Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**
Title	**End-to-End Throughput Metrics for QoS Management in 802.16j MR Systems**
Date Submitted	**2006-11-07**
Source(s)	Ozgur Oyman Voice: +1 (408) 653-5789 Email : ozgur.oyman@intel.com Sumeet Sandhu Nageen Himayat Intel Corporation 2200 Mission College Blvd. Santa Clara, CA 95054, U.S.A.
Re:	Response to a call for contributions on 2006-10-15 for the Relay TG, see C80216j-06/027.pdf
Abstract	This document proposes metrics to quantify end-to-end throughput of a routing path over multi-hop relay (MR) systems. These metrics capture the instantaneous physical layer channel quality over each hop and enable the MR BS to keep track of route qualities and end-to-end link quality of service (QoS) over multi-hop routes under varying physical channel conditions. Their computation is performed during uplink and downlink message exchanges among terminals (BS, RS, MS) in the MR network.
Purpose	To propose throughput metrics for end-to-end QoS management
Notice	This document has been prepared to assist IEEE 802.16. It is offered as a basis for discussion and is not binding on the contributing individual(s) or organization(s). The material in this document is subject to change in form and content after further study. The contributor(s) reserve(s) the right to add, amend or withdraw material contained herein.
Release	The contributor grants a free, irrevocable license to the IEEE to incorporate material contained in this contribution, and any modifications thereof, in the creation of an IEEE Standards publication; to copyright in the IEEE's name any IEEE Standards publication even though it may include portions of this contribution; and at the IEEE's sole discretion to permit others to reproduce in whole or in part the resulting IEEE Standards publication. The contributor also acknowledges and accepts that this contribution may be made public by IEEE 802.16.
Patent Policy and Procedures	The contributor is familiar with the IEEE 802.16 Patent Policy and Procedures <http://ieee802.org/16/ipr/patents/policy.html>, including the statement "IEEE standards may include the known use of patent(s), including patent applications, provided the IEEE receives assurance from the patent holder or applicant with respect to patents essential for compliance with both mandatory and optional portions of the standard." Early disclosure to the Working Group of patent information that might be relevant to the standard is essential to reduce the possibility for delays in the development process and increase the likelihood that the draft publication will be approved for publication. Please notify the Chair <mailto:chair@wirelessman.org> as early as possible, in written or electronic form, if patented technology (or technology under patent application) might be incorporated into a draft standard being developed within the IEEE 802.16 Working Group. The Chair will disclose this notification via the IEEE 802.16 web site <http://ieee802.org/16/ipr/patents/notices>.

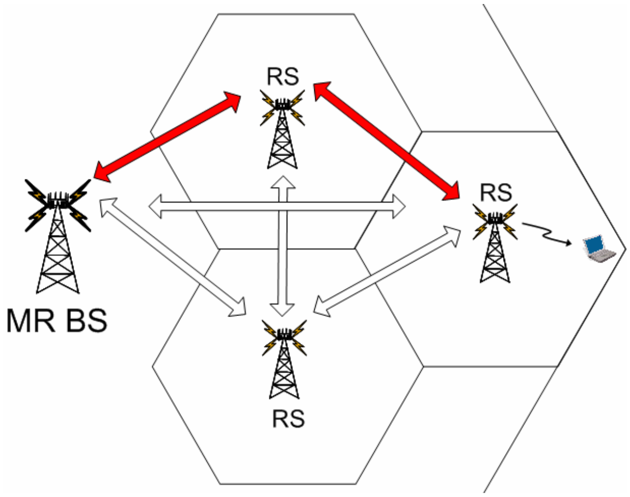
End-to-End Throughput Metrics for QoS Management in 802.16j MR Systems

Ozgur Oyman, Sumeet Sandhu and Nageen Himayat

Intel Corporation

1. Summary of Contribution

The goal of this contribution is to propose throughput metrics for end-to-end quality of service (QoS) management in 802.16j multi-hop relay (MR) systems. In particular, these metrics enable the computation of end-to-end throughput during uplink and downlink message exchanges among terminals (BS, RS, MS) in the MR network.



Knowledge of such end-to-end QoS metrics reflecting the physical channel conditions over multiple hops enables the MR BS to

- Manage end-to-end quality of service (QoS) requirements (e.g., throughput, reliability, latency etc.),
- Coordinate centralized functionalities (e.g., execution of scheduling algorithms, routing algorithms, other MAC and higher layer functions).

