

基于无线网络编码的信道利用率研究

徐 奎, 戴 彬, 黄本雄, 张 波

(华中科技大学 电子与信息工程系, 湖北 武汉 430074)

摘 要: 为了使无线随机网络编码策略更好适用于较大数据量的传输要求, 在无线信道的广播特性的基础上加入了将数据分块处理, 并考虑了额外开销对信道利用率的影响. 通过对分块传输模型的分析, 确认信道利用率与分块的大小紧密相关, 为了在通信前选择一个适当的分块大小以获得更高的信道利用率, 给出了一种计算最优数据块大小的方法. 仿真结果表明: 该方法简单可行, 使用该方法选择数据块大小后, 信道利用率将十分接近其上限值.

关键词: 网络编码; 无线网络; 信道利用率; 分块大小

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2010)01-0127-04

Channel Utilization Based on Wireless Network Coding

XU Kui, DAI Bin, HUANG Ben-xiong, ZHANG Bo

(Department of Electronics and Information Engineering,
Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: To make the wireless network coding scheme fit for large data transmission well, the processing of dividing data into blocks is added into the scheme. After the analyzing the transmission model of the blocks, it is found that channel utilization and the size of the block is closely related. In order to transmit the data with an appropriate block size to get higher channel utilization, a block size selection approach is given to maximize the channel utilization. The approach is simple and feasible, and by the simulations, after using this approach, the channel utilization will be very close to its upper limit.

Key words: network coding; wireless networking; channel utilization; block size

1 引言

网络编码(Network Coding)是由 R. Ahlswede 等人在 2000 年提出的一种概念^[1], 结合了编码与路由的思想. 随后, Koetter 等人通过代数方法提出了随机网络编码^[2], 将网络编码的应用延伸到任意网络结构中. M. Medard 等人^[3-4]充分结合了无线信道的广播特性, 提出了一套高效、可靠的无线网络编码策略. 文献[5-6]则推荐将文件先分成较小的数据块, 然后再对各个块分别进行随机网络编码的操作.

文中考虑到额外开销对信道利用率的影响, 从理论上分析了基于分块的随机编码策略下的信道利用率, 着重研究了数据块大小与信道利用率之间的

关系, 并给出了一种计算数据块大小使得信道利用率最大化的方法.

2 随机网络编码策略

利用网络编码策略通信之前, 首先要选择一个适当的子网 H 作为通信子网. 为了刻画无线信道的广播特性, 通信网络建模成一个有向超图(Directed hypergraph) $H = (N, A)$, 其中 N 为所有节点的集合, 而 A 为所有超弧(Hyperarc)的集合. 超图和普通图的不同主要在于超图拥有超弧, 超弧可以表示为一个二元组 (i, J) , 其中 i 是 N 的一个元素, 表示弧的起点, J 则是 N 的一个子集, 表示弧的终点的非空集合. 显然, 一般的图可以看成超图的一个特例. 每条超弧 (i, J) 可以表示一条从节点 i 到集合 J 的一

条广播链路. 设 $A_{iJK}(\tau)$ 为经过时间 τ , 集合 $K \subset J$ 通过超弧 (i, J) 收到的数据量. 定义 A_{iJK} 的平均速率为:

$$f_{iJK} = \lim_{\tau \rightarrow \infty} A_{iJK}(\tau) / \tau \quad (1)$$

可利用线性规划的方法找到流向量 f 使得源节点到目的节点的传输速率达到 R . 中继子网的选择根据具体对数据传输需求不同, 在式(2)的基础上可再添加各种约束条件, 例如, 可以通过选取适当的通信子网来达到节省网络能量的目的^[7].

$$\sum_{j|j|(i,j) \in A} \sum_{j \in J} f_{ijj} - \sum_{j|j|(j,I) \in A, i \in I} f_{jji} = \begin{cases} R, & i = S \\ 0, & \text{其他} \\ -R, & i = T \end{cases} \quad (2)$$

确定通信子网和通信流向量后 f , 就在指定的通信子网上按照再对利用随机网络传送各块的编码策略. 在传输之前, 首先将要传输的文件分为 l 个数据块 $\{B_1, B_2, \dots, B_l\}$, 其中每个数据块又由 K 个数据包组成, K 称为数据块的“尺寸”, 假设所有的数据块的尺寸相等, 数据包的随机编码方式可表述为:

$P'_m = \sum_{i=1}^K c_{mi} P_i$, 其中 $c_m = [c_{m1}, c_{m2}, \dots, c_{mK}]$ 称为数据包 P'_m 的“全局编码向量”(Global Coding Vector), 向量的所有元素都依均匀分布随机从有限域 F_q 内选取. 在传输过程中数据包中将需要在包头添加字段用于记录其相应的全局编码向量, 这些附加字段的长度为 $K \log_2 q$ 比特, 这些字段的线性相关性等价于数据包的线性相关性^[5]. 中继节点缓存接收到的线性无关数据包, 一旦链路条件允许就会将之前缓存的数据包进行编码后在其邻域中广播. 而目的节点能解码时就会发送确认包来通知源节点开始下一数据块的发送.

