

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Traffic Control for Composite Wireless Access Route of IEEE802.11/16 Links

Yasuhisa Takizawa
Kansai University
Japan

1. Introduction

The expansion and diversification of wireless communications are proceeding rapidly with the diffusion of cellular phones, WiFi and WiMAX. However, concern is increasing that the growth of wireless systems will exhaust finite wireless resources. Cognitive radio technology (Mitorall & Maguire, 1999; Mitorall, 1999; Harada, 2005), which has been proposed as a solution to this problem, aims to optimize the utilization of diverse wireless resources. Furthermore, AIPN (All-IP Network) (3GPP, 2005) and NGN (Next Generation Network) (ITU, 2006) investigate the network architecture that accommodates diverse communication media. Accordingly, we expect that in the near future, wireless access networks will be composed of diverse wireless medias.

To exploit wireless media diversity in expected access networks, some bandwidth-aggregation methods in wireless media have recently been proposed. Bandwidth-aggregation combines diverse communication links in parallel and suitably distributes packets to communication links. The works (Phatak & Goff, 2002; Snoeren, 1999; Shrama et al., 2007) aggregate wireless links in IP to improve IP throughput. The work (Chebrou & Rao, 2006) also aggregates wireless links in IP to decrease IP delay based on wireless media that provide a bandwidth guarantee. The works (Hsieh et al., 2004; Zhang et al., 2004) aggregate communication links in a transport layer to improve TCP throughput. Meanwhile, wireless access networks process traffic of diverse application, and the traffic is classified by the following two types of application traffic:

- Traffic of throughput-oriented application such as FTP and Web on TCP.
- Traffic of delay-oriented application such as VoIP and Video Conference on UDP.

Therefore, wireless access networks are required to provide high throughput and low delay by diverse applications. The above works do not consider delay except for the work (Chebrou & Rao, 2006), and the work (Chebrou & Rao, 2006) does not consider IEEE802.11 that no bandwidth guarantee is provided. Furthermore, the works (Phatak & Goff, 2002; Snoeren, 1999; Shrama et al., 2007; Chebrou & Rao, 2006) improve IP performance, but can not provide effective improvement of application performance because they do not consider out-of-order packets which occur by the packet distribution to multiple links. The works (Hsieh et al., 2004; Zhang et al., 2004) consider the out-of-order packet, and can improve the performance of TCP application, but can not improve that of UDP application such as VoIP and Video Conference.

| | 802.11a/b | 802.16 |
|---------------------|----------------------------|--------------------------|
| Transmission Rate | 54Mbps/11a, 11Mbps/11b | 75Mbps |
| Coverage | 50m/11a, 100m/11b | 1000m |
| Access Control | CSMA/CA (Decentralized) | TDD/FDD (Centralized) |
| Bandwidth Guarantee | No | Yes |

Table 1. Performance of wireless systems.

In this chapter, assuming the expected wireless access network to be composed of IEEE802.11, which is a popular wireless system, and IEEE802.16, which is expected to spread, a IP packet distribution on the access route, which combines IEEE802.11-link and IEEE802.16-link in parallel, is proposed to improve the application performance. The proposed packet distribution increases IP throughput and decreases IP delay. Furthermore, it reduces out-of-order packets and provides high throughput and low delay to both UDP applications and TCP applications simultaneously.

Our works(Takizawa et al., 2008; Takizawa, 2008) have proposed the packet distribution for combining IEEE802.11/16 wireless upload links. We expand the above packet distribution to reduce out-of-order packets and to apply download traffic, and show its essential characteristics of packet distribution for composite wireless access route of IEEE802.11/16-links (call M-route) , then propose a packet distribution method for M-route. Furthermore, we evaluate the method’s performance by multiple application traffic on both UDP and TCP in a wireless access network composed of 802.11a, 802.11b and 802.16, which have the different characteristic from each other (see Table 1).

The configuration of wireless access networks by wireless media diversity is assumed as follows (see Fig. 1).

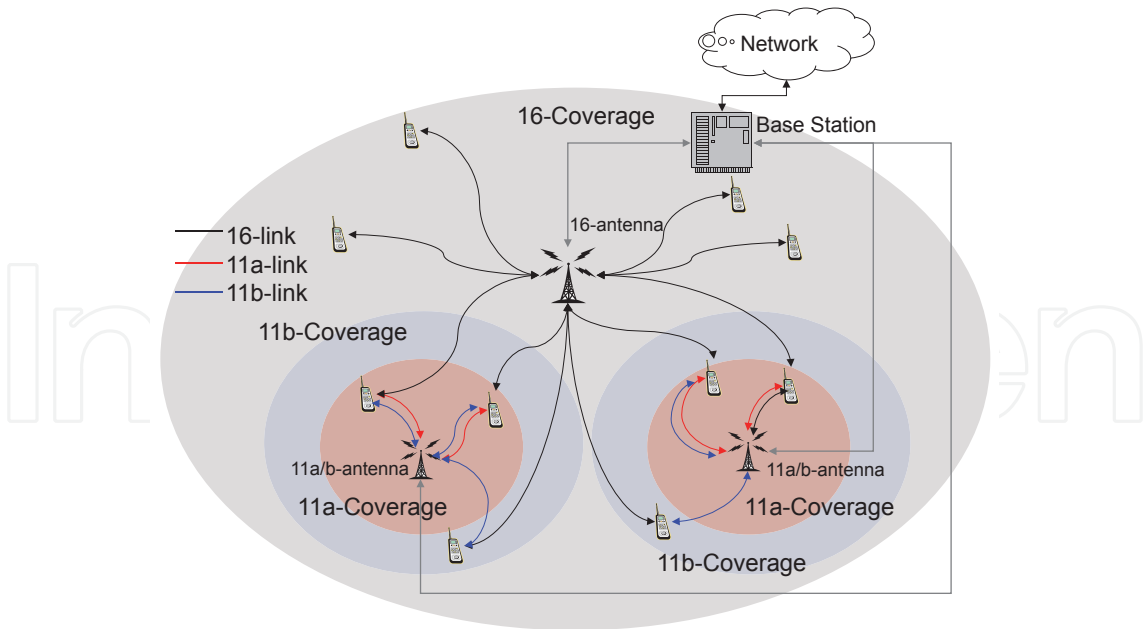


Fig. 1. Assumed wireless access network.

- Base station provides an access point function of IEEE802.11a/b-wireless systems and a base station function of 16-wireless system, and accommodates IEEE802.11a/b-antennas and an IEEE802.16-antenna by wired connecting. It also provides the function of gateway.

- Each terminal is equipped with IEEE802.11a/b-interfaces and IEEE802.16-interface, and can communicate with base station by using each interface.
- IEEE802.11a/b-antennas and terminals are randomly deployed within coverage of IEEE802.16-antenna.
- The access network is IP network.

2. Characteristics of IEEE802.11 link for packet distribution

In this section, based on Media Access Control (MAC) of IEEE802.11 DCF, the characteristics of IEEE802.11 wireless link (11-link) for packet distribution is analyzed.

2.1 IEEE802.11 link cost

Based on queuing theory (Gross & Harris, 1985), a link load is shown as the number of packets in a link, including waiting packets in the queue and the currently processed packet. $d_{(i,k)}$, which is cost of link k between a terminal i and a base station, is defined as the link load, and it is expressed using Little's theorem (Little, 1961) as follows.

$$d_{(i,k)} = F_{(i,k)} \cdot T_{(i,k)} \quad (1)$$

where $F_{(i,k)}$ is the packet arrival rate of link k in terminal i and $T_{(i,k)}$ is the average delay of link k in terminal i . Delay is the time from packet arrival at the terminal to completion of packet transmission, therefore the delay is composed of a waiting delay in queue and an air time. The air time is composed of MAC delay and transmission delay, which take the MAC retransmission into consideration.

Based on Eq. (1), $T_{(i,k)}$ decreases if $d_{(i,k)}$ decreases on constant $F_{(i,k)}$ and on maximum of $d_{(i,k)}$, that is, link capacity, $F_{(i,k)}$ can increase if $T_{(i,k)}$ decreases. $F_{(i,k)}$ corresponds to a throughput on condition that no packet loses. Therefore, when $d_{(i,k)}$ decreases, a throughput increases and a delay decreases on a link.

The dependence of the link cost on the packet arrival rate, which corresponds to the number of distributed packets in unit time to a link, is shown. Based on Eq. (1), the link cost depends on the average delay. The average delay is composed of the waiting delay in queue and the packet service time. Therefore, in regard with 11-link, the dependence of the above elements on the packet arrival rate are shown, and in summarizing them, the dependence of the link cost on the packet arrival rate is shown.

2.1.1 Dependence of packet service time on packet arrival rate

In (Bianchi, 2000), throughput analysis of IEEE802.11 DCF is shown, and in (Carvalho & Garcia, 2003), the packet service time analysis of that is shown based on (Bianchi, 2000). According to these, the dependence of the average packet service time on the packet arrival rate is shown.

DCF adopts an exponential backoff scheme, and employs a discrete-time backoff timer. The timer immediately following a Distributed InterFrame Space (DIFS) starts, and a terminal, which is a terminal or a base station, is allowed to transmit only at the beginning of each Slot Time. The Slot Time size σ is set equal to the time needed at any terminal to detect the transmission of a packet from any other terminal. At each packet transmission, the backoff timer is randomly chosen in the range $(0, CW - 1)$. CW is called Contention Window, and depends on the number of transmissions failed for the packet. At the first transmission attempt, CW is set equal to CW_{min} called minimum contention window. After each failed transmission, CW is doubled, up to a maximum value $CW_{max} = 2^r CW_{min}$ (r is a maximum

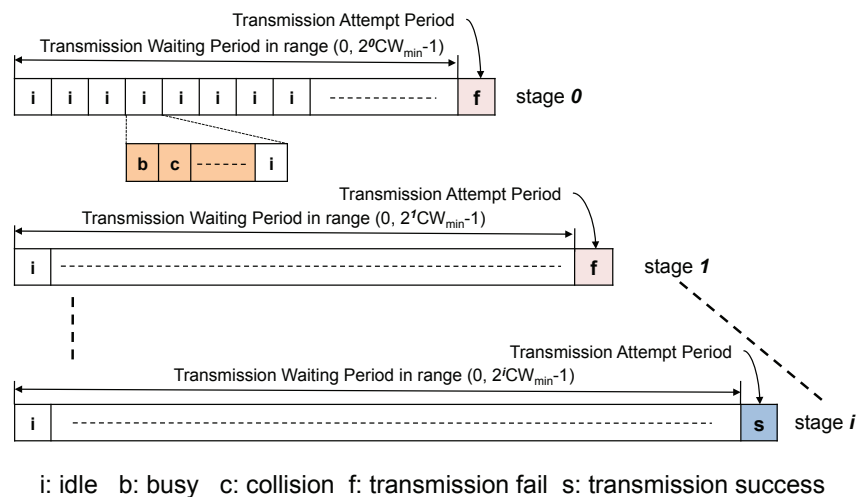


Fig. 2. Exponential binary backoff in IEEE802.11.

number of retransmissions). Each transmission attempt is referred to as a backoff stage. The packet service time is the sum of time for each backoff stage. Each backoff stage is composed of the transmission waiting period and the transmission attempt period (see Fig. 2). The backoff stage starts in the transmission waiting period, and the backoff timer is initialized to a random value in the range $(0, CW_i - 1)$ at the backoff stage i start. CW_i is the contention window size of the backoff stage i . In the period, the backoff timer is decremented only when the channel is idle, and it is frozen when the channel is busy. The duration of the period is the time until the backoff timer becomes zero from initial value. The transmission attempt period starts when the backoff timer reaches zero, and a packet transmission takes place. The duration of period is the time to transmit a packet. In the model of (Bianchi, 2000) and (Carvalho & Garcia, 2003), a fixed number of terminals is assumed, and the backoff stage is repeated until a packet transmission success using CW_i until stage r and using CW_r beyond stage r . The stage r is called maximum backoff stage. Furthermore, using the probability τ that a terminal transmits in a randomly chosen slot time, the following probabilities in an exponential backoff scheme are expressed.

$$\begin{aligned}
 p_{tr} &= 1 - (1 - \tau)^{n-1} \\
 p_{suc} &= \frac{(n-1)\tau(1-\tau)^{n-2}}{p_{tr}} \\
 p_i &= 1 - p_{tr} \\
 p_s &= p_{tr} \cdot p_{suc} \\
 p_c &= p_{tr}(1 - p_{suc}) \\
 q &= (1 - \tau)^{n-1}
 \end{aligned} \tag{2}$$

where n is the number of terminal in the channel coverage, p_{tr} is the probability that there is at least one transmission in the slot time of the transmission waiting period, p_{suc} is the probability that a transmission occurring on the channel is successful, p_i is the probability that the slot time is idle in the transmission waiting period, p_s is the probability that the channel is busy due to a packet transmission success in the transmission waiting period, p_c is the probability that the channel is busy due to a collision in the transmission waiting period, and q is the probability that a packet transmission success in the transmission attempt period. Let B be the average time which the transmission waiting period takes until a packet transmission succeeds, and let A be the average time which the transmission attempt period takes until a

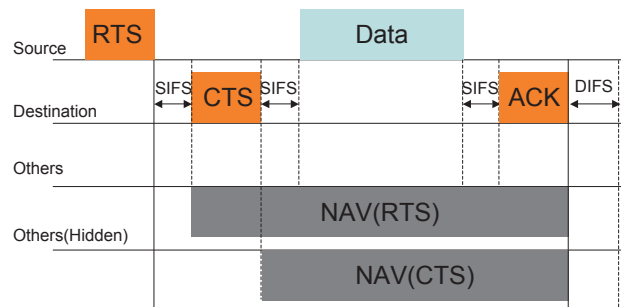


Fig. 3. RTS/CTS access control sequence in IEEE802.11.

packet transmission succeeds, B and A are derived from a binary exponential backoff scheme as follows (Carvalho & Garcia, 2003). (Note: In this section, "time" is the duration in slot time units σ of IEEE802.11)

$$B = \frac{t_b(\eta CW_{min} - 1)}{2q} \quad (3)$$

$$t_b = p_i t_i + p_s t_s + p_c t_c$$

$$\eta = \frac{q - 2^r (1 - q)^{r+1}}{1 - 2(1 - q)} \quad (4)$$

$$A = \frac{1 - q}{q} t_c + t_s \quad (5)$$

$$t_i = 1$$

$$t_s = RTS + SIFS + \delta + CTS + SIFS + \delta + H$$

$$+ PL + SIFS + \delta + ACK + DIFS + \delta \quad (6)$$

$$t_c = RTS + DIFS + \delta$$

where t_i is the time of idle (i.e., one backoff slot), t_s is the average time that the channel is sensed busy due to a packet transmission success, t_c is the average time that the channel is busy due to a collision in the channel, RTS , CTS and ACK are time that RTS , CTS and ACK frame is transmitted respectively, $SIFS$ and $DIFS$ are the interval time (see Fig.3), δ is the propagation delay, H is the time that a packet header is transmitted, and PL is the time the payload is transmitted. According to Eq. (2), $q = 1 - p_{tr}$, therefore, t_b/q expresses the average time that the backoff timer is decreased by one, and $(\eta CW_{min} - 1)/2$ expresses the average of sum of backoff timer in all stage. In Eq. (5), $(1 - q)/q$ expresses the average number of collision in the transmission attempt period.

