QoS_GTE: A Centralized QoS Guaranteed Throughput Enhancement Scheduling Scheme for Relay-Assisted WiMAX Networks

Chung-Ju Chang[*], Chih-Ming Yen[*], Fang-Ching Ren[†], and Chia-Hsuan Chuang[*]

[*]Department of Communication Engineering

National Chiao Tung University

Hsinchu, Taiwan 300

Corresponding author: cjchang@cc.nctu.edu.tw

[†]Computer and Communication Research Laboratories

Industrial Technology Research Institute

Hsinchu, Taiwan 300

Abstract − **A centralized QoS guaranteed throughput enhancement (QoS_GTE) scheduling scheme for downlink relay-assisted WiMAX networks is proposed. The QoS_GTE consists of a transmission time based path selection algorithm (TT_PSA), a service order based resource allocation algorithm (SO_RAA), and a transmission concurrency decision algorithm (TCDA). The TT_PSA selects the path with the minimal transmission time for packets, which takes path loss, shadow fading, and interference into consideration. The SO_RAA gives high priority to urgent users and maximizes throughput under QoS satisfaction. The TCDA carries out flexible resource reuse by deciding which relay station (RS) can transmit concurrently using the same frequency and time slots. Simulation results show that the proposed QoS_GTE outperforms the scheme of without relay by an amount of 63% and that of the LMP by an amount of 28%, in the system throughput. The QoS_GTE can also guarantee QoS requirements of real time services by 97%.**

I. INTRODUCTION

In cellular networks, there has been a concern about the limited bandwidth and the weak SINR at edges of cells due to buildings, distance, interference, and so on. In order to extend the cell coverage and maintain the SINR with low upfront investment, WiMAX networks are designed with the assistant of relay technique, where the traffic signal can be relayed through one or more relay station (RS) as intermediate node(s). It is recognized that this relay structure combines advantages of both cellular and ad-hoc networks. The profit of using a RS is the lower complexity, the less power compared with a BS, and the absence of need for wired backhaul connection [1]. The most important cause of using the relaying technique comes from the reduction in the overall path loss between a BS and a MS. Another is the path diversity gain which can be achieved by selecting the most favorable path in shadowed environments.

Choosing an optimal RS can have effective impacts on the overall performance improvement. In [2],[3], effects of two-hop RS selection strategies on system performance had been studied. The authors proposed several different schemes for RS selection: selection based on distance, selection based on path loss, and then random selection. These schemes can improve the SINR, and the order from best to worst is: selection based on path loss, random selection, and selection based on physical distance. From the path loss based schemes, the least maximum path loss relaying node selection (LMP) can choose good channels with the percentage of 98%. In [4], the distance based selection was adopted under the assumption that the path loss exponent from the BS to the MS is larger than the path loss exponent from the BS to the RS. If the distance from the BS to the MS is greater than a certain distance, the BS will use the relay path for its transmission. However, if the path loss exponent from RS to MS is large enough, the total performance of relay path may not be good as it is expected.

Substantial throughput improvement can be additionally obtained by operating concurrent relaying transmission. The resource reuse is very crucial in the realization of the full potential of relay networks [5]-[6]. The previous schemes maximize the total transmission rate by finding a set of active links. However, it needs a complicated computation to get the set of active links because the mobility of relay node in mesh or ad-hoc networks. Also, there is a tradeoff between system throughput and the quality of service (QoS) in cellular networks. Challa and Cam [7] simulated the relay performance with only one traffic type, hence the quality of service (QoS) for different traffic type is ignored. However, Cho and Hass [5] had shown that in the multi-hop cellular networks, BS can choose to utilize the multi-hop relaying instead of the single hop direct transmission. Such a hybrid operation can be exploited to mitigate the unfairness in QoS among users. Therefore, QoS as well as system throughput in cellular network can be improved through the use of the relaying.

