Analyses for Optimal Control of Discrete Time-Delay Systems Based on ADP Algorithm with Finite-Horizon Performance Index

Ruizhuo Song, Shi Xing

1. School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, 100083, China

Abstract: An optimal control scheme is obtained for a class of nonlinear discrete time-delay systems in this paper. Based on adaptive finite-horizon dynamic programming (ADP) algorithm, we focus on two different initial control conditions, and then the value iteration method is used to get the iterative performance index function and iterative control. Two neural networks, which are named as critic network and action network are used to approximate the performance index function and compute the optimal control policy. The iteration termination error and the performance index function are analyzed based on different initial states. In the simulation study, we analysis the control results of different initial states and different maximal iterative steps.

Key Words: Adaptive dynamic programming, optimal control, nonlinear systems, approximate dynamic programming

1 INTRODUCTION

If delays are contained in control systems, the control efficiency will be degradation or even the systems are instability [1]. So many works in various research areas are proposed for systems with time delays, such as electrical systems, chemical engineering systems and networked control systems [2]. We know that the controllability of linear time-delay systems are studied in [3, 4]. But, since the nonlinear systems are too complex to design the optimal control. Therefore, we want to develop a new method to obtain the optimal control of time-delay systems based on finite ADP algorithm.

Adaptive dynamic programming (ADP) is used to solve the optimal Hamilton-Jacobi-Bellman (HJB) equation for nonlinear systems [5–7], which is often unable to commutate to use dynamic programming [8, 9]. Up to now, ADP algorithms are successfully applied in multi-objective optimal control problems [10], self-learning optimal control problems [11], multi-battery coordination control problems [12, 13], optimal control problems with finite approximation errors [14], etal. Note that, in [15], based on a novel HDP algorithm for a class of nonlinear discrete-time systems with time delays, an optimal tracking controller was proposed. In [16], according to discrete-time nonlinear systems ADP algorithm, a finite-horizon optimal control was solved. Above the research results of [16], optimal control problems for nonlinear systems are discussed in [17, 18] by time-delay HJB equations. In [19], neurocognitive psychology is used to deveiop a novel controller based on multiple actor-critic structures for unknown systems. This controller traded off fast actions based on stored behavior patterns with real-time exploration using current input-output data.

In this paper the optimal controller is designed based on the original time-delay systems, directly. Consider the two different initial control conditions, two iterative processed are established. They are have iterative control, iterative state and iterative performance index. Neural networks are used to get the approximate performance index function and the control policy. We emphasis the control effects by the different initial states and different maximal iterative steps.

This paper is organized as follows. In Section 2, the problem formulation is presented. In Section 3, the nearly finite-horizon optimal control scheme is developed based on iteration ADP algorithm. In section 4, the convergence proof is given. In Section 5, two examples are given to demonstrate the effectiveness of the proposed control scheme. In Section 6, the conclusion is drawn.

2 Problem Statement

Consider a class of deterministic time-delay nonaffine nonlinear systems

$$\begin{aligned} s(t+1) =&\mathcal{F}(s(t), s(t-l_1), \cdots, s(t-l_L), h(t)), \\ s(t) =&\chi(t),\ -l_L \le t \le 0 \end{aligned} \tag{1}$$

where $s(t) \in \Re^n$ is state and $s(t-l_1), \cdots, s(t-l_L) \in \Re^n$ are time-delay states. $h(t) \in \Re^m$ is the system input. $\chi(t)$ is the initial state, l_i, $i = 1, 2, \cdots, L$, is the time delay, set $0 < l_1 < l_2 < \cdots < l_L$, and they are nonnegative integer numbers. $\mathcal{F}(s(t), s(t-l_1), s(t-l_2), \cdots, s(t-l_L), h(t))$ is known function. $\mathcal{F}(0, 0, \cdots, 0) = 0$.

