

IEEE 802.16 竞争解决方案的性能分析*

陆文彦¹⁺, 贾维嘉^{1,2}, 杜文峰³, 张历卓^{1,2}

¹(中南大学 信息科学与工程学院, 湖南 长沙 410083)

²(香港城市大学 计算机科学系, 香港 九龙)

³(深圳大学 信息工程学院, 广东 深圳 518060)

Performance Analysis of the Contention Resolution Scheme in IEEE 802.16

LU Wen-Yan¹⁺, JIA Wei-Jia^{1,2}, DU Wen-Feng³, ZHANG Li-Zhuo^{1,2}

¹(School of Information Science & Engineering, Central South University, Changsha 410083, China)

²(Department of Computer Science, City University of Hong Kong, Kowloon, China)

³(School of Information Engineering, Shenzhen University, Shenzhen 518060, China)

+ Corresponding author: Phn: +86-731-8877711, E-mail: luwenyan@gmail.com, http://www.csu.edu.cn

Lu WY, Jia WJ, Du WF, Zhang LZ. Performance analysis of the contention resolution scheme in IEEE 802.16. *Journal of Software*, 2007,18(9):2259-2270. <http://www.jos.org.cn/1000-9825/18/2259.htm>

Abstract: At present, the mandatory contention resolution that shall be supported by IEEE 802.16 is based on truncated binary exponential backoff algorithm. In this paper, the differences of contention mechanism between IEEE 802.16 and IEEE 802.11 are analyzed. The calculations of performance metrics such as the utilization of transmission opportunity u , the delay of bandwidth request d and the drop probability of bandwidth request p_d are presented. Based on simulation results, the effects of contention parameters such as the initial window W , the number of subscriber station n and the number of transmission opportunity N_{to} are discussed. Some policies to adjust the parameters are also provided. These policies are useful for BS to schedule its uplink bandwidth resource.

Key words: IEEE 802.16; contention resolution scheme; binary exponential backoff; performance analysis; bandwidth request

摘要: 目前, IEEE 802.16 标准推荐采用基于截断二进制指数后退算法的竞争解决方案. 分析了该方案同 IEEE 802.11 竞争机制的区别, 给出了传送机会利用率 u 、带宽请求延时 d 以及带宽请求丢失率 p_d 等性能指标的计算方法. 通过性能模拟, 讨论了初始化后退窗口 W 、用户站数目 n 以及单位时间帧内传送机会数目 N_{to} 等参数对性能指标的影响, 进而得出基站调整性能参数的一般策略. 这些策略对于基站进行上行带宽资源的分配具有指导意义.

关键词: IEEE 802.16; 竞争解决方案; 二进制指数后退; 性能分析; 带宽请求

中图法分类号: TP301 文献标识码: A

随着网络技术的快速发展, 无线宽带接入技术逐渐成为当前无线通信技术领域最活跃的因素. WiMAX

* Supported by the National Basic Research Program of China under Grant No.2003CB317003 (国家重点基础研究发展计划(973)); the Strategy Grant of City University of Hong Kong of China under Grant Nos.7001709, 7001777 (香港城市大学战略发展计划)

Received 2006-04-04; Accepted 2006-06-30

(world interoperability for microwave access)作为一项新兴的无线宽带接入技术,能够向用户提供高带宽、远距离的无线访问连接.它在高速Internet接入、军事、医疗等领域具有非常广阔的应用前景^[1].2004年10月,IEEE在IEEE 802.16(2001)和IEEE 802.16a的基础上正式发布了IEEE Std. 802.16-2004^[2],为WiMAX网络的媒体接入控制层和物理层提供了详细的规范.

IEEE 802.16标准支持PMP(point-to-multipoint)和Mesh两种模式.在PMP模式下,无线城域网由一个基站(base station,简称BS)和多个用户站(subscriber stations,简称SS)组成,基站控制网络中所有数据的收发,用户站要发送数据,首先必须通过初始化维护(initial maintenance)与基站建立连接,然后向基站发送带宽请求(bandwidth request),基站成功收到带宽请求后,根据特定的资源调度算法给用户站分配带宽资源.此时,用户站才能发送数据.初始化维护和带宽请求两个阶段都涉及到多个用户站对上行信道的竞争,因而有效的竞争解决方案对于整个网络系统的性能具有至关重要的意义.

由于初始化维护和带宽请求两个阶段采用相同的竞争解决方案,本文以带宽请求为例,分析带宽请求竞争解决方案的实施性能.一方面,可以得到网络运行的性能指标,如带宽请求的延时、带宽请求的丢弃概率等,找出影响这些指标的参数,进而优化这些参数可以得到更好的系统性能;另一方面,竞争解决方案的性能也是基站分配带宽的重要依据.在一个时间帧(time frame)内,带宽请求部分和数据发送部分共享上行子帧,它们之间也是一种“竞争关系”(如图1所示),基站只有知道竞争部分的性能后才能给出合理的上行信道分配策略.

目前,IEEE 802.16标准推荐采用基于截断二进制指数回退算法的竞争解决方案^[1].二进制指数回退算法是一种分布式函数(distributed coordinated function).已有众多的文献讨论了二进制指数回退算法的性能和改进.在无线局域网中,文献[3]给出了基于Markov模型的饱和性能分析,讨论了无线局域网络吞吐量与二进制指数回退参数的关系;文献[4]在文献[3]的基础上考虑了二进制指数回退算法的截断特性,给出了更准确的性能计算;文献[5]将性能分析推广到非饱和状态,给出了饱和与非饱和状态下的性能分析;文献[6]对二进制指数回退算法进行改进,提出了一种自适应的回退算法;而文献[7]则给出了一般意义上的指数回退算法的分析.然而,由于IEEE 802.16采用了完全不同于IEEE 802.11的MAC结构,其竞争机制也具有很大的不同,目前还很少有文献讨论IEEE 802.16竞争解决方案的性能.文献[8]中给出的帧时隙(framed slotted)类似无线城域网的MAC结构,文献[9]给出了一种无线城域网中带宽请求分析模型,但是它们都没有基于二进制指数回退算法来分析.

