

# 网络能耗系统模型及能效算法

张 法<sup>1)</sup> Antonio Fernandez Anta<sup>4)</sup> 王 林<sup>1),2)</sup> 侯晨颖<sup>1),2)</sup> 刘志勇<sup>1),3)</sup>

<sup>1)</sup>(中国科学院计算技术研究所 北京 100190)

<sup>2)</sup>(中国科学院研究生院 北京 100190)

<sup>3)</sup>(中国科学院计算技术研究所计算机体系结构国家重点实验室 北京 100190)

<sup>4)</sup>(马德里高等研究院网络所 马德里 28918 西班牙)

**摘 要** 网络能耗问题是当前网络研究和发展的重大问题,影响着网络的设计、应用和发展.当前网络能效算法的研究大多从网络局部角度关注于网络某(些)设施的能耗问题,缺乏从网络全局的角度研究网络整体能耗的算法和策略.文中从网络全局角度研究网络的能耗模型和算法问题.文中基于 3 种基本网络数据传递模式和 3 种基本网络能耗机制,构建了网络能耗系统优化模型,对其中 5 种关键系统模型给出了详细的描述.针对几个关键系统模型,文中给出了形式化描述并且提出了相应的优化的节能路由算法.文中给出的网络系统能耗优化模型有助于对于网络能耗的理解和分析,所提出的能耗优化的网络数据包路由算法可以从网络路由的角度有效降低网络能耗.

**关键词** 能耗;系统模型;能效算法;最优化;网络延迟;绿色计算

**中图法分类号** TP393 **DOI 号:** 10.3724/SP.J.1016.2012.00603

## Network Energy Consumption Models and Energy Efficient Algorithms

ZHANG Fa<sup>1)</sup> Antonio Fernandez Anta<sup>4)</sup> WANG Lin<sup>1),2)</sup> HOU Chen-Ying<sup>1),2)</sup> LIU Zhi-Yong<sup>1),3)</sup>

<sup>1)</sup>(*Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190*)

<sup>2)</sup>(*Graduate University of Chinese Academy Sciences, Beijing 100190*)

<sup>3)</sup>(*State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190*)

<sup>4)</sup>(*Institute IMDEA Network, Madrid 28918, Spain*)

**Abstract** Energy consumption is a momentous problem that severely challenges further design and application of networks. While most researches work on a local view of some aspects (e. g. some devices used in networks) of the energy consumption problems in networks, there has been scarce research on a global view to reduce the amount of energy consumed at a network level (e. g. routing, network deployment). Energy consumption problem is investigated from network routing aspect in this paper. Energy consumption optimization strategies are developed from the aspect of network routing on the network system level. Combining three traffic arrival modes and three energy adaptation modes, optimized network energy consumption models are presented first. Further some energy efficient routing algorithms are developed for specific system models including the Continuous Flow with Speed Scaling model with bandwidth constraint, and the Continuous Flow with Rate Adaptation model. A model and corresponding algorithm for bi-criteria system are also developed so that a trade-off can be made between energy consumption and network delay. While the models can help understand the energy consumption optimization

收稿日期:2011-08-09;最终修改稿收到日期:2012-01-12. 本课题得到国家自然科学基金重大国际合作项目(61020106002)、NSFC-RGC 联合项目(61161160566)和创新研究群体项目(60921002)资助. 张 法,男,1974 年生,博士,副研究员,主要研究方向为网络能效算法、高性能计算. Antonio Fernandez Anta,男,1965 年生,博士,教授,博士生导师,主要研究领域为网络通信、分布式计算、计算机算法. 王 林,男,1989 年生,博士研究生,主要研究方向为网络能效算法、绿色网络. 侯晨颖,女,1988 年生,博士研究生,主要研究方向为网络能效算法、绿色网络. 刘志勇(通信作者),男,1946 年生,博士,研究员,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为高性能算法与体系结构、并行处理与网络. E-mail: zyliu@ict.ac.cn.

problems from the aspect of network routing on the network system level, the energy efficient routing algorithms can significantly reduce the energy consumed for network packet transmission.

**Keywords** energy consumption; system model; energy efficient algorithm; optimization; network latency; green computing

## 1 引 言

互联网络在过去 20 多年中飞速发展,对整个社会产生了深远的影响,已经从最初的仅供学术研究的实验平台发展成为对全球社会和经济有着至关重要影响力的通信媒介<sup>[1]</sup>. 统计数据显示<sup>[2]</sup>,2010 年全球以太网的接入节点数已超过 30 亿,而与之相连接的交换机、路由器等设备的数量也同样庞大. 根据 CNNIC(中国互联网络信息中心)最新公布的《第 28 次中国互联网络发展状况统计报告》显示<sup>①</sup>,“截止到 2011 年 6 月,我国网民规模已达到 4.85 亿,互联网普及率达到 36.2%”,而随着网络成本的降低与应用的普及,这些数字仍在快速持续地增长. 正如微软的首席技术官 Ray Ozzie 所讲:“我们已经进入了一个以互联网为中心的新时代”<sup>[3]</sup>.

长久以来,人们普遍认为信息系统对全球能源消耗的影响微乎其微,然而事实并非如此. 美国环保署(EPA)公布的一项数据表明:当前美国数据中心所消耗的电力已占到美国全部电力使用量的 2%,并且其电力需求正在以 12% 的速度递增,预计到 2020 年,这一数字将激增到 9%<sup>[4]</sup>. 另外数量庞大的网络设备也带来了巨额的能耗,2008 年服务于互联网的路由器、服务器、交换机、冷却设施、数据中心等各种设施总共消耗约 8680 亿度电,占全球总耗电量的 5.3%<sup>[5]</sup>,而且按照目前的增长趋势,到 2025 年,信息和通信技术行业(Information Communication Technology, ICT)平均耗能将达到 2006 年的 5 倍,网络领域的能耗更会达到 13 倍<sup>[6]</sup>. 伴随着能耗的持续增长,ICT 行业造成的碳排放量也在急剧增长,2007 年全球 ICT 行业的碳排放量占全球总排放量的 2%,并且这一比例到 2020 年将翻一翻<sup>[7]</sup>. 众所周知,节能减排已经成为全球最受关注的焦点,我国在 2009 年哥本哈根气候峰会上承诺到 2020 年二氧化碳排放比 2005 年下降 40%~45%,并且节能减排已经成为“十二五”发展规划的重要内容. 因此,网络设备和计算设备对能源的消耗,严重阻碍着节能型社会的创建,已成为一个急需解决的对于环境和

经济发展具有重要影响的重大社会问题.

然而传统网络系统设计的两个原则,即超额资源供给和冗余设计,有悖于低碳节能的目标<sup>[5]</sup>,造成当前网络的能耗效率极其低下. 各种网络通信设备在最初设计时并没有充分考虑能耗问题,而是以解决通信拥塞和保障服务质量为设计目标,通常考虑如何满足未来网络流量的增长、提高网络传输的可靠性、应对突发故障、方便维护等因素,因此这些设备浪费了大量的能量<sup>[8]</sup>,而且随着网络规模的扩大和网络设施的不断更新,网络系统的高能耗、低效率等问题日益显著. 例如,目前网络繁忙时的最大平均链路利用率不足 30%,很多网络的空闲时链路利用率甚至不足 5%,几乎所有网络设备的能耗都是由峰值带宽决定的,且绝大部分设备为 7×24 小时全速工作,而网络用户真正需要最高带宽的工作时间不足设备运行时间的 5%<sup>[9]</sup>. 这就意味着,即使在网络空闲的情况下,网络设备依旧按照峰值带宽的标准消耗能量. 如果网络设备的能耗能够与其负载流量成正比,保守估计<sup>[7]</sup>,超过 80% 的网络接入层能耗和 40% 的网络设备总能耗将被节省下来,至少将减少 15% 的二氧化碳排放量,这样在相同能源成本的情况下,我们可以部署更多的网络设施,从而提高整个网络系统的稳定性和可靠性. 目前许多政府机构和产业组织也纷纷制定标准和法规来提倡绿色网络. 例如美国贝尔实验室倡导构建的全球性联盟“绿色沟通”(Green Touch),旨在通过技术创新实现千倍提升通信网络能效的目标;电子电气工程师协会(IEEE)正式批准了 802.3az EEE 规范以降低网络设备的能耗,其中 EEE 为高效节能以太网(Energy Efficient Ethernet)的缩写<sup>②</sup>. 因此,构建绿色网络,降低网络能耗已经成为当前计算机网络领域一个意义重大、迫切需要解决的关键问题.