Then, the average packet service time S is argued using the above analysis. S is shown as follows.

$$S = B + A \quad (7)$$

When the number of terminal is constant, the dependence of S on τ is shown using the first and second derivative of S at τ as follows.

$$\frac{dS}{d\tau} > 0 \quad \frac{d^2S}{d\tau^2} > 0 \quad (8)$$

Therefore, S is a convex monotonically increasing function of τ . Figure 4(a) illustrates the dependence of S on τ by using Eq. (7) in 11b MAC parameter, transmission rate 11Mbps, a

number of terminals 10~40, maximum backoff stage 5, and the payload size 1500 bytes, and it also shows the same characteristics.

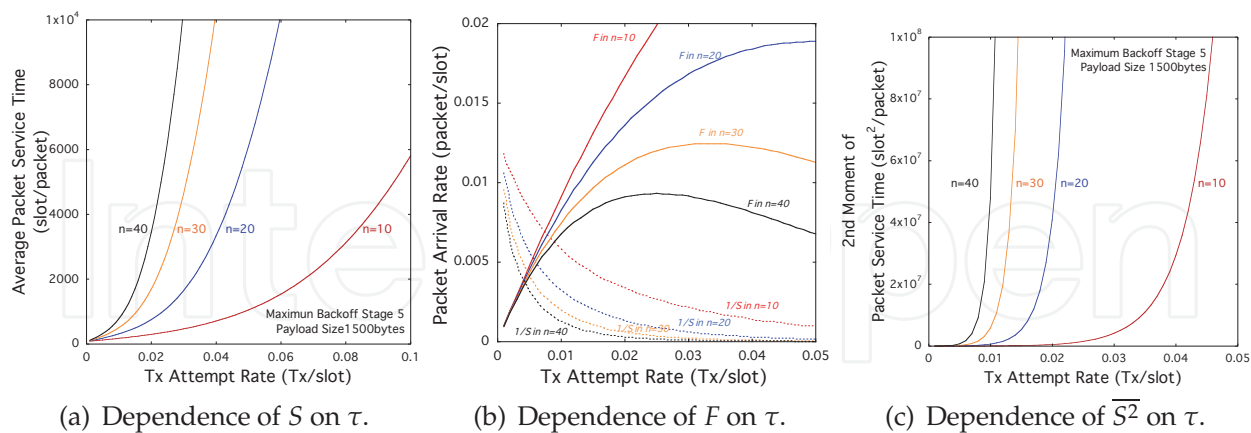


Fig. 4. Dependence of each element on τ .

In (Bianchi, 2000) and (Carvalho & Garcia, 2003), the transmission queue is assumed to be always non-empty, thus, the dependence of τ on the packet arrival rate F is not considered. Let F be the number of arrival packets at a link in a slot time, the dependence is argued. The average number of arrival packets in period S is FS , and the average number of transmission attempts on a successfully transmitted packet is $(1 - q)/q + 1$. Then, the average number of that a packet transmission attempts in period S is FS/q . Therefore, τ is shown as follows.

$$\tau = \frac{FS}{qS} = \frac{F}{q} \quad (9)$$

Figure 4(b), which illustrates the dependence of F on τ using Eq. (9) in the same parameter as Fig. 4(a). In Fig. 10, when $F < 1/S$ ($1/S$ is the packet service rate), that is, when the load does not exceed the link capacity, and when the number of terminal is constant, F for τ is concavely and monotonically increasing. Therefore, within link capacity, the dependence of F on τ is shown using the first and second derivative of F at τ as follows.

$$\frac{dF}{d\tau} > 0 \quad \frac{d^2F}{d\tau^2} < 0 \quad (10)$$

Furthermore, the first and second derivative of S on F is shown using Eqs.(8), (10) as follows.

$$\frac{dS}{dF} = \frac{dS}{d\tau} \frac{d\tau}{dF} = \frac{dS}{d\tau} \frac{1}{\left(\frac{dF}{d\tau}\right)} > 0 \quad (11)$$

$$\frac{d^2S}{dF^2} = \frac{d^2S}{d\tau^2} \left(\frac{d\tau}{dF}\right)^2 - \frac{dS}{d\tau} \frac{1}{\left(\frac{d^2F}{d\tau^2}\right)} > 0 \quad (12)$$

Therefore, within link capacity, S is a convex monotonically increasing function of F .

2.1.2 Dependence of waiting delay in queue on packet arrival rate

The dependence of W which is the waiting delay in queue on F is argued. N_Q , which is the number of waiting packets in queue, is $F \times W$ using Little's theorem. W is composed of the

packet service time for N_Q packets and R , which is the sum of the residual service time in each packet arrival. Consequently, W is shown as follows.

$$W = N_Q \cdot S + R = F \cdot W \cdot S + R \quad (13)$$

Each residual service time in a packet arrival is $\overline{S^2}/2S$ (Bertsekas & Gallager, 1992), where $\overline{S^2}$ is the second moment of S . The average number of packet arrivals in S is FS ; accordingly, R is $F\overline{S^2}/2$. Applying the above relations to Eq. (13), W is given as

$$W = \frac{F\overline{S^2}}{2(1 - FS)} \quad (14)$$

Let $V[S]$ be the variance of S , and it is shown as follows (Carvalho & Garcia, 2003)

$$V[S] = \left[\frac{t_b(CW_{min}\gamma - 1)}{2} + t_c \right]^2 \frac{1 - q}{q^2} \quad (15)$$

$$\gamma = \frac{[2q^2 - 4q + 1 - r(-1 + 2q)q][2(1 - q)]^r + 2q^2}{(-1 + 2q)^2}$$

Using Eq. (15), $\overline{S^2}$ is shown as follows.

$$\overline{S^2} = S + V(S) \quad (16)$$

Furthermore, using Eq. (16), the first and second derivatives of $\overline{S^2}$ at τ are shown, respectively, as follows.

$$\frac{d\overline{S^2}}{d\tau} > 0 \quad \frac{d^2\overline{S^2}}{d\tau^2} > 0 \quad (17)$$

Figure 4(c) illustrates the dependence of $\overline{S^2}$ on τ using Eq. (16) in the same parameter as Fig. 4(a), and it also shows the same characteristics. Furthermore, applying Eq. (10) to Eq. (17), the first and second derivatives of $\overline{S^2}$ at F are shown, respectively, as follows.

$$\frac{d\overline{S^2}}{dF} > 0 \quad \frac{d^2\overline{S^2}}{dF^2} > 0 \quad (18)$$

Using Eqs. (14) (18), the first and second derivatives of W at F are shown, respectively, on the condition of $FS < 1$, as follows.

$$\frac{dW}{dF} > 0 \quad \frac{d^2W}{dF^2} > 0 \quad (19)$$

$FS < 1$, that is, $F < 1/S$ expresses the condition that a link load is with a link capacity. Therefore, within a link capacity, W is also a convex monotonic increasing function of F .

2.1.3 Dependence of 11-link cost on packet arrival rate

Finally, the dependence of the 11-link cost on the packet arrival is argued. The average delay T is also a convex monotonic increasing function of F because of $T = W + S$. Applying the dependence of T on F to Eq. (1), the first and second derivatives of a 11-link cost d at F are as follows.

$$\frac{dd}{dF} > 0 \quad \frac{d^2d}{dF^2} > 0 \quad (20)$$

Consequently, a 11-link cost d is also a convex monotonic increasing function of F within a link capacity and in a fixed number of terminals.

2.2 Cost of M-route compositing multiple 11-links for upload traffic

On communications using a M-route which aggregates multiple 11-links from terminal to a base station, the cost of M-route for upload traffic is the sum of cost of each 11-uplink composing M-route because the number of packets in a M-route is the sum of the number of packets in each link composing M-route. Therefore, m_i which is the cost of M-route for upload traffic in terminal i is shown as follows (see Fig. 5(a)).

$$m_i = \sum_{x \in U_i} d_{(i,x)} \quad (21)$$

U_i is the set of an uplink which is provided by a 11-wireless interface equipped with terminal i . Here, in steady packet arrival rate, the packet distribution from an 11-uplink k to an 11-uplink j in M-route of terminal i , is argued. In this case, the packet distribution to the other 11-uplinks is constant, thus the dependence of $F_{(i,j)}$ on $F_{(i,k)}$ is shown as follows.

$$\frac{dF_{(i,j)}}{dF_{(i,k)}} = -1 \quad \frac{d^2F_{(i,j)}}{d(F_{(i,k)})^2} = 0 \quad (22)$$

Using Eqs.(20) and (22), the first and second derivatives of $d_{(i,j)}$ at $F_{(i,k)}$ are shown as follows.

$$\begin{aligned} \frac{dd_{(i,j)}}{dF_{(i,k)}} &= \frac{dd_{(i,j)}}{dF_{(i,j)}} \frac{dF_{(i,j)}}{dF_{(i,k)}} = -\frac{dd_{(i,j)}}{dF_{(i,j)}} < 0 \\ \frac{d^2d_{(i,j)}}{d(F_{(i,k)})^2} &= \frac{d^2d_{(i,j)}}{d(F_{(i,j)})^2} \left(\frac{dF_{(i,j)}}{dF_{(i,k)}} \right)^2 + \frac{dd_{(i,j)}}{dF_{(i,j)}} \frac{d^2F_{(i,j)}}{d(F_{(i,k)})^2} > 0 \end{aligned} \quad (23)$$

Consequently, $d_{(i,j)}$ is a convex monotonically decreasing function of $F_{(i,k)}$. According to Eq.(21), m_i is the sum of $d_{(i,k)}$, which is a convex monotonically increasing function of $F_{(i,k)}$, and $d_{(i,j)}$, which is a convex monotonically decreasing function of $F_{(i,k)}$, and the uplink cost of the others, which are constant for $F_{(i,k)}$. Therefore, m_i is a convex function of $F_{(i,k)}$ (see Fig.5(b)), and m_i has an optimal solution for $F_{(i,k)}$.

Because m_i is a convex function of $F_{(i,k)}$, the optimal solution can be searched by the packet distribution which aim to descend the gradient in the convex function. When packets are distributed from a 11-uplink k to 11-uplink j in M-route, the condition of the gradient descent on M-route cost is shown as follows using Eq.(22).

$$\frac{dm_i}{dF_{(i,k)}} = \frac{dd_{(i,k)}}{dF_{(i,k)}} - \frac{dd_{(i,j)}}{dF_{(i,j)}} > 0 \quad (24)$$

Applying Eq.(1) to Eq.(24), and transforming Eq.(24) into difference equation, thus the first derivative of m_i at $F_{(i,k)}$ is shown as follows.

$$\frac{dm_i}{dF_{(i,k)}} = \lim_{\Delta F_{(i,k)} \rightarrow 0} \left(\left(T_{(i,k)} + F_{(i,k)} \frac{\Delta T_{(i,k)}}{\Delta F_{(i,k)}} \right) - \left(T_{(i,j)} + F_{(i,j)} \frac{\Delta T_{(i,j)}}{\Delta F_{(i,j)}} \right) \right) > 0 \quad (25)$$

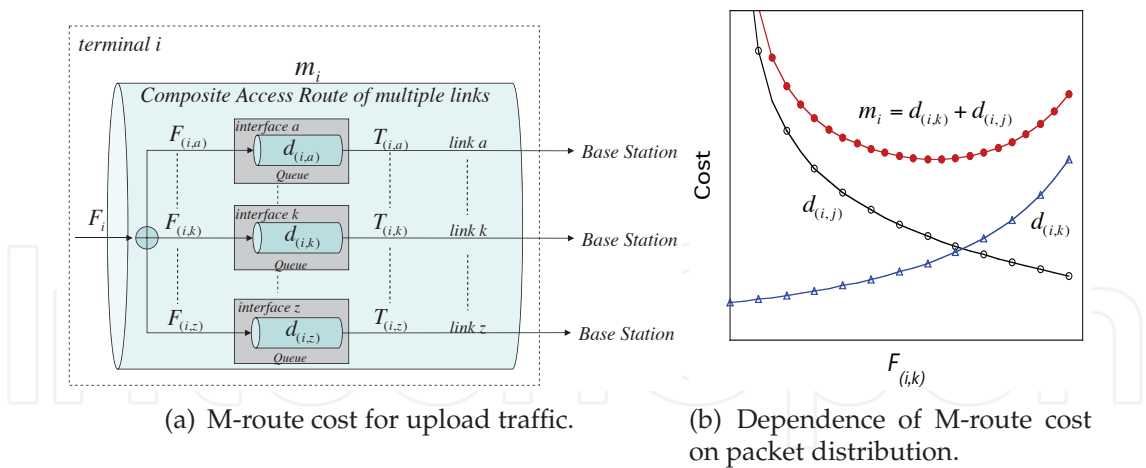


Fig. 5. M-route for upload traffic.

