An Efficient Packet Scheduling Algorithm for IEEE 802.16 BWA System

Xiuyuan Li[1, 2]
1: School of Information Science and Engineering
Shandong University
Jinan 250100, China
lixiuyuansdu@yahoo.cn

Xiaojuan Wu[1], Wenming Li[1], Yuanyuan Zhang[1]
2: School of Computer Science and Technology
Shandong Economy Institute
Jinan 250014, China
{xiaojwu, wenmingli007, zhangyy}@sdu.edu.cn

***Abstract*—Channel quality has a great influence on Bit Error Rate (BER) of received signal. The scheme to efficiently ensure four kinds of data services provided by IEEE 802.16 standards is an interesting research issue. In this paper, we propose an adaptive scheduling algorithm cross physical (PHY) layer and MAC layer for IEEE 802.16 BWA system. The Signal to Noise Ratio (SNR) on wireless fading channels determines the modulation and coding (AMC) scheme and the quantity of transmitted packets at the PHY layer. The objective of the AMC coding scheme is maximize packets throughput while maintaining the BER requirement. At the MAC layer, the algorithm scheduling the packets to ensure diverse QoS requirements for multiple connections, e.g., delay for rtPS, throughput for nrtPS, and fairness for BE. To obtain optimal scheduling, we define the cost function for each connection based on channel condition, service status, throughput or deadline. We also verify the analytic results using computer simulation. The simulation results show that our proposed scheduler can provide diverse QoS guarantees, maximum throughput and achieve an optimal tradeoff between throughput and fairness.**

Keywords-cross-layer; scheduling algorithm; BWA; QoS; IEEE 802.16

I. INTRODUCTION

One of the main advantages of IEEE 802.16[1] is its accommodating a wide range of data transmission, with different traffic characteristics and quality of service (QoS) requirements. However, the standard does not mention the realized algorithm.

The advised proposals on how to deliver QoS in BWA systems are various. Among these mechanisms, scheduling plays an important role because it selects the data for a particular frame/bandwidth allocation. Traditional scheduling algorithms for wireless networks only consider the characteristics at the MAC layer or PHY layer. Recently, there are some researches based on cross-layer. Qingwen Liu[2] and Chong Tian[3] have proposed the similar scheduling algorithm. They adopt the adaptive modulation and coding (AMC) scheme at the PHY layer and the diversity QoS control scheme at the MAC layer. Unfortunately, their algorithms have some disadvantages, e.g., the average transmission rate [2, (8)] and [3, (5)] shall be adjusted by a coefficient, which will be more reasonable. Both of the quantification function [2, (10)] and the fairness indicator [3, (8)] do not connect the packets generate rate with the packets transmission rate, which is not feasible. Besides that, they do not consider the situation that several SSs have four types of data service and the selection of modulation mode is more complicated.

In this paper, we apply a cross-layer design approach to design a scheduling algorithm that considers multiple SSs, each has diversity QoS requirements. The remainder is organized as follows: In section II, we describe the system model and the algorithm including AMC scheme at the PHY layer and the hierarchic scheduling scheme at the MAC layer. In section III, simulation results of this proposed algorithm is displayed in detail. And Section IV concludes the paper.

II. SYSTEM MODEL AND ALGORITHM

A. System Model

The IEEE 802.16 standards prescribe the PHY layer and MAC layer profile in detail [1]. According to the standard, the MAC supports a primary point-to-multipoint architecture.

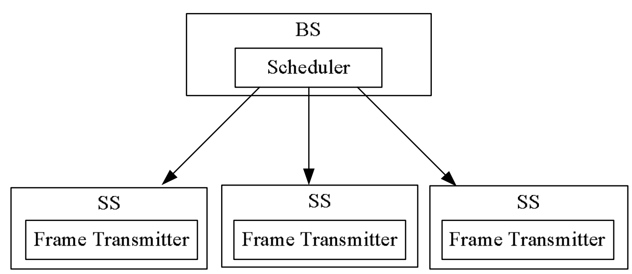

Figure 1. Architecture of BS and SSs

A example of three subscribe stations (SS) which connected to one base station (BS) or relay station over wireless channels can be showed in Fig.1. The downlink (from BS to SS) uses broadcast method and is relatively simple. So this paper will focus on the uplink direction. In uplink direction (from SS to BS), SSs utilizes time-division method to share bandwidth and timeslots. BS scheduler adaptively adjusted the allocated bandwidth and time slots to each SS according to the channel quality. Frame Transmitter in each SS will send frames to channel in the negotiated bandwidth and time slots.