绿色网络是一个系统级的研究课题,涉及网络

① 中国互联网络信息中心.《第 28 次中国互联网络发展状况统计报告》,2011 年 7 月. [www.cnnic.net.cn/dtygg/dtgg/201107/W0201107195217\\_25234632.pdf](http://www.cnnic.net.cn/dtygg/dtgg/201107/W0201107195217_25234632.pdf)

② IEEE P802.3az Energy Efficient Ethernet. Task force public area. <http://grouper.ieee.org/groups/802/3/az/public/index.html>, 2008

体系结构、网络协议、网络设备结构创新和算法设计与优化等网络领域的核心科学问题<sup>[5]</sup>. 针对网络能效算法, 人们进行了深入研究. 速度缩放和休眠是当前两种动态适应能量消耗的主要机制<sup>[10-15]</sup>. 速度缩放机制是根据网络的负载流量动态调整网络设备的传输速度, 从而达到节能的目的, 休眠机制通过关闭相应的网络设备以实现降低网络能耗的效果. 围绕这两种机制, 人们已经取得了大量的研究成果. 但是, 从网络全局的路由分配角度分析, 如何描述网络中消息传送的能耗优化问题以及如何为消息分配路由以使得能耗最小化, 仍然没有得到充分研究. 其中包括, 如何结合网络数据不同的传输模式、采用不同的节能策略(如休眠、速度缩放、速率适应等)为能耗最优化建立模型, 分析其计算的难度(复杂性)以及设计可行的路由分配算法等问题. 现有研究中存在的主要问题包括: (1) 缺乏降低网络全局整体能耗的算法研究. 目前大多数的研究工作只关注于网络(有线或无线)某一(些)方面的能耗问题, 例如如何降低处理器的能耗以延长移动设备的电池寿命, 如何减少网络系统某一接口设备(CPU、网络接口、路由器、交换机等)的能耗等等, 缺乏从网络全局的角度研究网络整体能耗的算法和策略. (2) 缺乏针对网络能效算法理论分析的研究, 包括算法正确性证明及其理论性的性能分析和评价等等.

本文从网络全局数据包调度和路由的角度研究能耗优化的模型和算法. 从数据传递模式以及网络能耗机制角度, 构建了网络通信能耗优化的系统模型, 并且给出了 5 种关键系统模型的详细描述. 在此基础上, 给出了部分系统模型的形式化描述和降低网络能耗的优化算法. 本文主要分为以下 4 个部分: 第 2 节综合分析当前提高网络能效的算法和策略; 第 3 节基于数据的传递模式和网络能耗机制, 对网络通信能耗的优化问题进行了系统建模, 并且对其中 5 种关键系统模型进行了详细的形式化描述; 第 4 节给出基于部分系统模型的网络能效算法及相应的实验结果; 最后对全文进行总结, 提出了进一步的研究方向.

## 2 提高网络能效的算法和策略

由于 CPU 是现代计算系统中能耗最高的部件, 关于能效算法的研究最初主要是针对处理器的电源管理<sup>[10]</sup>, 其主要目的是为了延长笔记本电脑或其它计算系统的电池使用寿命. 围绕 CPU 电源管

理的研究可分为如下两类: 速度缩放机制(speed scaling)和休眠机制(powering down). 而目前网络能效算法的研究策略是根据实际的流量负载动态调整网络的传输速率, 从而达到降低网络能耗的目的. 这方面的研究大多都借鉴了 CPU 电源管理的研究策略, 同样也大致分为速度缩放机制和休眠机制两类. 在速度缩放机制中每一个网络设备都具有调节自身传输速度的能力, 它是根据网络流量负载动态调节处理速度以实现节约能源的目的. 在休眠机制中每一个网络设备或者工作在全功耗的激活(active)模式下, 或者工作在低功耗的休眠(sleeping)模式下, 它是通过使相应的网络设备休眠以达到节约能源的目的.

围绕上述两种机制, 人们针对单一计算系统或该系统的某一方面(如 CPU 或某一网络接口设备)的能耗进行了深入研究, 例如针对 CPU 任务调度的能源效率研究<sup>[16-24]</sup>, 针对单一网络接口设备节能的研究<sup>[10-15]</sup>.

Yao 等人<sup>[16]</sup>对处理器的速度缩放问题进行了系统阐述, 将处理器的能量消耗  $P$  定义为一个同处理速度  $s$  相关的多项式函数,  $P(s) = s^\alpha$ , 其中常数  $\alpha \geq 2$ . 能耗函数  $P(s)$  是严格意义上的凸函数, 也就是说, 任务处理的速度越快, 能耗增加得越快. 这样给定一组任务集, 并且明确指定每一任务的执行时间和被处理的速度, 则处理器能耗最小化问题就转化为最优任务调度问题. 并且 Yao 等人给出了一个离线的多项式时间内的贪心算法和两个在线的竞争算法. 围绕速度缩放机制, 研究人员进行了深入研究并取得了大量研究成果, 例如文献<sup>[17-20]</sup>. 此外 Wierman 等人<sup>[20]</sup>针对 Intel PXA270、Pentium M770 和 TCP/IP Offload Engine 进行了功耗研究, 分别将常数  $\alpha$  降低到 1.11、1.62 和 1.66.

Irani 等人<sup>[21]</sup>明确定义了休眠(powering down)的工作机制, 并且提出了一种基于概率的能耗最小算法<sup>[22]</sup>. Irani 定义的休眠机制如下<sup>[21]</sup>: 在处理器运行过程中存在很多未知的空闲时段, 在此期间处理器没有任何任务处理, 因此在这些空闲时段可以将处理器处于休眠状态, 直到某一任务需要处理器处理为止. 对于处理器的非空闲时间, 根据能耗将其划分成  $k$  个运行状态,  $S_0, \dots, S_k$ , 其对应的能耗分别为  $P_0, \dots, P_k$ , 且  $P_0 \geq P_1 \geq \dots \geq P_k$ . 处理器从状态  $S_i$  转变到  $S_j$  需要转换时间  $t_{ij}$  和转换能耗  $e_{ij}$ . 此外休眠状态结束时, 处理器将被重置回  $S_0$  状态. 为了进一步降低 CPU 的能耗, Irani 等人<sup>[23-24]</sup>尝试着将速度缩

放机制和休眠机制相结合,提出了一个 $(3+2\sqrt{2})$ -competitive 的算法.他们将处理器的运行过程划分成激活和休眠两种状态.在激活状态下,处理器可以以任意速度运行,且其能耗函数是一个与处理速度相关的凸函数;在休眠状态下,处理器不消耗任何能量,但是从休眠状态过渡到激活状态需要一个固定的能耗.

此外,通信的能量消耗也一直是无线网络研究的热点问题,尤其是围绕 ad hoc 无线网络和传感器网络中的能耗问题,人们进行了大量的研究工作<sup>[25-26]</sup>,例如 Jones 等人<sup>[26]</sup>概括总结了不同协议层面上的各种降低能耗的技术,提出了一种根据邻居状态的分布式休眠算法.

Gupta 和 Singh 首次强调了通用网络能耗问题的重要性<sup>[10-12]</sup>.他们发现:当前大部分以太网接口设备都具备低能耗的工作模式,但是大部分链接在空闲时并没有被置于该状态.他们提出了通过改变网络协议将网络设备中的部分组件置于休眠状态(或节能状态)来降低能耗的方法.在此基础上,Christensen 等人<sup>[13-15]</sup>提出了通过代理因特网适配器的途径来减少因特网边缘能耗的方法.同时他们对减少 PC、以太网链接以及边缘路由器的能耗进行了研究,其中“在可能的情况下减少以太网链接中的数据传递速度”的思想被进一步发展为自适应链接速率(Adaptive Link Rate, ALR)技术<sup>[15]</sup>:基于动态速度缩放机制,通过自主调节全双工传输的以太网链路交换速率从而适应不同的链路利用率,使链路交换速率与链路负载成比例.目前对于该技术的研 究主要集中在速率切换(包括切换的条件和时机、最优速率的选择以及如何快速切换)及其对性能的影响等方面<sup>[5]</sup>.

在以太网工业界,速率适应和链接休眠这两种方法受到高度的重视<sup>[9]</sup>,速度适应(rate adaptation)的方法已被应用到 ALR 技术中<sup>[15]</sup>,链接休眠的方法也已经被 IEEE 802.3az 节能以太网工作组定为低功耗空闲技术的标准.此外其它的链接标准 ADSL2 和 ADSL2+ 也支持多种链接速度和能耗模式<sup>[27]</sup>.