Furthermore, applying finite difference approximation to Eq.(25), the following is derived.

$$\frac{dm_i(n)}{dF_{(i,k)}} \approx T_{(i,k)}(n+1) - T_{(i,j)}(n+1) > 0 \tag{26}$$

Where, $m_i(n)$ is M-route cost of terminal i in packet distribution of n time and $T_{(x,y)}(n+1)$ is average delay of 11-link y in terminal x in packet distribution of $n+1$ time. Consequently, when the packet distribution meets Eq.(26) which means the average delay of source 11-uplink on packet distribution becomes larger than that of destination 11-uplink on packet distribution, the M-route cost for upload traffic decreases and approaches the optimal solution. Such packet distribution is repeated with the decrease in the amount of the distributing packets ($\Delta F_{(i,k)} \rightarrow 0$), and finally the average delay of source 11-uplink becomes equal to that of destination 11-uplink, the M-route cost for upload traffic reaches its optimal solution.

Furthermore, the search for the optimal solution of M-route cost has the additional effectiveness which decreases the arrival of out-of-order packets because of the equalization of the delay of source 11-link and destination 11-link.

2.3 Cost of M-route compositing multiple 11-Links for download traffic

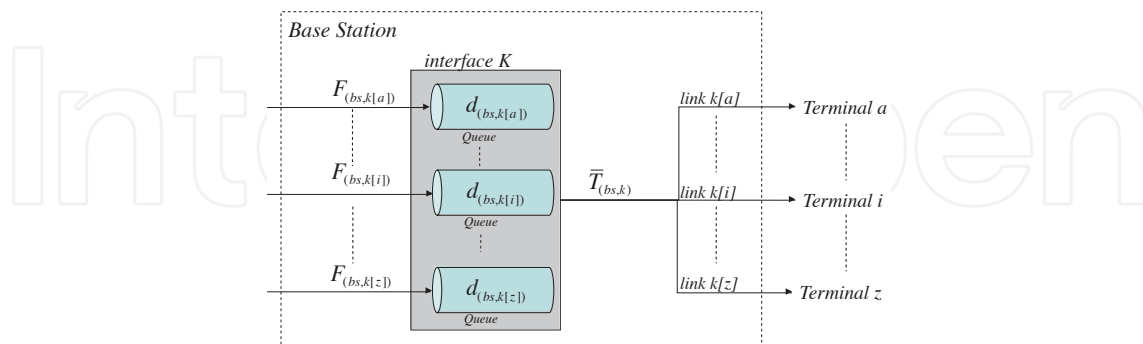
A base station associates its 11-interface with multiple terminals in its coverage. Thus, its interface is composed of multiple 11-downlinks according to multiple terminals in its 11-coverage, that is, its topology is point-to-multipoint. In this subsection, the cost of M-route for download traffic (i.e. in a base station) in steady packet arrival rate is argued. In queueing theory, a link has a queue of packets to be transmitted, and has an independent server on other links within the same interface. However, an 11-downlink is different from a link reserved the resource such as WiMAX (TDD or FDD) link and CDMA link, and an 11-downlink shares the resource of interface among other downlinks within the same interface. Conceptually, we can also view an 11-downlink within an interface as follows.

- Each 11-downlink has a queue which is independent on the other downlinks.
- Each 11-downlink has a common server as an interface among the other downlinks.

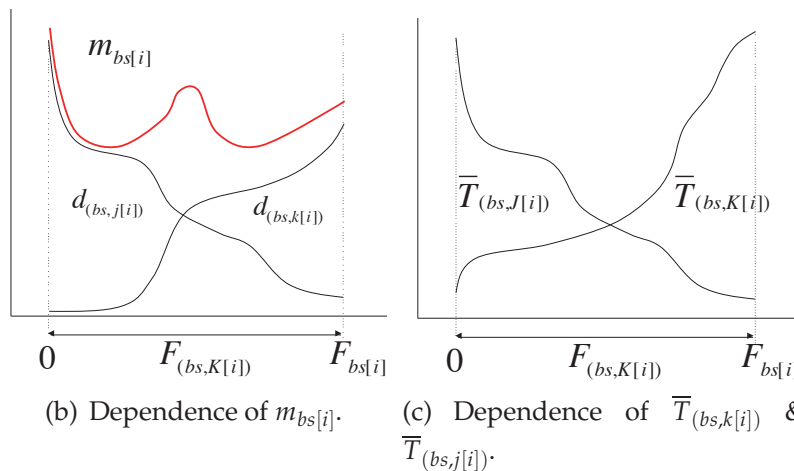
That is, in 11-downlink $k[i]$ to terminal i , which is provide by interface k of base station, $F_{(bs,k[i])}$ which is packet arrival rate of link $k[i]$ in base station, is independent on the others,

and $T_{(bs,k[i])}$ which is average delay of link $k[i]$ in base station, is common among the others. Therefore, $d_{(bs,k[i])}$ which is cost of 11-link $k[i]$ in base station, is shown as follows (see Fig. ??).

$$d_{(bs,k[i])} = F_{(bs,k[i])} \cdot \bar{T}_{(bs,k[i])} \quad (27)$$



(a) Downlink cost associated by 11-interface.



(b) Dependence of $m_{bs[i]}$. (c) Dependence of $\bar{T}_{(bs,k[i])}$ & $\bar{T}_{(bs,j[i])}$.

Fig. 6. Downlink and M-route for download traffic.

where $\bar{T}_{(bs,k[i])}$ is the average delay of 11-interface k providing downlink $k[i]$ in base station. That is, $\bar{T}_{(bs,k[i])}$ is the average delay based on all the packets which are distributed to 11-interface k .

To argue the dependence of $d_{(bs,k[i])}$ on $F_{(bs,k[i])}$, the first derivative of $d_{(bs,k[i])}$ at $F_{(bs,k[i])}$ is shown. Using Eq.(27), it is shown as follows.

$$\frac{dd_{(bs,k[i])}}{dF_{(bs,k[i])}} = \bar{T}_{(bs,k[i])} + F_{(bs,k[i])} \frac{d\bar{T}_{(bs,k[i])}}{dF_{(bs,k[i])}} \quad (28)$$

It is difficult to derive $\frac{d\bar{T}_{(bs,k[i])}}{dF_{(bs,k[i])}}$, which is the dependence of $\bar{T}_{(bs,k[i])}$ on $F_{(bs,k[i])}$, because $\bar{T}_{(bs,k[i])}$ is dependent on not only $F_{(bs,k[i])}$ but also the packet distribution of the other downlinks provided by 11-interface k . To simplify this difficulty, the following condition is assumed.

$$\frac{d\bar{T}_{(bs,k[i])}}{dF_{(bs,k[i])}} > 0 \quad (29)$$

According to the condition Eq.(29), its is $\frac{dd_{(bs,k[i])}}{dF_{(bs,k[i])}} > 0$, then $d_{(bs,k[i])}$ is a monotonically increasing function of $F_{(bs,k[i])}$.

In the condition, the packet distribution from a 11-downlink $k[i]$ to a 11-downlink $j[i]$, is argued. These 11-downlinks are contained in the M-route which aggregates 11-downlinks to terminal i , and are respectively provided by different 11-interface (11-interface k and j). The same as the packet distribution of M-route for upload traffic, the packet distribution to the other 11-downlinks to terminal i , which is respectively provided by different 11-interface except for 11-interface k and j , is constant, thus the dependence of $F_{(bs,j[i])}$ on $F_{(bs,k[i])}$ is shown as follows.

$$\begin{aligned} \frac{dF_{(bs,j[i])}}{dF_{(bs,k[i])}} &= -1 \\ \frac{d^2F_{(bs,j[i])}}{d(F_{(bs,k[i])})^2} &= 0 \end{aligned} \quad (30)$$

Therefore, the first derivative of $d_{(bs,j[i])}$ at $F_{(bs,k[i])}$ in the condition Eq.(29) is shown as follows.

$$\frac{dd_{(bs,j[i])}}{dF_{(bs,k[i])}} = -\frac{dd_{(bs,j[i])}}{dF_{(bs,j[i])}} < 0 \quad (31)$$

Consequently, in the condition Eq.(29), $d_{(bs,j[i])}$ is a monotonically decreasing function of $F_{(bs,k[i])}$. Because Eq.(21) can be applied to M-route for download traffic, $m_{bs[i]}$ which is the cost of M-route to terminal i is the sum of $d_{(bs,k[i])}$, which is a monotonically increasing function of $F_{(bs,k[i])}$, and $d_{(bs,j[i])}$, which is a monotonically decreasing function of $F_{(bs,k[i])}$, and the 11-downlink cost of the others, which is constant for $F_{(bs,k[i])}$. Therefore, $m_{bs[i]}$ is a multi-optimization function of $F_{(bs,k[i])}$, and it has some local minimums for $F_{(bs,k[i])}$ (see Fig. 6(b)).

Here, argue the dependence of $m_{bs[i]}$ on $F_{(bs,k[i])}$. it is shown as follows using Eq.(30) and (31).

$$\frac{dm_{bs[i]}}{dF_{(bs,k[i])}} = \frac{dd_{(bs,k[i])}}{dF_{(bs,k[i])}} - \frac{dd_{(bs,j[i])}}{dF_{(bs,j[i])}} \quad (32)$$

Furthermore, Eq.(32) is transformed into difference equation, and is applied finite difference approximation based on Eq.(29), then the condition that the M-route cost for download traffic decreases is shown as follows.

$$\frac{dm_{bs[i]}(n)}{dF_{(bs,k[i])}} \approx \bar{T}_{(bs,k[i])}(n+1) - \bar{T}_{(bs,j[i])}(n+1) > 0 \quad (33)$$

Where, $m_{bs[i]}(n)$ is the cost of M-route to terminal i from base station in packet distribution of n time and $\bar{T}_{(bs,y[x])}(n+1)$ is average delay of interface y in packet distribution of $n+1$ time, and the interface y provides 11-downlink to terminal x .

Furthermore, based on Eqs. (29) and (30), the dependence of $T_{(bs,j[i])}$ on $F_{(bs,k[i])}$ is shown as follows.

$$\frac{d\overline{T}_{(bs,j[i])}}{dF_{(bs,k[i])}} < 0$$

(34)

That is, $\overline{T}_{(bs,j[i])}$ is a monotonically decreasing function of $F_{(bs,k[i])}$. Therefore, a cost of each link in M-route should be considered a monotonically increasing function of the packet arrival rate, and the cost of M-route is the sum of each link cost, is a multioptimization function of $F_{(bs,k[i])}$ (see Fig.6(b)). That is, $m_{bs[i]}$ has some local minimums for $F_{(bs,k[i])}$ and the packet distribution meeting Eq.(33) may not bring $m_{bs[i]}$ to the optimal solution.

On the other hand, $\overline{T}_{(bs,k[i])}$ and $\overline{T}_{(bs,j[i])}$ is respectively a monotonically increasing/decreasing function for $F_{(bs,k[i])}$, and then, in $0 \leq F_{(bs,k[i])} \leq F_{bs[i]}$, the number of solutions which makes $\overline{T}_{(bs,k[i])}$ equal to $\overline{T}_{(bs,j[i])}$ is 1 in the maximum (see Fig. 6(c)). Consequently, the packet distribution which meets Eqs.(29) and (33) is repeated, and finally it reaches $\overline{T}_{(bs,k[i])}(n + 1) - \overline{T}_{(bs,j[i])}(n + 1) = 0$, then the M-route cost $m_{bs[i]}$ reaches its optimal solution. Furthermore, the search for the optimal solution of M-route cost $m_{bs[i]}$ has the additional effectiveness which decreases the arrival of out-of-order packets because of the equalization the delay of source 11-link and destination 11-link.

3. Characteristics of IEEE802.16 link for packet distribution

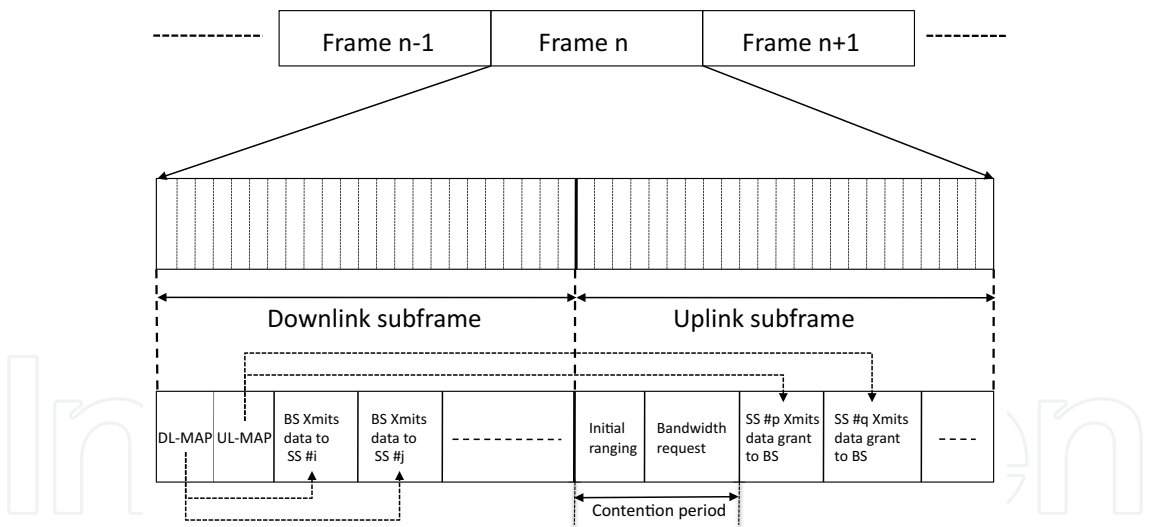


Fig. 7. IEEE802.16 MAC frame.