For any time step k, the performance index function for state $s(q)$ under the control sequence $U(q, N+q-1) = (h(q), h(q+1), \cdots, h(N+q-1))$ is defined as

$$\begin{aligned} P(s(q), U(q, N+q-1)) = \sum\nolimits_{j=q}^{N+q-1} \{ & s^T(j)Bs(j) \\ & + h^T(j)Rh(j)\}, \end{aligned} \tag{2}$$

where B and R are positive definite constant matrixes.

According to the theory of dynamics programming, the optimal performance index function is defined as

$$P^*(s(q)) = \min_{U(q,N+q-1)} P(s(q), U(q, N+q-1)) \\ = \min_{h(q)} \left\{ s^T(q)Bs(q) + h^T(q)Rh(q) + P^*(s(q+1)) \right\}, \quad (3)$$

and the optimal control policy is

$$h^*(q) = \arg\min_{h(q)} \left\{ s^T(q)Bs(q) + h^T(q)Rh(q) + P^*(s(q+1)) \right\}, \quad (4)$$

so the state under the optimal control policy is

$$s^*(t+1) = \mathcal{F}(s^*(t), s^*(t-l_1), \cdots, s^*(t-l_L), h^*(t)), t = 0, 1, \cdots, k, \cdots, \quad (5)$$

and then, the HJB equation is written as

$$P^*(s^*(q)) = P(s^*(q), U^*(q, N+q-1)) \\ = (s^*(q))^T Bs^*(q) + (h^*(q))^T Rh^*(q) + P^*(s^*(q+1)). \quad (6)$$

3 ADP Iteration Algorithm with Different Initial Situations

In this section, we will give the novel iteration ADP algorithm in detail. For the state $s(q)$ of system (1), there exists two cases. Case 1: $\exists U(q,k)$ which makes $s(q+1) = 0$. Case 2: $\exists U(q, q+m), m > 0$, which makes $s(q+m+1) = 0$. In the following part, we will discuss the two cases, respectively.

There exists finite-horizon admissible control sequence $U(q, q+m) = (\beta(q), \cdots, \beta(q+m))$ which makes $s_f(s(q), U(q, q+m)) = 0$. We suppose that for $s(q+1)$, there exists optimal control sequence $U^*(q+1, q+j-1) = (h^*(q+1), h^*(q+2), \cdots, h^*(q+j-1))$. For time step k, the iteration ADP algorithm between

$$h^{[1]}(q) = arg \min_{h(q)} \{ s^T(q)Bs(q) + h^T(q)Rh(q) + \Theta^{[0]}(s(q+1)) \}, \quad (7)$$

and

$$\Theta^{[1]}(s^{[1]}(q)) = (s^{[1]}(q))^T Bs^{[1]}(q) + (h^{[1]}(q))^T Rh^{[1]}(q) \\ \Theta^{[0]}(s^{[0]}(q+1)), \quad (8)$$

where $\forall s(q+1)$, $\Theta^{[0]}(s(q+1))$ in (7) is obtained as

$$\Theta^{[0]}(s(q+1)) = P(s(q+1), U^*(q+1, q+j-1) \\ = P^*(s(q+1)). \quad (9)$$

In (8), $\Theta^{[0]}(s^{[0]}(q+1))$ is obtained by the similar Equation (9). The states in (8) are obtained as

$$s^{[1]}(t+1) = \mathcal{F}(s^{[1]}(t), s^{[1]}(t-l_1), s^{[1]}(t-l_2), \cdots, s^{[1]}(t-l_L), h^{[1]}(t)), \quad (10)$$

where $t = 0, 1, \cdots, q-1$,
and

$$s^{[0]}(t+1) = \mathcal{F}(s^{[1]}(t), s^{[1]}(t-l_1), s^{[1]}(t-l_2), \cdots, s^{[1]}(t-l_L), h^{[1]}(t)), \quad (11)$$

where $t = q, q+1, \cdots$
For the iteration step $\gamma = 1, 2, \cdots$, the iteration algorithm will be implemented as follows:

$$h^{[\gamma+1]}(q) = arg \min_{h(q)} \{ s^T(q)Bs(q) + h^T(q)Rh(q) + \Theta^{[\gamma]}(s(q+1)) \}, \quad (12)$$