978-1-4244-2108-4/08/$25.00 © 2008 IEEE

Authorized licensed use limited to: KTH THE ROYAL INSTITUTE OF TECHNOLOGY. Downloaded on March 2, 2009 at 02:37 from IEEE Xplore. Restrictions apply.

At the MAC layer, the transmission signals are packets. At the PHY layer, packets will convert into frames. Fig.2 illustrates the adopted cross-layer scheduling model between the PHY and MAC layers in each SS. As depicted in Fig.2, the scheduler is implemented at the MAC layer and operates in a hierarchic quality control scheme to accommodate a variety of services, including priority, throughput or delay. Based on channel detector, the modulation-coding controller updates the transmission mode and determines the modulation-coding mode. The MAC supports multiple PHY specifications. We will focus on the time division multiplexing (TDM) in the uplink here, although our results can be extended to the downlink as well.

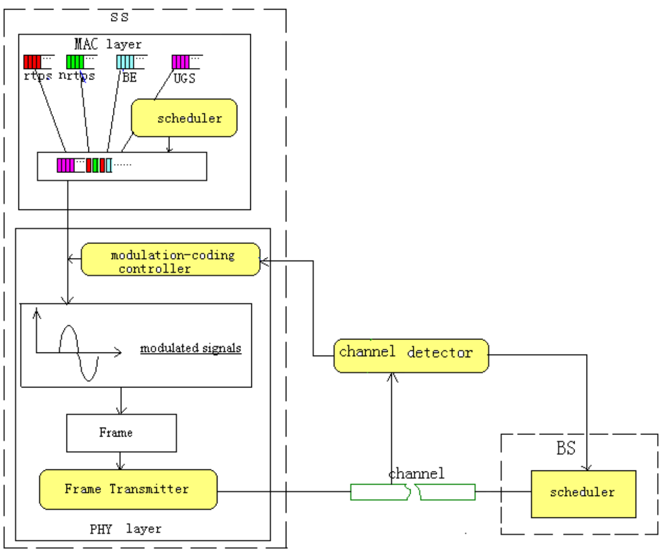

Figure 2. Cross-Layer Scheduling System Model

B. *PHY Layer Model*

In wireless networks, each user's packet data are ultimately transmitted through the time varying fading channel. Channel Quality determines Bit Error Rate (BER) and ultimately affects QoS. Specially, when the wireless channel experiences deep fades, the prescribed BER may not be satisfied, and the allocated bandwidth is wasted. So we adopt AMC schemes that adjust transmission parameters to the channel variation adaptively.

As in [2] and [3], let N denote the total number of transmission mode available. We partition the entire signal-to-noise ratio (SNR) range into N+1 non-overlapping consecutive interval, with the boundary points denoted as $\{\gamma_n\}_{n=0}^{N+1}$. When $\gamma \in [\gamma_n, \gamma_{n+1})$, mode n is chosen. Note that no data will be sent when $\gamma \in [\gamma_0, \gamma_1)$ in order to avoid deep channel fades.

We first adopt the Nakagami-m distribution model [4] to simulate the actual channel, because the Nakagami-m distribution has been found to be a very good fitting for the wireless channel [5][[6][7]. The Nakagami-m probability density function p(γ) of the envelope is given by:

$$p_\Omega(r) = \frac{2m^m \cdot r^{2m-1}}{\Omega^m \cdot \Gamma(m)} \exp(-\frac{mr^2}{\Omega}) \tag{1}$$

Where r is the amplitude of signal envelope; m ($m \geq 1/2$) is Nakagami parameter and $\Omega = 2E[r^2/2]$ which are determined by the hardware of wireless equipments and propagation conditions of radio waves; and $\Gamma(m) = \int_0^\infty t^{m-1} e^{-t} dt$ is the Gamma function. If we assume the average noise lever is stable in the channel during the whole transmission for simplicity, equation (1) can be written as:

$$p_r(\gamma) = \frac{m^m \cdot \gamma^{m-1}}{\bar{\gamma}^m \cdot \Gamma(m)} \exp(-\frac{m\gamma}{\bar{\gamma}}) \tag{2}$$

Where: γ is the instantaneous SNR; $\bar{\gamma} = E\{\gamma\}$ is the average channel SNR.