本文分析了IEEE 802.16和IEEE 802.11竞争机制的差异,并基于这些差异定义了传送机会(transmission opportunity,简称TO)利用率 u 、带宽请求延时 d 等性能评价指标.通过计算任意一个用户站在任意一个传送机会发送带宽请求的概率 τ ,分析了单位时间帧内基站提供的传送机会数目 N_{to} 对竞争解决方案性能的影响.在此基础上,给出各个性能指标的计算方法,进而讨论了初始化回退窗口 W 、单位时间帧内传送机会数目 N_{to} 、用户站数目 n 等参数之间的关系以及这些参数对系统性能的影响.通过模拟,验证了理论分析并得出了基站调整竞争参数的一般性策略.

1 IEEE 802.16中的竞争解决方案

1.1 IEEE 802.16中的竞争过程

在IEEE 802.16中,数据收发以时间帧为单位周期进行.每一时间帧由下行子帧(downlink subframe)和上行子帧(uplink subframe)两部分组成,基站通过下行子帧将数据发给用户站,如图1所示.用户站则通过上行子帧向基站发送带宽请求和数据.每一上行子帧包括初始化维护、带宽请求竞争以及数据发送三部分.具体到带宽请求竞争部分,则又包含多个传送机会.传送机会是带宽请求竞争的基本单位.各个用户站通过UL_MAP消息获取时间帧的结构信息,然后各自独立地发送自己的带宽请求.

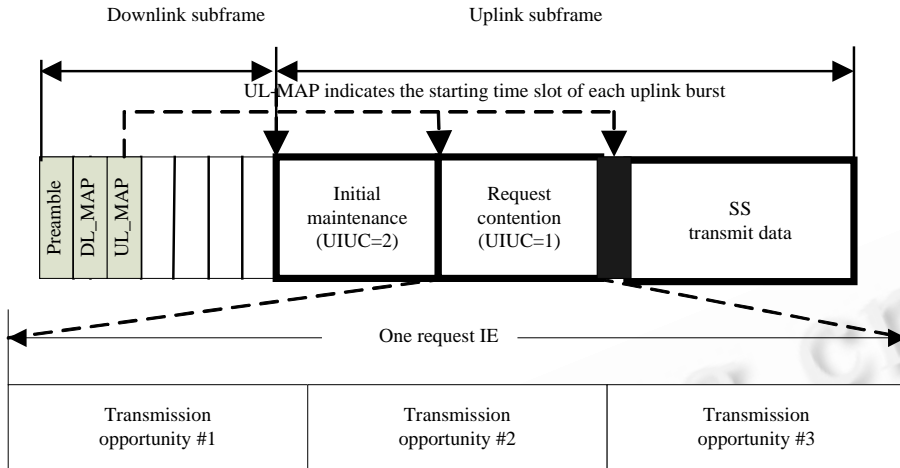


Fig.1 Structure of transmission opportunity

图 1 传送机会结构

在发送带宽请求之前,用户站通过读取基站发送的上行链路描述(uplink channel describer)消息获取初始化回退窗口 W_{min} 和最大回退窗口 W_{max} 等参数,然后它在 $[0, W_{min}]$ 范围内随机地选取一个回退数(backoff value) B_i ,每经过一个传送机会, B_i 减 1,当 B_i 减到 0 时,用户站将在下一个传送机会发出带宽请求.各个用户站带宽请求之间的竞争可以通过如图 2 所示的例子来说明.

设基站在 3 个连续的时间帧内分别给出了 3,4,3 个传送机会(图中数字 1~10).假定 $W_{min}=15$,按随机选取的策略,设用户站 SS_1 选择的回退数为 7,则表示它要放弃 7 次传送机会,也就是说,在前两个时间帧它不发送请求,一直要到第 3 个时间帧的第 1 个传送机会(对应图中第 8 个传送机会)才发送.设用户站 SS_2 在第 2 个时间帧开始参与竞争,假定它选择的回退数为 5,则它将放弃从第 2 个时间帧开始的 5 次传送机会,即将在图中第 9 个传送机会发送带宽请求.设用户站 SS_3 选择的回退数是 8,则它也在第 9 个传送机会发送请求.不难看出,在第 8 个传送机会,由于只有 SS_1 发送带宽请求,故其请求能被正确接受,而在第 9 个传送机会,由于 SS_2 和 SS_3 同时发送请求,它们的请求将发生冲突,冲突发生后,卷入冲突的用户站将按二进制指数回退过程来扩大竞争窗口.

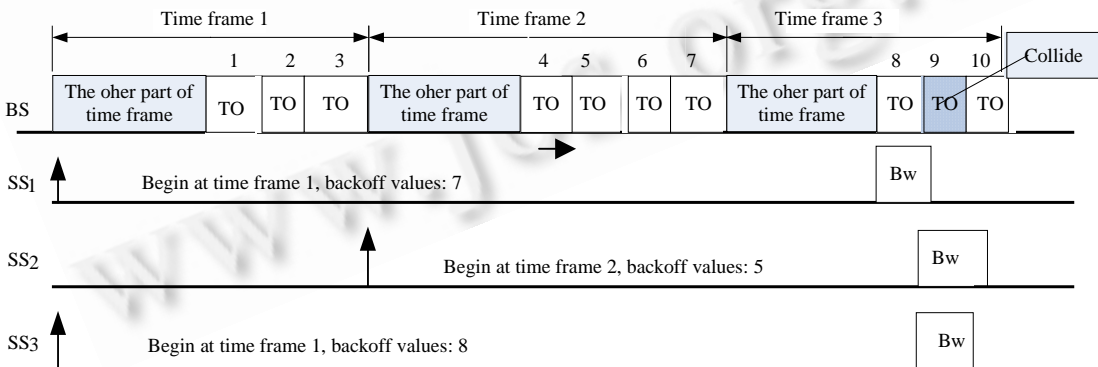


Fig.2 Example of bandwidth contention process

图 2 带宽请求竞争示意图

设 i 是用户站的带宽请求被成功发送之前发生冲突的次数,则新的回退窗口可按式(1)来设置.

$$W_i = \begin{cases} 2^i W & i \leq m' \\ 2^{m'} W & m' < i \leq m \end{cases} \quad (1)$$

其中, $W=W_{min}+1, 2^{m'} W=W_{max}+1$.