and

$$\Theta^{[\gamma+1]}(s^{[\gamma+1]}(q)) = (s^{[\gamma+1]}(q))^T Bs^{[\gamma+1]}(q) + (h^{[\gamma+1]}(q))^T Rh^{[\gamma+1]}(q) + \Theta^{[\gamma]}(s^{[\gamma]}(q+1)), \quad (13)$$

where $\Theta^{[\gamma]}(s(q+1))$ in (12) is updated as

$$\Theta^{[\gamma]}(s(q+1)) = \min_{h(q+1)} \{ s^T(q+1)Bs(q+1) + h^T(q+1)Ru(q+1) + \Theta^{[\gamma-1]}(s(q+2)) \}, \quad (14)$$

and the states in (13) are obtained as

$$s^{[\gamma+1]}(t+1) = \mathcal{F}(s^{[\gamma+1]}(t), s^{[\gamma+1]}(t-l_1), s^{[\gamma+1]}(t-l_2), \cdots, s^{[\gamma+1]}(t-l_L), h^{[\gamma+1]}(t)), \quad (15)$$

where $t = 0, 1, \cdots, q-1$, and

$$s^{[\gamma]}(t+1) = \mathcal{F}(s^{[\gamma+1]}(t), s^{[\gamma+1]}(t-l_1), s^{[\gamma+1]}(t-l_2), \cdots, s^{[\gamma+1]}(t-l_L), h^{[\gamma+1]}(t)), \quad (16)$$

where $t = q, q+1, \cdots$.
This completes the iteration algorithm. From the above algorithm we can see that, if $\Theta^{[0]} = 0$ in (8), then Case 1 is a special one of Case 2.

4 Convergence Analyses

Theorem 1 *For system (1), the states of the system are driven by a given initial state $\chi(t), -l_L \leq t \leq 0$, and the initial finite-horizon admissible control policy $\beta(t)$. The iteration algorithm is as in (7)-(16). For time step k, $\forall$ $s(q)$ and $U(q, q+\gamma)$, we define$U(q, q+\gamma) = (h^{[\gamma+1]}(q), h^{[\gamma]}(q+1), \cdots h^{[1]}(q+\gamma))$. We define*

$$\Omega^{[\gamma+1]}(s(q), U(q, q+\gamma)) \\ = s^T(q)Bs(q) + h^T(q)Ru(q) + s^T(q+1)Bs(q+1) + h^T(q+1)Ru(q+1) + \cdots \\ + s^T(q+\gamma)Bs(q+\gamma) + h^T(q+\gamma)Ru(q+\gamma) + \Theta^{[0]}(s(q+\gamma+1)), \quad (17)$$

where $\Theta^{[0]}(s(q+\gamma+1))$ as in (9) and $\Theta^{[\gamma+1]}(s(q))$ is updated as (14). Then we have

$$\Theta^{[\gamma+1]}(s(q)) = \min_{U(q,q+\gamma)} \Omega^{[\gamma+1]}(s(q), U(q, q+\gamma)). \tag{18}$$

Proof:
From (14) we have

$$\begin{aligned}
&\Theta^{[\gamma+1]}(s(q)) \\
=&\min_{h(q)}\{s^T(q)Bs(q) + h^T(q)Rh(q) \\
&+ \min_{h(q+1)}\{s^T(q+1)Bs(q+1) + h^T(q+1)Rh(q+1) \\
&+ \cdots + \min_{h(q+\gamma)}\{s^T(q+\gamma)Bs(q+\gamma) \\
&+ h^T(q+\gamma)Rh(q+\gamma)\} \\
&+ \Theta^{[0]}(s(q+\gamma+1))\}\cdots\}\}.
\end{aligned} \tag{19}$$

