ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.17, No.4, April 2006, pp.838−844
DOI: 10.1360/jos170838

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

基于原始对偶方法求解网络流量监测集算法*

刘湘辉[1,2+]， 殷建平[1]， 卢锡城[1]， 蔡志平[1]， 赵建民[3]

[1](国防科学技术大学 计算机学院,湖南 长沙 410073)
[2](国防科学技术大学 电子科学与工程学院,湖南 长沙 410073)
[3](浙江师范大学 计算机学院,浙江 金华 321004)

Algorithm for the Network Flow Monitoring Set Based on Primal-Dual Method

LIU Xiang-Hui[1,2+], YIN Jian-Ping[1], LU Xi-Cheng[1], CAI Zhi-Ping[1], ZHAO Jian-Min[3]

[1](School of Computer, National University of Defense Technology, Changsha 410073, China)
[2](School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)
[3](School of Computer Science, Zhejiang Normal University, Jinhua 321004, China)
+ Corresponding author: Phn: +86-731-4573675, Fax: +86-731-4573675, E-mail: LiuXh@tom.com

Liu XH, Yin JP, Lu XC, Cai ZP, Zhao JM. Algorithm for the network flow monitoring set based on primal-dual method. *Journal of Software*, 2006,17(4):838−844. http://www.jos.org.cn/1000-9825/17/838.htm

Abstract: Efficiently monitoring the network flow is important to a variety of network applications. Considering the equation of flow conservation, the problem of efficient monitoring is regarded as the problem of finding out the minimum weak vertex cover set and the minimum weak vertex cover set based on flow partition for a given graph $G(V,E)$ which are all proved NP-Complete. Firstly the constraints of weak vertex cover set are analyzed and the integer programming formulation for it is given. Next an approximation algorithm for the minimum weak vertex cover set problem is constructed and the approximation ratio of 2 by primal-dual method is analyzed. Finally the approximation algorithm for the minimum weak vertex cover set is analyzed based on the maximal flow partition.
Key words: weak vertex cover; NP-hard; approximation algorithm; flow conservation

摘 要: 考虑网络节点的流守恒特性,网络流量的有效监测问题可抽象为求给定图 $G(V,E)$的最小弱顶点覆盖集的问题和基于流划分的最小弱顶点覆盖集的问题,这是 NP 难的问题.首先分析了弱顶点覆盖集的约束关系,并给出了问题的整数规划形式.然后利用原始对偶方法构造了求解最小弱顶点覆盖集的近似算法,并分析了算法的比界为 2.进一步分析了求解基于最大流划分的最小弱顶点覆盖集的近似算法.
关键词: 弱顶点覆盖;NP 难的;近似算法;流守恒
中图法分类号: TP301 文献标识码: A

随着 Internet 的日益膨胀,各种各样的网络服务也层出不穷,对网络管理者来说,就需要了解各个节点之间的网络流量,以支持这些可区分的服务.这些及时的流量信息对于许多网管业务,如主动式和被动式的资源管理、流量工程、端到端的服务质量保证,显得尤为重要.特别是现代的网络管理系统注重于服务级、应用级的

* Supported by the National Natural Science Foundation of China under Grant No.60373023 (国家自然科学基金)
Received 2004-11-22; Accepted 2005-07-11

管理,流量的测量过程需要更大的数据量和更高的数据采集频率.然而,目前网络流量的测量方法需要人工合理地针对特定的、感兴趣的链路来规划网络的观测节点,并在这些节点上安装特定的测量软件.这种方法难以扩展,不便于动态适应网络的变化,而且,可能因设置过多的观测节点也会增加网络的额外负担[1,2].

网络流量有效监测方法的关键是,既要准确获取网络流量参数,又要尽量减少由于数据收集对实际网络传输数据的影响[2].给定一个无向图 $G=(V,E)$,其中 V 是顶点集,E 是边集,W 是 V 的子集,若根据与 W 中顶点相关联的各条边的流量,可以确定 E 中任意边的流量.我们称 W 是图 G 关于流量的监测集.有效监测问题的目标是求给定图 $G=(V,E)$关于流量的最小监测集.如果监测代理是路由器或交换机等交换设备,令 $d(v)$表示图 $G=(V,E)$中顶点 v 的度,那么还可挖掘以下两条约束:(1) 对图 G 的顶点集 V 中的任意顶点 v,其度 $d(v)\geq 2$;(2) 对图 G 的顶点集 V 中的任意顶点 v,满足流守恒方程,即流进=流出.故提出以下无向图弱顶点覆盖集的定义[1,2].

