

缓存门限自适应调整的智能手机节能研究

蒋 波¹ 李陶深^{1,2} 葛志辉^{1,2}

(广西大学计算机与电子信息学院 南宁 530004)¹

(广西高校并行与分布式计算技术重点实验室 南宁 530004)²

摘 要 为了解决 LTE 无线流媒体终端能量利用率低的问题,提出一种新的基于动态缓存门限调整的流媒体传输控制算法(DBTA)。该算法根据监测到的网络带宽和流媒体编码率信息,对终端缓存区下限阈值做出自适应的调整,使缓存空间大小更契合当前带宽和流媒体编码率,提高了终端的休眠工作时间比,进而延长了终端的续航时间。仿真结果表明,该算法可有效解决 LTE 无线流媒体终端能耗过快的的问题,最多可降低约 22% 的能耗。

关键词 LTE, 移动多媒体流, 节能, 自适应缓存

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.1.031

Research of Smartphone Energy Saving Based on Buffer Threshold Adaptive Adjustment

JIANG Bo¹ LI Tao-shen^{1,2} GE Zhi-hui^{1,2}

(School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China)¹

(Guangxi Colleges and Universities Key Laboratory of Parallel and Distributed Computing Technology, Nanning 530004, China)²

Abstract To solve the low energy efficiency problem of LTE wireless streaming media terminal, a new energy-efficient streaming media transmission control method based on dynamic buffer threshold adjustment (DBTA) was proposed. This algorithm adaptively adjusts the buffer lower threshold of receiver according to the monitored network bandwidth and streaming media encoding rate information, making the size of the buffer space more fit the current bandwidth and streaming media encoding rate, and improves the time ratio of dormant to work, thereby extending the life of the terminal. Simulation results show that this algorithm (DBTA) can effectively resolve the problems of the excessive energy consumption of LTE wireless streaming media terminal, and achieve up to 22% energy saving.

Keywords LTE, Mobile multimedia streaming, Energy saving, Buffer adaptive

1 引言

智能手机能耗过快是亟待解决的问题^[1]。随着硬件工艺的改进和无线网络技术的成熟,在智能手机上观看流媒体视频已成为一种广泛应用的无线数据业务。然而,视频等多媒体业务的广泛应用在大大提升手机用户性能体验的同时,也急剧地消耗着手机的电量,尤其是流媒体大数据长时间持续传输的特性会严重缩短手机续航时间,影响用户使用满意度。

近年来,人们为智能手机设计了许多节能优化方案,但针对 LTE 流媒体数据传输的解决方案很少,相关的研究成果不多。移动流媒体与其它 Web 服务相比,具有其明显的通信特性^[2],即每次会话分组传输量大,数据帧间间隔固定,具有较大的时间延迟容忍度等等。针对移动流媒体通信能耗过高的问题,人们提出了一些改进机制。文献[3]提出一个 MAC 层的优化方案,将多个待传输视频帧绑定为一个大的视频块传输,利用帧聚合增大每次传输的数据量,获得了一定的节能效果;文献[4]提出一种利用 DRX (discontinuous reception) 降低用户设备功耗的视频流调度算法,其在满足数据包时延要

求的同时,减少了 Inactivity Timer 的开启次数,增加了用户设备的休眠时间。文献[5]把能耗优化问题归结成最优资源分配问题,提出的算法根据终端软硬件能力分配稀缺资源来提高总体用户的能耗利用率,但该算法缺乏公平性考虑,可能造成单个终端一直得不到所需资源而饿死;文献[6]提出一种基于能量感知的自适应动态速率算法,即通过选择合适的流媒体编码速率保证流媒体能在所剩电量可持续的时间内流畅播完;文献[7]针对 LTE 系统恒定功率休眠模式中存在的低业务、高功耗等缺点,提出一种变功率休眠算法。然而,上述优化机制都未考虑用户设备硬件缓存能力有限的特性,节能和保质往往顾此失彼,在节省电量的同时不可避免地降低了性能体验。

为适应智能手机硬件受限的特性,研究者们提出利用缓存节能的思想。文献[8,9]针对速率不匹配的组件,提出在组件间插入缓存,使得速率较快的组件周期性休眠以降低能耗,并从理论上推导出需插入缓存的最佳大小。但此类方法对于缓存大小已经固定的移动终端并不适用。文献[10]提出了一种新的自适应缓存节能解决方案(AB-PSM),即通过在服务