So we can further obtain

$$\begin{aligned}
&\Theta^{[\gamma+1]}(s(q)) \\
=&\min_{U(q,q+\gamma)}\{s^T(q)Bs(q) + h^T(q)Rh(q) \\
&+ s^T(q+1)Bs(q+1) + h^T(q+1)Rh(q+1) \\
&+ \cdots + s^T(q+\gamma)Bs(q+\gamma) \\
&+ h^T(q+\gamma)Rh(q+\gamma) + \Theta^{[0]}(s(q+\gamma+1))\},
\end{aligned} \tag{20}$$

Thus we can have

$$\Theta^{[\gamma+1]}(s(q)) = \min_{U(q,q+\gamma)} \Omega^{[\gamma+1]}(s(q), U(q, q+\gamma)). \tag{21}$$

Based on Theorem 1, we give the monotonicity theorem about the sequence of performance index functions $\Theta^{[\gamma+1]}(s^{[\gamma+1]}(q)), \forall s^{[\gamma+1]}(q)$.

Theorem 2 *For system (1), let the iteration algorithm be as in (7)-(16). Then we have $\Theta^{[\gamma+1]}(s^{[\gamma]}(q)) \leq \Theta^{[\gamma]}(s^{[\gamma]}(q))$, $\forall\gamma > 0$, for Case 1; $\Theta^{[\gamma+1]}(s^{[\gamma]}(q)) \leq \Theta^{[\gamma]}(s^{[\gamma]}(q))$, $\forall\gamma \geq 0$, for Case 2.*

Proof: We first give the proof for Case 2. Define $\hat{U}(q, q+\gamma) = (h^{[\gamma]}(q), \cdots, h^{[1]}(q+\gamma-1), h^*(q+\gamma))$, then according to the definition of $\Omega^{[\gamma+1]}(s(q), \hat{U}(q, k+\gamma))$ in (17), we have

$$\begin{aligned}
&\Omega^{[\gamma+1]}(s(q), \hat{U}(q, q+\gamma)) \\
=&s^T(q)Bs(q) + (h^{[\gamma]}(q))^T Rh^{[\gamma]}(q) \\
&+ \cdots + s^T(q+\gamma-1)Bs(q+\gamma-1) \\
&+ (h^{[1]}(q+\gamma-1))^T Rh^{[1]}(q+\gamma-1) \\
&+ s^T(q+\gamma)Bs(q+\gamma) + (h^*(q+\gamma))^T Rh^*(q+\gamma) \\
&+ \Theta^{[0]}(s(q+\gamma+1)).
\end{aligned} \tag{22}$$

From (9), we get

$$\begin{aligned}
&\Theta^{[0]}(s(q+\gamma)) \\
=&P^*(s(q+\gamma)) \\
=&s^T(q+\gamma)Bs(q+\gamma) + (h^*(q+\gamma))^T Rh^*(q+\gamma) \\
&+ P^*(s(q+\gamma+1)) \\
=&s^T(q+\gamma)Bs(q+\gamma) + (h^*(q+\gamma))^T Rh^*(q+\gamma) \\
&+ \Theta^{[0]}(s(q+\gamma+1)).
\end{aligned} \tag{23}$$

On the other side, from (14), we have

$$\begin{aligned}
&\Theta^{[\gamma]}(s(q)) \\
=&s^T(q)Bs(q) + (h^{[\gamma]}(q))^T Rh^{[\gamma]}(q) \\
&+ \cdots + s^T(q+\gamma-1)Bs(q+\gamma-1) \\
&+ (h^{[1]}(q+\gamma-1))^T Rh^{[1]}(q+i-1) \\
&+ \Theta^{[0]}(s(q+\gamma)).
\end{aligned} \tag{24}$$

So according to (23), we obtain

$$\Omega^{[\gamma+1]}(s(q), \hat{U}(q, q+i)) = \Theta^{[\gamma]}(s(q)). \tag{25}$$

From Theorem 1, we can get

$$\Theta^{[\gamma+1]}(s(q)) \leq \Omega^{[\gamma+1]}(s(q), \hat{U}(q, q+\gamma)). \tag{26}$$