在针对全局网络最优能耗的策略和算法研究方面,Nedevschi 等人<sup>[28]</sup>对速度自适应、链接休眠和部分网络设备休眠技术在降低全局网络能耗方面进行了探索,并在真实的拓扑结构上对这几种方法进行了对比.Chabarek 等人<sup>[1]</sup>评估了两个标准路由器的能量需求,创建了一个能耗模型,并通过求解混合的

整数规划的方法实现了能耗最优.

综上所述,当前大多数的研究工作侧重于关注网络能耗问题的某一(些)具体设施,例如针对某单一计算系统,或者该系统某一设备(CPU 或是网络接口)的能耗问题,缺乏从网络全局的角度研究网络整体能耗的算法和策略.如何从网络全局范围的角度(例如网络的部署、网络数据传输的调度、路由,等)分析网络的能耗行为,建立网络能耗的系统模型,设计高效率低能耗算法,是影响网络能耗的具有挑战性的问题.

### 3 网络能耗系统模型

本节从网络数据的传递模式以及网络能耗模式的角度,对网络通信能耗模型进行分析,归纳出几种系统模型,并且对其中部分系统模型给出了形式化描述和详细阐述.

#### 3.1 网络能耗系统模型的构建

从网络全局的角度考虑,通用的网络框架体系是由一些路由器和链路构成的,该网络系统为一些数据提供传输服务,且提供服务所消耗的能量取决于选择的路径以及每个链路上数据包的调度策略.在此我们定义数据传递的源点为  $s$ , 终点为  $t$ .

网络数据的传递模式可以分为如下 3 种:

(1) 连续流模式(continuous flow). 假设存在一个源点终点对  $(s_i, t_i)$  和某一传输请求  $d_i$ , 使得有一个连续的, 大小为  $d_i$  的流量从  $s_i$  传递到  $t_i$ .

(2) 周期性包到达模式(periodic packet arrival). 假设存在一个源点终点对  $(s_i, t_i)$  和传输速率  $p_i$ , 使得在每  $T$  个单位时间内, 存在有  $p_i$  个数据包从  $s_i$  传递到  $t_i$  (我们可以将该模式视为连续流模式的一个特例, 即  $d_i = p_i / T$ ).

(3) 对抗性包到达模式(adversarial packet arrival). 基于对抗性排队理论<sup>[29-30]</sup>, 在该模式中, 我们引入了一个受约束的对抗量(adversary)到被传递的数据包中, 即数据包的传输是由受约束的对抗量控制的. 同样每个被传递的数据包有一个源点  $s_i$  和终点  $t_i$ . 根据每个数据包是否已经事先分配了一条路径, 我们可以建立两个不同的传递模式. 此外, 对抗量的引入必须保证网络传递的稳定.

除了网络数据传递模式外, 我们引入了 3 种网络能耗机制. 假设任一链接  $e$  都对应一个能耗函数  $f_e(s)$ , 它表示该链接以速率  $s$  传输数据时所消耗的能量. 这 3 种机制是:

(1) 速度缩放机制(speed scaling). 每个网络链接可以根据网络的传输负载情况在 $[0, R]$ 速率区间中选择它的实际传输速率 $s$ ,  $0 \leq s \leq R$ . 当 $R = \infty$ 时为特殊情况.

(2) 休眠机制(powering down). 每个网络链接都有两种工作状态(激活状态和休眠状态). 在激活状态时, 网络的传输速度为全速 $R$ , 其对应的能耗为 $f(R)$ ; 在休眠状态时, 网络的传输速度为 $0$ , 其对应的能耗也为 $0$ . 此外, 网络由休眠状态转为激活状态需要花费 $\delta$ 时间, 在这段时间内网络的传输速度为 $0$ , 但能耗为 $f(R)$ .

(3) 速率自适应机制(rate adaptation). 网络工作时存在有多种工作状态( $S_0, \dots, S_k$ ), 其中包含休眠状态, 且每一状态对应一传输速率. 每一网络链路可以根据网络传输负载情况选择某一工作状态. 对于任一状态 $S_i$ ,  $0 \leq i \leq k$ , 其传输速率为 $R_i$ ,  $0 \leq R_i \leq R$ , 对应的能耗为 $f(R_i)$ . 网络由状态 $S_i$ 转换为 $S_j$ 需要花费 $\delta$ 时间, 期间网络的传输速率为 $0$ , 能耗为 $f(\max(R_i, R_j))$ .

围绕上述 3 种网络数据传递模式和三种网络能耗机制, 可得到 9 种可能的网络通信能耗系统模型, 其中绝大部分系统模型还没有被研究过. 我们对其中更具有典型意义的 5 种网络能耗系统模型进行了构建和定义, 这 5 种网络能耗系统模式如表 1 所示.

表 1 5 种网络通信能耗系统模型

	连续流	周期性包到达	对抗性包到达
速度缩放机制	CFSS		
休眠机制		PPPD	
速率自适应机制	CFRA	PPRA	AARA

(1) 连续流速度缩放模型(Continuous Flow with Speed Scaling, CFSS). 在该模型中, 数据以稳定且连续的形式传输. 网络中消耗的能量可以通过路由网络中的数据流量进行调节, 也可以通过降低每个链路上的传输速度来调节.

(2) 连续流速度自适应模型(Continuous Flow with Rate Adaptation, CFRA). 该模型和 CFSS 模型有着相同的数据传递模式, 同样可以通过路由来降低能源的消耗. 但是在该模型中, 链路上的数据传递速度不能取任意值, 而只能从预先设定的、可能的传递速度中选取.

(3) 周期性包到达休眠模型(Periodic Packet Arrival with Powering-Down, PPPD). 该模型中, 数据以分批分组的形式周期性传递, 同组被传递的数据有着相同的源点和终点, 且可能沿着同样的传

输路径. 可以通过对数据包的相应调度, 使链路在较长时间内处于休眠状态, 从而达到降低能耗的目的.

(4) 周期性包到达速率自适应模型(Periodic Packet Arrival with Rate Adaptation, PPRA). 该模型同 PPPD 模型有着相同的数据传递模式, 但是其能耗的减少并不是通过将链路处于休眠状态, 而是通过对数据包的调度和路由选择, 使链路暂时工作于低速率/低能耗的状态.

(5) 对抗性包到达速率自适应模型(Adversarial Packet Arrival with Rate Adaptation, AARA). 该模型中数据包以对抗性形式传递, 且通过对数据包的调度和路由选择, 使链路暂时转换到低速/低能耗的工作状态, 以实现能源消耗最小(注: 此模型主要为分析网络节能策略和算法的最差性能而做出的一种抽象, 本文将不详细论述).

### 3.2 5 种基本网络能耗系统模型描述

通过对真实网络中的路由和数据传输特点的分析, 结合当前网络技术的特性和约束条件, 以及系统模型中的主要组成部分, 如网络拓扑结构、流量和能耗函数等, 我们给出上述 5 种基本网络能耗系统模型的具体描述.

用一个无向图 $G=(V, E)$ 来表示一个网络, 其中 $V$ 代表网络设备(主机、路由器、交换机等),  $E$ 表示网络中的链接. 假设每一个网络设备(顶点或边)都有一个能耗函数 $f(s)$ , 可以定义该网络设备传输流量 $s$ 所消耗的能源.

(1) CFSS 模型. 在该模型中, 请求从源点到终点以稳定且连续的形式传输. 给定一组请求 $D=(d_1, d_2, \dots, d_k)$ , 假设请求 $i$ 沿一条单路径从源点 $s_i$ 路由至终点 $t_i$ , 请求 $i$ 的大小为 $d_i$ . 此能量优化问题的最终的目标是在满足每个请求 $i$ 时使链路能耗最小, 也就是, 寻找满足请求 $i$ 的路径, 使 $\sum_{e \in E} f(x_e)$ 的值最小, 其中 $x_e$ 表示所有请求在链路 $e$ 上的总流量. 那么, 在 CFSS 模型中能耗最小问题就可以转换为一个如下的最优化问题:

$$\begin{aligned}
 (P_0) \quad & \min \sum_e f(x_e) \\
 \text{subject to:} \quad & \\
 & x_e = \sum_i y_{i,e} d_i \quad \forall e \\
 & y_{i,e} \in \{0, 1\} \quad \forall i, e \\
 & y_{i,e}: \text{flow conservation (流守恒)}.
 \end{aligned}$$

其中,  $y_{i,e}$ 的值表示请求 $i$ 是否通过边 $e$ . 我们定义 $I_i(v)$ 和 $O_i(v)$ 分别表示请求 $i$ 流入顶点 $v$ 的量和流

出顶点  $v$  的量. 令  $F_i(v) = O_i(v) - I_i(v)$ , 流守恒即为对于源点  $s_i$  和终点  $t_i$  有  $F_i(s_i) = d_i$ ,  $F_i(t_i) = -d_i$ , 对于其它顶点  $v$ ,  $F_i(v) = 0$ . 对于大多数能耗函数  $f(x)$  而言, 这一问题为 NP 难问题, 通过对该问题进行适当的约束, 利用启发式算法可以得到不同形式能耗函数  $f(x)$  的近似解.