2. Technical Description of Proposed Metrics

The following PHY abstraction relationship is the key formula for quantifying the end-to-end quality of an N-hop routing path in terms of throughput in an MR network:

$$End\_to\_End\_Throughput = \left(\sum_{n=1}^{N} \frac{1}{Per\_Link\_Throughput\_over\_Hop\_n} \right)^{-1}.$$

In other words, based on this PHY abstraction, MR BS obtains an estimate of the end-to-end throughput by computing the harmonic mean of the per-link throughputs over the individual hops.

To motivate the quantification of end-to-end throughput as the harmonic mean of the throughputs over the individual wireless links, consider an N-hop routing path such that the transmission rate at hop n is R_n bits/second/Hertz and all hops operate over the common bandwidth W but orthogonally time-share the channel. If the transmitted packet contains B bits of information, then the required transmission time at hop n is $t_n = B/WR_n$ seconds (ignoring retransmissions and channel overheads for now). Thus, the end-to-end latency T (i.e., the total time required to transmit this packet in multiple hops over the routing path), and the resulting end-to-end throughput (in bits/second/Hertz) can be calculated as

$$T = \sum_{n=1}^{N} t_n = \sum_{n=1}^{N} \frac{B}{WR_n} \qquad \rightarrow \qquad Throughput = \frac{B}{WT} = \left(\sum_{n=1}^{N} \frac{1}{R_n} \right)^{-1},$$

which validates the harmonic mean relationship. Alternatively, we can consider an N-hop routing path such that the transmission rate at hop n is R_n bits/second/Hertz, all transmissions simultaneously occur over the time duration of T seconds and the bandwidth is allocated orthogonally among hops. Hence the channel over hop n should be allocated $W_n = B/TR_n$ Hertz of bandwidth, and accordingly the total required bandwidth W and the resulting end-to-end throughput can be calculated as

$$W = \sum_{n=1}^{N} W_n = \sum_{n=1}^{N} \frac{B}{TR_n} \qquad \rightarrow \qquad Throughput = \frac{B}{WT} = \left(\sum_{n=1}^{N} \frac{1}{R_n} \right)^{-1},$$

yielding the identical harmonic mean relationship, which therefore holds for any orthogonal sharing of resources (in time or frequency) among hops over a given routing path.

In the following, we shall provide two approaches for designing per-link metrics to allow the MR BS estimate the end-to-end throughput over a multi-hop link. It must be emphasized that other per-link metrics for quantifying throughput over the individual hops can be designed as well; and one can still estimate end-to-end throughput in terms of per-link throughputs through the harmonic mean formula.

A. Capacity-based Approach:

The capacity-based approach computes the harmonic mean of the capacities (only depends on the signal-to-interference-and-noise ratio (SINR)) over individual wireless links and obtains an end-to-end PHY abstraction given as follows:

$$SINR_{eff} = 2^C - 1, \quad where \qquad C = \frac{1}{\sum_{n=1}^{N} \frac{1}{C_n}} = \frac{1}{\sum_{n=1}^{N} \frac{1}{\log_2(1 + SINR_{eff,n})}}$$

Where:

- $SINR_{eff}$: Effective SINR for the multi-hop route
- C : Effective end-to-end capacity for the multi-hop route
- C_n : Effective capacity over hop n, $n = 1,..., N$
- $SINR_{eff,n}$: Effective SINR over hop n
- N : Number of hops over the established route between MR BS and MS

Further analysis on this capacity-based end-to-end PHY abstraction metric can be found in [1]-[4]. The effective SINR parameter, $SINR_{eff,n}$, is determined on a per-hop basis over each individual OFDMA PHY link, using any PHY abstraction methodology. In this respect, for an OFDMA system with K subcarriers, denoting the SINR over the k-th subcarrier and n-th hop by $SINR_{n,k}$, $n = 1,..., N$, $k = 1,..., K$, example PHY abstraction metrics could be

- the mean-capacity metric proposed by [5]-[7], where

$$\log_2(1+SINR_{eff,n}) = \frac{1}{K}\sum_{k=1}^{K}\log_2(1+SINR_{n,k}) \qquad \Rightarrow SINR_{eff,n} = 2^{\frac{1}{K}\sum_{k=1}^{K}\log_2(1+SINR_{n,k})} - 1,$$

- or the exponential effective SNR mapping (EESM) metric proposed by [8]-[11], where, for a given choice of parameter β,

$$SINR_{eff,n} = -\beta\log\left(\frac{1}{K}\sum_{k=1}^{K}\exp\left(-\frac{SINR_{n,k}}{\beta}\right)\right).$$

B. Throughput-based Approach:

This PHY abstraction approach involves more computation but is also more accurate than the capacity-based approach in terms of quantifying the end-to-end throughput performance. Unlike the capacity-based approach, the per-link throughput estimation accounts for losses in data rate due to link errors, finite modulation and coding schemes (MCSs) and overheads associated with channel access and protocols. In this setting, the end-to-end throughput estimation is based on computing the harmonic mean of achievable throughputs over individual wireless links.

