Lyapunov-based Traffic Scheduling and Optimization in Smart Distribution Grid Communication Networks

Si-Ya Xu[1], Shao-Yong Guo[1], Zhu Liu[2], Lei Wei[3], Luo-Ming Meng[1]

[1] State Key Laboratory of Network and Switching Technology, Beijing University of Posts and Telecommunications, China

[2] Information & Telecommunication Branch of Nanjing Power Supply Company of State Grid Jiangsu Electric Power Company, China

[3] State Grid Information & Telecommunication Co., Ltd., China

xusiyaxsy@hotmail.com, syguo@bupt.edu.cn, liuzhu_558@163.com, nj_weilei@163.com, lmmeng@bupt.edu.cn

Abstract

Saving financial cost is a challenge in heterogeneous smart distribution grid communication networks for different operation modes and differentiated quality of service (QoS). In this paper, financial cost and transmission quality are integrated into the cost-optimal distributed control mechanism to guide network controller to make traffic scheduling decisions. Firstly, a dedicate system architecture is established to model the characteristics of traffic and the dynamic evolution of hybrid access networks. And then a novel distributed control strategy based on *Lyapunov* theories is designed to optimize the traffic scheduling by taking output network access control and service rate adjustment control. Performance of the proposed scheduling strategy is evaluated by using *MATLAB* and *OPNET*, which shows that it can lower financial cost than conventional strategies while ensuring transmission quality.

Keywords: Smart distribution grid, Heterogeneous access networks, Lyapunov theory, Financial cost optimization

1. Introduction

With the rapid development of smart grid, more and more sub-systems need to achieve information sensing, transmitting and processing to ensure safety and reliability of the power grid operation [1]. Diverse QoS requirements (e.g., delay, rate, and reliability) must be satisfied in heterogeneous communication networks [2], [3-4]. For instance, video surveillance service demands high bandwidth, while distributed control service requires the exchange of packet with critical delay and reliability constrains [5]. Compared to the backbone communication networks with relatively stable operation quality, the heterogeneous smart distribution grid (SDG) communication network has many weaknesses, such as the complex network structure and the differentiated transmission quality. Since heterogeneous service traffic has diverse requirements for QoS, there is no single technology that can solve all the needs [6]. A prime architecture of heterogeneous SDG communication networks is illustrated in Figure 1. It shows that a variety of communication technologies, including EPON, PLC, WIMAX and LTE, are applied to constitute the existing communication network to provide ubiquitous low-cost connections for SDG communication networks jointly. Integration of these communication technologies is non-trivial due to the distinct differences in QoS and financial cost. Moreover, with the aim of forming a green and efficient network, the financial cost will no longer be ignored. Therefore, it is an important problem to design an effective traffic scheduling mechanism for controlling cost and providing reliable QoS in SDG communication networks.

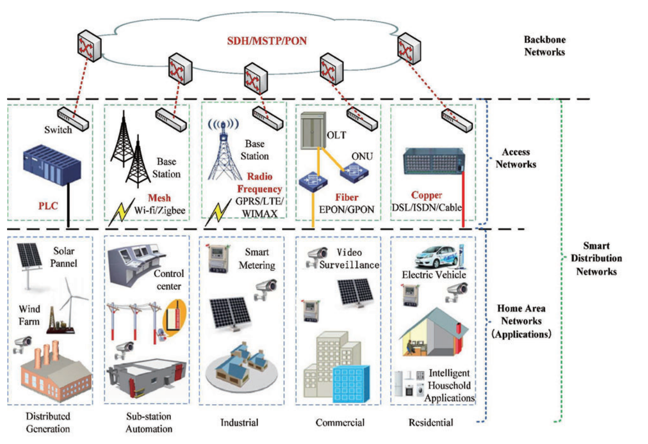

Figure 1. The architecture and applications of heterogeneous SDG communication networks

Traffic scheduling across SDG communication networks poses a novel research problem. It becomes considerably more challenging than that in traditional

*Corresponding Author: Shao-Yong Guo; E-mail: syguo@bupt.edu.cn

DOI: 10.3966/160792642018011901013

grid for distributed production and priced-based local consumption [7]. The research on traffic scheduling in heterogeneous networks has attracted lots of attentions and gained many achievements [8-9]. A sub-optimal distributed control algorithm is presented in wireless networks to efficiently support QoS through channel control, flow control, scheduling and routing decisions in [3]. A queuing-based dynamic channel selection strategy for heterogeneous applications is presented in [10], but the scope is limited to certain applications, and only suitable for wireless communication networks. Similarly, another throughput algorithm is proposed in [11] that iteratively increases the rate of each flow until it converges to the optimal rate of all the flows. However, it is not to be neglected that few research on financial cost control exits in lots of routing protocols for smart grid communication networks. In conclusion, the exiting routing protocols pertaining to QoS support can hardly solve the cost-efficient scheduling problems well [12-13].

To address the problems above, we should construct a cost-optimal distributed control mechanism to accomplish two goals: to optimize financial cost and to satisfy QoS requirements. Firstly, a queuing model for SDG is established according to the characteristics of heterogeneous services and output networks. Then, a distributed scheduling algorithm is designed to dynamically allocate heterogeneous service data to yield minimum financial cost and guarantee the required QoS. The scheduling algorithm presented in this paper is based on the *Lyapunov* theory [14] which converts all QoS and cost constrains into queues stability problems. Furthermore, we optimize the cost by adjusting network parameters (e.g., service rate) to adapt with different traffic intensities.

The outline of this paper is organized as follows. In Section 2, we construct the network model, including detailed system model considered herein and general system dynamic evolution models. Section 3 defines system delay and economic costs as performance parameters in our system. The optimization algorithm and the cost-optimal distributed control strategy (CDCS) are presented in Section 4 and Section 5 respectively. Section 6 describes simulation environments and illustrates performance evaluation results. Section 7 concludes the paper.

