDP and its resolution by using the value iteration algorithm was also shown. The problem domain was described, as well as the formalization of an MDP and its application to solve the problem. The results obtained were also commented through six test cases using the system.

Some advantages of using Markov Models are the efficient way to determine an optimal policy in small state spaces. This representation also allows to consider the implicit uncertainty in the operation of control valves, and the use of joint actions to correct the transient. However, other components not considered in this demonstration that complicate the problem solution should be taken into account. Consequently, it will be necessary to include other complementary techniques to factorize the state and action space, and the value function, since the computational complexity can be so high that the problem could become intractable. Some authors suggest to use methods of local optimization through the global selection of the joint action. In this technique, the concept of coordination graph is used to reduce the number of elements that participate in the process of the utility maximization [Guestrin, Koller, Parr; 2002].

The next step will be the implementation of the MDP factorization techniques and the test of the whole system in a combined power plant simulator. After that, it will be possible to compare the control system performance versus the recommendations presented to the plant operators by the intelligent system.
References

6. Guestrin, Koller, Parr. 2002 ; Multiagent Planning with Factored MDPs; Stanford and Duke Universities.