
Rong Yu, Member, IEEE, Weifeng Zhong, Student Member, IEEE, Shengli Xie, Senior Member, IEEE, Yan Zhang, Senior Member, IEEE, and Yun Zhang

Abstract—As the next-generation power grid, smart grid will be integrated with a variety of novel communication technologies to support the explosive data traffic and the diverse requirements of quality of service (QoS). Cognitive radio (CR), which has the favorable ability to improve the spectrum utilization, provides an efficient and reliable solution for smart grid communications networks. In this paper, we study the QoS differential scheduling problem in the CR-based smart grid communications networks. The scheduler is responsible for managing the spectrum resources and arranging the data transmissions of smart grid users (SGUs). To guarantee the differential QoS, the SGUs are assigned to have different priorities according to their roles and their current situations in the smart grid. Based on the QoS-aware priority policy, the scheduler adjusts the channels allocation to minimize the transmission delay of SGUs. The entire transmission scheduling problem is formulated as a semi-Markov decision process and solved by the methodology of adaptive dynamic programming. A heuristic dynamic programming (HDP) architecture is established for the scheduling problem. By the online network training, the HDP can learn from the activities of primary users and SGUs, and adjust the scheduling decision to achieve the purpose of transmission delay minimization. Simulation results illustrate that the proposed priority policy ensures the low transmission delay of high priority SGUs. In addition, the emergency data transmission delay is also reduced to a significantly low level, guaranteeing the differential QoS in smart grid.

Index Terms—Adaptive dynamic programming (ADP), cognitive radio (CR), quality of service (QoS), smart grid communications network, transmission scheduling.

I. INTRODUCTION

The traditional power grid is a centrally controlled system with unidirectional information collection and unidirectional electricity delivery, which causes many problems and challenges, such as energy wasting, pollution, electricity demand prediction, integration of distributed energy resources, and security [1]–[3]. Therefore, the evolution of power grid has become an urgent worldwide concern. The smart grid is regarded as the next-generation power grid, providing with renewable energy, intelligent control, high efficiency, and high reliability [2], [4]–[6]. In order to implement the desirable functions of smart grid, novel information technologies should be integrated, including embedded sensing, wireless communications, pervasive computing, and adaptive control [7]–[9].

To balance the energy supply and demand, efficient two-way communications between the customers and the utilities are required for information exchange in real time [10]. The huge amount of data of control commands, monitoring, and smart meters reading will be transmitted through the smart grid communications networks [11]. To cover a large number of network nodes in a wide area, wireless communications is a preferred option for the smart grid communications networks. However, the free industrial, scientific and medical (ISM) frequency band has become crowded without the guarantee of quality of service (QoS). Purchasing licensed frequency bands will increase the burden of utilities [12]. Meanwhile, other licensed frequency bands are used in a fixed and an inefficient way [13]. For all these challenges, cognitive radio (CR), as a promising technology, improves the spectrum resource utilization and provides considerable bandwidth of large-scale data transmission.

The CR technology allows two radio systems, i.e., the primary system and the secondary system, to share the same frequency band. Primary system has the license and priority to use the legal spectrum. Secondary system has no priority to use the spectrum, but could opportunistically access to the spectrum when they are not occupied by the primary system. The user in the primary system is called as primary user (PU) and the user in the secondary system is called as secondary user. By introducing CR technology, the smart grid users (SGUs) can use potentially all available spectrum resources and improve the spectrum utilization, supporting the huge data traffic in smart grid.

In the study of the CR-based smart grid communications networks [14]–[17], spectrum access schemes are designed to optimize the QoS in smart grid. Different services in smart grid have different QoS requirements. For example,
the substations transform the high-voltage electricity into the low-voltage electricity and distribute it to the neighborhood distribution grid. The SGUs, who require high QoS, like substations, should use the available spectrum for vital data transmission with high priority. The priority-based optimization in CR-based smart grid communications networks is studied in [18]. The meter reading service requires low QoS in smart grid, since there is no need to upload the energy consumption data in real time. However, more than meter reading, the household smart meter also can detect and report faults. If there is an emergency, e.g., device damage, the smart meter has to report the emergency, which requires rather high QoS for the reliability of smart grid. To guarantee the differential QoS, we readjust the priorities of SGUs according to their situations. The available channels are allocated to the SGUs based on their priorities.