简单地说,就是发生冲突后,用户站将其回退窗口扩大为原来的 2 倍,然后再在 $[0, W_i]$ 范围内随机选取回退数,当回退数递减到 0 后,用户站再次发送带宽请求,如果带宽请求又发生了冲突,则继续按式(1)扩大回退窗口.当回退窗口达到 W_{\max} 后,回退窗口将保持不变,如此反复,直到冲突次数达到最大重传次数 m ,用户站或丢弃该带宽请求,或交由高层协议处理为止.

1.2 IEEE 802.16与IEEE 802.11竞争机制的差别

由IEEE 802.16 的带宽请求竞争过程可以看出,虽然IEEE 802.16 采用类似于IEEE 802.11 中的竞争冲突解决策略,但是,由于它们MAC层的差异,两者的竞争机制存在一些差别^[10].

(1) IEEE 802.16 中引入了帧的概念,虽然用户站竞争的是传送机会,但用户站是以时间帧为周期来参与竞争的.当用户站在某个时间帧内产生带宽需要时,它至少要等到该时间帧结束以后才能发送带宽请求,并且基站在每个时间帧内提供的传送机会数目 N_{to} 对系统也有较大的影响.而IEEE 802.11 不需要这种等待,当某个终端准备好数据时,它就马上参与到竞争过程中.

(2) IEEE 802.11 的终端在发送数据前会检测信道是否空闲,如果信道被占有,则该终端的回退数就保持不变;只有信道在 DIFS 时间内为空闲时,终端才将回退数减 1.而 IEEE 802.16 没有信道检测过程,每经过一个传送机会,回退数就减 1,当回退数为 0 后,用户站不管信道是否处于空闲状态,都会发送带宽请求.显然,检测信道是否占用,可以降低发生冲突的概率.

(3) 两者的载荷不同.IEEE 802.11 每次发送都是一个数据包,这个数据包的大小与具体应用有关,当数据包较大时,其发送时间和竞争时间的比值就相对较大,因而吞吐量就比较大.特别是 IEEE 802.11 还可以采用 RTS/CTS 模式,通过 4 次握手过程,让两个终端在一段时间内独占整个信道,从而减少竞争的开销比例.而在 IEEE 802.16 中,参与竞争的带宽请求是一个固定大小的小数据包.

(4) 冲突的检测方法不同.IEEE 802.11 终端接收到数据后,会马上给发送方一个 ACK 信号,以表示数据成功接收.如果在规定的时间内,发送方没有收到 ACK 信号,则发送方认为发送冲突.而 IEEE 802.16 中用户站发送带宽请求后不能马上得到确认信息,它需要等到下一个时间帧,通过读取 UL_MAP 查看基站是否给自己分配带宽,才知道上次请求是否成功.

(5) 控制方法不同.在IEEE 802.11 中,各个用户站都分布式地进行通信,因而控制信号和数据没有分开,各个终端的初始回退窗口事先指定.而IEEE 802.16 中,控制信号和数据信号是分离的^[10],带宽请求的竞争是针对控制信号进行的,竞争的冲突不直接影响数据信道,因此,可以单独分析控制信号的性能.此外,IEEE 802.16 中的基站具有集中调控功能,用户站和基站之间可以保持同步关系,基站可以修改接入参数调整用户站带宽请求的发送,从而可以更加灵活地控制用户站之间的竞争.

2 性能分析

2.1 性能评价指标

由于 IEEE 802.16 和 IEEE 802.11 在 MAC 层存在着较大的差异,因此我们也必须采用不同的性能指标来衡量 IEEE 802.16 的实时性能.从基站的角度看,它关心的是发出的传送机会能否正确接收带宽请求;而对用户站而言,它则强调能否尽快将带宽请求发送给基站以尽量避免带宽请求的丢失.针对基站和用户站不同的侧重点,可定义以下性能指标:

(1) 传送机会利用率 u :指一个传送机会能正确接收到用户站带宽请求的概率.由于用户站各自独立地发送带宽请求,一个传送机会收到的带宽请求有 3 种可能:0 个带宽请求、1 个带宽请求和多于 1 个带宽请求.如果带宽请求数为 0,则该传送机会被浪费了;如果带宽请求数大于 1,则多个带宽请求发生冲突,基站也不能正确收到带宽请求.因此,传送机会利用率 u 就是任意一个传送机会恰好收到一个带宽请求的概率.

(2) 带宽请求延时 d :指从用户站产生带宽请求到带宽请求被成功发送到基站所经历的时间.它反映了带宽请求的接入性能.由于带宽请求竞争只是时间帧的一部分,与时间帧相比,传送机会是一个非常小的数据,并且

在一个时间帧内,无论用户站选择哪个传送机会,它都要等到下一个时间帧才知道竞争结果.因此,我们忽略带宽请求在时间帧内的延时,以时间帧为单位来衡量一次带宽请求所需要的时间.

(3) 带宽请求丢弃率 p_d :指用户站的带宽请求由于冲突次数超过某特定值而被丢弃的概率.当系统中同时有太多的带宽请求时,让一些用户站放弃带宽请求发送,可以增加其他用户站带宽请求成功的机会.因此,请求丢弃率 p_d 是基站控制带宽请求接入的一种手段.

这些性能指标都是随机变量,其瞬时值与具体过程有关,计算比较困难,并且瞬时的值对系统意义不大.在本文中,我们主要考虑系统处于稳定运行状态时的性能,即考察当竞争参数给定时,网络系统长时间运行时各个性能指标的平均值.

2.2 不考虑 N_{th} 时用户站的发送概率

为了借鉴IEEE 802.11 中的研究结果,我们先考虑一种极端情况.假定一个时间帧内基站只提供一个传送机会,则IEEE 802.16 的竞争过程演化为一般意义下的二进制指数回退过程.类似于文献[3-5]中的分析,当系统运行在一定的参数时,假定无论用户站是第一次发送带宽请求还是经过几次冲突后再发送带宽请求,该带宽请求发生冲突的概率都是 p_c .这里, p_c 是用户站数目 n 的函数,用户站数目 n 越大,则冲突的概率越大.通过这一假定,用户站数目 n 对系统性能指标的影响将通过 p_c 来体现.根据二进制指数回退算法的竞争过程,任意一次带宽请求有两种可能,或是经过 $i(i=0,1,\dots,m)$ 次冲突后成功发送,或是经过 $m+1$ 次冲突后被丢弃.不难发现,丢弃的概率 p_d 为