To find an appropriate γ_n, we consider five kinds of modulation-coding mode. They are BPSK, QPSK (1/2), QPSK (3/4), 16-QAM and 64-QAM [1]. The signal error rate of modulation-coding modes is:

BPSK [8]: $$P_E = \frac{1}{2} erfc(\sqrt{\gamma}) \tag{3}$$

QPSK [8]: $$P_E = 1 - [1 - \frac{1}{2} erfc(\sqrt{\frac{\gamma}{2}})]^2 \tag{4}$$

QAM [9]: $$P_E = 2[1 - \frac{1}{\sqrt{M}}] erfc\left(\sqrt{\frac{3}{2(M-1)} k\bar{\gamma}}\right) \times \left[1 - \frac{1}{2}(1 - \frac{1}{\sqrt{M}}) erfc(\sqrt{\frac{3}{2(M-1)} k\bar{\gamma}})\right] \tag{5}$$

Where M is the modulation order, k is the binary digit number of signal and *erfc* is the error function complement.

We assume:

At the MAC layer, each service queue has finite capacity of packets. Each packet includes fixed number of signals, which is 1024 bits.

Given a fixed symbol rate, the frame duration (T_f) is fixed at the PHY layer. One symbol can stand for different bits according to modulation mode. Therefore, we can derive the BER and packet error rate from eq. (3-5). Here, we adopt the similar table [3, table (1)].

C. *MAC Layer Model*

At the MAC layer, four scheduling services are provided by the IEEE 802.16 standards [1]. They are:

1) Unsolicited Grant Service (UGS) is designed to support real-time data streams consisting of fixed-size data packets

Authorized licensed use limited to: KTH THE ROYAL INSTITUTE OF TECHNOLOGY. Downloaded on March 2, 2009 at 02:37 from IEEE Xplore. Restrictions apply.

issued at periodic intervals. This service provides guarantees on throughput, latency and jitter.

2) Real-time Polling Service (rtPS) is designed to support real-time data streams consisting of variable-sized data packets that are issued at periodic intervals. It provides guarantees on throughput and latency, and the QoS metrics are the PER and maximum delay.

3) Non-Real-time polling service (nrtPS) is designed to support delay-tolerant data streams consisting of variable-sized data packets. So it can tolerate longer delays and is rather insensitive to delay jitter.

4) Best Effort (BE) services provides neither throughput nor delay guarantees. A prescribed PER should be guaranteed for BE connections over wireless channels although no QoS parameter is specified.

In our cross-layer scheduler, the UGS will be given higher priority than the other three QoS class and deployed fixed number of time slots in the frame. Our algorithm will only consider scheduling for rtPS, nrtPS and BE connections.

Each user is allocated a fixed number of time slots per frame. We focus on three users here. Except for UGS service, the residual time slots are scheduled for other three QoS service in the frame.

For each rtPS connection, the scheduler timestamps each arriving packet according to its arrival time and defines its maximum latency (deadline) as the QoS scheduling parameter.

For each nrtPS connection, guaranteeing the minimum reserved packet rate means the average packet transmission rate should be greater than it. In practice, we should define the average packet transmission rate estimated over a window size as the QoS scheduling parameter.

We will just take priority and fairness about BE connection into accounts.

D. MAC Scheduling Scheme

The scheduling algorithm includes the packets scheduler algorithm and the built frame controller algorithm depicted as in Fig.3.

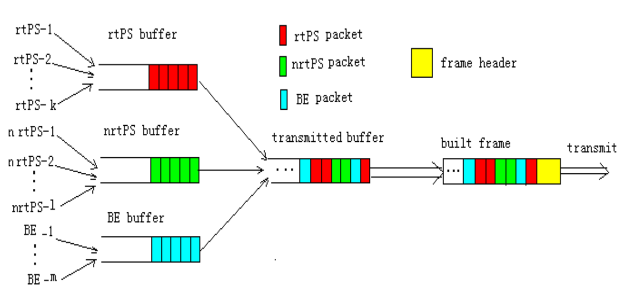

Figure 3. Packets Scheduling Scheme Graph

In Fig. 3, rtPS packets are fetched into rtPS buffer according to the maximum delay; nrtPS packets are fetched into nrtPS buffer according to the minimum prescribed packets throughput; BE packets are fetched into BE buffer according to FIFO. Firstly, rtPS packets and nrtPS packets are fetched into transmitted buffer according to their QoS guarantee function. Secondly, all the diverse kinds of packets according to different transmitted parameters and fairness indicator.