In this paper, we investigate the transmission scheduling of SGUs in CR-based smart grid communications network satisfying the potential requirement of differential QoS. The main contributions of this paper are as follows.

1) A QoS differentiation policy with diverse priorities is devised. First, the SGUs are divided into different classes based on their roles in the smart grid. The higher class SGUs have higher priority to use the available spectrum resources for ensuring their QoS. Second, to enhance the reliability of smart grid, an SGU is allowed to upgrade its priority whenever it has emergency information to report, even if it belongs to the lowest class. As a consequence, the priorities of SGUs jointly depend on their roles and the current situations.

2) The scheduler will manage spectrum resources and allocate channels to SGUs based on their current priorities, minimizing the long-term transmission delay. The entire transmission scheduling process is slotted and formulated as the semi-Markov decision process (SMDP). The methodology of adaptive dynamic programming (ADP) is employed to estimate the long-term transmission delay (corresponding to system cost) by learning from the PUs’ and SGUs’ activities. Through tuning the scheduling decision, the long-term delay is minimized.

3) Extensive simulation experiments are carried out to evaluate the proposed QoS differential scheduling scheme. The numeric results show that the proposed priority policy can keep the low transmission delay for high class SGUs. Furthermore, the significantly low transmission delay could be achieved when emergency events occur to SGUs. This fact verifies that the proposed scheduling scheme is able to provide satisfying differential QoS.

The remainder of this paper is organized as follows. The smart grid communications network structure and the QoS differential transmission scheduling scheme are introduced in Section II. In Section III, the transmission scheduling problem is formulated within the framework of SMDP. Section IV discusses the ADP solution for the formulated SMDP problem. The simulation results are shown in Section V. Finally, the conclusion is given in Section VI.
to allocate channels (spectrum resources) to the SGUs. The utility providers will purchase licensed spectrum bands to provide hard QoS for smart grid communications network, and rent other spectrum bands from primary systems, since there are large-scale energy consumption data transmission from countless nodes. In this paper, we discuss the transmission scheduling in the case of unlicensed channels. This means that the scheduler allocates the limited unlicensed channels to SGUs, then the SGUs, as the unlicensed users, perform data transmission without interfering the PUs. Once the PUs, as the licensed users, appear, the scheduler will withdraw the corresponding channels from SGUs, leading to the transmission delay of SGUs.

The transmission delay of SGUs may be also caused by spectrum sensing. SGUs may spend specific time for channel detection. The development of the IEEE 802.22 WRAN standard proposes the spectral awareness by the geolocation/database methods [20]. Hence, we consider that the spectrum manager (scheduler) will obtain the information of the spectrum resources from the spectrum database without performing local sensing. The scheduler has the full knowledge of PU occupation, and allocates the channels to SGUs efficiently. Therefore, the transmission delay of SGUs is mainly caused by the arrivals and departures of PUs. Meanwhile, the channel availability for SGUs is also determined by the PUs’ activities in this paper.

In smart grid communications networks, most nodes are static, unlike the mobile networks. Thus, we consider that the condition of the wireless links between SGUs and access points are static, and the transmission rate is stable and constant in the scheduling process. For the SGUs, the regular data packets will be generated from time-to-time. The packets will stack in the buffer following the first-in-first-out (FIFO) rule and wait for transmission scheduling. However, the emergency data packets will be generated whenever they are needed. The emergency packets will directly move to the head of the buffer queue and have higher priority for transmission than regular packets. Thus, the buffer queue system is a priority-based FIFO system.