$$p_d = P_c^{m+1} \tag{2}$$

设 P_i 为带宽请求发送 i 次冲突后,在 $i+1$ 次发送成功的概率.由于 P_i 是没有发生丢弃的条件概率,因此有

$$P_i = \frac{(1-p_c)}{1-p_c^{m+1}} p_c^i \tag{3}$$

记用户站发生 i 次冲突后再次参与竞争为用户站处在状态 i .当用户站处在状态 i 时,它将在 $[0, W_i]$ 的范围内随机地选取一个回退数,因此平均说来,用户站处在状态 i 时,其带宽请求需要经过 $(W_i+1)/2$ 个传送机会才能发送.换句话说,用户站平均在状态 i 停留 $(W_i+1)/2$ 个传送机会.记 S_i 为用户站处于状态 i 的概率,则 S_i 可以计算为

$$S_i = \frac{P_i d_i}{\sum_{j=0}^m P_j d_j} \tag{4}$$

式中, d_i 为用户站在状态 i 的平均停留时间.根据二进制回退过程, $P_i d_i$ 可以计算为

$$P_i d_i = \begin{cases} \frac{(2^i W + 1) p_c^i (1 - p_c)}{2(1 - p_c^{m+1})}, & 0 \leq i \leq m' \\ \frac{(2^{m'} W + 1) p_c^i (1 - p_c)}{2(1 - p_c^{m+1})}, & m' < i \leq m \end{cases} \tag{5}$$

将式(5)代入 $\sum_{j=0}^m P_j d_j$,有

$$\sum_{j=0}^m P_j d_j = \sum_{j=0}^{m'} \frac{(2^j W + 1) p_c^j (1 - p_c)}{2(1 - p_c^{m+1})} + \sum_{j=m'+1}^m \frac{(2^{m'} W + 1) p_c^j (1 - p_c)}{2(1 - p_c^{m+1})} \tag{6}$$

化简整理有

$$\sum_{j=0}^m P_j d_j = \frac{W(1 - p_c)(1 - (2p_c)^{m'+1}) + (1 - p_c^m)(1 - 2p_c) + 2^{m'} p_c^{m'+1} W(1 - 2p_c)(1 - p_c^{m-m'})}{2(1 - 2p_c)(1 - p_c^{m+1})} \tag{7}$$

记 $s_{i,k}$ 为用户站处在状态 i 时,回退状态值为 k 的概率.注意,回退状态值和回退数不是同一个概念,当用户站选择某个回退数 B_i 后,每经过一个传送机会,回退状态值减1,也就是回退状态值的取值为 $B_i, B_i-1, B_i-2, \dots, 0$.由于 B_i 是在 $[0, W_i]$ 内均匀取值,当 B_i 取值分别为 $0, 1, \dots, W_i$ 时,相应的回退状态值为 $(0), (1, 0), (2, 1, 0), \dots, (k, k-1, \dots, 0), \dots, (W_i, W_i-1, \dots, 0)$.将不同 B_i 值时所有相同的状态值累加,可得到回退状态值分别为 $0, 1, \dots, W_i$ 时各个状态值出现的次数为 $W_i, W_i-1, \dots, 0$,即状态值为 k 出现的次数为 W_i-k .计算各种状态值占有所有状态值的比例就可

得到

$$s_{i,k} = S_i \frac{W_i - k}{\sum_{i=1}^{W_i} i} = \frac{2(W_i - k)}{W_i(W_i + 1)} S_i \tag{8}$$

将 $k=0$ 代入,则有

$$s_{i,0} = \frac{2}{(W_i + 1)} S_i = \frac{P_i}{\sum_{j=0}^m P_j d_j} = \frac{(1 - p_c) p_c^i}{\sum_{j=0}^m P_j d_j (1 - p_c^{m+1})} \tag{9}$$

$s_{i,0}$ 表示用户站发生了*i*次冲突后,在*i+1*次竞争过程中回退状态值到了0,在下一个传送机会,用户站将发送带宽请求.令 τ 表示对于任意传送机会,用户站发送带宽请求的概率.由于 $s_{i,0}(i=0,1,\dots,m)$ 是独立事件,将不同状态时用户站发送概率累加就可以得到 τ .

$$\tau = \sum_{i=0}^m s_{i,0} = \sum_{i=0}^m \frac{(1 - p_c) p_c^i}{\sum_{j=0}^m P_j d_j (1 - p_c^{m+1})} = \frac{1}{\sum_{j=0}^m P_j d_j} \tag{10}$$

将式(7)代入式(10)有

$$\tau = \frac{2(1 - 2p_c)(1 - p_c^{m+1})}{W(1 - p_c)(1 - (2p_c)^{m+1}) + (1 - p_c^m)(1 - 2p_c) + 2^{m'} p_c^{m+1} W(1 - 2p_c)(1 - p_c^{m-m'})} \tag{11}$$

这一结果与文献[4]中的基于 Markov 过程分析的结论是一致的.

2.3 考虑 N_{to} 时用户站的发送概率

公式(11)给出了 $N_{to}=1$ 时任意用户站在任意一个传送机会发送带宽请求的概率.增大 N_{to} 但是保持 $N_{to} \leq W$,此时,虽然用户站是以时间帧为单位进行竞争,但对于任意一个时间帧内的任意传送机会,任意一个用户站都可能发送带宽请求,因此,这里仍然可以假定用户站在每一个传送机会发送带宽请求的概率为常数.然而,当 $N_{to} > W$ 时,就不能假定为常数概率了.

如图3所示,将 N_{to} 按 $W, 2W, 4W, \dots, 2^L W$ 划分为 $[0, W], [W+1, 2W], [2W+1, 4W], \dots, [2^{L-1}W+1, 2^L W], [2^L W+1, N_{to}]$,共 $L+1$ 个区间,其中, $L = \lfloor \log_2(N_{to}/W) \rfloor$.对于区间 $[0, W]$,无论用户站的竞争窗口是多少,它的带宽请求都有可能落在该区间内的任意一个传送机会,因此,任意用户站在该区间内的任意一个传送机会发送带宽请求的概率 τ 仍然可以按公式(11)来计算.然而在区间 $[W+1, 2W]$ 范围内,对于那些第一次参加竞争的用户站来说,由于其竞争窗口是 W ,它的带宽请求不可能落到 $[W+1, N_{to}]$,即它们不可能在序号大于 W 的传送机会发送带宽请求.同样,经历一次冲突的带宽请求也不可能落在 $[2W+1, N_{to}]$.如图3所示,用户站的发送概率将随传送机会的序数增大而逐渐减少.