E. MAC Packet Scheduler Algorithm

Step 1: Define $RT_i\ (t+1)$ as the QoS guarantee function of packet *i* which represents the QoS guarantee level at time *(t+1)* if the rtPS packet *i* will not be transmitted at time *t*.

$$RT_i(t+1) = L_i - \hat{W}_i(t+1) + 1 \tag{6}$$

Where L_i is the deadline of the packet *i* and $\hat{W}_i(t+1)$ is the estimation of the packet waiting time. To avoid $RT_i\ (t+1)$ to become a negative number, we add one in equation (6). $\hat{W}_i(t+1)$ can be given as:

$$\begin{cases} W_i(0)=0 \\ \hat{W}_i(t+1)=W_i(t)+T_f \end{cases} \tag{7}$$

T_f (T_f can be 0.5 or 1ms) denotes the frame duration time, and $W_i(t)$ is the practical waiting time of packet *i*. If $RT_i\ (t+1)$ <=1, the packet *i* should be sent into the transmitted buffer immediately, except that the transmitted buffer is full, in order to avoid QoS dissatisfaction at time *(t+1)*.

If multiple rtPS packets have their $RT_i\ (t+1)$ <=1 at the same time, all these rtPS packets will be stochastically derived into the transmitted buffer.

If $RT_i\ (t+1)$> 1, the QoS requirement is satisfied even if packet *i* may not be transmitted at time t.

Step2: Define $NRT_j\ (t+1)$ represents the QoS guarantee level at time *(t+1)* if the nrtPS packets will not be transmitted at time *t*, which is the QoS guarantee function of the packets belong to the nrtPS connection *j*.

$$NRT_j(t+1) = \frac{\hat{N}_j(t+1)}{N_T} + \varepsilon \tag{8}$$

Where $\hat{N}_j(t+1)$ is the estimation of the packet throughput at time *(t+1)*, which is given by the number of packets per second. N_T is the average prescribed packet throughput. We add ε ($\varepsilon \in (0,1)$) in (8) to allow floating in a small scope. And $\hat{N}_j(t+1)$ can be given as:

$$\hat{N}_j(t+1) = N_j(t) \bullet (1 - T_f) \tag{9}$$

Where $N_j(t)$ is the packet throughput at time *t*. We assume T_f=0.001s. Supposed the packet transmission rate is $N_j(t)$ during (1-T_f) and there are no plackets transmitted during *(t+1)*. $N_j(t)$ can be given as:

$$N_j(t) = \begin{cases} 0 & t = 0 \\ N_j(t-1) \bullet (1 - T_f) + CN_j(t) \bullet T_f & t \neq 0 \end{cases} \tag{10}$$

Authorized licensed use limited to: KTH THE ROYAL INSTITUTE OF TECHNOLOGY. Downloaded on March 2, 2009 at 02:37 from IEEE Xplore. Restrictions apply.

Where $CN_j(t)$ is the number of packets belonging to the connection j at time t. Equation (10) has the similar assumption as eq. (9): during (1- T_f), the packet transmission rate is $N_j(t-1)$,and during T_f, there are $CN_j(t)$ plackets transmitted. So the whole transmitted packets number is $N_j(t)$.

NRT_j $(t+1)$ has the same scheduling effects on nrtPS packets as RT_i $(t+1)$ on rtPS packets.

Step3: Let μ_{rtPS} , μ_{nrtPS} and μ_{BE} are the transmitted parameters ($\mu_{rtPS}+\mu_{nrtPS}+\mu_{BE}=1$). The transmission buffer fetches the packets from the rtPS, nrtPS and BE connections according the parameter μ_{rtPS} , μ_{nrtPS} and μ_{BE} .

The role of μ_{rtPS} , μ_{nrtPS} and μ_{BE} is to provide different priorities for different QoS class. The coefficients can be set under the constraint $\mu_{rtPS}>\mu_{nrtPS}>\mu_{BE}$. This priority scheme can provide comparable priorities among connections with different kinds of services, which enable exploiting multi-user diversity by similarly conduct.