B. QoS Differential Transmission Scheduling

The SGUs are categorized into different classes based on their roles in NAN. Fig. 1 shows two classes of SGUs, c1 and c2. The SGUs, who provide vital messages of control, protection, management, belong to the higher class (c1). Those, who carry meter reading or transmit delay-tolerant data, belong to the lower class (c2). This means that the data transmission of low-class SGUs always lags behind that of high-class SGUs. However, when some emergencies happen, such as the damage or regular hard check of devices, the smart meters should have high priority for transmission to report the emergencies. Therefore, emergency priority of SGUs should be beyond the SGU class. But the SGUs of c1 still has higher priority than the SGUs of c2, when both of them have emergencies.

For the SGUs satisfaction, interruption during data transmission is more unexpected than being blocked occasionally on a new transmission. An interrupted SGU will be more eager to finish its interrupted transmission than a newly arrived SGU. Therefore, in the same class, the interrupted SGUs should have higher priority than the newly arrived SGUs, decreasing the transmission delay of the interrupted SGUs. Similarly, the interrupted emergency SGUs have higher priority than the newly arrived emergency SGUs. It should be noticed that the interrupted emergency SGUs of c2 have higher priority than the interrupted SGUs of c1.

As shown in Fig. 2, there are five priority queues for the differential QoS. Being sorted from high to low priority, they are PUs, interrupted emergency SGUs, newly arrived emergency SGUs, interrupted SGUs, and newly arrived SGUs. Before transmission, SGUs should queue according to their priorities. In each priority queue, SGUs are sorted by their classes in the decreasing order. The SGUs with high class move to the head of the queue. The SGUs with low class move to the tail of the queue. The scheduler preferentially allocates channels to the users in the highest priority queue, which means that the SGU has chance for data transmission only when its above queues are empty. The transmission of low priority SGUs will be interrupted at anytime when the PUs or high-priority SGUs arrive. The interrupted SGUs will go back to their corresponding queues and wait for being assigned.

The above QoS-aware priority policy guarantees the differential QoS provision in smart grid. Based on this priority policy, the scheduler will adjust the allocation decision at each time slot to minimize the transmission delay of SGUs according to the channel’s state and SGU’s state. In Section III, we formulate the transmission scheduling procedure as an SMDP problem.

III. SEMI-MARKOV DECISION PROCESS FORMULATION

In this section, the problem of transmission scheduling in CR-based smart grid communications networks is formulated as an SMDP. For the SMDP, the transmission scheduling process is divided into a series of stages, and the number of
stages cannot be limited. The optimal transmission scheduling decision is made stage-by-stage. Hence, the optimization of the entire process can be achieved.

A. Components of SMDP

For transmission scheduling, the scheduler will make a decision on the channel allocation to SGUs and minimize the system cost according to the knowledge of the communications networks states, including the availability of channels (the activities of PUs) and SGUs conditions. The components of the SMDP corresponding to the transmission scheduling problem are as follows.

1) Stage: The scheduling process is naturally divided into a series of stages. Stages are indexed by integer \( k = 1, 2, \ldots \). The duration of one stage is identically defined as \( \Delta \tau \). The scheduler makes the channel allocation decision at the beginning of each stage. PUs and SGUs will come at the beginning of each stage and leave at the end of each stage after services.

2) State: The current system state consists of the channel availability and the conditions of SGUs. Let \( ch_n(k) \) denote the availability of channel \( n \) at stage \( k \), \( n = 1, 2, \ldots, N \). \( ch_n(k) = 1 \) means channel \( n \) is available for SGUs, and \( ch_n(k) = 0 \) means channel \( n \) is occupied by a PU, which depends on the PU’s activity. For the state of SGUs, there are \( M \) SGUs opportunistically accessing to \( N \) channels, \( m = 1, 2, \ldots, M \). \( pm(m(k)) \) denotes the number of packets in the buffer of SGU \( m \) at stage \( k \). \( Im_m(k) \) is the interruption indicator. \( Im_m(k) = 1 \) denotes SGU \( m \) is interrupted at stage \( k \). Otherwise, \( Im_m(k) = 0 \). \( Em_m(k) \) is the emergency indicator. \( Em_m(k) = 1 \) denotes that SGU \( m \) has emergency packets at stage \( k \). Otherwise, \( Em_m(k) = 0 \). Let \( x_m(k) \) presents the states of SGU \( m \) at stage \( k \). We have

\[
x_m(k) = (pm(m(k)), Im_m(k), Em_m(k)). \tag{1}
\]

