Simulative Investigation on VOIP over WiMAX Communication Network

Ambica
BGIET,Sangrur

AvinashJethi
BGIET,Sangrur

ABSTRACT
The objective of this research paper is to provide brief overview on simulative investigation on VOIP over WiMAX communication network for static environment. The mobile WiMAX (IEEE802.16e) support VOIP traffic under different scenario, QOS, the system throughput, the packet delay,the signaling overhead,various design choices and performance analysis. This paper reports about the various VOIP application in the IEEE802.16e including all traffic contract were made correctly by considering overheads and performance of voice in mobile environment.

Keywords
voice over internet protocol(VOIP), IEEE802.16e , WiMAX

1. INTRODUCTION
WiMAX is the Next Generation of Wireless Broadband based on IEEE 802.16 standard. It provides broadband connectivity anywhere, anytime, for any device and on any network and can connect to the internet in faster speed and wider coverage. WiMAX is most suitable for home users, individual, small office and home office etc. It erases the suburban and rural blackout areas that currently have no broadband Internet access.

Number of users and their location from the Base Station play an important role in the network performance. Thus, in this paper, the simulative investigation includes the variation in the number of subscriber stations requesting for VoIP traffic and the distance between the subscriber station and base station.The focus of the paper is on the support of VoIP over mobile

WiMAX (802.16e). Delay sensitive applications, such as VoIP, pose significant challenges to the design of wireless network infrastructures due to their stringent QoS requirements and to the typical hostile radio environments; the mobility of users is obviously an extra-dimension that adds to the overall complexity, as it is more difficult to predict the radio channel conditions, and hence to allocate resources efficiently at high users speeds. VoIP applications are characterized by tight air-interface delay budgets (usually in the order of tens of ms) and a typical QoS target of less than 1%-2% packet loss for 95% - 98% of active VoIP users. Another major challenge is posed by the amount of overhead required to support many simultaneous VoIP connections, which unless controlled in an efficient and intelligent manner, may grow unacceptably large. In WiMAX, the overhead includes the downlink map (DL-MAP) and the uplink map (UL-MAP) overheads. DL-MAP is a Medium Access Control (MAC) message that defines burst start times for a subscriber station on the downlink. UL-MAP supports a set of information that defines the entire access for a scheduling.

1.2 VOIP DESIGN CHOICES, RESOURCES

ALLOCATION AND MODELING ASSUMPTIONS
he speech traffic is generated according to a two-state Adaptive Multi-Rate (AMR) speech codec. Depending on the air-interface delay budget, VoIP packet bundling can be considered at the base station, i.e., when creating MAC PDUs. This form of bundling offers the advantage of reducing overhead at the expense of some additional delay given by N*20 ms where N is the number of VoIP packets that are bundled into a single packet. Bundling offers the benefit that a single MAC header may be used for a bundled VoIP packet containing payload from multiple VoIP packets; however, the compressed header is still needed for each constituent packet.

Packet bundling also allows the MAP overhead to be reduced since fewer bursts may need to be scheduled for transmission over the air interface. Silence suppression is achieved through blanking out the eight-rate speech frames, which in turn facilitates statistical multiplexing of VoIP users and also plays a significant role in reducing the interference levels. Each VoIP packet/bundle is mapped to a physical Orthogonal Frequency Division Multiple Access (OFDMA) burst and corresponds to a MAC Protocol Data Unit (PDU). This allows the Hybrid Automatic Request (HARQ) mechanism to operate on a per-burst basis, since the Cyclic Redundancy Check (CRC) is embedded in the MAC PDU. The MAC PDU overhead is 8 bytes, resulting from 6 bytes MAC header overhead and 2 bytes CRC for HARQ bursts. Robust header compression (ROHC) is employed to reduce the original 40 bytes of Real Transport Protocol/User Datagram Protocol/Internet Protocol (RTP/UDP/IP) overhead to 3 bytes only. Partially Used Sub-channelization (PUSC) based on distributed sub-carriers allocation to sub-channels is employed. To support the uplink traffic, control slots are allocated for ranging, channel quality indication and downlink HARQ acknowledgments, as part of the uplink sub- frame. OFDMA symbols were considered sufficient to carry out the uplink overhead in the context of an Extended Real Time Polling Service (eRTPS) like VoIP [1-2]. In the downlink sub-frame, control slots are allocated for overhead including preamble (1 symbol), Frame Control Header (4 OFDMA slots) and MAP messages. MAP messages constitute the majority of the overhead and are used for allocating resources to VoIP packets (i.e. number of sub-channels and symbols, and their respective allocation offsets). Combining packets bundling and silence suppression features may offer desirable benefits in terms of reducing the total overhead.