We denote the expected transmission time (ETT) (first proposed by [12] for 802.11s mesh standard) over hop n as ETT_n , which represents the overall airtime cost in terms of the amount of channel resources consumed by transmitting the packet over the particular link and includes the cost of data transmission as well as cost of necessary retransmissions to recover from packet decoding errors and cost of overhead. Accordingly, the end-to-end throughput for a given multi-hop route of length N can be expressed as

$$Throughput = \frac{B}{T} = \frac{B}{\sum_{n=1}^{N} ETT_n} \qquad where \qquad ETT_n = \left[T_{overhead} + \frac{B}{R_n} \right] ETX_n .$$

Where:

ETX_n : Expected number of packet transmissions until successful reception over hop n subject to instantaneous channel conditions

R_n : Aggregate data rate per packet over hop n based on the MCS chosen by the link adaptation algorithm while satisfying a certain target packet error rate (PER) subject to instantaneous channel conditions

B : Number of bits per packet

$T_{overhead}$: Latency cost per link due to fixed channel access and protocol overheads

Note that the total sum of ETTs has a physical meaning as well; it is an estimate of the total end-to-end latency $T = \sum_{n=1}^{N} ETT_n$ experienced by a packet traveling along that path.

In the ETT formula, R_n and ETX_n both depend on the instantaneous channel realizations over hop n; in particular they depend on the vector of received SINRs over the K OFDMA subcarriers given by $\{SINR_{n,k}\}_{k=1}^{K}$ or effectively they depend on a single measure $SINR_{eff,n}$ computed using the per-link PHY abstractions.

To compute ETX_n more explicitly, we consider a hybrid automatic repeat request (HARQ) mechanism over each link, which requires packet retransmissions upon decoding failure, continued until successful reception of each packet. Thus, the definition of ETT also incorporates the impact of retransmissions upon erroneous reception of transmitted frames/packets and hence takes into account the additional transmission time necessary until successful delivery to the destination. Now,

- If the HARQ protocol discards erroneous packets completely, we have

$$ETX_n = \frac{1}{1 - PER_n},$$

where PER_n denotes the packet error rate over hop n determined based on the chosen MCS and channel conditions over the physical layer given by the vector of SINRs in the set $\{SINR_{n,k}\}_{k=1}^{K}$.

- If the HARQ protocol stores previously received erroneous packets and uses them during later decoding attempts (e.g. as in chase combining), the PER will improve upon retransmissions, which leads to

$$ETX_n = 1 + \sum_{m=1}^{M} \prod_{j=1}^{m} PER_{n,j},$$

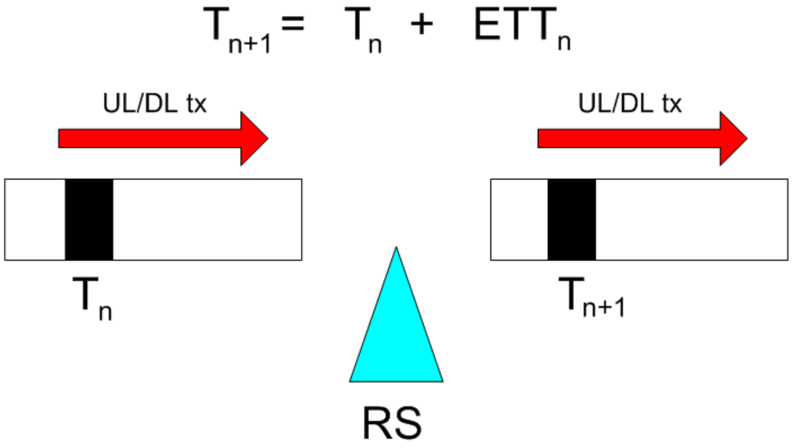
where $PER_{n,j}$ denotes PER over hop n during transmission j and M is the maximum number of allowed transmissions ($M-1$ retransmissions allowed).