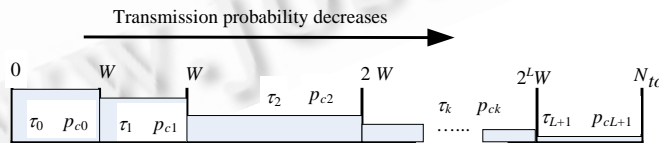


Fig.3 Transmission probability decrease when $N_{to} > W$

图3 当 $N_{to} > W$ 时发送概率逐渐减少

由于用户站在各个区段内的发送概率不一样,则每区段内发生冲突的概率也不一样,但是在一个区段内,仍假定任意一次带宽请求发生冲突的概率为 p_{ck} ,记 τ_k 为用户站的带宽请求在 $[2^{k-1}W+1, 2^k W]$ 发送的概率,则有

$$\tau_k = \sum_{i=k}^m s_{i,0} = \sum_{i=k}^m \frac{(1 - p_{ck}) p_{ck}^i}{\sum_{j=0}^m P_j d_j (1 - p_{ck}^{m+1})} = \frac{(1 - p_{ck}^{m+1-k}) p_{ck}^k}{(1 - p_{ck}^{m+1})} \tau \tag{12}$$

当 n 个用户站以概率 τ_k 在 $[2^{k-1}W+1, 2^k W]$ 内发送带宽请求时,某个用户站在某个传送机会发送请求而不发生

冲突的概率是其他 $n-1$ 个用户站都不发送请求的概率,即 $(1-\tau_k)^{n-1}$,也就是发生冲突的概率为

$$p_{ck}=1-(1-\tau_k)^{n-1} \tag{13}$$

注意,式(13)中,当 $k=0$ 时可得到 p_c 和 τ 的关系,即 $p_c=1-(1-\tau)^{n-1}$,将其代入式(11)就可以解出 p_c 和 τ ,将 τ 代入式(12),联立式(12)和式(13)可以求解出 τ_k 和 p_{ck} .

2.4 传送机会利用率 u

对任意一次传送机会,当有且仅有一个用户站发送带宽请求时,它能成功收到带宽请求.因此,从 n 个用户站中任取一个发送带宽请求,其他 $n-1$ 个用户站都不发送的概率就是该传送机会的利用率,对于区间 $[2^{k-1}W+1, 2^k W]$,其传送机会利用率 u_k 可计算为

$$u_k = C_n^1 \tau_k (1-\tau_k)^{n-1} = n \tau_k (1-\tau_k)^{n-1} \tag{14}$$

当 $N_{to} \leq W$ 时,式(14)得到的是单一值,而当 $N_{to} > W$ 时,由式(14)得到的是分段的利用率.将各个段的利用率求和平均,可得到一个时间帧内的带宽请求平均利用率.综合两种情况有

$$u = \begin{cases} u_0, & N_{to} \leq W \\ \frac{Wu_0 + \sum_{k=1}^L 2^{k-1} Wu_k + (N_{to} - 2^L W)u_{L+1}}{N_{to}}, & N_{to} > W \end{cases} \tag{15}$$

其中, L 为对 N_{to} 的分段数.

为了求出传送机会利用率的上限,我们对式(14)关于 τ_k 求导并令导数为 0,可得当 $\tau_k=1/n$ 时, u_k 取最大值,此时对应的最大 u_k 为 $(1-1/n)^{n-1}$,将其代入式(15),很容易得到, u 的最大值也是 u_k 的最大值,即有

$$u_{\max}=(1-1/n)^{n-1} \tag{16}$$

$(1-1/n)^{n-1}$ 是关于 n 的递减函数,它表明了用户数目与传送机会利用率之间的关系,用户数目越大,传送机会最大利用率越小.当 n 趋近无穷时,它的极值为 $1/e$.从这个数值来看,IEEE 802.16 的竞争解决方案中传送机会的利用率并不算高,如当 $n=5$ 时,这一数值就已经降到 0.409 6.

2.5 延时 d 的计算

对单个用户站来说,延时 d 就是用户站发送带宽请求时需要放弃的传送机会的期望值,这个期望值与带宽请求成功发送之前所经历的冲突次数有关,而冲突的次数又取决于冲突的概率.根据前面的分析,当 $N_{to} \leq W$ 时,冲突的概率可假设为单一值,而当 $N_{to} > W$ 时,考虑到在时间帧的不同区段用户站的带宽请求发生冲突的概率不同,为了简便起见,这里对区段冲突概率求平均值,类似于带宽请求平均利用率,有

$$p_c = \frac{Wp_{c0} + \sum_{k=1}^L 2^{k-1} Wp_{ck} + (N_{to} - 2^L W)p_{cL+1}}{N_{to}} \tag{17}$$

将式(17)代入式(3)可得到带宽请求需要经历 i 次冲突才成功的概率 P_i .当冲突次数为 0 时,由于用户站的回退数是在 $[0, W]$ 的范围内随机选取的,因而回退数的期望值为 $(W+1)/2$;而当冲突次数为 1 时,用户站发送了两次带宽请求,第一次发送时用户站的期望回退数为 $(W+1)/2$,而第 2 次发送时由于回退窗口变成了 $2W/2$,它回退的期望值为 $(2W+1)/2$,即总的回退期望值为 $(W+1)/2+(2W+1)/2$.考虑最大回退窗口 W_{\max} 和有限重传次数 m ,当冲突次数为 i 时,回退数的期望值为

$$E(B(i)) = \begin{cases} \sum_{k=0}^i \frac{2^k W + 1}{2}, & i \leq m' \\ \sum_{k=0}^{m'} \frac{2^k W + 1}{2} + \frac{2^{m'}(i-m')W + 1}{2}, & i > m' \end{cases} \tag{18}$$