定义 **1**. 假定无向图 $G=(V,E)$满足对任意 $v\in V$ 有 $d(v)\geq 2$,称 $W\subseteq V$ 是图 G 的弱顶点覆盖集,当且仅当执行以下操作能使 E 中所有边可以被标记.

(1) 标记所有与 W 中顶点相关联的边.

(2) 若某个顶点 v 的 $d(v)-1$ 条相关联的边已被标记,则标记剩下的那条相关联的边.

(3) 重复第(2)步,直到不能再标记新的边为止.

图 G 的弱顶点覆盖集 W 就是在流守恒约束下的图 G 关于流量的监测集.首先,与集合 W 中顶点相关联的边的流量可通过监测手段来获取.其次,如果 $v\notin V$,且 v 的 $d(v)-1$ 条边的流量已获得,那么根据流守恒方程,可以计算出另外一条边的流量.反复应用流守恒方程,可以计算出图 G 中所有边的流量.

定义 1 利用了节点的流守恒性质,我们通过进一步挖掘流信息,例如在 IP 网络中采用 OSPF[3]协议的链路状态信息或 MPLS[4]中的标记交换路径信息,我们可以获取节点连接链路之间的相互关系,以进一步减少监测节点数目或数据收集时的数据量[5].

假设 $E(v)$是图 $G=(V,E)$中与顶点 v 相关联的边的集合.对任意 $e_v\in E(v)$,定义链路 e_v 的相关链路集 $R(e_v)=\{e\mid e\in E(v)$并且通过链路 e_v 的数据包也通过链路 $e\}$,使用 OSPF 协议和路由表可以确定节点 v 的各条链路 e_v 的相关链路集 $R(e_v)$.

定义 **2**. 给定无向图 $G=(V,E)$中任意顶点 v 的各条边 e_v 的相关链路集 $R(e_v)$,那么顶点 v 的流划分是边集 $E(v)$的一个划分: $E(v)=E_1(v)\cup E_2(v)\cup\ldots\cup E_k(v)$ 且对任意 $1\leq i\neq j\leq k$ 有 $E_i(v)\cap E_j(v)=\varnothing$ 并满足非重叠流特性,即不存在 $1\leq i\neq j\leq k$ 使得有 $e_v\in E_i(v)$, $e'_v\in E_j(v)$ 而 $e_v\in R(e'_v)$ 或 $e'_v\in R(e_v)$.其中 $E_i(v)(1\leq i\leq k)$ 称为顶点 v 的流划分元素.k 取最大值的流划分称为顶点 v 的最大流划分,k 取最小值 1 的流划分称为顶点 v 的最小流划分.

由定义 2 可知:对图 $G=(V,E)$的顶点集 V 中的任意顶点 v,若 $E_i(v)$是其流划分元素,则关于 $E_i(v)$中的各边流守恒方程成立,即流进=流出.这样,对每个顶点可能可以建立多个约束方程,基于这些约束方程,也许能够进一步减少监测节点数量或监测链路数量.

定义 **3**. 假定无向图 $G=(V,E)$满足对任意 $v\in V$ 有 $d(v)\geq 2$,称 $W\subseteq V$ 是图 G 的基于最大流划分的弱顶点覆盖集,当且仅当执行以下操作能使 E 中所有边可以被标记.

(1) 标记所有与 W 中顶点相关联的边.

(2) 若某个顶点 v 的某个最大流划分元素中只有一条边尚未被标记,则标记剩下的这条边.

(3) 重复第(2)步,直到不能再标记新的边为止.

所以,求图 G 关于链路实际使用带宽的测量集的问题便归结为求图 G 的基于最大流划分的弱顶点覆盖集的问题.基于最大流划分的弱顶点覆盖的规模就是 S 中的顶点数.基于最大流划分的弱顶点覆盖问题就是寻求给定图中的最小规模的基于最大流划分的弱顶点覆盖问题.因此,有效监测问题的目标就是求图的最小基于最大流划分的弱顶点覆盖集,即求解基于最大流划分的弱顶点覆盖问题.

但是已经证明求解弱顶点覆盖问题和基于最大流划分的弱顶点覆盖问题是一个 NP 难的问题[5],至今尚无多项式时间的求解算法.令 d 是图 G 中顶点的最大度,我们利用贪婪策略给出了一个求最小弱顶点覆盖集的近似算法,并证明了该算法具有比界 $2(\ln d+1)$,时间复杂性为 $O(|V|^2)$[1].同样利用贪婪策略,Yong Zhang 和 Hong

Zhu 根据有关环路空间的性质,得到了近似程度为 $\ln d+1$ 的算法[6].本文利用原始对偶方法设计了新的求解弱顶点覆盖问题算法,证明了该算法具有比界为 2,时间复杂性为 $O(|V|^2)$.并进一步构造和分析了求解基于最大流划分的弱顶点覆盖问题的算法,证明了算法的比界为 $2D$,其中 D 是图 G 节点包含的最大流划分个数.

利用原始对偶方法进行近似算法分析和设计,首先描述问题的整数规划形式,然后利用松弛条件分析问题的原始可行解和对偶可行解.对于原始规划形式目标为求极小的情形,如果我们可以证明问题原始可行解 $Answer_P$ 和对偶可行解 $Answer_D$,满足约束 $Answer_D/Answer_P \le \alpha(\alpha>1)$,则我们可以构造一个比界为 α 的近似算法[7].

本文第 1 节研究有关弱顶点覆盖的性质并引入一些基本定义和定理,证明弱顶点覆盖集合的一些不等式约束关系,这些约束关系用来导出问题的原始和对偶规划形式.第 2 节基于原始对偶方法构造求解最小弱顶点覆盖集的近似算法,并证明近似算法比界为 2.第 3 节将算法用于求解基于最大流划分的最小弱顶点覆盖集问题,并给出算法的比界 $2d$.最后是总结.

1 弱顶点覆盖集的性质与约束关系

对 $G=(V,E)$而言,考虑节点集合 $S(S\subseteq V)$,记 $E(S)=\{(u,v)|(u,v)\in E,u,v\in S\}$, $G(S)=(S,E(S))$.显然 $E(V)=E$, $G(V)=(V,E)$.

引理 **1**. 假定无向图 $G=(V,E)$满足对任意 $v\in V$ 有 $d(v)\ge 2$,则 $W\subseteq V$ 是图 G 的弱顶点覆盖集,当且仅当 $G(V')$ 是森林,其中 $V'=V-W$(证明见参考文献[1]).

对图 $G=(V,E)$而言,称集合 $\{W|W\subseteq V, W\text{是图}G=(V,E)\text{的弱顶点覆盖集合}\}$为图 G 的弱顶点覆盖集族,记为 $\Omega(G)$.若集合 $W\in\Omega(G)$,对任意节点 $v\in W$ 满足 $W-\{v\}\notin\Omega(G)$,我们称 W 是图 $G=(V,E)$的极小弱顶点覆盖集.若 $W_{\min}\in\Omega(G)$,且满足 $\forall W\in\Omega(G)\to|W|\ge|W_{\min}|$,则我们称 $W_{\min}$ 是图 $G=(V,E)$的最小弱顶点覆盖集,并记 $\beta(V)=|W_{\min}|$.

引理 **2**. 假定 $W\in\Omega(G)$,若顶点 $v_x\in W$ 不属于图 $G=(V,E)$的圈中,则 $W-\{v_x\}$ 也是图 G 的弱顶点覆盖集.

证明:考虑图 G 的子图 $G(V')=(V',E')$,其中 $V'=V-W$, $E'=\{(u,v)|(u,v)\in E\wedge u\in V'\wedge v\in V'\}$,根据引理 1,要证明 $W-\{v_x\}$ 也是 G 的弱顶点覆盖集,只需证明 $G(V-(W-\{v_x\}))=G(V'+\{v_x\})=(V_x,E_x)$ 是森林,其中顶点 $v_x\in W$ 不属于图 $G=(V,E)$的圈中.

由于顶点集 W 是图 $G=(V,E)$的弱顶点覆盖集,所以 G 的子图 $G(V')=(V',E')$ 是森林,其中 $V'=V-W$, $E'=\{(u,v)|(u,v)\in E\wedge u\in V'\wedge v\in V'\}$,注意到 $V_x=V'+\{v_x\}$, $E_x=E'+\{(v_x,v')|(v_x,v')\in E, v'\in V'\}$,并且 $G(V')$ 不含圈,若 $G(V'+\{v_x\})$ 含圈,则 v_x 必属于图 (V_x,E_x) 的圈中,又 (V_x,E_x) 是图 $G=(V,E)$的子图,所以顶点 v_x 属于图 $G=(V,E)$的圈中,矛盾.故 (V_x,E_x) 是森林,根据引理 1, $W-\{v_x\}$ 也是 $G=(V,E)$的弱顶点覆盖集. □

引理 **3**. 假定 $W\in\Omega(G)$,若 W 是图 $G=(V,E)$的极小弱顶点覆盖集,则具有性质:对任意顶点 $v(v\in W)$,存在基本圈 C_v,满足 $V(C_v)\cap W=\{v\}$.

证明:反之,对任意顶点 $v(v\in W)$ 假设不存在基本圈 C_v 使得 $V(C_v)\cap W=\{v\}$,则可能有两种情况:

情况 1:节点 v 不存在基本圈 C_v,根据引理 2,图 $G(V-W+\{v\})$ 是树林,根据引理 1, $W-\{v\}\in\Omega(G)$,与 W 是图 $G=(V,E)$的极小弱顶点覆盖集矛盾.

情况 2:包含节点 v 的任意基本圈 C_v^i,满足 $V(C_v^i)\cap W=\{v\}\cup T^i$,其中 $T^i\ne\varnothing$ 且 $T^i\subset W-\{v\}$,则删除顶点集合 $W-\{v\}$ 及其关联边后,生成图 $G(V-W+\{v\})$ 是树林,若不是树林,则存在圈 C,满足 $C\cap W=T\ne\varnothing$,根据 W 是图 $G=(V,E)$的弱顶点覆盖集以及节点 v 的特点, $T\subseteq W-\{v\}$,矛盾.所以根据引理 1, $W-\{v\}\in\Omega(G)$,与 W 是图 $G=(V,E)$的极小弱顶点覆盖集矛盾. □

我们称节点 $v(v\in W)$ 满足 $V(C_v)\cap W=\{v\}$ 的基本圈为标志圈,并记 $C(v)$ 是节点 $v(v\in W)$ 的标志圈族.

引理 **4**. 设 $W\in\Omega(G)$ 是图 $G=(V,E)$的弱顶点覆盖集合,对任意 $S, S\subseteq V$, $E(S)\ne\varnothing$,则 $W\cap S$ 是图 $G(S)=(S,E(S))$ 的弱顶点覆盖集.

证明:由于 $W\in\Omega(G)$,根据引理 1,图 $G(V-W)=(V-W,E(V-W))$ 是树林.由于 $(V-W)\cap S=S-W\cap S$,所以,

图 $G(V-W)$ 的子图 $G((V-W)\cap S)=G(S-W\cap S)$ 也是树林.即 $W\cap S$ 是图 $G(S)=(S,E(S))$ 的弱顶点覆盖集. □

考虑图 $G=(V,E)$,对任意 $S,S\subseteq V$, $E(S)\neq\varnothing$,记 $\alpha(S)=|E(S)|-|S|+1$,则 $\alpha(V)=|E|-|V|+1$.

定理 1. 若 $W\in\Omega(G)$,则 $\sum_{v\in W}d(v)\geq\alpha(V)+|W|$.

证明:若 $W\in\Omega(G)$,令 $V'=V-W$,根据引理 1, $G(V')=(V',E')$ 是森林,所以 $|E'|\leq|V'|-1$,其中:

$$E'=\{(u,v)|(u,v)\in E\wedge u\in V'\wedge v\in V'\}.$$

注意到, $\sum_{v\in W}d(v)=\sum_{v\in V-V'}d(v)=\sum_{v\in V}d(v)-\sum_{v\in V'}d(v)=2(|E|-|E'|)-\sum_{v\in W}d_{V'}(v)\geq 2(|E|-|E'|)-\sum_{v\in W}d(v)$,即 $\sum_{v\in W}d(v)\geq|E|-|E'|$,由 $|E'|\leq|V'|-1$ 可知: $\sum_{v\in W}d(v)\geq|E|-|V'|+1=|E|-|V|+1+(|V|-|V'|)=\alpha(V)+|W|$. □

若 $W_{\min}$ 是图 $G=(V,E)$的最小弱顶点覆盖集,并记 $\beta(V)=|W_{\min}|$.根据定理 1 有

$$\sum_{v\in W}d(v)\geq\alpha(V)+\beta(V) \tag{1}$$

定理 2. W 是极小弱顶点覆盖集,则 $\sum_{v\in W}d(v)\leq 2(\alpha(V)+\beta(V))-2$.

证明:记 $\chi(G)=\{(u,v)|(u,v)\in E(G),u\in W,v\in V-W\}$,令 $V'=V-W$, δ 表示图 $G(V')=(V',E')$ 的连通度,满足 $\delta=|V'|-|E'|$,则 $\sum_{v\in W}d(v)=\sum_{v\in V}d(v)-\sum_{v\in V-W}d(v)=\sum_{v\in V}d(v)-\sum_{v\in V-W}d_{V-W}(v)-|\chi(G)|=2|E|-2|E'|-|\chi(G)|$.

而 $2(\alpha(V)+\beta(V))-2=2|E|-2|V|+2\beta(V)$,故只需证明 $|\chi(G)|\geq 2|W|+2\delta-2\beta(V)$.

根据引理 3,对任意顶点 $v(v\in W)$,存在节点 v 的标志圈族 $C(v)(C(v)\neq\varnothing)$.考虑子图 $H=C\cup G(V-W)$,其中 $C=\sum_{v\in W}\left(\cup\sum_{c\in C(v)}c\right)$,记 $\chi(H)=\{(u,v)|(u,v)\in E(H),u\in W,v\in V-W\}$,由于 $H\subseteq G$,所以 $\chi(H)\leq|\chi(G)|$,故只需证明 $|\chi(H)|\geq 2|W|+2\delta-2\beta(V)$.注意到, C 是节点 v 的标志圈集合,令 $C_{dis}\subseteq C$ 是非相交标志圈的并,对应节点集合记为 W_{dis},则 $|W_{dis}|\leq\beta(V)$.所以我们只需证明 $|\chi(G)|\geq 2|W-W_{dis}|+2\delta$.注意到, C_{dis} 应该与 $G(V-W)$ 各个连通子图相连,所以 $|\chi(G)|\geq 2|W-W_{dis}|+2\delta$. □

根据引理 4,对图 $G=(V,E)$而言,若 $W\in\Omega(G)$,则对任意 $S,S\subseteq V$, $E(S)\neq\varnothing$, $W\cap S$ 是图 $G(S)=(S,E(S))$ 的弱顶点覆盖集,结合定理 1 有

$$\sum_{v\in W\cap S}d_S(v)\geq|E(S)|-|S|+1+\beta(S)=\alpha(S)+\beta(S) \tag{2}$$

2 弱顶点覆盖集的整数规划描述与算法分析构造

对图 $G=(V,E)$而言,记 $V=\{v_1,v_2,v_3,...,v_n\}$,引入变量 $x_v\in\{0,1\}(v\in V)$,对于 $W\in\Omega(G)$,当 x_v 取值 1 时表示节点 v 属于弱顶点覆盖集 W,当 x_v 取值 0 时表示节点 v 不属于弱顶点覆盖集 W.

根据式(1),构造求解图 $G=(V,E)$最小弱顶点覆盖集合的原始规划形式如下:

$$\begin{aligned}&\text{Min}\sum_{v\in V}x_v\\&\text{st}:\sum_{v\in S}d_S(v)x_v\geq\alpha(S)+\beta(S)(\forall S\subseteq V,|E(S)|\neq 0)\quad\text{(Primal)}\\&x_v\in\{0,1\}(v\in V)\end{aligned}$$

若原始的整数规划问题的可行解不是弱顶点覆盖集合,则存在图 $G=(V,E)$的某个圈 C_x 满足:存在节点 v, $v\in C_x$ 使得 $x_v=0$.对于 $C_x(C_x\subseteq V,|E(C_x)|\neq 0)$,约束方程式 $\sum_{v\in C_x}d_S(v)x_v\geq\alpha(C_x)+\beta(C_x)$ 的左边为 0,右边因为 C_x 含圈,故 $\alpha(C_x)\geq 1$,矛盾.所以原始整数规划问题(primal)的可行解是图 $G=(V,E)$的弱顶点覆盖集合.

根据对偶定理,问题原始整数规划的对偶形式如下:

$$\begin{aligned}&\text{Max}\sum_{S}(\alpha(S)+\beta(S))y_S\\&\text{st}:\sum_{S:v\in S}d_S(v)y_S\leq 1(v\in V)\quad\text{(Dual)}\\&y_S\geq 0(S\subseteq V,|E(S)|\neq 0)\end{aligned}$$

构造算法如下:

ComputeSetByPrimalDual(G=(V,E))

{

(1) $Y\leftarrow 0$

(2) $i\leftarrow 0$

(3) $W\leftarrow\varnothing$ /*表示初始弱顶点覆盖输出集为空*/

(4) $w(v)\leftarrow 1$ ($\forall v\in V$) /*对任意属于 V 中的顶点 v 分配初始权重*/

(5) $G_i\leftarrow G$ ($G_i=(V_i,E_i)$)

(6) while W is not feasible for G /*判断 W 是不是 G 弱顶点覆盖集合*/

{

(*) Increase y_{V_i} until $\exists v_i\in V_i:\sum_{S:v_i\in S} d_S(v_i)y_S=w(v_i)$

(10) $W\leftarrow W\cup\{v_i\}$ /*添加节点 v_i 到集合 W */

(11) 对 G_i 反复删除度不超过 1 的所有顶点及其相关联的边,直到不能再删除新的顶点和边.

(12) $i\leftarrow i+1$

}

for $j\leftarrow i$ downto 1 /*保证 W 是极小弱顶点覆盖集*/

(13) if $W-\{v_i\}$ 是弱顶点覆盖集合则 $W\leftarrow W-\{v_i\}$

return W

}

假设 while 循环执行 n 次,令集合 $W=\{v_1,v_2,\dots,v_n\}$,由于初始令 $Y\leftarrow 0$,方程式 $\exists v_i\in V_i:\sum_{S:v_i\in S} d_S(v_i)y_S=w(v_i)$ 中我们不需要枚举所有的 V 的子集,只需考虑集合 V_i ($1\le i\le n$),即 $\exists v_i\in V_i:\sum_{j=1}^{i} d_{V_j}(v_i)y_{V_j}=w(v_i)$.

所以步骤(*)等价于

(7) $\varepsilon=\min\{w_i(v)/d_{G_i}(v)|v\in V_i\}=w_i(v_i)/d_i(v_i)$

(8) $\forall v\in V_i\rightarrow w_i(v)=w_i(v)-\varepsilon * d_i(v)$

(9) $y_{V_i}=\varepsilon$

定理 4. 假设 U^* 是图 G=(V,E)的最小弱顶点覆盖集, U 是算法输出的 U^* 的近似最优解,那么 $|U|<2|U^*|$.

证明:不失一般性,假定问题的原始规划形式最优解为 OPT^P,对偶形式的最优解为 OPT^D,根据对偶定理(duality theorem)有 $OPT^P=OPT^D$.

假定 $y_S\ge 0$ ($S\subseteq V,|E(S)|\ne 0$),则 $|U|=\sum_{v\in V}x_v=\sum_{v\in V:x_v>0}x_v=\sum_{v\in V:x_v>0}\left(\sum_{S:v\in S}d_S(v)y_S\right)x_v=\sum_{S:y_S>0}y_S\left(\sum_{v\in V:x_v>0}d_S(v)x_v\right)$.

根据式 2,对任意 $S,S\subseteq V$ 考虑子图 $G(S)=(S,E(S))$,有 $\sum_{v\in W\cap S}d_S(v)\le 2(\alpha(S)+\beta(S))-2$,

所以 $|U|=\sum_{v\in V}x_v\le\sum_{S:y_S>0}y_S(2(\alpha(S)+\beta(S))-2)<2\sum_{S:y_S>0}(\alpha(S)+\beta(S))y_S\le 2OPT^D=2OPT^P=2|U^*|$. □

定理 5. 在输入为 G=(V,E)时,算法的时间复杂性为 $O(|V|^2)$.

证明:假设集合 W 用链表 L 表示,向量 Y 用动态数组表示 A,图 G_i 用邻接矩阵 M 表示,并用数组 D 记录图中各顶点的度,用链表 Q 记录图中度为 1 的顶点的编号,并引入中间整数变量 d 记录循环中 $\sum_{v\in W}d(v)-|W|$ 的数值.那么,语句(1)~(3)需常数时间;语句(4)的时间复杂性为 $O(|V|)$;语句(5)初始化矩阵 G_i 的时间复杂性为 $O(|V|^2)$;初始化链表 Q 需常数时间.因为每次循环至少去掉一个顶点,所以循环次数最多为 $O(|V|)$.借助中间变量 d,根据定理 1 判断 W 是不是 G 弱顶点覆盖集合只需常数时间,即语句(6)循环是否结束的判断需要常数时间.借助数组 D,语句(7),(8)执行一次的时间复杂性为 $O(|V|)$.语句(9),(10)执行一次需常数时间.语句(11)的执行包括当链表 Q 非空时反复从中取出 j,置 $D[j]=0$,然后扫描邻接矩阵 A 的第 j 行,对任意 $1\le k\le |V|$,若 $M[i,k]=1$,则置 $M[i,k]=M[k,j]=0$ 并置 $D[k]=D[k]-1$,这时若有 $D[k]-1$,则将 k 加入链表 Q.对从 Q 中取出的每个 j,执行上述工作的时间复杂性为 $O(|V|)$.另一方面,在算法的整个执行过程中,每个顶点的编号最多加入链表 Q 一次,所以从 Q 中取出的编号的总数最多为 $|V|$,因此执行语句(11)的总时间复杂性为 $O(|V|^2)$.语句(12)执行一次需常数时间.语句(13)的时间复杂性

为 $O(|V|)$,且至多循环$|V|$次.综上所述,算法的时间复杂性为 $O(|V|^2)$.