We call the system state at the beginning of stage \( k \) as the state of stage \( k \), denoted by \( x(k) \). Thus, \( x(k) \) is defined by

\[
x(k) = (ch_n(k), x_m(k)) | n = 1, \ldots, N, m = 1, \ldots, M. \tag{2}
\]

The value \( x(k) \) does not change in the short duration of one stage. The set of all possible states is called as state space, denoted by \( X \).

3) Decision: At stage \( k \), the scheduler should make a decision \( u(k) \) according to the state \( x(k) \). The decision \( u(k) \) is defined by

\[
u(k) = (u_m(k)) | m = 1, 2, \ldots, M \tag{3}
\]

where \( u_m(k) \) denotes the channel allocation of SGU \( m \) at stage \( k \). If \( u_m(k) = n \), this means that the SGU \( m \) is permitted to access to channel \( n \). \( u_m(k) = 0 \) denotes SGU \( m \) has no channel for packet transmission. Before SGU \( m \) is allocated a channel, i.e., \( u_m(k) > 0 \), its packets must be nonempty, i.e., \( pm(m(k)) > 0 \). The decision space denoted by \( U \). The subset \( U[x(k)] \) contains all possible decisions given the system state \( x(k) \).

The transmission rate mainly depends on the wireless channel condition. We set the transmission rates of all channels are the same, and the \( Ps \) number of packets just can be successfully delivered in one stage. Hence, we have \( pm(m) = pm(m) - Ps \), if \( u_m(k) > 0 \) and \( pm(m) > Ps \).

4) Policy: Policy is a sequence of decision functions. That is \( \pi = (\mu(1), \mu(2), \ldots, \mu(k), \ldots) \). If for all stages \( k \), there is \( \mu(k) \equiv \mu \), then decision functions do not change with stages and the policy is stationary. Since we only consider the stationary policy in this paper, each decision function \( \mu(k) : X \mapsto U \) is a mapping from the state space \( X \) to the decision space \( U \). The stationary policy is denoted by \( \mu \). The decision at stage \( k \) is also represented by \( u(k) \equiv \mu(x(k)) \).

5) Utility Function: The utility function of stage \( k \) is determined by state \( x(k) \) and decision \( u(k) \), described as \( U[x(k), u(k)] \). The packet transmission delay is the measurement to estimate the QoS performance in CR-based smart grid communications networks. Hence, let the utility function have the following form:

\[
U[x(k), u(k)] = \sum_{m=1}^{M} W_m(k) \tau_m(k) 1_{[pm(m)>0]} \tag{4}
\]

where \( 1_A \) is an indicator function, providing that \( 1_A = 1 \) when \( A \) is true, and \( 1_A = 0 \) when \( A \) is false. In (4), the transmission delay is computed only when the SGU has packets for transmission. \( \tau_m(k) \) denotes the transmission delay of SGU \( m \) at stage \( k \), which is caused by blocking and interruption. The transmission delay of SGU \( m \) is calculated by

\[
\tau_m(k) = \begin{cases} 
\Delta \tau, & u_m(k) = 0 \\
0, & u_m(k) > 0.
\end{cases} \tag{5}
\]

At stage \( k \), when the SGU \( m \) is blocked or interrupted, i.e., SGU \( m \) has packets in buffer but it has no channel for transmission, SGU \( m \) has to wait in its queue in this stage, and the duration of stage is \( \Delta \tau \). Thus, the delay time \( \tau_m(k) = \Delta \tau \). Otherwise, the SGU can transmit packets on the given channel, or the SGU is idle and has no packet at stage \( k \). Thus, we have \( \tau_m(k) = 0 \).