2. Network Model

2.1 System Model

Consider the system model depicted in Figure 2 in three parts: input queue set, output queue set and time-varying fading channels between these two queue sets. Input queues regarded as buffers associated with different QoS requirements and characteristics are used to store input traffic data. Assume that traffic flows of heterogeneous services are properly differentiated into M classes through aggregation based on their QoS constraints (e.g., delay tolerance, rate and failure probability). And those flows are assigned priorities for classification. Each traffic priority corresponds to a dedicated input queue, which means that input queue i only admits the arrival of traffic flow with priority i. Therefore, the number of input queues is M. The N output queues represent diverse choices for delivery and link to N output networks under different communication technologies including EPON, WCDMA, WIMAX, PLC, and etc.. In each time slot t, new data randomly arrives to input queues, and awaits to transmit from input queues to output queues, and then be delivered into output networks. The network controller adopts scheduling strategies to decide which packet to be served, which output network to be connected, and how much data to be transmitted in each scheduling period. Table 1 lists the definition of variables used in our network model. These definitions are observed at time slot t and expressed in units corresponding to buffer slots.

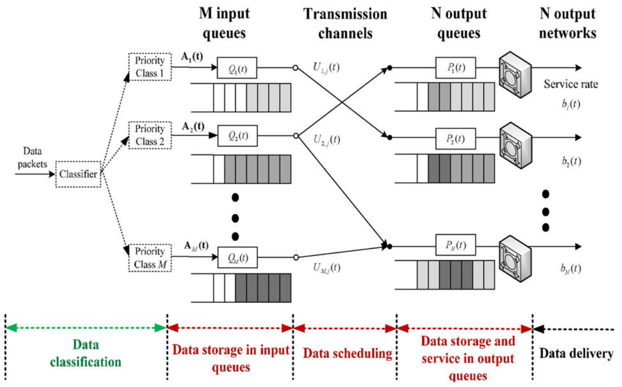

Figure 2. System model

Input queues. Consider the traffic flows with similar QoS requirements inject into the same input queue. When a network controller makes scheduling decisions, it cares the service priority rather than the actual size of a packet. Let $\boldsymbol{A}_i(t),\ i = 1, 2, ..., M$ be the packet set of class i which arrives in time slot t, and $|\boldsymbol{A}_i(t)|$ be the packet number. If $S_{i,x}$ is the actual size of packet x in units, the average packet size of class i is $S_i = E\,|S_{i,x}|$ units. The arrivals of class i on slot t are $\alpha_i(t) = |\boldsymbol{A}_i(t)| * S_i$. So the time average arrival rate of class i denoted by λ_i can be calculated as

$$\lambda_i = \lim_{T\to\infty} \sup \sum_{t=0}^{T-1} S_i\,|\boldsymbol{A}_i(t)| \tag{1}$$

We denote $Q_i(t)$ as the backlogs in input queue i at the beginning of time slot t, which means the amount of data need to be transmitted. Then, the vector of backlogs in all input queues over integer time slot $t \in \{0, 1, 2, ...\}$ can be expressed as $\boldsymbol{Q}(t) = (Q_1(t), Q_2(t), ... Q_M(t))$.

Table 1. Definitions of variables used in system model

Symbol	Description
$Q_i(t)$	Backlog of input queue i
$P_j(t)$	Backlog of output queue j
$\boldsymbol{A}_i(t)$	Packet set of class i arrives in time slot t
$\boldsymbol{A}_{i,j}(t)$	Packet set of class i transmits from input queue i to output queue j
$S_{i,x}$	The actual size of packet x
$\alpha_i(t)$	Traffic arrival rate in input queue i
λ_i	The time average arrival rate of class i
$b_j(t)$	Service rate of output queue j
$\tilde{b}_j(t)$	Actual service rate of output queue j
$\boldsymbol{U}(t)$	Scheduling strategies for all queues over time slot t
$U_{i,j}(t)$	Amount of traffic transmitted from input queue i to output queue j
$T_{i,j}(t)$	The average assignment ratio between input queue i and output queue j
$C_{i,j}^{CAP}(t)$	Capacity of the channel connected input queue i and output network j
d_j	Delivery delay associated with output network j
$Z_k(t)$	Amount of past cost exceeding the required cost bound
$H_i(t)$	Amount of past queuing time exceeding the required delay bound for service class i

Output queues and output networks. The hybrid access network consists of N individual output networks. All of them have ability to deliver SDG communication traffic to its destination in the same area. At each interface of output networks, there exits an output queue buffering packets transmitted from input queues. Assume that all output queues are continuous-working, strict priority, and non-preemptive. The service rate $b_j(t)$ for output queue j on slot t is associated with the characteristics of output networks and greatly influenced by bandwidth. The different delivery delay among different communication technologies is a major characterization for output networks.

Let $P_j(t)$ represent the backlog of output queue j on slot t and $\boldsymbol{P}(t) = (P_1(t), P_2(t), \ldots P_N(t))$ be vector of current backlogs in all output queues.

Transmission channels. To reflect fading coefficient and/or noise ratios in transmission channels, channel conditions are assumed to be constant for the duration of a slot, but varying from slot to slot. Let $U_{i,j}(t)$ denote the amount of data transmitted from input queue i to output queue j on slot t, and $C_{i,j}^{CAP}(t)$ is the current channel capacity. Then, we have

$$U_{i,j}(t) \le \min[Q_i(t), C_{i,j}^{CAP}(t)] \tag{2}$$

The scheduling strategies $U(t)$ determine the amount of packets transferred from input queues to output queues in each time slot. We can define $U(t)$ as a vector

$$\boldsymbol{U}(t) = (U_{1,1}(t), U_{1,2}(t), \ldots U_{M,N}(t)),\ t \in \{0, 1, 2, \ldots\} \tag{3}$$

2.2 System Dynamic Evolution Model

Input queues dynamic evolution. In time slot t, network controller selects $U_{i,j}(t)$ units of data to be removed from input queue i to output queue j. Future states of input queue i are driven by stochastic arrival and scheduling process according to the following dynamic equation:

$$Q_i(t+1) = Q_i(t) + \alpha_i(t) - u_i(t) \tag{4}$$

where $\alpha_i(t)$ is traffic arrival rate in input queue i over slot t. For $u_i(t)$ is the total transmission rate from input queue i to all output networks, it can be expressed as

$$u_i(t) = \sum_{j=1}^{N} u_{i,j}(t) \tag{5}$$

Output queues dynamic evolution. Similarly, $P_j(t)$ represents the backlog of output queue j on slot t and $u_j(t)$ is the quantity of data injecting into output queue j. Then we have

$$u_j(t) = \sum_{i=1}^{M} U_{i,j}(t) \tag{6}$$

Therefore, the update rule for output queue j can be described as

$$\begin{aligned} P_j(t+1) &= \max\left[P_j(t) - b_j(t), 0\right] + u_j(t) \\ &= P_j(t) + u_j(t) - \tilde{b}_j(t), \text{for } t \in \{0,1,2\ldots\} \end{aligned} \tag{7}$$

where $\tilde{b}_j(t)$ is the actual service rate of output queue j over slot t.