到稿日期:2014-11-09 返修日期:2015-04-01 本文受国家自然科学基金项目(61363067),广西自然科学基金资助项目(2012GXNSFAA053226)资助。

蒋 波(1989-),男,硕士生,CCF 会员,主要研究方向为移动计算;李陶深(1957-),男,博士,教授,主要研究方向为无线 Mesh 网络、分布式数据库、云计算等,E-mail:tsli@gxu.edu.cn;葛志辉(1978-),男,博士,教授,主要研究方向为无线网络、网络性能优化。

器应用层额外设置一个较大的缓存,所有发往移动基站的数据先在此缓存中暂存,等该缓存填满后再将聚合的数据分组一次性发往基站,以此减少基站发送唤醒终端的 beacon 信令帧的频率。但该方案忽略了硬件的成本问题,因此难以得到推广。文献[11]提出一种基于接收端 TCP 缓存意识的跨层优化方案 EStreamer,即依据从模型和客户端播放缓存状态得出的启发来确定突发数据尺寸,在不危害流畅播放的前提下,实现尽可能小的电量消耗。

造成智能手机流媒体传输能耗严重的主要原因是流媒体需要传输的数据量大,而每次传输的数据量有限,使得无线数据端口长时间处于活跃状态或频繁地进行状态切换。目前,流媒体传输为平滑带宽变动对播放造成影响,通常在用户端设置一个较大的缓存区,由两个阈值门限触发数据传输,且两个阈值门限在整个传输会话期间保持不变。文献[12]指出固定缓存大小策略是以确保平滑播放为导向,往往比较保守,而并非能耗最优,动态调整缓存大小可获得能耗收益,但却没有给出具体的实现方案。显然,动态缓存自适应调整在无线流媒体节能应用中还处于初步研究阶段,有很多问题亟待解决。本文在此背景下研究设计一种新的基于动态缓存下限阈值调整的移动终端节能算法,该算法在保证用户业务体验的同时,降低了手机的能耗。

2 问题的描述

如图 1 所示,客户端缓存采用循环队列结构来存储数据包;缓存中的数据包先进先出,前端按客户端的处理速率即视频解码器的解码速率 $D(t)$ 从缓存取出数据,后端按网络上实际到达的数据速率 $A(t)$ 把接收的数据读入缓存;通过丢弃已经播放的内容,可以重新利用空出的缓存空间来暂存后继尚未播放的内容。

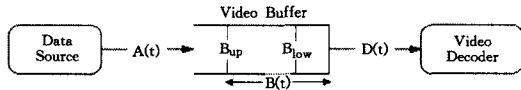


图 1 视频缓存模型

当前,移动流媒体普遍使用一种基于阈值驱动的内容传输触发机制,其主要原理是:播放终端维持两个缓存阈值 (B_{up}, B_{low}),同时时刻监控自身缓存占有率。当缓存量 $B(t)$ 低于下限阈值 B_{low} 时,发出数据传输请求,开始接收数据;当缓存量达到上限阈值 B_{up} 时,发出终止请求,终止数据传输。数据传输呈开关(ON-OFF)特性。

图 2 是从播放请求开始至播放达到稳定期间缓存数据量的理论变化情况示意图。图中,横坐标表示时间(单位 s),纵坐标表示当前时刻的缓存数据量(单位 MB)。

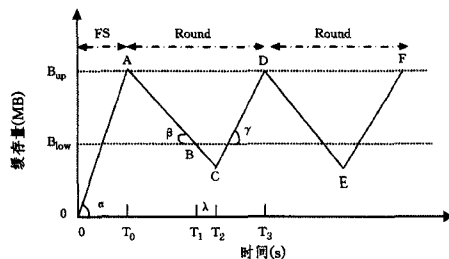


图 2 缓存数据量随时间的变化

整个数据传输过程分为两个阶段:FS 和 Round。

快传输(FS)阶段: $0 \sim T_0$,持续时间 T_0 。缓存量从 0 到达 B_{up} (图中的 A 点),此阶段播放尚未开始,忽略背景流等额外开销,图中的角 α 约等于缓存分组到达速率 $A(t)$ 。