To guarantee the differential QoS of SGUs, the scheduler distributes the channels based on the priorities of SGUs. \( W_m(k) \) is the weight of SGU \( m \) at stage \( k \), which is related to its role class, interruption, and emergency situations. Let

\[
W_m(k) = W_{b,m} + W_{I} 1_{[Im_m(k)=1]} + W_{E} 1_{[Em_m(k)=1]} \tag{6}
\]

where \( W_{b,m} \) is the basic weight of SGU \( m \) according to its class in smart grid communications networks. If SGU \( m \) corresponds to class \( i \) in smart grid, we have \( W_{b,m} = W_{c,i}, i = 1, 2, \ldots, ) \). \( W_{c,i} \) is the weight of class \( i \). There are total \( \mathbb{I} \) classes \( (I \ll M) \) and class \( 1 \) (c1) has the highest priority with the highest weight, i.e., \( W_{c,1} > W_{c,2} > \cdots > W_{c,\mathbb{I}} \). \( W_I \) and \( W_E \) are interruption weight and emergency weight, respectively. For the above weights, the following constraints must be satisfied:

\[
W_{c,i} > W_{c,i+1} + W_I \tag{7}
\]

\[
W_{c,i} + W_I < W_{c,i+1} + W_E. \tag{8}
\]

Constrain (7) means a regular SGU in class \( i \) still has higher priority than an interrupted SGU in class \( i+1 \), keeping the interrupted priority within one class. Constrain (8) means
The SMDP problem is to find the optimal policy \( \mu \) to minimize the system cost, denoted by \( J \). The optimal policy leads to an interrupted SGU in class \( i \), which permits the communications of the emergency SGU beyond the limit of class.

The utility function describes the transmission delay at each stage. The transmission delay of the entire scheduling process is defined as the system cost, i.e., the long-term transmission delay. Under policy \( \mu \), the system cost is provided by

\[
J[x(k)] = \sum_{j=k}^{\infty} U(x(j), \mu(x(j))). \tag{9}
\]

The number of stages can be infinite. The target of the SMDP problem is to find out the optimal policy \( \mu^* \) to minimize the system cost. The optimal policy leads to minimum system cost, denoted by \( J^* \).

**IV. ADAPTIVE DYNAMIC PROGRAMMING SOLUTION**

In this section, the methodology of ADP is introduced to solve the corresponding problems in SMDP. The ADP is the powerful tool for controlling nonlinear systems with extensive use. Optimal control schemes are developed for the unknown discrete-time nonlinear systems using ADP [21]–[23]. Newly developed ADP approaches are proposed to obtain near optimal control [24]–[26]. The ADP method is also applied to design decentralized controller for large-scale nonlinear systems [27].

In this paper, we employ the heuristic dynamic programming (HDP) that is the basic architecture of ADP [28]. A typical HDP contains action, model, and critic networks, which adopt neural networks for function approximation. For transmission scheduling problem, the HDP can learn from the activities of PUs and SGUs, and the QoS differential priority policy to estimate the optimal long-term system costs by online network training, even if the scheduler has no full knowledge of the system model.

**A. Bellman Equation**

By solving the Bellman optimality equation, the optimal policy \( \mu^* \) of the SMDP problem could be obtained. The optimal system cost from stage \( k \) is described by

\[
J^*[x(k)] = \min_{u(k)} [U(x(k), u(k)) + J^*[x(k+1)]]. \tag{10}
\]

The principle of optimality is to minimize \( J \) in immediate future and subsequently minimize the sum of \( U \) over all stages. If there is an available way to compute the optimal cost \( J^*[x(k+1)] \), the optimal transmission scheduling policy can be obtained by

\[
u^*(k) = \arg \min_{u(k)} [U(x(k), u(k)) + J^*[x(k+1)]]. \tag{11}\]

However, due to the huge size of state space and the overwhelming computational requirement, the accurate solution of \( J^*[x(k+1)] \) in (11) becomes unrealistic. Hence, we present an HDP architecture for transmission scheduling problem to produce an approximating optimal solution.