综合式(3)和式(18),可以得到用户站成功发送带宽请求前需要放弃的传送机会的期望值:

$$E(TO) = \frac{(1-p_c)}{1-p_c^{m+1}} \left(\sum_{i=0}^{m'} p_c^i \sum_{k=0}^i \frac{2^k W + 1}{2} + \sum_{i=m'+1}^m p_c^i \left(\sum_{k=0}^{m'} \frac{2^k W + 1}{2} + \frac{2^{m'}(i-m')W + 2}{2} \right) \right) \quad (19)$$

而最终的延时还取决于基站在每个时间帧内给出的传送机会数目 N_{to} 。目前,IEEE 802.16 标准并未给出传送机会数目的设定方法,因此,我们采用平均值来计算。记 \tilde{N}_{to} 为一个时间帧内基站提供的传送机会平均数目,则平均延时为

$$\tilde{d} = \frac{(1-p_c)}{(1-p_c^{m+1})\tilde{N}_{to}} \left(\sum_{i=0}^{m'} p_c^i \sum_{k=0}^i \frac{2^k W + 1}{2} + \sum_{i=m'+1}^m p_c^i \left(\sum_{k=0}^{m'} \frac{2^k W + 1}{2} + \frac{2^{m'}(i-m')W + 2}{2} \right) \right) \quad (20)$$

3 模拟结果

由于推出IEEE 802.16 标准时间不长,目前还没有专门的工具来模拟和分析。我们用C++编写IEEE 802.16 带宽请求竞争过程的模拟程序。它主要由基站、用户站和Simulator这 3 类组成,基站在每个时间帧发出 N_{to} 个传送机会;用户站产生带宽请求后,按二进制指数回退算法参与竞争;Simulator负责生成基站和用户站对象、协调基站和用户站之间的同步并记录模拟结果。为了降低其他因素对模拟结果的影响,类似于文献[3,7]中的方法,模拟程序采用理想无线信道,即没有考虑信道错误和俘获效应。

考虑到竞争的随机性,所有的模拟都经过 1 000 000 个时间帧,每次模拟主要记录以下数据:1) 基站发出的传送机会的总数 S_{to} ;2) 基站成功收到的带宽请求总数 S_s ;3) 所有成功带宽请求花费时间的总和 T ;4) 所有用户站丢弃的带宽请求总数 S_d ;5) 所有用户站产生的带宽请求的总数 S_{bw} 。根据这些数据,各个性能参数的模拟结果计算方法为

$$u = \frac{S_s}{S_{to}}, \quad d = \frac{T}{S_s}, \quad p_d = \frac{S_d}{S_{bw}} \quad (21)$$

3.1 初始化窗口 W 对性能的影响

从第 3 节的分析可知, u 与 W, n, m, m' 以及 N_{to} 相关。考虑到用户站发送带宽请求时需要知道初始化窗口 W ,我们首先考察 W 对系统性能的影响。在评价 W 作用时为了尽量降低 N_{to} 的影响,我们固定 $N_{to}=4$,并且让 $N_{to}<W$,而对于其他参数,由于目前IEEE 802.16 并没有给出标准值,我们参考了IEEE 802.11 中的典型值,固定 $m'=5, m=8$,分别取 n 为 10,50 和 200 这 3 组数据,让 W 从 4 变化到 1 024,图 4~图 6 分别给出了 u, d 和 p_d 的模拟结果以及计算的结果。

从图 4~图 6 可知,当 $N_{to}<W$ 时,通过式(11)~式(14)和式(20)计算得到的理论值与模拟结果十分吻合。从曲线的变化趋势可以得出以下结论:1) 带宽请求的延时与活动用户站数目 n 以及 W 直接相关,活动用户站数目 n 越大,延时 d 越大,而当活动用户站数目一定时,增大窗口 W ,延时 d 也随之增大;2) 当其他参数给定时,存在一个优化窗口 W_{opt} ,使得传送机会的利用率最大,而且 W_{opt} 与 n 相关, n 增大, W_{opt} 也相应地需要增大。当 $N_{to}<W$ 时,将发送概率 $\tau_k=1/n$ 代入到式(11)和式(13),可得到 W_{opt} 和 n 之间满足以下关系:

$$\begin{cases} \frac{1}{n} = \frac{2(1-2p_c)(1-p_c^{m+1})}{W_{opt}(1-p_c)(1-(2p_c)^{m+1}) + (1-p_c^m)(1-2p_c) + 2^{m'} p_c^{m'+1} W_{opt}(1-2p_c)(1-p_c^{m-m'})} \\ p_c = 1 - (1-1/n)^{n-1} \end{cases} \quad (22)$$

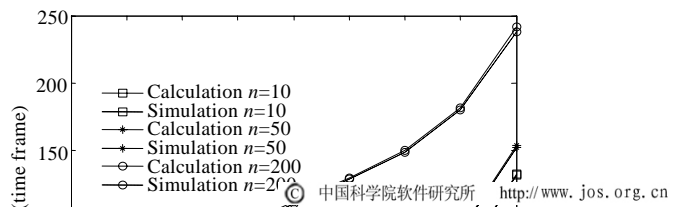
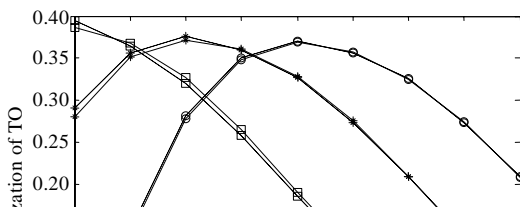