Round 阶段:正常播放阶段,定义缓存量从 $B_{up} \rightarrow B_{low} \rightarrow B_{up}$ 为一个优化周期(Round),对应图 2 中的 A-B-C-D 线段。在 AB 段,数据端口关闭,缓存消耗速率 $D(t)$ 可用角 β 近似表示。AB 段对应时间跨度为 $(T_1 - T_0)s$; T_1 时刻,当前缓存量 $B(t) = B_{low}$,触发接收端数据端口重新开启,由于端口重启和网络传输有一定的时延,缓存不会立即接收到数据,而在此期间播放仍一直进行,故缓存量还会继续下降,此过程中的缓存变化情况如图 2 所示的 BC 段。经过时延 λs ,缓存开始接收到新的数据;再经过 $(T_3 - T_2)s$,缓存到达 B_{up} 。角 γ 可近似表征缓存量的变化率 $D(t) - A(t)$ 。整个数据传输过程如此反复,直至所有内容传输完毕或用户人为中断。

以往有关缓存的研究主要关注于如何平滑网络带宽的抖动,降低网络拥塞概率,提升缓存利用率等,对终端能耗关注较少。由于无线信道的时变性和随机性,往往将下限阈值 B_{low} 设定得比较高,以确保不会出现缓存空的情况。从缓解因分组突发造成缓存队列过度拥塞或过度空闲方面来说,设置一个较高的下限阈值是合适的。但是,从能耗的角度来看,过高的下限阈值会导致频繁的端口开启,甚至会使得端口持续处于开启状态,产生不必要的状态转换能耗。

分析可知,缓存队列的长度应由当前的分组到达速率、视频播放速率、时延参数共同决定。在分组到达速率变大、播放速率不变的情况下,适当减小 B_{low} 并不会带来缓存空的危险,同时又可以延长端口的休眠时间,带来可观的能耗收益。

3 算法设计与描述

针对 LTE 流媒体的节能需求,本文提出动态缓存门限调整算法(Dynamic Buffer Threshold Adjustment, DBTA)。算法的设计思路是:实时监测网络带宽、流媒体编码率和缓存占有量,根据监测到的网络带宽和流媒体编码率信息,对接收终端缓存区下限阈值做出自适应的调整,使缓存空间大小更契合当前带宽和流媒体编码率,提高终端的休眠工作时间比,进而延长终端的续航时间。DBTA 算法包括数据监测、自适应缓存下限调整两个部分。

3.1 数据监测

算法需要监测的数据有 3 个:网络带宽、流媒体编码率和重启时延。网络带宽决定了缓存分组到来的速率 $A(t)$,流媒体编码率决定了数据流出缓存的速率 $D(t)$,重启时延与下限阈值的设定紧密相关。为了顺畅播放,流媒体要求视频帧播放的速率必须等于其编码率,而视频块的编码率对终端来说是已知的,故只需要记录视频块的编码率及视频块的数目,就可以得到缓存流出速率 $D(t)$ 。另外,接收终端的时延 λ 一般可以认为是一个与特定系统有关的值,这里把它看作一个常量。关键是需要监测缓存数据到来的速率 $A(t)$ 。

媒体流的瞬时速率是时刻变动的,记录一定时间 t 内接收的数据包个数及其大小 b ,就可计算出媒体流的平均速率 $A(t)$ 。考虑网络具有自相似性,采用加权平均预测算法得到下一测量周期内的媒体流平均速率。为此,引入一个关于测

量周期内媒体流历史速率的变量 \overline{A}_h ,用于记录近 num 次(本文取 $num=6$)测量周期数据接收速率的均值,然后同本周期内速率 \overline{A}_n 进行加权求和,来预测下一周期缓存分组到达速率。设第 $n+1$ 个测量周期内媒体流的平均速率为 \overline{A}_{n+1} ,则 \overline{A}_{n+1} 可表示为:

$$\overline{A}_{n+1} = (1-\delta)\overline{A}_h + \delta\overline{A}_n \quad (1)$$

