软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2017,28(7):1815−1834 [doi: 10.13328/j.cnki.jos.005111]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

一种大规模 IP 网络多链路拥塞推理算法*

陈 宇[1], 温欣玲[1], 段哲民[2], 李宇翀[3]

[1](郑州航空工业管理学院 电子通信工程学院,河南 郑州 450015)

[2](西北工业大学 电子信息学院,陕西 西安 710072)

[3](中国人民解放军 32147 部队,河南 商丘 476000)

通讯作者: 陈宇, E-mail: chenyu3440@gmail.com

摘 要: 基于最小集覆盖理论的拥塞链路推理算法,仅对共享瓶颈链路进行推理,当拥塞路径存在多条链路拥塞时,算法的推理性能急剧下降.针对该问题,提出一种基于贝叶斯最大后验(Bayesian maximum a-posterior,简称 BMAP)改进的拉格朗日松弛次梯度推理算法(Lagrange relaxation sub-gradient algorithm based on BMAP,简称 LRSBMAP).针对推理算法中链路覆盖范围对算法推理性能的影响,以及探针部署及额外 E2E 路径探测发包的开销问题,提出设置度阈值(degree threshold value,简称 DTV)参数预选待测 IP 网络收发包路由器节点,通过引入优选系数 ρ,在保证链路覆盖范围的基础上,兼顾开销问题,确保算法的推理性能.针对大规模 IP 网络多链路拥塞场景下,链路先验概率求解方程组系数矩阵的稀疏性,提出一种对称逐次超松弛(symmetry successive over-relaxation,简称 SSOR)分裂预处理共轭梯度法(preconditioned conjugate gradient method based on SSOR,简称 PCG_SSOR)求解链路先验概率近似唯一解的方法,防止算法求解失败.实验验证了所提算法的准确性及鲁棒性.

关键词: 拥塞链路推理;tomography;贝叶斯网模型;拉格朗日松弛;贝叶斯最大后验(BMAP)准则

中图法分类号: TP393

中文引用格式: 陈宇,温欣玲,段哲民,李宇翀.一种大规模 IP 网络多链路拥塞推理算法.软件学报,2017,28(7):1815−1834. http://www.jos.org.cn/1000-9825/5111.htm

英文引用格式: Chen Y, Wen XL, Duan ZM, Li YC. Algorithm for large scale IP network multiple link congestion inference. Ruan Jian Xue Bao/Journal of Software, 2017,28(7):1815−1834 (in Chinese). http://www.jos.org.cn/1000-9825/5111.htm

Algorithm for Large Scale IP Network Multiple Link Congestion Inference

CHEN Yu[1], WEN Xin-Ling[1], DUAN Zhe-Min[2], LI Yu-Chong[3]

[1](School of Electronics and Communication Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450015, China)

[2](Institute of Electronic Information, Northwestern Polytechnical University, Xi'an 710072, China)

[3](Troop 32147, Chinese People's Liberation Army, Shangqiu 476000, China)

Abstract: Congested link inference algorithms only infer the set of share links based on methods of smallest set coverage. When some congested path contains more than one congested link, the inference performance is obviously descending. Aiming at this problem, a version of Lagrange relaxation sub-gradient algorithm based on Bayesian maximum a-posterior (LRSBMAP) is proposed. Aiming at the impacts of congested link inference performance in the different link coverage, and the cost problems of probe deployments and additional E2E active detection, the paper proposes a preliminary selection method for transceiver nodes by optimally selecting degree threshold value (DTV) parameter of IP networks. Through introducing the optimization coefficient ρ, problems of cost and link coverage can be both considered to ensure the performance of inference algorithm. In addition, according to the sparsity of coefficient matrix in link prior

* 基金项目: 国家重点基础研究发展计划(973)(2012CB315901, 2013CB329104); 河南省高等学校重点科研项目(18A510019)

Foundation item: National Basic Research Program of China (973) (2012CB315901, 2013CB329104); Colleges and Universities Key Research Project of He'nan Province (18A510019)

收稿时间: 2016-02-23; 修改时间: 2016-04-01; 采用时间: 2016-06-02; jos 在线出版时间: 2016-10-11

CNKI 网络优先出版: 2016-10-12 16:26:53, http://www.cnki.net/kcms/detail/11.2560.TP.20161012.1626.020.html

probability solution equations, a preconditioned conjugate gradient method based on symmetry successive over-relaxation (PCG_SSOR) is proposed to obtain approximate unique solutions, helping to avoid the solution failures in large scale IP networks under the scenarios of multiple link congestion. Experiments demonstrate that the algorithms proposed in this paper have higher accuracy and robustness.

Key words: congestion link inference; tomography; Bayesian network model; Lagrange relaxation; BMAP criterion

随着 IP 网络规模的不断扩大,网络终端接入数量剧增,路由器/交换机参与数目越来越多,除了物理链路切断造成的网络阻塞外,复杂的网络结构和不合理的路由原则均会造成网络多链路拥塞现象的发生.从而带来网络整体性能和服务质量的急剧下降.多链路拥塞造成的 IP 网络高网络时延和高网络丢包率等现象还可能是因为涉及违反了 SLA(service-level agreement)等相关服务等级协定造成的[1].所以,网络管理者需要及时、准确地定位网络中发生拥塞的链路并进行处理.

目前,国内外对 IP 网络内部链路性能推理主要借助主动检测及被动检测两种方法.其中,借助少量端到端(end-to-end,简称 E2E)路径探测(tomography)推理链路性能的主动检测[2−5]方法因具有不涉及用户隐私、实时性好、设备开销小,且对网络性能影响不大等优点,而被网络运营商及国内外研究学者所青睐.目前,借助 Boolean 代数,基于最小链路覆盖集(smallest consistent failure set,简称 SCFS)理论[3,6]进行拥塞链路推理的主动检测方法 Boolean tomography,当待测 IP 网络拥塞链路比例增大时,特别是当某条拥塞路径中除共享瓶颈链路外,仍存在其他拥塞链路时,因算法理论本身存在的缺陷,将会造成算法推理性能的急剧下降.另外,有部分文献对主动检测方法中的探针部署点[7,8]及 E2E 发包路径优化[9,10]等方面进行了研究,在尽量减小开销的基础上,尽可能多地覆盖待测 IP 网络中的链路范围.但并未研究算法链路覆盖范围变化对算法推理性能带来的影响.此外,Boolean tomography 利用链路拥塞先验概率,借助贝叶斯理论进行拥塞链路推理的 CLINK 算法,能够有效地避免单时隙 E2E 路径探测对时间强关联性的依赖.但是,对大规模 IP 网络多链路拥塞场景,由于链路先验概率求解系统方程组系数矩阵的稀疏性,易造成求解失败,目前未见文献提出好的解决方法.

针对上述问题,本文基于 Boolean tomography 的实用性,针对大规模 IP 网络多链路拥塞场景,提出一种基于贝叶斯最大后验概率(Bayesian maximum a-posterior,简称 BMAP)改进的拉格朗日松弛次梯度算法(Lagrange relaxation sub-gradient algorithm based on BMAP,简称 LRSBMAP).考虑到算法链路覆盖范围对推理性能的影响.首先,根据待测 IP 网络各路由器度值,提出一种通过设置度阈值参数(degree threshold value,简称 DTV)进行 IP 网络链路覆盖的方法(link cover method based on DTV,简称 LCDTV). LCDTV 引入优选系数ρ进行待测 IP 网络各 E2E 发包路径及发包路由器探针部署节点的优选.考虑到网络用户及管理者对 IP 网络拥塞的容忍程度,在链路先验概率学习的过程中,引入了路径拥塞次数参数(path congestion time,简称 PCT),对 N 次 E2E 性能探测过程中拥塞次数小于 PCT 的 E2E 路径视为正常路径,通过去除正常路径及途经链路,在链路拥塞先验概率学习过程中,对待测 IP 网络进行拥塞选路矩阵及拥塞贝叶斯网模型的构建.针对大规模 IP 网络多链路拥塞场景下,先验概率求解系统线性方程组系数矩阵的稀疏性,提出一种收敛速度快且运行时间短的基于对称逐次超松弛(symmetry successive over-relaxation,简称 SSOR)分裂预处理的共轭梯度法(preconditioned conjugate gradient method based on SSOR,简称 PCG_SSOR),迭代求解链路拥塞先验概率近似唯一解.在拥塞链路推理过程中,首先,对当前推理时刻正常路径及途经链路从链路拥塞先验概率学习过程中构建的拥塞贝叶斯网模型中去除,并由此得到拥塞链路推理过程中的剩余拥塞选路矩阵.最后,基于 BMAP 准则,将 IP 网络拥塞链路推理问题转化为集合覆盖问题(set cover problem,简称 SCP),利用本文提出的 LRSBMAP 算法在多项式时间内完成此 SCP 的迭代求解.

通过模拟不同拓扑类型及规模的 IP 网络模型,利用检测率(detection rate,简称 DR)及误报率(false positive rate,简称 FPR)对本文提出算法与 Boolean tomography 经典算法 CLINK 进行性能比较评价,验证本文提出算法的准确性及鲁棒性.另外,本文提出利用 LCDTV 方法改变推理算法的链路覆盖范围,实验验证了该方法的有效性.在 LCDTV 算法中,通过引入优选系数ρ,在兼顾开销的基础上,有效地保证了拥塞链路推理算法的推理性能.在大规模 IP 网络多链路拥塞场景下,利用本文提出的 PCG_SSOR 算法与传统的求解方法分别进行链路先验概率迭代求解,验证了本文提出的 PCG_SSOR 方法的收敛性及稳定性.

本文的创新点主要体现在以下 4 个方面.

(1) 根据待测 IP 网络中各路由器度值,引入度阈值参数 DTV,提出一种链路覆盖方法 LCDTV 算法;算法根据优选系数ρ确定最优 DTV,兼顾了链路覆盖率、额外 E2E 路径探测发包数及探针部署开销;

(2) 根据路径拥塞容忍程度引入路径拥塞次数参数 PCT,根据拥塞路径及途经链路关系,借助矩阵性质,构建链路拥塞先验概率学习过程及拥塞链路推理过程中的化简拥塞选路矩阵及剩余拥塞贝叶斯网模型,简化了推理复杂度;

(3) 针对大规模 IP 网络多链路拥塞场景,先验概率求解系统线性方程组中系数矩阵的稀疏性问题,提出 PCG_SSOR 方法迭代求解各链路先验概率近似唯一解;

(4) 根据拥塞链路推理时刻各 E2E 路径性能测量结果,基于推理时刻剩余拥塞选路矩阵及 BMAP 准则,提出一种多链路拥塞场景下的拥塞链路集合推理算法 LRSBMAP.

本文第 1 节综述相关工作.第 2 节简述本文算法流程.第 3 节介绍 LCDTV 算法.第 4 节介绍拥塞贝叶斯网模型及拥塞选路矩阵的构建方法.第 5 节介绍 PCG_SSOR 算法.第 6 节介绍 LRSBMAP 算法.第 7 节对本文所提算法进行实验验证评价.第 8 节总结全文并对未来加以展望.

1 相关工作

目前,对 IP 网络内部拥塞链路状态推理主要借助主动检测及被动检测两种方法.其中,被动检测是基于 SNMP(simple network management protocol)协议,通过采集 IP 网络中各网络设备(如:路由器,交换机等)上的数据信息推理网络内部各链路连接性能.理论上,被动检测方法对 IP 网络不产生附加流量,不增加网络负担.但是,此方法适合单个设备的监测,对待测 IP 网络规模较大,特别是需要进行大范围的拥塞链路推理、分析时,需采集 IP 网络中各路由器的日志数据,数据量巨大,实时性较差,且存在用户数据泄漏等安全性问题.而主动检测方法通过在网络各发包路由器上部署探针,基于 ICMP(Internet control messages protocol),利用测量工具(如:Ping)向目的端路由器发送 TCP(UDP)探针包,根据 E2E 路径丢包率,借助网络拓扑结构等,推理 IP 网络内部链路性能,不涉及用户隐私.

早期的主动检测方法利用多播探测发包方式进行多路 E2E 性能检测(snapshots)[11,12].通过单时隙 E2E 路径性能探测,推理 IP 网络内部各链路丢包率,进而推理链路性能,但因支持多播方式的路由器在大规模 IP 网络中并未普及而使多播方法受到限制[13].利用单播背靠背[5]与单播模拟多播的包群测量方式[14]以及用于测量共享路径长度的三明治报文[15],虽然弥补了多播探测的缺陷,但对时间关联性要求较高,推理精度较难保证,需要投入复杂的基础设施建设及较高的管理费用[4,5,10].并且,在对各链路丢包率求解时,涉及复杂的线性方程组矩阵求逆运算,对大规模 IP 网络,特别是多链路拥塞场景,容易陷入维数灾难,导致算法求解失败.因此,为了避免单时隙性能探测对时间强相关性的要求,Nguyen 和 Thiran 等人提出多时隙 E2E 路径性能探测方法,将路径及链路状态通过 Boolean 代数模型进行表示[16],并借助 Bayes 定理,通过 N 次 E2E 路径性能测量结果,学习 IP 网络链路拥塞先验概率,基于 BMAP 准则推理当前推理时刻拥塞链路[4],且较早期推理方法假设 IP 网络内部各链路具有一致先验概率的 SCFS 算法[3]或不借助先验概率的 MCMC(Monte Carlo Markov chain)算法[2],在推理性能上有了较大程度的提高.此外,Ghita 等人将 CLINK 方法[4]推广到更一般的场景中,发现并证明,当网络中某些链路的状态不相互独立时,链路先验拥塞概率可辨识的充要条件[17],提出一种只需少量 E2E 路径测量即可求得链路先验拥塞概率的方案[18].

除了对 E2E 测量方式的研究外,还有一些关于如何减少 E2E 路径探测报文数目[9,10]、减少部署测量监控节点数目[7,8]等方面的讨论.但有关推理算法链路覆盖范围的改变对推理算法性能的影响未进行研究.另外,为解决网络 tomography 方法本身 E2E 路径数过少造成的系统方程组系数矩阵欠定问题,Augustin[19]、潘胜利[1]等人提出单源多径路由等方法.此外,当两条路径经历公共拥塞链路时,路径性能将会具有很强的相关性.Kim 等人[20]基于时延相关性,提出小波降噪方法以识别共享拥塞链路.Zarifzadeh 等人[6]借助 Boolean tomography 的实用性,通过 Analog tomography 方法提高了链路性能分辨率.但当拥塞路径途经拥塞链路数过多时,现有的基于

SCFS 理论的拥塞链路集合推理算法[3,4,6]的性能下降得较为明显.

2 LRSBMAP 算法

本文提出一种大规模 IP 网络多链路拥塞推理算法 LRSBMAP,主要包括 3 部分:(1) E2E 路径及探针部署优选.在保证链路覆盖率的基础上,根据待测 IP 网络拓扑,进行 E2E 发包探测路径及发包路由器探针部署位置的优选;(2) 链路拥塞先验概率学习.根据 N 次 E2E 路径性能测量结果,学习算法覆盖的各链路拥塞先验概率;(3) 当前时刻拥塞链路推理.根据当前推理时刻各 E2E 路径的拥塞情况,基于 BMAP 准则,推理当前 IP 网络最有可能发生拥塞的链路集合.算法原理框图如图 1 所示.

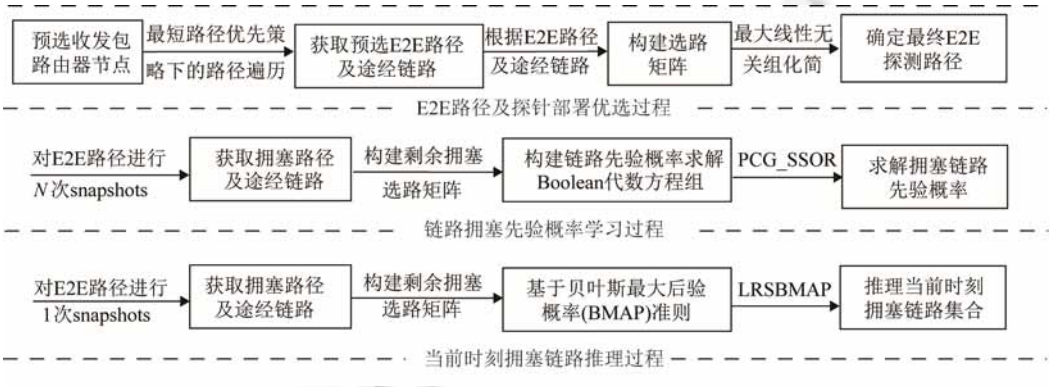


Fig.1 Principle frame of LRSBMAP algorithm

图 1 LRSBMAP 算法原理框图

3 E2E 路径及探针部署优选方法 LCDTV

3.1 预选收发包路由器节点

借助 tomography 方法进行 IP 网络拥塞链路推理时,对尽可能少的 E2E 路径利用 Ping 发包探测,对尽可能多的途经链路进行性能推理.目前,IP 网络拥塞链路推理算法通常选取叶子节点(端主机)作为收发包路由器端点[2−4],如选取叶子节点向其他各叶子节点进行 E2E 路径发包探测,虽然 E2E 路径数较少,但因路由算法最短路径优先原则,遍历的链路数不高,无法进行整个待测 IP 网络中各链路的性能推理.另外,如果 E2E 路径较长,途经的链路数较多,对推理模型构建的系统线性方程组因系数矩阵欠定[4],各链路性能无法精确求解,导致算法失败.虽然 Nguyen 等人提出利用两条路径相关补满秩的方法扩展矩阵行来解决此问题,但是,当两条路径相关无法对欠定矩阵补满秩时,使用更多路径相关补满秩通常也会造成扩展路径行线性相关,导致系数矩阵欠定.

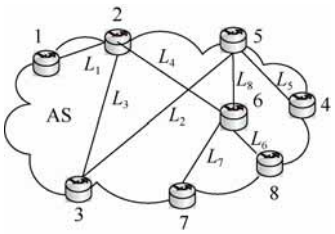


Fig.2 Some IP network topology simulation model

图 2 某 IP 网络拓扑仿真模型

为便于算法说明,模拟一个小型 IP 网络自制系统(autonomous system,简称 AS),如图 2 所示.其中,路由器节点有 8 个,分别用数字 1~8 进行表示;链路有 8 条,叶子节点(路由器度值=1)分别为 1,4,7,8.如仅对节点 1 进行发包探针部署,向其他叶子节点发送探针包,因路由算法中最短路径优先策略,路径数仅为 3 条.另外,由于在实际 IP 网络中,通过 Traceroute 获取到的路径途经链路信息为各路由器网卡的 IP 地址,各网卡 IP 地址可通过 sr-all 工具对应到所属的同一个路由器,为了便于程序编写,将各路由器间连接链路利用路由器节点标号进行表示,如:1|2 代表路由器 1 与路由器 2 间的连接链路.传统

的链路覆盖方法以叶子节点作为收发包路由器节点,如:链路 2|6 较 2|3 优选,3 条 E2E 路径分别为 P_1:1|2,2|6,6|5,5|4,P_2:1|2,2|6,6|7,P_3:1|2,2|6,6|8,覆盖的链路数为 6 条,链路 2|3 及 3|5 未能被 E2E 路径所覆盖.

在根据 E2E 路径探测性能结果进行途经链路性能推理时,需借助 E2E 路径及途经链路构建链路先验概率求解线性方程组,若路径数过少,链路数远远超出路径数,则将导致方程组无唯一解.特别是对大规模、不同拓扑类型的IP网络,拥塞链路推理时,如部分E2E路径性能探测正常,正常路径及途经各链路将不被作为可能拥塞的链路集合参与推理,而每条路径至少不小于 1 条链路.因此,去除 1 条路径将去除多条链路,路径数减小幅度通常超过链路数.因此,本文根据待测 IP 网络中各路由器节点度值,引入 DTV 参数,由最优 DTV 参数决定 IP 网络预选收发包节点.具体方法如下:对 IP 网络中路由器度值小于等于 DTV 的路由器,作为获取 E2E 路径中的预选收发包路由器节点.如图 2 所示的 IP 网络,如设置 DTV=1,即以度值小于等于 1(叶子节点)的路由器作为预选收发包路由器节点.路由器节点 1,3,4,7,8 将作为预选收发包路由器节点.

3.2 预选E2E路径

预选收发包路由器节点选定后,基于最短路径优先策略,利用贪心搜索算法[21]中最小生成树 Prim 理论,以路径途经链路数最少作为权值对所有预选收发包路由器节点间所有可能到达的 E2E 路径进行轮询遍历,获取待测 IP 网络中各 E2E 路径及途经链路.搜索出途经链路数最少的 E2E 路径.起始路由器节点相同(途经链路方向相反)的路径保留其中一条.由此,得到预选 E2E 路径.图 3 所示为图 2 预选 E2E 路径获取示意图.

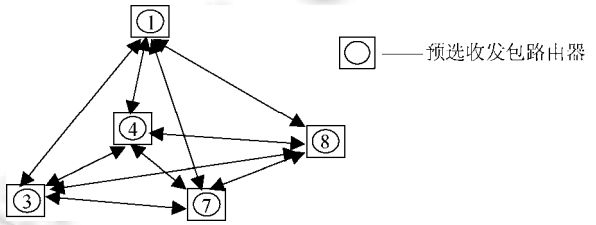


Fig.3 Preliminary selection method of E2E paths in IP network of Fig.2

图 3 图 2 所示 IP 网络 E2E 路径预选方法

在图 2 所示的 IP 网络中,选择路由器节点 1,3,4,7,8 作为预选收发包路由器节点进行 E2E 路径轮询遍历,可获得 5 条最短 E2E 路径.P_1:1/2,2/3,P_2:1/2,2/6,6/8,P_3:1/2,2/6,6/7,P_4:1/2,2/6,6/5,5/4,P_5:3/5,5/4.覆盖待测 IP 网络中全部 8 条链路.

3.3 选路矩阵构建

各预选 E2E 路径 P_1~P_5 及途经各链路按照跳数(hop)级别关系依次排列,见表 1.

Table 1 Preliminary selection of E2E paths and traversing links in IP network of Fig.2

表 1 图 2 所示 IP 网络各预选 E2E 路径及途经链路

路径 $P_i(\rightarrow d)$	IP 地址对(链路)			
	1 跳(Hop1)	2 跳(Hop2)	3 跳(Hop3)	4 跳(Hop4)
$P_1(1\rightarrow 3)$	1|2(L_1)	2|3(L_3)		
$P_2(1\rightarrow 8)$	1|2(L_1)	2|6(L_4)	6|8(L_6)	
$P_3(1\rightarrow 7)$	1|2(L_1)	2|6(L_4)	6|7(L_7)	
$P_4(1\rightarrow 4)$	1|2(L_1)	2|6(L_4)	6|5(L_8)	5|4(L_5)
$P_5(3\rightarrow 4)$	3|5(L_2)	5|4(L_5)		

为了便于描述,对各 IP|IP 对应的链路按照跳数顺序依次利用 L_1~L_8 进行表示.根据 IP 网络拓扑结构中各 E2E 路径及途经链路的关系,选路矩阵(依赖矩阵)D 的构建方法如下定义.

定义 1. IP 网络的选路矩阵 D 的各行为 E2E 探测路径 P_i(i=1,2,…,n_p),各列为 IP 网络中所有链路

L_j(j=1,2,…,n_c)按照 Hop 级别从小到大依次排列;当某条 E2E 探测路径 P_i 途经某条链路 L_j 时,选路矩阵对应列的位置处元素值 D_{ij}=1,否则,D_{ij}=0.

对表 1 所示的 IP 网络拓扑,根据各 E2E 探测路径源端 s 到目的端 d 的 Hop 级别关系(Hop1,Hop2,...),可依此列写出 D 中各列所代表的链路,见表 2.

Table 2 Construction method of routing matrix D in IP network of Fig.2

表 2 图 2 所示 IP 网络选路矩阵 D 构建方法

IP 地址对(链路) P_i(s→d)	1|2(L_1)	3|5(L_2)	2|3(L_3)	2|6(L_4)	5|4(L_5)	6|8(L_6)	6|7(L_7)	6|5(L_8)
P_1(1→3)	1	0	1	0	0	0	0	0
P_2(1→8)	1	0	0	1	0	1	0	0
P_3(1→7)	1	0	0	1	0	0	1	0
P_4(1→4)	1	0	0	1	1	0	0	1
P_5(3→4)	0	1	0	0	1	0	0	0

根据定义 1,由表 2 可构建出图 2 所示 IP 网络的选路矩阵 $D(n_p \times n_c)$. 如式(1)所示.

$$D=\begin{array}{c}\begin{matrix}L_1 & L_2 & L_3 & L_4 & L_5 & L_6 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}1&0&1&0&0&0&0&0\\1&0&0&1&0&1&0&0\\1&0&0&1&0&0&1&0\\1&0&0&1&1&0&0&1\\0&1&0&0&1&0&0&0\end{pmatrix}\begin{matrix}P_1\\P_2\\P_3\\P_4\\P_5\end{matrix}\end{array} \tag{1}$$

其中, n_p 代表待测 IP 网络预选 E2E 路径总数, n_c 为各预选 E2E 路径途经的链路总数.

3.4 E2E探测发包路径及探针部署节点获取

对选路矩阵 D 进行最大线性无关组化简可得矩阵 D',D'中各行即为实际需要发包探测的 E2E 路径;根据矩阵性质,去除线性相关路径行并不影响矩阵列数,即覆盖链路数不发生改变,去相关化简后的矩阵 D'中的各矩阵行,即需进行 Ping 发包探测的 E2E 路径,路径途经的各链路即为算法覆盖链路;E2E 路径源端即为需要进行探针部署的路由器节点位置.由于式(1)构建的图 2 所示的 IP 网络选路矩阵 D 本身已为最大线性无关矩阵,故 D'=D. 如式(2)所示.

$$D'=\begin{array}{c}\begin{matrix}L_1 & L_2 & L_3 & L_4 & L_5 & L_6 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}1&0&1&0&0&0&0&0\\1&0&0&1&0&1&0&0\\1&0&0&1&0&0&1&0\\1&0&0&1&1&0&0&1\\0&1&0&0&1&0&0&0\end{pmatrix}\begin{matrix}P_1\\P_2\\P_3\\P_4\\P_5\end{matrix}\end{array} \tag{2}$$

因此,图 2 所示的 IP 网络实际需要进行的 E2E 发包探测路径数为 5 条,需进行探针部署的发包路由器节点分别为路由器节点 1 和路由器节点 3.

3.5 探针部署优化

通过分析发现,为减少探针部署开销,在对大型 IP 网络进行拥塞链路推理时,E2E 探测路径及探针部署点选取可进一步进行优化.

(1) 如某预选收发包路由器端点与叶子节点(度值=1)直接相连,则此预选收发包路由器端点即便不作为预选收发包路由器端点,根据 E2E 路径探测策略,由于叶子节点必定仅连接了与之相连的预选收发包路由器端点,因此,路径探测时,要到达此叶子节点必将途经连接彼此预选收发包路由器端点间的链路.故将与叶子节点相连的预选收发包路由器端点去除,这并不影响 IP 网络 E2E 探测路径数及覆盖的链路数,减少了探针部署开销;

(2) 考虑到若两个预选收发包路由器端点彼此相连,根据路由算法中最短路径优先原则,E2E 路径势必以

两个预选收发包路由器直接相连的链路作为最短路径进行,如均部署探针,两个探针间仅包含一条链路,则准确率能够保证,但增加了探针部署开销.因此,对相邻预选收发包路由器端点中度值小的路由器不进行探针部署;

(3) 如叶子节点数较多,叶子节点既作为发包点也作为接收包点,则算法对 E2E 路径获取复杂度较大.因此,在通过探针部署对 IP 网络进行 E2E 路径获取及链路覆盖时,可将叶子节点仅作为接收包点,如图 4 所示.○表示叶子节点,□表示度值>1 的路由器节点.依据最短路径优先原则完成 E2E 路径获取及链路覆盖.

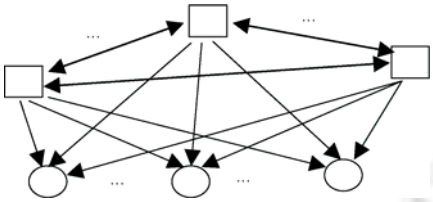


Fig.4 Schematic diagram of optimizing E2E measurement paths

图 4 E2E 探测路径获取优化方法示意图

通过实验,对大规模 IP 网络借助 LCDTV 优化算法进行 E2E 路径及链路覆盖,可大大缩短 E2E 路径及探针部署的计算时间,且链路覆盖率变化不明显.

3.6 DTV选取策略

在借助 DTV 参数进行 E2E 探测路径获取时,路由器探针部署节点的数目对最终确定的主动 E2E 探测路径数及链路覆盖率有较大影响.因此,为了兼顾探针部署开销及链路覆盖率,LCDTV 算法中引入性能优选参数ρ对探针部署及 E2E 路径获取算法中的 DTV 参数进行优选.优选系数的计算方法如式(3)所示.

$$\rho = \text{探针部署数目} \times \text{E2E探测路径数目} \times \text{未覆盖链路率} \tag{3}$$

优选系数ρ选取时,综合考虑如下 3 个因素:(1) 尽可能少地对待测ρ IP 网络下的路由器进行探针部署;(2) 尽可能地减少对 IP 网络的额外流量注入,对尽可能少的 E2E 路径发包探测;(3) 尽可能多地覆盖待测 IP 网络中的链路.本文提出一种 DTV 的选取策略:在保证链路覆盖率的情况下,以参数ρ取得最小值时对应的 DTV 值作为获取 E2E 探测路径及探针部署节点的依据.即:

$$DTV_{\text{optimal}} = \min(DTV|_{\arg\min(\rho)}) \tag{4}$$

另外,可以根据实际情况对式(3)中的 3 个参数进行优选.例如:当以覆盖最大范围的链路为目标时,应该以参数“未覆盖链路数/率”最小作为最重要的选取依据.

4 拥塞贝叶斯网模型及拥塞选路矩阵构建

贝叶斯网模型是一个有向无环图(directed acyclic graph,简称 DAG),可用式(5)表示.

$$G = (\nu, \varepsilon) \tag{5}$$

其中,ν代表节点,ε代表连接节点的有向边.在贝叶斯网中,每个节点储存一个条件概率表,当该节点为已知的证据节点时,条件概率表为节点的先验概率分布.根据图中的因果关系以及一致的条件概率和先验概率,可通过证据节点推理未知的隐藏节点状态.对 IP 网络进行贝叶斯网模型构建时,E2E 路径的状态变量集合 $\mathbf{Y} = (y_1,\ldots, y_i,\ldots,y_{n_p})$ 为贝叶斯网中的观测节点.各 E2E 路径途经链路的状态变量 $\mathbf{X} = (x_1,\ldots,x_j,\ldots,x_{n_c})$ 为隐藏节点.为了进行拥塞链路推理,需要对待测 IP 网络构建推理时刻的拥塞贝叶斯网模型.

定义 2. E2E 路径 P_i 拥塞,其状态变量 y_i=1;正常,y_i=0.同理,链路拥塞,其状态变量 x_j=1;正常,x_j=0.

对图 2 所示 IP 网络构建的贝叶斯网推理模型如图 5 所示.

其中,对待测 IP 网络进行拥塞链路推理时,由于拥塞链路所在路径必为拥塞路径,为简化推理过程,可不考虑 IP 网络中的正常路径及途经链路.

定义 3. 对待测 IP 网络构建的贝叶斯网模型中去除各 E2E 探测正常路径(观测节点)及途经链路(隐藏节点)

以及连接有向边,即得到待测 IP 网络的拥塞贝叶斯网模型.

本文关于拥塞链路的推理过程包括对两个拥塞贝叶斯网模型的构建过程.分别为 IP 网络链路拥塞先验概率学习过程中的拥塞贝叶斯网模型构建以及拥塞链路推理过程中的拥塞贝叶斯网模型的构建.

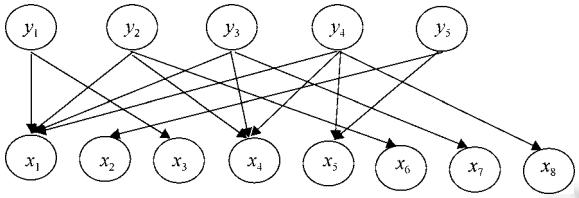


Fig.5 Bayesian network model in IP network of Fig.2

图 5 图 2 所示 IP 网络贝叶斯网推理模型

4.1 学习过程中拥塞选路矩阵构建

在链路拥塞先验概率学习过程中,借助拥塞选路矩阵作为系统线性方程组中的系数矩阵.因此,需要对学习过程中贝叶斯网模型及拥塞选路矩阵 D''进行构建.通过对待测 IP 网络各 E2E 路径进行 N 次 snapshots,当路径拥塞次数不超过设定阈值 PCT(path congestion time)时,则该路径正常,途经链路也均正常.反之,该路径拥塞.参数 PCT 的大小可根据网络用户或管理者对待测 IP 网络拥塞容忍程度进行设置,如对网络性能要求很高,则可设置 PCT=0.即 N 次 snapshot 中只要有 1 次探测中路径发生拥塞,则路径拥塞.将正常路径及途经链路从图 3 所示 IP 网络贝叶斯网模型中去除,即可得到链路拥塞先验概率学习过程中的拥塞贝叶斯网模型.

将 N 次 E2E 路径 snapshots 中正常路径及途经链路对应的矩阵行及列从线性无关化简选路矩阵 D'中移除,并再次进行线性无关化简,即可得到待测 IP 网络在链路拥塞先验概率学习过程中的拥塞选路矩阵 D''.如对图 2 所示 IP 网络进行 N=30 次 E2E 路径探测,路径 P_2 始终保持正常,则路径 P_2 及途经链路 L_1,L_4 及 L_6 对应的状态变量 x_1,x_4,x_6 及连接有向边可从图 3 所示的 IP 网络贝叶斯网模型中移除,移除后可得到先验概率求解过程中的拥塞贝叶斯网模型,如图 6 所示.

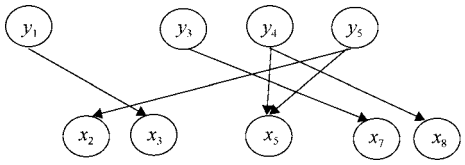


Fig.6 Rest Bayesian network model in IP network of Fig.2

图 6 图 2 所示 IP 网络拥塞贝叶斯网模型

同样,对去相关化简矩阵 D'中移除拥塞路径对应的矩阵行及列,移除后的矩阵为 D_1'.

$$D_1'=\begin{array}{c}\begin{matrix}L_1 & L_2 & L_3 & L_4 & L_5 & L_6 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}1&0&1&0&0&0&0&0\\1&0&0&1&0&1&0&0\\1&0&0&1&0&0&1&0\\1&0&0&1&1&0&0&1\\0&1&0&0&1&0&0&0\end{pmatrix}\begin{matrix}P_1\\P_2\\P_3\\P_4\\P_5\end{matrix}\end{array}\Rightarrow D_1'=\begin{array}{c}\begin{matrix}L_2 & L_3 & L_5 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}0&1&0&0&0\\0&0&0&1&0\\0&0&1&0&1\\1&0&1&0&0\end{pmatrix}\begin{matrix}P_1\\P_3\\P_4\\P_5\end{matrix}\end{array} \tag{6}$$

对矩阵 D_1'去相关化简后可得 D'',本例中 $D''=D_1'$,如式(7)所示.

$$D''=\begin{array}{c}\begin{matrix}L_2 & L_3 & L_5 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}0&1&0&0&0\\0&0&0&1&0\\0&0&1&0&1\\1&0&1&0&0\end{pmatrix}\begin{matrix}P_1\\P_3\\P_4\\P_5\end{matrix}\end{array} \tag{7}$$

从图 6 可以看出,在推理拥塞链路时,如路径 P_1 拥塞,是由链路 L_3 拥塞造成的.同样,路径 P_3 拥塞,是由链路

L_7 拥塞造成的.拥塞贝叶斯网模型的构建能够有效地减小拥塞链路推理的复杂度.

4.2 推理过程中拥塞选路矩阵构建

在拥塞链路推理过程中,需要构建剩余拥塞选路矩阵 D_d,对链路拥塞先验概率学习过程中的拥塞贝叶斯网模型中对应的拥塞路径进行 1 次 E2E 性能探测 snapshots,得到各 E2E 路径的性能结果,并从学习过程中的拥塞贝叶斯网模型中移除探测结果为正常的路径及途经链路对应的节点及有向边,即可获取当前拥塞链路推理时刻剩余拥塞贝叶斯网模型.同样,将链路拥塞先验概率学习过程中构建的拥塞选路矩阵 D''中移除推理时刻正常路径及途经链路对应的矩阵行及矩阵列,则线性无关化简后即可得到待测 IP 网络在拥塞链路推理过程中的剩余拥塞选路矩阵 D_d.如图 2 所示的 IP 网络,在推理过程中,如测得路径 P_4 为正常路径,则剩余拥塞选路矩阵 D_d 如式(8)所示.

$$D_d=\begin{array}{c}\begin{matrix}L_2 & L_3 & L_5 & L_7 & L_8\end{matrix}\\ \begin{pmatrix}0&1&0&0&0\\0&0&0&1&0\\0&0&1&0&1\\1&0&1&0&0\end{pmatrix}\end{array}\begin{matrix}P_1\\P_3\\P_4\\P_5\end{matrix}\Rightarrow D_d=\begin{array}{c}\begin{matrix}L_2 & L_3 & L_7\end{matrix}\\ \begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\end{array}\begin{matrix}P_1\\P_3\\P_5\end{matrix} \tag{8}$$

5 链路拥塞先验概率求解方法 PCG_SSOR

本文基于 Boolean tomography 拥塞链路推理框架,对当前推理时刻拥塞链路推理前,先对待测 IP 网络多链路拥塞场景下的各链路进行拥塞先验概率学习.提出利用矩阵特性,对 IP 网络构建的选路矩阵 D 进行去相关化简,减少矩阵行数,即:减少了 E2E 发包探测路径数.化简后,矩阵 D'列数不变,链路覆盖范围保持不变.根据 N 次 snapshots,移除正常路径及途经链路对应的矩阵行及列后,再次化简后得到矩阵 D''.矩阵 D''将作为链路拥塞先验概率求解线性方程组中的系数矩阵,经两次化简操作,大大降低了 Boolean 线性方程组求逆计算的复杂性.

IP 网络拥塞路径的传输率等于该路径途经各链路传输率的乘积[22].如式(9)所示.

$$\Psi_i=\prod_{j=1}^{n_c}\varphi_j^{D''_{ij}} \tag{9}$$

其中,Ψ_i 为第 i 条路径整体传输率,φ_j 为该路径下途经的第 j 条链路的传输率,$D''_{ij}|_{D_{ij}=\{0,1\}}$ 为矩阵 D''_{ij} 第 i 行第 j 列的元素值.借助二进制最大化操作“$\vee$”,可将路径拥塞与链路拥塞关系表示为式(10).

$$y_i=\bigvee_{j=1}^{n_c}x_j\cdot D''_{ij} \tag{10}$$

对式(10)两边取数学期望 E 转换后可得路径拥塞与链路拥塞关系表达式如下.

$$E[y_i]=P\left(\bigvee_{j=1}^{n_c}x_j\cdot D''_{ij}=1\right)=1-P\left(\bigvee_{j=1}^{n_c}x_j\cdot D''_{ij}=0\right)=1-\prod_{j=1}^{n_c}(1-p_j)^{D''_{ij}} \tag{11}$$

路径拥塞的期望值 $E[y_i]$ 可通过 N 次路径探测得到的每次路径拥塞变量 y_i=$\{0,1\}$求和取平均后得出,用 $\overline{y}_i$ 进行表示.为方便求解,对式(11)两边同时取对数,可得出拥塞贝叶斯网模型下拥塞链路先验概率求解Boolean代数方程组,如式(12)所示.

$$\lg(1-\overline{y}_i)=\lg(1-P_i)=\sum_{j=1}^{n_c}D''_{ij}\cdot[\lg(1-p_j)] \tag{12}$$

将各 E2E 路径拥塞概率 P_i 及拥塞选路矩阵 D''_{ij} 带入式(12),即可求得拥塞路径途经各链路的拥塞先验概率 p_j. 当式(12)的系数矩阵 D''_{ij} 非奇异时,方程组有唯一解.而由于 E2E 探测路径途经的链路数不止一条,当探测路径条数过少时,拥塞选路矩阵 D''_{ij} 可能欠定.根据拥塞选路矩阵各行为拥塞路径,各列为拥塞路径途经链路,如在拥塞链路先验概率学习时,将 N 次 E2E 探测得出的路径性能情况作为一个整体,任意两条拥塞路径可关联为一条路径,作为拥塞选路矩阵补满秩所需的扩展路径行.并通过 Boolean 二进制最大化操作构建拥塞选路矩阵关联路径行.扩展矩阵用 D''_{kj} 进行表示.将去相关后的拥塞选路矩阵 D''_{ij} 与其合并,并再次去相关后可构建系数矩阵

满秩的链路拥塞先验概率求解 Boolean 代数方程组.扩展后的拥塞选路矩阵 D'' 为秩为 n_ε 的方阵,k 值大小取决于 $n_\varepsilon - n'_\theta$ 的值, n_ε 为拥塞路径途径的所有链路数, n'_θ 为拥塞选路矩阵去相关后的路径行数.

$$D''' = \begin{pmatrix} D''_{ij} \\ D''_{kj} \end{pmatrix}_{n_\varepsilon \times n_\varepsilon} \tag{13}$$

如式(7)即为对图 2 所示 IP 网络构建的 4×5 矩阵 D',需要扩展一条路径行完成满秩矩阵 D'' 的构建,对式(7)补满秩的结果如式(14)所示.其中, D'' 的扩展路径行为路径 P_4 及 P_5 通过 Boolean 最大化操作获取.

$$D''' = \begin{array}{c} \begin{matrix} L_2 & L_3 & L_5 & L_7 & L_8 \end{matrix} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \begin{matrix} P_1 \\ P_3 \\ P_4 \\ P_5 \\ P_{45} \end{matrix} \tag{14}$$

由式(14)可知,拥塞选路矩阵 D'' 各行分别代表路径 P_1,P_3,P_4,P_5 及扩展路径 P_{45}.各列分别为链路 L_2,L_3,L_5,L_7, L_8.求解各 E2E 拥塞路径途经链路的拥塞先验概率唯一解 Boolean 代数方程组可转化为式(15)的形式.

$$\lg(1-P_i) = \sum_{j=1}^{n_\varepsilon} D''' \cdot [\lg(1-p_j)] \tag{15}$$

其中,扩展路径行 $\bar{y}'_{il}$ 可通过对两条拥塞路径 P_i 及 P_l 根据每次 snapshot 探测得出的拥塞变量 y_i 及 y_l 进行二进制最大化操作结果求和取平均后获得.为方便表示,其向量表达如式(16)所示.

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}' \end{bmatrix} = D'''[\lg(1-\mathbf{p})]^T \tag{16}$$

式(16)中,

$$\left.\begin{aligned} \mathbf{y} &= [\lg(1-\bar{y}_1), \lg(1-\bar{y}_2), ..., \lg(1-\bar{y}'_{n_\theta})]^T, \\ \mathbf{y}' &= \left[\lg(1-\bar{y}_{12}), ..., \lg(1-\bar{y}_{1n'_\theta}), \lg(1-\bar{y}_{23}), ..., \lg(1-\bar{y}_{2n'_\theta}), ..., \lg\left(1-\bar{y}_{\frac{(n'_\theta-1)n'_\theta}{2}}\right) \right]^T \\ \lg(1-\mathbf{p}) &= [\lg(1-p_1), \lg(1-p_2), ..., \lg(1-p_{n_\varepsilon})]^T \end{aligned}\right\} \tag{17}$$

但是,通过随机选取任意两条路径行关联获取扩展路径行补满秩时,有可能造成补满秩后的拥塞选路矩阵仍存在线性相关行,而再次导致式(16)系数矩阵欠定.因此,在进行方程组求解时,选择拥塞选路矩阵补满秩后为最大线性无关矩阵所对应的关联路径行进行补满秩操作,以保证链路拥塞先验概率有解且唯一.

对线性方程组求解的直接解法主要包括:高斯消元法、克拉默法则、直接三角形法、平方根法、追赶法等.但是,由于直接解法的存储量和计算量很大,特别是在本文对求解链路拥塞先验概率的线性方程组中.当 IP 网络规模较大,主元 $D_{kk}^{(k)} = 0$ (k=1,2,…, $n_{\varepsilon'}$)时,则无法计算.即使主元 $D_{kk}^{(k)} \neq 0$, 但当其绝对值 $|D_{kk}^{(k)}|$ 很小时,对拥塞选路矩阵求逆时作为除数,因舍入误差的存在也会带来较大误差.由于 IP 网络构建的拥塞选路矩阵作为线性方程组系数矩阵,为 Boolean 代数模型矩阵,且非零值较少,为典型的稀疏矩阵.因此,本文将稀疏矩阵系数线性方程组求解方法中的迭代法引入到 IP 网络链路拥塞先验概率求解中.通常,迭代法求解线性方程组主要包括:雅克比(Jacobi)迭代法、高斯-赛德尔(Gauss-Seidel)迭代法、逐次超松弛(SOR)迭代法、共轭梯度(conjugated garden,简称 CG)迭代法以及加预处理的共轭梯度法(preconditioned conjugated garden,简称 PCG)等.而对求解大型线性方程组以 CG 迭代法和 PCG 迭代法为主,其具有存储量少、计算量小的优点.由于 CG 算法的收敛速度与系数矩阵的条件数密切相关,条件数越小,收敛性越好,获取的近似解精度越高.但当系数矩阵的条件数很大时,收敛速度就会变慢.由于对称逐次超松弛(symmetric successive over relaxation,简称 SSOR)分裂中具有对称因子,可用于加速共轭梯度法.因此,本文提出采用 SSOR 分裂预处理的共轭梯度(preconditioned conjugated garden based on SSOR,简称 PCG_SSOR)迭代算法,通过引入预处理矩阵 M,使链路拥塞先验概率求解线性方程组中的特征值

分布更集中,从而降低矩阵条件数,提高共轭梯度迭代法的收敛速度和求解精度.

首先,将链路拥塞先验概率求解线性方程组中去相关,补满秩后的拥塞选路矩阵系数 D''' 分裂为 $D''' = D_{sd} - C_L - C_L^T$. 其中, $D_{sd} = \mathrm{diag}(D'_{11}, D'_{22}, ..., D'_{kk}, ..., D'_{n_\varepsilon n_\varepsilon})$, C_L 为严格的下三角矩阵,则预处理矩阵 C 为

$$C = \frac{1}{\sqrt{\omega(2-\omega)}}(D''' - \omega C_L)D'''^{-\frac{1}{2}} \tag{18}$$

其中, $\omega(0 \le \omega \le 1)$ 是参数,所得 SSOR 预处理矩阵 M 如式(19)所示.

$$M \approx CC^T = \frac{1}{\sqrt{\omega(2-\omega)}}(D_{sd} - \omega C_L)D_{sd}^{-1}(D_{sd} - \omega C_L)^T \tag{19}$$

在链路拥塞先验概率的求解过程中,对图 2 所示 IP 网络优选出的各 E2E 路径进行 30 次性能探测,如本例中:各 E2E 路径拥塞次数分别为 13,1,0,16,15,由此可得各 E2E 路径拥塞概率 P_1=0.247,P_2=0.0125,P_4=0.331,P_5=0.301,将各路径拥塞概率及式(13)所得 D'' 代入式(16),利用 PCG_SSOR 算法求出链路 L_2,L_3,L_5,L_7 及 L_8 的拥塞先验概率:p_2=0,p_3=0.4338,p_5=0.5,p_7=0.0339,p_8=0.0667.

6 拥塞链路推理算法 LRSBMAP

在对当前时刻 IP 网络拥塞链路推理时,根据 1 次 snapshots 获取各 E2E 路径性能,将正常路径及途经链路从拥塞贝叶斯网模型中去除,可得到剩余拥塞贝叶斯网模型及剩余拥塞选路矩阵.将拥塞链路集推理过程归纳为 SCP 的求解过程,由于 SCP 是 NP Hard 问题[23],本文提出一种基于贝叶斯最大后验概率(BMAP)改进的拉格朗日次梯度算法 LRSBMAP 以求解拥塞链路集合最优解.

6.1 推理模型建立

在对当前 IP 网络进行拥塞链路推理时,根据主动 E2E 发包探测获取当前时刻拥塞路径及途经链路,当前时刻 IP 网络拥塞链路推理问题可以表述为:根据当前时刻各 E2E 路径探测结果,确定可能发生拥塞的链路集合 $\mathbf{X} \subseteq \mathbf{S}$, 使其后验拥塞概率最大,即求解 argmax$P(\mathbf{X}|\mathbf{Y})$.

$$\arg\max P(\mathbf{X} \mid \mathbf{Y}) = \arg\max \frac{P(\mathbf{X}, \mathbf{Y})}{P(\mathbf{Y})} \tag{20}$$

式(20)中,由于 $P(\mathbf{Y})$ 仅与网络状态及 E2E 探测结果有关,与链路选取无关,可简化为

$$\mathrm{argmax}\, P(\mathbf{X} \mid \mathbf{Y}) = \mathrm{argmax}\, P(\mathbf{X}, \mathbf{Y}) = \mathrm{argmax}\left\{\prod_{i=1}^{n_{\theta'}} P[(y_i \mid p_a(y_i)]\prod_{j=1}^{n_{\varepsilon'}} P(x_j)\right\} \tag{21}$$

其中, $n_{\theta'}$ 为推理时刻剩余拥塞路径总数, $n_{\varepsilon'}$ 为推理时刻剩余拥塞路径途经链路总数, $p_a(y_i)$ 为拥塞贝叶斯网中 y_i 的父节点. $P(x_j)$ 为链路 x_j 先验概率.由于各 E2E 路径性能测量结果正常时,途经所有链路状态均正常;各 E2E 路径性能测量结果拥塞时,路径途经链路至少有一条发生拥塞.从而有,E2E 路径状态变量与途经各链路状态变量间存在如下概率关系:

$$P(y_i = 0 \mid p_a(y_i) = \{0, ..., 0\}) = 1, \quad P(y_i = 1 \mid \exists x_i = 1 \cap x_i \in p_a(y_i)) = 1 \tag{22}$$

对当前 E2E 路径探测结果进行全网各链路拥塞联合概率求解表达式服从贝努利概型中二项概率公式:

$$P_n(k) = C_n^k p^k (1-p)^{(n-k)} \tag{23}$$

由于当前时刻仅对待测 IP 网络各 E2E 路径进行 1 次 snapshots 性能测量,n=1.且各链路状态概率均为独立分布,由式(22)、式(23),借助第 4 节求得待测 IP 网络各链路拥塞先验概率 p_j,可得式(24).

$$\arg\max P(\mathbf{X} \mid \mathbf{Y}) = \arg\max P(\mathbf{X}) = \arg\max \prod_{j'=1}^{n_{\varepsilon'}} p_j^{x_j} \cdot (1-p_j)^{(1-x_j)} \tag{24}$$

为研究方便,对式(24)两边取对数,可得式(25).

$$\mathrm{argmax} \sum_{j=1}^{n_{\varepsilon'}} [x_j \cdot \lg p_j + (1-x_j) \cdot \lg(1-p_j)] = \mathrm{argmax} \sum_{j=1}^{n_{\varepsilon'}} \left[x_j \cdot \lg \frac{p_j}{1-p_j} + \lg(1-p_j) \right] \tag{25}$$

其中,第 2 项 $\lg(1-p_j)$ 的取值与链路状态 x_j 的取值无关,故 $\arg\max P_p(X|Y)$可由式(26)求得.

$$\arg\max P_p(\mathbf{X}\mid\mathbf{Y})=\arg\max\sum_{j=1}^{n_{\varepsilon'}}\left(x_j\cdot\lg\frac{p_j}{1-p_j}\right)=\arg\min\sum_{j=1}^{n_{\varepsilon'}}\left[x_j\cdot\left(-\lg\frac{p_j}{1-p_j}\right)\right] \tag{26}$$

可将式(26)归纳为 SCP,求解式(27).

$$\begin{cases} z_{sc}=\arg\min\sum_{j=1}^{n_{\varepsilon'}}c_jx_j, & c_j=-\lg\dfrac{p_j}{1-p_j} \\ \text{s.t.}\sum_{j=1}^{n_{\varepsilon'}}D_{d_{ij}}x_j\geq 1, & i=1,2,...,n_{\theta'} \\ x_j\in\{0,1\}, & j=1,2,...,n_{\varepsilon'} \end{cases} \tag{27}$$

6.2 LRSBMAP算法

式(27)的求解问题转化为 NP Complete 问题[21],对于大规模 IP 网络,当拥塞选路矩阵维数过大时,现有的最优化算法均无法在多项式时间内得到最优解.一种折衷的方法是利用启发式算法在可接受的计算时间内得到问题的近优解.本文基于 BMAP 改进拉格朗日松弛次梯度算法 LRSBMAP,对 IP 网络拥塞链路覆盖集进行推理.

根据拉格朗日松弛理论[24,25],式(27)可松弛为如下优化问题.

$$\begin{cases} z_{LRSC}(\mathbf{u})=\arg\min\sum_{j=1}^{n_{\varepsilon'}}d_jx_j+\sum_{i=1}^{n_{\theta'}}u_i, & j=1,2,...,n'_{\varepsilon} \\ x_j\in\{0,1\}, \end{cases} \tag{28}$$

式中,拉格朗日乘子 $d_j=c_j-\sum_{i=1}^{n_{\theta'}}u_iD_{d_{ij}}$, $u_i\geq 0, i=1,2,...,n_{\theta'}$. 对初始 SCP 优化算法中的一个有效的下界是当 $d_j\leq 0$ 时, $x_j=1$, 反之, $x_j=0$. 即:

$$Z_{LB}=\sum_{j=1}^{n_{\varepsilon'}}\min(0,d_j)+\sum_{i=1}^{n_{\theta'}}u_i \tag{29}$$

令 z_{LD}, z_{UB}, z_{LB} 分别代表算法已搜索到的下确界值、式(27)最优可行解对应的最优值、式(28)某可行解对应的最优值.本文提出 LRSBMAP 算法伪代码如下.

LRSBMAP 算法.

输入:推理时刻拥塞路径集合 P_i 及途经链路集合 L_j,拥塞选路矩阵 D_d;

L_j 中各链路 l_j 先验拥塞概率 p_j.

Step 1: 初始化 $z_{LD}=-\infty$, $z_{UB}=\infty$, $u_i=\min[c_j\mid D_{d_{ij}}=1, j=1,2,...,n'_{\varepsilon}]$, $i=1,2,...,n'_{\theta}$;

Step 2: 用当前的 $\mathbf{u}$ 求解式(28),令其最优值为 z_{LB}, 更新 $z_{LD}=\max(z_{LD},z_{LB})$;

Step 3: 构造原始SCP的一个可行解:

(1) 令 $S=[j\mid x_j=1,2,...,n'_{\varepsilon}]$;

(2) 对未覆盖的行i(即: $\sum_{j=1}^{n'_{\varepsilon}}D_{d_{ij}}x_j=0$), 将对应于 $\min[j\mid D_{d_{ij}}=1, d_j<\infty, j=1,2,...,n'_{\varepsilon}]$ 的j添加到S;

(3) 将$j\in S$按降序排列.如: $S-[j]$ 仍然是SCP的可行解, $S=S-[j]$;

(4) 更新 $z_{UB}=\min\left(z_{UB},\sum_{j\in S}c_j\right)$.

Step 4: $z_{LD}=\max z_{LRSC}(\mathbf{u})$, If $z_{LD}=z_{UB}$, 转**Step 9**;

Step 5: 计算次梯度 $g_i=1-\sum_{j=1}^{n'_{\varepsilon}}D_{d_{ij}}x_j$, $i=1,2,...,n'_{\theta}$. If $\sum_{i=1}^{n'_{\theta}}(g_i)^2=0$. 转**Step 9**;

Step 6: 计算迭代步长 $\delta=f(1.05z_{UB}-z_{LB})\Big/\sum_{i=1}^{n'_{\theta}}(g_i)^2$, 其中,初始值 $f(0)=2$, If z_{LD} 连续30次迭代中没有增加,

将f减半;

Step 7: If $f \le 0.005$, 转**Step 9**;

Step 8: 更新拉格朗日乘子 $u_i = \max[0, u_i + \delta g_i],\ i = 1,2,...,n'_\theta$, 转**Step 2**;

Step 9: **输出** $L_j = \{l_j \mid x_j = 1\}$.

7 实验验证

对 Internet 拓扑结构的研究,前人已经做了很多工作,从最初的随机图模型代表 Waxman[26]到 Barabási, Albert 等人[27,28]提出的无标度网络模型 BA,再到 Bu 等人提出的基于节点度的幂率分布特征模型 GLP[29],人们都试图去发现和解析 Internet 拓扑演化的规律.因此,为了验证推理算法的有效性及准确性,本文借助 Brite 拓扑生成器[30],分别生成 Waxman、BA 及 GLP 这 3 种不同类型、不同规模的 IP 网络拓扑.其中,Waxman 模型是基于随机图模型的代表,模型中的节点度数值随着节点数量的增加而增加,但随机图模型无法生成节点众多但节点平均度值较小的网络.因 IP 网络规模的不断扩大,当新的路由器节点加入 Internet 网络时,通常倾向于与具有高度数值的"大节点"连接.基于这两个特征构造具有度分布呈幂率特征的无标度网络模型 BA 及 GLP.3 种拓扑网络模型均体现了 Internet 特性,为了更好地验证本文提出算法在不同 Internet 环境中的拥塞链路推理性能,分别在 3 种网络拓扑模型下进行了算法比较实验验证.

通过 Eclipse 平台将拓扑模型导入完成待测网络构建,利用各推理算法进行拥塞链路推理实验验证.在进行拥塞链路推理实验时,模拟实际 IP 网络路由算法最短路径优先原则,模拟 ICMP 协议分别进行 snapshots(包括 Traceroute 及 Ping),获取 E2E 路径及途经链路以及各 E2E 路径性能测量值,本文利用随机数模型 RNM(random number model)模拟待测 IP 网络每次 snapshots 中算法覆盖链路产生的拥塞事件.

7.1 算法参数设置及评价指标

本文提出算法中各参数设置主要包括:DTV——路由器度值≤DTV 时,该路由器作为预选收发包路由器,根据ρ值自动优选 DTV;PCT——N 次 E2E 路径探测中,拥塞次数≤PCT 的路径正常.算法默认设置 N=30,PCT=0;LCR(link congestion ratio)——LRSBMAP 算法模拟实验中设置参数.即:拥塞链路占算法覆盖链路的比例,取值范围为[0,1].通过选取各链路随机数赋值由大到小根据 LCR 得到每次 snapshots 的拥塞链路.利用检测率 DR 及误报率 FPR,对本文提出的 LRSBMAP 算法推理拥塞链路的结果进行评估.为了减小随机数模型对算法推理性能的影响,每组实验中,DR 及 FPR 均为各参数设置不变情况下 10 次实验结果取平均值后所得结果.

DR 及 FPR 计算公式如式(30)所示.

$$DR = \frac{F \cap X}{F},\ FPR = \frac{X \setminus F}{X} \tag{30}$$

其中,F 为实际拥塞链路,X 为算法推理出的拥塞链路.

7.2 模拟实验流程

模拟实验流程如下文图 7 所示.

7.3 LCDTV实验结果分析

为了验证本文提出的 LCDTV 算法在 E2E 路径及探针部署获取中的有效性.利用 Brite 拓扑生成器模拟不同类型、不同规模的 IP 网络模型.传统算法中,E2E 路径的收发包路由器为叶子节点,通过 E2E 路径的发包探测,对链路进行覆盖,其方法相当于 LCDTV 算法中 DTV=1 时覆盖的 E2E 路径及链路.

(1) Waxman 模型

利用 Brite 拓扑生成器以默认参数生成 150 个节点,300 条链路的 Waxman 模型,通过选取不同 DTV 值,基于最短路径优先原则进行 E2E 路径及发包探针获取,从 E2E 路径中去除方向相反链路所在路径,不同 DTV 下,通过本文提出的优选策略获取的 E2E 路径数、覆盖链路数(覆盖率)、发包探针部署数及优选参数ρ的计算结果见表 3.

如表 3 所示,*DTV*=4 时,ρ取得最小值,此时,覆盖链路 285 条,覆盖率 95%,E2E 探测发包路径数为 276 条,需要部署发包探针 30 个.因 Waxman 模型路径较长,途经链路较多,E2E 路径数较链路数少,需要借助本文提出方法对链路拥塞先验概率求解 Boolean 线性方程组补满秩.由于默认参数生成的 Waxman 网络模型中没有度值为 1 的节点,因此传统算法以度值最小的端主机作为 E2E 路径及链路覆盖的路由器,链路覆盖率最低.根据 Waxman 模型结构特点,利用 LCDTV 算法在优选系数ρ取得最小值时,链路覆盖范围最广,且兼顾了硬件探针部署开销及发包探测对 IP 网络的额外流量注入.

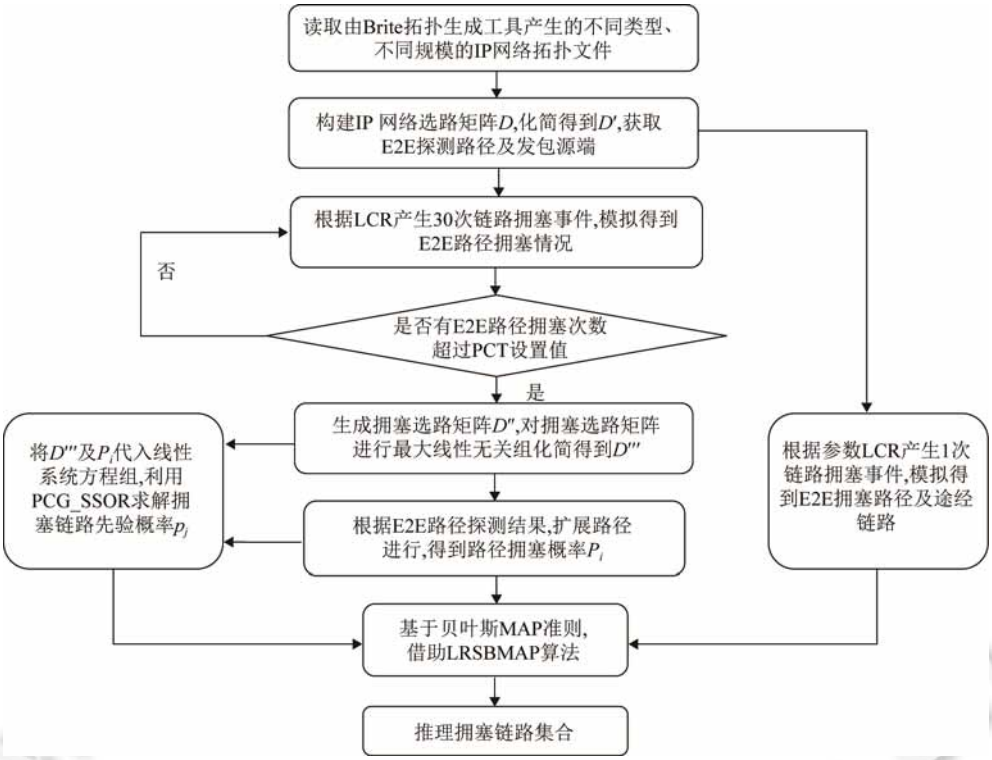


Fig.7 Flow diagram of analog experiment

图 7 模拟实验流程

Table 3 Performance comparision of LCDTV under the different DTV (Waxman model)

表 3 不同 DTV 下 LCDTV 算法性能比较(Waxman 模型)

DTV	E2E 发包探测路径数	覆盖链路数(覆盖率)	发包探针部署数	ρ (10^3)
2(文献[2−4])	179	251 (84%)	15	0.429 6
3	266	277 (92%)	26	0.553 3
4	276	285 (95%)	30	0.414 0
5	294	281 (94%)	29	0.511 6
6	272	272 (91%)	27	0.661 0
7	286	280 (93%)	32	0.640 6
8,9,10	284	275 (92%)	29	0.658 9
11	283	274 (91%)	27	0.687 7
12,13,14	253	277 (92%)	25	0.506 0

(2) GLP 模型

由于 IP 网络规模的不断扩大,当前 IP 网络拓扑结构表现出较强的幂率特性.因此,利用 Brite 拓扑生成器以默认参数生成 200 个节点的 GLP 模型(链路 354 条,叶子节点 148 个,节点度数最大值 47).利用传统算法及本文提出的 LCDTV 算法分别进行 E2E 路径及链路获取,通过实验,算法优化前后的 E2E 路径数、覆盖链路数(覆盖率)、发包探针部署数及优选参数ρ的计算结果见表 4.

Table 4 Performance comparisons of LCDTV under the different DTV (GLP model)

表 4 不同 DTV 下算法 LCDTV 优化前后性能比较(GLP 模型)

DTV	探针部署数		E2E 路径数		覆盖链路数(覆盖率)		运行时间(ms)		$\rho(10^3)$ (优化后)
	优化前	优化后	优化前	优化后	优化前	优化后	优化前	优化后	
1(文献[2−4])	1	–	147	–	188(58%)	–	17 704	–	–
2	4	4	248	250	220(67.9%)	220(67.9%)	23 548	5 833	0.321
3	7	7	378	393	247(76.2%)	245(75.6%)	26 005	7 692	0.671
4	9	8	494	387	270(83.3%)	253(78.1%)	27 319	8 047	0.678
5,6,7	8	8	442	454	260(80.2%)	260(80.2%)	20 455	8 494	0.719
8	9	8	389	416	247(76.2%)	247(76.2%)	19 768	7 224	0.792
9~47	10	10	352	352	245(75.6%)	245(75.6%)	20 461	7 339	0.859

由表 4 可知,对 200 个节点的 GLP 模型进行链路覆盖,由于强幂率模型下的 IP 网络,路径途径链路数较少,因此,传统算法仅通过叶子节点进行链路覆盖时,因路由算法中最短路径优先原则,不能有效覆盖待测 IP 网络中尽可能多的链路,链路覆盖率仅为 58%.LCDTV 算法通过改变 DTV 参数,覆盖链路有一定程度的提高.随着 IP 网络规模的激增,实验发现,利用 LCDTV 改进方法,运行时间明显缩短,但链路覆盖率变化不大.当 DTV=3 时,$\min(\rho|_{DTV=3})=0.671$, 此时,链路覆盖率为 75.6%,较传统算法提高了 17.6%.虽然 DTV=2 时,ρ值更小,但是链路覆盖率不超过 70%,为了覆盖 IP 网络中更多的链路,可选择ρ次小值作为 LCDTV 算法中最优 DTV 值选取的依据.通过对不同模型下参数ρ值计算结果的分析,当ρ取得最小或次小值时,能够兼顾链路覆盖率、探针部署及 E2E 发包路径数,验证了本文 E2E 路径及探针部署优化中参数ρ设置的有效性.

7.4 PCG_SSOR实验结果分析

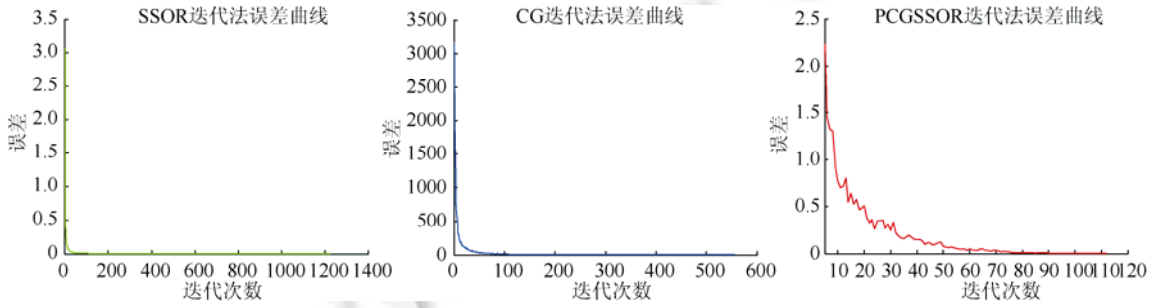
为了验证 PCG_SSOR 迭代算法在一定规模 IP 网络下的链路拥塞先验概率求解性能,模拟 500 个节点的 Waxman(1 000 条链路),BA(997 条链路)及 GLP(923 条链路)IP 网络模型.其中,CLR=0.1.首先,对 3 种不同模型基于最优度值进行探针部署,分别覆盖链路数 958,937 及 703 条.利用误差曲线对 PCG_SSOR,CG 及 SSOR 算法进行比较.误差限均设置为 10^{-3}.在不同 IP 网络模型下,模拟 30 次链路拥塞事件,利用 3 种不同迭代算法求解链路拥塞先验概率,各算法的误差曲线如图 8 所示.

表 5 为 500 个节点规模、不同类型的网络模型下,拥塞先验概率迭代求解运行时间结果.实验结果均是在 i7-5600U CPU,8G 内存,64 位 Win7 操作系统 LenovoX250 下运行所得.

Table 5 Comparisons between iterations and operation time under the different algorithms

表 5 不同算法迭代次数及运行时间比较

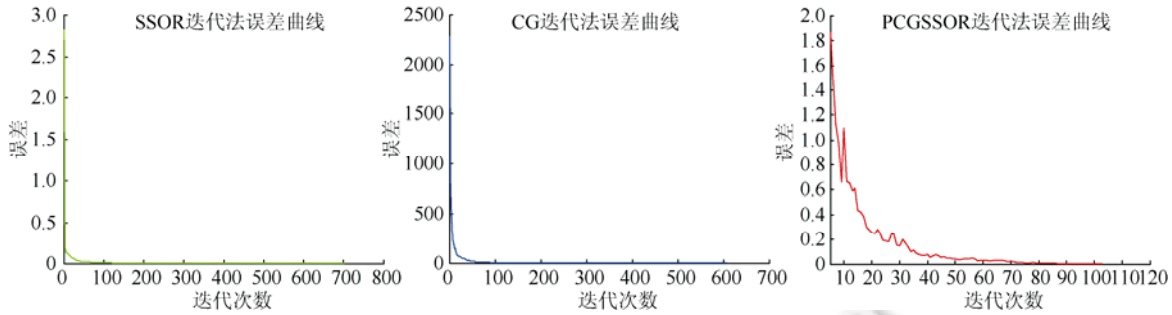
迭代算法	SSOR		CG		PCG_SSOR	
模型类型	迭代次数	运行时间(s)	迭代次数	运行时间(s)	迭代次数	运行时间(s)
Waxman	1 228	2.717 8	556	2.000 3	112	1.238 5
BA	702	1.544 9	607	2.066 4	103	1.184 7
GLP	661	0.852 9	401	0.627 9	88	0.564 8



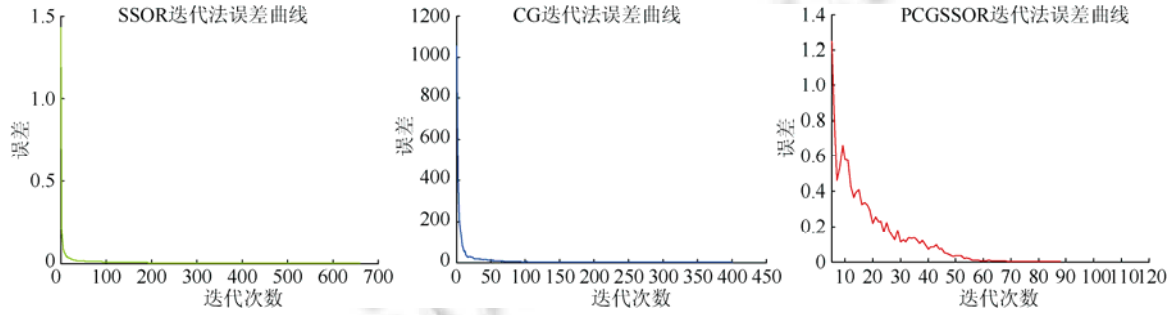
(a) Waxman 模型

Fig.8 Comparisons of iteration algorithm error curves under the different models

图 8 不同模型下各迭代算法误差曲线比较



(b) BA 模型



(c) GLP 模型

Fig.8 Comparisons of iteration algorithm error curves under the different models (Continued)

图 8 不同模型下各迭代算法误差曲线比较(续)

由表 5 可知,在相同误差限 10^{-3} 下,在 3 种不同类型的网络模型下,SSOR,CG 及 PCG_SSOR 迭代算法均能在 2 000 次迭代内实现算法收敛.PCG_SSOR 算法迭代次数最少,收敛速度最快,运行时间最短.通过实验发现,在一定规模的 IP 网络多链路拥塞场景下,传统 Gauss 消元法及 Jacobi 迭代算法不能有效地实现链路拥塞先验概率的求解.

7.5 LRSBMAP实验结果比较分析

为了验证本文提出的 LRSBMAP 算法在拥塞链路推理中的有效性及准确性.利用 Brite 拓扑生成器以默认参数模拟不同类型,规模的 IP 网络模型 Waxman、BA 及 GLP,并与 CLINK 算法进行推理性能比较.

(1) 不同 CLR 对算法推理性能的影响

对于 150 个节点规模的 IP 网络模型,CLR 从 0.05~0.6 发生变化,两种算法在最优 DTV 下 DR 及 FPR 如图 9 所示.

在不同类型的 IP 网络模型下,LRSBMAP 算法的推理性能优于 CLINK 算法.随着 CLR 的增大,DR 均呈下降趋势.两种算法的 DR 均在 GLP 模型下最高,其次是 BA 及 Waxman 模型.这与 IP 网络模型拓扑结构有直接关系,因 Waxman 为随机模型,路径较长,而 BA 及 GLP 为幂率模型,网络中部分路由器度值较大,共享链路数较 Waxman 模型中要多.因此,在 Waxman 模型下,DR 较 GLP 及 BA 模型有明显下降.当 $CLR<0.2$ 时,LRSBMAP 及 CLINK 算法在 GLP 及 BA 模型下的推理性能相差不大.但是,当 $CLR>0.2$ 时,LRSBMAP 算法在 Waxman,BA 及 GLP 模型下,其推理性能均优于 CLINK 算法,并且随着 CLR 的增大,推理性能的优势更加明显.体现出 LRSBMAP 算法在多链路拥塞下的性能优势.随着 CLR 的增大,LRSBMAP 算法的性能下降较 CLINK 算法趋缓.当 $CLR=0.5$ 时,CLINK 算法在 GLP 及 BA 模型下的 DR 不足 55%,在 Waxman 模型下仅有 40%;LRSBMAP 算法在 GLP 模型下的 DR 仍保持在 75%左右,在 BA 及 Waxman 模型下分别为 65%和 55%以上.两种算法均在 GLP 模型下的 FPR 最低,其次是 BA 及 Waxman 模型.随着 CLR 的增大,FPR 均先呈缓慢上升趋势,当 CLR 达到一定

比例时,FPR 呈下降趋势.

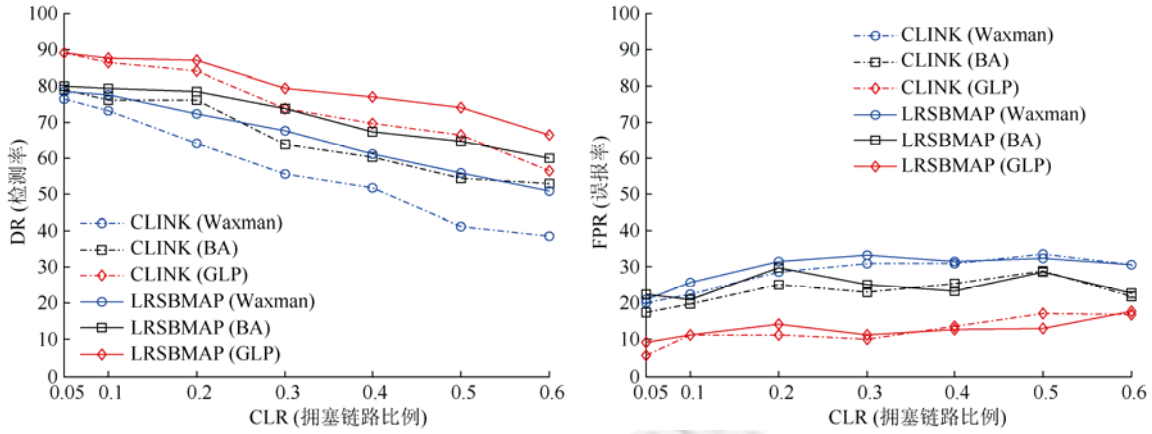


Fig.9 Inference performance comparisons under the different CLR (*node number*=150)

图 9 不同 CLR 下,两种算法的推理性能比较(节点数=150)

(2) 不同网络规模对算法的影响

为了验证算法在不同 IP 网络类型、不同规模下的推理性能.利用 Brite 生成节点数 50~500 的 Waxman、BA 及 GLP 网络拓扑模型.设置多链路拥塞场景,*CLR*=0.5.LRSBMAP 算法及 CLINK 算法的 DR 及 FPR 如图 10 所示.

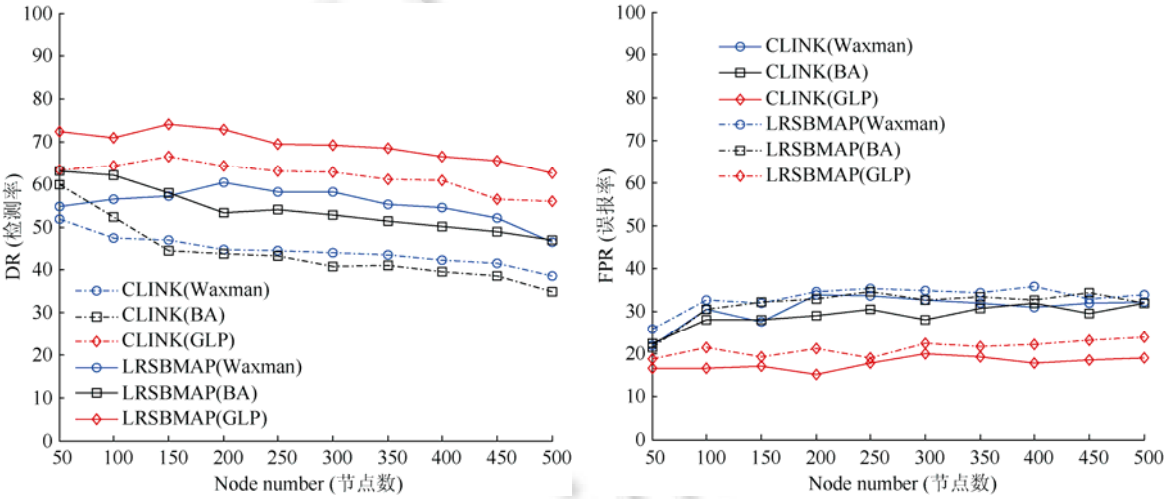


Fig.10 Inference performance comparisons under the different network scales (*CLR*=0.5)

图 10 不同网络规模下,两种算法的推理性能比较(*CLR*=0.5)

由图 10 可见,在不同类型、不同规模的 IP 网络模型下,两种算法的推理性能均随网络规模的扩大呈缓慢下降趋势.其中,LRSBMAP 算法对 Waxman,BA 及 GLP 模型推理性能均优于 CLINK 算法,在 GLP 模型下,DR 最高,其次是 BA 及 Waxman 模型.在 GLP 模型下,FPR 最低,其次是 BA 及 Waxman 模型.两种算法在 Waxman,BA 及 GLP 模型下的推理 FPR 均随着 IP 网络规模的扩大,基本保持稳定.在 GLP 模型下的 FPR 均较 BA 及 Waxman 模型下要低.在 3 种不同网络模型下,LRSBMAP 与 CLINK 算法的拥塞链路推理 FPR 差距不大,当 *CLR*=0.5 时,LRSBMAP 算法的 FPR 平均值略高于 CLINK 算法.

(3) 不同 DTV 对算法推理性能的影响

假设 IP 网络 CLR=0.5,通过本文优选 E2E 探测路径及部署点方法,DTV 从模型最小度值增大到待推理 IP 网络中路由器最大度值.得到不同的链路覆盖范围,借助CLINK及LRSBMAP算法分别进行拥塞链路推理.在利用 Brite 拓扑生成器以默认参数生成的 Waxman 模型中,路由器度值范围为 2~14,BA 模型中,路由器度值范围为 2~33,GLP 模型中,路由器度值范围为 1~33.因此,选取 $DTV_{\text{Waxman}}\in[2,14]$,$DTV_{\text{BA}}\in[2,33]$,$DTV_{\text{GLP}}\in[1,33]$.按照本文对 E2E 路径获取优化策略进行 E2E 路径获取及链路覆盖.对 Waxman 模型,当 DTV=8~10 及 12~14 时,E2E 探测路径数及覆盖链路保持不变;对 BA 模型,当 DTV=12~15 及 16~33 时,E2E 探测路径数及覆盖链路保持不变;对 GLP 模型,当 DTV=5~7 及 9~33 时,E2E 探测路径数及覆盖链路不变.因此,为了便于在同一坐标系下比较 LRSBMAP 算法在 3 种不同网络模型下的拥塞链路推理性能,以 DTV 选取 1~16 作为横轴坐标刻度点.当 CLR=0.5 时,在 3 种不同的网络模型下,LRSBMAP 及 CLINK 算法的推理性能如图 11 所示.

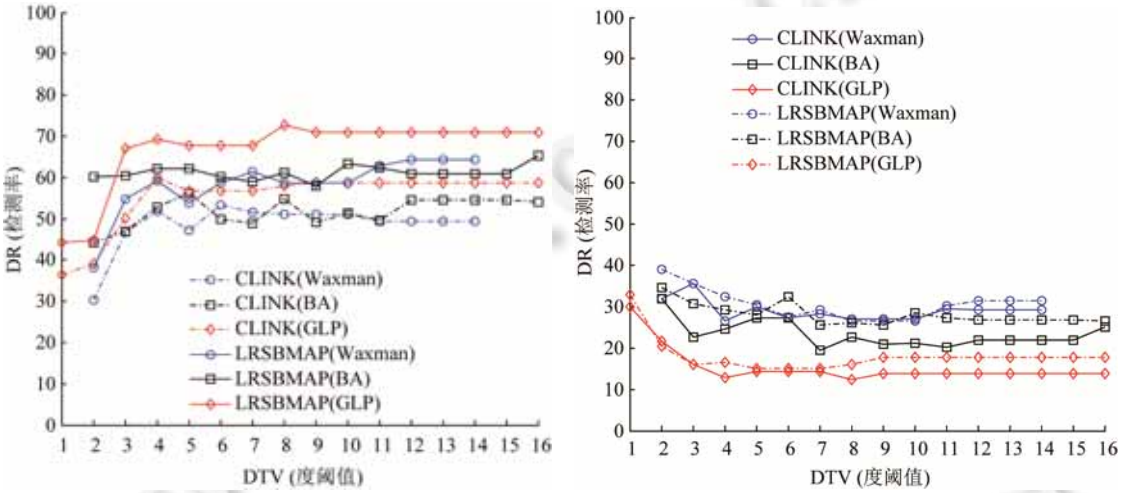


Fig.11 Inference performance comparisons under the different DTV (CLR=0.5)

图 11 不同 DTV 下,两种算法的推理性能比较(CLR=0.5)

由图 11 可见,当 DTV 选取叶子节点或端主机进行探针部署,通过 snapshots 对各 E2E 路径发包探测时,对 GLP 及 Waxman 模型的链路覆盖范围不足,DR 下降得明显,FRP 较高.而对于 150 个节点的 BA 模型,当进行 DTV 优选时,因 DTV_{optimal}=2,是 BA 模型中路由器度值的最小值.此时,链路覆盖率较高,算法推理性能不受影响.通过对不同 DTV 的选取发现,当推理算法对链路覆盖比例达到一定范围后,推理算法对 IP 网络拥塞链路推理性能均趋于稳定,算法推理性能与算法链路覆盖范围之间有一定的联系.