**B. HDP Architecture**

The HDP architecture for transmission scheduling is shown in Fig. 3, including the scheduler, the system model, and the critic network. If the number of stages is infinite, the long-term system cost is also infinite according to (9), since the transmission delay always occurs with the PUs’ activities. Hence, we rewrite the system cost (9) as the following form:

\[
J(k) = \sum_{j=k}^{\infty} \gamma^{j-k} U(j). \tag{12}\]

For presentation, let \( U(k) = U[x(k), u(k)] \) and \( J(k) = J[x(k)] \). \( \gamma \) is the discount factor (0 \( \leq \gamma \leq 1 \)). \( \gamma = 0 \) means that only the current utility value \( U(k) \) is considered in system cost, and the future values are neglected. \( \gamma = 1 \) means all future utility values are equally considered in the infinite range. The long-term system cost can be estimated only when \( \gamma < 1 \). In Fig. 3, the scheduler will give a decision \( u(k) \) based on the system state \( x(k) \) and approximating system cost \( \hat{J}(k) \). The scheduler behavior can be described by the following equation:

\[
u(k) = \arg \min_{u(k)} [U(k) + \gamma \hat{J}(k+1)]. \tag{13}\]

The \( U(k) \) is computed by \( x(k) \), so that the scheduler also has an input of \( x(k) \).

The system model in Fig. 3 describes the activities of PUs and SGUs. Its inputs are scheduler decision \( u(k) \) and system state \( x(k) \), and the output is the system state of the next stage \( \hat{x}(k+1) \). The system state consists of channel...
state and SGU state. The channel state changes according to the PUs’ activities, without the effect of SGUs and scheduler. For SGU state, the numbers of packets from SGU m at the next stage is provided as follows:

\[ p_m(k + 1) = \begin{cases} 
    p_m(k), & u_m = 0 \\
    p_m(k) - P_s, & u_m > 0.
\end{cases} \tag{14} \]

The interruption indicator of SGU m at the next stage is provided by

\[ I_m(k + 1) = \begin{cases} 
    1, & p_m(k) > 0 \text{ and } u_m = 0 \\
    0, & \text{or}.
\end{cases} \tag{15} \]

The emergency indicator of SGU will change according to the appearance of emergency packets. The emergency packets will be generated based on smart grid service needs.

For the critic network in Fig. 3, we employ the back propagation neural network (BPNN). The input of critic network is the system state. The output is the approximating system cost \( \hat{J}(k + 1) \) approaching the optimal system cost \( J^* \) by network training. The aim of the critic network training is to minimize the following error function:

\[ \| E_c \| = \sum E_c(k) = \frac{1}{2} \sum \varepsilon_c^2(k) = \frac{1}{2} \sum_k [\hat{J}(k) - U(k) - \gamma \hat{J}(k + 1)]^2 \tag{16} \]

where \( \hat{J}(k) = \hat{J}(x(k), u(k), W_{c1} \text{ and } W_e \text{ is the critic network weight parameters. When } E_c(k) = 0 \text{ is obtained for stage } k, \text{ we get the following derivation from (16)}:

\[ \hat{J}(k) = U(k) + \gamma \hat{J}(k + 1) \]
\[ = U(k) + \gamma [U(k + 1) + \gamma \hat{J}(k + 2)] \]
\[ = \sum_{j=k}^{\infty} \gamma^{j-k} U(j) \tag{17} \]

which is the same to the system cost in (12). Therefore, we train a neural network by minimizing the error function (16). Then, we can obtain the approximating value estimating the system cost defined in (12) for stage \( k + 1 \).