(2) CFRA 模型. 在该模型中, 每一网络设备都存在有多种工作状态  $S_0, S_1, \dots, S_k$ , 其对应的传输速率分别为  $R_0, R_1, \dots, R_k$  (本文称之为有效速率), 且  $R_0 \leq R_1 \leq \dots \leq R_k$ , 有效速率  $R_i$  的能耗为  $f(R_i)$ . 任一链接的传输速率只从这些有效速率中选择. 假设能耗函数  $f(\cdot)$  是非递减的, 则 CFRA 模型类似于 CFSS 模型, 其能耗最小问题可以转化为如下的最优化问题:

$$(P_1) \quad \min \sum_e f(x_e)$$

subject to:

$$x_e = \sum_i y_{i,e} d_i \quad \forall e$$

$$x_e \leq z_e \quad \forall e$$

$$z_e \in \{R_1, R_2, \dots, R_k\} \quad \forall e$$

$$y_{i,e} \in \{0, 1\} \quad \forall i, e$$

$y_{i,e}$ : flow conservation (流守恒).

其中,  $z_e$  代表链接  $e$  所选择的有效速率, 且  $x_e \leq z_e$ , 即  $z_e$  可以确保流量  $x_e$  在链路上的稳定传输. 同 CFSS 模型类似, 对于大多数能耗函数  $f(x)$  而言, 这一问题同样是 NP 难问题.

(3) PPPD 模型. 在这个模型中, 由源点  $s_i$  路由至终点  $t_i$  的请求  $d_i$  是周期性传递的, 即每  $T$  个单位时间内, 有  $p_i$  个数据包从  $s_i$  传递到  $t_i$ . 网络链路有两种工作状态: 激活状态和休眠状态, 其中激活状态, 网络链路以全速工作, 休眠状态不消耗能量. 该模型中降低能耗的唯一方法是在没有包到达时尽可能使网络链路处于休眠状态. 因此能耗最小问题就转化为如何对每个链路中的数据包进行批处理调度, 以使链路能够尽可能长时间地处于休眠状态. 此外, 休眠状态向激活状态的转换会产生大量能耗, 因此休眠状态应该尽可能地保持连续. 同时, 在该模型中, 数据的传输延迟也需要限定在一个可以接受的范围内.

此类问题的困难在于, 如何在满足传输延迟要求的条件下, 对数据包进行调度和路由, 以使得链路尽可能地处于较长时间的休眠状态, 休眠状态又尽可能地连续以尽量避免状态转换引起的能量消耗. 该模型的能耗最小问题可以分为两种情况进行分

析: (1) 由  $s_i$  至  $t_i$  的数据传输路径是已知的, 此时需要解决的问题仅仅是数据包的调度问题; (2) 由  $s_i$  至  $t_i$  的数据传输路径是未知的, 此时我们可以自由选择路径, 这时可以首先对简单无环的拓扑结构, 如直线、树等进行有效地调度, 然后在这些路径中寻找一个生成树, 以实现网络能耗最小.

(4) PPRA 模型. 该模型的数据传输模式与 PPPD 模型类似, 由源点  $s_i$  路由至终点  $t_i$  的请求  $d_i$  是周期性传递的; 但是其网络能耗机制同 CFRA 模型类似, 每一网络设备都存在有多种工作状态  $S_0, S_1, \dots, S_k$ , 其对应的有效传输速率分别为  $R_0, R_1, \dots, R_k$ , 且  $R_0 \leq R_1 \leq \dots \leq R_k$ , 有效速率  $R_i$  的能耗为  $f(R_i)$ . 该模型是通过选择不同的有效传输速率来实现降低能耗的目的. 然而比 CFRA 模型更复杂的是, 在不同的传输周期, 链路的传输速率是可以动态变化的.

(5) AARA 模型. 在该模型中, 数据的传输是按照对抗性包到达模式进行的, 即数据的传输由一个受约束的对抗量控制, 它决定了数据到达网络的时间以及传输的源点和终点. 在数据的传输过程中, 传输时间被分为离散的时间间隔, 在每一时间间隔中对抗量会以一定平均速率  $r$  动态生成包注入到网络中, 这些包的源点和终点已知, 其路径可能已经由对抗量给定, 或是路径未知需要在对抗量注入后由网络进行动态自适应调度.

网络链路可以根据网络传输负载情况选择某一有效工作速率, 不同工作速率对应有不同的能量消耗, 由于网络中的对抗量包是在每个时间间隔内动态注入的, 所以对对抗量包的注入具有一定的破坏性, 会影响当前网络中的能量消耗. 因此, 该模型中能耗最小问题可以看成是对抗量和调度协议的博弈. 即对于每一个链路, 在每个时间间隔内选择合适的工作速率和数据的调度协议, 以使能耗达到最小. 由于数据传输不是周期性的, 所以解决方案很可能不是周期性的, 而且它更有可能随着分组到达的方式进行调整.

## 4 网络能效算法的设计

针对上述 5 种基本系统能耗模型及其最小能耗问题, 在设计相应能效算法时, 应充分考虑以下按照优先级排列的几个特性:

(1) 网络链路的稳定性. 网络传输流量在任何时候都是有限制的, 因此链接不允许过载.

(2) 能耗最小. 设计的能效算法要使得系统消

耗的总能量尽可能少。

(3) 延迟最小. 设计的能效算法要保证数据包的传输延迟尽可能小。

#### 4.1 基于 CFSS 模型的能效算法

在 CFSS 模型中, 对于大多数能耗函数  $f(x)$  而言, 最优化问题( $P_0$ )是 NP 难的. 然而, 如果能耗函数  $f(x)$  是次可加的, 即  $f(a+b) \leq f(a) + f(b)$ , 此类问题可以归结为 Buy-at-Bulk 网络设计问题. 网络设计的 Buy-at-Bulk 问题已经被广泛研究, 并且有一个对数性能比的近似算法<sup>[31]</sup>; 如果能耗函数  $f(x)$  是凸函数, 并且允许部分路由 (fractional routing), 那么上述最优化问题就变成了一个可以在多项式时间内最优解决的凸规划问题. 文献<sup>[32]</sup>针对两种非次可加的能耗函数, 利用随机舍入的思想, 给出了 CFSS 模型下的能效路由算法:

(1) 对于多项式形式的能耗函数  $f(x) = \mu x^\alpha$ , 针对最优化问题( $P_0$ ), 首先将整数变量  $y_{i,e}$  松弛到区间  $[0, 1]$  上, 然后通过求解松弛后的规划问题, 获得分数的近似解  $y_{i,e}^*$ , 最后利用 Raghavan-Thompson 的方法<sup>[33]</sup>进行随机舍入得到整数解. 对于请求  $d_i$ , 通过不断寻找从源点  $s_i$  到终点  $t_i$  的路径, 并且取每条路径上的瓶颈作为该路径的权值. 当找到所有连接源点  $s_i$  到终点  $t_i$  的路径后, 利用每条路径的权值作为其选取概率, 进行随机选择, 最终每个请求都将由唯一的路径路由。

(2) 对于带启动能量的能耗函数

$$f(x) = \begin{cases} 0, & x = 0 \\ \sigma + \mu x^\alpha, & x > 0 \end{cases}$$

可以证明其近似算法的理论上界为  $\Omega(\log^{1/4} N)$ , 并且利用随机舍入的方法可以获得  $O(K)$  的近似比, 其中  $K$  为请求的数目。