It is necessary to ensure that MAP messages are reliably received by a very large fraction of users since VoIP packets cannot be successfully decoded without first decoding the MAP messages. MAP messages require the use of QPSK Rate 1⁄2 encoding; however, the reliability may be further improved through slot repetition factors of 2, 4 or 6 –

however, the higher the repetition factor, the larger the overhead grows. In order to avoid having to send MAP messages to all users with repetition factors of 4 or 6 at QPSK Rate 1⁄2, telescopic MAP usage is employed as a way of reducing MAP overhead. With the telescopic MAP approach, DL/UL-MAP messages in the frame are broken up into one or more sub-MAPs (overhead portion of the downlink sub-frame), and instead of transmitting the whole MAP at one low data rate (e.g. QPSK Rate 1⁄2 with repetition factor of 4), only the compressed MAP is transmitted with repetition factors and the sub-MAPs are transmitted at higher data rates (QPSK Rate 1⁄2 or higher). For this VoIP analysis, we assume the transmission of a single compressed DL/UL-MAP message that is sent using QPSK Rate 1⁄2 with a repetition factor of 4 followed by a sub- DL-UL-MAP message that is sent using QPSK Rate 1⁄2. The compressed DL/UL-MAP carries information elements for resource allocation to about a quarter of the users while the sub-DL-UL-MAP is used to allocate resources to the rest of the users. MAP messages are constructed such that the messages may be reliably received by a large fraction of the mobile station.

2. VOICE TRAFFIC MODEL OF USERS

An exponentially distributed on-off model is used to characterize traffic from a single voice source, where the mean on-time is $\alpha^{-1} = 1$ s and the mean off-time is $\beta^{-1} = 1.5$ s . The extended real-time polling service (ertPS) is assumed to be used for the uplink scheduling, where the BS allocates a fixed amount of bandwidth at periodic intervals during the on state. The user informs the BS of its transition between the on state and the off state by either using a piggyback request field or sending a codeword over a channel quality indicator channel (CQICH) . The codeword minimizes the signaling overheads and prevents delays in the scheduling of uplink requests and grants. Hence, for the sake of simplicity, this paper ignores the uplink request and grant procedure.Each user is assumed to generate VoIP traffic every p frames in the on state, where the size of a VoIP protocol data unit is fixed and the size is denoted by L. For example, the G.723.1 codec generates a 24 byte encoded voice frame every 20 ms Assuming the size of a VoIP packet is fixed, a BS can calcu- late the number of packets to be transmitted in the uplink from the amount of requested bandwidth or the amount of allocated bandwidth. However, the use of the variable rate codec such as the adaptive multi-rate (AMR) codec makes it difficult to count the number of packets from the bandwidth request. The case of variable-rate VoIP codecs can be studied in future works. This paper assumes that a VoIP packet has a fixed size The VoIP traffic requested from N_v users is aggregated in the initial transmission queue at the BS. The operation of the initial transmission queue can be modeled as a two-state Markov- modulated Poisson process (MMPP) with the transition rate ma- trix, R, and Poisson arrival rate matrix, Λ .

3. PERFORMANCE ANALYSIS

No error in the transmission of the signaling message is assumed because the signaling message is transmitted with a low MCS level. In the CC-based HARQ transmissions, the probability of successful decoding for a VoIP packet transmitted with MCS level i at the kth transmission is given by

$$\mathbf{P}_i(k|\gamma) = (1 - PER_i(k\gamma)) \$ PER_i(n\gamma)$$

where $PER_i(\gamma)$ is the associated PER of MCS level i at SNR γ. The probability of a VoIP packet being transmitted at the kth transmission after $k - 1$ failures is given by

$$Q_i(k|\gamma) = \$ PER_i(n\gamma)$$

Let α_i be defined as the probability of a packet modulated with MCS level i being successfully transmitted in N_{max} transmissions. The value of α_i can be expressed as the sum of the probabilities of successful decoding at the kt transmission. Thus,

$$\alpha_i = \{P_i(k|\gamma) f_\gamma / P_\gamma(i) d\gamma\}$$

$$= 1/P_\gamma(i) P_i(k|\gamma) f_\gamma(\gamma) d\gamma \leq 1$$

where the denominator, $P_\gamma(i)$, is a normalized factor. Then, out of x_i packets initially transmitted with MCS level i, the total number of packets successfully transmitted in N_{max} is

$$u_i = \alpha_i x_i .$$

4. QUALITY OF SERVICE (QoS) IN IEEE 802.16e

Originally, four different service types were supported inthe 802.16 standard: UGS, rtPS, nrtPS and BE.The UGS (unsolicated grant service) is similar to the CBR(constant bit rate) service in ATM which generates a fixed size burst periodically .this service can be used to replace T1/E1 wired line or a constant rate services.it can also be used to support real time applications such as Voip or streaming applications. Even though the UGS is simple ,it may not be the best choice for the VoIP in that it can waste bandwidth during the off period of voice calls.

