Cross-layer Scheduling with Secrecy Demands in Delay-aware OFDMA Network

Xingzheng Zhu[1], Bo Yang[1,2,*] and Xinping Guan[1]
[1]Department of Automation, Shanghai Jiao Tong University
and Key Laboratory of System Control and Information Processing, Ministry of Education of China
[2] Key Laboratory of Education Ministry for Image Processing and Intelligent Control
* Corresponding Author
Email: {wendyzhu, bo.yang, xpguan}@sjtu.edu.cn

***Abstract*—An Orthogonal Frequency Division Multiple Access (OFDMA) downlink system with secure transmissions and delay constraints is investigated, in which a base station (BS) transmits both open data with delay limitations and private data with secrecy demands to each user. The time-varying channel is modeled as slow-fading and all channel state information (CSI) is assumed to be known to BS. An online cross-layer scheduling algorithm composed of flow control and resource allocation is proposed in this paper to maximize the downlink throughput with delay and power constraint. Furthermore, the algorithm makes decisions on current CSI instead of channel statistics. In addition, virtual queues of delay and power are constructed to track them so that those time-average constraints are fulfilled. In the end, it is proven by means of Lyapunov optimization technique that our algorithms will obtain a performance which can be extremely close to optimality and reduce operation complexity significantly.**

I. Introduction

In recent decades, OFDMA has become a practical technology in high speed wireless communication networks. Because OFDM allows that the subcarrier power and modulation scheme can be individually controlled for each carrier, OFDMA networks have perfect characteristics of high speed efficiency and multipath tolerance [1]. Many researches focus on resource allocation in OFDMA networks to maximize the network capacity [2]–[6]. [2] has found that the data rate of a OFDMA system is maximized when each subcarrier is assigned to the user with the best channel gain for that subcarrier. [3] solves OFDMA downlink resource allocation problem under total power constraint using Lagrange dual decomposition. [4] studies the resource allocation policy joint with flow rate control with secrecy demands in OFDMA networks. In this paper, we consider a more practical communication situation with secrecy transmissions and delay limitations in OFDMA networks. Besides, time-average power constraint is also proposed in consideration of energy saving, and our goal is to maximize the total weighted throughput of open data and private data.

This work was supported partially by the NSFC under Grant 61174127, 61104033, 60974123, 61273181, 61203104, the NSFC key Grant with No. 60934003, the National Basic Research Project of China under Grant 2010CB731800, the Research Found for the Doctoral Program of Higher Education under Grant 20110073120025, 20110073130005, the Scientific research plan projects of Hebei Education Department with grant No. Q2012088 and the Cyber Joint Innovation Center.

Security has always been an important issue in wireless communication. Shannon's information theory lays the foundation for information-theoretic security [7]. [8] proposes the concept of wire-tap channel and eavesdropper. Furthermore, it also proves that if the receiver enjoys a better channel condition than passive eavesdropper there would exist a positive data rate within which receiver's data can be transmitted securely [8]. As a result, many works such as [9] have used artificial noise to degrade eavesdropper's channel state. Taking artificial noise generation for secrecy transmissions into account, [10] studies the resource allocation in multi-user multi-subcarrier system. Encoding is another effective way to exploit secrecy capacity. [11] proposes a joint encoding of private and open message model and achieve an optimal network capacity. [12] investigates the power and subcarrier allocation policy for OFDMA networks where open and secrecy transmission coexist while emphasized users' secrecy demands from the aspect of secrecy capacity.

In practice, when a user asks BS for data, it is significant to consider delay performance in downlink transmission to ensure user experience. [13] summarizes three approaches to deal with delay-aware resource allocation in wireless networks and their performance. Those three methods are based on large deviation theory, Markov decision theory and stochastic Lyapunov drift. As to the first two methods, they have to know some statistics knowledge on channel state and random arrival data rate to design algorithm, while these prior knowledge is expensive to get, even unavailable. To overcome this problem, many authors pay attention to Lyapunov drift method. [14] and [15] investigate scheduling in multi-hop wireless networks and resource allocation in cooperative communications respectively as two typical applications of Lyapunov optimization and virtual queue technique in delay-limited system .