其中, δ 是加权系数,且 $0 \leq \delta < 1$ 。测量周期为 2s 时,取 $\delta=0.25$ 。

3.2 自适应缓存下限调整

对接收端来说,假定缓存上限阈值固定,也就是说在会话开始时播放器分配了一个固定大小的缓存空间,下限阈值决定了每次下载的内容块大小、下载周期(ON)频率和持续时间。下限越小,缓存数据再装事件发生的频率越低,平均消耗的电量越少。但是,下限阈值设置得越低,播放暂停风险越高,尤其是再装阶段网络出现低带宽的情形。因此,下限缓存阈值的选取需在降低缓存下溢风险和节省能量间进行折中权衡。

为此,定义变量 k 表示缓存下溢风险, k 满足:

$$k = \sqrt{\frac{D\lambda}{B_{low}}} \quad (2)$$

其中, \overline{D} 表示平均播放速率, λ 表示端口开启至缓存实际读入数据之间的时延, B_{low} 为下限阈值。 k 的取值范围为 $k \in (0, 1]$,当 $k > 1$ 时,视作 $k=1$ 。 k 值越大,表示缓存下溢的可能性越大。根据实验,将 k 划分成 4 个级别来表示不同的危险程度,不同危险程度采用不同的调节因子以防止缓存下溢,如表 1 所列。

表 1 下溢风险等级

k	等级
(0.8, 1]	严重危险
(0.5, 0.8]	中度危险
(0.3, 0.5]	轻度危险
(0, 0.3]	安全

下限阈值调整公式如下:

$$B_{low}^{new} = B_{low}^{old} + \omega * \frac{B_{up} - B_{low}^{old}}{D} \quad (3)$$

$$\omega = \frac{\sum_{i=n-2}^n \overline{A}_i - \overline{A}_{i-1}}{\overline{A}_{i-1}} \quad (4)$$

其中, B_{low}^{new} 为计算得到的新下限阈值; B_{low}^{old} 为上一轮(Round)的下限阈值; ω 为增量因子(可为正数或负数),用以调整阈值增加或减少的程度; B_{up} 为缓存上限阈值,整个实验中保持不变; \overline{D} 为平均播放速率。式(4)表明增量因子 ω 用近 3 个测量周期缓存数据读入速率相对变化情况表示。

出于无线网络经常变动和避免缓存下溢的考虑,需要对使用式(3)和式(4)计算出来的下限阈值做出修正,修正的依据是新下限阈值对应的下溢风险,风险越大,修正的步长越大。将修正后的下限阈值作为下一轮最终的缓存下限阈值。修正公式如下:

$$B_{low}^{new} = \begin{cases} B_{low}^{new}, & k \in (0, 0.3] \\ \beta_1 B_{low}^{new}, & k \in (0.3, 0.5] \\ \beta_2 B_{low}^{new}, & k \in (0.5, 0.8] \\ \beta_3 B_{low}^{new}, & k \in (0.8, 1.0] \end{cases} \quad (5)$$

式中, $\beta_1, \beta_2, \beta_3$ 为不同危险情况下的修正步长,且 $1 < \beta_1 < \beta_2 <$

β_3 ,步长既不能太大也不能太小,实验中取 $\beta_1=1.5, \beta_2=2, \beta_3=2.5$ 。

3.3 算法的描述

DBTA 算法描述如下:

Step1 初始化参数。

Step2 监测缓存分组到达速率 $A(t)$ 、离开速率 $D(t)$ 及缓存占有量 $B(t)$ 。

Step3 根据式(3)、式(4)计算出优化的缓存下限阈值。

Step4 根据式(2)计算由 Step3 得到的缓存下限阈值所对应的危险系数 k 。

Step5 根据危险系数 k 的范围,对计算出的下限阈值进行修正:若 k 落在安全的范围,用计算所得的缓存下限阈值更新 B_{low}^{new} ;若 k 落在危险的范围内,则根据危险情况对计算所得的下限值做不同程度的调整更新,并返回 Step4。

4 仿真结果与性能分析

4.1 仿真实验环境设置

实验平台为 Windows7+VMware9.0+ Ubuntu 12.04+ ns3.19,仿真拓扑结构如图 3 所示。



图 3 仿真拓扑结构

为简化起见,设置 4 个节点,其中 S 为源端,业务流使用固定码流(CBR)流量模式产生数据,最大数据包为 1500B;N2 节点为基站,与源端 S 通过有线链路经 N1 连接;D 为客户端,N2 和 D 间通过无线链路连接。假设有线链路带宽足够大,不构成瓶颈,链路丢包率足够小,可以忽略。