The rtPS(real time polling service) is for a variable bit rate real time service such as VoIP. Every polling interval ,BS polls a mobile and the polled mobile transmit bw request (bandwidth request) if it has datato transmit. The BS grants the data burst using UL-MAP-IE upon its reaction.ThenrtPS (non-real-time polling service) is very similar to the rtPS except that it allows contention based polling.

The BE (Best Effort) service can be used for applications such as e-mail or FTP, in which there is no strict latency requirement. The allocation mechanism is contention based using the ranging channel. Another service type called ertPS (Extended rtPS) was introduced to support variable rate real-time services such as VoIP and video streaming. It has an advantage over UGS and rtPS for VoIP applications because it carries lower overhead than UGS and rtPS.

5. PACKET DELAY ANALYSIS

The average packet delay, which is the sum of the average waiting time in the queue, W_q , and the average transmission time in the air, T_{tr} , can be expressed as follows

$$D = W_q + T_{tr}.$$

The queue waiting time includes the waiting time in the initial transmission queue and the waiting time in the retransmission queue. However, the waiting time in the retransmission queue can be neglected because the BS schedules users firstly from the retransmission queue and then from the initial transmission queue. The average length of the initial transmission queue is given by

$$L_q = \quad k\pi(k).$$

The average waiting time in the initial transmission queue can be expressed in terms of Little's theorem. Thus,

$$W_q = L_q / \lambda_e$$

where λ_e , which is the effective packet arrival rate at the initial transmission queue during the frame, is equal to the average number of scheduled packets: That is,

$\lambda_e = x$.

When a packet is successfully transmitted at the kth transmission, the total transmission time is given

$TR(k) = (k - 1)RTT + T_f$.

5.1 Packet Delay Analysis In A Virtually Dedicated Resource Allocation Region

A. System Model

Although there may exist several VDRA regions in a system, we first analyze a single VDRA region case, because one VDRA region can be independent of the other regions. The DL packet transmission model consists of the following five component**s:**

1) L buffers for storing the arriving packets for the corresponding L users;

2) a VDRA module that performs the operation mentioned in Section III at each frame;

3) N_{ru} resource units for packet transmissions;

4) L receivers, which decode the received packets, check whether an error occurs, and generate ACK/negative acknowledge (NACK) signals for retransmission;

5) a delay element (DE), which stores ACK/NACK signals generated at the receiver and feeds them back to the transmitter. This overall packet transmission model can be simplified as a model for a single user as shown in Fig. 5(b), where a packet is served with a probability of μ(L) in a frame. μ(L) is affected by the buffer empty probability of all the buffers in the VDRA region because of resource collisions.

The VDRA region has the following parameters.

1) M : the MCS level supported by the VDRA region;

2) S_{ru} : the size of the resource unit;

3) B_{ru} : the number of bits which can be transmitted in one resource unit. This value depends on the M and S_{ru} ;

4) N_{ru} : the number of resource units in the VDRA region;

5) L: the number of users accommodated in the VDRA region;

6) λ: the packet arrival probability in a frame time. We assume that packet arrivals follow a Bernoulli process with a packet arrival probability of λ in a frame time T ;

7) α: the probability that the transmitted packet is negatively acknowledged (NACK) because of a transmission error;

8) D_r : time duration (in frame times) that ACK/NACK signal for a transmitted packet is fed back to the BS;

9) $U_l(F_N)$: the resource unit number generated by the PHP

In addition, we will derive the following analytical results.

1) $P_S(N_D, N_R|m)$: the probability that N_D and N_R users transmit packets using the predetermined and redirected

resource units, respectively, when m buffers have pending packets to transmit;

2) μ_m : the probability that a packet is transmitted to a specific user, regardless of the use of predetermined or redirected resource unit, where m buffers have pending packets;

3) μ(L): the probability that a packet is transmitted to a specific user in a frame when L users are accommodated in the VDRA region;

4) $E[N_D]$: the mean number of predetermined resource units used in the VDRA region;

5) $E[N_R]$: the mean number of redirected resource units used in the VDRA region;

6) $\varphi_0(L)$: the probability that a buffer is empty when L users are accommodated in the VDRA region;

7) P: the state transition probability matrix for the number of packets in a buffer;

8) Π: the steady-state probability for the number of packets in a buffer;

9) $P_d(D|L)$: the probability that packet delay is D frame times when L users are accommodated in the VDRA region.

6. SIGNALLING SCHEMES OF BS

6.1 Conventional Allocation

In conventional mobile WiMAX systems, the characteristic that the BS broadcasts a signaling message for every frame generates a substantial overhead for every frame. Hence, the size of the signaling overhead is directly proportional to the number of scheduled users.