So we have $\forall s(q)$,

$$\Theta^{[\gamma+1]}(s(q)) \leq \Theta^{[\gamma]}(s(q)), \tag{27}$$

i.e., for $s^{[\gamma]}(q)$,

$$\Theta^{[\gamma+1]}(s^{[\gamma]}(q)) \leq \Theta^{[\gamma]}(s^{[\gamma]}(q)). \tag{28}$$

For Case 1, we set $\Theta^{[0]} = 0$, the proof is similar with Case 2.
From the two theorem, we can see that the proposed algorithm is convergent. For implementation of the algorithm, two neural networks, which are critic and action networks, are used to approximate the performance index function and control policy, respectively. The detail approximate process can be seen in [20–22].

5 Simulation study

We take the example in [16] with modification:

$$s(t+1) = s(t-2) + sin(0.1s^2(t) + h(t)). \tag{29}$$

We give the initial states as $\chi_1(-2) = \chi_1(-1) = \chi_1(0) = 1.5$, and the initial control policy as $\beta(t) = sin^{-1}(s(t+1) - s(t-2)) - 0.1s^2(t)$. We implement the proposed algorithm at the time instant $k = 3$.
First, according to the initial control policy $\beta(t) = sin^{-1}(s(t+1) - s(t-2)) - 0.1s^2(t)$ of system (29), we give fist group of state data: $s(1) = 0.8$, $s(2) = 0.7$, $s(3) = 0.5$, $s(4) = 0$. We also can get the second group of state data: $s(1) = 1.4$, $s(2) = 1.2$, $s(3) = 1.1$, $s(4) = 0.8$, $s(5) = 0.7$, $s(6) = 0.5$, $s(7) = 0$. Obviously, for the first sequences of states we can get the optimal controller by Case 1 of the proposed algorithm. For the second

one, the optimal controller can be obtained by Case 2 of the proposed algorithm, and the optimal control sequence $U^o(q+1, q+j+1)$ can be obtained in the first group of state data. We select $B = R = 1$.

The three-layer BP neural networks are used to approach the critic network and the action network with the structure $2-8-1$ and $6-8-2$, respectively. The iteration times of the weights updating for two neural networks are 200. The initial weights are chosen randomly from $[-0.1, 0.1]$, and the learning rates are $\alpha_a = \alpha_c = 0.05$. We select maximal iteration step γ_{max} is 5, 15, 25 and 50 to analyse the effect of the iteration times.

Under the same neural network weights, we get $P^*(s^*(3))$ for four different γ_{max} in Table. 1, and iteration termination error ε in Table. 2. From Table 1, we can see that $P^*(s^*(3))$ decreases as γ_{max} increases. It has been proved in Theorem 2. According to Table 2, the iteration termination error is smaller when we choose larger γ_{max}. When $i \to \infty$, the optimal controller will be reached. The performance index trajectories for the first and second state data group show in Fig. 1 and Fig. 2, respectively. According to Theorem 2, for the first state group, the performance index is decreasing as $\gamma > 0$. For the second state group, the performance index is decreasing as $\gamma \geq 0$. The state trajectory and the control trajectory of the second state data are shown in Fig. 3 and Fig. 4. From the figures, we can see that the system is asymptotically stable no mater what the initial states are selected. The simulation study shows the new iteration ADP algorithm is very feasible.

State	$\gamma_{max} =$ 5	$\gamma_{max} =$ 15	$\gamma_{max} =$ 25	$\gamma_{max} = 50$
$x = 0.5$	1.2339	1.2338	1.2338	1.2338
$x = 1.1$	5.5173	5.4075	5.4048	5.4048

Table 1: $P^*(s^*(3))$

State	$\gamma_{max} =$ 5	$\gamma_{max} =$ 15	$\gamma_{max} =$ 25	$\gamma_{max} =$ 50
$s = 0.5$	$1.1287e-004$	$5.5713e-012$	0	0
$s = 1.1$	0.0776	0.0019	$4.7227e-005$	$4.4787e-009$

Table 2: iteration termination error ε

6 Conclusion

This paper proposed a novel ADP algorithm to deal with the finite-horizon optimal control for a class of deterministic nonaffine time-delay nonlinear systems. For determining the optimal state, the state updating was contained in the novel ADP algorithm. The results of theorems showed the proposed iteration algorithm was convergent. The simulation study has demonstrated the effectiveness of the proposed control algorithm.