实验的主要目的是测试缓存下限阈值调整对终端的节能效果及所提新算法的有效性。为此,设置了两个实验场景。在场景 1 中,节点位置相对静止,设置 4 种不同的缓存数据到达速率和消耗速率,研究不同下限阈值设定对能耗的影响情况。在场景 2 中,S,N1 和 N2 相对静止,客户端 D 在 N2 节点信号覆盖范围内随机移动,D 距离 N2 越近,信号越强,缓存分组读入速率越大,反之亦然。在相同条件下,分别对静态缓存门限(SBT)和动态缓存门限调节(DBTA)进行仿真,性能评价指标为传输播放同等大小流媒体消耗的电量。仿真时间为 300s。

B_{up} 和 B_{low} 的初值由经验值决定, B_{up} 在实验运行过程中保持不变,而 B_{low} 可动态地根据网络状况做出自适应的调整。本文实验中设定 $B_{up}=0.9B_{max}$, B_{low} 的初值为 $0.5B_{max}$ 。带宽检测时间间隔为 2s。其他实验参数如表 2 所列。

表 2 实验参数

参数名称	参数说明	参数值
B_{max}	缓存总大小	40MB
B_{up}	缓存上限	36MB
$A(t)$	缓存分组到达速率	3Mbps, 6Mbps
$D(t)$	缓存分组消耗速率	400Kbps, 2Mbps
λ	端口开启时延	1.8s

本文采用如下简化的能耗模型计算能耗^[13]:

$$E(t) = t_{on} * P_{data} + t_{off} * P_{nodata} + E_{switch}$$

其中, t_{on}, t_{off} 分别为无线端口处于开启、关闭的时间; P_{data} ,

P_{data} 为对应状态单位时间的能耗, E_{switch} 为状态转换能耗。

4.2 仿真结果与分析

图 4 示出场景 1 下下限阈值对能耗影响的变化曲线。横坐标是下限阈值 B_{low} 占缓存总大小 B_{max} 的比重, 取值范围为 $[0.15, 0.85]$; 纵坐标是消耗的电量占总电量的比值。从图中可以看出, 终端能耗随缓存下限减低而减低, 在 $0.25 \sim 0.3$ 倍的 B_{max} 时, 各接收速率和消耗速率下的能耗最小。但是, 当缓存下限低于一定的值时, 能耗反而会急剧增加。这可能是由于缓存下限设置得过小, 导致缓存下溢, 出现了播放中断所引起的。与此同时, 增大缓存下限会使得能耗增大, 但当下限增大到一定的值时, 能耗并不再随着缓存下限的增大而增加, 这是因为此时的下限阈值已经使得终端数据端口处于持续工作的状态。此外, 缓存接收速率和消耗速率对能耗有着重要的影响。由图 4 可知, 合理调节缓存下限阈值可带来可观的能量收益, 其原因就在于降低了下限阈值, 使得用户设备有更长时间的休眠。

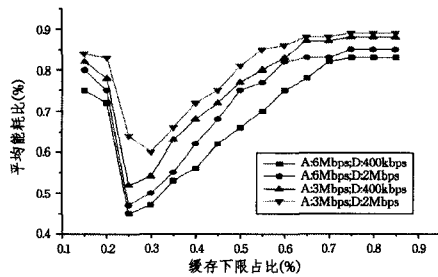


图 4 下限阈值对能耗的影响

图 5 示出场景 2 下传统静态缓存门限算法 (SBT) 和本文动态缓存门限调节算法 (DBTA) 的电量消耗变化曲线。相对于 SBT 来说, DBTA 的节能效果明显优于 SBT, 最多可节省约 22% 的能量。产生这种效果的原因是 DBTA 可以根据分组到达速率、视频播放速率和时延参数变化动态调整缓存下限, 避免数据端口频繁开关, 充分挖掘出缓存的节能潜力。