The performance of IEEE802.16 is actively analyzed. (Nakaya & Hossain, 2006) investigates the delay analysis based on queueing theory, but it does not consider MAC of IEEE802.16. (Cho et al., 2005; Lin et al., 2007; Iyengar et al., 2005; He et al., 2007; Ni et al., 2007) investigate the performance analysis based on MAC of IEEE802.16. Cho et al. (2005) analyzes the utilization and throughput and (Lin et al., 2007) analyzes the utilization for BW request based on polling. These analyses do not investigate the delay. On the other hand, (Iyengar et al., 2005; He et al., 2007; Ni et al., 2007) analyze the delay, but does not consider waiting time in queue. In this section, in regard with IEEE802.16 link (16-link), considering the waiting time in queue and MAC of IEEE802.16, the dependence of average delay on traffic is analyzed in

accordance with its four QoS classes. Furthermore, based on the analyzed dependence, the characteristics of 16-link for packet distribution is shown.

Figure.7 shows 16-frame in TDD. The frame consists of DL-subframe and UL-subframe. Each subframe consists of time slots. Base station (BS) sends DL-MAP and UL-MAP in DL-subframe, and all terminals listen to the DL-subframe, and know that they should listen to slots in DL-subframe, and know that they should use slots in UL-frame to transmit data. In such communications between BS and terminals, IEEE802.16(IEEE std. 802.16-2004, 2004; IEEE std. 802.16e-2005, 2005) supports four class for QoS, which are UGS, rtPS, nrtPS, BE. In UGS class, BS assigns fixed-size periodic data grants to both of uplink and downlink in terminals. In rtPS class and nrtPS class, BS assigns data grants to downlink, and polls to terminals in accordance with the reserved capacity for uplink in each terminal, and in nrtPS class, terminals are additionally allowed to use contention requests for uplink bandwidth (BW). In BE class, terminals are allowed to use contention requests only for both of uplink and downlink, and BS does not poll to terminals.

On the analysis, the assumptions are as follows.

- 16-frame length is constant.
- The multiplexing is TDD.
- The DL-subframe and DL-subframe length in frame is the ratio of 1:1.
- The modulation for each link is unchanged after the communication is arranged
- A time is normalized by slot.

3.1 16-link in UGS

In UGS class, BS assigns fixed-size periodic data grants to both of uplink and downlink in terminals. The fixed-sized periodic data grants is slots of which map is in DL-MAP or UL-MAP. The data arrival process at slot can be approximated to poisson process(Bertsekas & Gallager, 1992) (Note. data arrival at link means transmission data occurrence in link). Based on the above, argue the average time that a packet waits in queue of downlink, which is $W_{dl.UGS}$. $W_{dl.UGS}$ consists of the follows.

- The average residual time $R_{dl.UGS}$. When a new packet arrives at 16-downlink, a 16-frame is already being processed. $R_{dl.UGS}$ is a remaining average time until the current 16-frame is processed completely.
- The queued packet average processing time for UGS of downlink, $Q_{dl.UGS}$. $\bar{Q}_{dl.UGS}$ is a average time to process the all queued packets in UGS of downlink on a packet arrival.
- The average advance time $A_{dl.UGS}$. In 16-frame, $A_{dl.UGS}$ is a average time to process the other packets before a packet in USG of downlink is processed.

$R_{dl.UGS}$ consists of $R_{ds.UGS}$, which is the average residual time for the packet in USG of downlink, and R_{other} , which is the average residual time for the packet in frame except for UGS of downlink. Let $C_{dl.UGS}$ be the reserved slots in frame for UGS of downlink, $R_{ds.UGS}$ is $C_{dl.UGS}/2$. Let $\bar{V}_{dl.UGS}$ and $\bar{V}_{dl.UGS}^2$ be respectively the first and second moment of process time for a packet in frame except for UGS of downlink, R_{other} is $\bar{V}_{dl.UGS}^2/2\bar{V}_{dl.UGS}$. Let L_F be the number of slots in 16-frame, then $R_{dl.UGS}$ is derived as follows.

$$\begin{aligned} R_{dl.UGS} &= \frac{C_{dl.UGS}}{L_F} R_{ds.UGS} + \left(1 - \frac{C_{dl.UGS}}{L_F}\right) \bar{R}_{other} \\ &= \frac{C_{dl.UGS}^2}{2L_F} + (L_F - C_{dl.UGS}) \frac{\bar{V}_{dl.UGS}^2}{2\bar{V}_{dl.UGS}L_F} \end{aligned} \quad (35)$$

Argue $Q_{dl.UGS}$. Based on Little's theorem (Gross & Harris, 1985), the number of queued packets in UGS of downlink, which is $N_{dl.UGS}$, is derived as follows.

$$N_{dl.UGS} = F_{dl.UGS} \cdot W_{dl.UGS} \quad (36)$$

Where $F_{dl.UGS}$ is a packet arrival rate at UGS of downlink, which is average number of arrival packets within a slot in UGS of downlink, $W_{dl.UGS}$ is the average time that a packet waits in queue in UGS of downlink. Let m be a data grants period which is expressed by the number of frames, and $Q_{dl.UGS}$ is derived as follows.

$$Q_{dl.UGS} = F_{dl.UGS} \cdot W_{dl.UGS} \cdot m \cdot L_F \quad (37)$$

$A_{dl.UGS}$ is equal to the residual time of DL-subframe, and is derived as follows.

$$A_{dl.UGS} = \frac{C_{dl.UGS}^2}{2L_{dl}} + (L_{dl} - C_{dl.UGS}) \frac{\overline{V}_{ds.UGS}^2}{2\overline{V}_{ds.UGS}L_{dl}} \quad (38)$$

L_{dl} is the number of slots in DL-subframe, $\overline{V}_{ds.UGS}$ and $\overline{V}_{ds.UGS}^2$ are respectively the first and second moment of process time of a packet in DL-subframe except for UGS. Accordingly, $W_{dl.UGS}$ is expressed as follows.

$$\begin{aligned} W_{dl.UGS} &= \frac{R_{dl.UGS} + F_{dl.UGS} \cdot W_{dl.UGS} \cdot m \cdot L_F + A_{dl.UGS}}{1 - mF_{dl.UGS}L_F} \\ W_{dl.UGS} &= \frac{R_{dl.UGS} + A_{dl.UGS}}{1 - mF_{dl.UGS}L_F} \end{aligned} \quad (39)$$

Based on Eq.(39), the average delay in UGS of downlink, which is $T_{dl.UGS}$, is derived as follows.

$$T_{dl.UGS} = W_{dl.UGS} + C_{dl.UGS} \quad (40)$$

Assuming the modulation for each link to be unchanged, $C_{dl.UGS}$, $\overline{V}_{dl.UGS}$, $\overline{V}_{dl.UGS}^2$, $\overline{V}_{ds.UGS}$, and $\overline{V}_{ds.UGS}^2$ are constant even if $F_{dl.UGS}$ changes, and they are independent on $F_{dl.UGS}$. That is, $R_{dl.UGS}$ and $A_{dl.UGS}$ are independent on $F_{dl.UGS}$. Therefore, using Eq. (40), the first and second derivative of $T_{dl.UGS}$ at $F_{dl.UGS}$ are derived respectively as follows.

$$\frac{dT_{dl.UGS}}{dF_{dl.UGS}} > 0 \quad \frac{d^2T_{dl.UGS}}{dF_{dl.UGS}^2} > 0 \quad (41)$$

Consequently, $T_{dl.UGS}$ is a convex monotonically increasing function of $F_{dl.UGS}$.

Argue $W_{ul.UGS}$, which is the average time that a packet waits in queue of uplink. Similar to $W_{dl.UGS}$, $W_{ul.UGS}$ consists of $R_{ul.UGS}$, which is the average residual time for frame on a packet arrival at USG of uplink, $Q_{ul.UGS}$, which is the queued packet processing time for UGS of uplink, and $A_{ul.UGS}$ which is the average advance time for UGS of uplink. $R_{ul.UGS}$ is common to $R_{dl.UGS}$, and $Q_{ul.UGS}$ is $F_{ul.UGS}W_{ul.UGS}mL_F$ based on Little's theorem. $A_{ul.UGS}$ is the sum of L_{dl} and the residual time for UL-subframe because UL-subframe is arranged to following DL-subframe. Let $T_{ul.UGS}$ and $\overline{C}_{ul.UGS}$ be respectively the average delay in USG of uplink and the number of reserved slots for UGS of uplink, $W_{ul.UGS}$ and $T_{dl.UGS}$ are respectively derived as follows

$$\begin{aligned} W_{ul.UGS} &= \frac{R_{ul.UGS} + A_{ul.UGS}}{1 - mF_{ul.UGS}L_F} \\ T_{ul.UGS} &= W_{ul.UGS} + \overline{C}_{ul.UGS} \end{aligned} \quad (42)$$

Similar to downlink, $R_{ul.UGS}$ and $A_{ul.UGS}$ are independent on $F_{ul.UGS}$. Accordingly, $T_{ul.UGS}$ is a convex monotonically increasing function of $F_{ul.UGS}$.

3.2 16-downlink in rtPS and nrtPS

In rtPS, BS periodically assigns data grants to downlink of terminals based on the reserved capacity for the link. Similar to UGS, delay of 16-downlink in rtPS, which is $T_{dl.rtPS}$, is derived as follows.

$$\begin{aligned} R_{dl.rtPS} &= \frac{\overline{X}_{dl.rtPS}^2}{2L_F} + (L_F - \overline{X}_{dl.rtPS}) \frac{\overline{V}_{dl.rtPS}^2}{2\overline{V}_{dl.rtPS}L_F} \\ A_{dl.rtPS} &= \frac{\overline{X}_{dl.rtPS}^2}{2L_{dl}} + (L_{dl} - \overline{X}_{dl.rtPS}) \frac{\overline{V}_{ds.rtPS}^2}{2\overline{V}_{ds.rtPS}L_{dl}} \\ W_{dl.rtPS} &= \frac{R_{dl.rtPS} + A_{dl.rtPS}}{1 - mF_{dl.rtPS}L_F} \\ T_{dl.rtPS} &= W_{dl.rtPS} + \overline{X}_{dl.rtPS} \\ \overline{X}_{dl.rtPS} + \overline{V}_{dl.rtPS} &= L_F \\ \overline{X}_{dl.rtPS} + \overline{V}_{ds.rtPS} &= L_{dl} \end{aligned} \quad (43)$$

$\overline{X}_{dl.rtPS}$ and $\overline{X}_{dl.rtPS}^2$ are respectively the first and second moment of the number of granted slots, which is a process time of a packet, for rtPS of downlink, $\overline{V}_{dl.rtPS}$ and $\overline{V}_{dl.rtPS}^2$ be respectively the first and second moment of process time of a packet in frame except for rtPS of downlink, $\overline{V}_{ds.rtPS}$ and $\overline{V}_{ds.rtPS}^2$ be respectively the first and second moment of process time of a packet in DL-subframe except for rtPS, $F_{dl.rtPS}$ is a rtPS packet arrival rate at 16-downlink, and $W_{dl.rtPS}$ is the average time that a packet waits in queue in rtPS of downlink, Argue the dependence of \overline{X}_{rtPS} and $\overline{X}_{dl.rtPS}^2$ on $F_{dl.rtPS}$. Assuming the modulation for each link to be unchanged, \overline{X}_{rtPS} increases in the linear for the increase in $F_{dl.rtPS}$. Therefore, the dependence of \overline{X}_{rtPS} and $\overline{X}_{dl.rtPS}^2$ on $F_{dl.rtPS}$ are respectively expressed as follows.

$$\begin{aligned} \frac{d\overline{X}_{dl.rtPS}}{dF_{dl.rtPS}} &> 0 & \frac{d^2\overline{X}_{dl.rtPS}}{dF_{dl.rtPS}^2} &= 0 \\ \frac{d\overline{X}_{dl.rtPS}^2}{dF_{dl.rtPS}} &> 0 & \frac{d^2\overline{X}_{dl.rtPS}^2}{dF_{dl.rtPS}^2} &= 0 \end{aligned} \quad (44)$$

Based on Eq. (43) and (44), the dependence of $T_{dl.rtPS}$ on $F_{dl.rtPS}$ is derived as follows.

$$\frac{dT_{dl.rtPS}}{dF_{dl.rtPS}} > 0 \quad \frac{d^2T_{dl.rtPS}}{dF_{dl.rtPS}^2} > 0 \quad (45)$$

The difference of nrtPS form rtPS is the length of data grant periods, and the data grants period in nrtPS is longer than that in rtPS. Then the depenadence of delay $T_{dl.nrtPS}$ on $F_{dl.nrtPS}$, which is nrtPS packet arrival rate at uplink, is the same as that in rtPS. Consequently, $T_{dl.rtPS}$ and $T_{dl.nrtPS}$ are a convex monotonically increasing function of the each packet arrival rate.

3.3 16-uplink in rtPS

In rtPS, BS periodically polls to terminals in accordance with the reserved capacity for uplink, and terminals reply by sending BW requests with allocated space (i.e., contention free). In next frame, BS assigns data grants which is mapped by UL-MAP to terminals, and terminals use data grant to transmit data. The difference of rtPS of uplink from that of downlink is that

two frames is necessary to transmit a packet. Let $R_{ul.rtPS}$, $A_{ul.rtPS}$, m , and $F_{ul.rtPS}$, $\bar{X}_{ul.rtPS}$ be respectively the average residual time for rtPS packet of uplink, the average advance time for rtPS packet of uplink, the polling period in rtPS, the packet arrival rate at rtPS of uplink, and the average process time for packet in rtPS of uplink, $W_{ul.rtPS}$, which is the queued packet processing time for UGS of uplink, and $T_{ul.rtPS}$, which is the average delay in rtPS of uplink, are respectively expressed as follows.

$$\begin{aligned} W_{ul.rtPS} &= \frac{R_{ul.rtPS} + A_{ul.rtPS}}{1 - 2mF_{ul.rtPS}L_F} \\ T_{ul.rtPS} &= W_{dl.rtPS} + \bar{X}_{ul.rtPS} \end{aligned} \quad (46)$$

$R_{ul.rtPS}$ is common to $R_{dl.rtPS}$, and $A_{ul.rtPS}$ is the sum of $A_{dl.rtPS}$ and L_F because the rtPS of uplink is necessary to additional a frame to poll to terminal and to request BW to BS with contention free. Therefore, $R_{dl.rtPS}$ and $A_{dl.rtPS}$ are independence on $F_{ul.rtPS}$, and then $T_{ul.rtPS}$ is a convex monotonically increasing function of $F_{ul.rtPS}$ the same as rtPS of downlink.