3 数据块尺寸与信道利用率

没有使用分块策略时, 所有的中间节点将需要大量的缓存来存放接收到的组成整个文件的全部 K 个线性无关数据包, 这会造成中间节点缓存资源需求过大的问题. 而使用分块策略后, 中间节点每次只用缓存一个数据块的而不是整个文件, 缓解了缓存资源需求过大的问题. 但是, 数据块的尺寸不能选择过大, 数据块越大, 传送该数据块所需要的数据包会越多(K 的值越大), 对应的, 每个数据包需要在其包头用于记录全局编码向量的字段也会更长($K \log_2 q$ 比特), 由此额外的开销在数据包中占用的比例也会增大. 无线随机网络编码策略要利用无线

信道的广播特性, 因此编码需要在链路层进行. 由于通信链路层的限制, 数据包的长度往往都有上限要求. 例如, 在 802.11b 中, 为了不使数据包因过长而使其在传输过程中的出错概率过大, 数据包的最大长度限定为 2346 比特, 若此时要用某数据块传输大小为 1MB 的数据, 编码有限域大小为 255. 若该数据块尺寸 K 的值取为 1786, 经过计算可得, 这时有有效信息的长度为 560 byte ($1786 \times 560 = 10^6$, 且 $1786 + 560 = 2346$), 即每个数据包的包头需要加入长度为 1786 比特的字段来记录全局编码向量. 此时数据包的额外开销占总长度的比例高达 76.2%, 这是对信道资源极大的浪费.

然而, 选择的数据块尺寸也不能过小, 随着尺寸的变小, 要传输的文件需要分成的数据块会更多. 由于在每解码一个数据块之后目的节点都需要向源节点发送一个 ACK 确认信息, 过多的确认信息同样会降低信道利用率. 例如, 在极端情况下, 当 $K = 1$ 时, 数据块只由一个数据包 P_1 组成, 利用随机编码后, 在传送这个数据块时, 所有的编码后的数据包都将表示成单个数据包 P_1 与自身的线性组合 $c_1 P_1$. 这时, 每个数据包都需要确认一次, 随机网络编码策略将退化成低效率的 ARQ 协议, 无线网络编码的信道利用率也会变得很低.

因此, 信道利用率与 K 值的选择紧密相关, 数据分块尺寸的选择为一个十分重要的问题. 记信道利用率为 $U(K)$, 基于分块的随机网络编码, 于是, 计算数据块尺寸 $\tilde{K}$ 使得信道利用率最大化, 称 $\tilde{K}$ 为数据块的“最优数据块尺寸”, 求解最优数据块尺寸的目标方程可表示为:

$$\tilde{K} = \arg \max_K [U(K)] \quad (3)$$

4 最优数据块尺寸

4.1 信道利用率

确定通信子网, 通信流向量 f , 以及通信速率 R 后, 利用随机网络编码策略将数据块 B_i 从源节点 S 发送到目的节点 T 的过程可以用图 1 表示.

图 1 中, t_i^B 为源节点 S 发送组成数据块 B_i 的 K 个数据包所需要的时间, 且有 $E(t_i^B) = K t^P$, 其中 $t^P = L/R$, L 为数据包的长度, 只考虑数据包头中用于记录全局编码向量的字段, 而包头中其它字段的长度则认为很小并可忽略不计. 在选定通信子网和流向量后, 传输速率 R 为定值.

t_i^D 是传播时延, 为信号在链路上传播造成和各

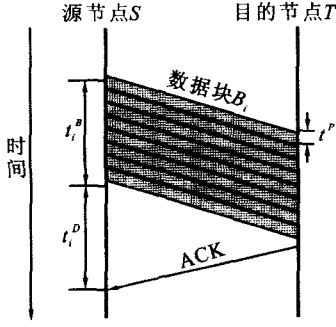


图1 数据块 B_i 发送到目的节点 T 的过程

中间节点造成的时延,其大小主要由链路状态(如:拥塞程度和节点的移动性)决定.假设链路状态稳定,且 $E(t_i^D)$ 为常数.