Fig.4 Effect of TO utilization on initial window

图 4 TO 利用率和初始化窗口的关系

Fig.5 Effect of delay on initial window

图 5 延时和初始化窗口的关系

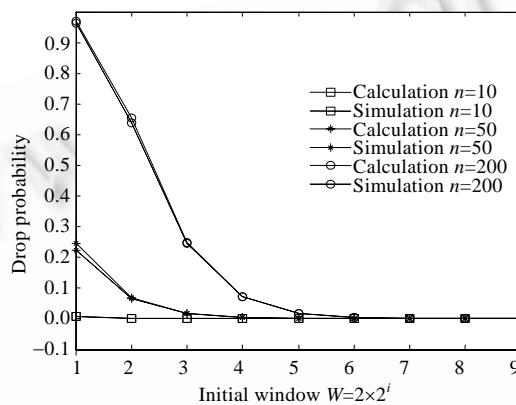


Fig.6 Effect of drop probability on initial window

图 6 带宽请求丢失率和初始化窗口的关系

3.2 传送机会数目 N_{to} 对性能的影响

图 4、图 5 也表明当活动用户站数目 n 比较大时,过小的 N_{to} 导致了较大的延时.例如,当 $n=200, W=64$ (此时对应的利用率最大) 时,带宽请求的平均延时接近 100 个时间帧,显然这不是一个小数值.从式(20)可以看出,延时 d 和 N_{to} 呈反比关系,为了降低延时,基站需要在单一时间帧内提供更多的传送机会,而增大 N_{to} 则会对利用率 u 带来一定的影响.根据式(22),当 n 分别为 50 和 200 时,相应地, W 取 32 和 64 时传送机会利用率比较高.因此,在考察 N_{to} 影响时,我们取 $n=50, W=32$ 和 $n=200, W=64$ 两组数据,同样固定 $m'=5, m=8$, 让 N_{to} 从 4 变化到 1 024.图 7 是 u 随 N_{to} 的变化曲线.为了便于比较,图 7 中也给出没有考虑 N_{to} 影响时的两组计算值.

从图 7 中可以看出,当没有考虑 N_{to} 影响时,利用率是平直线.不难发现,当 $N_{to} > W$ 时,计算值与模拟结果之间存在较大的误差;而考虑 N_{to} 的影响后,理论计算和模拟结果基本吻合.从理论分析和模拟结果的曲线来看,当 $N_{to} < W$ 时,增大 N_{to} , 利用率基本保持不变;而当 $N_{to} > W$ 后,增大 N_{to} , 利用率开始下降.因此,盲目地增大 N_{to} 并不一定能提高系统的性能.

利用率的降低显然不是系统所期望的.根据前面的分析,由于初始窗口 W 比 N_{to} 小时,那些序数大于 W 的传送机会被用户站选中的机会偏小,最终导致了 N_{to} 增大时利用率 u 减少.为此,可以考虑增大 N_{to} 的同时增大初始窗口.图 8 和图 9 分别给出了 $n=50, W=16, 32, 64, 128$ 时 u 和 d 的变化曲线(其他参数与上面的分析一样).由图 8 可知,当 N_{to} 大到一定程度时,适当地增大 W 可以得到更高的利用率,如当 $N_{to}=64$ 时(图中 $i=5$ 时),在利用率上, $W=32$ 和 $W=64$ 比 $W=16$ 要好,而当 $N_{to}=128$ 时, $W=64$ 和 $W=128$ 利用率比 $W=16$ 和 $W=32$ 利用率要高.从图 9 可以看出,增大 W 对延时影响不是很大.如当 $N_{to}=64$ 时,不同 W 情况下延时基本差不多.

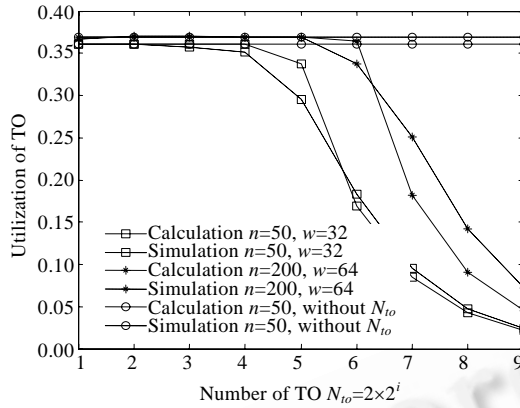


Fig.7 Effect of N_{to} on utilization

图7 N_{to} 对利用率的影响

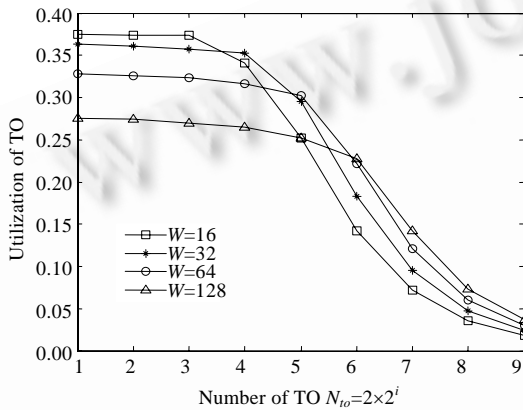


Fig.8 Effect of TO Utilization on N_{to}

图8 TO利用率和 N_{to} 的关系

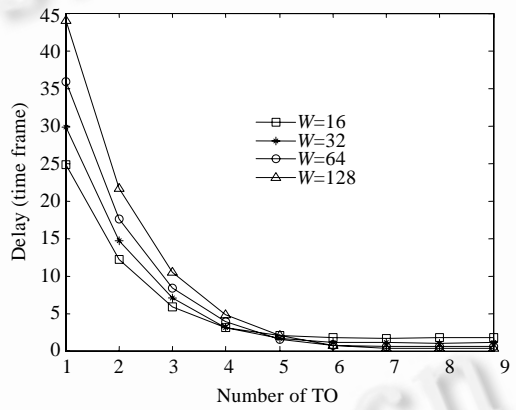


Fig.9 Deffect of Delay on N_{to}

图9 延时和 N_{to} 的关系

3.3 系统性能与用户站数目的关系

对实际应用来说,人们更关心用户站的数目与系统性能的关系.固定 $N_{to}=W$, W 分别取 8,16,32,图 10 和图 11 分别是利用率 u 和延时 d 随用户数目的变化曲线.对于一个给定的 W ,由图 10 可知,利用率会先增大,到达一个最大值后,逐渐减小.从图 10 中也可以看出 W 和 n 之间的关系,当 n 比较小时,小的 W 具有较好的利用率;而 n 较大时,选择大的 W 利用率更高.由图 11 可知,随着 n 的增大,延时也逐渐增大,对于同一个 n , W 越大,延时越小.究其原因,是因为 $N_{to}=W$,延时更多地取决于冲突的概率, W 增大,冲突的概率减小,从而延时会随 W 的增大而减小.

上面讨论了 W,N_{to} 以及 n 与 u 和 d 之间的关系,从整个竞争过程来看, m 和 m' 也对系统的性能产生影响,特别是对带宽请求丢弃率 p_d .但当 N_{to} 固定后, m 和 m' 的分析过程与IEEE 802.11 中的分析并没有实质上的差别,因此具体的分析过程可以参考文献[3-5].