3 Performance Parameters

3.1 System Delay

The system delay denoted by D, consists of three parts: input queuing delay, output queuing delay, and delivery delay in the output network [15].

Input queuing delay. Input queuing delay D^{in} represents the waiting time of a packet before transmitting to output queues. By *Little*'s law, the average waiting time in input queue i expressed in time slots is

$$\overline{D_i^{in}} = \lim_{T\to\infty} \sup \frac{1}{\lambda_i T} \sum_{t=0}^{T-1} E[Q_i(t)] \tag{8}$$

Output queuing delay. Output queuing delay D^{out} is the waiting time in output queues before being served. Once a packet injects into output network *j*, D^{out} is determined by the total transmission rate $u_j(t)$ and service rate $b_j(t)$. The average assignment ratio $T_{i,j}(t)$ between input queue *i* and output queue *j* is

$$T_{i,j}(t) = \lim_{T\to\infty} \sup \sum_{t=0}^{T-1} E[\frac{U_{i,j}(t)}{\alpha_i(t)}] \quad \textbf{(9)}$$

For $P_j(t)$ is the backlog of output queue *j* at time *t*, the waiting time D_j^{out} is

$$D_j^{out} = \frac{P_j(t)}{b_j(t)} \quad \textbf{(10)}$$

From (1), (9) and (10), the time average waiting time in output queues can be described as

$$\overline{D_i^{out}} = \lim_{T\to\infty} \sup \frac{1}{\lambda_i T} \sum_{t=0}^{T-1} \sum_{j=1}^{N} E[\frac{U_{i,j}(t)P_j(t)}{b_j(t)}] \quad \textbf{(11)}$$

Output network delivery delay. Once a packet transferred to an output network, the output network is in charge of forwarding the packet to its final destination. Transmitting through different output networks will result in different delivery delays. Assume that delivery delay in output network *j* is d_j, we can calculate the mean delivery delay by

$$\overline{D_i^{tr}} = \lim_{T\to\infty} \sup \frac{1}{\lambda_i T} \sum_{t=0}^{T-1} \sum_{j=1}^{N} d_j E[U_{i,j}(t)] \quad \textbf{(12)}$$

Overall system delay. The overall system delay D is the sum of all delays accumulated in input queues, output queues, and output networks. Hence,

$$\overline{D_i} = \overline{D_i^{in}} + \overline{D_i^{out}} + \overline{D_i^{tr}} \quad \textbf{(13)}$$

From the expressions, we can find that the mean output queuing delay $\overline{D_i^{out}}$ and the mean delivery delay $\overline{D_i^{tr}}$ are all functions of scheduling strategies $\boldsymbol{U}(t)$.

3.2 Financial Cost

Diverse access networks will not only provide users with differentiated QoS, but also cause different financial costs. Define $\boldsymbol{T}(t) = (y_0(t), y_1(t), \ldots y_k(t))$ as the vector indicating all cost functions in our system. In order to maintain system stability and balance the cost in all output networks, we set a predefined constant $\overline{y_k}$ as control parameter. Define $\overline{y_k}(t)$ as the average expectation of financial cost over the first *t* slots, we have

$$\overline{y_k} \triangleq \frac{1}{t} \sum_{\tau=0}^{t-1} E[y_k(\tau)] \quad \textbf{(14)}$$

If we set per units cost of transmission, financial cost can change with scheduling strategies $\boldsymbol{U}(t)$.

4 Lyapunov Optimization

4.1 Optimization Problem Model

To optimize financial cost, we formulate scheduling strategies by applying the *Lyapunov* theories to our queuing system [14]. Let $y_k(t)$ be the cost function, the optimization problem can be formulated as

$$\text{min:}\ \lim_{t\to\infty} \sup \overline{y_0}(t) \quad \textbf{(15)}$$

Subject to

$$\lim_{t\to\infty} \sup \overline{y_k}(t) - \overline{y_k} \le 0,\ k = 1, 2, \ldots L \quad \textbf{(16)}$$

$$\overline{D_i} \le d_i^{\lim},\ i = 1, 2, \ldots M \quad \textbf{(17)}$$

all $Q_i(t)$ and $P_j(t)$ queues are mean rate stable **(18)**

$$U_{i,j}(t) \le C_{i,j}^{CAP}(t),\ t \in \{0, 1, 2, \ldots\} \quad \textbf{(19)}$$

To solve problems given in (16)-(19), we first transform all inequality and equality constraints into queue stability problems. Specially, we define virtual queues $Z_k(t)$ for each $k \in \{0, 1, 2, \ldots L\}$ to accumulate the past cost exceeding cost constraints (store as virtual queue backlogs) with update equations:

$$Z_k(t+1) = \max[Z_k(t) + y_k(t) - \overline{y_k}, 0] \quad \textbf{(20)}$$

The virtual queues $Z_k(t)$ are used to enforce the constraint given by (16). To clearly explain our system and optimization theory, we give a proof in Appendix. Similarly, we define virtual queues $H_i(t)$ to monitor the amount of past observed delay violating delay constraints in each traffic priority class. Assume that $H_i(0)$ is non-negative and finite, and $H_i(t)$ is finite for $i \in \{1, 2, \ldots, M\}$. The update equations of $H_i(t)$ are calculated according to

$$H_{i,j}(t) = \max[H_i + \sum_{x \in A_{i,j}(t)} (W_{i,x} - d_{i,j}^*), 0] \quad \textbf{(21)}$$

where $\boldsymbol{A}_{i,j}(t)$ is the packet set of class *i* removed from input queue *i* to output queue *j*. Define $d_{i,j}^* = \{d_i^{\lim} - d_j \mid x \in \boldsymbol{A}_{i,j}(t)\}$ as the total queuing delay bound for packets with priority *i* before being served by output network server. Based on *Lyapunov* theories, to maintain the stability of virtual queues $H_i(t)$ is to satisfy the delay constraint as shown in (17). If the following equations (22), (23) are true for all values of *M* and *N*, the system can be stable and the constraint described in (18) can be met.