In the same service class of rtPS or nrtPS, we denote $f_i(t)$ as the fairness indicator belong to connection i. $f_i(t)$ can be defined as:

$$f_i(t)=\frac{N_{iG}(t)-N_{iT}(t)}{N_{iG}(t)} \tag{11}$$

Where $N_{iG}(t)$ is the number of generated packets, and $N_{iT}(t)$ is the number of transmitted packets from the beginning to time t. This fairness indicator is related with the burden of each service queue, which can avoid the situation that some connections are given higher priority when they have few number of packets generated to be transmitted. The buffer in each kind of services will range the packets with descending order according to the fairness indicator.

III. SIMULATION RESULTS

In our simulation, we consider three users that each transmits one rtPS connections, one nrtPS connections and one BE connections in the time varying fading channel. The channel, QoS and traffic parameters are respectively:

The frame duration T_f=0.001s. The packet length at MAC layer is fixed as 1024bits. The average SNR is $\bar{\gamma}$ =15dB. The Nakagami parameter is m=2.

User A. rtPS-1: the deadline L_1=10ms; the throughput T_1=2Mbps; the generate probability λ_1 =0.2; nrtPS-1: the throughput T_1=4Mbps; the generate probability λ_1=0.3; BE-1: the generate probability λ_1 =1.

User B. rtPS-2: the deadline L_2=20ms; the throughput T_2=1Mbps; the generate probability λ_2 =0.2; nrtPS-2: the throughput T_2=3Mbps; the generate probability λ_2 =0.3; BE-2: the generate probability λ_2 =1.

User C. rtPS-3: the deadline L_3=30ms; the throughput T_3=2Mbps; the generate probability λ_3 =0.2; nrtPS-3: the throughput T_3=3Mbps; the generate probability λ_3 =0.3; BE-3: the generate probability λ_3 =1.

We adopt Nakagami-m channel mode to describe channel SNR distribution, which is remain invariant during a frame. Fig. 4 depicts the variability of the SNR in the frequency-flat fading channel.

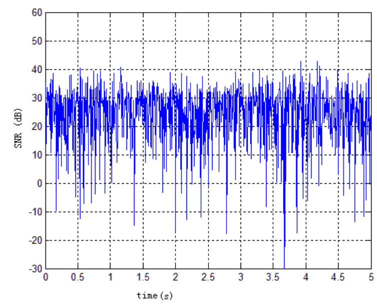

Fig. 4. SNR of Time-Fading Channel

The rtPS service requirement is evaluated by the delay. Fig.5 depicts the rtPS-User A, rtPS-User B and rtPS-User C transmitted delay. All the delay of three users are almost less than 10 ms during the most simulation time, no matter what prescribed deadline for different users. So the rtPS service can ensure perfect QoS requirements by our algorithm.

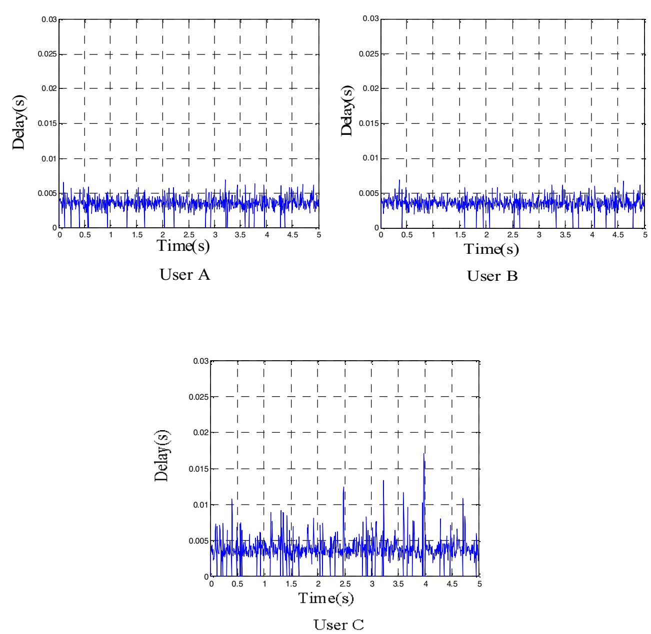

Fig. 5. Delay of rtPS

Authorized licensed use limited to: KTH THE ROYAL INSTITUTE OF TECHNOLOGY. Downloaded on March 2, 2009 at 02:37 from IEEE Xplore. Restrictions apply.