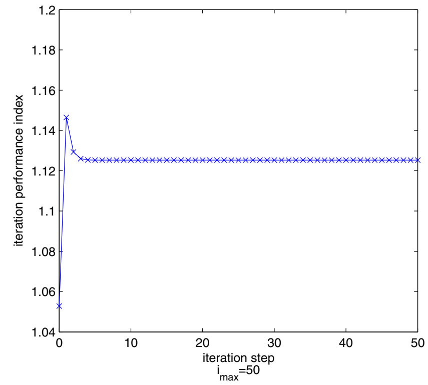

Figure 1: The performance trajectory for $s(3) = 0.5$ and $\gamma_{max} = 50$.

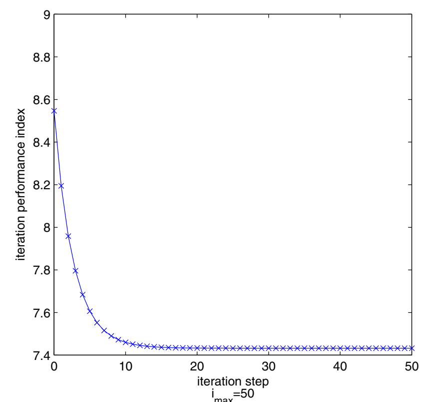

Figure 2: The performance trajectory for $s(3) = 1.1$ and $\gamma_{max} = 50$.

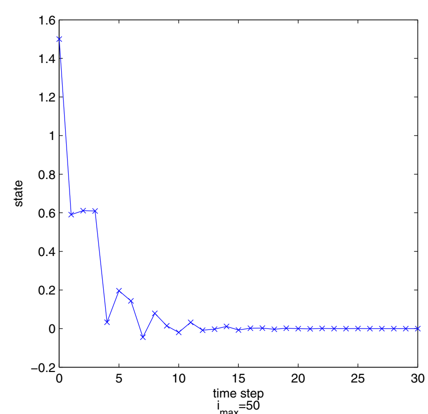

Figure 3: The state trajectory for the second state data and $\gamma_{max} = 50$.

REFERENCES

[1] K. Gu, V. Kharitonov, and J. Chen, *Stability of Time-Delay Systems*, Birkhäuser, Boston, 2003.

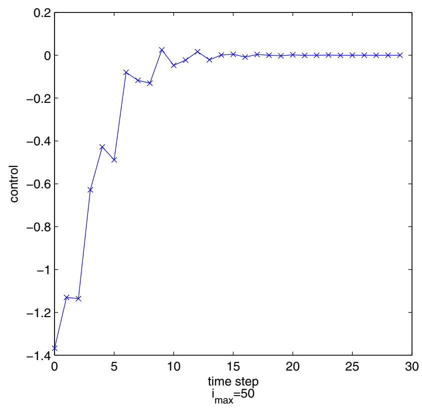

Figure 4: The control trajectory for the second state data and $\gamma_{max} = 50$.

[2] R. Song, H. Zhang, Y. Luo, and Q. Wei, "Optimal control laws for time-delay systems with saturating actuators based on heuristic dynamic programming," *Neurocomputing*, vol. 73, no. 16-18, pp. 3020-3027, 2010.

[3] D. Chyung, "Discrete optimal systems with time delay," *IEEE Transactions on Automatic Control*, Vol. 13, No. 1, pp. 117, 1968.

[4] D. Chyung and E. Lee, "Linear optimal systems with time delays," *SIAM J. Control.*, Vol. 4, No. 3, Nov. 1966.

[5] Q. Wei and D. Liu, "Neural-network-based adaptive optimal tracking control scheme for discrete-time nonlinear systems with approximation errors", *Neurocomputing*, Vol. 149, No. A, pp. 106-115, 2015.