In the paper, we propose a downlink centralized QoS guaranteed and throughput enhancement scheduling (QoS_GTE scheduling) scheme in WiMAX relay-assisted network. It is proven that the two-hop path is most efficient and easiest to implement in multi-hop relaying [5][8]. Thus the 2-hop relay paths are considered in the paper. The QoS_GTE scheduling scheme includes three algorithms. First, a transmission time based path selection algorithm (TT_PSA) finds a best path for each MS based on the minimum transmission time. The TT_PSA makes a deliberate judgment in path selection by the transmission time, which is determined by the pathloss, the shadow fading and the interference. For QoS guarantee, a service order parameter is defined, which takes both priority and urgency into account as the foundation of the packets' service order. Based on service order parameter, a service order based resource allocation algorithm (SO_RAA) manages the resource allocation, which not only guarantees the QoS requirements but also restrains the decline of the system throughput as possible as it can when QoS guarantee is done. The SO_RAA gives priority to MSs whose traffic constraint is going to be violated for the purpose of QoS guarantee. Later, MSs with best channel condition will get the resource for the purpose of the raise in the system throughput. Finally, a transmission concurrency decision algorithm (TCDA) compresses the used-resource by means of allowing several RSs transmit at the same time so as to make full utilization of radio resources. The TCDA uses the RSs location concept to determine which RSs can transmit at the same time without difficult computation.

Simulation results show that the proposed QoS_GTE outperforms the without relay scheme and the LMP in terms of system. The QoS_GTE also guarantees the QoS requirements of real time services. The SO_RAA guarantees the QoS requirements and raises the throughput compared with that of the P_RAA at the same time, which balances the tradeoff between the QoS guarantee and system throughput.

II. SYSTEM MODEL

Consider a WiMAX relay_assisted cellular network in the urban area, where nineteen cells are wrapped around to imitate the condition with interference from cells. The frequency reuse

factor in the wrapping model is four. As shown in Fig. 1, there are one base station (BS) at the center of each cell and several wireless fixed RSs surrounding the BS uniformly. All RS positions are also on the crossroads. A regular Manhattan type of grid in each cell is considered, where solid line represents the street and the intersections are the crossroads. Mobile stations (MSs) are uniformly distributed on the street and the snap-shot method is used to monitor the location of MSs. Assume that all the proposed algorithms are done in the BS; and RSs merely forward the information received from the BS to the MSs. Each RS can communicate only with BS and MSs in the same cell. The communication between any two RSs in the same cell is not allowed. Moreover, transmitting and receiving actions do not occur at the same time in both BS and RSs. Since the sub-carriers in a sub-channel is assumed to be nonadjacent, the channel condition of each sub-channel can be regarded the same, and the frequency selective condition does not happen. Also, we choose equal power for all sub-channels. It is because some literatures [9], [10] have proved that the resource allocation algorithm with equal power distribution over all sub-carrier can perform as well as the optimal power and sub-channel allocation.
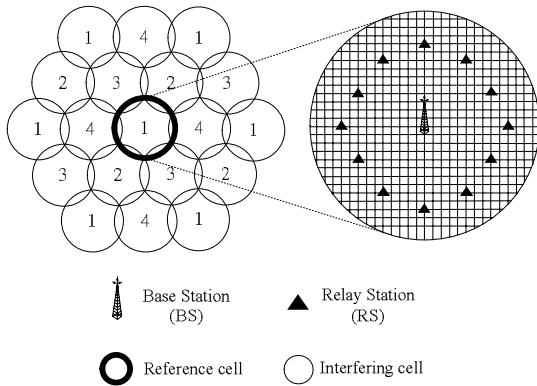

Figure 1. The system model

A two-hop relay-assisted WiMAX network is considered. In the network, the downlink transmission may involve a direct path or a relay path. The transmission of a direct path is that a BS transmits directly to a MS without any intermediate node. The transmission of a relay path is that a BS transmits information to a RS and the RS forwards the received information to the MS. BS and RSs share the same spectrum through a centralized scheduling for channel access in the TDMA manner. At the start of a downlink frame, BS transmits first, and then RSs transmit for the rest of the frame time. The transmission time for BS and RSs is variable in order to achieve good resource usage.

Assume each BS has complete transmission information at the start of each time frame. The information consists of (i) the transmission rates on all links during the frame that will be scheduled, (ii) the channel state information which includes path loss and shadow fading, and (iii) QoS requirements of each MS. There are four service types of traffic from high priority to low priority, which are voice, video, HTTP and FTP in order. Each traffic type is given an integer priority value $p(x)$. The $p(x)$ from high to low is voice, video, HTTP, and FTP in sequence. Assume that each MS can request only one type of traffic service at each time. The transmission is supposed to be with equal power, and the channel state information is assumed constant in a frame but various frame by frame. The channel is

affected by thermal noise, slow fading which includes path loss, and shadow fading. The modulation order M_j for each transmission path can be decided [12] as follows:

$$m = K \times SINR + 1, \text{ with } K = \frac{-1.5}{\ln(5 \times \varepsilon)}, \quad (1)$$

where ε is the bit error rate requirement. For a given value m, we can find its corresponding M_j from Table 1 thus available data rate is obtained as:

$$R = M_j. \quad (2)$$

Table 1. Modulation Order Mapping

m	$0 \leq m < 4$	$4 \leq m < 16$	$16 \leq m < 64$	$64 \leq m$
M_j	0	4	16	64

III. QOS_GTE SCHEDULING SCHEME

The proposed QoS_GTE scheduling scheme contains three parts: the transmission time based path selection algorithm (TT_PSA), the service order based resource allocation algorithm (SO_RAA), and the transmission concurrency decision algorithm (TCDA). First, the TT_PSA takes channel state information of all MSs to selects the path with the minimal transmission time for each MS. Under the given path, the SO_RAA allocates the resource to the MSs according to either the QoS requirements or the channel condition. Eventually, the TCDA checks if RSs' transmissions are able to be combined together to compress the used-resource. If there is any transmission concurrency occured, the occupied resource will be released. This released resource can be assigned to other MSs by the SO_RAA again.

A. Transmission Time Based Path Selection Algorithm (TT_PSA)

The TT_PSA is based on the total transmission time that a path experiences, which takes the distance, the path loss, the shadow fading, and the interference all into consideration. To make the system throughput high, it is necessary to transmit the information with a minimal transmission time as much as possible. This TT_PSA just intends to find a path with the minimum transmission time for each MS.

The TT_PSA first computes the transmission time for each possible path, including direct path and relay paths. Denote $R'_{m,i}$ the allocated data rate for MS_m from BS to RS_i, and $R''_{m,i}$ is the allocated data rate for MS_m from RS_i to MS_m. Then the total transmission time from BS through RS_i to MS_m, denoted by, $t_{m,i}$, is obtained by

$$t_{m,i} = \frac{1}{R'_{m,i}} + \frac{1}{R''_{m,i}}, \ i \neq 0. \quad (3)$$

Similarly, denote $R_{m,0}$ the allocated data rate through direct path, the transmission from BS to MS_m directly. The transmission time from BS to MS_m $t_{m,o}$ can be given by

$$t_{m,0} = \frac{1}{R_{m,0}}. \quad (4)$$

Finally, the TT_PSA selects the path with minimal transmission time as the transmission path for MS_m.

B. Service Order Based Resource Allocation Algorithm (SO_RAA)

In order to maintain the QoS requirement, we define a service order parameter which combines the priority and urgency together. The service order parameter of MS_m at the beginning of the xth-frame, denoted by $S_m(x)$, is defined as

$$S_m(x) = \frac{p(x)}{p^*} + \frac{u(x)}{u^*}, \quad (5)$$

where $p(x)$ is the priority value of the packet for MS_m at the beginning of the xth-frame, $u(x)$ is the urgency value of the packet for MS_m at the beginning of the xth-frame, p^* is the maximum value of $p(x)$ and u^* is the maximum value of $u(x)$. Usually, real-time packets such as voice or video have higher priority value than non-real-time packets. Also, in order to avoid the violation of QoS requirements of low priority packets, we advocate that the low priority traffic deserves being served with high service order when it is in an urgent condition. We given the low priority packet high urgency value. We take both the priority value and the urgency value together to decide which packet should be served immediately.

In detail the urgency value $u(x)$ is brought out and is defined as

$$u(x) = l(x) + t(x),\qquad(6)$$

where the $l(x)$ is the *urgency* value for a packet contributed by the remaining bits which are not transmitted yet at the beginning of the xth –frame, and the $t(x)$ is the urgency value for a packet contributed by the delay time at the beginning of the xth –frame. The $l(x)$ is given by

$$l(x) = \frac{L - C_{x-1}}{L},\qquad(7)$$

where L is the size of the packet and C_{x-1} is the total bits that have been transmitted at the end of the $(x-1)^{th}$–frame. The value of $t(x)$ is determined according to the delay time that the packet undergoes. D^* denotes maximum delay tolerance in the unit of frame. Suppose the number of level of $t(x)$ is U^* and we divide the D^* into U^* levels. The $t(x)$ is set to be one when the delay of the packet at the xth –frame, denoted as $d(x)$, is not greater than $D^*/2$, which implies this packet is the less urgent one. Then we divide the other $D^*/2$ range into (U^*-1) parts equally and the values of $t(x)$ are set to be two, three, …, and U^*. The packet with $t(x)=U^*$ is approaching the delay upper bound D^* and is so urgent to be served immediately. Notice that the method above only works when the delay constraint D^* is larger than U^*, the level of $t(x)$. This concept can be expressed in mathematic equations given below.
Where $D^* > U^*$