通过随机舍入的方法, 能够解决 CFSS 模型下非次可加能量函数的路由问题, 从而达到降低网络能耗的目的。

##### 4.1.1 基于 time-slot 思想的能效算法

但是上述方法没有考虑网络中的带宽限制. 在实际网络中, 每条链路的带宽都是有限的, 如果只把网络中能耗总和最小作为路由的唯一目标, 则可能会造成某些链路过载, 最终导致拥塞或数据丢失. 针对这一问题, 本文在考虑了网络链路带宽限制的基础上提出一种全局规划的能效路由算法, 该算法首先保证每条链路的流量不超过其容量上界, 再以全局能耗最小化为目标进行路由。

将网络的拓扑结构抽象为一个无向图  $G = (V, E)$ ,  $G$  中每条边  $e$  都有一容量值  $c_e$ , 能耗函数  $f_e(x)$  表示边  $e$  上的流量为  $x$  时消耗的能量. 此外还定义一个流量矩阵  $T$ , 其中  $T_{i,j}$  表示点  $i$  与点  $j$  之间的流量. 假设在该模型中允许部分路由, 则能耗最小问题可转化为如下最优化问题:

$$(P_2) \quad \min \sum_e f_e(x_e)$$

subject to:

$$x_e = \sum_i y_{i,e} d_i \quad \forall e$$

$$x_e \leq c_e \quad \forall e$$

$$y_{i,e} \in [0, 1] \quad \forall i, e$$

$$y_{i,e}: \text{flow conservation (流守恒)}.$$

由于允许部分路由的存在, 即  $y_{i,e}$  可以取 0 到 1 之间的分数值, 采用文献<sup>[32]</sup>中的方法, 该线性规划问题可以在多项式复杂度时间内获得最优解。

当数据传输为非连续时, 上述能效方法仍然存在节能的空间. 由于网络是一个异步的系统, 在实际网络中, 不同数据包可能存在不同的路径, 造成的延迟也不相同. 如果按照最优化问题( $P_2$ )的路由策略, 链路并不是时刻都处在满负荷状态, ( $P_2$ ) 仅为非连续流问题提供了一个下界, 保证了链路中不会出现拥塞, 但与此同时也带来了链路的带宽浪费和能量消耗. 因此, 本文提出一种基于 time-slot 思想的节能调度算法. 在调度时, 将所有数据包的调度分时槽进行, 将不同的数据包分配到不同的时槽里进行调度以提高链路带宽的利用率. 在时槽的分配上, 可以采用多种方法, 例如平均分配、随机选择等. 在本文中, 我们采用最优的分配时槽策略, 则能耗最小问题可转化为如下最优化问题(以时槽数  $k=2$  为例):

$$(P_3) \quad \min \sum_e f_e(x_{e,k}) \quad k=0, 1$$

subject to:

$$x_{e,k} = \sum_i y_{i,e,k} d_i \quad \forall e \quad k=0, 1$$

$$x_{e,k} \leq c_e \quad \forall e \quad k=0, 1$$

$$y_{i,e,k} \in [0, 1] \quad \forall i, e \quad k=0, 1$$

$$y_{i,e,k}: \text{flow conservation (流守恒)}.$$

由于随着时槽数的增加, 网络的拥塞程度会加大, 数据包的延迟会变长, 因此最优时槽数的确定取决于能耗与延迟之间的权衡关系. 在本研究中时槽数定为 2。

##### 4.1.2 实验结果

本文采用了 Rocketfuel 数据集集中的美国 ISP



拓扑结构(Rocketfuel AS3967)为数据测试集<sup>①</sup>. 采用文献[34]中的方法, 首先提取出原始拓扑结构的最大连通子图, 抽象出路由器图与城市图, 如图 1 所示. 然后根据城市图中每个城市的路由器数量以及城市间的地理距离, 按照公式  $T[a, b] = \frac{|routers(a)| \times |routers(b)|}{e^{\text{distance}(a,b)/1500}}$  (其中  $|routers(a)|$  为城市  $a$  的路由器数量) 计算出流量矩阵  $T$ . 最后在残缺图上(即去掉一个点的图, 以此模拟网络中可能出现的突发情况) 采用最短路径方法路由流量矩阵  $T$  来估算每条边的容量. 实验中能耗函数  $f_e(x)$  分别采用线性和凸函数形式, 即  $f_e(x) = \frac{c_e - \sigma}{c_e} \cdot x$  与  $f_e(x) = \frac{c_e - \sigma}{c_e} \cdot x^2$ , 其中  $\sigma = \beta \cdot c_e$ ,  $\beta$  反映了网络设备随流量大小调整传输速率的能力. 在此基础上, 本文对比了最短路径方法<sup>[32]</sup>、全局规划的能效路由算法( $P_2$ )和基于 time-slot 思想的能效算法( $P_3$ )的节能效果, 实验结果如表 2 所示.

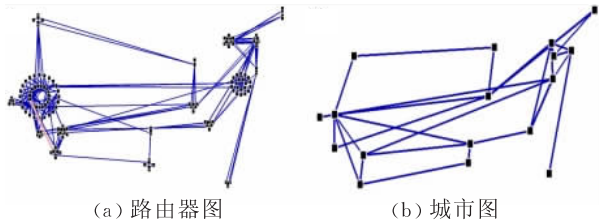


图 1 Rocketfuel AS 3967 路由器图和城市图

表 2 Rocketfuel AS 3967 数据集上各种方法的能耗对比

能耗函数	最短路径算法	全局规划能效算法( $P_2$ )	基于 time-slot 的能效算法( $P_3$ )
$f_e(x) = x$	45848	45848	45848
$f_e(x) = 1/2 c_e x$	56198419	47905680	46000000
$f_e(x) = 1/4 c_e x$	28090610	23945050	22900000
$f_e(x) = 1/(2c_e) x^2$	20563	12947	9410
$f_e(x) = 1/(4c_e) x^2$	10205	6474	4710

由表 2 可知, 在节能方面, 基于 time-slot 的能效算法要优于全局规划能效算法, 而后者要优于最短路径法, 并且节能的多少与能耗函数的表现形式有关. 本文的能耗函数选取依据文献[34], 利用凸函数来描述能耗随传输速率呈超线性增长这一特性, 因此本实验中凸函数形式的能耗函数具有更好的节能效果.

## 4.2 基于 CFRA 模型的能效算法

### 4.2.1 算法描述

CFRA 模型可以适应网络中的设施可以工作在非连续的速度可调整的状态, 从而描述网络中的设施可以工作在预先设定的阶梯型的速率上. 在 CFRA 模型中, 对于大多数能耗函数  $f(x)$  而言, 最

优化问题( $P_1$ )是 NP 难的; 并且在该模型中, 网络设备的能耗由其对应的工作状态决定, 而这些状态又已经被预先给定. 因此在( $P_1$ )中, 能耗函数  $f(x)$  是阶梯函数. 阶梯函数的离散性大大增加了( $P_1$ )问题的求解难度, 这使得仅利用 CFSS 模型中对  $y_{i,e}$  松弛的方法不能够使问题( $P_1$ )的难度降低到多项式时间内可解.

为了解决这样一个困难的路由问题, 我们需要对能耗函数  $f(x)$  进行预处理, 将其拟合为连续性的能耗函数, 以简化问题( $P_1$ )的难度.

对于阶梯函数的拟合, 最常用的方法有 midpoint 插值和采样插值法. midpoint 插值法取分段函数每段的中点, 然后利用这些点进行最小二乘拟合. 由于没有考虑每段的长度, 此方法存在很大的局限性. 采样插值法通过在分段函数上采样获取一系列的样本点, 并利用这些样本点进行最小二乘拟合, 此方法在很大程度上依赖于使用的采样方法, 效果无法保证. 本文提出一种基于积分最小化的插值方法, 利用原函数和目标函数的均方误差最小作为指导原则进行插值, 可克服上述两种方法的不足, 求得能够准确描述原函数特征的目标函数.