[6] R. Song, W. Xiao, Q. Wei, and C. Sun, "Neural-Network-Based Approach to Finite-Time Optimal Control for a Class of Unknown Nonlinear Systems", *Soft Computing*, vol. 18, no. 8, pp. 1645-1653, Aug. 2014.

[7] R Song, W. Xiao, H. Zhang, and C. Sun, "Adaptive Dynamic Programming for a Class of Complex-Valued Nonlinear Systems", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1733-1739, Sep. 2014.

[8] R. Bellman, "Dynamic programming," Princeton, New Jersey: Princeton University Press, 1957.

[9] P. Werbos, "Advanced forecasting methods for global crisis warning and models of intelligence," *General Systems Yearbook*, Vol. 22, pp. 25-38, 1977.

[10] R. Song, W. Xiao, and H. Zhang, "Multi-objective optimal control for a class of unknown nonlinear systems based on finite-approximation-error ADP algorithm," *Neurocomputing*, vol. 119, no. 7, pp. 212-221, 2013.

[11] R. Song, W. Xiao, and C. Sun, "A new self-learning optimal control laws for a class of discrete-time nonlinear systems based on ESN architecture", *Science China (Information Sciences)*, vol. 57, no. 6, pp.1-10, 2014.

[12] Q. Wei, D. Liu, G. Shi, and Y. Liu, "Optimal multi-battery coordination control for home energy management systems via distributed iterative adaptive dynamic programming", *IEEE Transactions on Industrial Electronics*, Vol. 62, No. 7, pp. 4203-4214, 2015.

[13] Q. Wei, D. Liu, and G. Shi, "A novel dual iterative Q-learning method for optimal battery management in smart residential environments", *IEEE Transactions on Industrial Electronics*, Vol. 62, No. 4, pp. 2509-2518, 2015.

[14] Q. Wei, F. Wang, D. Liu, and X. Yang, "Finite-approximation-error based discrete-time iterative adaptive dynamic programming," IEEE Transactions on Cybernetics, vol. 44, no. 12, pp. 2820-2833, Dec. 2014.

[15] H. Zhang, R. Song, Q. Wei, and T. Zhang, "Optimal Tracking Control for a Class of Nonlinear Discrete-time Systems with Time Delays Based on Heuristic Dynamic Programming," *IEEE Transactions on Neural Networks*, Vol. 22. No. 12, pp. 1851-1862, 2011.

[16] F. Wang, N. Jin, D. Liu, and Q. Wei, "Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ϵ-error bound," *IEEE Transactions on Neural Networks*, Vol. 22, pp. 24-36, 2011.

[17] R. Song, H. Zhang, "The finite horizon optimal control for a class of time-delay affine nonlinear system," *Neural Computing and Applications*, vol. 22, no. 2, pp. 229-235, Feb. 2013.

[18] R. Song, Q. Wei, Q. Sun, "Nearly finite-horizon optimal control for a class of nonaffine time-delay nonlinear systems," Neurocomputing, vol. 156, no. 25, pp. 166-175, May 2015.

[19] R. Song, F. Lewis, Q. Wei, H. Zhang, Z. Jiang, D. Levine, "Multiple actor-critic structures for continuous-time optimal control using input-output data", *IEEE Transactions on Neural Networks and Learning Systems*, vol.26, no. 4, pp. 851-865, April 2015.

[20] Q. Wei and D. Liu, "Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification", *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 4, pp. 1020-1036, Oct. 2014.

[21] Q. Wei and D. Liu, "Neural-network-based adaptive optimal tracking control scheme for discrete-time nonlinear systems with approximation errors", *Neurocomputing*, vol. 149, no. 3 pp. 106-115, Feb. 2015.

[22] R. Song, W. Xiao, C. Sun, and Q. Wei, "Approximation-Error-ADP-Based Optimal Tracking Control for Chaotic Systems with Convergence Proof", *Chinese Physics B*, vol.22, no.9, pp. 090502, Sep. 2013.