6.2 Persistent Allocation

Persistent allocation is a technique used to reduce the signaling overhead for connections, e.g., VoIP services, that have a periodic traffic pattern and a relatively fixed payload size .A high-level concept of a persistent mapping scheme and a conventional mapping scheme. In the persistent allocation scheme, the BS allocates a persistent resource to a user at frame t and the allocated resource is valid in a periodic sequence of future frames, namely, frame t + p, frame t + 2p, without notification of a signaling message. However, if the optimized modulation and coding scheme (MCS) level at the current frame is different from the latest MCSlevel indicated by the BS, the BS may transmit a signaling message in order to adjust the attributes of a persistently allocated resource because the MCS mismatch causes a link adaptation error.

7. THROUGHPUT ANALYSIS

A discrete-time MMPP can be equivalent to an MMPP in continuous time. The system state is defined as the number of packets in the initial transmission queue. The arrival and service process of the initial transmission queue is depicted .

The average packet arrival rate at the initial transmission queue during the frame is expressed as follows

$\rho = s(k D_k)$

where 1 is a column matrix of ones. The matrix $s = [s_1\ s_2]$ is obtained by solving $sU = s$ and $s_1 + s_2 = 1$, where U, which is the phase transition probability matrix in the MMPP, is expressed as $U = (\Lambda - R)^{-1} \Lambda$. Additionally, D_k is the diagonal probability matrix in which each diagonal element is the probability of k packets arriving at the BS for the frame, retransmission attempts along with the transmissions per packet.

(λ i T f) k e −λ i T f /k! for i = 1, 2.

8. RESULTS

Table 1 :system simulations assumptions and paramters

Parameter	Assumption
Network Topology	19 cloverleaf cells with three sectors with wrap-around enabled
Duplexing Scheme	TDD
Bandwidth	10 MHz
FFT size	1024
ISD (Inter-Site Distance)	500 m (ITU Ped. B 3Km/h); 1500m (ITU Veh.A 60Km/h)
Propagation Formula	128.1 + 37.6*log10(d) dB, d inKm, for 2GHz
Carrier Frequency	2.5 GHz
Number of VoIP sources/users per sector simulated	Variable, subject to capacity
VoIP Capacity Criteria	8% of users satisfied. Users with less than 2% packet loss are declared satisfied
Reference VoIP Traffic Source	AMR 12.2 kbps; AMR 7.95 kbps
Transmission Method	MIMO, STC (2x2, DL)
Maximum number of HARQ transmissions (1 st trans + retransof same packet)	4
MAC Header	48 bits
Total Packet Overhead	88 bits

The following performance metrics are used for evaluating the VoIP users in this paper:

VoIP Packet Loss: it captures packet losses at the receiver due to residual post-HARQ errors (after a total of maximum allowed number of transmissions) or due to the discarding of packets that have exhausted their delay budgets

at the transmitter.

VoIP Packets Transmmission Statistics: it captures the percentage of packet transmissions across succesive retransmission attempts along with the average number of transmissions per packet. These characteristics are driven by the H-ARQ mechanisms.

9. CONCLUSION

In this paper, we covered performance analysis,voice traffic model of users, QOS, packet delay analysis which includes the different procedures set for all the system simulation. We verified the analytical results by comparing them with simulation result. The HARQ mechanism may deteriote the QOS of real time VoIP services because of the increment in the transmission delay. Hence, the maximum number of retransmissions should be appropriately configured with due consideration of allowable transmission delay. The results also show that signaling overhead is considerable and can cause serious spectral inefficiency for VoIP services. The persistent allocation scheme significantly decreases the signaling overhead by allocating resources persistently without giving any notification of the signaling information.

10. REFERENCES

[1] M.-H. Fong, R. Novak, S. McBeath, and R. Srinivasan, “Improved VoIP capacity in mobile WiMAX systems using persistent resource allocation,” IEEE Commun. Mag., pp. 50–57, Oct. 2008.

[2] DoruCalin Bell Laboratories, Alcatel-Lucent,VoIP over Realistic IEEE 802.16e System Scenarios: The Uplink Direction 2011.

[3] D. Calin, “VoIP over Realistic IEEE 802.16e System Scenarios: The Uplink Direction”, in Proc. of IEEE Globecom 2012, pp. 3384 – 3388, Anaheim, December 2012.

[4] DoruCalin Bell Laboratories, Alcatel-Lucent, Design and Analysis of a VoIP Based IEEE 802.16e System,2013.

[5] S. Alshomrani, S. Qamar, S. Jan, I. Khan and I. A. Shah,QoS of VoIP over WiMAX Access Networks,april 2012.

[6] JoãoHenriques, Vitor Bernardo, Paulo Simões, Marilia Curado Center for Informatics andSystems,VoIP performance over Mobile WiMAX: An Urban Deployment Analysis,2012.