3.4 16-uplink in nrtPS and 16-link in BE

In 16-uplink of nrtPS and 16-link of BE, also the arrival packets are enqueued and wait to be processed with FCFS. Let the waiting time be W_{bw} (argue later in detail). The packet is dequeued with FCFS, and then, is processed. The packet processing in nrtPS is based on the polling from BS the same as uplink of rtPS. Furthermore, uplink of nrtPS is additionally allowed to use contention BW request. In BE, the link is allowed to use contention BW request only. In such contention mode, terminals send BW request during the contention period in UL-subframe. Depending on the number of contention BW request, the collision of BW request occurs. In contention BW request, each terminal resolves and avoids the collision as follows.

- Each terminal waits the random number of slots before sending BW request in the contention period. The number of waiting slots, which is back-off counter, is generated based on exponential binary backoff mechanism.
- The backoff counter is decreased during the contention period.
- When the counter is zero, terminal sends BW request in the contention period.
- The terminal sending BW request waits data grants in DL/UL-map from BS.
- When the terminal does not receive data grants from BS in duration of the timer, terminal increases the contention window size, and generates the backoff counter based on exponential binary backoff mechanism, and then waits the opportunity sending BW request when the counter is zero. That is the retransmission process.

The contention BW request is analyzed based on the following model.

- The packet processing time consists of BW request opportunity waiting period, BW request attempt period, and packet transmission period.
- A BW request opportunity waiting period is the number of slots to be spent until the back-off counter becomes zero.
- A BW request attempt period is the number of slots to be spent by BW request transmission. In BW request attempt period, BW request transmission succeeds or collides. The collision causes the timeout in receipt of data grant, and spends the number of slots corresponding to the timeout. The success spends the number of slots to be spent from BW request accepted by BS to complete transmission of a packet in terminal.

- In each terminal, Let τ_{bw} be the BW request attempt rate (req/slot) in the contention period of UL-subframe, and then the probability q_{bw} that BW request is transmitted successfully is $(1 - \tau_{bw})^{n-1}$, where n is the number of terminals transmitting BW request.
- In each terminal, the packet arrival process (i.e., upload traffic) and packet request process (i.e., download traffic) is poisson process (Bertsekas & Gallager, 1992). Let F_{bw} be an packet arrival/request rate (packets/slot), which need the contention BW request.
- The contention period ratio, which is the ratio of the number of slots in the contention period in a frame, is constant. Let U_c be the contention period ratio.
- The process of the BW request that BS receives is assumed to FCFS, and the allocating data grants rate (slot/packet) in DL-subframe or UL-subframe for BW request in BS is S_{dg} , and is constant.

The contention BW request process is the same as the model described in 2.1.1 except for t_b in Eq.(4), t_s and t_c in Eq.(6). t_b is 1 because the contention BW request process decrements the backoff counter without carrier sensing. t_c is the number of slots to be spent by timeout of data grant receipt from BS, and is a constant. t_s is the number of slots to be spent from the success transmitting of BW request to the complete transmission of packet, and then it depends on F_{bw} . t_s is divided into t_{ss} , which is the air time of BW request from terminal to BS, and t_{bs} , which is the time from the receipt of BW request in BS to the complete transmission of packet in terminal, and t_{ss} is a constant.

Here, argue the dependence of t_{bs} on F_{bw} . In S_{bw} which is the average time from first transmission attempt of contention BW request to successful transmission of that, the average number of arrival/request packets for contention BW request is $F_{bw}S_{bw}$, and, in S_{bw} , the average number of BW request transmission attempts is $(1 - q_{bw})/q_{bw} + 1$. Therefore, τ_{bw} is expressed as follows.

$$\tau_{bw} = \frac{F_{bw}S_{bw}}{U_c S_{bw} q_{bw}} = \frac{F_{bw}}{U_c q_{bw}} \quad (47)$$

And, based on Eqs.(3), (4) and (5), S_{bw} is shown as follows.

$$\begin{aligned} S_{bw} &= B_{bw} + A_{bw} \\ B_{bw} &= \frac{\eta CW_{min} - 1}{2q} \\ A_{bw} &= \frac{1-q}{q} t_c + t_{ss} \end{aligned} \quad (48)$$

Furthermore, let F_{bw_bs} be the arrival rate of BW request at BS, F_{bw_bs} is shown as follows.

$$F_{bw_bs} = q_{bw} n F_{bw} \quad (49)$$

Based on Eqs.(47), (48) and (49), on condition of $F_{bw} < 1/S_{bw}$, the dependence of F_{bw} , F_{bw_bs} and S_{bw} on τ_{bw} is respectively shown as follows.

$$\frac{dF_{bw}}{d\tau_{bw}} > 0 \quad \frac{dF_{bw_bs}}{d\tau_{bw}} > 0 \quad \frac{dS_{bw}}{d\tau_{bw}} > 0 \quad (50)$$

Figure 8(a) and 8(b) respectively illustrates the dependence of F_{bw} and F_{bw_bs} on τ_{bw} by using Eqs. (47), (49), and each also shows the same characteristics. Therefore, on condition of $F_{bw} < 1/S_{bw}$, the dependence of F_{bw_bs} and S_{bw} on F_{bw} is respectively shown, by using Eq.(50), as follows.

$$\frac{dF_{bw_bs}}{dF_{bw}} > 0 \quad \frac{dS_{bw_bs}}{dF_{bw}} > 0 \quad (51)$$

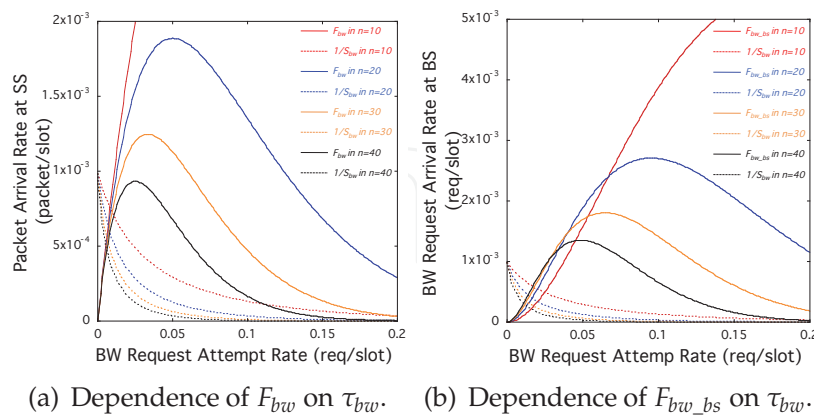


Fig. 8. Dependence of each element on τ_{bw} .

Argue W_{bw_bs} which is the waiting time in queue of BS for data grant. The process of the received BW requests in BS is assumed to be FCFS, and conceptually it can be view as queueing system of which the packet arrival rate is F_{bs_bs} and the packet service rate is S_{dg} . Therefore, W_{bw_bs} is shown, based on Eq.(14), as follows

$$W_{bw_bs} = \frac{F_{bw_bs} \overline{S_{dg}^2}}{2(1 - F_{bw_bs} S_{dg})} \quad (52)$$

S_{dg} is constant for F_{bw_bs} , and then, on condition of $F_{bw_bs} < 1/S_{dg}$, the dependence of W_{bw_bs} on F_{bw_bs} is shown as follows.

$$\frac{dW_{bw_bs}}{dF_{bw_bs}} > 0 \quad (53)$$

t_{bs} is the sum of W_{bw_bs} and S_{dg} , and then, based on Eqs.(51) and (56), the dependence of t_{bs} on F_{bw} on condition of $F_{bw} < 1/S_{bw}$ and $F_{bw_bs} < 1/S_{dg}$, that is, within the link capacity, is shown as follows.

$$\frac{dt_{bs}}{dF_{bw}} > 0 \quad (54)$$

Argue W_{bw} which is the packet waiting time in queue of terminal. According to the exponential binary backoff model described in 2.1.1, and applying $t_b = 1$ and the constance of t_c for τ_{bw} to Eq.(15), W_{bw} is derived as follows.

$$W_{bw} = \frac{F_{bw} \overline{S_{bw}^2}}{2(1 - F_{bw} S_{bw})} \quad (55)$$

According to Eqs.(16), (50), on condition of $F_{bw} < 1/S_{bw}$, that is, within link capacity, the dependence of W_{bw} on F_{bw} is derived as follows.

$$\frac{dW_{bw}}{dF_{bw}} > 0 \quad (56)$$

Finally, T_{bw} , which is the average delay for contention BW request, is the sum of W_{bw} , S_{bw} , t_{ss} and t_{bs} , the dependence of T_{bw} on F_{bw} is derived, based on Eqs.(51), (54) and (56), as follows.

$$\frac{dT_{bw}}{dF_{bw}} = \frac{d}{dF_{bw}}(W_{bw} + S_{bw} + t_{ss} + t_{bs}) > 0 \quad (57)$$

Therefore, T_{bw} is monotonically increasing function of F_{bw} .

3.5 Packet distribution for 16-link

The average delay on 16-link, except uplink in nrtPS and link in BE, is a convex monotonically increasing function of packet arrival rate, therefore, its characteristics on packet distribution corresponds to that of 11-downlink. On the other hand, 16-uplink in nrtPS and 16-link in BE are a monotonically increasing function of packet arrival rate, therefore, their characteristics on packet distribution corresponds to that of 11-uplink.

4. IP packet distribution for M-route compositing IEEE802.11/16 links

Based on the analyzed characteristics of 11/16-link for packet distribution, the characteristics of the access route compositing multiple 11-links or 16-links is the same. Therefore, the characteristics of M-route compositing 11links and 16-links for the packet distribution corresponds to that of the access route compositing multiple 11-links or 16-links.

According to the above, IP packet distribution method for the M-route compositing 11-links and 16-links be described.

The characteristics of M-route compositing 11/16-link for the packet distribution corresponds to that of the access route compositing multiple 11-links or 16-links because that of 11-link and 16-link are the same.

4.1 Restriction condition

According to Eqs.(26) and (33), the optimal solution of the M-route cost can be searched by the repeating packet distribution that the average delay of distribution source link becomes larger than that of distribution destination link, and that the average delay of both source link and destination link become equal finally. Additionally, the packet distribution for download traffic needs to meet the condition Eq.(29) when the source link on the packet distribution is an 11-link.

Here, argue the condition Eq.(29). Transforming Eq.(29) into finite difference approximation, it is shown as follows.

$$\frac{d\bar{T}_{(bs,k[i])}}{dF_{(bs,k[i])}} \approx \frac{\Delta\bar{T}_{(bs,k[i])}}{\Delta F_{(bs,k[i])}} > 0 \quad (58)$$

Because 11-link $k[i]$ is a source link on packet distribution, $\Delta F_{(bs,k[i])} < 0$. Therefore, to meet Eq.(58), $\Delta\bar{T}_{(bs,k[i])} < 0$. In other words, it is that the average delay of source 11-interface on packet distribution decreases. The increase in average delay of source 11-interface k does not meet the condition and it occurs in the following unsteady state.

- The packet arrival rate at other links provided by 11-interface k increases.
- The number of links provided by 11-interface k increases.

The first item in the above list means the increase in contention with other terminals, thus it also causes the increase in average delay of source link when source link is 11-uplink or 16-uplink in nrtPS or 16-link in BE. The second item means the increase in a number of

terminals, thus it causes the increase in average delay of source link because of the same reason as the first item. Then it also occurs when source link on packet distribution is 11-uplink or 16-uplink in nrtPS or 16-link in BE. In above cases, M-route cost also loses the monotonically increasing characteristics for packet arrival rate. Therefore, in consideration of the unsteady state that traffic fluctuates, the restriction condition which is the decrease in the average delay of source link on packet distribution is a necessary condition to bring the M-route cost to the optimal solution.

4.2 Search for optimal solution of M-route cost with packet distribution

Argue the search for optimal solution of M-route cost with Packet Distribution in unsteady state by the following packet distribution.

- $M_{(x,y)}$ is a M-route from x to y . On x and y , one is a base station and the other is a terminal.
- Packets transmitted to y at x are distributed.
- K denotes a source interface on the packet distribution. J denotes a destination interface on the packet distribution.
- K is either an 11-interface or 16-interface and J is also either an 11-interface or 16-interface.
- $(x, Z[y])$ denotes a certain link to y in x , which link is provided by a certain interface Z .
- $F_{(x,Z[y])}$ denotes a packet arrival rate at $(x, Z[y])$
- $T^*(p_{(x,Z[y])})$ denotes interface average delay $\bar{T}(p_{(x,Z[y])})$ if Z is 11-interface, and denotes link average delay $T(p_{(x,Z[y])})$ if Z is 16-interface.

Based on subsection 4.1, the search for optimal solution of M-route cost in unsteady state is the search for the packet distribution meeting the following conditions.

$$T^*_{(x,K[y])}(n) - T^*_{(x,J[y])}(n) > 0 \quad \Delta T^*_{(x,K[y])}(n) < 0 \quad (59)$$

where $\Delta T^*_{(x,K[y])}(n)$ denotes the difference between $T^*_{(x,K[y])}(n)$ and $T^*_{(x,K[y])}(n-1)$. According to Eq. (59), the proposed packet distribution method implements the search for the optimal solution in IP layer using the measured average delay in MAC layer as the following iteration.

Step1: In the initial period, packets are distributed equally to each link in M-route with a round robin manner.