Due to the stationarity of the terminals over the wireless backhaul (i.e. MR BS and RSs), we expect that the channels experienced over the hops will be slow-fading (except for the last hop involving the MS) and each node will be able to track its transmit/receive channels to perform link adaptation and estimate supportable throughput.

Considering the cost of each link to be equal to ETT_n, the throughput-maximizing path (i.e., the path that maximizes end-to-end throughput or minimizes end-to-end latency T) is the path that minimizes total end-to-end routing cost given by $T = \sum_{n=1}^{N} ETT_n$ for a path of length of N hops. The use of such end-to-end link QoS metrics toward designing centralized scheduling algorithms for MR systems have been studied in [13].

3. Computation of End-to-End Metrics during UL/DL Message Exchanges

The computation of the end-to-end link cost metric $\sum_{n=1}^{N} \frac{1}{C_n}$ in the capacity-based approach or $\sum_{n=1}^{N} ETT_n$ in the throughput-based approach requires a field in the uplink (UL) or downlink (DL) frame to store the end-to-end accumulated route costs, i.e. the summation of the reciprocal capacities or ETTs over various hops, which is to be propagated and updated during uplink (UL) or downlink (DL) message exchanges, as depicted below for the throughput-based approach. As the UL/DL frame gets passed among RS terminals over the multi-hop route between MR BS and SS, this field carrying the information on the end-to-end link cost will be updated using the estimated routing metric (i.e. over hop n, this routing metric is the estimated $1/C_n$ for the capacity-based approach and estimated ETT_n for the throughput-based approach) to account for the throughput/latency of transmissions over the current hop.



It should be emphasized that this section presented one approach to compute the end-to-end link QoS metric on an aggregate basis in a distributed fashion to reduce feedback overhead complexity. However, alternative approaches exist toward the computation of the end-to-end metrics; for instance another option may consider feeding back individual per-hop link quality metrics to the MR BS separately and performing the end-to-end link quality computations at the MR BS in a centralized fashion.

4. References

[1] O. Oyman and S. Sandhu, “Throughput Improvements in Micro-cellular Multi-hop Networks”, 802.16 MR Study Group Contribution, Nov. 2005

[2] O. Oyman and S. Sandhu, “A Shannon-Theoretic Perspective on Fading Multi-hop Networks”, Proc. CISS, Princeton, NJ, March 2006

[3] O. Oyman, “Reliability Bounds for Delay-Constrained Multi-hop Networks”, Proc. Allerton Conference, Monticello, IL, Sep. 2006

[4] O. Oyman, “End-to-End Throughput and Latency Measures for Multi-hop Routing in Relay-Assisted Broadband Cellular OFDM Systems”, Proc. IEEE Radio and Wireless Symposium (RWS), Long Beach, CA, Jan. 2007.

[5] “TGn Sync TGn Proposal MAC Simulation Methodology”, Sony, Intel, IEEE 802.11-04/895r2, November 2004.
[6] A. Poloni, S. Valle and G. Villa, “Time Correlated Packet Errors in MAC Simulations”, ST Micro-Electronics, IEEE Contribution, 802.11-04-0064-00-000n, Jan. 2004.
[7] “Unified Black Box PHY Abstraction Methodology”, Atheros, Mitsubishi, ST Micro-Electronics and Marvell Semiconductors, IEEE Contribution 802.11-04/0218r1, March 2004.
[8] “Feasibility Study for OFDM for UTRAN enhancement,” 3GPP TR 25.892 V2.0.0.
[9] “System level evaluation of OFDM- further considerations”, Ericsson, TSG-RAN WG1 #35, R1-03-1303, November, 2003
[10] “Effective SIR Computation for OFDM System-Level Simulations,” Nortel, TSG-RAN WG1 #35, R03-1370, November 2003.
[11] “OFDM Exponential Effective SIR Mapping Validation, EESM Simulation Results for System-Level Performance Evaluations,” Nortel, 3GPP TSG-RAN1 Ad Hoc, R1-04-0089, January, 2004.
[12] R. Draves, J. Padhye and B. Zill, “Routing in Multi-Radio Multi-hop Wireless Mesh Networks”, ACM Mobicom, Philadelphia, PA, Sep. 2004

[13] O. Oyman, “OFDM^{2}A: A Centralized Resource Allocation Policy for Cellular Multi-hop Networks”, Proc. IEEE Asilomar Conference, Monterey, CA, Oct. 2006