$$t(x) = \begin{cases} U^*, & \text{if } D^* - \left(\frac{1}{U^*-1} \times \frac{D^*}{2}\right) < d(x) \leq D^* \\[2mm] U^*-1, & \text{if } D^* - \left(\frac{2}{U^*-1} \times \frac{D^*}{2}\right) < d(x) \leq D^* - \left(\frac{1}{U^*-1} \times \frac{D^*}{2}\right) \\[2mm] U^*-2, & \text{if } D^* - \left(\frac{3}{U^*-1} \times \frac{D^*}{2}\right) < d(x) \leq D^* - \left(\frac{2}{U^*-1} \times \frac{D^*}{2}\right) \\[2mm] \vdots \\[2mm] 2, & \text{if } D^* - \left(\frac{U^*-1}{U^*-1} \times \frac{D^*}{2}\right) < d(x) \leq D^* - \left(\frac{U^*-2}{U^*-1} \times \frac{D^*}{2}\right) \\[2mm] 1, & \text{if } d(x) \leq \frac{D^*}{2}. \end{cases} \qquad(8)$$

If we use the same mechanism to divide the urgency levels when the delay constraint D^* is not larger than U^*, it happens that the range of some level is a fraction of a frame. Since our minimum unit in scheduling is one frame, a fraction of a frame will never occur. For this reason we modify our mechanism when D^* is not larger than U^*. This concept can be expressed in mathematic modeling as below.
When $D^* \leq U^*$

$$t(x) = \begin{cases} U^*, & \text{if } d(x) = D^* \\ U^*-1, & \text{if } d(x) = D^*-1 \\ \vdots \\ U^*-D^*, & \text{if } d(x) = 0. \end{cases} \qquad(9)$$

As D^* decreasing by one frame, $t(x)$ also decreases one, $t(x)=U^*-1$, and so on until x = 0. There are total D^*+1 levels in $t(x)$ and the difference of delay constraint between any two adjacent levels is one frame.

The QoS constraint of HTTP is the minimum required transmission rate R^*. If a packet with size L is transmitted at the rate R^*, the transmission time for that packet is,

$$D^* = \frac{L}{R^*} \times N \quad \text{(frames)},\qquad(10)$$

where N is the number of frame per second. In other words, D^* is the maximum transmission time for a packet with size L. After having D^*, all the process of obtaining $t(x)$ is the same as real time traffic. Since FTP has no QoS requirements, its $t(x)$ is set to be the lowest.

For the purpose of QoS guarantee, the packets can be served first when their $S_m(x)$ are larger than service order threshold S_{th}. Also, the service order of these high $S_m(x)$ MSs is according to the $S_m(x)$ in descending way. After QoS guarantee, the best first resource allocation [11] assigns the residual resource to other packets with good channel condition to enhance the total throughput.

C. Transmission Concurrency Decision Algorithm (TCDA)

The TCDA aims to find more than one RSs which can transmit concurrently so that the other original occupied resource can be removed for other transmitters to improve the spectrum efficiency. The steps of finding concurrency RSs for the reference RS (ref. RS) are as follows. First, the TCDA chooses active RSs whose angle separated from the ref. RSs are not less than $120°$ as the ref. RS's concurrency candidates. After the candidate selection, if these RSs which are $120°$ apart from the ref. RS, the TCDA will enable these two RSs to transmit with the ref. RS concurrently with the transmission time which is maximum among those of these candidate RSs and ref. RS.

Figure 2(a) shows an OFDMA frame before the TCDA operated, in which suppose RSa, RSi, and RSj can transmit concurrently. After the TCDA operation, as Figure 2(b) shown, RSa, RSi, and RSj transmit at the same time, and it makes original-occupied resources released. Note that the transmission time T is set to be TRa, where TRa is the maximum transmission time among TRa, TRi and TRj. The released resources can be utilized again by the SO_RAA.