$$\lim_{t\to\infty}\frac{E\,|\,Q_i(t)\,|}{t}=0,\ i=1,2,\ldots M \tag{22}$$

$$\lim_{t\to\infty}\frac{E\,|\,P_j(t)\,|}{t}=0,\ i=1,2,\ldots M \tag{23}$$

4.2 Lyapunov Drift-plus-penalty Optimization

Let $\Theta(t)=(\boldsymbol{Q}(t),\boldsymbol{P}(t),\boldsymbol{Z}(t),\boldsymbol{H}(t))$ be a concatenated vector of all actual and virtual queues, and define the *Lyapunov* function as

$$L(\Theta(t))\triangleq\frac{1}{2}\begin{pmatrix}\sum_{i=1}^{M}\alpha Q_i(t)^2+\sum_{j=1}^{N}\beta P_j(t)^2\\ +\sum_{k=1}^{L}\gamma Z_k(t)^2+\sum_{i=1}^{M}\delta H_i(t)^2\end{pmatrix} \tag{24}$$

where the weighting coefficients α,β,γ and δ are assigned to intensify and balance each of constraints.

Define $\Delta\Theta(t)$ as the conditional *Lyapunov* drift in slot t:

$$\Delta\Theta(t)\triangleq E[L\Theta(t+1)-L(\Theta(t))\,|\,\Theta(t)] \tag{25}$$

where $E[.]$ is the expectation with respect to channel states and the control actions made in response to these channel states that relays on control strategies $\boldsymbol{U}(t)$. Instead of taking a control action by formulating scheduling strategies to minimize a bound on $\Delta\Theta(t)$, we minimize a bound on the following drift-plus-penalty expectation:

$$\Delta(\Theta(t))+VE\{y_0(t)\,|\,\Theta(t)\} \tag{26}$$

where $E\{y_0(t)\,|\,\Theta(t)\}$ is the financial cost in our system over time slot t, and $V\geq 0$ is a control parameter chosen to represent how much we emphasize the cost minimization and to tradeoff between financial costs and QoS constrains. Such a control decision can be motivated as follows: not only to make $\Delta\Theta(t)$ small to push queue backlogs toward a lower congestion state, but also to make $E\{y_0(t)\,|\,\Theta(t)\}$ small so that the system do not incur a huge financial cost. We can bound the drift-plus-penalty by

$$\begin{aligned}&\Delta(\Theta(t))+VE\{y_0(t)|\Theta(t)\}\leq B+VE[y_0(t)|\Theta(t)]\\&+\sum_{i=1}^{M}Q_i(t)E\{a_i(t)-u_i(t)|\Theta(t)\}\\&+\sum_{j=1}^{N}P_j(t)E\{u_j(t)-\tilde{b}_j(t)|\Theta(t)\}\\&+\sum_{k=1}^{L}Z_k(t)E\{y_k(t)-\overline{y_k}|\Theta(t)\}\\&+\sum_{i=1}^{M}H_i(t)E\left\{\sum_{x\in\mathbf{A}_{i,j}(t)}\left(W_{i,x}(t)-d_{i,j}^{*}\right)|\Theta(t)\right\}\end{aligned} \tag{27}$$

where there is also a bound on B:

$$B\geq\frac{1}{2}\begin{bmatrix}\sum_{i=1}^{M}E\{a_i(t)^2+u_i(t)^2|\Theta(t)\}\\+\sum_{j=1}^{N}E\{u_j(t)^2+\tilde{b}_j(t)^2|\Theta(t)\}\\+\sum_{k=1}^{L}E\{(y_k(t)-\overline{y_k}^2)|\Theta(t)\}\\+\sum_{i=1}^{M}E\left\{\sum_{x\in\mathbf{A}_{i,j}(t)}\left(W_{i,x}(t)-d_{i,j}^{*}\right)^2|\Theta(t)\right\}\end{bmatrix} \tag{28}$$

We need to develop strategy $\boldsymbol{U}(t)$ to achieve the minimum bound of *Lyapunov* drift-plus-penalty greedily, while keeping system stable. Let y_0^{opt} be the optimal value, and assume that $E[L\Theta(0))]<\infty$, we have

$$\lim_{t\to\infty}\sup\frac{1}{t}\sum_{\tau=0}^{t-1}y_0(\tau)\leq y_0^{opt}+\frac{B}{V}+\frac{E[L(\Theta(0))]}{Vt}\leq y_0^{opt}+O(1/V) \tag{29}$$

As (29) shows, any feasible scheduling strategies can help us to get a value $0(1/V)$ away from the optimal financial cost y_0^{opt}. We can approach the optimal value y_0^{opt} by amplifying V, which may cause the enlargement of queue backlog in return. Furthermore, if the service rate $b_j(t)$ can be adjusted in reaction to the changing of queue backlogs, the financial cost can also be effected by $b_j(t)$. The heavier load will cause a higher cost in this system. In other words, the financial cost in our system can vary by two factors which are assignment ratio $T_{i,j}(t)$ and service rate $b_j(t)$.

5 Cost-optimal Distributed Control Strategy (CDCS)

As above investigated, the average financial cost and system delay are related to the selected access networks and the assigned service rates. To deal with this complex optimization problem, we decompose it into two separate sub-problems. That is, the CDCS minimizes $\lim_{T\to\infty}\overline{y_0}(t)$ by output network access control and service rate adjustments. These two sub-problems are solved at fixed and dynamic service rates separately.

5.1 Output Network Access Control

When a packet need to be sent to an output network, the network controller should make output network selection under the assumption that the service rate $b_j(t)$ of each output network j is a fixed value, and equal to a rated speed b_j unit/slot. Each output queue in our system can be seen as an *M*/*M*/1 queue. Through output network access control process, all the packets transmitted from input queues have accessed into appropriate output queues where the requirements of transmission delay can be satisfied at a rated speed b_j.

During time slot *t*, $U_{i,j}(t)$ units of data are selected to be transferred to output queue *j*, where $U_{i,j}(t)$ is the optimal solution to

$$\min:\ \lim_{T\to\infty} \overline{y_0}(t)$$

Subject to

$$\lim_{T\to\infty} \overline{y_k}(t) - \overline{y_k} \le 0,\ k = 1, 2, ..., L$$
$$U_{i,j}(t) \le C_{i,j}^{CAP}(t),\ t \in \{0, 1, 2, ...\}$$

5.2 Service Rate Adjustment Control

With changeable service rates, the value of $b_j(t)$ is decided by network controller according to load conditions in output queues. The boundary values b_j^{min} and b_j^{max} limit rate assignment and are limited by characteristics of output networks. The variable $\tilde{b}_j(t)$ will differ financial cost $y_k(t)$ and effect system delay. Each of the output queues becomes an *M*/*G*/1 queue, which has a poly-matroidal cost region. We can design adaptive online service rate adjustment polices to minimize the objective function [17]. Accordingly, $\tilde{b}_j(t)$ can be assigned by network controller as

$$\tilde{b}_j(t) = b_j^{min}$$
$$\tilde{b}_j(t)^{max} = \min[b_j^{min}, \max[u_j(t), b_j(t)]] \quad \textbf{(30)}$$

For the close relationship between financial costs $y_k(t)$ and service rate $b_j(t)$, we consider a cost function similar to the quadratic cost function of queue lengths given in [16] as

$$F = b_j(t)^2 * 10^{-4} \quad \textbf{(31)}$$

where $\tilde{b}_j(t)$ can take values in continuous interval $[b_j^{\min}, \min[b_j^{min}, max[u_j(t), b_j(t)]]]$ and be the solution to

$$\min:\ \lim_{t\to\infty} \sup \overline{y_0}\ (\tilde{b}_j(t)) \quad \textbf{(32)}$$

Subject to

$$\lim_{t\to\infty} \sup \overline{y_k}\ (\tilde{b}_j(t))\ - \overline{y_k} \le 0,\ k = 1, 2, ..., L \quad \textbf{(33)}$$
$$\overline{d_i} \le d_i^{\lim}, i = 1, 2, ..., M.$$

5.3 The CDCS Process Description

At the beginning of each time slot *t*, traffic data is classified into *M* classes and sent into input queue buffers according to priorities. Then, network controller checks the system to find out whether there are packets need to be sent by input queue priority order. By taking output network access control, $U_{i,j}(t)$ units of data are selected to transfer to output queue *j* based on delay and cost constrains, as well as the current capacity $C_{i,j}^{CAP}(t)$ of transmission channels. Next, network controller adjusts service rate $b_j(t)$ depending on the real-time backlogs in output queues. After these two steps, the average system cost achieves the optimal value. At the boundary of every scheduling process, all queues update according to system dynamic evolution models are given in (4), (7), (20) and (21) finally. The pseudo-code for CDCS can be described as below:

	Working process in each scheduling period under CDCS
(1)	At every scheduling time *t* initialize $Q_i(t), P_i(t), Z_k(t)$ and $H_i(t)$; // set all actual and virtual queues
(2)	check the system to obtain $C_{i,j}^{CAP}(t)$;
(3)	classify data into *M* classes;
(4)	differentiated data arrives in the corresponding input queues;
(5)	while (*i*<*M*); // check input queues by priority order
(6)	*i*=*i*+1
(7)	if (input queue *i* packet!=0)
(8)	for(*j*=1: *N*)
(9)	send $0 < U_{i,j}(t) < C_{i,j}^{CAP}(t)$ data to output queue *j*; // $U_{i,j}(t)$ is the solution to min: $\lim_{t\to\infty} = \sup \overline{y_0}(t)$ with rated service rate b_j and subject to $\lim_{t\to\infty} = \sup \overline{y_k}(t) - \overline{y_k} \le 0$
(10)	end for
(11)	else $U_{ij}(t) = 0$; // input queue *i* is empty
(12)	end if
(13)	end while

(14)	for (*j*=*1*: *N*); // check output queues
	adjust $\tilde{b}_i(t) \in [b_i^{min}, \min[b_i^{min}, max[u_j(t), b_j(t)]]$
	and ensure $\overline{D_i} \le d_i^{lim}$;

(15)	// $\tilde{b}_i(t)$ is the solution to minimize $\lim_{t\to\infty} \sup \overline{y_0}(\tilde{b}_i(t))$ and subject to $\lim_{t\to\infty} = \sup \overline{y_k}(t)(\tilde{b}_i(t)) - \overline{y_k} \le 0$
(16)	end for
(17)	return $U_{i,j}(t), \tilde{b}_i(t)$ and $y_0(t)$;
(18)	update all queues according to evolution models

6 Performance Evaluation

6.1 Network Settings

6.1.1 Data and Input Queues Classification

We set three input queues based on delay requirements given in IEEE 1646 standard [17] and ITU-T recommendation G.114 [18] to store heterogeneous traffic in communication networks, they are:

Input queue 1. Stores communication traffic with extremely strict delay requirement, such as emergency response and SCADA (Supervisory Control and Data Acquisition). The traffic belongs to this set are generated and transmitted to control or protect the power gird, which usually has the highest priority lever.

Input queue 2. Stores communication traffic with normal delay requirement and second priority level, such as automated demand response.

Input queue 3. Stores non-real-time traffic which is assigned the lowest priority, such as power usage information acquisition.

6.1.2 Output Network Classification

We choose three typical kinds of access networks to form our hybrid output networks as following:

Public wired networks. Have much larger bandwidth and smaller transmission delay than PLC network, such as EPON. Transmission through this kind of access network will cause financial cost for they are constructed and operated by communication operators.

Public wireless networks. Be hired from communication operators and can always offer a bandwidth larger than PLC network while smaller than EPON. The most common type is CDMA networks which can provide transmission at a rate of hundreds kilobytes.

Private networks. PLC networks belong to the power companies themselves, and can provide traffic flow with bandwidth of merely dozens of kilobytes. The advantage of this option is free for use.

6.2 Result Analysis

Due to the lack of cost-constrained routing protocols in SDG communication networks, we compare the performance of CDCS with a random scheduling strategy (RSS), and a traditional Lyapunov-based scheduling strategy (LSS) with the same objective as CDCS.

6.2.1 Overall System Cost

In the first scenario, we value these three scheduling strategies by varying arrival rates λ_i of each traffic class *i*. By comparing the results in Figure 3(a-c), we can observe the improvement: the total financial cost in RSS mode is much higher than in LSS and CDCS. Besides, the accumulation of cost in CDCS and LSS modes is proved to be fairly smoother than in RSS mode. This is because packets are routed arbitrarily in RSS mode, which results in uncontrollable costs. The strategies derived from the *Lyapunov* theories can optimize the financial cost throughout the system evolution process.

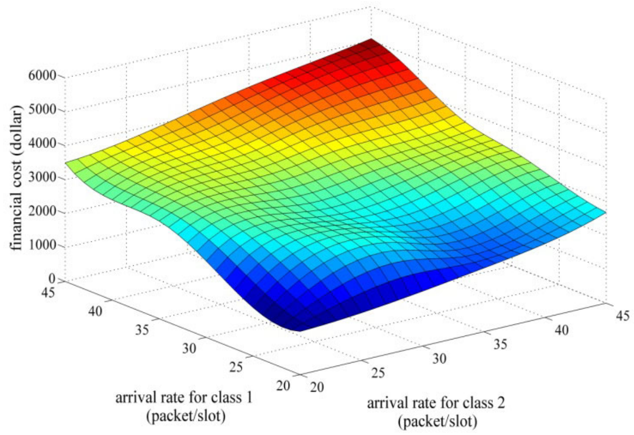

Figure 3. (a) Overall system cost under RSS

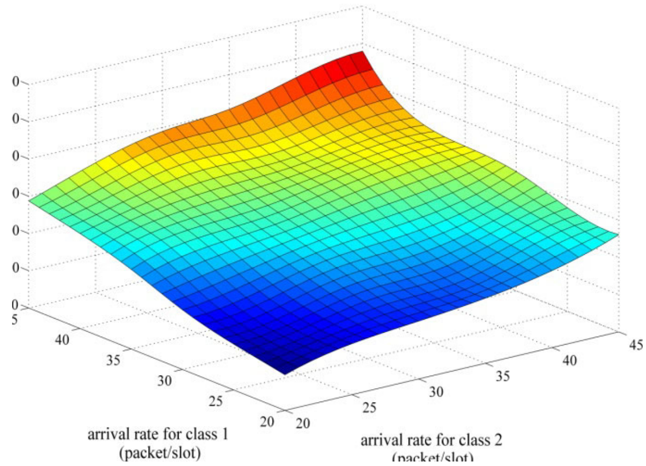

Figure 3. (b) Overall system cost under LSS

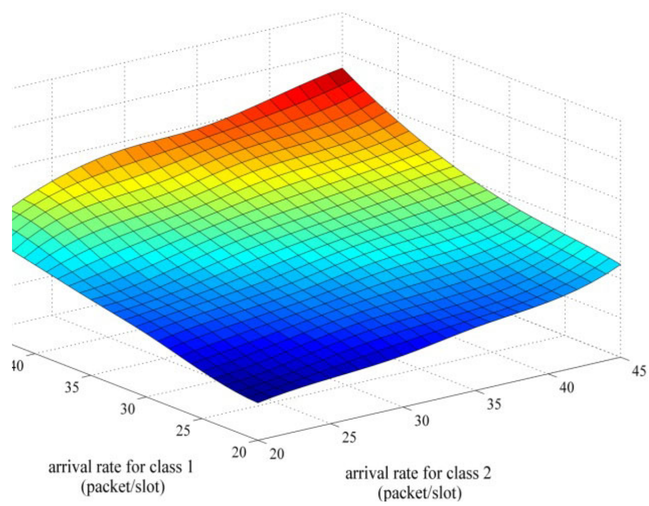

Figure 3. (c) Overall system cost under CDCS

To observe it more clearly, we compare these three scheduling strategies by reporting the overall cost with fixed arrival rates: $\lambda_1 = 40$, $\lambda_2 = 40$, $\lambda_3 = 4$, packet/slot. The arrival rate set in the second scenario can make the system stable while keeping the system offering continuous service in most of time. It can be seen from the Figure 4 that the accumulation of cost is uneven in operation process. The CDCS outperforms LSS by approximately 5 percent with the introduction of service rate adjustment control. It is because that $b_j(t)$ change adaptively with time-varying output queue backlogs, which leads to the lower average rate under CDCS than LSS.

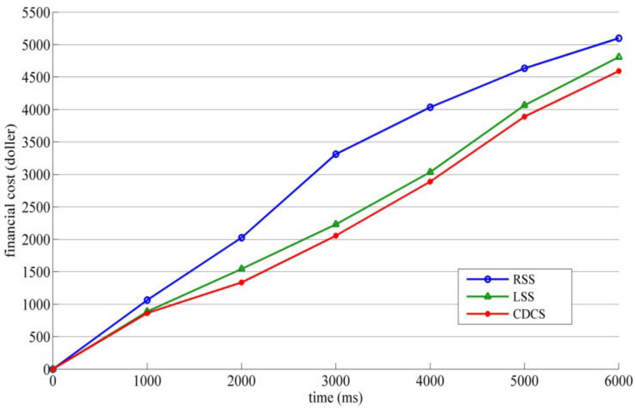

Figure 4. Overall system cost

6.2.2 Average System Delay

As discussed in network settings, each service in SDG communication networks has its own delay requirement listed in Table 2. The impact on delay of these three classes is compared among different scheduling strategy modes. Figure 5 demonstrates the average system delay at fixed arrival rates: $\lambda_1 = 40$, $\lambda_2 = 40$, and $\lambda_3 = 4$.