假设 CFRA 模型中阶梯形能耗函数  $f(x)$  的定义如下:

$$f(x) = \begin{cases} y_1, & x_0 < x \leq x_1 \\ y_2, & x_1 < x \leq x_2 \\ \vdots & \\ y_m, & x_{m-1} < x \leq x_m \end{cases},$$

其中  $y_i = f(R_i)$  ( $1 \leq i \leq m$ ), 为每一个工作状态的能耗,  $x_i = R_i$ ,  $x_{i+1} = R_{i+1}$  ( $1 \leq i \leq m$ ), 表示每个工作状态有效速率的上下界. 假设拟合得到的连续性目标函数为  $g(x) = \mu x^\beta$ , 则积分最小化可描述为

$$G(\mu, \beta) = \int [f(x) - g(x | \mu, \beta)]^2 dx$$

$$= \sum_{i=1}^m \int_{x_{i-1}}^{x_i} [y_i - \mu x^\beta]^2 dx.$$

由于  $g(x)$  为非线性函数, 此最小化问题求解难度很大. 对  $g(x)$  两边做对数变换则成为线性, 即  $\log(g(x)) = \log \mu + \beta \log x$ . 令  $z_i = \log y_i$ ,  $w_i = \log x_i$  和  $\mu' = \log \mu$ , 则上式变为

$$G'(\mu, \beta) = \sum_{i=1}^m \int_{x_{i-1}}^{x_i} [\log y_i - (\log \mu + \beta \log x)]^2 dx \\ = \sum_{i=1}^m \int_{w_{i-1}}^{w_i} [z_i - (\mu' + \beta w_i)]^2 dx,$$

① Rocketfuel: An ISP topology mapping engine. <http://www.cs.washington.edu/research/networking/rocketfuel/>



由于  $G'$  为标准的二次型, 选取合适的参数  $\mu$  和  $\beta$  使  $G'$  取得最小值的必要条件为  $G'$  的二次偏导均非负, 且一次偏导为零, 则有

$$\begin{cases} \frac{\partial G'}{\partial \mu'} = \sum_{i=1}^m \int_{w_{i-1}}^{w_i} -2[z_i - \mu' - \beta w] dw = 0 \\ \frac{\partial G'}{\partial \beta} = \sum_{i=1}^m \int_{w_{i-1}}^{w_i} -2w[z_i - \mu' - \beta w] dw = 0 \end{cases}$$

通过求解此方程组, 得到参数  $\mu$  和  $\beta$  的值, 进而可得到拟合后的连续性能耗函数  $g(x)$ . 则优化问题  $(P_1)$  转化成如下的优化问题  $(P'_1)$ .

优化问题  $(P'_1)$  为标准的整数规划问题, 采用文献[32]中的方法予以求解. 将规划问题中的整数变量松弛到分数区间, 松弛后的部分路由问题可以由求解凸规划得到分数精确解. 然后利用随机舍入的方法, 对于每个流量请求将多路径的流量统一到单一路径上. 至此问题  $(P'_1)$  得到近似求解, 对于每个流量请求分配单一路径路由.

$$(P'_1) \quad \min \sum_e g(x_e)$$

subject to:

$$x_e = \sum_i y_{i,e} \quad \forall e$$

$$y_{i,e} \in \{0, 1\} \quad \forall i, e$$

$y_{i,e}$ : flow conservation (流守恒).

#### 4.2.2 实验结果

本文在 Geant2 网络拓扑上测试了 CFRA 模型能效算法的性能. Geant2 为欧洲教育骨干网拓扑<sup>①</sup>, 包含有 34 个节点和 50 个链接, 如图 2 所示.

实验中能耗函数  $f(x)$  满足超可加的约束, 每一工作状态对应的能耗按照指数  $ax^b$  随机生成, 且  $b \in (2, 3)$ [32], 则每一工作状态对应的能耗如表 3 所示.

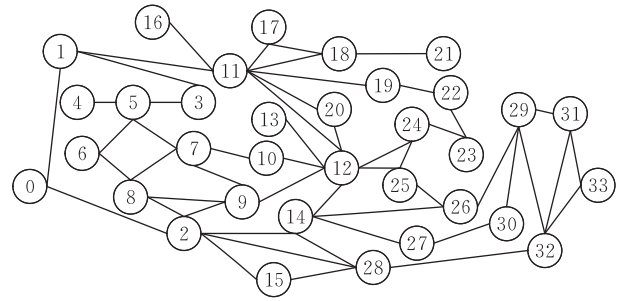


图 2 Geant2 网络拓扑图

表 3 超可加约束条件下随机生成的能耗函数  $f(x)$

工作状态	通过的请求数	能耗
$S_1$	(0, 2]	1
$S_2$	(2, 5]	16
$S_3$	(5, 10]	120
$S_4$	(10, 20]	764
$S_5$	(20, 30]	2885
$S_6$	(30, 40]	6918
$S_7$	(40, 50]	13298
$S_8$	(50, 60]	22406

利用积分最小化的差值拟和方法, 可得到如下方程组:

$$\begin{cases} 4.09 \log \mu + 8.38 \beta = 20.18 \\ 8.38 \log \mu + 22.88 \beta = 56.32 \end{cases}$$

求解可得  $\mu = 0.664$ ,  $\beta = 2.62$ , 即拟合后的能耗函数  $g(x) = \mu x^\beta = 0.664x^{2.62}$ .

表 4 给出了分别以  $f(x)$  和  $g(x)$  为能耗函数时的能量消耗以及它们之间的误差比率. 实验中流量请求随机生成并且数量以 10 的增量从 10 增加到 60. 对于每次路由, 随机生成一定数量的流量请求, 然后在网络拓扑上路由这些请求. 表 4 中的平均值为重复进行 10 次实验后的平均能耗.

表 4 能耗函数  $f(x)$  和  $g(x)$  在不同流量请求下的能量消耗及误差比率

请求数目	10		20		30		40		50		60	
	$f(x)$	$g(x)$	$f(x)$	$g(x)$	$f(x)$	$g(x)$	$f(x)$	$g(x)$	$f(x)$	$g(x)$	$f(x)$	$g(x)$
第 1 轮	53	47.5	297	337.2	635.5	575	1620	1475.3	3675	2876.5	5580	5401.2
第 2 轮	89	90.4	384	366.2	620	659.8	2586	2264.2	2606	2581.4	3914	3711.7
第 3 轮	65	62.6	251	221.1	900	1013.3	1965	2025.4	3386	3469.5	7097	6884.8
第 4 轮	96	95.4	215	279.4	816	859.5	2163	1949.0	1271	1199.7	5208	5664.9
第 5 轮	49	51.3	313	305.7	892	937.6	2452	2146.9	1730	1869.6	5406	5627.7
第 6 轮	50	54.4	153	162.5	979	937.2	1803	1698.4	3676	3862.2	5774	5989.9
第 7 轮	111	106	492	517.9	1015	951.1	2050	2251.8	2596	2458.6	4311	4635.5
第 8 轮	48	61.1	294	333.2	873	939.3	1577	1687.7	3791	3527.6	5765	5212.0
第 9 轮	49	57.9	509	560.5	455	485.2	1568	1518.9	2163	2374.5	4256	4978.3
第 10 轮	69	65.9	289	311.8	788	822.1	984	1136.4	3306	3517.1	4105	4660.7
平均值	<b>68</b>	<b>68.9</b>	<b>284.1</b>	<b>299.7</b>	<b>843.7</b>	<b>869.3</b>	<b>1900.3</b>	<b>1833.3</b>	<b>2971</b>	<b>2922.4</b>	<b>5293.4</b>	<b>5377.8</b>
平均误差	<b>1.36</b>		<b>5.50</b>		<b>3.03</b>		<b>3.53</b>		<b>1.65</b>		<b>1.60</b>	

① Geant2: Lighting the pathway for European Research and Education. <http://www.geant2.net>

基于积分最小化的插值方法将分段函数上的每个点都纳入考察,因此拟合后的能耗函数  $g(x)$  能够准确描述原能耗函数  $f(x)$  的特征. 此外,拟合中出现的误差对网络总能量计算的影响是两方面的,当  $f(x) > g(x)$  时能量被减少,而  $f(x) < g(x)$  时能量被增加,两个方面的作用也降低了实际误差. 由表 4 可知,拟合过程产生的误差维持在 2% 左右,最多不超过 6%. 由于 CFRA 模型中假设网络设备只工作在一定的状态下,对于当前的网络设备更具有普遍适用性,因此在实际网络中 CFRA 模型较 CFSS 模型具有更广的应用范围.

#### 4.3 双度量网络设计问题

传统路由协议只采用单度量,力求使延迟最小而不考虑其它因素(如网络容量等),如基于最短路径的路由算法等. 由于当前网络规模的扩大,网络环境日趋复杂,在路由问题上部分协议可同时考虑多个度量. 但是,同时将能量节省和网络延迟最小作为网络设计度量的研究还很少,尤其是对于适用性更强的超可加分段能耗函数,此问题还没有被研究过. 从本文前述的能耗模型可以看出,能耗最优化的路径分配依赖于能耗函数的性质,当能耗函数为线性函数(满足可加性)时,通常的最短路径(shortest-path)路由就是全局能耗最优化的路由;然而当能耗函数为超可加或者次可加函数时,简单的最短路径方法并不能获得全局最优路由. 针对现实世界网络中的能耗函数并不是线性函数的特性我们可以在系统能耗和传输延迟之间做出权衡,设计双度量的目标函数. 针对降低网络能耗和减小网络延迟的双度量网络设计问题,通过在目标函数中引入参数  $\alpha$ ,以凸组合的方式将上述两个目标结合起来,我们提出一种多项式时间的能效算法,能够在较低延迟情况下,得到较高的能量节省.