Step2: At end of the initial period, $T^*_{(x,Z[y])}(0)$ of each link in M-route is derived, and $(x, Max[y])(0)$ which has maximum average delay in the initial (0-th) period, and $(x, Min[y])(0)$ which has minimum average delay in the initial (0-th) period, is respectively selected in $M_{(x,y)}$. On the packet distribution, $(x, Max[y])(0)$ and $(x, Min[y])(0)$ is respectively assigned to the source link $(x, K[y])(1)$ in the next (1-th) period and the destination link $(x, J[y])(1)$ in that period. $\Delta F_{(x,K[y])}(1)$, which is the amount of packet distribution from $(x, K[y])(1)$ to $(x, J[y])(1)$ in the next (1-th) period, is derived as follows. where $r_{(x,y)}$ denotes the packet distribution rate of $M_{(x,y)}$, and r_0 denotes the initial packet distribution rate.

$$\begin{aligned} \Delta F_{(x,K[y])}(1) &= r_{(x,y)}(1) \cdot F_{(x,K[y])}(0) \\ r_{(x,y)}(1) &= r_0 \end{aligned} \quad (60)$$

Step3: According to $\Delta F_{(x,K[y])}(1)$, the packet distribution in the 1-th period is carried out.

Step4: At end of n -th period ($n \geq 1$), $T_{(x,Z[y])}^*(n)$ of each link in $M_{(x,y)}$ is derived. The delay of each packet is a period when the packet arrives at IP layer, and is enqueued in queue of an interface, and is dequeued by an interface, and is sent and interface receives its ACK based on the media access control. Therefore, it can be measured within packet distributing side x . Based on the relation of $T_{(x,K[y])}^*(n)$ and $T_{(x,J[y])}^*(n)$, $\Delta F_{(x,K[y])}(n+1)$ is derived as follows.

- In $T_{(x,K[y])}^*(n) > T_{(x,J[y])}^*(n)$ and in $\Delta T_{(x,K[y])}^*(n) < 0$, Eq. (59) is met. Therefore, $\Delta F_{(x,K[y])}(n+1)$ is allocated the same as $\Delta F_{(x,K[y])}(n)$, and it is shown as follows.

$$\begin{aligned}\Delta F_{(x,K[y])}(n+1) &= r_{(x,y)}(n+1) \cdot \Delta F_{(x,K[y])}(n) \\ r_{(x,y)}(n+1) &= r_{(x,y)}(n)\end{aligned}\quad (61)$$

- In $T_{(x,K[y])}^*(n) < T_{(x,J[y])}^*(n)$ and in $\Delta T_{(x,K[y])}^*(n) < 0$, $M_{(x,y)}$ cost goes beyond the optimal solution and ascends the gradient. Because it is caused by the excessive packet distribution from source link to destination link, $\Delta F_{(x,K[y])}(n+1)$ is allocated smaller than $\Delta F_{(x,K[y])}(n)$ as follows. where α is the decrement rate ($0 < \alpha < 1$).

$$\begin{aligned}\Delta F_{(x,K[y])}(n+1) &= r_{(x,y)}(n+1) \cdot \Delta F_{(x,K[y])}(n) \\ r_{(x,y)}(n+1) &= \alpha \cdot r_{(x,y)}(n)\end{aligned}\quad (62)$$

- In $\Delta T_{(x,K[y])}^*(n) > 0$, the traffic among the source link increases as shown in subsection 4.1. Because $\Delta F_{(x,K[y])}(n)$ is underestimated, and because the monotonically increasing characteristics of the source link cost for the packet distribution is regained, $\Delta F_{(x,K[y])}(n+1)$ is allocated larger than $\Delta F_{(x,K[y])}(n)$ as follows. where β is the increment rate ($\beta > 1$).

$$\begin{aligned}\Delta F_{(x,K[y])}(n+1) &= r_{(x,y)}(n+1) \cdot \Delta F_{(x,K[y])}(n) \\ r_{(x,y)}(n+1) &= \beta \cdot r_{(x,y)}(n)\end{aligned}\quad (63)$$

Step5: $(x, \text{Max}[y])(n)$ and $(x, \text{Min}[y])(n)$ are respectively selected in $M_{(x,y)}$, and are respectively assigned to $(x, K[y])(n+1)$ and $(x, J[y])(n+1)$. According to $(x, K[y])(n+1)$, $(x, J[y])(n+1)$, and $\Delta F_{(x,K[y])}(n+1)$, the $(n+1)$ -th packet distribution is carried out, then return to Step4.

In each M-route of both a base station and terminals, the above iteration gradually updates the amount of packet distribution, and brings M-route cost to the optimal solution, reducing the out-of-order packets occurred by distributing packets to multiple links.

5. Performance evaluation

In this section, the simulation evaluation of the packet distribution method for M-route compositing 11/16-links is shown.

5.1 Simulation scenario

For the simulation evaluation, OPNET 12.0A PL3 was used, and the network configuration was as follows (see Fig. 9):

- Base station is equipped with an 16-interface and 4×11a/b-interfaces. 16-interface and 11a/b-interface respectively connects to 16-antenna and 11/ab-antenna.
- The number of terminals is 100, and each terminal is equipped with 16-interface and 11a/b-interface.

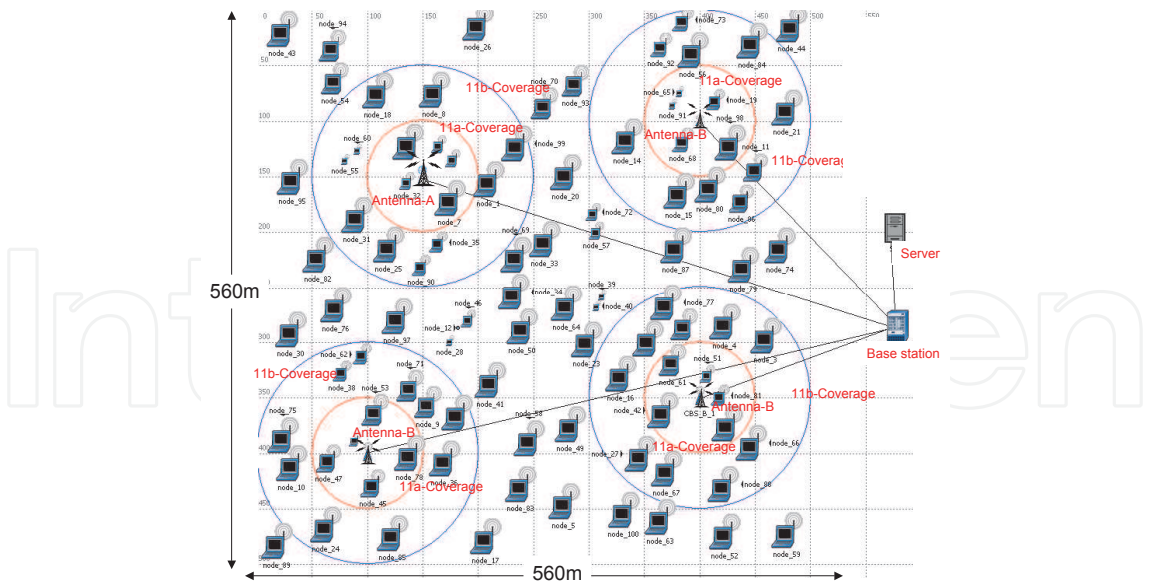


Fig. 9. Example of access network topology.

- An antenna-A which equips with 16- and 11a/b-antena, three antenna-B which equip with 11a/b-antena, and 100 terminals without mobility are randomly deployed in 560m × 560m space with a 1/10 scale of 16-coverage with 1000m radius.
- A FTP server and a Video Conference (VC) server, which are outside the wireless access network, are connected to the base station by a wired network.

In the above access network, M-route between each terminal and a base station combines available links as follows.

- The M-route between a base station and a terminal in 11a-coverage (area-A) combines 11a/b-link and 16-link.
- The M-route between a base station and a terminal in 11b-coverage and outside 11a-coverage (area-B) combines 11b-link and 16-link.
- The M-route between a base station and a terminal outside 11b-coverage (area-C) uses only 16-link.

The performance of 11a/b-wireless system and 16-wireless system shown in Table 1 is applied, and each the capacity reservation of 16-link is shown in Table 2. Assuming the evaluation environment to be a suburban area in line of sight, the 11a/b-radio propagation model is a two-ray model and Ricean fading with Ricean factor 6dB(Takada, 2004), and the 16 radio propagation model is a Erceg (TerrainA).

According to (3GPP2, 2006), the VC traffic on UDP is generated at each terminal as follows:

- The average video rate in the incoming and the outgoing is 32 Kbps.
- The distribution in video rate is a truncated pareto distribution with maximum 8Kbits
- The frame rate in the incoming and outgoing is 10fps. A frame corresponds to a data packet in VC.
- As the sequence control of frame, VC waits for the frame with expected sequence number for a period of 100 msec that is equal to frame interval. The frame that arrived on excess of the period is destroyed.

- In 16-link, VC is mapped to rtPS for QoS class.

Furthermore, FTP traffic on TCP is also generated at each terminal as follows:

- In 10 sec period, FTP session which transfers a file of the size of 1K~400Kbytes starts.
- 50% of the FTP sessions are download session.
- Each FTP session is established between each terminal and a FTP server.
- In 16-link, FTP is mapped to nrtPS for QoS class.

The evaluation items are as followings.

- IP average delay (sec/packet), is the average delay between terminal and servers in an IP packet.
- IP throughput (bps), is the average arrival amount of IP packets at terminals and servers during a unit time.
- FTP response time (sec/file), is the average delay to transfer a file in end-to-end between a terminal and an FTP server.
- FTP throughput (bytes/sec), is the average amount of arrival data packets at terminals and an FTP server during a unit time.
- VC average delay (sec/frame), is the average delay of end-to-end between terminal and a VC server in a data frame.
- VC throughput (bytes/sec), is the average arrival amount of data frames at terminals and a VC server during a unit time.

The end-to-end delay is composed of the delay in wireless access network and that in wired communication between the base station and server. The delay in wired communication is common without depending on any packet distribution in wireless access network because the wired communication is out of scope of wireless access network. Therefore, the delay in wired communication can be assumed to be constant to any packet distribution in wireless access network, and the delay in wireless access network depends on packet distribution in wireless access network. In viewpoint of packet distribution, the trend of the end-to-end delay corresponds with that of the delay in wireless access network. Thus the delay in wired communication can be logically ignored. Furthermore, assuming the access speed of a future core network to be Gigabits order (Konishi et al., 2008), the delay in WiFi corresponds to $10^2 \sim 10^3$ order of that in wired core network because the bandwidth of WiFi is Mbps. Then, the delay between the base station and server is left out of consideration because it is independent on the performance of the wireless access network. Furthermore, to demonstrate the effectiveness of the proposed method, it is compared with the following methods.

- Single link (SL) uses a link. The terminals in area-A use 11a-link, the terminals in area-B use 11b-link, and the other terminals use 16-link.
- Round robin (RR) uses available links and distributes packets equally to each link.
- Actual transmission rate (TR) uses available links and distributes packets to each link in proportion to the measured transmission rate at each link in every 10 sec.

In the search for minimal solution, r_0 is 0.1, α is 0.5, β is 1.5, and the update period of packet distribution is 10 sec.

Furthermore, the link combination in IP is transparent to the upper layer. Therefore, the upper layer is provided with the M-route as a single link view.

| QoS Class | rtPS | nrtPS |
|-------------------------------------|---------|---------|
| Maximum Sustained Transmission Rate | 384Kbps | 384Kbps |
| Minimum Reserved Transmission Rate | 80Kbps | 1Kbps |

Table 2. Capacity reservation for 16-link.

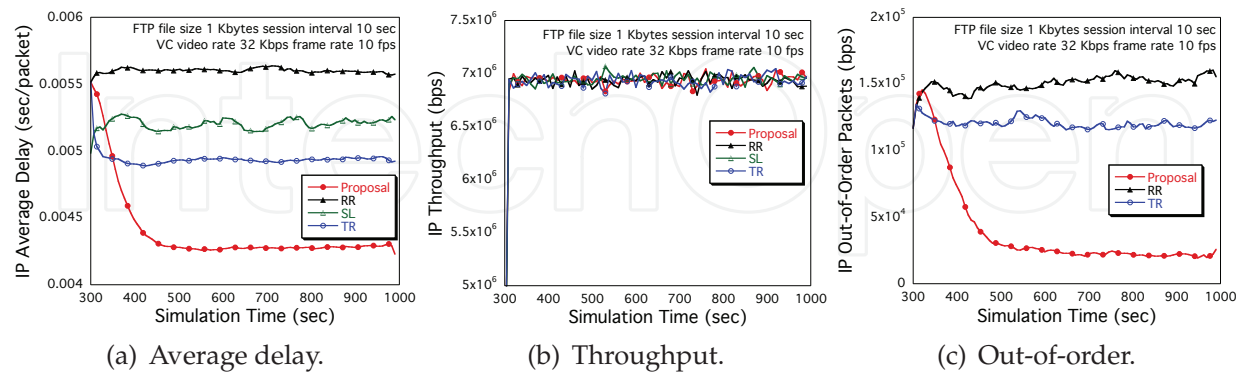


Fig. 10. Transition of IP on FTP file size 1K bytes.

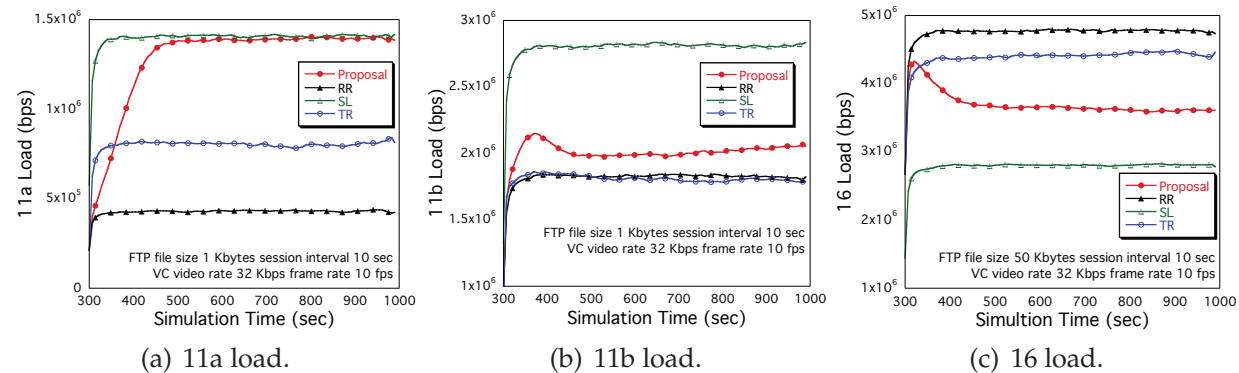


Fig. 11. Distributed traffic load to each wireless system on FTP file size 1K bytes.