C. Critic Network Training

For the weight update of the critic network, the gradient-based adaptation is adopted to minimize the error function (16). The weight update formulation of critic network is presented as follows:

\[ \Delta W_e(k) = l_e(k) \left[ -\frac{\partial E_e(k)}{\partial W_e(k)} \right] = l_e(k) \left[ -\frac{\partial E_e(k)}{\partial \hat{J}(k)}, \frac{\partial \hat{J}(k)}{\partial W_e(k)} \right] \tag{18} \]

\[ W_e(k + 1) = W_e(k) + \Delta W_e(k) \tag{19} \]

where \( l_e \) is the positive learning rate of the critic network, which usually decreases with time to a small value for control system. \( W_e \) consists of \( W_{c1} \) and \( W_{c2} \). \( W_{c1} \) denotes the weight matrix between the input and the hidden layer, and \( W_{c2} \) denotes the weight matrix between the hidden layer and the output, which are derived by

\[ \Delta W_{c1} = -l_c e_c x_T \left[ W_{c2}^T \cdot (1 - c_{i2} \cdot c_{o2}) \right] \tag{20} \]

\[ \Delta W_{c2} = -l_c e_c x_T \tag{21} \]

where \( c_{i1} \) and \( c_{i2} \) are the input matrix and output matrix of hidden layer. Updating the weight parameters by the above formulations, the output of the critic network will approach to the approximating value of the system cost defined in (12). By adopting this HDP architecture in transmission scheduling problem, the long-term system cost is estimated and minimized. The optimal scheduling policy can be obtained by HDP with a well-trained critic network.

V. SIMULATION RESULTS

A. Simulation Setup

In the simulations, we consider the cognitive NAN in smart grid. The utility provider rents a spectrum band, which consists of four orthogonal channels with identical bandwidth. These channels are shared by their licensed PUs, \( N = 4 \). There are eight SGUs opportunistically accessing to the licensed channels, \( M = 8 \). There are two classes of SGUs, \( c_1 \) and \( c_2 \). SGUs 1–4 belong to \( c_1 \), and SGUs 5–8 belong to \( c_2 \).

The transmission scheduling process has 1000 stages, and the duration of one stage \( \Delta t \) is 1. The PUs arrive as the Poisson process with arrival rate of \( \lambda_{PUs} \). The PUs occupy the channel for the duration of \( 10 \Delta t \) each time. The activities of PUs are known and recorded in the spectrum database. By accessing to the database, the scheduler can allocate the channels to SGUs when the channels are not occupied by PUs. The SGUs also arrive as Poisson process with arrival rate of \( \lambda_{SGUs} \) and bring three packets each time when they arrive. Each packet can be transmitted successfully at one stage, i.e., \( P_s = 1 \).

There are four channels and eight SGUs. Each SGU has 3 states at each stage. Thus, the system state \( x \) have \( 4 \times 8 \times 3 = 28 \) components at each stage. The decision \( u \) has eight components. Therefore, in the proposed HDP architecture, the scheduler has 28 + 1 = 29 inputs of state and eight outputs of decision. The system model has 28 + 8 = 36 inputs and 28 outputs of next state. The critic network is chosen as 28-35-1 structure of BPNN, including 28 input neurons, 35 hidden layer neurons, and one output neuron. The output layer uses purelin function. The hidden layer uses the sigmoidal function, provided by

\[ y = \frac{1 - e^{-x}}{1 + e^{-x}}. \tag{22} \]

The weight parameters of critic network are initialized randomly, which is similar to the training of regulate BPNN for function approximation. Then, the gradient-based method is employed for the weight update to approach the minimum system cost in (18)–(21). The simulation results are presented using the self-developed simulator based on MATLAB.
B. Convergence Performance

The convergence performance of the proposed HDP architecture is presented to solve the transmission scheduling problem. The learning rate \( l_c \) is set to 0.2, which is kept in the whole network training process. Because the PUs and SGUs states are changing continuously, the network should keep this learning ability in the whole scheduling process. With the discount factor \( \gamma = 0.9 \), Fig. 4 shows the output of the critic network \( \hat{J} \), which is called the approximating optimal system cost. We find that our HDP architecture can estimate the \( \hat{J} \) approaching to the optimal system cost \( J^* \) by network training. Fig. 5 shows the approximating optimal system cost with \( \gamma = 0.7 \). Compared with \( \gamma = 0.9 \), the small discount factor \( \gamma \) improves the convergence speed. Because a smaller \( \gamma \) means that the future utility values are less considered in the system cost in network training. If the system cost only contains the utility value in current stage (\( \gamma = 0 \)), the scheduler can give an optimal decision immediately, and there is no need to train the critic network. A higher \( \gamma \) is corresponding to the long-term system cost, which considers more utility values in future stages with the sacrifice of convergence speed.