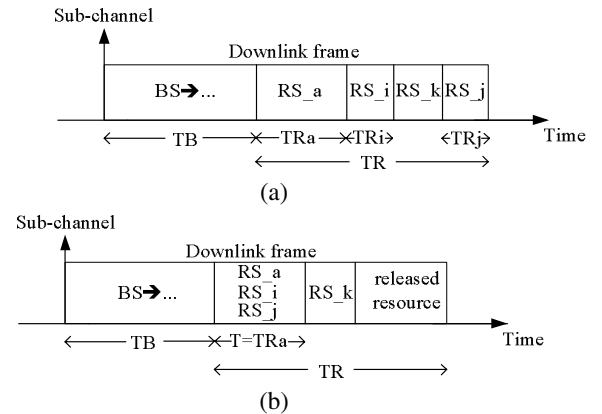Figure 2. Transmission concurrency example. (a) an OFDMA frame before concurrency (b) an OFDMA frame after concurrency

IV. SIMULATION RESULTS AND DISCUSSION

System parameters of a relay-assisted WiMAX network for simulations are listed in Table 2. Scalable parameters in the physical layer are modified according to the suggested values in [13]. Equal power is supposed for each sub-channel. Also, table 3 lists the QoS requirements of each traffic type. A

Manhattan street scenario is considered in which the width of each street is 30 meters. The separation of two adjacent parallel streets is 200 meters. Twelve RSs are uniformly located around the BS in a circle at the cell radius of three fourths. When the desired site of three-fourth cell radius is not exactly at an intersection, the RS would be placed on the nearest intersection from the desired site.

For fair comparison, the resource allocation of each relay scheme is SO_RAA. The level of $t(x)$, denoted as U^*, is equal to 6. The threshold value S_{th} is set to be the average of maximum and minimum S_m of HTTP traffic (S_{th}=1.0714). Although HTTP is the third priority traffic, HTTP should also own a service right to guarantee the QoS as possibly as it can when HTTP is facing some kind of time emergency. This is why we set S_{th} as the average of maximum and minimum S_m of HTTP traffic. In the simulations, the number of user is composed of four service traffic types with equal percentage. We define the traffic load as the ratio of the total average rate of all users over the system maximum transmission rate. The maximum transmission rate is achieved when Q users are multiplexed for each sub-channel and the highest modulation order is used for all users. It is equal to 11.74 Mbps in the simulation environment of this paper. Note that the average arrival rate of voice, video, HTTP, and FTP users is equal to 5.2 Kbps, 64 Kbps, 14.5 Kbps, and 88.9 Kbps, respectively.

We compare the TT_PSA with without relay (w/o Relay) and the least maximum path loss (LMP) relaying node selection scheme [2][3], which is based on a route that has the lowest bottleneck (in terms of path loss). SO_RAA is compared with priority based resource allocation algorithm (P_RAA), which allocates the resource according to the packets' priority value.

Table 2. System Parameters

Parameters	Value	Parameters	Value
BS cell size	1500 m	Guard time	11.2 μs
Number of data subcarriers	384	RS cell size	500 m
Number of antenna at BS	1	Number of subchannels	24
Number of OFDMA symbol for downlink transmission per frame	24	Number of data subcarriers per subchannel	16
System bandwidth	5 MHz	Frame duration	5 ms
Carrier frequency (f_c)	3.5 GHz	FFT size	512
Thermal noise density (N_0)	-174 dBm/Hz	Power of BS	25 W
Sampling frequency spacing	11.16 kHz	Power of RS	5 W
OFDMA symbol duration	100.8 μs	Guard time	11.2 μs
Useful symbol time	89.6 μs		

Table 3. The QoS Requirements of Each Traffic Type

Figure 3 shows the system throughput for several schemes. First, we analyze the difference between the QoS_GTE and the TT_PSA+P_RAA+TCDA. We observe that the system throughput of the QoS_GTE is greater than that of the P_RAA by about 33%. This is because the P_RAA allocates the resource based only on priority value so that the good channel condition users may have no chance to be served, which results in throughput decrease. As for the SO_RAA, it not only put efforts on QoS guarantee but also tries to increase the throughput by using the best first resource allocation. We also find that QoS_GTE outperforms the other four schemes: QoS_GTE w/o TCDA, LMP schemes, and w/o Relay, which use the SO_RAA and/or TCDA. Because the TT_PSA in the QoS_GTE chooses the efficient and good channel condition path by means of picking the minimum transmission time path,