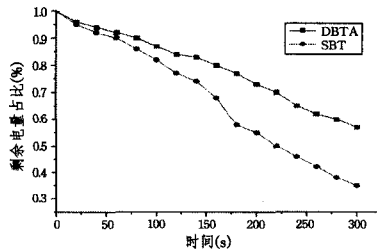


图 5 两种算法能耗的对比

综上所述, DBTA 算法可在节能和保证用户顺畅观看之间取得平衡, 能够在不明显增加延迟的情况下降低终端的能耗。

结束语 针对传统流媒体固定下限阈值能耗利用率不高的问题, 本文提出一种新的基于动态缓存门限调整的流媒体传输控制算法 (DBTA)。算法包括数据监测、自适应缓存下限调整两个部分, 根据监测到的网络带宽和流媒体编码率信息, 对终端缓存区下限阈值做出自适应的调整, 避免数据端口频繁开关, 提高终端的休眠工作时间比, 进而延长终端的续航时间。仿真实验证明, 相比于传统静态缓存门限 (SBT) 算法, DBTA 可在保证顺畅播放的前提下降低终端的电量消耗。

参考文献

- [1] Wang X, Vasilakos A V, Chen M, et al. A survey of green mobile networks: Opportunities and challenges [J]. *Mobile Networks and Applications*, 2012, 17(1): 4-20
- [2] Falaki H, Lymberopoulos D, Mahajan R, et al. A first look at traffic on smartphones [C] // *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*. ACM, 2010: 281-287
- [3] Rajaraman S V, Siekkinen M, Virkki V, et al. Bundling frames to save energy while streaming video from LTE mobile device [C] // *Proceedings of the Eighth ACM International Workshop on Mobility in the Evolving Internet Architecture*. ACM, 2013: 35-40
- [4] Tian Jian-jun, Qiu Ling. UE power saving scheduling method using DRX for video streaming in LTE [J]. *Journal of Graduate University of Chinese Academy of Sciences*, 2012, 29(4): 518-523 (in Chinese)
田建军, 邱玲. LTE 系统中一种利用 DRX 降低 UE 功耗的视频流调度方法 [J]. *中国科学院研究生院学报*, 2012, 29(4): 518-523
- [5] Khan S, Schroeder D, El Essaili A, et al. Energy-efficient and QoE-driven adaptive HTTP streaming over LTE [C] // *IEEE Wireless Communications and Networking Conference (WCNC 2014)*. Istanbul, Turkey, 2014: 8-14
- [6] Tian Gui-bin, Liu Yong. On Adaptive HTTP Streaming to Mobile Devices [C] // *2013 20th International Packet Video Workshop*. San Jose, CA, 2013: 1-8
- [7] Mai Yue-bo, Lin Xiao-hui. Research on Terminal Electricity Saving Mechanism of LTE Wireless System [J]. *Computer Engineering*, 2011, 37(15): 279-282 (in Chinese)
麦岳波, 林晓辉. LTE 无线系统终端省电机制研究 [J]. *计算机工程*, 2011, 37(15): 279-282
- [8] Pettis N, Cai L, Lu Y H. Dynamic power management for streaming data [C] // *Proceedings of the 2004 International Symposium on Low Power Electronics and Design*. ACM, 2004: 62-65
- [9] Cai L, Lu Y H. Dynamic power management using data buffers [C] // *Proceedings of the Conference on Design, Automation and Test in Europe - Volume 1*. IEEE Computer Society, 2004: 10526
- [10] Adams J, Muntean G M. Adaptive-buffer power save mechanism for mobile multimedia streaming [C] // *IEEE International Conference on Communications, 2007 (ICC '07)*. IEEE, 2007: 4548-4553
- [11] Hoque M A, Siekkinen M, Nurminen J K. TCP receive buffer aware wireless multimedia streaming: an energy efficient approach [C] // *Proceeding of the 23rd ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*. ACM, 2013: 13-18
- [12] Hoque M A, Siekkinen M, Nurminen J K, et al. Mobile multimedia streaming techniques: QoE and energy saving perspective [J]. *Pervasive and Mobile Computing*, 2014(5): 1574-1589
- [13] Deng S, Balakrishnan H. Traffic-aware techniques to reduce 3g/lte wireless energy consumption [C] // *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies*. ACM, 2012: 181-192