信道利用率(Channel utilization)也可称为带宽利用效率(Bandwidth utilization efficiency),定义为有效数据的吞吐量与网络传输比特率的比值.进行线性随机网络编码的有限域为 F_q ,数据块尺寸为 K .观察图1,随机网络编码策略下的信道利用率可表示为^[8]:

$$U(K) = \left[\frac{(L - K \log_2 q)}{L} \right] \left[\frac{E(t_i^B)}{E(t_i^B + t_i^P)} \right] \quad (4)$$

记 $a = \log_2 q$ 和 $E(t_i^D) = D$.则式(4)可写为:

$$U(K) = \frac{(L - aK)Kt^P}{L(Kt^P + D)} \quad (5)$$

从式(5)可以看出,传输文件采取基于分块的随机网络编码策略以后,信道利用率与数据块的尺寸紧密相关而待传输的文件总大小没有直接的联系.

4.2 最优数据块尺寸的计算

对式(5)中的 K 求一阶、二阶导数后可得:

$$\frac{\partial U}{\partial K} = \frac{-(at^P K^2 + 2aDK - DL)t^P}{(Kt^P + D)^2} \quad (6)$$

$$\frac{\partial^2 U}{\partial K^2} = -2t^P \frac{D(Kt^P + D)(aD + t^P L)}{(Kt^P + D)^4} \quad (7)$$

观察式(7),当 $K > 0$ 时,有 $\partial^2 U / \partial K^2 < 0$.因此, $K > 0$ 时, $U(K)$ 存在极大值.

进而,求解方程 $\partial U / \partial K = 0$ 即可求得 $U(K)$ 的极值点,舍去其中 $K < 0$ 的解后,可得到满足式(3)要求的最优数据块尺寸 $\tilde{K}$:

$$\tilde{K} = \frac{\sqrt{a^2 D^2 + aDLt^P} - aD}{at^P} \quad (8)$$

综合前面的推导, $K > 0$ 时,有且仅有一个可通过式(8)计算出来的最优数据块尺寸 $\tilde{K}$.利用式(8)求最优数据块尺寸 $\tilde{K}$ 需要输入的参数有 a 、 D 、 t^P 、 L .

4.3 参数选择

对于参数 a , T. Ho 等人证明,在随机编码过程中,目的节点成功解码的概率会随着有限域的减小而变小,为了保证解码成功率,有限域 F_q 的大小 q 不能太小,因此,一般不建议通过减小 a 的途径来提高信道利用率.

对于参数 L 和 t^D ,由式(5),提高数据包长度 L 是能提高信道利用率的.由于无线链路对数据包的长度的限制.而且在通信子网确定之后,数据传输率 R 将随之确定,由 t^P 的定义,这时 t^P 由数据包只由长度 L 确定.因此,建议为了获得更高的信道利用率,数据包的长度 L 应设置为该无线网络允许的最大包长.

式(8)的各参数中,只有传播时延的期望 D 是不能直接设定的.从图1可以看出,传播时延 t^D 可以近似等于一个数据包的全往返时间(RTT)减去该数据包的发送时间.由于讨论的链路均状态稳定,于是,在实际操作中, D 的值可以通过在通信之前从源节点向目的节点发送 N 个探测包,并利用探测包全的往返时间使用式(9)来进行估计.

$$\hat{D} = \frac{1}{N} \sum_{i=1}^N (t_i^{\text{RTT}} - L_{\text{prob}}/R) \quad (9)$$

其中, L_{prob} 为探测包的长度, t_i^{RTT} 为探测包 i 的全往返时间.

5 仿真与结果

为了研究数据块尺寸与文件传输的吞吐量之间的关系,并讨论求最优数据块尺寸时的参数选择问题.在仿真过程,使用 802.11b 为 MAC 层,并将网络编码的数据包长度设为 802.11b 的最大包长: $L = 2346$ 字节.编码有限域 F_q 的大小设为 2^8 ,即随机网络编码的基本编码符号为 $a = 1$ 字节.根据式(2)建立通信子网 H 和通信流向量 f 后,可使通信的速率 R 达到 2Mb/s.