which takes the efficiency of using resource into account. Moreover, the TCDA make the several RSs transmit concurrently so that the resource can be fully used. Since the LMP selects the path only based on the bottleneck path-loss ignoring the overall path condition and the disadvantages on system while using relay path, it has poor performance in system throughput compared with the TT_PSA. Note that throughputs of the QoS_GTE w/o TCDA are less than that of w/o Relay case. This is because that twelve RSs share the resource in TDMA mode. In other words, each RS transmits the information in different symbol time. If the sub-channels in a symbol time for a certain RS cannot be used entirely, the spare resource is not allowed to be used by other RSs. This results in resource consuming. Therefore, TCDA is undoubtedly the main factor to increase the system throughput in using relay. By several RSs transmitting concurrently, TCDA advances the system efficiency and improves the system throughput.

As the traffic load keeps increasing, the throughputs of all algorithms start to decrease. The reason is that resource is used for QoS guarantee first. When the traffic becomes heavy, more urgent packets are waiting to be transmitted. Using the SO_RAA, the urgent packets whose S_m is greater than S_{th} will be served no matter how its channel condition is.
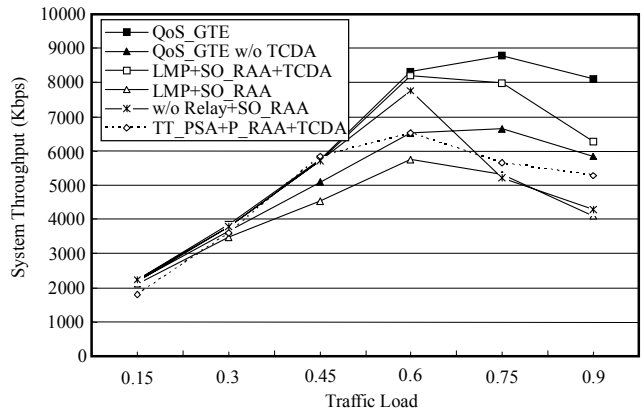

Figure 3. System Throughput

Note that the voice packet dropping rates are almost zero despite the varying of traffic load. The reason is that as the voice packets are urgent, they have the highest S_m so the SO_RAA will first serve these voice packets as soon as possible. The P_RAA also first serves the voice packets because voice packets' priority value is the highest.

Figure 4 shows the video packet dropping rate. The video packet dropping rate of the P_RAA starts to increase when the traffic load is at 0.45 but that of the QoS_GTE increases until the traffic load is at 0.75. This is because the P_RAA begins to serve the video packets when the voice packets are served entirely letting the urgency of video packets aside. When the amount of video packets becomes larger, the probability of serving video packets becomes lower and the dropping rate intuitively increases. The SO_RAA gives the prior to the packets with higher S_m in spite of its traffic type so the video dropping rate of SO_RAA can be suppressed.

Figure 5 illustrates the unguaranteed ratio for HTTP users. Unlike the real time traffic, packets of HTTP users will not be dropped but still waiting for service when they cannot reach the minimum transmission rate. We can observe that the unguaranteed ratio of the P_RAA is much higher than QoS_GTE. For the same reason above, the priority of HTTP packets is the third one which means that HTTP packets have

to wait until all real time packets are served. For the SO_RAA of QoS_GTE, since we set S_{th} equal to the average of HTTP maximum and minimum S_m, about half of the HTTP packets with S_m higher than S_{th} are probably be served. By this way, the situation of resource being occupied by certain traffics can be avoided and the HTTP users can be guaranteed to be served.
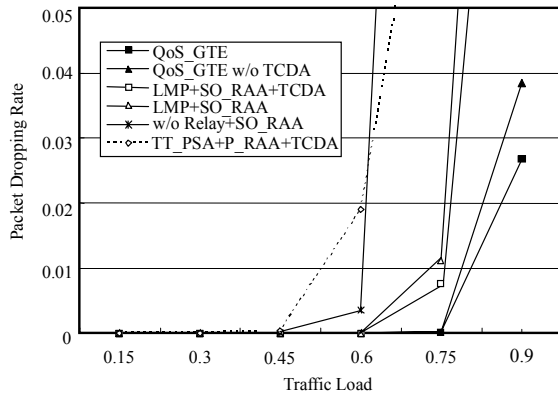
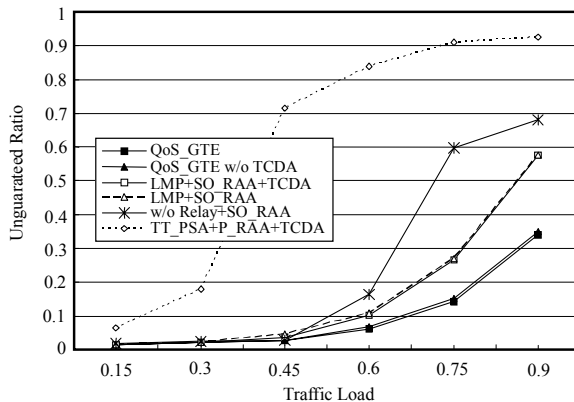