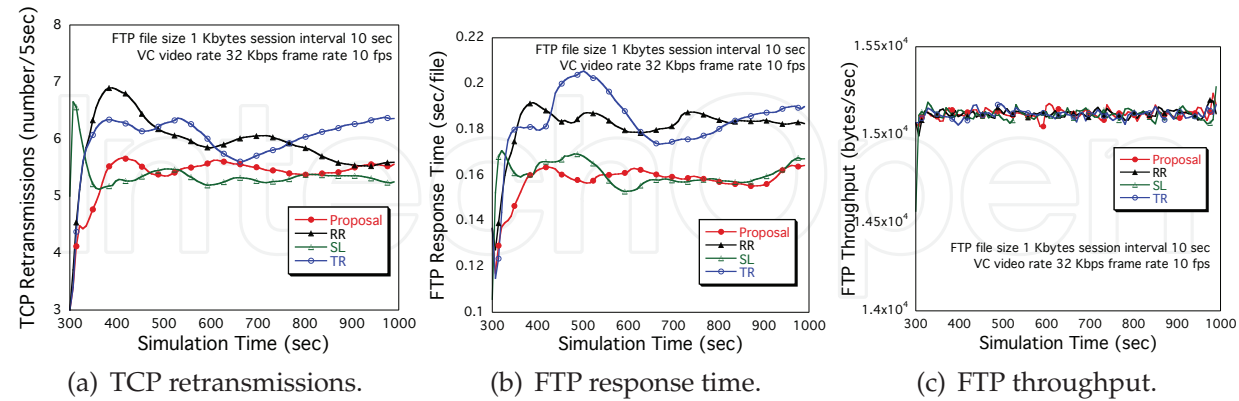


Fig. 12. Transition of TCP and FTP on FTP file size 1K bytes.

5.2 Transition of delay and throughput in low traffic load

Figures 10(a) and 10(b) show, respectively, the transition of IP average delay and IP throughput, when file size in FTP is 1K bytes. As the packet distribution proceeds, the IP average delay of the proposal decreases rapidly, and becomes much lower than that of the

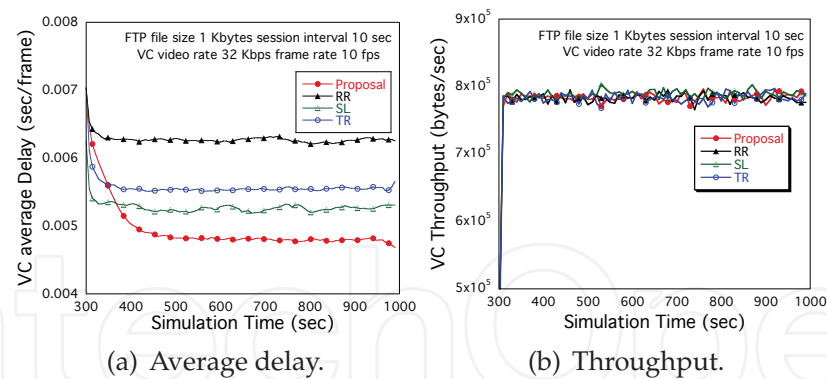


Fig. 13. Transition of VC on FTP file size 1K bytes.

others. Figures 11(a), 11(b) and 11(c) show, respectively, the transition of distributed load to 11a-wireless system (11a-load), that to 11b-wireless system (11b-load) and that to 16-wireless system (16-load), when file size in FTP is 1K bytes. The decrease in IP average delay of the proposal corresponds to the increase in 11a-load of the proposal (see Fig. 10(a) and Fig. 11(a)). In area-A, 11a accommodates a few terminals because of its narrow coverage, and the proposal distributes almost packets to 11a-link the same as SL, and saves the capacity of 11b and 16 for many terminals outside area-A. RR and TR in the area distributes packets to other link as well, thus RR and TR can not use 11a capacity effectively to save the capacity of 11b and 16. Consequently, RR and TR bring the large load to 16 (see Fig. 11(c)), which of links have low transmission rate (see Tab. 2), and it causes the inferior IP average delay of RR and TR to that of the proposal. In area-B, SL distributes all packets to 11b-link (see Fig. 11(b)), and then the packet collision in 11b occurs frequently. Thus, it causes the inferior IP average delay of SL to that of the proposal. In comparison with SL, the packet distribution of the proposal and TR improve IP performance, but that of RR lowers IP performance.

The IP out-of-order packets of the proposal decreases the same as the decrease in its IP average delay, consequently, its out-of-order packets becomes much lower than that of RR and TR (see Fig. 10(c)). Therefore, its packet distribution effects the decrease in IP average delay and the decrease in out-of-order packets. Figures 12(a) shows the number of TCP retransmissions for a period of 5 sec. The TCP retransmissions of the proposal is nearly equal to that of SL and RR, and that of TR is larger than that of the others. The cause of TCP retransmission in SL is packet loss. In area-B, SL distributes all packets to 11b, thus the packet collision occurs frequently in 11b and then it causes the TCP retransmission. The cause of TCP retransmission in the proposal, RR and TR is out-of-order packets. The number of TCP transmissions in RR is lower than that of TR. RR loads larger mount of packets with 16 than the others (see Fig. 11(c)). Because the 16-link has the low transmission rate, the IP average delay of RR is inferior to that of the others (Fig. 10(a)). Then TCP congestion window size of RR is smaller than that of TR and the proposal, and the amount of distributed packets to multiple links for a period is fewer than that of TR and the proposal, thus the probability of occurrence of out-of-order packets is lower. Consequently, the TCP retransmissions of RR is lower than that of TR. That of the proposal is also lower than that of TR, then the delay equalization between multiple links in the proposal effects the decrease in the occurrence of out-of-order packets, and effects the decrease in TCP retransmissions.

Figures 12(b) and 12(c) show, respectively, the transition of FTP response time and FTP throughput. The FTP response time of SL and the proposal are superior to that of RR and TR. The IP average delay of TR is superior to that of SL, however, the FTP response time of TR is inferior to that of SL. The inversion is caused by the large number of TCP retransmissions in

TR, and the packet distribution of TR lowers the FTP performance. The cause of the inferior FTP response time of RR to that of SL is not the TCP retransmissions, but is the small amount of TCP flow based on TCP congestion window size, then the packet distribution in RR distributes the large number of packets to 16-link, which is narrow bandwidth, and originally lowers IP performance. The number of TCP retransmissions and the FTP response time of the proposal is the same as those of SL. As the above mentioned, the cause of TCP retransmission in SL is the packet loss in 11b-link, but the cause of that in the proposal is the out-of-order packet, that is, the proposal offsets the improvement of IP performance against the out-of-order packets, and does not improve the FTP performance, but does not lower it. Figures 13(a) and 13(b) show, respectively, the transition of VC average delay and VC throughput. The VC average delay of SL is equal to the IP average delay because a VC frame corresponds to a IP packet and because out-of-order packet does not occur. In the proposal, RR, and TR, the VC average delay is larger than that of IP because the sequence control in VC waits for frame with the expected sequence on the occurrence of out-of-order packet. Therefore, VC average delay of TR is higher than that of SL though IP average delay of TR is lower than that of SL, i.e., the packet distribution of TR lowers the VC performance. On the other hand, that of the proposal is lower than that of SL, therefore, the effect of the packet distribution in the proposal overcomes the ill of it, and can improve the VC performance. That of RR is higher than that of the others because RR originally lowers IP performance.

5.3 Transition of delay and throughput in high traffic load

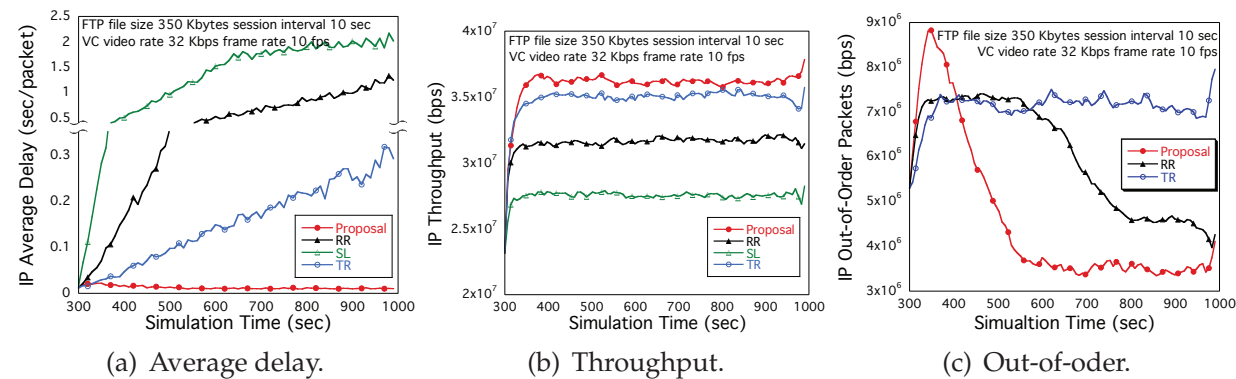


Fig. 14. Transition of IP on FTP file size 350K bytes.

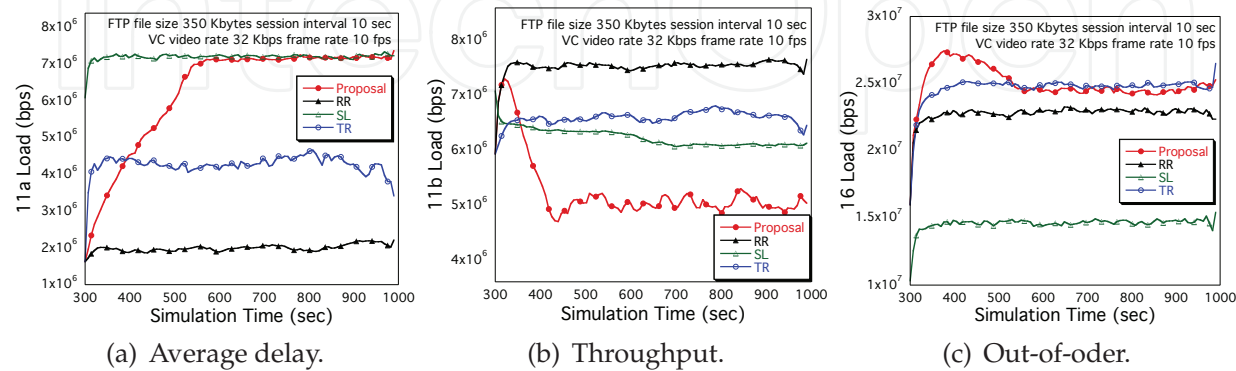


Fig. 15. Distributed traffic load to each wireless system on FTP file size 350K bytes.

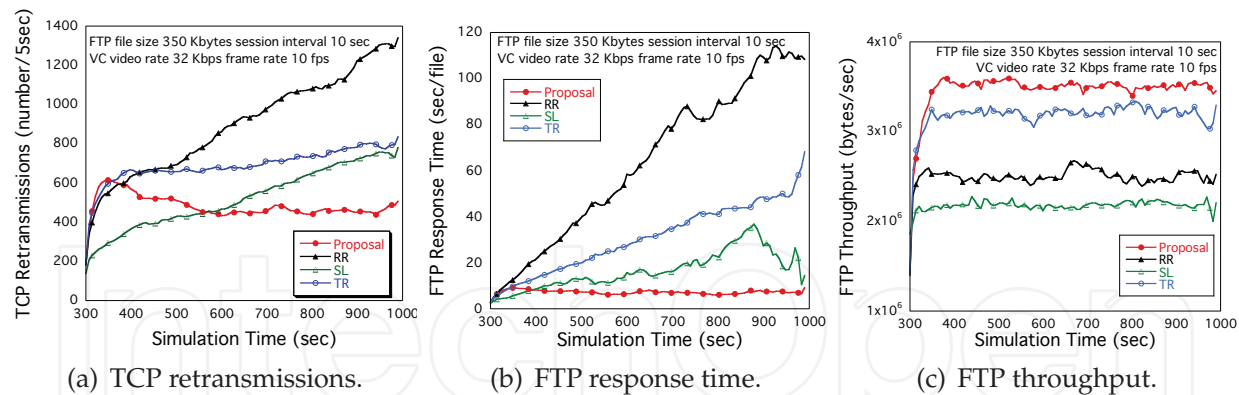


Fig. 16. Transition of TCP and FTP on FTP file size 350K bytes.

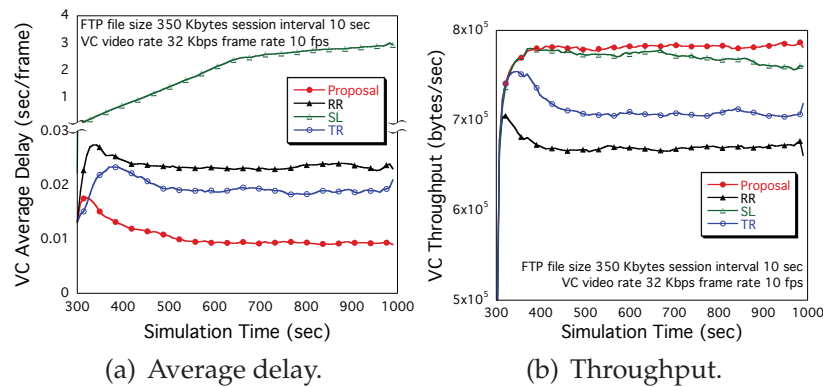


Fig. 17. Transition of VC on FTP file size 350K bytes.