3 求解基于最大流划分的弱顶点覆盖集

对图 $G=(V,E)$而言,已知集合 V 中的任意顶点 v 的流划分元素,考虑图 $G_p=(V_p,E_p)$, $V_p=\{p_1,p_2,p_3,...,p_m\}$ 表示图 G 的所有流划分元素.边集 E 到 E_p 上的映射 f,满足:若 $e\in E(e=(u,v),e\in E(u)\wedge e\in E(v))$,其中 $E(u)(E(u)\in E_p)$和 $E(v)(E(v)\in E_p)$,分别是节点 u 和 v 的流划分元素,则 $e_p=f(e)=(E(u),E(v))$.显然,映射过程可以在多项式时间内实现.我们称图 G_p 是图 G 的基于最大流划分的构造图.定义映射 $g:V_p\to V$,对于 $V_p(v_p\in V_p)$,$g(v_p)$表示流划分元素 v_p 对应于图 G 中的顶点.并记 $P(v)(v\in V)$表示顶点 v 最大流划分中包含的所有流划分元素.

对图 $G=(V,E)$的基于最大流划分的构造图 $G_p=(V_p,E_p)$而言,引入变量 $x_p\in\{0,1\}(v\in V_p)$,对于图 G 的基于最大流划分的弱顶点覆盖集 W 而言,当 x_p 取值为 1 时表示流划分元素 p 对应的节点 $g(p)$属于集合 W,当 x_p 取值为 0 时表示流划分元素 p 对应的节点 $g(p)$不属于集合 W.引入变量 $y_v\in\{0,1\}(v\in V_p)$,当 y_v 取值为 1 时表示节点 v 属于集合 W,当 y_v 取值为 0 时表示节点 v 不属于集合 W.

根据式(1),构造求解图 $G=(V,E)$基于最大流划分最小弱顶点覆盖集合的原始规划形式如下:

$$\begin{array}{l}\mathrm{Min}\sum_{v\in V}y_v\\ \mathrm{st}:\sum_{p\in S}d_S(p)x_p\ge\alpha(S)+\beta(S)(\forall S\subseteq V_p,|E_p(S)|\neq 0)\ (1)\\ x_p\le y_v(\forall v\in V,\forall p\in P(v))\ (2)\\ x_p\in\{0,1\}(p\in V_p),y_v\in\{0,1\}(v\in V)\end{array}\quad(\mathrm{Primal})$$

约束条件 1 保证选择的流划分元素集是图 G_p 的弱顶点覆盖集合.约束条件 2 保证若节点 $V_p(v_p\in V_p)$属于图 G_p 的弱顶点覆盖集合,则对应的节点 $v(v=g(v_p))$应属于图 G 基于流划分的弱顶点覆盖集.

根据对偶定理,问题原始整数规划(primal)的对偶形式如下:

$$\begin{array}{l}\mathrm{Max}\sum_{S}(\alpha(S)+\beta(S))a_S\\ \mathrm{st}:\sum_{S:p\in S}d_S(p)a_S-b_v\le 0(\forall p\in V_p)\\ b_v\le 1(\forall v=V)\\ a_S\ge 0(S\subseteq V_p,|E_p(S)|\neq 0)\end{array}\quad(\mathrm{Dual})$$

根据问题的对偶形式,可以构造求解基于最大流划分的弱顶点覆盖算法如下,算法分为两个步骤:首先应用算法 ComputeSetByPrimalDual 求图 G 的基于最大流划分的构造图 G_p 的弱顶点覆盖集合,记 U_p^* 是图 G_p 的最小弱顶点覆盖集,U_p 是算法输出的关于图 G_p 最小弱顶点覆盖集的近似最优解.根据定理 4,且满足:$|U_p|<2|U_p^*|$.然后对图 G 节点集合 V 结合流划分元素集合 U_p 进行扫描,若节点 $v(v\in V)$存在流划分元素 $p_v\in U_p$,则将节点 v 列入图 G 基于最大流划分的弱顶点覆盖集 W.根据定义 3,扫描集合 V 和 U_p 得到的 W 是图 G 基于流划分的弱顶点覆盖集.