没有选用最优数据块传输数据时,选用不同的数据块大小 K 将在传输过程中对应不同的信道利用率 $U(K)$.如图2所示,与第4.2节推导的结论一致: $K > 0$ 时, $U(K)$ 存在极大值.而且 $K < 179$ 时,信道利用率会随 K 的增大能得到快速的提升.在 $K = 179$ 时信道利用率达到峰值 0.85,随后,信道利用率将逐渐降至 0,即当 $K = 2346$,数据包需要用等于包长的 2346 字节用于记录全局编码,这时数据包中没有任何有效负载.

利用式(8)得到最优数据块尺寸 $\tilde{K} = 185$.仿

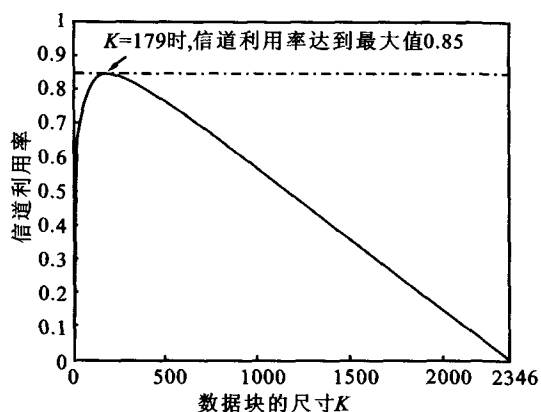


图2 信道利用率与数据块尺寸

真中,探测包的个数 $N = 100$,并在传输一个大小为1GB的文件时,利用这些探测包的全往返时间,利用求出的最优数据块尺寸为185时的平均信道利用率达为0.84,只是略小于最大信道利用率。

由式(9),传播时延的期望 D 可以通过选择用不同数量的探测包来估计。图3为选择不同数量探测包时文件传输的信道利用率。在 $N < 15$ 时,吞吐量会随 N 的增大而增大,随后的信道利用率几乎不受 N 值大小的影响,而且总体来说,信道利用率随 N 的变化非常小,最大信道利用率和最小信道利用率之差不到0.003。因此,在当前仿真环境下, N 的值选择在20左右就已经足够。

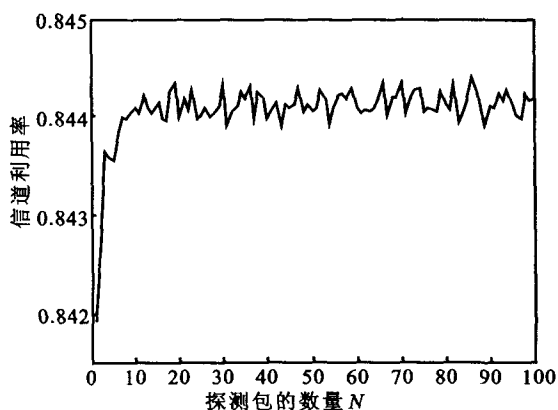


图3 探测包的数量和信道利用率

参考文献:

- [1] Ahlswede R, Cai N, Yeung R, et al. Network information flow[J]. IEEE Transactions on Information Theory, 2000, 46(4):1204-1216.
- [2] 葛晓虎,朱耀庭. 无线局域网帧数据流量研究[J]. 微电子学与计算机, 2008, 25(1):13-16.
- [3] Lun S, Medard M, Koetter R, et al. On coding for reliable communication over packet networks[J]. Physical Communication, 2008, 1(1):3-20.
- [4] Lun S, Medard M, Koetter R, et al. Further results on coding for reliable communication over packet networks[C]// Proceeding of ISIT. Adelaide, Australia: IEEE, 2005:1848-1852.
- [5] Philip Chou, Yunnan Wu. Network coding for the internet and wireless networks[R]. Microsoft Science Report, 2007.
- [6] Sachin Katti, Dina Katabi. Wireless network coding: opportunities and challenges[C]// Proceeding of MILCOM. USA:Florida, AFCEA, 2007:1-8.
- [7] Tao Cui, Lijun Chen, Tracey Ho. Energy efficient opportunistic network coding for wireless networks[C]// IEEE INFOCOM. California, 2008:361-365.
- [8] 段中兴,张德运. 基于模糊逻辑的无线网络流量自适应控制[J]. 微电子学与计算机, 2006, 23(1):31-35.

作者简介:

徐奎男, (1979-), 博士. 研究方向为无线传感器网络以及网络编码相关技术.

戴彬男, (1977-), 博士. 研究方向为无线网络、P2P技术.

黄本雄男, (1966-), 教授, 博士生导师. 研究方向为通信理论、通信信号处理、通信对抗.

张波男, (1984-), 博士. 研究方向为认知无线电以及网络编码相关技术.