从整个模拟的结果来看,竞争解决方案的性能由多个参数决定,由于IEEE 802.16 中的基站可以灵活地调整这些参数,因此,有效的调整策略对系统的运行具有重要的意义.目前,IEEE 802.16 并没有提供具体的调整策略,我们初步总结为:首先,估算系统中需要发送数据的用户站的数目,然后根据用户站的数目以及式(22)大致确定 W 的数值.这里, m 和 m' 可先按典型值设置.其次,综合考虑各个用户站的带宽请求延时 d 和带宽请求丢弃率 p_d .如果要求带宽请求延时较小,则可以考虑减小 W ,同时也可以考虑减小 m ;反之,如果要求带宽请求延时较大,则可以

考虑采用较大的 W 和 m .最后,根据基站提供的 N_{to} 数目对 W 进行调整,这里假定 N_{to} 是由其他因素如带宽分配算法决定的,调整的策略是:如果 N_{to} 比 W 大,则将 W 设为与 N_{to} 一样大.

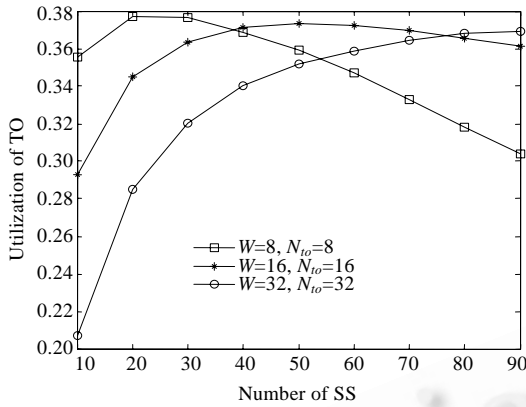


Fig.10 Effect of TO utilization on n

图 10 TO 利用率和用户站数目 n 的关系

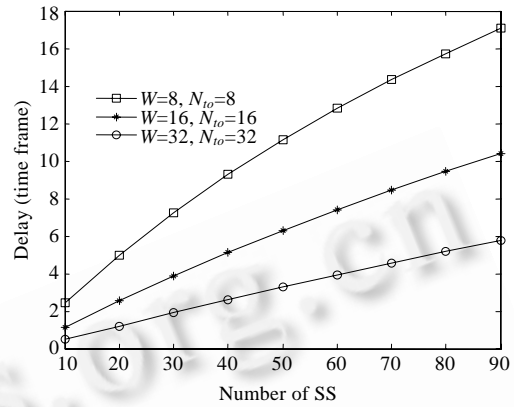


Fig.11 Effect of delay on n

图 11 延时和用户站数目 n 的关系

4 结论

由于IEEE 802.16 的竞争解决方案结合了二进制指数回退算法和Framed Slotted ALOHA的特点,因此需要采用与分析IEEE 802.11 不同的方法进行分析.本文通过计算任意一个用户站在任意一个传送机会发送带宽请求的概率,分析了单位时间内基站提供的传送机会数目对性能的影响,对IEEE 802.16 竞争解决方案的基本性能进行了分析,并给出了 N_{to} 与 W 以及 n 之间的制约关系.同时,通过分析带宽请求发送概率,本文得到各个性能参数对性能指标的影响,并给出了性能参数调整的基本策略.这些策略对于基站进行上行信道的资源分配具有现实的指导意义.

从IEEE 802.16 的带宽请求竞争过程来看,采用基于二进制指数回退算法的竞争解决方案,传送机会的利用率并不是很高.因此,结合其他带宽请求发送机制如单播轮询来提高传送机会利用率将是本文下一步研究的重点.同时,针对无线信道,如MIMO 以及俘获效应等进行分析和模拟也将是进一步研究的重点.

References:

- [1] Ghosh A, Wolter DR, Andrews JG, Chen RH. Broadband wireless access with WiMax/802.16: Current performance benchmarks and future potential. IEEE Communications Magazine, 2005,43(2):129-136.
- [2] IEEE 802.16-2004 IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. IEEE Press, 2004.
- [3] Bianchi G. Performance analysis of the IEEE 802.11 distributed coordination function. IEEE Journal on Selected, Areas in Communications, 2000,18(3):535-547.
- [4] Wu H, Pang Y, Long K, Cheng S, Ma J. Performance of reliable transport protocol over IEEE 802.11 wireless LAN: Analysis and enhancement. In: Kermani P, Bauer F, Morreale P, eds. Proc. of the 21st Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM 2002). New York: IEEE Press, 2002. 599-607.
- [5] Lin LD, Fu HH, Jia WJ. An efficient admission control for IEEE 802.11 networks based on throughput analyses of (un)saturated channel. Proc. of the IEEE Globecom, 2005,1(5):3017-3021.
- [6] Peng Y, Cheng SD. A self-adaptive wireless LAN protocol. Journal of Software, 2004,15(4):604-615 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/15/604.htm>
- [7] Kwak BJ, Song NO, Miller LE. Performance analysis of exponential backoff. IEEE/ACM Trans. on Networking, 2005,13(2): 343-355.

- [8] Wieselthier JE, Ephremides A, Michales LA. An exact analysis and performance evaluation of framed ALOHA with capture. IEEE Trans. on Communications, 1989,37(2):125-137.
- [9] Oh SM, Kim JH. The analysis of the optimal contention period for broadband wireless access network. In: Proc. of the the 3rd Int'l Conf. on Pervasive Computing and Communications Workshop. 2005. 215-219.
- [10] Cao M, Ma WC, Zhang Q, Wang XD, Zhu WW. Modeling and performance analysis of the distributed scheduler in IEEE 802.16 mesh mode. In: Proc. of the Int'l Symp. on Mobile Ad Hoc Networking and Computing. 2005. 78-89.

附中文参考文献:

- [6] 彭泳,程时端.一种自适应无线局域网协议.软件学报,2004,15(4):604-615. <http://www.jos.org.cn/1000-9825/15/604.htm>



陆文彦(1973-),男,湖南岳阳人,博士生,主要研究领域为无线城域网,接入控制,SIP协议.



杜文峰(1977-),男,博士,主要研究领域为无线网络,Handoff,带宽调度算法.



贾维嘉(1957-),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为组播,选播,路由,无线宽带网络,移动多媒体通信,分布式系统.



张历卓(1974-),男,博士生,主要研究领域为多媒体通信.