In this paper, we consider a joint model to describe the simultaneous transmission of private and open data . On the other hand, considering time-average power limitation and delay constraints, the virtual power queue and the virtual delay queues are constructed. In the meantime, Lyapounov optimization is introduced to design the dynamic cross-layer optimization algorithms. In this way, the weighted throughput maximization problem is decoupled into a two-layer optimization algorithms composed of flow control and resource

978-1-4673-5939-9/13/$31.00

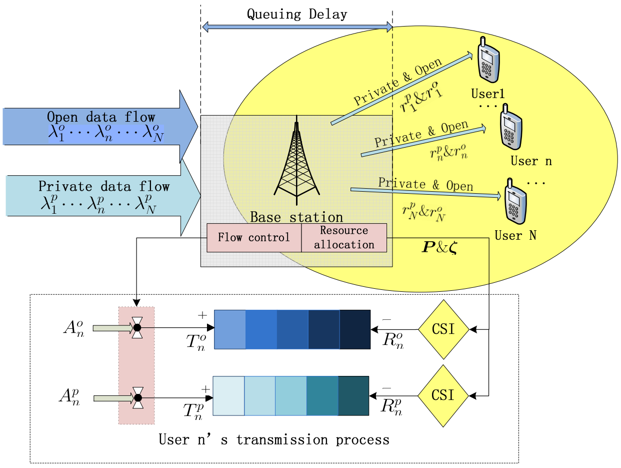

Fig. 1. System model

allocation, and this algorithms can operate using current CSI and queue state information instead of statistics.

This paper is organized as follows. In Section II, we introduce the system model and relevant constraints in detail. Section III formulates the problem. In Section IV, we introduce our cross-layer optimization algorithm. We give the performance bound and stability result in Section V. In Section VI, some simulation results are shown. Finally, we conclude this paper in Section VII.

II. System Model

We consider a downlink system that consists of one BS and N users which can be regarded as mobile terminals in an OFDMA network as shown in Fig. 1. The BS transmits both private and open data to each user. The private data has security requirement and open data has delay constraint. The total bandwidth is divided into M orthogonal subcarriers using OFDM. The users and subcarriers are indexed respectively as $n \in \{1, 2, \cdots, N\}$ and $m \in \{1, 2, \cdots, M\}$.

This system operates in slotted time. The length of a timeslot is T and we use slot t to denote the time interval $[tT, (t+1)T)$ for $t \in \{1, 2, \cdots\}$. Let $\alpha_{n,m}(t)$ denote the channel-to-noise ratio (CNR) of user n in subcarrier m in slot t and the set $\boldsymbol{\alpha}(t) = \{\alpha_{n,m}(t), \forall n, \forall m\}$ represents the system CSI. All channels are assumed to be slow fading [16], thus $\boldsymbol{\alpha}(t)$ remains fixed during one slot and changes between two. We assume that perfect CSI can be got by the BS at the beginning of every timeslot.

In the side of BS, the amount of open data of user n, $A_n^o(t)$, and private data, $A_n^p(t)$ that arrive at BS during slot t are independent identically distributed (i.i.d) Bernoulli processes with average arrival rates λ_n^o and λ_n^p, and their upper bounds are μ_{max} and A_{max}, respectively. The amount of open and private data admitted by BS entering the queue are denoted by $T_n^o(t)$ and $T_n^p(t)$. BS is privileged to make decision on the values of $T_n^o(t)$ and $T_n^p(t)$ according to a certain principle which would be specified in section IV.

BS is also in charge of resource allocation. Let $\Omega_n(t) = \{\omega_{n,m}(t), \forall m\}$ denote the subcarrier assignment policy of user n, where $\omega_{n,m}(t)$ is either 1 representing subcarrier m is assigned to user n, or 0 otherwise. Then let $\Omega(t) = \{\Omega_n(t), \forall n\}$ be the overall policy. We set limitations as follows to ensure that each subcarrier can be assigned to no more than one user:

$$0 \leq \sum_{n=1}^{N} \omega_{n,m}(t) \leq 1, \qquad \forall m \tag{1}$$

$\boldsymbol{P}(t) = \{P_{n,m}(t), \forall n, \forall m\}$ denotes the overall system power allocation policy and $P_{n,m}$ represents the power allocated by BS to user n in subcarrier m.

A. Power consumption constraint

In the following discussion, the time index (t) is omitted for simplicity. We define $E = \sum_{\forall n, \forall m} P_{nm}$ as total power consumption of the whole system in one time slot. There exists a physical peak power limitation P_{max} that E cannot exceed:

$$0 \leq E \leq P_{max} \tag{2}$$

The time-average power consumption also has an upper bound P_{avg}, which is proposed for energy conservation:

$$e \quad \leq \quad P_{avg} \tag{3}$$

where $e = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E(\tau)$

B. Security Model

To user n, the achievable transmission data rate is: $R_n = \sum_{m\in\Omega_n} R_{nm}$ where $R_{nm} = \log(1 + \alpha_{nm} P_{nm})$.

In this paper, we consider a situation that open and private data of one user can be transmitted simultaneously. For each user, BS makes decision if his secure data could be transmitted in this slot and this decision is expressed as the secure transmission control vector $\boldsymbol{\zeta} = \{\zeta_n, \forall n\}$. The indicator variable $\zeta_n = 1$ implies that private and open messages are encoded at rate $\hat{R}_n^p$ and $R_n - \hat{R}_n^p$ respectively in timeslot t and $\zeta_n = 0$ means that only open messages can be transmitted at rate R_n.

When user n is transmitting messages, all the other users except itself are treated as potential eavesdroppers. According to [11], subject to perfect privacy of user n, the *instantaneous privacy rate* of user n on subcarrier m is: $\hat{R}_{nm}^p = [\log(1 + \alpha_{nm} P_{nm}) - \log(1 + \beta_{nm} P_{nm})]^+$, where $[\cdot]^+ = \max\{\cdot, 0\}$ and $\beta_{nm} = \max\limits_{n', n' \neq n} \alpha_{n',m}$. Obviously, $\hat{R}_n^p = \sum\limits_{m\in\Omega_n} \hat{R}_{nm}^p$. Thus the achievable privacy rate of user n is:

$$R_n^p = \zeta_n \hat{R}_n^p$$

and directly the open rate of user n is: $R_n^o = R_n - R_n^p$

C. Flow rate model

In BS, there exists actual data queue of open and private data which are presented by Q_n^o and Q_n^p for all $n \in \{1, \cdots, N\}$. These queues are updated as follows:

$$Q_n^o(t+1) = [Q_n^o(t) - R_n^o(t)]^+ + T_n^o(t) \tag{4}$$
$$Q_n^p(t+1) = [Q_n^p(t) - R_n^p(t)]^+ + T_n^p(t) \tag{5}$$

All Q_n^o and Q_n^p have initial values of zero. We define $t_n^o = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{T_n^o(\tau)\}$, $t_n^p = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{T_n^p(\tau)\}$ as time-average admission rates of open data and private data respectively. The time-average service rate of open and private data are also defined as: $r_n^o = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{R_n^o(\tau)\}$, $r_n^p = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{R_n^p(\tau)\}$.

As a discrete time process, $Q_n(t)$ which is similar to either $Q_n^o(t)$ or $Q_n^p(t)$ is *strongly stable* if:

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t-1}\mathbb{E}\{Q_n(\tau)\} \le \infty \tag{6}$$

In particular, a multi-queue network is stable when all queues of the network are *strongly stable*. Thus, $Q_n^o(t)$ and $Q_n^p(t)$ should be kept *strongly stable* in order to ensure the system stability. Furthermore, according to *Rate Stability Theorem* in [17], Q_n^o and Q_n^p are stable if and only if $t_n^o \le r_n^o$ and $t_n^p \le r_n^p$ hold.

To stabilize the network system, the key point is to avoid blocking of data at BS. BS should take full use of the data rate supplied by physical layer (PHY) with security concerns to send out the data admitted by itself. However, if the downlink system can't supply sufficient data rate under poor channel conditions while the external data is arriving at BS continuously, BS should be wise enough to adjust admission rate to avoid blocking.

In this paper, virtual queues of open date, $X_n^o(t)$, and private data $X_n^p(t)$ are introduced as (7) and (8) to assist in developing our algorithms that guarantee the actual queues Q_n^o and Q_n^p are bounded deterministically in the worst case.

$$X_n^o(t+1) = [X_n^o(t) - T_n^o(t)]^+ + \mu_n^o(t) \tag{7}$$
$$X_n^p(t+1) = [X_n^p(t) - T_n^o(t)]^+ + \mu_n^p(t) \tag{8}$$

Denote μ_n^o and μ_n^p as the virtual admission rate of open data and private data and we upper bound them by processes A_n^o and A_n^p respectively. Notice that X_n^o, X_n^p, μ_n^o and μ_n^p don't stand for any actual queue and data. They are only controlled by the proposed algorithms. According to queuing theory, when X_n^o and X_n^p are stable, the time-average value of μ_n^o and μ_n^p would satisfy:

$$\nu_n^o = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{\mu_n^o(\tau)\} \le t_n^o \tag{9}$$

$$\nu_n^p = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{\mu_n^p(\tau)\} \le t_n^p \tag{10}$$

D. Delay-limited model

Each user has a time-average queuing delay ρ_n^o for it's open data. The queuing delay is defined as the time a packet waits in a queue until it can be transmitted. To each user, it proposes a delay constraint ρ_n as in (11) for its open data.

$$\rho_n^o \le \rho_n \tag{11}$$

III. Problem Formulation

From above descriptions, especially in (9) and (10), we should notice that when virtual queues X_n^o and X_n^p are stable, the virtual admission rate ν_n^o and ν_n^p are the lower-bound of admission rate t_n^o and t_n^p, respectively. Our optimization objective is to maximize the sum weighted admission rates of all users by maximizing the weighted virtual admission rates and stabilizing X_n^o and X_n^p at the same time. Let θ_n and φ_n for all n be a collection of no-negative weights of private and open data throughput respectively. Basically, we think private data is more important than open data to every user, thus it is set that $\theta_n > \varphi_n$ by default. Then the optimal problem can be formulated as:

$$\begin{aligned}
\text{Maximize} \quad & \sum_{n=1}^{N}\{\theta_n\nu_n^p + \varphi_n\nu_n^o\} \qquad (12)\\
\text{Subject to:} \quad & 0 \le \nu_n^p \le \lambda_n^p, \forall n\\
& 0 \le \nu_n^o \le \lambda_n^o, \forall n\\
& (3), (11)\\
& \mathbf{r} = \{\nu_n^p, \nu_n^o : \forall n\} \in \Lambda
\end{aligned}$$

where Λ is the *network capacity region*. When BS takes a kind of control policy under a certain channel condition, the network system will have a certain network capacity and the *network capacity region* is the set of network capacities under all possible control policies and all channel conditions. In the proposed system, the control policy should fulfill subcarrier assignment rule (1), peak power constraint (2) and stabilizing all queues including actual queues and virtual queues.

Theoretically, we can get the optimal solution to (12) if we get the distribution of CSI and external data arrival rate beforehand. However, it will be a huge payment to get those statistics. In this paper an online algorithm requiring only current information of queue state and channel state is proposed and will be described in detail then.

IV. Online Control Algorithm

Before we introduce the control algorithm in detail, it's worth to notice that the original problem has a time-average limitation on power consumption and queuing delay. Using the technique similar to [17], we construct power virtual queue Y and delay virtual queue Z_n to track the power consumption and queuing delay respectively. These virtual queues don't exist in practice, and they are just generated by the iterations of (13) and (14):

$$Y(t+1) = [Y(t) - P_{avg}]^+ + E(t) \tag{13}$$
$$Z_n(t+1) = [Z_n(t) - \rho_n\mu_n]^+ + Q_n^o(t) \tag{14}$$

Similar to actual queues, Y and Z_n have initial values of zero. According to queuing theory, if Y is stable, constraint (3) is satisfied. In addition, by Little's Theorem and queuing theory, when Z_n is stable, the delay constraint (11) would be achieved. It will be proven that the proposed optimal control algorithm can stabilize these queues in section V, that is to say the time-average constraints are fulfilled.

Using virtual queues X_n, Z_n and Y, we decouple problem (12) into two parts: one is *flow control algorithm* which decides the admission of data, and another is *resource allocation algorithm* in charge of subcarrier assignment, power allocation and secure transmission control in every slot.

A. Flow control algorithm

When external data arrives at BS, BS will decide whether to admit it according to queue lengthes. Let $q^o_{max} \geq \mu_{max}$ and $q^p_{max} \geq A_{max}$. They are actually the deterministic worst case upper bound of relative queue length to be shown later. The flow control rules of open data and private data are obtained by solving (15) and (16) respectively:

$$\begin{aligned} &\text{Minimize} && T^o_n[Q^o_n - q^o_{max} + \mu_{max}] && (15)\\ &\text{Subject to:} && 0 \leq T^o_n \leq A^o_n \\ &\text{Minimize} && T^p_n[Q^p_n - q^p_{max} + A_{max}] && (16)\\ &\text{Subject to:} && 0 \leq T^p_n \leq A^p_n \end{aligned}$$

The corresponding solutions to (15) and (16) are easy to get:

$$T^o_n = \begin{cases} 0 & \text{if} \quad Q^o_n - q^o_{max} + \mu_{max} \geq 0 \\ A^o_n & \text{otherwise} \end{cases} \tag{17}$$

$$T^p_n = \begin{cases} 0 & \text{if} \quad Q^p_n - q^p_{max} + A_{max} \geq 0 \\ A^p_n & \text{otherwise} \end{cases} \tag{18}$$

Here we can have an intuitive explaination on flow control rules. They work like valves. When any actual data queue exceeds some threshold, the corresponding valve would turn off and no data would be admitted.

As to virtual variables μ^o_n and μ^p_n, there are also their respective virtual flow control algorithms (19) and (20) so as to update virtual queue X^o_n and X^p_n which will play an important role in resource allocation. Let V be a fixed non-negative control parameter:

$$\begin{aligned} &\text{Minimize} && \mu^o_n[\frac{q^o_{max} - \mu_{max}}{q^o_{max}} X^o_n - \rho_n Z_n - V\varphi_n] && (19)\\ &\text{Subject to:} && 0 \leq \mu^o_n \leq A^o_n \\ &\text{Minimize} && \mu^p_n[\frac{q^p_{max} - A_{max}}{q^p_{max}} X^p_n - V\theta_n] && (20)\\ &\text{Subject to:} && 0 \leq \mu^p_n \leq A^p_n \end{aligned}$$

Solutions to (19) and (20) are (21) and (22) respectively:

$$\mu^o_n = \begin{cases} 0 & \text{if} \quad (\frac{q^o_{max} - \mu_{max}}{q^o_{max}} X^o_n - \rho_n Z_n - V\varphi_n) \geq 0 \\ A^o_n & \text{otherwise} \end{cases} \tag{21}$$

$$\mu^p_n = \begin{cases} 0 & \text{if} \quad (\frac{q^p_{max} - A_{max}}{q^p_{max}} X^p_n - V\theta_n) \geq 0 \\ A^p_n & \text{otherwise} \end{cases} \tag{22}$$

B. Resource allocation algorithm

The resource allocation policy can be found in solving the convex optimization problem PS.

$$\text{PS: Maximize} \sum_{n=1}^{N} \frac{X^o_n Q^o_n}{q^o_{max}} R^o_n + \sum_{n=1}^{N} \frac{X^p_n Q^p_n}{q^p_{max}} R^p_n - YE$$
$$\text{Subject to:} (1), (2)$$

At the beginning of every slot, all $X^o_n, X^p_n, Q^o_n, Q^p_n$ and Y can be regarded as constants because they all have been decided in the previous slot. Notice that, the solution to PS is determined at the beginning of every slot and all queues are updated at the end of every slot. In addition, PS can be solved by using dual decomposition, which is omitted here due to space limitation and the main results are given below. The detail process can be found in [18].

Let δ denote the nonnegative Lagrange multiplier for the peak power constraint in problem PS. When δ is fixed, the optimality condition of secrecy transmission control is:

$$\zeta^*_n = \begin{cases} 1 & \text{if} \quad (\frac{X^p_n Q^p_n}{q^p_{max}} - \frac{X^o_n Q^o_n}{q^o_{max}}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

The optimal power allocation result is:

$$P^*_{nm} = [\frac{\sqrt{\triangle} - A_3(\alpha_{nm} + \beta_{nm}) + (A_2 - A_1)\alpha_{nm}\beta_{nm}}{2 \cdot A_3 \alpha_{nm}\beta_{nm}}]^+ \tag{24}$$

where $A_1 = \frac{X^o_n Q^o_n}{q^o_{max}} + (\frac{X^p_n Q^p_n}{q^p_{max}} - \frac{X^o_n Q^o_n}{q^o_{max}})\zeta^*_{nm}$, $A_2 = (\frac{X^p_n Q^p_n}{q^p_{max}} - \frac{X^o_n Q^o_n}{q^o_{max}})\zeta^*_{nm}$, $A_3 = (\delta + Y)\ln 2$, $\triangle = [[A_3(\alpha_{nm} + \beta_{nm}) - (A_1 - A_2)\alpha_{nm}\beta_{nm}]^2 + 4A_3\alpha_{nm}\beta_{nm}(A_3 - A_1\alpha_{nm} + A_2\beta_{nm})]^+$.

Then substituting (23) and (24) into $j_{nm}(\delta, \boldsymbol{P}, \zeta) = \frac{X^o_n Q^o_n}{q^o_{max}} R^o_{nm} + \frac{X^p_n Q^p_n}{q^p_{max}} R^p_{nm} - (Y + \delta)P_{nm}$. For any subcarrier m, it will be assigned to the user who has the biggest $j_{n,m}$. Let n^*_m be the result of subcarrier m's assignment which is given by:

$$n^*_m = \arg\max_n j_{n,m}, \forall n \text{ and } \omega^*_{n,m} = \begin{cases} 1 & \text{if } n = n^*_m \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

Let $E^* = \sum_{n=1}^{N} \sum_{m=1}^{M} P^*_{n,m}\omega^*_{n,m}$. As to the value of δ, we use subgradient method to update it as in (26),

$$\delta(i+1) = [\delta(i) - \varsigma(P_{max} - E^*(t,i))]^+ \tag{26}$$

where ς is step size which should be a small positive constant. In addition, index i stands for iteration number.

The resource allocation is completed when the subgradient method converges. Table I show the whole process of the proposed algorithm. We use a small positive number δ_c denote the converge condition of subgradient on δ.

From the above description, we can find some principles of resource allocation. In (23), virtual queues of open as well as private data reflect the gap between the corresponding user's demand on data rate and the data rate that system can provide. Thus, $\frac{X^o_n Q^o_n}{q^o_{max}}$ and $\frac{X^p_n Q^p_n}{q^p_{max}}$ can be regarded as the transmission urgency of open data and private data, respectively. To each user, we can find that the transmissions of open and private data share the limited resource. Only when the transmission urgency of private data exceed open data, the BS would allocate some resource to transmit private data, otherwise, the BS would use the user's entire resource to transmit open data due to delay constraint. In (24), it is easy to find that a bigger Y results in less power allocated to every user which will reduce the system power consumption.

TABLE I

ALGORITHM IMPLEMENTATIONS

Proposed online control algorithm in timeslot t
1) Flow control:
Use (17),(18),(21),(22) to calculate T_n^o, T_n^p, μ_n^o and μ_n^p respectively.
2) Resource allocation:
a) Set the Lagrange multiplier $\delta = \delta_{ini}$, (δ_{ini}: A initial value of δ).
b) For each (n, m)
i) Use (23) to calculate ζ_n^*.
ii) Use (24) to calculate P_{nm}^*.
iii) Use (25) to calculate ω_{nm}^*.
c) Use (26) to update δ.
d) If $\lvert\delta\rvert > \delta_c$, goto **b)**, else proceed.
3) Update the queues:
Use (4),(5),(7),(8),(13) and (14) to update all queues including Q_n^o, Q_n^p, X_n^o, X_n^p, Y, Z_n.

V. ALGORITHM PERFORMANCE

First, it's necessary to introduce $\mathbf{r}^* = (\nu_n^{p,*}, \nu_n^{o,*})$ and $\mathbf{r}^*(\epsilon) = (\nu_n^{p,*}(\epsilon), \nu_n^{o,*}(\epsilon))$ for analysis, they are the respective solutions to (27) and (28):

$$\max_{\mathbf{r}:\mathbf{r}\in\Lambda} \sum\nolimits_{n=1}^{N} \theta_n \nu_n^p + \varphi_n \nu_n^o \quad \text{subject to} \quad e \le P_{avg} \tag{27}$$

$$\max_{\mathbf{r}:\mathbf{r}+\epsilon\in\Lambda} \sum\nolimits_{n=1}^{N} \theta_n \nu_n^p + \varphi_n \nu_m^o \quad \text{subject to} \quad e \le P_{avg} \tag{28}$$

According to [19], it is true that:

$$\lim_{\epsilon\to 0} \sum_{n=1}^{N} \{\theta_n \nu_n^{p,*}(\epsilon) + \varphi_n \nu_n^{o,*}(\epsilon)\} = \sum_{n=1}^{N} \{\theta_n \nu_n^{p,*} + \varphi_n \nu_n^{o,*}\} \tag{29}$$

Theorem 1 and *Theorem 2* state the algorithm performance.

Theorem 1: Employing the proposed algorithm, both actual queues of open data $Q_n^o(t)$ and privacy data $Q_n^p(t)$ in BS have deterministic worst-case bounds:

$$Q_n^o(t) \le q_{max}^o, Q_n^p(t) \le q_{max}^p, \forall t, \forall n \tag{30}$$

Theorem 2: Given that

$$q_{max}^o > \mu_{max} + \frac{{C_{max}^o}^2 + \mu_{max}^2}{2\epsilon}, \tag{31}$$

$$q_{max}^p \ge A_{max} + \frac{{C_{max}^p}^2 + A_{max}^2}{2\epsilon}, \tag{32}$$

$$\rho_n > \frac{q_{max}^o}{\nu_n^{o,*}(\epsilon)}, \forall n \tag{33}$$

where ϵ is positive and can be chosen arbitrarily close to zero. Our algorithm performance will be bounded by:

$$\sum_{n=1}^{N} \{\theta_n \nu_n^p + \varphi_n \nu_n^o\} \ge \sum_{n=1}^{N} \{\varphi_n \nu_n^{o,*}(\epsilon) + \theta_n \nu_n^{p,*}(\epsilon)\} - \frac{B}{V} \tag{34}$$

where B is a positive constant independent of V and it's expression can be found in [18].

In addition, our algorithm also ensures that the time-average sum of virtual queues X_n^o, X_n^p, Z_n and Y has an upper bound:

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t-1} \{\sum_{n=1}^{N} (X_n^o + X_n^p + Z_n) + Y\} \le \frac{B + V \sum\limits_{n=1}^{N} \{[\theta_n \nu_n^{p,*} + \varphi_n \nu_n^{o,*}]\}}{\sigma} \tag{35}$$

where $0 \le \sigma \le \epsilon$. The proof of Theorem 1, Theorem 2 and the definition of σ can be found in [18] and omitted here.

Remark 1 (Network stability): According to the definition of *strongly stable* as shown in (6), (30) and (35) indicate the stabilities of all queues in the network system. As a result, the network system is stabilized and the time-average constraints of delay and power are satisfied. Furthermore, (30) states that all the actual queues of open data and private data have deterministic upper bounds, and this characteristics ensures finite buffer sizes for all users' open and private data in BS.

Remark 2 (Optimal throughput performance): (34) states a lower-bound on the weighted throughput that our algorithm can achieve. Since B is a constant independent of V, our algorithm would achieve a weighted throughput arbitrarily close to $\sum_{n=1}^{N}\{\varphi_n \nu_n^{o,*}(\epsilon) + \theta_n \nu_n^{p,*}(\epsilon)\}$ for some $\epsilon \ge 0$. Furthermore, given any $\epsilon \ge 0$, we can get a better algorithm performance by choosing a larger V without improving the buffer sizes. Additionally, as is shown in (29), when ϵ tends to zero, our algorithm would achieve a weighted throughput arbitrarily close to $\sum_{n=1}^{N}\{\varphi_n \nu_n^{o,*} + \theta_n \nu_n^{p,*}\}$ with a tradeoff in queue length bounds and time-average delay constraints as shown in (31)-(33). Thus, with some certain finite buffer sizes, our algorithm can achieve an arbitrarily-close-to-optimal performance by choosing a large V, and V's influence on queue length is shifted from actual queues to virtual queues.

VI. SIMULATIONS

We simulate our algorithms on an example system consisting of one BS, eight users and 64 subcarriers. All weights of open data are set to be 0.8 and private data are 1. The main algorithm parameters are set as: $P_{avg} = 0.8W$, $P_{max} = 1W$, $\rho_n = 60, \forall n$, $q_{max}^o = q_{max}^p = 10000$, $\mu_{max} = 8$, $A_{max} = 5$ and $\lambda_n^o = \lambda_n^p = n * 0.1, \text{for } n \in \{1, 2, \cdots, 8\}$. And the processes of arrival data are simulated as Bernoulli process.

Fig. 2(a) and Fig. 2(b) show the dynamics of actual queues Q^o, Q^p and virtual data queues X^o, X^p of user 8 with $V = 50$, respectively. As to the other users with different arrival rate, their queues' evolutions are similar to user 8. From the figures, we see that all queues are bounded and as a result, the network system is stablized and the time-average constraints of delay and power are satisfied. Fig. 2(c) and Fig. 2(d) show eight users time-average admitted and service rate of open data and private data, respectively. Notice that, all users admitted rate is smaller than service rate and this promises the stabilities of actual data queues as well as the stability of the network.

Table II illustrates the queuing delays of all users. It can be found that all users can meet the time-average queuing delay

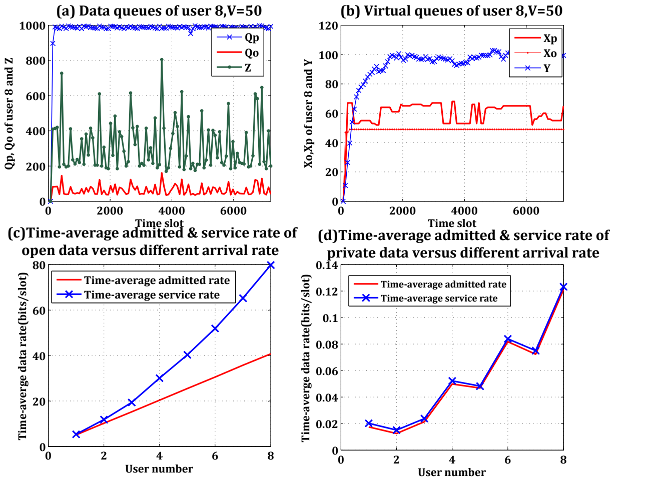

Fig. 2. Queue evolutions and all users' time-average admission/service rate

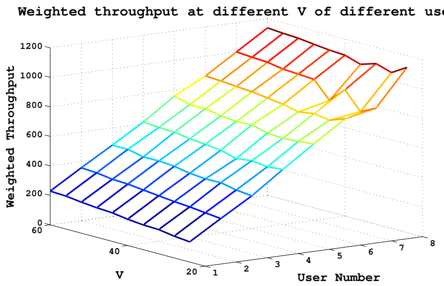

Fig. 3. User $1 \rightarrow 8$'s throughputs at different V

constraints. The reason that the user who has the bigger arrival rate has a smaller queuing delay is due to that all of them have the same deterministic queue upper bound. Thus, the user with larger arrival rate will be allocated more resource resulting a larger service rate and a less queuing delay.

Fig. 3 demonstrates the relationship between the weighted admission rate of each user versus V. In general, the bigger parameter V results in the higher weighted throughput under the same external data arrival rate. This can be proven by *Theorem 2*. Besides, for users in the same system, the one who has higher external arrival rates has a higher throughput.

VII. Conclusion

In this paper, we proposed a cross-layer scheduling algorithm for an OFDMA network in which both open data with delay constraints and private data can be transmitted at the same time. Only when the transmission urgency of private data exceeds open data, the system would allocate some resource to transmit private data. Otherwise, the system would use the entire resource to transmit open data due to delay constraint.

We show the tradeoff between the delay bound and the weighted throughput. While with some certain delay constraints, the weighted throughput can be close to optimality with finite buffer size since the influence on actual queue length is shifted from actual queues to virtual queues. Furthermore, with the help of virtual queues to track power consumption and delay, our algorithms keep the time-average power consumption and delay meet the constraints.

TABLE II

Queuing delay of different user, $\rho_n = 60, \forall n$

User number	**1**	**2**	**3**	**4**	**5**
Queuing delay	59.98	35.83	26.64	21.261	18.3624
User number	**6**	**7**	**8**		
Queuing delay	16.467	14.7776	13.76		

References

[1] E. Lawrey, "Multiuser ofdm," in *Signal Processing and Its Applications, 1999. ISSPA'99. Proceedings of the Fifth International Symposium on*, vol. 2. IEEE, 1999, pp. 761–764.

[2] J. Jang and K. Lee, "Transmit power adaptation for multiuser ofdm systems," *Selected Areas in Communications, IEEE Journal on*, vol. 21, no. 2, pp. 171–178, 2003.

[3] K. Seong, M. Mohseni, and J. Cioffi, "Optimal resource allocation for ofdma downlink systems," in *Information Theory, 2006 IEEE International Symposium on*. IEEE, 2006, pp. 1394–1398.

[4] X. Zhu, J. Yue, B. Yang, and X. Guan, "Flow rate control and resource allocation policy with security requirements in ofdma networks," in *Intelligent Control and Automation (WCICA), 2012 10th World Congress on*. IEEE, 2012, pp. 1020–1025.

[5] Z. Shen, J. Andrews, and B. Evans, "Adaptive resource allocation in multiuser ofdm systems with proportional rate constraints," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 6, pp. 2726–2737, 2005.

[6] L. Georgiadis, J. Michael, and R. Tassiulas, "Resource allocation and cross-layer control in wireless networks," in *Foundations and Trends in Networking*, 2006.

[7] C. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[8] L. Ozarow and A. Wyner, "Wire-tap channel ii," in *Advances in Cryptology*. Springer, 1985, pp. 33–50.

[9] X. Zhou and M. McKay, "Secure transmission with artificial noise over fading channels: achievable rate and optimal power allocation," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 8, pp. 3831–3842, 2010.

[10] D. Ng and R. Schober, "Resource allocation for secure ofdma communication systems," in *Communications Theory Workshop (AusCTW), 2011 Australian*. IEEE, 2011, pp. 13–18.

[11] C. Koksal, O. Ercetin, and Y. Sarikaya, "Control of wireless networks with secrecy," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*. IEEE, 2010, pp. 47–51.

[12] X. Wang, M. Tao, J. Mo, and Y. Xu, "Power and subcarrier allocation for physical-layer security in ofdma networks," in *Communications (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.

[13] Y. Cui, V. Lau, R. Wang, H. Huang, and S. Zhang, "A survey on delay-aware resource control for wireless systemsłlarge deviation theory, stochastic lyapunov drift, and distributed stochastic learning," *Information Theory, IEEE Transactions on*, vol. 58, no. 3, pp. 1677–1701, 2012.

[14] D. Xue and E. Ekici, "Delay-guaranteed cross-layer scheduling in multihop wireless networks," *arXiv preprint arXiv:1009.4954*, 2010.

[15] R. Urgaonkar and M. Neely, "Delay-limited cooperative communication with reliability constraints in wireless networks," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 2561–2565.

[16] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[17] M. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[18] X. Zhu, B. Yang, and X. Guan, "Cross-layer scheduling with secrecy demands in delay-aware ofdma network." Tech. Rep. [Online]. Available: http://wicnc.sjtu.edu.cn/techreport/zhu2012.pdf

[19] A. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm," *Queueing Systems*, vol. 50, no. 4, pp. 401–457.