The fluctuation of output is caused by the dynamic arrivals and departures of PUs and SGUs. Therefore, it is not appropriate for critic network training to set the learning rate as a too low or too high value. For a low learning rate, the user’s dynamical activities will overwhelm the weak learning ability of the critic network. A high learning rate will make the network parameters change dramatically, and eventually miss the optimal convergence value.

C. Scheduling Performance

The transmission delay is presented to show the performance of the proposed QoS differential scheduling. The delay and the transmitted packet number of each SGU are summed up in the whole process. Thus, we have the overall delay and the throughput. Let the overall delay divide the throughput, which is the delay per packet. The delay per packet is calculated after 100 experiments.

Fig. 6 shows the mean and standard deviation of delay per packet of SGU 1, with fixed \( \lambda_{PU} \), when the smart grid communications network introduces the QoS differential priority policy, SGU 1 has the highest priority in \( c_1 \). Thus, we find the SGU 1 has lower delay than that without priority. Even the arrival rate \( \lambda_{SGU} \) of all SGU increase, leading to heavy SGU traffic, the SGU 1 still has low transmission delay guaranteeing the QoS of high priority SGUs. Fig. 7 shows the delay per packet of SGU 1 when all SGUs have fixed \( \lambda_{SGU} \). While the PU traffic is heavy, the transmission delay of SGUs will also be high, in the case of priority policy, since the PUs have the absolute priority to use the licensed channels regardless of the highest priority of SGU 1. But the proposed priority policy still decreases the transmission delay of high priority SGUs compared with the case with no priority.

In the experiments, we let SGUs generate emergency packets with the probability of 10%. Figs. 8 and 9 show the
Fig. 7. Delay per packet of SGU 1 in terms of PU arrival rate.

Fig. 8. Delay per emergency packet of SGUs in terms of SGU arrival rate.

Fig. 9. Delay per emergency packet of SGUs in terms of PU arrival rate.

delay per emergency packet of SGUs with fixed $\lambda_{PU}$ and fixed $\lambda_{SGU}$, respectively. In our priority policy, the SGU 8 belongs to the lowest class with lowest priority. In both Figs. 8 and 9, we find that the proposed QoS differential priority policy reduces the delay of emergency packets of SGU 1 and SGU 8 significantly. Actually, delays of emergency packets for all SGUs are similarly low in our policy. Because the probability of emergency packets is low, the probability of two or more SGUs simultaneously appearing emergency packets is even lower. In general, if one SGU has emergency packets, it will have the highest priority with very high probability in the scheduling. Therefore, all SGUs have the similar low delay levels. Since the emergency packets are generated randomly, all the standard deviations are rather high, which are not marked in figures for presentation.

In general, the QoS differential priority policy assigns the priority levels to the SGUs based on their roles in smart grid and their situations of interruption and emergency, which guarantees the QoS of high-priority SGUs and all emergency SGUs.

VI. CONCLUSION

In this paper, the QoS differential transmission scheduling is proposed in the CR-based smart grid communications networks. A QoS-aware priority policy is devised for differential QoS provision. The SGUs are assigned to have different classes according to their roles in the smart grid. The priorities of SGUs will be readjusted depending on their current situations of interruption or emergency. Based on the proposed priority policy, the scheduler makes decision of channel allocation for the minimum transmission delay of SGUs. The entire transmission scheduling problem is formulated as an SMDP with infinite stages and solved by the methodology of ADP. An HDP architecture is established, which has the ability of online network training, learning from the activities of PUs and smart grid users, and adjusting scheduling decision to achieve transmission delay minimization. Simulation results illustrate that the proposed priority policy is able to guarantee low transmission delay for high priority SGUs. Meanwhile, the transmission delay of emergency data decreases to a satisfying low level for all users, ensuring the differential QoS provision in smart grid.

REFERENCES