To study the network throughput which is scheduled according to our algorithm, we do not display the throughput graph of rtPS, nrtPS and BE in details. Fig. 6,7 shows the throughput of nrtPS and BE. We can see that the packet throughput varies around the prescribed throughput, although the throughput of BE fluctuate more fiercely than the throughput of nrtPS. It shows BE connection has the lowest priority, which QoS may not ensure during some transmission time.

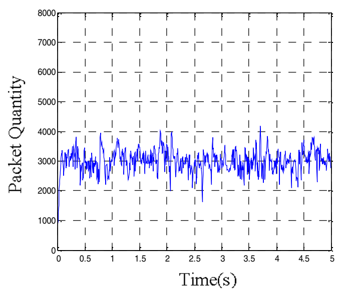

Fig. 6. Throughput of nrtPS

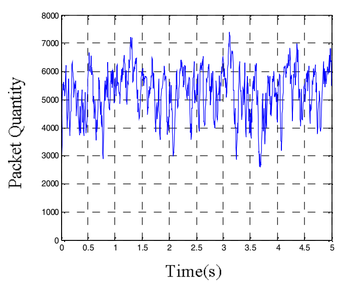

Fig.7. Throughput of BE

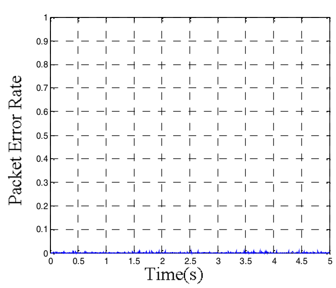

Fig. 8. Packets Error Rate Per Frame

Fig. 8 descript packets error rate per frame. During the whole simulation time, the number of error packets is no more than one packet.

IV. CONCLUSIONS

In this paper, a cross-layer model for QoS support in wireless networks is developed. The focus is on adaptive modulation and coding at the PHY layer and diverse QoS requirements for multiple connections at the MAC layer are ensured. The key advantages of the proposed model are illustrated by some simulation results, which are displayed the characters of generality, simplicity and practicality.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose comments helped to improve this manuscript. This work has been given significantly helpful advices by Dr. Maode Ma (Nanyang Technological University) and Chong Tian (Shandong University), we appreciate them.

REFERENCES

[1] IEEE Standard 802.16 Working Group, "IEEE standard for local and metropolitan area networks part 16: air interface for fixed broadband wireless access systems (revision of IEEE standard 802.16-2001) ," 2004

[2] Q. Liu, S. Zhou, and G. B. Giannakis, " Cross-Layer Modeling of Adaptive Wireless Links for QoS Support in Multimedia Networks," In: proc. 1st int. conf. QSHINE. Dallas, TX, 2004, pp. 65-75. Oct. 18-20

[3] Chong Tian, Dongfeng Yuan, "A Novel Cross-Layer Scheduling Algorithm for IEEE 802.16 WMAN," In: 1st int. Workshop on Cross Layer Design (IWCLD07) conf. JINAN, CHINA, Sep.2007, pp.20-21

[4] M. Nakagami, "the m-Distribution-A General Formula of Intensity Distribution of Rapid Fading," In: Statistical Methods in Radio Wave Propagation, W. C. Hoffman, Ed. Elmsford, NY: Pergamon, 1960

[5] Michel Daoud Yacoub, Jose Edson Vargas Bautista, and Leonardo Guerra de Rezende Guedes, " On Higher Order Statistics of the Nakagami-m Distribution," In: IEEE Transactions on Vehicular Technology, May. 1999, Vol. 48, no. 3, pp. 790-794

[6] H. Suzuki, "A Statistical Model for Urban Radio Propagation," In: IEEE Trans. Commun. , 1975, vol. COM-25, no. 7, pp. 673-679

[7] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-Layer Combining of Adaptive Modulation and Coding With Truncated ARQ Over Wireless Links," In: IEEE Trans. Wireless Communication, Sep. 2004, vol. 5, pp. 1746-1755

[8] Changxin Fan, "Communication Principle," Defense Industry Press of China, May, 2001, pp.167-185

[9] Eric Newman, "Analysis and Optimization of Receiver Performance with Error Vector", http://www.eetchian.com, Oct. 2005

Authorized licensed use limited to: KTH THE ROYAL INSTITUTE OF TECHNOLOGY. Downloaded on March 2, 2009 at 02:37 from IEEE Xplore. Restrictions apply.