优化问题( $P'_1$ )只考虑了降低网络能耗,通过引入参数  $\alpha$  对能量消耗和网络延迟两个目标进行凸组合,则总的消耗可表示为

总消耗  $Cost = \alpha \times \text{能量消耗} + (1-\alpha) \times \text{网络延迟代价}$ .

凸组合可以保持两个目标原有的凹凸性,使组合后的问题与原问题可以用相同的方法求解. 通过调节参数  $\alpha$  可以获得不同的问题描述:当  $\alpha=0$  时,此模型退化成基于最短路径的路由问题;当  $\alpha=1$  时,模型退化成 CFRA 模型;当  $\alpha$  在  $(0,1)$  之间取值时,可以获得能耗和延迟的折中考虑.

网络延迟代价可通过对 CFRA 模型的扩展予以形式化描述:在图 G 的每条边上引入一个新的特征参数  $l_e$  代表边  $e$  的延迟,即边  $e$  所抽象的网络设

备的延迟. 这样总的消耗可以表示为

$$Cost = \alpha \times \sum_e g(x_e) + (1-\alpha)C(P_1, P_2, \dots, P_k),$$

其中  $P_i$  表示给请求  $i$  分配的路径,  $C(P_1, P_2, \dots, P_k)$  表示所有请求的路由路径的延迟代价,可以取所有路径延迟的最大值,或所有路径延迟的平均值. 如果选取所有路径延迟的平均值作为网络延迟,则总消耗可以表示为

$$\begin{aligned} Cost &= \sum_e \alpha g(x_e) + (1-\alpha)AVG(\omega_i) \\ &= \sum_e \alpha g(x_e) + (1-\alpha) \frac{1}{m} \sum_i \sum_e y_{i,e} l_e \\ &= (1-\alpha) \sum_e l_e x_e + \alpha \sum_e \mu x_e^\beta, \end{aligned}$$

其中  $y_{i,e}$  表示请求  $i$  是否经过链路  $e$ ,  $\omega_i$  表示请求  $i$  的传输路径上延迟,  $AVG(\omega_i)$  表示所有路径延迟的平均值. 因此,给定  $\alpha$  后,优化问题( $P'_1$ )可转换为如下形式( $P_4$ ).

根据凸组合的性质可知,能耗和延迟两个凸函数结合,其目标函数仍然是凸函数. 这样采用 CFSS 模型中求解凸规划和随机舍入的方法,可获得多项式复杂度时间内的最优解.

$$(P_4) \quad \min \alpha \sum_e g(x_e) + (1-\alpha)AVG_i(\omega_i)$$

subject to:

$$x_e = \sum_i y_{i,e} d_i \quad \forall e$$

$$\omega_i = \sum_e y_{i,e} * l_e \quad \forall i$$

$$x_e \leq z_e \quad \forall e$$

$$z_e \in \{R_1, R_2, \dots, R_k\} \quad \forall e$$

$$y_{i,e} \in \{0, 1\} \quad \forall i, e$$

$y_{i,e}$ : flow conservation (流守恒).

本文在 Geant2 网络拓扑上对参数  $\alpha$  和算法性能进行了测试. 图 3 给出了当  $\alpha$  取  $(0,1)$  间的值时,对于给定能耗函数和随机生成 60 个流量请求的能

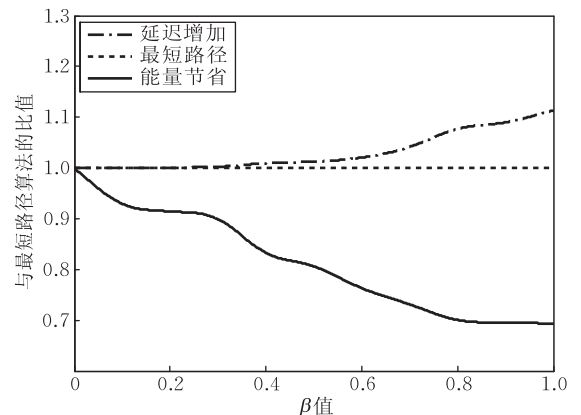


图 3 网络能耗与通信延迟随参数  $\alpha$  的变化情况

耗与延迟的变化情况. 随着  $\alpha$  的增大, 延迟将逐渐增大, 但能耗将逐渐降低. 当  $\alpha$  在 1 附近逐渐变小时, 网络延迟会迅速减小, 而  $\alpha$  靠近 0 时, 网络延迟的变化相对平稳. 由图 3 可知当  $\alpha \in [0.4, 0.7]$  时, 能够保证在较低的延迟下获得较高的能量节省.

表 5 给出了不同参数  $\alpha$  下网络能耗与延迟的性能对比. 实验中原始能耗函数  $f(x)$  如表 3 所示,  $g(x) = 0.664x^{2.62}$ .

表 5 不同参数  $\alpha$  下网络能耗与延迟的性能对比

参数	能耗降低/%	延迟增加/%
$\alpha = 0.99$	56.7	35.8
$\alpha = 0.9$	56.9	26.8
$\alpha = 0.8$	52.2	18.7
$\alpha = 0.7$	51.4	13.5
$\alpha = 0.6$	46.3	7.1
$\alpha = 0.5$	43.5	5.3
$\alpha = 0.4$	36.1	4.6
$\alpha = 0.3$	24.7	1.5

表 5 进一步验证了当  $\alpha \in [0.4, 0.7]$  时, 能够保证在较低延迟下获得较高的能量节省. 例如当  $\alpha = 0.6$  时, 能耗降低达到了 46% 而延迟的增加仅为 7.1%.

## 5 总 结

网络能耗问题是影响网络系统设计、使用和发展的重大问题. 人们已经从局部的、具体的网络设施的角度对节能策略和算法进行了大量研究, 而缺乏从网络全局的角度(如网络通讯信息包的调度、路由以及网络的总体部署等)对于网络能耗问题研究. 本文从网络全局的路由角度研究了网络节能问题, 给出了具体的能耗模型和优化方法. 本文首先基于 3 种基本网络数据传递模式和 3 种基本网络能耗机制, 构建了网络通信能耗优化的系统模型, 并且对其中 5 种关键系统模型给出了详细的形式化描述. 然后给出了 CFSS 和 CFRA 系统模型下的能效算法. 其中, 针对 CFSS 系统模型, 我们提出了一种有限容量网络中的全局规划的能效路由算法和基于 time-slot 思想的节能调度算法. 实验表明基于 time-slot 思想的节能调度算法, 可以提高链路带宽的利用率, 并且能够有效降低网络能耗. 针对 CFRA 模型, 我们利用基于积分最小的插值拟合和随机舍入的方法, 提出了一种能效路由策略. 同时考虑网络能耗效率和通讯延迟, 提出了双度量网络设计问题, 并且给出了一种启发式算法, 通过调整参数可以获得节能和延迟权衡. 实验表明在延迟增加不超过 10% 的情

况下, 可以降低超过 40% 的能耗.

本文主要贡献在于, 第一, 对于具有固定拓扑结构的网络系统的能耗问题从网络全局路由角度抽象出了具体模型并给出了计算路由的算法; 第二, 给出了可在多项式复杂度时间内计算的路由分配方法(路由算法). 其意义在于: 第一, 作为从全局路由的角度研究网络能耗的模型, 它提供了对于从全局路由的角度对于网络能耗的理解认识. 比如, 他可以解释为什么常用的最短路径路由方法不能给出最优的能耗以及如何通过调整对延迟和能耗的权衡的参数可以在最短路径和能耗之间取得所期望的权衡或达到整体上“最好的”(最理想的)效果. 第二, 所给出的路由算法难于直接在分布式路由器的网络中直接实现, 但本文的模型和算法可以作为评价和比较不同路由算法在能量消耗方面的一个标准.

