Real-time-service-based Distributed Scheduling Scheme for IEEE 802.16j Networks

Kuo-Feng Huang, Shih-Jung Wu
Taipei College of Maritime Technology/Visual Communication Design, New Taipei City, Taiwan
Tamkang University/Innovative Information and Technology, New Taipei City, Taiwan
Email: sailer5069@gmail.com, wushihjung@mail.tku.edu.tw

Abstract—Supporting Quality of Service (QoS) guarantees for diverse multimedia services is the primary concern for IEEE 802.16j networks. A scheduling scheme that satisfies the QoS requirements has become more important for wireless communications. We proposed an adaptive nontransparent-based distributed scheduling scheme (ANDS) for IEEE 802.16j networks. ANDS comprises three major components: Priority Assignment, Resource Allocation, Preserved Bandwidth Adjustment. Different service-type connections primarily depend on their QoS requirements to adjust priority assignments and dispatch bandwidth resources dynamically. Meanwhile, we promote the connections, which do not satisfy QoS requirements, to avoid the delay and starvation. Simulation results show that our APS methodology outperforms the representative scheduling approaches in both QoS satisfaction and maintains fairness in starvation prevention.

Index Terms—Distributed Scheduling, IEEE 802.16, Relay, Real-time, Non-Transparent

I. INTRODUCTION

The internet service has become the necessity of modern society. The demand of internet results in spreading internet constructions no matter in urban area or countryside. The wired network system meets more restrictions and suffers more difficulties. To save the cost of construction time and decrease the construction complexity when deploying wired network in a developed city, the wireless network system seems a better solution. WiMAX (Worldwide Interoperability for Microwave Access) system is the mainstream of wireless network technology [9-14]. The IEEE 802.16 standard is developed as the guideline of wireless network technology [4, 8]. The IEEE 802.16 standard is developed as the guideline of WiMAX system. The main object of the standard is to ensure that the device from different manufacturers won’t cause the compatibility problems [4, 8].

For the convenience of using internet resource, the goal of wireless network system is to provide network service as possible as it could. However, providing network service to a blind spot or a sparsely populated area by a BS usually substantially increases the cost to system suppliers. The RS architecture which is specified in the IEEE 802.16j, as the extension of IEEE 802.16, could overcome these problems by multi-hop relaying technology [6].

II. RELATED WORKS

To overcome the compatibility problems to the existing WiMAX system and unify the specifications from different manufacturers, the IEEE 802.16j working group is dedicated to establish the standard of multi-hop relay technology. The relay technology is a new issue because of the integration of relay station and the existing network system.

Relay stations can be classified into two classes by whether the Preamble and UL/DL MAP being broadcasted. The non-transparent RS supports the broadcasting of Preamble and UL/DL MAP but the transparent RS doesn’t.

A. Transparent RS

The frame structure of a transparent RS is based on the two-hop transparent relaying specified in the IEEE 802.16j standard. It includes the MR-BS frame structure and the transparent RS frame structure. One transmission frame could be divided into DL sub-frame and UL sub-frame. The DL sub-frame of MR-BS is divided into DL Access Zone for MS and DL Relay Zone. The UL sub-frame of MR-BS is divided into UL Access Zone for MS and UL Relay Zone. The DL sub-frame of transparent RS is divided into DL Access Zone in Receiving Mode and DL Relay Zone. The UL sub-frame of transparent RS is divided into UL Access Zone and UL Relay Zone in Transmitting Mode. The RS supports the relaying when the transmission between BS and MS is decided to use two-hop relaying. The packet from BS is delivered to RS by and relay to MS in downlink transmission, and the packet from MS is delivered to RS and relay to BS in uplink transmission.

B. Non-Transparent RS

The frame structure of a non-transparent RS is based on the two-hop transparent relaying specified in the IEEE 802.16j standard. It includes the MR-BS frame structure and the non-transparent RS frame structure. The DL sub-frame of MR-BS is divided into DL Access Zone and DL Relay Zone. The UL sub-frame of MR-BS is divided into UL Access Zone for MS and UL Relay Zone. The DL sub-frame of non-transparent RS is divided into DL Access Zone for MS and DL Relay Zone in Receiving Mode. The UL sub-frame of non-transparent RS is divided...
into UL Access Zone for MS and UL Relay Zone in Transmitting Mode.