Figures 14(a) and 14(b) show, respectively, the transition of IP average delay and IP throughput, when file size in FTP is 350K bytes, furthermore, Fig. 15(a), 15(b) and 15(c) show, respectively, the transition of 11a load, 11b load and 16 load, when file size in FTP is 350K bytes. The IP average delay of the proposal is low, and is stable. On the other hand, that of the others increase as linear, and become much higher than that of the proposal. Furthermore, their IP throughput are lower than that of the proposal. In area-A, the packet distribute to 11a-link brings low delay to IP because of wide bandwidth and few accommodated terminals in 11a, as mentioned in 5.2. In area-B, the packet collision and loss in 11b further increase because of the increase in traffic, and the large number of retransmissions in MAC brings the increase in delay to IP. Furthermore, the packet loss in 11b brings the decrease in throughput to IP. Each 16-link has the narrow bandwidth, but does not cause the collision because of TDD. i.e., The delay of 16-link is lower than that of 11b-link because of no retransmission process in MAC, which of delay in 11b exponentially increases based on a binary back-off mechanism. Therefore, the large number of packet distribute to 11b brings the increase in delay and the decrease in throughput to IP. Consequently, IP average delay of the proposal, which distributes the smaller number of packets to 11b than the others (see Fig. 15(b)), is lowest, and its IP throughput is highest.

Figures 14(c) and 16(a) show, respectively, the transition of IP out-of-order packets and TCP retransmissions, when file size in FTP is 350K bytes. The IP out-of-order packets of the proposal decreases rapidly as the packet distribute proceeds the same as the case that FTP file size is 1K bytes, i.e., the delay equalization between the multiple links in the proposal effects the decrease in IP out-of-order packets. That of RR also decreases, but the decrease in

the amount of TCP flow based on TCP congestion window size, which becomes small rapidly by the increase in IP delay of RR, brings it. TCP retransmission is caused by the IP packet loss and IP out-of-order packets. The TCP retransmissions in SL is caused only by IP packet loss, and IP packet loss is caused by the large number of distributed packets to 11b. That of RR, TR and the proposal is caused by IP packet loss and IP out-of-order packets. That of RR is caused largely by IP packet loss, because RR distributes the large number of packets to 11b and IP out-of-order packets decreases by the decrease in TCP flow. Therefore, the trend of TCP retransmissions of RR is similar to that of SL. TR also distributes the large number of packets to 11b, but distributes the larger number of packets than RR to 11a and 16, which of packet loss probability is much lower than 11b, i.e., the TCP retransmissions in TR is caused mainly by out-of-order packets and it reduces the upward trend of TCP retransmissions in comparison with SL and TR. On the other hand, the TCP retransmissions of the proposal is low stable in comparison with the others. The proposal distributes the much smaller number of IP packets than the others to 11b and reduces IP packet loss, furthermore, it equalizes the delay of each link in M-route, thus reduces also IP out-of-order packets. That brings the low and stable retransmissions to TCP.

Figures 16(b) and 16(c) show, respectively, the transition of FTP response time and FTP throughput, when file size in FTP is 350K bytes. The FTP response time of RR and TR increase as linear. In RR and TR, FTP session can not complete in a period of 10 sec, which is FTP session start interval, because the amount of TCP flow is restrained low by the large number of retransmissions. The active FTP session accumulates. Therefore, the access network causes the congestion. In the proposal, FTP session can complete within 10 sec, and the delay not increase and is stable. Furthermore, the throughput reaches the input load 4M bytes/sec. Therefore, the proposal controls avoids the congestion.

5.4 Dependence of delay on throughput

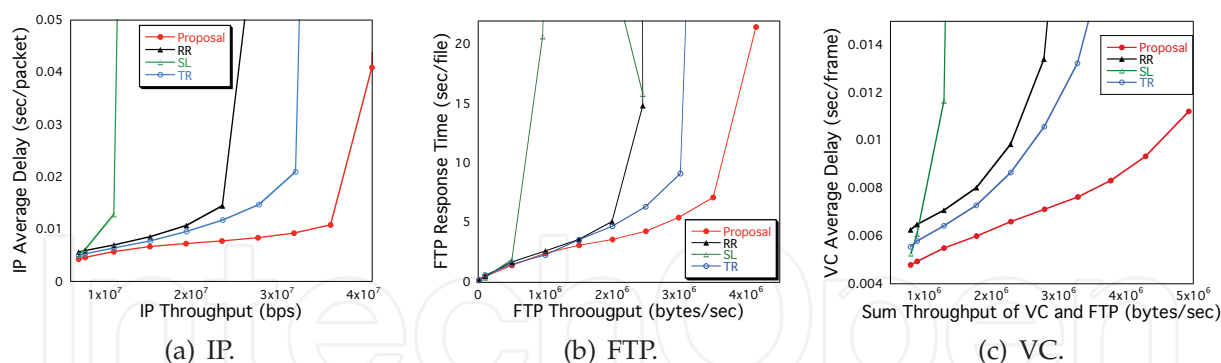


Fig. 18. Dependence of delay on throughput.

Figure 18(a), 18(b), and 18(c) shows, respectively, the dependence of IP average delay on IP throughput, the dependence of FTP response time on FTP throughput, and the dependence of VC average delay on VC throughput when FTP file size increases from 1K bytes to 400K bytes. The average delay and throughput are each the averages for 10 topologies in which the antennas and terminals are deployed randomly in the evaluation space.

When the FTP traffic is low, the performance of SL and the proposal is superior to that of RR and TR. In low load, if packets are distributed to a widest band link, that is, if the packet distribution is equalized to that of SL, the performance becomes high. The packet distribution of the proposal becomes equal to that of SL, but that of RR and TR do not. As FTP traffic

increases, the 11b-link load of M-route in 11b-coverage and outside 11a-coverage becomes high, then M-route including 11b-link needs to distribute packets to 11a-link or 16-link. SL can not distribute packets of 11b-link to other links, then SL is saturated first by the exhaustion of 11b-link capacity. By the same cause, RR and TR are saturated in FTP file size 300K bytes and 400K bytes respectively. The proposal distributes packets from 11b-link to 16-link and 11a-link, and avoids the saturation until FTP file size exceeds 400K bytes.

Summarizing, in any FTP traffic, the proposal can distribute packets effectively in comparison with other methods, and it produces low delay and high throughput on both TCP application and UDP application, and simultaneously.

6. Conclusion

In this chapter, the packet distribution characteristics in IEEE802.11-link and that in IEEE802.16-link was respectively shown, and, based on these characteristics, the packet distribution method for access route compositing IEEE802.11/16-links was proposed. Furthermore, its performance through evaluation with IEEE802.11a/b and IEEE802.16 was shown. Consequently, the proposed method was found to have the following effectiveness.

- It can greatly effectively distribute packets to IEEE802.11/16 links according to link load.
- And, it can also reduce out-of-packets caused by distributing packets to multiple links.
- Then, It can decrease delay and can increase throughput on both TCP application and UDP application, and simultaneously.

7. References

- Arkin, D. (2004). *My Life*, Arkin Publishing, Arkinson.
- Mitoralll, J. & Maguire, G. (1999). Cognitive Radio: Making Software Radios More Personal, *IEEE Pers. Comm.*, Vol. 6, No. 4, pp. 13–14 1999.
- Mitoralll, J. (1999). Cognitive Radio for Flixible Multimedia Communications, *Proc. MoMuC99*, pp. 3–10, 1999.
- Harada, H. (2005). A Study on a new wireless communications system based on cognitive radio technology, *IEICE Tech. Rep.*, Vol. 105, No. 36, pp. 117–124, 2005.
- 3GPP TS 22.258. (2005). Service Requirements for the All-IP Network (AIPN); Stage 1, v2.0.0, 2005.
- ITU-T. (2006). Y.2021: NGN Release 1, 2006.
- Phatak, D. S. & Goff, T. (2002). A novel mechanism for data streaming across multiple IP links for improving throughput and reliability in mobile environments, *Proc. IEEE INFOCOM*, pp. 773–781, 2002.
- Snoeren, A. C. (2002). Adaptive Inverse Multiplexing for Wide-Area-Wireless Networks, *Proc. IEEE GlobCom'99*, Vol.3, pp.1665–1672, 1999.
- Shrama, P.; Lee, S.; J. Brassil, J. & Shin, K. (2007). Aggregating Bandwidth for Multimode Mobile Collaborative Communities, *IEEE Tans. on MC*, Vol. 6, No. 3, pp. 280–296, 2007.
- Chebrolu, K. & Rao, R. (2006). Bandwidth Aggregation for Real-Time Applications in Heterogeneous Wireless Networks, *IEEE Tans. on MC*, Vol. 5, No. 4, pp. 388–403, 2006.
- Hsieh, H.; Kim, K. & Sivakumar, R. (2004). An end-to-end approach for transparent mobility across heterogeneous wireless networks, *Mob. Netw. Appl.*, Vol. 9, No. 4, pp. 363–378, 2004.

- Zhang, M.; Lai, J.; Krishnamurthy, A.; Peterson, L. & Wang, R. (2004). A Transport Layer Approach for Improving End-to-End Performance and Robustness Using Redundant Paths, *USENIX 2004*, 2004.
- D. Gross, D. & C. Harris, C. (1985). Fundamentals of Queueing Theory 2nd ed, *John Wiley & Sons*, 1985.
- Little, J. (1961). A Proof of the Queueing Formula $L = \lambda W$, *Oper Res J.*, 18:172–174, 1961.
- Bianchi, G. (2000). Performance analysis of the IEEE 802.11 distributed coordination function, *IEEE JSAC*, Vol. 18, No. 3, pp. 535–547, 2000.
- Carvalho, M. M. & Garcia-Luna-Aceves, J. J. (2003). Delay analysis of IEEE 802.11 in single-hop networks, *Proc. of ICNP*, pp. 146–155, 2003.
- Takizawa, Y.; Taniguchi, N.; Yamanaka, S.; Yamaguchi, A. & Obana, S. (2008). Characteristics of Packet Distribution in Wireless Access Networks Accommodating IEEE802.11 and IEEE802.16, *IPSSJ Journal*, Vol. 49, No. 9, pp. 3245–3256, 2008.
- Takizawa, Y.; Taniguchi, N.; Yamanaka, S.; Yamaguchi, A. & Obana, S. (2008). Packet Distribution Control for Wireless Access Networks Accommodating IEEE802.11 and IEEE802.16, *IPSSJ Journal*, Vol. 49, No. 10, pp. 3576–3587, 2008.
- Bertsekas, D. & Gallager, R. (1992). Data Networks, *Prentice Hall*, 1992.
- IEEE Std 802.16-2004. (2004). Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2004.
- IEEE Std. 802.16e-2005. (2005). Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access System, 2005.
- Takada, J. (2004). Radiowave Propagation for Mobile Satellite Communications, *Tech. Rep. of IEICE*, Vol. 104, No. 671, pp. 13–16, 2004.
- 3GPP2. (2006). C30-20060823-004A Evaluation methodology V4.0, 2006.
- Konishi, S.; Wang, X.; Kitahara, T.; Nakamura, H. & Suzuki, T. (2008). A Study on Ultra Low-Latency Mobile Networks, *Wireless Personal Comm.:An Int. Journal*, Vol. 44, No.1, pp.57–73, 2008.
- Nakayo, D. J. and Hossain, E. (2006). A Queuing-Theoretic and Optimization-Based Model for Radio Resource Management in IEEE 802.16 Broadband Wireless Networks, *IEEE Trans Comp.*, Vol. 55, No.11, pp. 1473–1488, 2006.
- Cho, D.; Song, J.; Kim, M. and Han, K. (2006). Performance Analysis of the IEEE 802.16 Wireless Metropolitan Area Network, *Proc. IEEE DFMA 2005.*, 2005.
- Lin, L.; Jia, W. and Lu, W. (2007). Performance Analysis of IEEE 802.16 Multicast and Broadcast Polling based Bandwidth Request, *Proc. IEEE WCNC 2007.*, 2007.
- Iyengar, R.; Iyer, P. and Biplab Sikdar, B. (2005). Delay Analysis of 802.16 based Last Mile Wireless Networks, *Proc. IEEE GlobeCom 2005.*, 2005.
- He, J.; Guild, K.; Yang, K. and Chen, H. (2007). Modeling Contention Based Bandwidth Request Scheme for IEEE 802.16 Networks, *IEEE Comm. Letter*, Vol. 11, No.8, pp. 698–700, 2007.
- Ni, Q.; Xiao, Y.; Turlikov, A. and Jiang, T. (2007). Investigation of Bandwidth Request Mechanisms under Point-to-Multipoint Mode of WiMAX Networks, *IEEE Comm. Magazine*, Vol. 4, No.4, pp. 477–486, 2007.



Recent Advances in Wireless Communications and Networks

Edited by Prof. Jia-Chin Lin

ISBN 978-953-307-274-6

Hard cover, 454 pages

Publisher InTech

Published online 23, August, 2011

Published in print edition August, 2011

This book focuses on the current hottest issues from the lowest layers to the upper layers of wireless communication networks and provides “real-time” research progress on these issues. The authors have made every effort to systematically organize the information on these topics to make it easily accessible to readers of any level. This book also maintains the balance between current research results and their theoretical support. In this book, a variety of novel techniques in wireless communications and networks are investigated. The authors attempt to present these topics in detail. Insightful and reader-friendly descriptions are presented to nourish readers of any level, from practicing and knowledgeable communication engineers to beginning or professional researchers. All interested readers can easily find noteworthy materials in much greater detail than in previous publications and in the references cited in these chapters.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yasuhisa Takizawa (2011). Traffic Control for Composite Wireless Access Route of IEEE802.11/16 Links, Recent Advances in Wireless Communications and Networks, Prof. Jia-Chin Lin (Ed.), ISBN: 978-953-307-274-6, InTech, Available from: <http://www.intechopen.com/books/recent-advances-in-wireless-communications-and-networks/traffic-control-for-composite-wireless-access-route-of-ieee802-11-16-links>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen