

绿色互联网中面向节能的分布式拓扑管理机制

张金宏¹⁾ 王兴伟²⁾ 黄敏³⁾ 马连博²⁾

¹⁾(东北大学计算机科学与工程学院 沈阳 110819)

²⁾(东北大学软件学院 沈阳 110819)

³⁾(东北大学信息科学与工程学院 沈阳 110819)

摘要 当前互联网存在着资源高冗余和高能耗等问题,随着网络规模的扩大和联网设备数量的激增,这些问题日益严重.文中在绿色互联网的背景下,着眼于主干网的拓扑管理和能耗控制,首先构建了主干网的网络模型,给出了节点和链路的结构模型,分析了网元功耗并建立了功耗模型;然后,基于上述模型,提出了一种绿色互联网中面向节能的分布式拓扑管理机制(Distributed Topology Management scheme for Energy saving, DTME),利用 VCG (Vickrey-Clarke-Groves) 机制对节点和链路进行配置管理,通过信息感知、流量预测、分布式拓扑决策和休眠控制之间的协同进行分布式拓扑管理,实现节能目标;最后,为了验证 DTME 在拓扑管理和节能方面的适用性和有效性,将其与采用 GRiDA(GREEN Distributed Algorithm)的拓扑管理机制在网络功耗和网络性能(收敛时间、平均延迟、分组丢失率和防策略性)等方面进行了仿真对比实验和性能评价,结果表明 DTME 具有良好的分布式拓扑管理能力和节能潜力.

关键词 绿色互联网;节能;分布式拓扑管理;信息感知;流量预测;休眠控制

中图分类号 TP393 **DOI号** 10.11897/SP.J.1016.2017.01517

A Distributed Topology Management Scheme for Energy Saving in Green Internet

ZHANG Jin-Hong¹⁾ WANG Xing-Wei²⁾ HUANG Min³⁾ MA Lian-Bo²⁾

¹⁾(School of Computer Science and Engineering, Northeastern University, Shenyang 110819)

²⁾(School of Software, Northeastern University, Shenyang 110819)

³⁾(School of Information Science and Engineering, Northeastern University, Shenyang 110819)

Abstract The current Internet is confronted with some severe problems, for example, high resource redundancy and excessive energy consumption. With the continuous expansion of network scale and the sharply increasing of the number of networking devices, these problems become more and more serious. In this paper, topology management and energy consumption control in the backbone networks are investigated in the context of green Internet. The backbone network model is built, and its node and link structure models are proposed. The power consumption of network element is analyzed and its model is devised. Based on the above models, a Distributed Topology Management scheme for Energy saving (DTME) in green Internet is devised. In DTME, node and link configurations are managed with VCG (Vickrey-Clarke-Groves) mechanism, and distributed topology management and energy saving are achieved via collaboration among information sensing, traffic prediction, distributed topology decision-making and sleeping control. To verify the applicability and effectiveness of DTME in terms of topology management and energy saving,

收稿日期:2015-09-21;在线出版日期:2016-05-23. 本课题得到国家自然科学基金(61572123)、国家杰出青年科学基金(71325002, 61225012)、辽宁省百千万人才工程项目(2013921068)、中国科学院科技创新交叉与合作团队计划面向智能电网安全监控及信息互动的网络与系统技术资助. 张金宏,男,1982年生,博士研究生,主要研究方向为节能路由和拓扑管理等. E-mail: neuzjh@aliyun.com. 王兴伟(通信作者),男,1968年生,博士,教授,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为未来互联网、云计算和网络空间安全等. E-mail: wangxw@mail.neu.edu.cn. 黄敏,女,1968年生,博士,教授,博士生导师,主要研究领域为物流建模与优化和供应链管理. 马连博,男,1981年生,博士,副教授,主要研究方向为计算智能与机器学习等.

simulated experiments and performance evaluations are carried out to compare DTME with the GRiDA (GReen Distributed Algorithm) based topology management mechanism on network power consumption and network performance (including convergence time, average delay, packet loss rate and strategy-proofness). Experiment results demonstrate that DTME has strong distributed topology management ability and good energy saving potential.

Keywords green Internet; energy saving; distributed topology management; information sensing; traffic prediction; sleeping control

1 引 言

近年来,随着互联网用户数量的增加、网络规模的扩大以及新型网络应用的涌现,网络能耗激增,互联网能耗问题受到学术界和产业界的广泛关注^[1-5].

在互联网设计之初,为了确保网络在峰值负载时仍能正常运行,网络通常采取超额资源供给的设计理念.在实际运行过程中,虽然这样的设计降低了网络故障率,增强了网络可靠性,但是由于峰值负载通常仅在很短的时段内出现,因此导致超额供给的网络资源在大部分时间里处于低负载全功率运行的状态,从而造成资源和能源的浪费.互联网主干网忙时最大平均链路利用率不足 30%,很多网络闲时平均链路利用率只有 5%^[1].当前网络存在的这些缺陷为绿色互联网提供了发展机遇.

Gupta 等人^[2]在 2003 年率先开展了关于互联网节能问题的研究工作.发展绿色互联网的目的是提高网络使用效率,降低网络能耗.由于采用休眠唤醒策略,网元的休眠和唤醒操作会导致网络拓扑变化,因此需要针对绿色互联网中的拓扑管理开展必要的研究,关键是要实现根据网络流量情况,在适当的时候休眠或者唤醒适当的网元.

网络拓扑管理可以分为集中式和分布式两种模式.集中式管理模式需要获取网络拓扑的全局信息,网元配置由集中控制单元决策,网元间无需协调.分布式管理模式无需全局信息,网元配置通过网元间的协调完成.尽管集中式管理模式对全局信息的掌握有利于做出更好的拓扑管理决策,但是缺点也很明显.首先,网络拓扑规模对管理机制性能的影响大,拓扑规模变大会加重信息获取的滞后性;其次,由于需要处理的数据多,计算量大,因而响应慢,难以应对状态快速变换的网络场景;最后,抗毁性差,集中处理节点的故障会导致整个系统崩溃.分布式

管理模式的优点在于适应性强,对不同规模的网络性能基本不变;基于局部信息(如自身和邻居节点信息)做出决策,不仅反应快,而且适合拓扑经常变化的网络.缺点是难以实现全局优化管理.

高效的拓扑管理需要建立在有效的网络流量感知的基础上.流量信息获取与流量预测对合理分配网络资源至关重要.早在 1993 年 Leland 等就开展了互联网流量分析与建模研究^[6].目前的研究主要涉及流量建模和流量预测两方面^[7].前者是通过建立数学模型来描述网络流量特性,后者是基于历史流量信息使用恰当的数学模型对未来流量进行预测.

由于主干网流量在一些时间段具有峰值和变化迅速的特点,因此假设其流量是静态的显然不切合实际,所以需要能够动态获知流量的变化.基于不同的动态流量获取方法,可以把拓扑决策分为反应式或前瞻式.反应式决策需要监测单元实时测量网络上的流量信息并将监测结果发给决策单元,供其进行拓扑管理,然而此过程所需的时间对系统性能的影响大,因此通常对反应时间有较为苛刻的要求,以减少因决策滞后性而对拓扑管理正确性产生的影响.前瞻式决策基于流量预测,可以更加及时地进行拓扑决策,避免了反应式固有的决策滞后问题.本文采用前瞻式决策进行拓扑管理.

回归模型、时间序列模型、灰色预测、神经网络、模糊预测、小波预测、统计学习预测等^[8],已经被广泛应用于网络流量预测.由于自回归和马尔可夫方法建模简单,计算快捷,而且可以进行多步预测,因此本文采用自回归和马尔可夫方法进行流量预测,根据它们的预测结果做出判定,以提高预测的可靠性.

在绿色互联网中,通过节能机制可以节省运营开支.因此,有些网元的拥有者有可能通过向节能机制提供虚假信息来骗取休眠,以降低自身的能耗,从

而降低成本.然而,这种个体的自私行为势必影响网络做出合理的拓扑管理决策,不利于全网的优化节能,因此需要能够合理分配网络资源且具有占优策略激励兼容特性的防策略机制.所谓占优策略是指在博弈中不论竞争对手的策略如何改变,参与者当前使用的策略都是最优的.对于每个参与者而言,如果其提供真实信息是占优策略,则该机制被称为满足占优策略激励兼容特性^[9].在非对称信息博弈中,如果所有参与者都没有动机对其他参与者说谎或是隐藏其私有信息,所有参与者的占优策略都是报出其真实成本,则称其是防策略机制^[10].本文针对网元的配置管理,采用微观经济学中的 VCG(Vickrey-Clarke-Groves)机制^[11]作为防策略机制.VCG 是一类在拟线性效用环境下满足个人理性与策略一致的机制,它能使提供真实报价成为占优策略,其核心在于支付函数的构造.每个参与者的支付值由其对网络的外在性确定^[12],即每个参与者的支付值等于其参与博弈前后其他参与者的价值和之差.

综上,基于主干网基础设施、网络拓扑和网络流量的特点,本文提出了一种绿色互联网中的面向节能的分布式拓扑管理机制(DTME),通过信息感知、流量预测、分布式拓扑决策和休眠控制之间的协同工作,满足主干网的节能需要.本文的主要贡献如下:面向绿色互联网的需要,构建了网络、节点、链路和功耗模型;提出了信息感知、流量预测、分布式拓扑决策和休眠控制一体化协同式的分布式拓扑管理机制;给出了基于自回归和马尔可夫方法的流量预测方法,既有利于避免信息获取的滞后性,也有利于做出合理的拓扑配置决策;给出了基于 VCG 机制的网元配置管理方法,不仅有利于实现网络节能最大化,而且防止了虚假信息可能导致的性能下降和管理机制的崩溃;给出了基于休眠唤醒的主动拓扑自适应方法,不仅可以满足流量需求,而且可以充分发挥网络节能潜力.

本文第 2 节简述网络拓扑管理、流量预测以及 VCG 机制等的研究与应用现状;第 3 节介绍本文构建的网络、节点、链路和功耗模型;第 4 节描述 DTME 的组成与工作机制;第 5 节对 DTME 进行性能评价;第 6 节得出结论.

2 相关工作

已有的集中式拓扑管理方法多集中在长期相对

静态的网络拓扑配置规划等上.文献[13]对网络拓扑进行集中规划,根据网络状况对网络拓扑进行配置,但是配置是反应式的,因此存在一定的滞后性,难以应对流量快速变化的情形,也不适用于高流量的大规模网络(如互联网的主干网)拓扑的动态配置.文献[14]在文献[13]的基础上,在集中式决策过程中引入服务质量和链路利用率作为参数,使网络拓扑能在节能和性能间达到平衡,但是其针对流量的变化仍然是反应式的,随着拓扑配置计算量的增加,难以迅速做出决策.文献[15]把网络节点划分为汇聚节点和边缘节点,并将网络节能问题归结为受限多商品最小流问题,提出了关闭最小流链路和随机选择链路两种启发式算法求解该问题,集中式地完成对节点间链路的配置以达到节能目的,但是该方法的可行性是建立在能够提前获取网络全局信息的前提下,而且仅限于求解已知静态流量下的网络拓扑设计问题,没有考虑流量动态变化情况下的网络拓扑配置.

已有的分布式拓扑管理方法大都基于静态网络流量而且不保证网络拓扑配置对网络节能管理是占优策略.此外,针对主干网的分布式拓扑管理研究也比较少.文献[16]对 OSPF 协议进行增强,使其具有能量感知能力,每个节点分布式地采用能量感知路由策略,确定迁移某个(些)链路上的流量对其他链路的影响,从而能在最低限度影响当前拓扑的情况下关闭低利用率的链路以实现节能,但不能保证针对某个(些)链路上的流量迁移是占优策略,也不能保证全网节能收益最大.文献[17]提出了一种分布式自适应节能机制,根据一天中不同时段流量变化情况动态开关路由器线卡,但是该方法假设流量在一个时段内是不变的.文献[18]基于图论和复杂网络理论,提出了一个基于代数连通性的节能算法,通过计算邻接矩阵的拉普拉斯矩阵第二特征值来衡量拓扑的连通性,并以此作为拓扑配置的依据.文献[19]通过求解多拓扑链路权重分配问题来实现 IP 网络节能的目的.它使用拉格朗日松弛方法将整数线性规划问题转化为其对偶问题,采用多种启发式算法对其进行求解,但是其采用的方法仅适用于在预先确定的多个不同时间段内的能量感知拓扑管理而不适用于随时间动态变化的能量感知拓扑管理.文献[20]提出了一种网络节能的分布式解决方案,给出了一个基于拓扑的启发式算法,通过识别并关闭可以休眠的链路,对拓扑进行修剪,以实现节

能,但是这种方案不是流量感知的,仍然是基于网络静态流量的假设,难以在实际网络中应用。

网络流量预测领域已有很多研究工作.文献[21]提出了一种改进的基于小波变换和分数自回归滑动平均模型的流量预测算法.文献[22]对具有长相关性的自相似流量的建模和预测进行了研究,提出了一种基于自回归滑动平均模型的经验模式分解预测自相似流量的方法.文献[23]针对时变的网络流量,提出了一种智能优化分布式网络流量预测方法.文献[24]提出了一种基于经验模式分解和时间序列分析的网络流量预测方法.文献[25]针对各时刻网络流量时间序列的非线性和多时间尺度特性,提出了把灰色系统理论和马尔可夫方法相结合的流量预测方法.相比较而言,采用自回归和马尔可夫方法的预测方法的时间复杂度更低,更适合动态流量预测。

VCG 机制具有抑制虚假信息且保证“说真话”是占优策略的特性,作为一种防策略机制,在竞标竞价、资源分配和路由选择等领域得到广泛应用.文献[26]针对认知无线网络中的动态频谱分配问题,建立了一种非合作博弈模型,基于 VCG 机制,提出了一种可以在满足主用户服务质量的前提下最大化网络系统吞吐量且保证公平性的分布式动态频谱分配方法.文献[27]根据移动自组织网络的特点,通过对恶意节点的路由欺骗行为进行分析,在 VCG 机制的基础上,提出了一种安全路由策略,可以在保证网络高效传输数据分组的前提下,杜绝恶意节点的欺骗行为.文献[28]提出了一种基于 VCG 机制的防策略信任机制,可以解决普适计算中存在的自私用户为最大化其利益而故意策略性谎报推荐信息的问题,用户只能在提供真实推荐时,才能实现其效用的最大化.文献[29]结合 VCG 机制与物联网的特点,给出了一种安全路由协议,可以抵御欺骗攻击和诽谤攻击.文献[30]提出了一种基于 VCG 机制的网络带宽分配机制,通过有效的支付惩罚,抑制自私用户的说谎动机,可以解决由于自私用户骗取资源而造成的分配不合理问题.上述工作采用 VCG 机制,实现了抑制虚假信息的目的,解决了其面对的特定领域的具体问题,但是由于应用环境与问题导向的不同,这些方法还不能直接应用于本文要解决的网络节能问题。

3 问题描述

本文将网络建模为连通图 $G(V, E)$, 其中, $V =$

(v_1, \dots, v_n) 表示节点集合, $E = (e_1, \dots, e_m)$ 表示边集合. 与节点 v_i 相邻的节点 v_j 称为 v_i 的邻居节点, v_i 的所有邻居节点组成其邻居节点集 $NEI_i, v_j \in NEI_i$; 为 v_i 转发分组的节点 v_f 称为 v_i 的转发节点; 若存在路径使得 v_i 的分组能够到达节点 v_g , 则称 v_g 为 v_i 的可达节点, v_i 的所有可达节点组成其可达节点集 $REA_i, v_g \in REA_i$; 若把节点 v_c 视为 v_i 的非转发节点将导致 v_i 对 v_g 不可达, 则称 v_c 为 v_i 和 v_g 的瓶颈节点. 每个节点都采用最短路径算法, 以链路功耗与流量的比值为权重, 计算到其他可达节点的最短路径。

节点结构如图 1 所示, 包括主控引擎、背板、底座、交换结构、线卡、转发引擎、复制引擎、端口等构件, 其中, 主控引擎是路由器的控制中心, 用于完成分组头部分析和路由表查找等功能; 背板由数据总线和交换结构组成, 是路由器内部数据交换通道; 底座用于承载线卡和交换结构, 为线卡提供连接槽位; 交换结构用于在路由器内部连接线卡的输入端口和输出端口; 线卡用于实现分组处理、队列调度和流量管理等功能; 转发引擎用于完成分组输入、存储与转发等功能; 复制引擎用于组播所需的分组复制; 端口用于连接路由器和外部线路, 并在两者之间进行数据传输。

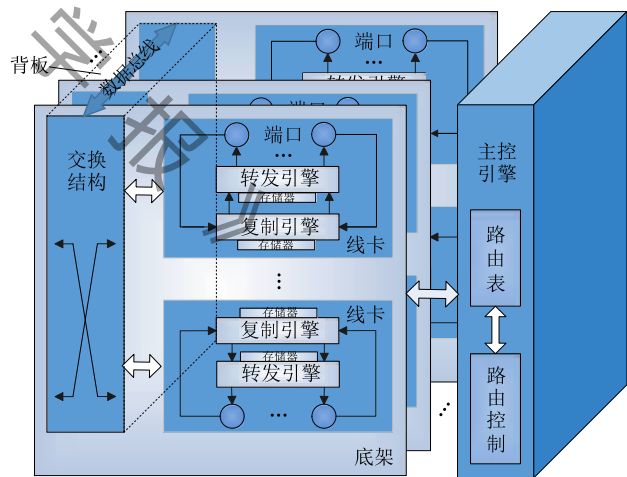


图 1 节点结构

链路结构如图 2 所示, 包括功率放大器、在线放大器、光再生器和前置放大器等中间设备. 其中, 功率放大器用来提高信号发送功率; 在线放大器用来延长信号传输距离; 光再生器用来对信号进行整形; 前置放大器用来改善接收端灵敏度。

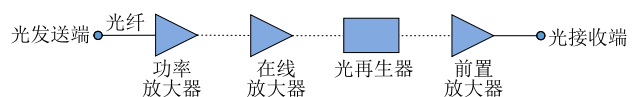


图 2 链路结构

根据节点内部各构件间的相互关系以及工作状态,可以得到节点功耗模型,如式(1)所示,其中, P_{cha}^i 表示节点 v_i 中底架(背板交换结构的功耗计入到对应的底架中,数据总线功耗忽略不计)的基准功耗(不插任何线卡时底架的功耗), P_{ctrl}^i 表示节点 v_i 中主控引擎的功耗, $P_{\text{lc_for}}^i$ 表示节点 v_i 中线卡 lc 上转发引擎的基准功耗, $P_{\text{lc_rep}}^i$ 表示节点 v_i 中线卡 lc 上复制引擎的基准功耗, $P_{\text{lc_mem}}^i$ 表示节点 v_i 中线卡 lc 上存储器的功耗, $P_{\text{lc_port}}^i$ 表示节点 v_i 中线卡 lc 上端口的基准功耗。 N_{cha}^i 表示节点 v_i 中底架的数量, N_{lcrd}^k 表示节点 v_i 中底架 k 上线卡的数量, $N_{\text{port}}^{\text{lc}}$ 表示节点 v_i 中底架 k 的线卡 lc 上端口的数量, trf_{lc} 表示节点 v_i 中经过线卡 lc 的流量, trf_p 表示节点 v_i 中经过端口 p 的流量; α 、 β 为常数,用来确定流量与功耗的关系; CS_i^k 表示节点 v_i 中底架 k 的状态, LSt_k^{lc} 表示节点 v_i 中底架 k 的线卡 lc 的状态, PSt_{lc}^p 表示节点 v_i 中底架 k 的线卡 lc 上端口 p 的状态,取值0或1,其中,0表示处于休眠状态,1表示处于工作状态。

$$P_i^{\text{node}} = P_{\text{ctrl}}^i + \sum_{k=1}^{N_{\text{cha}}^i} \left(P_{\text{cha}}^i \times \text{CS}_i^k + \sum_{lc=1}^{N_{\text{lcrd}}^k} \left((P_{\text{lc_mem}}^i + (P_{\text{lc_for}}^i + P_{\text{lc_rep}}^i) \times (1 + \alpha \cdot (\text{trf}_{lc})^\beta)) \times \text{LSt}_k^{\text{lc}} + \sum_{p=1}^{N_{\text{port}}^{\text{lc}}} (P_{\text{lc_port}}^i \times (1 + \alpha \cdot (\text{trf}_p)^\beta) \times \text{PSt}_{lc}^p) \right) \right) \quad (1)$$

基于本文给出的链路结构和文献[31]中给出的参考链路长度和参考链路基准功耗,可得链路功耗模型,如式(2)所示:

$$P_l^{\text{link}} = \lfloor \text{len}_l / \text{len}_{\text{ref}} \rfloor \times P_{\text{ref}_l} \times (1 + \alpha \cdot (\text{trf}_l)^\beta) \quad (2)$$

其中, P_l^{link} 为链路 l 的功耗, P_{ref_l} 为参考链路基准功耗, len_l 为链路 l 的实际长度, len_{ref} 为参考链路长度, trf_l 为链路 l 上的流量。

4 分布式拓扑管理

本文提出的面向节能的分布式拓扑管理机制由信息感知、流量预测、分布式拓扑决策和休眠控制等构件组成,通过它们的协同工作,支持主干网的节能。其中,信息感知构件感知节点状态和链路流量信息,调节信息感知频率;流量预测构件根据历史数据对链路流量进行预测,为休眠唤醒决策提供依据;分布式拓扑决策构件决定链路的休眠和唤醒;休眠控制构件负责链路休眠和唤醒操作。

4.1 信息感知

该构件负责处理与邻居节点的信息交换,定期

记录功耗与流量等数据,根据变化情况调整感知间隔。该构件包括感知管理、节点感知和流量感知3个子构件。

(1) 感知管理子构件

它负责监控节点感知子构件和流量感知子构件的运行,调节感知时间间隔。

节点感知和流量感知都是周期性进行的,两者采用一致的感知间隔。感知间隔长短与流量变化程度相关,若变化剧烈,则缩短感知间隔,否则延长感知间隔。流量变化程度采用式(3)计量:

$$\text{trf}_{\text{ch}} = \begin{cases} (|\text{trf}_{\text{last}} - \text{trf}_{\text{now}}|) / \text{trf}_{\text{last}}, & \text{trf}_{\text{last}} \neq 0 \\ 0, & (\text{trf}_{\text{last}} = 0) \wedge (\text{trf}_{\text{now}} = 0) \\ \epsilon, & (\text{trf}_{\text{last}} = 0) \wedge (\text{trf}_{\text{now}} \neq 0) \end{cases} \quad (3)$$

其中, trf_{ch} 为本次流量变化程度, trf_{last} 为上次流量, trf_{now} 为当前流量, ϵ 为常数, $0 < \epsilon < 1$ 。显然,流量变化越剧烈, trf_{ch} 越大。

感知间隔采用如下公式计算:

$$t_{\text{in}} = \begin{cases} \theta \cdot t_{\text{in_last}}, & (\text{trf}_{\text{ch}} \leq \gamma) \wedge (t_{\text{in_last}} < (t_{\text{in_sta}} / \theta)) \\ t_{\text{in_sta}}, & (\text{trf}_{\text{ch}} \leq \gamma) \wedge (t_{\text{in_last}} \geq (t_{\text{in_sta}} / \theta)) \\ t_{\text{in_rap}}, & \text{其他} \end{cases} \quad (4)$$

其中, t_{in} 为本次流量感知间隔; $t_{\text{in_last}}$ 为上次流量感知间隔; $t_{\text{in_sta}}$ 为稳定流量感知间隔,即流量变化较平稳时设置的流量感知间隔; θ 为间隔调节系数, $0 < \theta < 1$; γ 为流量变化程度衡量系数,若流量变化程度高于 γ ,则认为流量变化剧烈; $t_{\text{in_rap}}$ 为突变流量感知间隔,即流量变化剧烈时设置的流量感知间隔。

(2) 节点感知子构件

它负责收集、存储和处理节点功耗、链路功耗和节点资源使用信息,周期性感知节点各部件的工作情况,计算各部件功耗和资源使用情况,记录与节点相连的链路的休眠时间。

(3) 流量感知子构件

它负责感知经过端口的流量信息,周期性统计各端口流量,存储到端口流量信息表中,在流量预测时使用。

每个端口使用流量信息表记录相连链路的流量信息,以便进行流量预测。流量感知周期性进行,不保留所有的历史流量信息,在记录最新流量信息的同时,删除过期的流量记录。对此,本文采用滑动窗口机制,节点只记录最近 n 次感知的结果,淘汰此前的流量记录。

4.2 流量预测

流量预测构件根据端口中记录的历史流量数据 $\{x_h | h=1, \dots, N\}$,使用自回归和马尔可夫方法对端

口未来流量进行预测,帮助确定端口休眠时间。

对于自回归方法,本文使用的自回归模型表示如下:

$$x_k = \sum_{i=1}^r \varphi_i x_{k-i} + \delta \quad (5)$$

其中, φ_i 为第 i 个历史数据的权值, r 为自回归模型的阶数, δ 表示误差参数。

本文首先给历史流量数据分配权值,然后加权求和得出预测值,计算方法如式(6)~(8),其中, ρ_i 为样本自相关函数, $x'_k(s)$ 表示 x_k 的第 s 期预测结果。如果式(7)中由 ρ_i 组成的矩阵没有逆矩阵,则自回归方法失效。

$$\rho_i = \left(\sum_{k=i+1}^N x_k x_{k-i} \right) / \left(\sum_{k=1}^N x_k^2 \right) \quad (6)$$

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_r \end{bmatrix} = \begin{bmatrix} 0 & \rho_1 & \cdots & \rho_{r-1} \\ \rho_1 & 0 & \cdots & \rho_{r-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{r-1} & \rho_{r-2} & \cdots & 0 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_r \end{bmatrix} \quad (7)$$

$$x'_k(s) = \begin{cases} \sum_{i=1}^r \varphi_i x_{k-i+1}, & s=1 \\ \sum_{i=1}^{s-1} \varphi_i x'_k(s-i) + \sum_{i=s}^r \varphi_i x_{k+s-i}, & 1 < s \leq r \\ \sum_{i=1}^r \varphi_i x'_k(s-i), & s > r \end{cases} \quad (8)$$

依据马尔可夫方法,首先将历史流量数据划分为 a 种不同状态,如式(9)所示。

$$trfst_{x_h} = i, \frac{(i-1) \times bw}{a} \leq x_h < \frac{i \times bw}{a}, h=1, \dots, N \quad (9)$$

其中, $trfst_{x_h}$ 表示 x_h 所属的流量状态, bw 为端口链路带宽, $i \in \{1, \dots, a\}$ 。然后,根据历史流量数据的状态变化情况,计算出将来发生某种状态变化的概率,如式(10)和(11)所示,其中, $p_{ij}^{(b)}$ 为最近 b 期内流量从状态 i 转移到状态 j 的统计概率, $x(t)$ 表示第 t 期的流量值, $t \in \{1, \dots, s\}$ 。根据第 1 期的状态向量和状态转移概率矩阵,计算第 2 期的状态向量;重复此过程,直至计算出第 s 期的状态向量为止;将每期的状态向量中最大概率对应的状态作为当期的预测结果,若此最大概率不唯一,则马尔可夫方法失效。

$$\begin{bmatrix} p_1^{t+1} \\ \vdots \\ p_a^{t+1} \end{bmatrix}^T = \begin{bmatrix} p_1^t \\ \vdots \\ p_a^t \end{bmatrix}^T \begin{bmatrix} p_{11}^{(b)} & \cdots & p_{1a}^{(b)} \\ \vdots & \ddots & \vdots \\ p_{a1}^{(b)} & \cdots & p_{aa}^{(b)} \end{bmatrix} \quad (10)$$

$$p_i^t = \begin{cases} 1, & i = trfst_{x(t)} \\ 0, & i \neq trfst_{x(t)} \end{cases} \quad (11)$$

流量预测构件工作流程如下:

(1) 节点 v_i 监测端口,若发现当前时刻距上次预测时刻达到规定的时间间隔阈值 t_{th} ,则对端口启动本次流量预测;否则结束。

(2) 分别采用自回归方法和马尔可夫方法进行一步预测,若两种方法都失效,则预测失败,返回(1);若两种预测方法只有一个成功,则将其预测值作为本期预测结果;若两种预测方法都成功,则取它们的平均值作为本期预测结果。

(3) 判断预测是否执行了 s 期(s 根据预测准确度需求和预测间隔而定),若预测期数小于 s 期,则转(2)。

(4) 从本次的 s 期预测结果中,剔除值最高的前 5%,将剩余结果中的最大值作为本次预测结果。

每次预测剔除前 5% 的目的主要是为了防止出现个别强“突发”流量对预测结果准确度的干扰。

4.3 分布式拓扑决策

该构件定期从信息感知构件获取流量和功耗等信息,根据这些信息采用 VCG 机制对节点相连的链路进行配置规划,并对链路休眠唤醒进行管理。当节点感知到网络状况出现较大波动时,会触发该构件的运行。

该构件由配置子构件和休眠唤醒子构件组成。前者基于 VCG 机制对链路进行配置,并且抑制虚假信息的影响;后者管理休眠唤醒以及对突发事件进行处理。

本文采用 VCG 机制防止虚假信息对拓扑管理造成负面影响。假设在 VCG 机制中有 χ 个参与者,每个参与者有自己的私人信息 $pt_i \in T_i$, T_i 为参与者 i 的类型空间, $i=1, \dots, \chi$ 。所有参与者都知道 T_i ,但只有 i 知道 pt_i 。参与者选择参与策略,构成策略组 $pt = (pt_1, \dots, pt_\chi)$ 。把策略组作为输入,经计算得到输出结果 $o(pt)$,简记为 o ,作为分配规则。记 o 为所有分配规则组成的集合。设转移支付值为 $pay_i(pt)$,作为支付规则, $i=1, \dots, \chi$ 。不同参与者对结果有不同的评价,称其为结果 o 的价值 $value_i(pt, o)$ 。参与者获得的效用定义如下:

$$utility_i(pt, o) = value_i(pt, o) - pay_i(pt) \quad (12)$$

VCG 机制的分配策略用如下表示:

$$o^*(pt) = \arg \max_{o \in O} \sum_i value_i(pt, o) \quad (13)$$

其中, $o^*(pt)$ 表示期望结果。显然, VCG 机制实现的是社会福利函数,使所有参与者的价值之和最大。

VCG 的支付规则定义如下:

$$pay_i(pt) = \sum_{j \neq i} value_j(pt_j, o_{-i}^*) - \sum_{j \neq i} value_j(pt_j, o^*) \quad (14)$$

式(14)表示参与者 i 的支付值等于其参与博弈前后其他参与者的价值和之差。 o_{-i}^* 表示参与者 i 不参与博弈时的分配规则。

每个参与者最终获得的效用等于其给整个机制带来的价值增加值,定义如下:

$$utility_i(pt_i, o^*) = \sum_j value_j(pt_j, o^*) - \sum_{j \neq i} value_j(pt_j, o_{-i}^*) \quad (15)$$

该构件基于 VCG 机制为节点的每一条链路计算链路利益值和链路成本值. 链路利益值代表链路处于工作状态时的利益价值, 表示对该条链路的偏好. 链路成本值代表链路处于工作状态时的成本. 由于链路利益值和链路成本值的计算需要综合考虑链路流量的当前状况和未来变化趋势, 仅依据当前流量或预测流量进行计算不全面, 因此引入链路综合流量. 链路 l_{ij} 的利益值 $benefit_{l_{ij}}$ 、成本值 $cost_{l_{ij}}$ 、综合流量 $trf_{l_{ij}}^*$ 、价值 $value_{l_{ij}}$ 、支付值 $pay_{l_{ij}}$ 和效用值 $utility_{l_{ij}}$ 以及节点 v_i 的价值 $value_i$ 分别定义如下:

$$benefit_{l_{ij}} = \frac{Power_{l_{ij}}^{node} \times (trf_{l_{ij}} / trf_i)}{trf_{l_{ij}}^*} \quad (16)$$

$$cost_{l_{ij}} = Power_{l_{ij}}^{link} / trf_{l_{ij}}^* \quad (17)$$

$$trf_{l_{ij}}^* = trf_{l_{ij}} \times \left(1 + \lambda \cdot \frac{trf_{l_{ij}} - trf'_{l_{ij}}}{|trf_{l_{ij}} - trf'_{l_{ij}}|} (trf_{l_{ij}} - trf'_{l_{ij}})^\eta \right) \quad (18)$$

$$value_{l_{ij}} = benefit_{l_{ij}} - cost_{l_{ij}} \quad (19)$$

$$pay_{l_{ij}} = value_i^{-l_{ij}} - value_i \quad (20)$$

$$utility_{l_{ij}} = value_{l_{ij}} - pay_{l_{ij}} \quad (21)$$

$$value_i = \sum_{v_g \in REA_i} value_{path_{ig}} \quad (22)$$

其中, $trf_{l_{ij}}$ 为链路 l_{ij} 的当前流量, $trf'_{l_{ij}}$ 为链路 l_{ij} 的预测流量, trf_i 为节点 v_i 的当前流量, $Power_{l_{ij}}^{node}$ 为节点 v_i 的功耗, $Power_{l_{ij}}^{link}$ 为链路 l_{ij} 的功耗, λ 和 η 为常数, $\lambda, \eta \in (0, 1)$ 用于调节流量变化趋势对综合流量的影响程度, $value_i$ 为节点 v_i 的价值, 定义为该节点到其所有可达节点的路径的价值之和, $value_i^{-l_{ij}}$ 为从当前网络拓扑中去除链路 l_{ij} 后节点 v_i 的价值。

路径 $path_{ig}$ 的利益值 $benefit_{path_{ig}}$ 、成本值 $cost_{path_{ig}}$ 、价值 $value_{path_{ig}}$ 、支付值 $pay_{path_{ig}}$ 和效用值 $utility_{path_{ig}}$ 分别计算如下:

$$benefit_{path_{ig}} = \sum_{l_{path_{ig}} \in path_{ig}} benefit_{l_{path_{ig}}} \quad (23)$$

$$cost_{path_{ig}} = \sum_{l_{path_{ig}} \in path_{ig}} cost_{l_{path_{ig}}} \quad (24)$$

$$value_{path_{ig}} = \sum_{l_{path_{ig}} \in path_{ig}} value_{l_{path_{ig}}} \quad (25)$$

$$pay_{path_{ig}} = \sum_{l_{path_{ig}} \in path_{ig}} pay_{l_{path_{ig}}} \quad (26)$$

$$utility_{path_{ig}} = \sum_{l_{path_{ig}} \in path_{ig}} utility_{l_{path_{ig}}} \quad (27)$$

其中, $l_{path_{ig}}$ 为路径 $path_{ig}$ 上的链路。

路径利益矩阵和路径成本矩阵分别由网络中所有节点间的路径利益值和路径成本值组成, 分别定义如下:

$$BENE = \begin{bmatrix} benefit_{path_{11}} & \cdots & benefit_{path_{1n}} \\ \vdots & \ddots & \vdots \\ benefit_{path_{n1}} & \cdots & benefit_{path_{nn}} \end{bmatrix} \quad (28)$$

$$COST = \begin{bmatrix} cost_{path_{11}} & \cdots & cost_{path_{1n}} \\ \vdots & \ddots & \vdots \\ cost_{path_{n1}} & \cdots & cost_{path_{nn}} \end{bmatrix} \quad (29)$$

其中, $benefit_{path_{ii}} = 0$, $cost_{path_{ii}} = 0$, 且置不可达节点间的路径利益值和路径成本值为 0。

配置子构件工作流程如下:

(1) 当节点 v_i 从其信息感知构件得知网络状况发生波动时, 更新路径利益矩阵和路径成本矩阵。

(2) 节点 v_i 向所有邻居节点 $v_j \in NEI_i$ 发送路径成本矩阵。

(3) 邻居节点 $v_j \in NEI_i$ 接收到路径成本矩阵后, 对节点 v_i 的所有不是其邻居节点的可达节点 $v_g \notin NEI_i$, 计算 v_i 通过 v_j 到达 v_g 的路径成本 $cost_{path_{ig}}^j$, 如果 $cost_{path_{ig}}^j$ 小于原来 v_i 到 v_g 的路径成本 $cost_{path_{ig}}$, 则更新 $path_{ig}$ 中 v_i 的下一跳为 v_j 且更新 v_i 到 v_g 的路径成本 $cost_{path_{ig}}$ 为 $cost_{path_{ig}}^j$ 。

(4) 对所有邻居节点 $v_j \in NEI_i$, 节点 v_i 更新链路 l_{ij} 的支付值 $pay_{l_{ij}}$ 。如果 v_j 为瓶颈节点, 则 $pay_{l_{ij}} = pay_{l_{ij}} + benefit_{path_{ig}} - cost_{path_{ig}}$; 否则, $pay_{l_{ij}} = pay_{l_{ij}} + cost_{path_{ig}}^j - cost_{path_{ig}}$ 。

休眠唤醒子构件工作流程如下:

(1) 当节点 v_i 完成对所有链路 l_{ij} 的支付值 $pay_{l_{ij}}$ 的更新之后, 计算链路 l_{ij} 的效用值 $utility_{l_{ij}}$; 如果 $utility_{l_{ij}} < 0$, 则链路 l_{ij} 可转入休眠状态。

(2) 按式(30)计算最短休眠时间 t_{hb} 。

$$t_{hb} = \frac{t_{ch} \times (Power_{ch} - Power_{hb})}{\omega \times (Power_{wk} - Power_{hb})} \quad (30)$$

其中, t_{ch} 为端口苏醒所需的状态转换时间, $Power_{hb}$

为端口休眠时的待机功耗, $Power_{ch}$ 为端口苏醒瞬间的功耗, $Power_{wk}$ 为正常工作状态下的端口功耗, ω 为常数, 用来调节休眠时长, $0 < \omega < 1$.

(3) 调用休眠控制构件处理链路 l_{ij} 休眠.

(4) 更新链路 l_{ij} 利益值和链路 l_{ij} 成本值. 因为链路 l_{ij} 转入休眠, 其上流量值为 0, 定义休眠时链路利益值和成本值如下:

$$benefit_{l_{ij}} = \begin{cases} 0, & t_{l_{ij}} < t_{hb} \\ \left(\frac{Power_i^{node} + Power_{wk}}{trf_i + trf'_{l_{ij}}} \right)^{\zeta \times (t_{l_{ij}} - t_{hb} + 1)}, & t_{l_{ij}} \geq t_{hb} \end{cases} \quad (31)$$

$$cost_{l_{ij}} = \begin{cases} \infty, & t_{l_{ij}} < t_{hb} \\ (Power_{wk} / trf'_{l_{ij}})^{-\zeta \times (t_{l_{ij}} - t_{hb} + 1)}, & t_{l_{ij}} \geq t_{hb} \end{cases} \quad (32)$$

其中, $t_{l_{ij}}$ 为链路 l_{ij} 休眠的时间, ζ 为常数, $0 < \zeta < 1$. 式(31)表明, 在链路 l_{ij} 休眠初期(休眠时间 $t_{l_{ij}}$ 小于最短休眠时间 t_{hb} 时), 休眠收益为 0, 当休眠时间 $t_{l_{ij}}$ 已到或已过最短休眠时间 t_{hb} 时, 休眠收益随休眠时间的增加而增加. 式(32)表明, 在链路 l_{ij} 休眠初期, 唤醒成本大于休眠收益, 除遇突发事件外, 这种情况是应该避免的; 当休眠时间 $t_{l_{ij}}$ 已到或已过最短休眠时间 t_{hb} 时, 唤醒成本随休眠时间的增加而减小.

(5) 如果链路 l_{ij} 所属节点的其他工作链路超载或即将超载, 则按照所有当前已休眠链路的休眠时间从长到短依次激活休眠链路. 每激活一条链路, 就检测一次当前所有工作链路的负载情况, 直到超载情况得到缓解或者已经不存在休眠链路.

4.4 休眠控制构件

该构件负责控制链路和其两侧的节点端口的休眠和苏醒, 对即将休眠的链路的两侧节点端口进行休眠前协商, 对苏醒的链路的两侧节点端口进行苏醒通告, 将链路的休眠或唤醒状态通知分布式拓扑决策构件.

链路的一侧节点端口在正式进入休眠之前要通告其另一侧的对应端口(其邻居节点上的对应端口). 如果链路相连的任一侧节点上存在未休眠的端口, 则该节点上与之相关的系统部件(底架、主控引擎等)不能休眠, 需继续为未休眠的活跃端口提供服务; 如果相关的所有端口都休眠, 则部件进入休眠; 如果所有部件均休眠, 则该节点进入休眠.

链路的任一侧节点端口苏醒有两种情况: 休眠时间已过; 遇突发情况, 需要紧急启动. 休眠端口苏

醒时, 启动端口并发送苏醒请求; 邻居节点收到苏醒请求后启动相应端口, 回送苏醒确认; 请求苏醒的节点收到苏醒确认后, 把链路状态通知给分布式拓扑决策构件.

5 仿真实现与性能比较

本文提出的 DTME 在 Windows 7 professional 64 bits 平台下使用 Microsoft Visual C++ 6.0 进行仿真实现. 在分析与评价 DTME 的性能时, 本文基于文献[32]提供的网络流量数据, 选用如图 3 所示的较高且较平稳的流量(流量用例 1)、波动较大的流量(流量用例 2)、较低且较平稳的流量(流量用例 3)作为仿真流量用例. 本文选用中国教育和科研计算机网 CERNET 拓扑(36 个节点和 46 条链路)、第二代中国教育和科研计算机网 CERNET2 拓扑(20 个节点和 22 条链路)、美国国家科学基金会主干网 NSFNET 拓扑(14 个节点和 21 条链路)和美国下一代互联网 INTERNET2 拓扑(56 个节点和 61 条链路)作为仿真拓扑用例, 如图 4 所示.

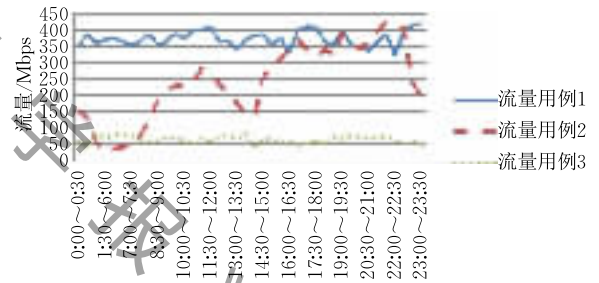


图 3 流量用例

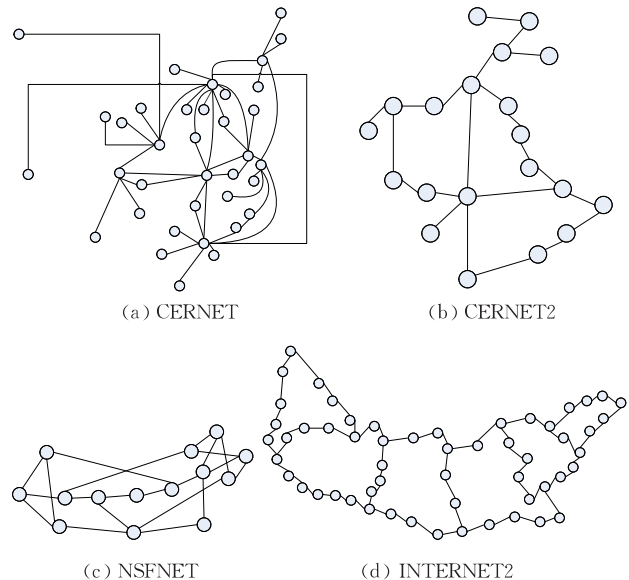


图 4 拓扑用例

5.1 对比机制

本文采用基于 GRiDA^[33] 的分布式拓扑管理机制(以下简称 GRiDA)作为对比机制来分析评价 DTME 的性能。

GRiDA 可以分布式在线控制网络链路是否进入休眠模式, 链路休眠唤醒决策取决于其所属节点, 由该节点基于其相连链路的当前负载、当前功耗以及由链路状态路由协议(OSPF 或 IS-IS)提供的当前拓扑知识进行控制. 每个节点不需要网络全局流量信息, 不需要与其他节点协作和同步. 由于 GRiDA 是面向主干网节能提出的分布式拓扑管理机制, 与本文提出的 DTME 在研究背景、应用场景以及拓扑管理方式上有可比性, 因此本文选取它作为对比机制。

5.2 性能评价

本文从网络功耗和网络性能等方面对比 DTME 和 GRiDA. 仿真中使用的参数参考思科 12000 系列路由器^①和文献[31]进行设置, 如表 1 所示。

表 1 仿真参数设置

参数	取值
节点中主控引擎功耗	356 W
节点中单个底架基准功耗	100 W
节点中单个复制引擎基准功耗	100 W
节点中单个转发引擎基准功耗	446 W
节点中线卡存储器功耗	12 W
节点中单个端口基准功耗	150 W
节点中底架数目	4
底架上的线卡数目	4
线卡上的端口数目	4
参考链路长度	80 km
参考链路基准功耗	1 kW
流量与功耗的关系常数 α	0.032
流量与功耗的关系常数 β	0.82
稳定流量感知间隔	5 s
突变流量感知间隔	1 s

5.2.1 网络功耗对比

在 4 个拓扑用例下分别采用 3 个流量用例测试在 DTME 和 GRiDA 下的网络功耗. 每隔 20 s 计算一次两种机制的网络功耗比, 统计时间持续 24 h, 取其间网络功耗比的算术平均值. 图 5 是两种机制下网络功耗的对比结果. 可以看出, 在不同拓扑和流量下, DTME 都有较好的节能效果. 特别是, 网络流量波动大时, DTME 的表现明显优于 GRiDA. 这主要是因为 DTME 有信息感知间隔自动调节和流量预测功能, 能在一定程度上避免因为流量波动造成的决策失误。

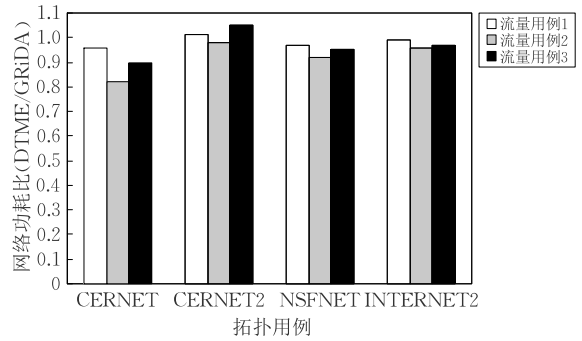


图 5 网络功耗

5.2.2 网络性能对比

本文比较 DTME 和 GRiDA 在收敛时间、平均延迟、分组丢失率和防策略性等性能指标方面的表现。

(1) 收敛时间

收敛时间是指两种机制达到最优网络功耗所需的时间. 由于流量用例 2 中流量波动较为剧烈且中午流量较大, 在此背景下对收敛时间测试更能反映机制收敛性的优劣, 因此采用流量用例 2 中 12:00~13:00 的流量对两种机制的收敛时间在 4 个网络拓扑上进行测试, 每隔 5 s 记录网络功耗, 直至网络功耗收敛至最小值, 结果如图 6. 可以看出, GRiDA 收敛速度慢于 DTME. 这主要是因为 DTME 中节点之间彼此交互休眠唤醒信息和拓扑管理决策信息等, 提高了收敛速度, 相比之下, 在 GRiDA 中节点之间的信息交互少, 导致其收敛缓慢. 此外, 网络拓扑中的节点和链路越多, 两种机制通常收敛越慢, 这是因为在这样的网络拓扑中节点之间需要交互的信息更多。

(2) 平均延迟

平均延迟主要受路径跳数和途经各节点的等待和处理时间的影响. 由于流量用例 2 中的流量波动较为剧烈, 在此背景下测得的是较差情况下的延迟, 因此采用其对 DTME 和 GRiDA 进行测试. 将两种机制分别在 4 个网络拓扑上运行, 每隔 15 min 统计一次数据分组到达目的节点的延迟时间, 统计时间持续 24 h, 取其间延迟的算术平均值, 结果如图 7 所示. 可以看出, GRiDA 的平均延迟高于 DTME. 这主要是因为 GRiDA 中节点之间的相互协调不如 DTME, 导致路径跳数较大, 经过的节点较多, 因而延迟较大. 此外, 对于节点和链路多的网络拓扑, 两者的平均延迟更接近, 这是因为在这样的网络拓扑中两者选取最短路径的可能性更接近。

① Cisco XR 12000 Series and Cisco 12000 Series Routers. <http://www.cisco.com/c/en/us/products/routers/12000-series-routers/datasheet-listing.html>

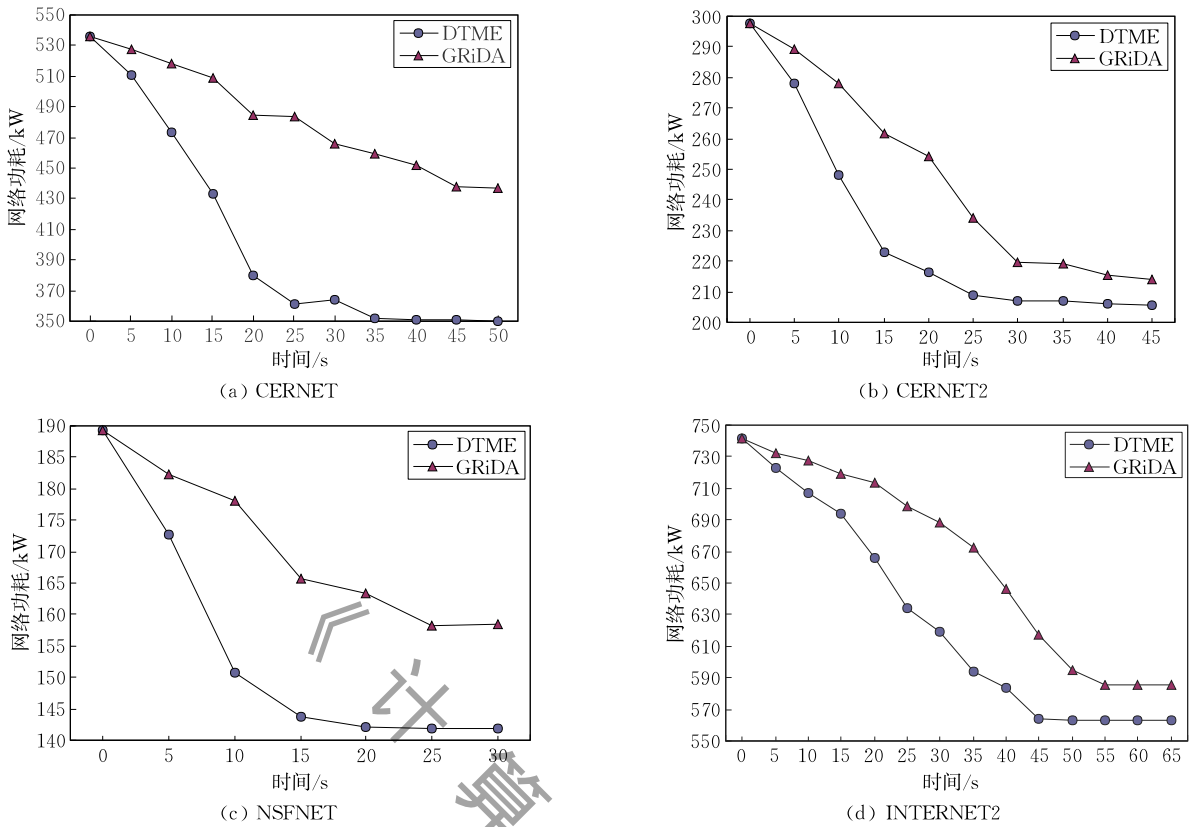


图6 收敛时间

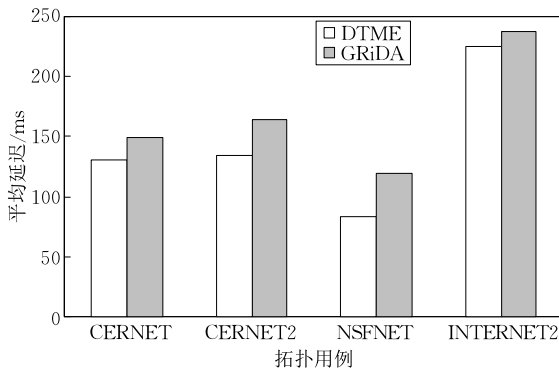


图7 平均延迟

(3) 分组丢失率

采用3个流量用例分别在4个网络拓扑上运行两种机制,对分组丢失率进行测试.每隔15min统计一次分组丢失率,统计时间持续24h,取其间分组丢失率的算术平均值,结果如图8所示.可以看出,在流量用例1下,较高且较平稳的流量使得两者的分组丢失率都高于其在流量用例2下的分组丢失率,这主要是因为流量用例1下处于拥塞状态的链路更多;在流量用例3中,很低且平稳的流量使得两者的分组丢失率极低且几乎相等,这主要是因为网络中多数链路的利用率低所以在很大程度上避免了拥塞的发生.在3个流量用例下,GRiDA的分组

丢失率均比DTME的低.这主要是因为GRiDA比DTME简单,不需要在流量预测和动态信息感知上花费时间,因此对突发流量处理更快,降低了分组丢失率.在节点和链路多的网络拓扑中,两者的分组丢失率更接近,这是因为这样的网络拓扑使得流量到来时有更多的路径可供选择,因而对网络拥塞具有更强的弹性.

(4) 防策略性

通过测试虚假信息对整个网络节能的影响程度来评价机制的防策略性.如前所述,若网络中出现恶意节点骗取节能,则会导致网络整体节能效果下降.记没有任何网元休眠时的网络功耗为 $Power_{net}^{full}$,恶意节点数目为 n_{mal} 时的网络功耗为 $Power_{net}^{n_{mal}}$,无恶意节点时的网络功耗为 $Power_{net}$,定义网络功耗节约量稳定程度 Δ 定义如下:

$$\Delta = \frac{Power_{net}^{full} - Power_{net}^{n_{mal}}}{Power_{net}^{full} - Power_{net}} \quad (33)$$

显然, Δ 值越大,机制的防策略性越好.

不同规模的网络对恶意节点提供虚假信息的容忍程度存在差异,节点和链路多的网络受其影响相对较小,相比之下,节点和链路少的网络受其影响相对较大.为了较为合理地评价DTME和GRiDA的防策略性,本文测试时选择节点和链路数相对适中

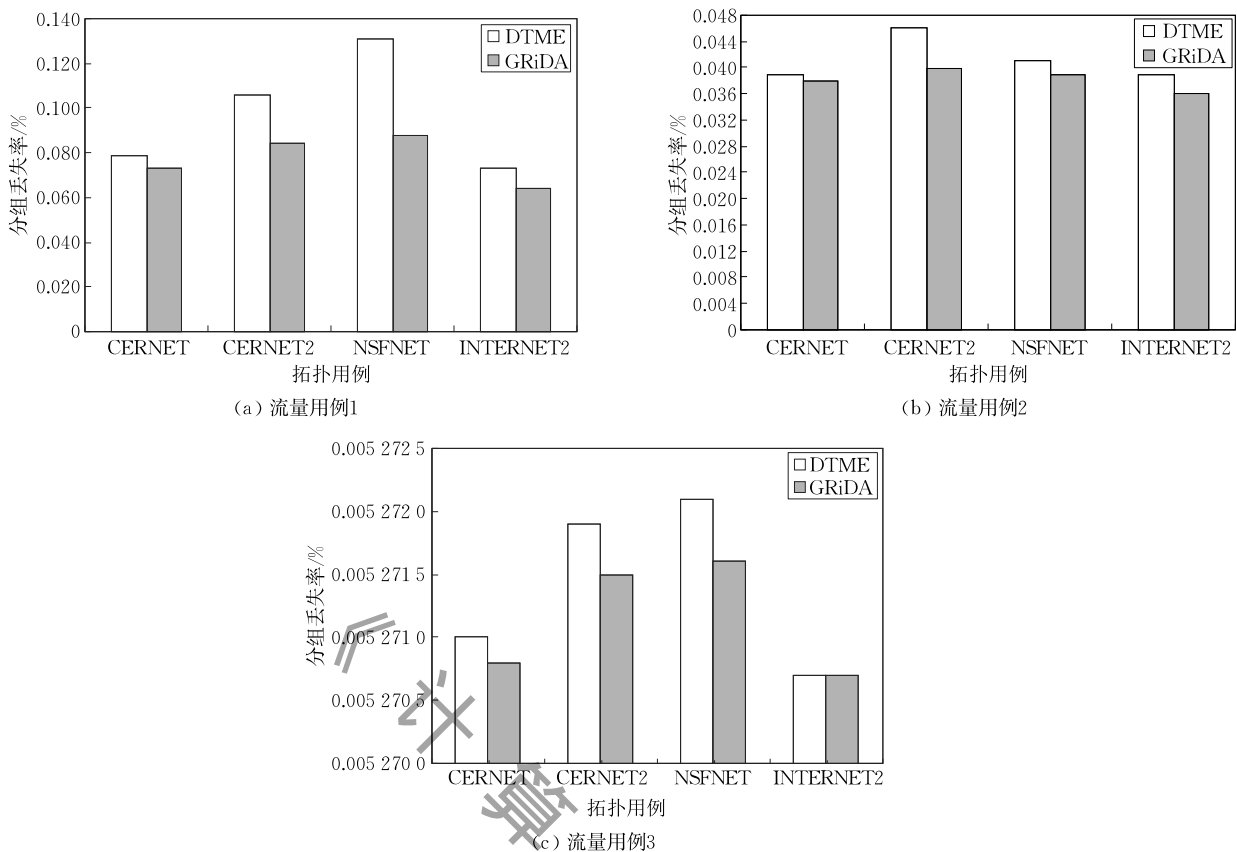


图 8 分组丢失率

的 CERNET. 由于流量用例 1 和流量用例 3 的流量波动较平缓, 节点间交互不活跃, 难以较好地反映恶意节点对网络的影响, 因此本文测试时选取波动剧烈的流量用例 2. 测试时, 逐渐增加提供虚假信息的节点个数, 每隔 15 min 计算一次 Δ 的值, 统计时间持续 24 h, 取其间 Δ 的算术平均值, 结果如图 9 所示. 可以看出, GRiDA 的节能效果受虚假信息的影响较大, 当提供虚假信息的节点数量逐渐增加时, 其节能效果迅速下降; 相比之下, DTME 的节能效果受虚假信息的影响较小, 随着提供虚假信息的节点数量的增多, 其节能效果仍然较好. 这是因为 DTME 引入了 VCG 机制, 具有防策略性, 从而做出错误决策的概率显著降低.

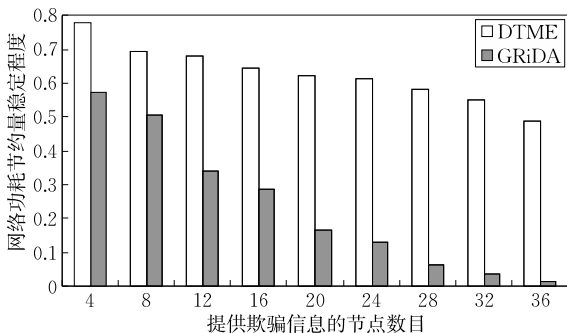


图 9 防策略性

6 总 结

在绿色互联网背景下, 本文构建了主干网的网络模型, 给出了节点和链路的结构模型, 建立了功耗模型. 在此基础上, 为了达到节能目的且避免集中式管理的缺陷, 本文提出了一种面向节能的分布式拓扑管理机制, 进行了仿真实验和性能评价. 实验结果表明, 本文提出的机制不仅具有良好的网络节能效果, 而且在收敛时间、平均延迟、分组丢失率和防策略性等性能指标上均表现良好. 在实际网络背景下进行原型实现和现场试验以进一步验证和提高本文机制的实用性以及进一步增强本文提出的机制各个组成部分的协同工作效率是未来研究工作的重点.

参 考 文 献

- [1] Guichard J, Le Faucheur F, Vasseur J P. Definitive MPLS Network Designs. Indianapolis, USA: Cisco Press, 2005
- [2] Gupta M, Singh S. Greening of the Internet//Proceedings of the 2003 ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Karlsruhe, Germany, 2003: 19-26

- [3] Wang Xing-Wei, Qu Da-Peng, Huang Min, et al. Multiple many-to-many multicast routing scheme in green multi-granularity transport networks. *Computer Networks*, 2015, 93(part 1): 225-242
- [4] Wang Xing-Wei, Cheng Hui, Li Ke-Qin, et al. A cross-layer optimization based integrated routing and grooming algorithm for green multi-granularity transport networks. *Journal of Parallel and Distributed Computing*, 2013, 73(6): 807-822
- [5] Wang Xing-Wei, Hou Wei-Gang, Guo Lei, et al. Energy saving and cost reduction in multi-granularity green optical networks. *Computer Networks*, 2011, 55(3): 676-688
- [6] Bolla R, Davoli F. The potential impact of green technologies in next-generation wireline networks: Is there room for energy saving optimization? *IEEE Communications Magazine*, 2011, 23(6): 234-243
- [7] Erramilli A, Roughan M, Veitch D, et al. Self-similar traffic and network dynamics. *Proceedings of the IEEE*, 2002, 90(5): 800-819
- [8] Dong Meng-Li, Yang Geng, Gao Xiao-Mei. Methods of network traffic prediction. *Computer Engineering*, 2011, 37(16): 98-100(in Chinese)
(董梦丽, 杨庚, 曹晓梅. 网络流量预测方法. *计算机工程*, 2011, 37(16): 98-100)
- [9] Shu J, Varaiya P. Smart pay access control via incentive alignment. *IEEE Journal on Selected Areas in Communications*, 2006, 24(5): 1051-1060
- [10] Wang Yang, Lin Chuang, Li Quan-Lin, et al. Non-cooperative game based research on routing schemes for wireless networks. *Chinese Journal of Computers*, 2009, 32(1): 54-68 (in Chinese)
(汪洋, 林闯, 李泉林等. 基于非合作博弈的无线网络路由机制研究. *计算机学报*, 2009, 32(1): 54-68)
- [11] Vickrey W. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 1961, 16(1): 8-37
- [12] Dimakis A, Jain R, Walrand J. Mechanisms for efficient allocation in divisible capacity networks//*Proceedings of the 45th IEEE Conference on Decision and Control*. San Diego, USA, 2006: 1264-1269
- [13] Yamanaka N, Shimizu S, Shan G. Energy efficient network design tool for green IP/Ethernet networks//*Proceedings of the 14th IEEE Conference on Optical Network Design and Modeling (ONDM)*. Kyoto, Japan, 2010: 1-5
- [14] Yonezu H, Kikuta K, Ishii D, et al. QoS aware energy optimal network topology design and dynamic link power management//*Proceedings of the 36th IEEE European Conference and Exhibition on Optical Communication (ECOC)*. Torino, Italy, 2010: 1-3
- [15] Chiaraviglio L, Mellia M, Neri F. Reducing power consumption in backbone networks//*Proceedings of the IEEE International Conference on Communications(ICC'09)*. Dresden, Germany, 2009: 1-6
- [16] Cianfrani A, Eramo V, Listanti M, et al. An OSPF enhancement for energy saving in IP networks//*Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Shanghai, China, 2011: 325-330
- [17] Coiro A, Iervini F, Listanti M. Distributed and adaptive interface switch off for Internet energy saving//*Proceedings of the 20th International Conference on Computer Communications and Networks(ICCCN)*. Hawaii, USA, 2011: 1-8
- [18] Cuomo F, Abbagnale A, Cianfrani A, et al. Keeping the connectivity and saving the energy in the Internet//*Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Shanghai, China, 2011: 319-324
- [19] Lee S S W, Tseng P K, Chen A. Link weight assignment and loop-free routing table update for link state routing protocols in energy-aware Internet. *Future Generation Computer Systems*, 2012, 28(2): 437-445
- [20] Cuomo F, Cianfrani A, Polverini M, et al. Network pruning for energy saving in the Internet. *Computer Networks*, 2012, 56(2012): 2355-2367
- [21] Chen Xiao-Tian, Liu Jing-Xian. Network traffic prediction based on wavelet transformation and FARIMA. *Journal on Communications*, 2011, 32(4): 153-157, 165(in Chinese)
(陈晓天, 刘静娴. 改进的基于小波变换和 FARIMA 模型的网络流量预测算法. *通信学报*, 2011, 32(4): 153-157, 165)
- [22] Gao Bo, Zhang Qin-Yu, Liang Yong-Sheng, et al. Predicting self-similar networking traffic based on EMD and ARMA. *Journal on Communications*, 2011, 32(4): 47-56(in Chinese)
(高波, 张钦宇, 梁永生等. 基于 EMD 及 ARMA 的自相似网络流量预测. *通信学报*, 2011, 32(4): 47-56)
- [23] Xiao Fu, Zhao Shuai-Shuai, Wang Shao-Hui, et al. Traffic prediction method used in distributed network based on intelligent optimization. *Journal of Beijing University of Posts and Telecommunications*, 2015, 38(Sup.): 45-48(in Chinese)
(肖甫, 赵帅帅, 王少辉等. 基于智能优化的分布式网络流量预测方法. *北京邮电大学学报*, 2015, 38(增刊): 45-48)
- [24] Tian Zhong-Da, Li Shu-Jiang, Wang Yan-Hong, Gao Xian-Wen. Network traffic prediction based on empirical mode decomposition and time series analysis. *Control and Decision*, 2015, 30(5): 905-910(in Chinese)
(田中大, 李树江, 王艳红, 高宪文. 经验模式分解与时间序列分析在网络流量预测中的应用. *控制与决策*, 2015, 30(5): 905-910)
- [25] Yao Qi-Fu, Li Cui-Feng, Ma Hua-Lin, Zhang Sen. Novel network traffic forecasting algorithm based on grey model and Markov chain. *Journal of Zhejiang University (Science Edition)*, 2007, 34(4): 396-400(in Chinese)
(姚奇富, 李翠凤, 马华林, 张森. 灰色系统理论和马尔柯夫链相结合的网络流量预测方法. *浙江大学学报(理学版)*, 2007, 34(4): 396-400)
- [26] Liu Jue-Fu, Zhu Bing-Hu, Wang Hua-Feng. Dynamic spectrum allocation algorithm based on VCG mechanism. *Computer Engineering and Design*, 2015, 36(6): 1415-1419, 1424(in Chinese)

- (刘觉夫, 朱丙虎, 王华锋. 基于 VCG 机制的动态频谱分配算法. 计算机工程与设计, 2015, 36(6): 1415-1419, 1424)
- [27] Wang Shu-Ning, Yu Xing-Guo. Improved routing strategy based on VCG mechanism in mobile Ad hoc network. *Computer Applications and Software*, 2008, 25(4): 172-173, 196(in Chinese)
- (王曙宁, 于行国. 基于 VCG 机制的移动 Ad hoc 网络路由改进策略. 计算机应用与软件, 2008, 25(4): 172-173, 196)
- [28] Wei Zhi-Qiang, Zhou Wei, Ren Xiang-Jun, et al. A strategy-proof trust based decision mechanism for pervasive computing environments. *Chinese Journal of Computers*, 2012, 35(5): 871-882 (in Chinese)
- (魏志强, 周炜, 任相军等 普适计算环境中防护策略的信任决策机制研究. 计算机学报, 2012, 35(5): 871-882)
- [29] Xia You-Hua, Xu Li, Lin Hui, Zhou Zhao-Bin. A VCG mechanism based secure routing protocol for Internet of Things. *Journal of Fujian Normal University (Natural Science Edition)*, 2015, 31(5): 21-28(in Chinese)
- (夏有华, 许力, 林晖, 周赵斌. 一种基于 VCG 机制的物联网安全路由协议. 福建师范大学学报(自然科学版), 2015, 31(5): 21-28)
- [30] Liu Zhi-Xin, Shen Yan-Yan, Guan Xin-Ping. A VCG-auction based distributed mechanism for network resource allocation. *Acta Electronica Sinica*, 2010, 38(8): 1929-1934(in Chinese)
- (刘志新, 申妍妍, 关新平. 一种基于 VCG 拍卖的分布式网络资源分配机制. 电子学报, 2010, 38(8): 1929-1934)
- [31] Vereecken W, Van H, Puype B, et al. Optical networks: How much power do they consume and how can we optimize this?//Proceedings of the 36th European Conference and Exhibition on Optical Communication. Torino, Italy, 2010: 1-4
- [32] Orłowski S, Wessály R, Pióro M, et al. SNDlib 1.0-Survivable Network Design library. *Networks*, 2010, 55(3): 276-286
- [33] Bianzino A P, Chiaraviglio L, Mellia M, et al. GRiDA: GRreen distributed algorithm for energy-efficient IP backbone networks. *Computer Networks*, 2012, 56(14): 3219-3232



ZHANG Jin-Hong, born in 1982, Ph. D. candidate. His research interests include energy-saving routing and topology management, etc.

WANG Xing-Wei, born in 1968, Ph. D., professor, Ph. D. supervisor. His research interests include future Internet, cloud computing and cyberspace security, etc.

HUANG Min, born in 1968, Ph. D., professor, Ph. D. supervisor. Her research interests include modeling and optimization for logistics and supply chain management, etc.

MA Lian-Bo, born in 1981, Ph. D., associate professor. His research interests include intelligence computing and machine learning, etc.

Background

Due to the explosive increment of Internet scale and the excessive underutilization of network resources, the power consumption issue of the Internet has become one of the hottest research topics extensively concerned by ICT (Information and Communication Technology) academy and industry in recent years.

To solve the above problem, the concept of green Internet was proposed, which aimed to achieve energy saving in the Internet. The Internet can be made green by various approaches, for example, by effectively and efficiently managing network topology to put appropriate nodes or links into sleeping. In this paper, a Distributed Topology Management mechanism for Energy saving (DTME) network-wide is proposed. It consists of information sensing, traffic prediction, distributed topology decision-making and sleeping control. Among them, information sensing is used to regularly collect information on node and link status; traffic prediction is based on historical

traffic data to predict future traffic and then provides basis for topology decision; distributed topology decision-making is used to determine which nodes and links are put into sleeping or waking-up; sleeping control is used to control node and link working status. By their collaboration, distributed topology management and energy saving are achieved in green Internet. Simulation results have shown that the proposed DTME is effective and efficient.

This research is supported by the National Natural Science Foundation of China (No. 61572123), the National Science Foundation for Distinguished Young Scholars of China (No. 71325002 and No. 61225012), and the Liaoning Bai Qian Wan Talents Program (No. 2013921068), the Cross-disciplinary Collaborative Teams Program for Science, Technology and Innovation of Chinese Academy of Sciences on Network and System Technologies for Security Monitoring and Information Interaction in Smart Grid.