实验结果表明,本文提出的 LRSBMAP 算法在推理性能上较 CLINK 算法有一定程度的提高.原因如下.

(1) CLINK 算法认为,链路包含在越多的 E2E 路径中,越有可能是发生拥塞的链路;

(2) CLINK 算法推理过程中,如已确定某 E2E 路径途经的某条链路为拥塞链路,则该路径及包含该链路的其他路径将从待推理的候选路径集合中去除.显然,这两个拥塞链路推理策略是存在一定问题的.

虽然共享路径最多的链路是最容易拥塞的链路(瓶颈链路).但是,如非瓶颈链路或共享路径较少的链路发生拥塞时,如其拥塞先验概率较大,CLINK 算法中通过引入共享数目|domain(e_k)|的权重,即便瓶颈链路的拥塞先验概率较小,加权后仍被认为是最容易拥塞的链路,推理结果可能存在误差.另外,由于确定某路径中最有可能发生拥塞的链路后,该路径将不再作为候选路径进行该路径下其他链路性能的推理,除非被去除的链路仍存在于其他待推理的候选路径中,否则,将也不再被作为候选链路被推理.因此,LRSBMAP 算法的推理精度较 CLINK 算法有一定程度的提高,特别是当 IP 网络 E2E 路径中同时存在多条链路拥塞的场景下,CLINK 算法的推理性能较本文提出的 LRSBMAP 算法有一定幅度的降低.

8 总结及展望

本文提出一种大规模 IP 网络多链路拥塞场景下的拥塞链路推理算法 LRSBMAP.通过度阈值优选,兼顾链路覆盖率,E2E 探测路径数及探针部署开销,尽可能多地覆盖待测链路;借助矩阵特性,基于 Boolean 代数线性方程组系数矩阵稀疏性提出 PCG_SSOR 算法,迭代求解各链路拥塞先验概率近似唯一解;基于推理时刻剩余拥塞选路矩阵及 BMAP 准则,提出利用改进的拉格朗日松弛次梯度算法推理最有可能发生拥塞的链路集合.实验验证了本文提出算法的准确性及鲁棒性.

考虑到当前 IP 网络中各 AS 区域多为动态路由算法,在可用带宽下降的情况下,将引起路由的动态变化.而当路由改变时,拥塞链路推理时拥塞路径途经的链路可能并不包含在链路拥塞概率学习时的链路集合中.因此,根据 IP 网络选路矩阵构建的贝叶斯网模型结构可能会发生变化,导致推理性能下降.如何准确学习动态路由下各链路的拥塞先验概率,并进行当前时刻拥塞链路定位推理和性能推断,是未来课题的研究方向.

References:

[1] Pan SL, Yang XR, Zhang ZY, Qian F, Hu GM. Congestion link identification under multipath routing for single-source networks. Journal of Electronics & Information Technology, 2015,37(9):2232–2237 (in Chinese with English abstract). [doi: 10.11999/JEIT150058]

[2] Padmanabhan VN, Qiu LL, Wang HJ. Server-Based inference of Internet performance. Technical Report, MSR-TR-2002-39, Redmond: Microsoft Corporation, 2002. 1–12.

[3] Duffield NG. Network tomography of binary network performance characteristics. IEEE Trans. on Information Theory, 2006, 52(12):5373–5388. [doi: 10.1109/TIT.2006.885460]

[4] Nguyen HX, Thiran P. The Boolean solution to the congested IP link location problem: Theory and practice. In: Proc. of the IEEE Int'l Conf. on Computer Communications, INFOCOM 2007. Alaska: IEEE, 2007. 2117–2125. [doi: 10.1109/INFCOM.2007.245]

[5] Coates M, Hero A, Nowak R, Yu B. Internet tomography. IEEE Signal Processing Magazine, 2002,19(3):47–65. [doi: 10.1109/79.998081]

[6] Zarifzadeh S, Gowdagere M, Dovrolis C. Range tomography: Combining the practicality of Boolean tomography with the resolution of analog tomography. In: Proc. of the ACM Conf. on Internet Measurement, IMC 2012. Boston: ACM, 2012. 385–398. [doi: 10.1145/2398776.2398817]

[7] Ma L, He T, Leung KK, Swami A, Towsley D. Monitor placement for maximal identifiability in network tomography. In: Proc. of the IEEE Conf. on Computer Communications, INFOCOM 2014. Toronto: IEEE, 2014. 1447–1455. [doi: 10.1109/INFOCOM.2014.6848079]

[8] Rong ZZ, Jin YH, Cui YD, Yang T. An optimization algorithm for measurement nodes' automatic deployment in distributed network measurement. Chinese High Technology Letters, 2014,24(11):1147–1152 (in Chinese with English abstract). [doi: 10.3772/j.issn.1002-0470.2014.11.008]

[9] Matsuda T, Nagahara M, Hayashi K. Link quality classifier with compressed sending based on ℓ_1-ℓ_2 optimization. IEEE Communications Letters, 2011,15(10):1117–1119. [doi: 10.1109/LCOMM.2011.082911.111611]

[10] Pepe T, Ericsson R, Italy P, Puleri M. Network tomography: A novel algorithm for probing path selection. In: Proc. of the IEEE Int'l Conf. on Communications, ICC. London: IEEE, 2015. 5337–5341. [doi: 10.1109/ICC.2015.7249172]

[11] Adams A, Bu T, Friedman T, Horowitz J, The use of end-to-end multicast measurements for characterizing internal network behavior. IEEE Communications Magazine, 2000,38(5):152–159. [doi: 10.1109/35.841840]

[12] Bu T, Duffield FL, Towsley PD. Network tomography on general topologies. In: Proc. of the SIGMETRICS 2002, ACM SIGMETRICS Int'l Conf. on Measurement and Modeling of Computer Systems. London: ACM, 2002,30(1). [doi: 10.1145/511334.511338]

[13] Lawrence E, Michailidis G, Nair V, Xi B. Network tomography: A review and recent developments. Frontiers in Statistics, 2005,345–366. [doi: 10.1142/9781860948886_0016]

[14] Duffield NG, Presti FL, Paxson F, Towsley D. Inferring link loss using striped unicast probes. In: Proc. of the IEEE Computer and Communications Societies, INFOCOM 2001. Anchorage: IEEE, 2001. 915–923. [doi: 10.1109/INFCOM.2001.916283]

[15] Malekzadeh A, MacGregor M. Network topology inference from end-to-end measurements. In: Proc. of the 27th IEEE Advanced Information Networking and Applications Workshops, WAINA. Barcelona: IEEE, 2013. 1101–1106. [doi: 10.1109/WAINA.2013.215]

[16] Duffield NG. Simple network performance tomography. In: Proc. of the ACM SIGCOMM Conf. on Internet Measurement, IMC

2003. Miami Beach: ACM, 2003. 210–215. [doi: 10.1145/948205.948232]

[17] Ghita D, Argyraki K, Thiran P. Network tomography on correlated links. In: Proc. of the ACM SIGCOMM Conf. on Internet Measurement, IMC 2010. Melbourne: ACM, 2010. 225–238. [doi: 10.1145/1879141.1879170]

[18] Ghita D, KaraKus C, Argyraki K, Thiran P. Shifting network tomography toward a practical goal. In: Proc. of the 7th Conf. on Emerging Networking Experiments and Technologies, CoNEXT. Tokyo: ACM, 2011. 1–12. [doi: 10.1145/2079296.2079320]

[19] Augustin B, Friedman T, Teixeira R. Measuring multipath routing in the Internet. IEEE/ACM Trans. on Networking, 2011,19(3): 830–840. [doi: 10.1109/TNET.2010.2096232]

[20] Kim MS, Kin T, Shin Y, Lam SS, Powers EJ. A wavelet-based approach to detect shared congestion. IEEE/ACM Trans. on Networking, 2008,16(4):763–776. [doi: 10.1109/TNET.2007.905599]

[21] Liu XH, Yin JP, Lu XC, Zhao JM. A monitoring model for link bandwidth usage of network based on weakvertex cover. Ruan Jian Xue Bao/Journal of Software, 2004,15(4):545–549 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/545.htm

[22] Shavitt Y, Sun X, Wool A, Yener B. Computing the unmeasured: An algebraic approach to internet mapping. IEEE Journal on Selected Areas in Communications, 2004,22(1):67–78. [doi: 10.1109/JSAC.2003.818796]

[23] Garey MR, Johnson DS. Computers and Intractability: A Guide to the Theory of NP-Completeness. New York: Freeman Press, 1979. 109–118.

[24] Xing WX, XIE JX. Modern optimization algorithm. 2th ed., Beijing: Tsinghua University Press, 2005. 210–236 (in Chinese).

[25] Shakeri M, Pattipati R, Raghavan V, Patterson-Hine A. Optimal and near-optimal algorithms for multiple fault diagnosis with unreliable tests. IEEE Trans. on Systems Man & Cybernetics (Part C: Applications & Reviews), 1998,28(3):431–440. [doi: 10.1109/ 5326.704583]

[26] Waxman BM. Routing of multipoint connections. IEEE Journal on Selected Areas in Communications, 1989,6(9):1617–1622. [doi: 10.1109/49.12889]

[27] Barabási AL, Albert R. Emergence of scaling in random networks. Science, 1999,286(5439):509–512. [doi: 10.1126/science.286.5439.509]

[28] Albert R, Barabási AL. Topology of evolving networks: Local events and universality. Physical Review Letter, 2000,85(24): 5234–5237. [doi: 10.1103/PhysRevLett.85.5234]

[29] Bu T, Towsley D. On distinguishing between Internet power law topology generators. In: Proc. of the IEEE Computer and Communications Societies, INFOCOM 2002. New York: IEEE, 2010. 638–647. [doi: 10.1109/INFCOM.2002.1019309]

[30] Medina A, Lakhina A, Matta I, Byers J. BRITE: An approach to universal topology generation. Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2001,28(2):346–353. [doi: 10.1109/MASCOT.2001.948886]

附中文参考文献:

[1] 潘胜利,杨析儒,张志勇,钱峰,胡光岷.单源多径路由网络拥塞链路识别.电子与信息学报,2015,37(9):2232–2237. [doi: 10.11999/JEIT150058]

[8] 荣自瞻,金跃辉,崔毅东,杨谈.分布式网络测量的测量节点自动部署优化算法.高技术通讯,2014,24(11):1147–1152. [doi: 10.3772/j.issn.1002-0470.2014.11.008]

[21] 刘湘辉,殷建平,卢锡城,赵建民.基于弱顶点覆盖的网络链路使用带宽监测模型.软件学报,2004,15(4):545–549. http://www/jos.org.cn/1000-9825/545.htm

[24] 邢文训,谢金星.现代优化计算方法.第 2 版,北京:清华大学出版社,2005.210–236.

陈宇(1978－),男,河南浚县人,博士生,副教授,主要研究领域为数据采集与信号处理,网络信息安全.

段哲民(1953－),男,教授,博士生导师,主要研究领域为电路与系统,集成电路分析设计.

温欣玲(1979－),女,副教授,主要研究领域为数据采集,信号处理.

李宇翀(1980－),男,博士,主要研究领域为网络测量,网络安全.