由于大量网络中具体的消息路由是由网络中多个没有集中控制机构的路由器以分布式的计算形式给出的, 本文所给出的路由算法难于直接在分布式路由器的网络中实现. 如何设计分布式路由器中易于实现的分布式高能效路由算法是值得进一步研究的重要问题. 此外, 针对 PPPD、PPRA 和 AARA 系统模型, 如何对它们进行形式化描述, 从而基于这三个系统模型开发设计实用的能效算法等, 也是值得进一步研究的重要问题.

**致 谢** 本文的研究中, 我们与美国贝尔实验室(Bell Labs) Matthew Andrews 研究员和 Lisa Zhang 研究员进行的多次讨论, 使我们获益匪浅, 在此深表感谢!

## 参 考 文 献

- [1] Chabarek J, Sommers J, Barford P et al. Power awareness in network design and routing//Proceedings of the Conference on Computer Communications, IEEE (INFOCOM). Phoenix, USA, 2008; 457-465
- [2] Christensen K, Reviriego P, Nordman B et al. IEEE 802.3az: The road to energy efficient ethernet. IEEE Communications, 2010, 48(11): 50-56
- [3] Waters R. Microsoft's ozzie declares end to PC era. Financial Times, July 2006
- [4] Koomey J G. Worldwide electricity used in data centers. Environmental Research Letters, 2008, 3: 034008
- [5] Lin Chuang, Tian Yuan, Yao Min. Green network and green evaluation: Mechanism, modeling and evaluation. Chinese Journal of Computers, 2011, 34(4): 593-612(in Chinese)

- (林闯, 田源, 姚敏. 绿色网络和绿色评价: 节能机制、模型和评价. 计算机学报, 2011, 34(4): 593-612)
- [6] Yun D, Lee J. Research in green network for future Internet. *Journal of KIISE*, 2010, 28(1): 41-51
- [7] Smart2020: Enabling the low carbon economy in the information age. [www.smartwowo.org](http://www.smartwowo.org), 2008
- [8] In Proceedings of the Vision and Roadmap Workshop on Routing Telecom and Data Centers Toward Efficient Energy Use, October, 2008
- [9] Prachi Patel-Predd. Energy-efficient ethernet. *IEEE Spectrum*, 2008, 5: 13
- [10] Gupta M, Singh S. Greening of the Internet//Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, ACM (SIGCOMM). Karlsruhe, Germany, 2003: 19-26
- [11] Gupta M, Singh S. Dynamic ethernet link shutdown for energy conservation on ethernet links//Proceedings of the International Conference on Communications, IEEE (ICC). Glasgow, Scotland, 2007: 6156-6161
- [12] Gupta M, Singh S. Using low-power modes for energy conservation in ethernet LANs//Proceedings of the Conference on Computer Communications, IEEE (INFOCOM). Anchorage, USA, 2007, 2451-2455
- [13] Christensen K, Gunaratne C, Nordman B, George A. The next frontier for communications networks: Power management. *Computer Communications*, 2004, 27(18): 1758-1770
- [14] Gunaratne C, Christensen K, Nordman B. Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed. *International Journal of Network Management*, 2005, 15(5): 297-310
- [15] Gunaratne C, Christensen K, Nordman B, Suen S. Reducing the energy consumption of ethernet with adaptive link rate (alr). *IEEE Transactions on Computers*, 2008, 57(4): 448-461
- [16] Yao F, Frances, Demers A J, Shenker S. A scheduling model for reduced cpu energy//Proceedings of the 36th Annual Symposium on Foundations of Computer Science, IEEE (FOCS). Milwaukee, USA, 1995: 374-382
- [17] Brooks D, Bose P, Schuster S et al. Power-aware microarchitecture: Design and modeling challenges for next-generation microprocessors. *IEEE Micro*, 2000, 20(6): 26-44
- [18] Bansal N, Kimbrel T, Pruhs K. Speed scaling to manage energy and temperature. *Journal of ACM*, 2007, 54(1): 1-39
- [19] Chan H L, Chan W T, Lam T W et al. Energy efficient online deadline scheduling//Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms, ACM (SODA). New Orleans, USA, 2007: 795-804
- [20] Wierman A, Lachlan L, Andrew H, Tang A. Power-aware speed scaling in processor sharing systems//Proceedings of the Conference on Computer Communications, IEEE (INFOCOM). Rio de Janeiro, Brazil, 2009: 2007-2015
- [21] Irani S, Gupta R, Shukla S. Competitive analysis of dynamic power management strategies for systems with multiple power savings states//Proceedings of the Conference on Design, Automation and Test in Europe, IEEE (DATE). Paris, France, 2002: 117-123
- [22] Irani S, Shukla S, Gupta R. Online strategies for dynamic power management in systems with multiple power-saving states. *ACM Transactions in Embedded Computing Systems*, 2003, 2(3): 325-346
- [23] Irani S, Pruhs K. Algorithmic problems in power management. *ACM SIGACT News*, 2005, 36(2): 63-76
- [24] Irani S, Shukla S, Gupta R. Algorithms for power savings//Proceedings of the 14th ACM-SIAM Annual Symposium on Discrete Algorithms, ACM (SODA). Philadelphia, USA, 2003: 37-46
- [25] Akyildiz I, Su W L, Sankarasubramaniam Y, Cayirci E. Wireless sensor networks: A survey. *Computer Networks*, 2002, 38(4): 393-422
- [26] Jones C, Sivalingam K, Agrawal P, Chen J C. A survey of energy efficient network protocols for wireless networks. *Wireless Networks*, 2001, 7(4): 343-358
- [27] Ginis George. Low-power modes for ADSL2 and ADSL2+. White Paper, Broadband Communications Group, Texas Instruments, 2005
- [28] Nedeveschi S, Popa L, Iannaccone G et al. Reducing network energy consumption via sleeping and rate-adaptation//Proceedings of the 5th USENIX Symposium of Networked Systems Design and Implementation. San Francisco, USA, 2008: 323-336
- [29] Andrews M, Awerbuch B, Fernandez A, Thomson Leighton F, Liu Z Y, Kleinberg J. Universal-stability results and performance bounds for greedy contention-resolution protocols. *Journal of ACM*, 2001, 48(1): 39-69
- [30] Borodin A, Kleinberg J, Raghavan P, Sudan M, Williamson D. Adversarial queuing theory. *Journal of ACM*, 2001, 48(1): 13-38
- [31] Andrews Matthew. Hardness of buy-at-bulk network design//Proceedings of the 45th Annual Symposium on Foundations of Computer Science, IEEE (FOCS). Rome, Italy, 2004: 115-124
- [32] Andrews M, Fernandez A, Zhang L, Zhao W. Routing for energy minimization in the speed scaling model//Proceedings of the Conference on Computer Communications, IEEE (INFOCOM). San Diego, USA, 2010: 2435-2443
- [33] Raghavan P, Thompson C D. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 1991, 7: 365-374
- [34] Antonakopoulos S, Fortune S, Zhang L. Power-aware routing with rate-adaptive network elements//Proceedings of the 2010 International Conference on Green Computing and Communications, IEEE (GreenCom). Hangzhou, China, 2010: 1428-1432



**ZHANG Fa**, born in 1974, Ph. D. , associate professor. His research interests include energy-aware algorithms and high performance computing.

**Antonio Fernandez Anta**, born in 1965, Ph. D. , senior researcher, Ph. D. supervisor. His research interests include communications and networks, distributed computing and computer algorithms.

### Background

Energy consumption is rapidly increasing with the expanding of network size. Methods to reduce the energy consumption of network elements have drawn significant research interest during the past few years. However, there has been scarce research on algorithms to globally reduce the amount of energy consumed at a network level. This paper constructs five energy consumption system models and presents some energy efficiency scheduling algorithms for some specific system models. The research results of this paper will be useful to devise the energy efficiency algorithms in network. This work is supported by National Natural of Sci-

**WANG Lin**, born in 1989, Ph. D. candidate. His research interests include energy-aware algorithms and green network.

**HOU Chen-Ying**, born in 1988, Ph. D. candidate. Her research interests include energy-aware algorithms and green network.

**LIU Zhi-Yong**, born in 1946, Ph. D. , professor, Ph. D. supervisor. His research interests include high performance algorithms and architectures, parallel processing and networks.

ence Foundation of China (Nos. 61020106002). This project aims to provide better energy efficiency and performance in computing networks and information systems. Our group has working on the energy efficiency in computer networks and datacenter. Many good papers have been published in respectable international conferences and journals, such as IEEE Symposium on Foundations of Computer Science (FOCS), IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC), Journal of ACM, IEEE Transaction serial, and Journal of Parallel and Distributed Computing.