定理**6**. 假设 U^* 是图$G=(V,E)$的基于流划分的最小弱顶点覆盖集,U 是算法输出的 U^* 的近似最优解,那么 $|U|<2d|U^*|$,其中 D 是图 G 节点包含的最多流划分个数.

证明:记 U_p^* 是图 G_p 最小弱顶点覆盖集合,记 P_{U^*} 是图 $G=(V,E)$的基于流划分的最小弱顶点覆盖集 U^* 在图 G_p 中对应的流划分元素集合,满足 $|P_{U^*}|\le D|U^*|$.又根据定义 3, P_{U^*} 是图 G_p 的弱顶点覆盖集合,有 $|U_p^*|\le|P_{U^*}|$,即 $|U_p^*|\le d|U^*|$.

注意到,U_p 是算法输出的关于图 G_p 最小弱顶点覆盖集的近似最优解,由于 $|U_p|<2|U_p^*|$,根据算法性质有 $|U|\le|U_p|$,故 $|U|\le|U_p|<2|U_p^*|\le 2D|U^*|$,所以 $|U|/|U^*|\le 2D$. □

若图 G 用邻接矩阵表示,构造其基于最大流划分的构造图 G_p 的时间复杂性为 $O(|V|^2)$,根据定理 5 求解 G_p 最小弱顶点覆盖集的近似最优解 U_p 的时间复杂性为 $O(|V|^2)$,对图 G 节点集合 V 结合流划分元素集合 U_p 采用链表进行扫描的时间复杂性至多为 $O(|V|^2)$,不难得出图 G 的基于最大流划分的弱顶点覆盖集的近似算法的时

间复杂性也为 $O(|V|^2)$.

4 总 结

注意到算法中我们给顶点分配初始权重为1,不难分析,对于求解以下形式的弱顶点覆盖问题[6](基于最大流划分的弱顶点覆盖问题),算法仍然有效.

已知图 $G=(V,E)$,满足对任意 $v\in V$ 有 $d(v)\geq 2$,记 $w(v)$是定义在顶点集合 V 上的非负实函数,Ω 是图 G 的弱顶点覆盖集族(基于最大流划分的弱顶点覆盖集族).定义 Ω 上的一个实函数 $w(X)=\sum_{v\in X} w(v)$.目标是确定一个 X_0 使得 $\sum_{v\in X} w(X_0)=\min_{X\in\Omega} w(X)$.

References:

[1] Breibart Y, Chan CY, Carofalakis M, Rastogi R, Silberszhatz A. Efficiently monitoring bandwidth and latency in IP network. In: Proc. of the IEEE Infocom. 2001. 933−942.

[2] Liu XH, Yin JP, Tang LL, Zhao JM. Analysis of efficient monitoring method for the network flow. Journal of Software, 2003,14(2):300−304 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/14/300.htm

[3] Shaikh A, Isett C, Greenberg A, Roughan M, Gottlieb J. A case study of OSPF behavior in a large enterprise network. In: Proc. of the ACM Internet Measurement Workshop 2002. Marseille, 2002. 217−230.

[4] Mortier R, Isaacs R, Fraser K. Switches and resource-assured MPLS networks. Technical Report, No.510, Cambridge University Computer Laboratory, 2000.

[5] Liu XH, Yin JP, Lu XC, Zhao JM. The monitoring model for the link bandwidth usage of network based on weak vertex cover. Journal of Software, 2004,15(4):545−549 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/15/545.htm

[6] Zhang Y, Zhu H. Approximation algorithm for weighted weak vertex cover. Journal of Computer Science and Technology, 2004,19(6):782−786.

[7] Hochbaum D. Approximation Algorithms for NP-Hard Problems. Boston: PWS Publishing Company, 1997.

附中文参考文献:

[2] 刘湘辉,殷建平,唐乐乐,赵建民.网络流量的有效测量方法分析.软件学报, 2003,14(2):300−304. http://www.jos.org.cn/1000-9825/14/300.htm

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

刘湘辉(1973－),男,湖南湘潭人,博士生,主要研究领域为计算机网络测量,无线传感器网络,网络算法设计与分析.

蔡志平(1975－),男,博士生,主要研究领域为计算机网络测量,近似算法.

殷建平(1963－),男,教授,博士生导师,CCF高级会员,主要研究领域为算法设计与分析,人工智能,模式识别,信息安全.

赵建民(1950－),男,教授,CCF 高级会员,主要研究领域为人工智能,模式识别,计算机网络.

卢锡城(1946－),男,教授,博士生导师,中国工程院院士,CCF 高级会员,主要研究领域为计算机系统结构,计算机网络.