In the network system with non-transparent RS architecture, the non-transparent RS supports the relaying when the transmission between BS and MS is decided to use two-hop relaying. The relay station architecture which is added to the new standard, IEEE 802.16j, gives more challenges to the scheduling issue. Because of the difference in RS’s functionality the scheduling scheme could be classified into two modes, that is centralized scheduling and distributed scheduling. In centralized scheduling mode, the BS needs to handle all of the scheduling information in a cell and decide the order how system serves each MS. However, the BS will share the scheduling overhead with RSs in distributed scheduling mode. For the network system with non-transparent RS, the system could schedule in both centralized and distributed mode since the non-transparent RS is capable of dealing with scheduling information. On the other hand, the network system with transparent RS should only schedules in centralized mode. Nevertheless, no matter which scheduling mode and RS is used, the BS is always has the authority to manage all of the MSs in a cell [1, 7].

The scheduling issue in wireless network system is close to resource management and the main purpose of scheduling is for QoS guaranteed. However, the IEEE 802.16j doesn’t make any provision about the scheduling mechanism hence the issue leaves discussion for later researches [2, 3, 5].

III. ADAPTIVE NONTRANSPARENT-BASED DISTRIBUTED SCHEDULING (ANDS)

![RTDS architecture](image)

Figure 1. RTDS architecture

Our research proposes an adaptive real-time-service-based distributed scheduling scheme (RTDS) for IEEE 802.16j network system. RTDS will guarantee the QoS of the users and enhance the efficiency of the network system by assigning and adjusting bandwidth dynamically. The main goal of RTDS is to satisfy all connections’ requests as far as possible.

Figure 1 is the scheduling architecture of RTDS. In this paper, we divide chapter three into three parts to make a detail description of RTDS. Part (A): Priority Assignment. Part (B): Resource Allocation. Part (C): Preserved Bandwidth Adjustment.

A. Priority Assignment

<table>
<thead>
<tr>
<th>Service Type</th>
<th>Requirement Calculation</th>
<th>Bandwidth Allocation</th>
<th>Bandwidth Status Report</th>
<th>Preserved Bandwidth Re-Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGS (Unsolicited Grant Service)</td>
<td>Maximum Sustained Traffic Rate</td>
<td>Minimum Reserved Traffic Rate</td>
<td>Maximum Latency</td>
<td>Unsolicited Grant Interval, Tolerated Jitter</td>
</tr>
<tr>
<td>erTPS (Extended Real-Time Polling Service)</td>
<td>Maximum Sustained Traffic Rate</td>
<td>Minimum Reserved Traffic Rate</td>
<td>Maximum Latency</td>
<td>Unsolicited Grant Interval, Tolerated Jitter</td>
</tr>
<tr>
<td>rtPS (Real-Time Polling Service)</td>
<td>Maximum Sustained Traffic Rate</td>
<td>Minimum Reserved Traffic Rate</td>
<td>Maximum Latency</td>
<td></td>
</tr>
<tr>
<td>nrtPS (Non-Real-Time Polling Service)</td>
<td>Maximum Sustained Traffic Rate</td>
<td>Minimum Reserved Traffic Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BE (Best Effort)</td>
<td>Maximum Sustained Traffic Rate</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1. IEEE 802.16J DEFINED SERVICE TYPE

RTDS quantifies five classes of traffic type which are specified in the IEEE 802.16j to give an initial priority value for scheduling order. Table 1 shows the initial priority value for five classes of traffic type. These values match our goal that is enhancing the QoS of real time service connections. There are five traffic types specified in the IEEE 802.16j, which are UGS, erTPS, rtPS, nrtPS and BE, and the initial priority value are 5, 4, 3, 2 and 1 respectively. Since UGS, erTPS and rtPS are real time services, the initial priority values of UGS, erTPS and rtPS are higher than nrtPS and BE.

After the initial priority value of each connection is set, RTDS adjust the initial priority value to the different condition and request of each connection. The section of priority adjustment is divided into three phases which are Priority Promotion by Packet Delay Tolerance Rating, Priority Promotion by Packet Critical Rating and Priority Diminution, respectively.

Phase one of priority adjustment is Priority Promotion by Packet Delay Tolerance Rating. The main idea of this phase is to promote the priority value of a connection by its case of packet delay.

![Packet delay time diagram](image)

Figure 2. Packet delay time diagram

Figure 2 depicts the delay time of a packet. The packet arrival time, denoted by $T^a$, means the time when a packet arrives in an access station, such as BS and RS. The calculation of $T^w$ is defined as:

$$T^w_i(j) = T^x_j - T^a_i(j)$$

$T^w_i(j)$ is packet waiting time of the j’th packet of connection i.

$T^x_j$ is the current time of the system.
is packet arrival time of the j’th packet of connection i.

Our research defines the packet remaining halt time, denoted by $T^r$, to indicate the legal halting time for a packet. Let the maximum latency of packet minuses its packet waiting time to obtain $T^h$, which means the packet will be out of date packet after the time $T^h$ and dropped. The calculation of $T^h$ is defined as:

$$T^h(j) = \omega_i - T^w(j)$$

(2)

$T^r(j)$: packet remaining halt time of the j’th packet of connection i.

$\omega_i$: maximum latency of service type i.

$T^w(j)$: number of packets which are waiting for the maximum latency of service type i.

$N_i$: number of packets from connection i.

In general speaking, the connections with a lower value of $R^c_i$ means most of the packets will be out of date soon and should be scheduled first. On the other hand, the connections with a higher value of $R^c_i$ means the connections could suffer more waiting time. Thus, our research defines a parameter for priority promotion, denoted by $P_i^c$. The function of normalization is defined as:

$$P_i^c = 1 - (R_i^d / R_{i,MAX}^d)$$

(3)

$N_i$: number of packets from connection i.

To obtain the final priority value, RTDS sums up all priority parameters in one total priority value. The final priority value of each connection i is defined as:

$$P_i = P_i^c + P_i^p + P_i^f - P_i^e$$

(9)

$P_i^c$: priority value to promote. The function is defined as:

$$P_i^c = P_i^c * \gamma$$

(8)

$\gamma$: fairness parameter, from 0% to 100%.

$P_i^p$: priority value to promote the priority value of the connections which get priority promoted twice by critical packet condition. The main idea of this phase is to offer fairness among the connections. The diminution function is defined as:

$$P_i^p = P_i^p + \beta_i * R_{i,MIN}^c * T^r$$

(10)

$P_i^f$: priority value to promote. The function is defined as:

$$P_i^f = P_i^f + \beta_i * R_{i,MIN}^c * T^r$$

(11)

$\beta_i$: size of a packet.

$R_{i,MIN}^c$: maximum sustained traffic rate of connection i.

$R_{i,MAX}^c$: minimum reserved traffic rate of connection i.

$T^r$: duration of a transmission frame.

If formula (5) stands, it means the packet will be out of date if it is not transmitted at the scheduling frame. And under this circumstance, the QoS of users will be decreased. In our research, we define these packets as critical packet. In general speaking, a connection with more critical packets will generate much more unsatisfied QoS and should be served earlier. We define critical packet rating to represent the critical degree of a connection, and as follows:

$$R^c_i = \sum_{j=1}^{N_i} T^w(j)$$

(6)

$N_i$: number of critical packets in connection i.

$R^c_i$: critical packet rating of connection i.

And we also normalize the critical packet rating to get the priority value to promote. The function is defined as:

$$P_i^c = R_i^c / R_{i,MAX}^c$$

(7)

$P_i^e$: priority value to promote. The function is defined as:

$$P_i^e = P_i^e + \beta_i * R_{i,MIN}^c * T^r$$

(12)

$P_i^f$: priority value to promote. The function is defined as:

$$P_i^f = P_i^f + \beta_i * R_{i,MIN}^c * T^r$$

(13)

$\beta_i$: size of a packet.

$R_{i,MIN}^c$: maximum sustained traffic rate of connection i.

$R_{i,MAX}^c$: minimum reserved traffic rate of connection i.

$T^r$: duration of a transmission frame.

The lower bound and upper bound of bandwidth request for all connections and the total bandwidth of the system will divide resource allocation into three conditions: (1) System bandwidth is greater than the upper bound. (2) System bandwidth is between the upper bound and the lower bound. (3) System bandwidth is smaller than the lower bound.
In situation (1), RTDS will allocate the upper bound of bandwidth to each connection since the system has plenty of resource.

In situation (2), RTDS allocates the lower bound of bandwidth to each connection to meet the basic requirement of QoS. After satisfying the basic requirement of QoS, RTDS will look for the connections which have got priority promoted by critical packet rating. For these connections, RTDS will allocate the upper bound of request by their priority value. Last, RTDS will allocate the remainder to each connection by the ratio of bandwidth request.

In situation (3), RTDS must make decision to sacrifice some connections’ QoS provision since the total system bandwidth couldn’t even afford the lower bound of bandwidth request of every connection. RTDS will allocate the lower bound of bandwidth request to each connection by its priority value until there is no longer any system bandwidth left.

C. Preserved Bandwidth Adjustment

The feature of distributed scheduling is that the BS doesn’t need to realize how RS making its own scheduling. This feature will bring lesser overhead to the BS but also bring some disadvantages such as the non-real time service handled by the BS may obtain more resource than the real time service handled by the RS since the BS only have the announcement of total bandwidth request for each RS. Therefore, RTDS builds a preserved bandwidth mechanism to protect the QoS of connections which are scheduled by RS.

RTDS defines the bandwidth status report which is send by the RS in order to notify the BS of the lower bound of bandwidth request for real time service connections. The BS collects all of the bandwidth status reports and then adjusts the preserved bandwidth of each RS by its ratio of bandwidth request.

IV. SIMULATION AND RESULTS ANALYSIS

The simulation model refers to the IEEE 802.16j standard and the simulation environment is in a cell. A BS is in the center of the cell and its transmission range is 8 km. 6 RSs are spreading among the cell and the transmission range of a RS is 3 km. The BS and RSs are in line of sight and at the distance of 5 km. The simulation model is depicted as figure 3.

We assume the buffer of the BS and the RS is unlimited and the packet size of UGS, eRTS, nRTS, nrtPS and BE are 160 bytes, 160 bytes, 240 bytes, 120 bytes and 120 bytes, respectively. The time duration of a transmission frame is 5 ms. The generation model of calls which are made by the MSs refers to Poisson Distribution function in order to meet the actual environment. The simulation time is 30000 frames, which is 150 seconds.

We use a simple call admission control (CAC) to decide whether a connection should be accepted by the system. If the system could not afford the minimum bandwidth request for a new connection, the system will reject the request for connecting.

Figure 4 shows the average delay time of real time service connections. RTDS suffers about 1.9 ms delay time in average. However, DFPQ and PQ suffer more on average delay time, about additionally 0.6 and 0.8 ms respectively.

Figure 5 shows the average delay time growth by different system load. PQ and DFPQ will obviously increase the delay time when the system load is at the percentage of 50%. However, RTDS increases the delay time until the system load is at the percentage of 70%

Figure 6 shows the packet drop rate growth according to different system load. PQ will dramatically increase the drop rate when the system load is at the percentage of 40%. DFPQ will dramatically increase the drop rate when the system load is at the percentage of 60%. RTDS also increases the drop rate when the system load is at the percentage of 70%. RTDS could afford much more
system load against to the drop rate and the growth curve is flatter than others.

![Diagram](image)

Figure 6. The packet loss arte growth by different system load

V. CONCLUSION AND FUTURE WORK

Our research proposed a distributed scheduling scheme, RTDS, for IEEE 802.16j networks. RTDS will allocate bandwidth dynamically for different types of connections to meet each connection’s QoS requirement. RTDS primary guarantees the QoS of real time service connections and additionally provides the fairness to every connection. The simulation results show that RTDS will suffer less packet delay time and packet loss rate than other representative researches.

In future works, we are going to take uplink scheduling into consideration to obtain a more completely and precisely scheduling scheme. Furthermore, we will provide more fairness to the non-real time service connections to enhance the overall QoS performance.

REFERENCES


Kuo-Feng Huang received the B.S., M.S., and Ph.D. degrees in Computer Science and Information Engineering from the Tamkang University in 2003, 2007, and 2011, respectively. From September 1st, 2011, he was an assistant professor in department of computer and communication engineering of Taipei College of Maritime Technology. His current research interests are Wireless Communication, Mobile Communication, Wireless Sensor Networks and Embedded System.

Shih-Jung Wu was born in Taipei, Taiwan (R.O.C.), on Oct. 25, 1976. He received his B.C. degree from Department of Business Administration, Yuan Ze University, Taiwan (R.O.C.) in 1998. He received M.S. degree from Department of Computer Science and Information Engineering, Tamkang University, Taiwan (R.O.C.) in 2001. And he received his Ph.D. degree in the Department of Computer Science and Information Engineering, Tamkang University, Taiwan (R.O.C.) in 2006. Presently, he is working at Department of Innovative Information and Technology, Tamkang University in Taiwan (R.O.C.). His major research interests in high speed communications, mobility, QoS guarantees, and parallel algorithms.