Figure 4. Video Packet Dropping Rate

Figure 5. Unguaranteed Ratio of HTTP User

IV. CONCLUDING REMARKS

In the paper, a downlink centralized QoS guaranteed and throughput enhancement (QoS_GTE) scheduling scheme for WiMAX relay-assisted network is proposed. The QoS_GTE scheduling scheme consists of TT_PSA, SO_RAA, and TCDA. The TT_PSA chooses the path with the minimal total transmission time as the transmission path. The SO_RAA gives high priority to the urgent users according to service order S_m at the current frame and dynamically adjusts the S_m of users frame by frame. The TCDA carries out resource reuse by deciding which RSs can transmit concurrently using the same frequency and time slots.

Simulation results show that the proposed QoS_GTE outperforms the without relay scheme with the improvement of 63% throughput and the LMP with the improvement of 28% in terms of system. The QoS_GTE also guarantees the QoS requirements of real time services by 97%. The TT_PSA increases the system modulation order by 14% compared with that of the LMP. The SO_RAA guarantees the QoS requirements and raises the throughput 50% compared with that of the P_RAA at the same time, which balances the tradeoff between the QoS guarantee and system throughput. The TCDA improves the throughput by 30%.

REFERENCES

[1] B. H. Walke, R. Pabst, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. P. Fettweis, "Relay-based deployment concept for wireless and mobile broadband radio," IEEE Commun. Mag., vol. 42, no.9, pp. 80-89, 2004.

[2] T. Hui, G. Xuelin, and Z. Ping, "The impact of relaying strategies on the performance in cellular system," Proc. ISCIT 2005, pp. 1357-1360.

[3] V. Sreng, H. Yanikomeroglu, and D. D. Falconer, "Relayer selection strategies in cellular networks with peer-to-peer relaying," Proc. IEEE VTC Fall 2003, pp. 1949-1953.

[4] J. Bae, R. Berry and M. L. Honig, "Power allocation, rate, and coverage for relay-assisted downlink data transmission," Proc. IEEE ICC June, 2006.

[5] J. Cho and Z. J. Hass, "On the throughput enhancement of the downstream channel in cellular radio networks through multihop relaying," IEEE Journal on Selected Areas in Communication, Vol. 22, no. 7, pp. 1206-1219, September 2004.

[6] L. Fu, Z. Cao, and P. Fan, "Spatial reuse in IEEE 802.16 based wireless mesh networks," Proc. ISCIT 2005, pp. 1311-1314.

[7] N. Challa, H. Cam, "Cost-aware downlink scheduling of shared channel for cellular network with relays," 2004 IEEE International Conference, Performance, Computing, and Communication, 2004, p.p 793-798.

[8] H. Wu, Ch. Qiao, S. De and O. Tonguz, "Integrated cellular and ad hoc Relaying system: iCAR," IEEE Journal on Selected Area in Communications, vol. 19, no. 10, pp. 2105-2115, October 2001.

[9] J. Jang and K. B. Lee, "Transmit power adaptive for multiuser OFDM systems," IEEE Journal on Selected Area in Communications, vol. 21, pp. 171-178, February 2003.

[10] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," Proc. IEEE VTC 2000, pp. 1085-1089.

[11] X. Zhang, E. Zhou, R. Zhu, S. Liu, and W. Wang, "Adaptive multiuser radio resource allocation for OFDMA system," IEEE Globecom 2005, pp. 3846-3850.

[12] A. J. Goldsmith and S. G. Chua, "Variable-rate rariable-power MQAM for fading channels," IEEE Transaction on Communication, vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[13] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," Intel Technology Journal, vol. 8, issue 3, 2004.

[14] V. Singh and V. Sharma, "Efficient and fair scheduling of uplink and downlink in IEEE 802.16 OFDMA networks," Proc. IEEE WCNC 2006, pp. 984-990.