When aiming at costs, the strategies will cause higher transmission delay since the network controller tends to allocate packets into lower-cost networks in spite of more transmission time. Meanwhile, packets randomly routed under RSS without concerning financial cost experience the lowest transmission delay. However, compared with the traditional LSS, the CDCS can lower transmission delay with the same optimization objective. It can be explained that the network controller allocates more packets to output networks with less backlogs, which will decrease load pressure and waiting time in output queues.

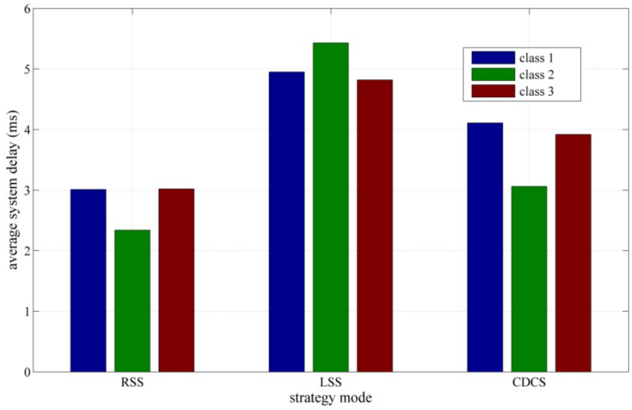

Figure 5. Average system delay

6.2.3 Influence of Various V

Given that LSS and CDCS are both based on *Lyapunov* theories, the predefined control parameter V can affect the average cost. Figure 6 reveals that the larger value assigned to V, the closer optimal value approched. However, if V is defined too large, the average backlog in our system will increase significantly as shown in Figure 7. So, a proper value for control parameter V is vital to balance the system operation and financial cost. The simulation results are well agreed with the *Lyapunov* optimization model proposed in Section 5.

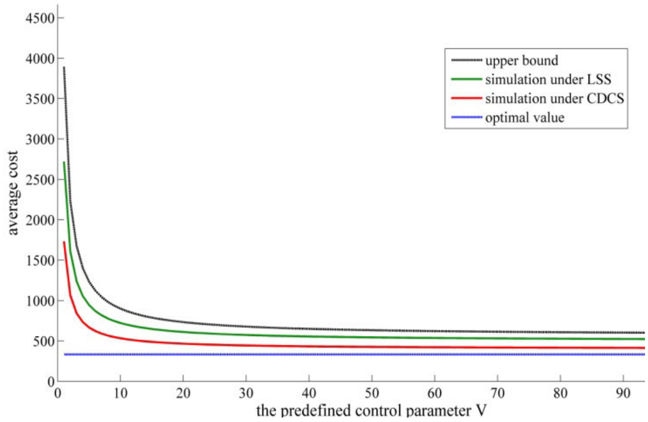

Figure 6. Average cost versus V

Table 2. System parameters for input queues

Input queue	Applications	Packet size (unit)	Priority	Arrival rate λ_i (packet/slot) (adjustable)	Arrival Probability	Delay Constraint (slot)
1	Tele-protection SCADA	5 4	1	40	0.4 0.6	10
2	Automated demand response	3	2	30	1	20
3	Smart metering	5	3	4	1	200

Table 3. System parameters for output networks

Output network	Networking technology	Delivery delay d_j (slot)	rated speed (unit/slot)	Financial Cost under rated speed b_j (dollar/slot)	Maximum service rate b_j^{max} (unit/slot)	minimum service rate b_j^{min} (unit/slot)
1	EPON	0.0325	200	4	220	180
2	CDMA	0.2	100	1	120	80
3	PLC	2	50	0	50	70

7 Conclusions

This paper presents a network model consisting of input queues and output queues that store heterogeneous service traffic and representing the different options for access networks respectively. The traffic scheduling problem is formulated as a Lyapunov-based optimization problem with the objective of minimizing the financial cost and be subject to delay constrains. The cost-optimal distributed control strategy (CDCS) is proposed to allocate packets to various output networks based on the current state of queue buffers, channel conditions, as well as transmission costs. The two control processes, which are output network selection and service rate adjustment, jointly optimize the routing and financial cost. Performance results prove that the proposed CDCS can save financial cost while meeting delay constrains.

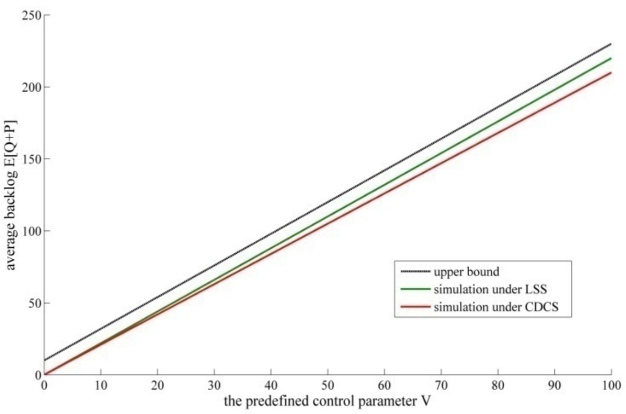

Figure 7. Average backlog versus V

Acknowledgements

This work is supported by the Science and Technology Project of State Grid Corporation of China (SGIT0000 KJJS1500008)

References

[1] E. Hossain, Z. Han, H. V. Poor, *Smart Grid Communications and Networking*, Cambridge University Press, 2012.

[2] G. A. Shah, V. C. Gungor, O. B. Akan, A Cross-layer Design for QoS Support in Cognitive Radio Sensor Networks for Smart Grid Applications, *Proc. IEEE ICC*, Ottawa, Canada, 2012, pp. 1378-1382.

[3] G. A. Shah, V. C. Gungor, O. B. Akan, A Cross-Layer QoS-Aware Communication Framework in Cognitive Radio Sensor Networks for Smart Grid Applications, *Industrial Informatics IEEE Transactions on*, Vol. 9, No. 3, August, 2013, pp. 1477-1485.

[4] Q.-D. Ho, Y. Gao, T. Le-Ngoc, Challenges and Research Opportunities in Wireless Communication Networks for Smart Grid, *IEEE Wireless Communications*, Vol. 20, No. 3, pp. 89-95, June, 2013.

[5] M. Levorato, U. Mitra, Optimal Allocation of Heterogeneous Smart Grid Traffic to Heterogeneous Networks, *Proc. IEEE SmartGridComm*, Brussels, Belgium, 2011, pp. 132-137.

[6] A. Zaballos, A. Vallejo, J. M. Selga, Heterogeneous Communication Architecture for the Smart Grid, *Network IEEE*, Vol. 25, No.5, pp. 30-37, September/October, 2011.

[7] G. Heydt, The Next Generation of Power Distribution Systems, *Smart Grid IEEE Transactions on*, Vol. 1, No. 3, pp. 225-235, December, 2010.

[8] M. J. Neely, E. Modrano, C.-P. Li, Fairness and Optimal Stochastic Control for Heterogeneous Networks, *IEEE/ACM Transactions on Networking*, Vol. 16, No. 2, pp. 396-409, April, 2008.

[9] L. Badia, M. Miozzo, M. Rossi, M. Zorzi, Routing Schemes in Heterogeneous Wireless Networks Based on Access Advertisement and Backward Utilities for QoS Support, *IEEE Communications Magazine*, Vol. 45, No. 2, pp. 67-73, May, 2007.

[10] H.-P. Shiang, M. van der Schaar, Queuing-Based Dynamic Channel Selection for Heterogeneous Multimedia Applications Over Cognitive Radio Networks, Multimedia IEEE *Transactions on*, Vol. 10, No. 5, pp. 896-909, August, 2008.

[11] Y. Shi, Y. T. Hou, A Distributed Optimization Algorithm for Multi-Hop Cognitive Radio Networks, *Proc. IEEE INFOCOM*, Phoenix, AZ, 2008, pp. 1292-1300.

[12] L. Zhou, B. Gever, X. Wang, A. Wei, B. Zheng, H.-C. Chao, Multi-user Video Streaming over Multiple Heterogeneous Wireless Networks: A Distributed, Cross-Layer Design Paradigm, *Journal of Internet Technology*, Vol. 10, No. 1, pp. 1-11, January, 2009.

[13] M. J. Neely, Delay-Based Network Utility Maximization, *IEEE/ACM Transactions on Networking*, Vol. 21, No. 1, pp. 41-54, June, 2013.

[14] M. J. Neely, *Stochastic Network Optimization with*

Application to Communication and Queueing Systems, Morgan & Claypool, 2010.
[15] C.-P. Li, M. J. Neely, Delay and Rate-optimal Control in A Multi-class Priority Queue with Adjustable Service Rates, *Proc. IEEE INFOCOM*, Rio de Janeiro, Brail, 2012, pp. 2976-2980.
[16] A. G. Marques, G. B. Giannakis, J. Ramos, Stochastic Cross-layer Resource Allocation for Wireless Networks Using Orthogonal Access: Optimality and Delay Analysis, *Proc. IEEE ICASSP*, Austin, TX, 2010, pp. 3154-3157.
[17] IEEE, 1646-2004. *IEEE Standard Communication Delivery Time Performance Requirements for Electric Power Substation Automation*, IEEE, 2005.
[18] International Telecommunication Union, *One-Way Transmission Time*, ITU-T Standard G.114, http://www.itu. int/rec/T-REC-G.114-200305-I/en.

Biographies

Si-Ya Xu received the B.E. degree from University of Science and Technology Beijing, China, in 2010. She is currently a Ph. D. student of Beijing University of Posts and Telecommunications, China. Her research interests include Smart Grid and Network Management. Email: xusiyaxsy@hotmail.com

Shao-Yong Guo received the Ph.D. degree from Beijing University of Posts and Telecommunication, Beijing, China, in 2013. He is currently a lecturer in Beijing University of Posts and Telecommunication. His research interests include device management, Internet of Things, Ubiquitous Network and Smart Grids. Email: syguo@bupt.edu.cn

Zhu Liu received the M.S. degree from Wuhan University of Technology, Wuhan, Hubei, in 2007. He is currently a senior engineer of State Grid Information & Telecommunication Co., Ltd., China. His research interests include Information Communication in Power System. Email: liuzhu_558 @163.com

Lei Wei is a senior engineer of Information & Telecommunication Branch of Nanjing Power Supply Company of State Grid Jiangsu Electric Power Company, China. His research interests include Information Communication in Power System and Smart Grid. Email: nj_weilei@163. com

Luo-Ming Meng received the M.S. degree from Tsinghua University, Beijing, China, in 1987. He is currently a professor and Ph.D. supervisor. He is the director of Communications Software Technical Committee of China Institute of Communications and the chairman of the National Network Management Standards Study Group. Email: lmmeng@bupt.edu.cn

Appendix

Theorem:

The virtual queues $Z_k(t)$ can enforce the $\lim_{t\to\infty} \overline{y_k(t)} - \overline{y_k} \le 0$ constraint.

Proof:

For any discrete time queuing system described by (20) and for two slots t_1 and t_2 such that $0 \le t_1 < t_2$, we have

$$Z_k(t_2) - Z_k(t_1) \ge \sum_{\tau=t_1}^{t_2-1} y_k(\tau) - \overline{y_k}$$

By substituting $t_1 = t$ and $t_2 = 0$, and dividing by t, we have the inequality below for any $t > 0$

$$\frac{Z_k(t_2)}{t} - \frac{Z_k(0)}{t} \ge \sum_{\tau=0}^{t-1} y_k(\tau) - \overline{y_k}$$

Suppose $Z_k(0)$ is non-negative and finite, and take expectation of the above and $t \to \infty$ shows

$$\lim_{t\to\infty} \frac{E(Z_k(t))}{t} \ge \lim_{t\to\infty} \sup \overline{y_k}(t) - \overline{y_k}$$

where we recall that $\overline{y_k}(t)$ is the time average expectation of $y_k(t)$ over $t \in \{0, 1, 2, ...\}$. Thus, if $Z_k(t)$ is mean rate stable, we have

$$\lim_{t\to\infty} \frac{E[Z_k(t)]}{t} = 0,\ k = 1, 2, ..., L$$

Hence

$$\lim_{t\to\infty} \sup \overline{y_k}(t) - \overline{y_k} \le 0,\ K = 1, 2, ..., L$$

This means that our desired time average constraint for $y_k(t)$ is satisfied. It turns the problem of satisfying a time average inequality constraint into a pure queue stability problem.

Proof completed.