

带宽延迟约束的层次化网络监测模型

刘湘辉^① 景宁^① 殷建平^② 李军^① 靳肖闪^①

^①(国防科技大学电子科学与工程学院 长沙 410073)

^②(国防科技大学计算机学院 长沙 410073)

摘要: 如何既准确获取监测数据, 又减少监测过程对实际网络传输数据的影响是网络研究的一个根本问题。该文提出了层次化网络测量模型, 强调数据聚集过程的传输延迟和监测数据流量占用链路的带宽都小于预设数值同时, 尽量追求聚集节点数目最小化。求解该模型的问题是 NP 难的, 在节点间路由唯一的情况下, 该文设计的近似算法的近似程度为 $\ln d + 1$, 其中 d 是监测节点的数量。

关键词: 网络监测; 监测模型; 近似算法; 层次化

中图分类号: TP393.06

文献标识码: A

文章编号: 1009-5896(2008)03-0712-05

The Hierarchical Network Monitoring Model with Bounded Bandwidth and Delay Constraints

Liu Xinag-hui^① Jing Ning^① Yin Jian-ping^② Li Jun^① Jin Xiao-shan^①

^①(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)

^②(College of Computer Science, National University of Defense Technology, Changsha 410073, China)

Abstract: Designing optimal monitoring infrastructure is a key step for network monitoring. In this paper the problem of optimizing a hierarchical monitoring system is to reduce the cost of deployment of the monitoring infrastructure by identifying a minimum aggregating set subject to bandwidth constraints on the individual links and delay constraint on the aggregating path. The problem is NP-hard and approximation algorithm is proposed with performance guarantee $\ln d + 1$ under unique aggregating route, where d is the number of monitoring object.

Key words: Network Monitoring; Monitoring model; Approximate Algorithm; Hierarchical

1 引言

随着 Internet 规模的急剧增长, 网络的拓扑变得越来越复杂和不规则, Internet 服务提供商和大型企业需要在网络内部部署一个全局的网络监测系统, 实时地获取网络的性能数据, 监测网络的性能和安全状况。由于现代的网络监测系统注重于服务级、应用级的管理, 监测过程需要更大的数据量和更高的数据采集频率。因此, 如何既能准确获取监测数据, 又能减少监测过程对实际网络传输数据的影响, 即如何构成一个有效的网络监测系统, 以保证各类网络应用对有限网络资源的共享, 成为网络研究的一个根本问题^[1-5]。

传统的集中式网络监测模型一般由一个中心节点和若干个监测代理两类实体组成。监测代理分布在路由器、交换机等关键节点上, 负责收集、管理本地的监测数据以及响应中心节点的命令, 而中心节点通过查询各个监测代理的数据来获取整个网络的全局信息。这种集中式网络监测模型有以下弱点: (1)在接近中心节点附近的链路上形成带宽瓶颈, 尤

其当监测过程有较大数据量和较高的数据收集频率时, 这些额外的数据流量对网络的冲击会非常明显, 并严重影响路由器的吞吐量和带宽利用率^[6]。(2)许多应用对数据的时间性敏感, 过时的信息同样将严重影响应用和管理的性能^[7], 从而限制了这种两层结构的集中式网络监测模型的应用。

注意到在大规模网络环境中将大量地分布监测代理, 最终用户直接获取这些监测代理的数据是非常困难的, 考虑系统灵活性和可以获取全局视图要求, 一种理想的监测结构是层次化结构。在拥有全局视图的中心节点和监测代理之间引入一层聚集节点。聚集节点分布在整个网络环境中, 负责收集、管理若干监测代理, 并对所收集的监测代理的数据进行缩减和综合等处理, 提交给中心节点。每个聚集节点对应管理范围称为聚集域。这样中心节点通过获取聚集节点所收集的各自聚集域内监测代理信息, 仍然可以在中心节点得到监测信息的全局视图。由于聚集节点负责收集聚集域内所有监测代理的监测信息, 这种方式同样面临以上两层结构的弱点, 因而聚集域的范围不可能无限大。

考虑到管理和协调聚集节点需要代价, 本文的层次化网络测量模型强调数据聚集过程的传输延迟和监测数据流量

2006-09-06 收到, 2007-04-30 改回

国家自然科学基金(60373023)和中国博士后科学基金(20060400257)资助课题

占用链路的带宽都小于预设数值同时, 尽量追求聚集节点数目最小化。Li 等^[6]在文献中研究了带宽约束的层次化监测模型, Thottan 等^[8]将 Li 的研究拓展到演化网络模型中, 我们在文献[7]中研究了延迟和延迟抖动约束的层次化网络监测模型, 在文献[9]中研究了演化网络中的延迟约束的层次化网络监测模型。本文对带宽延迟约束层次化网络监测模型做分析。文章的剩余部分内容安排如下: 第2节描述了带宽延迟约束的层次化网络监测模型, 第3节对模型的求解策略作了探讨, 第4节是实验测试, 最后叙述了进一步的研究工作。

2 模型描述

无向图 $G(V, E)$ 表示网络监测系统的拓扑结构, 称为监测域。集合 $V_m (V_m \subseteq V)$ 表示监测域中所有的监测节点, 集合 $V_a (V_a \subset V)$ 表示测量域中所有的聚集节点。

定义1 定义监测节点集到聚集节点集的映射 $\lambda: V_m \rightarrow V_a$ 。满足对任意监测节点 $v_m (v_m \in V_m)$, 如果 $\lambda(v_m) = v_a (v_a \in V_a)$ 则顶点 v_m 的监测信息被聚集节点 v_a 聚集。

记节点集合 $S(v_a) = \{v_m | \lambda(v_m) = v_a\}$ 表示聚集节点 v_a 对应的聚集域内所有监测节点。则 $V_m = \sum_{v_a \in V_a} S(v_a)$ 。

给定无向图 $G(V, E)$, 定义链路延迟函数 $D: E \rightarrow R^+$ 。记链路集合 $E(P_{u,v})$ 表示路径 $P_{u,v}$ 上所有相关链路, 则路径上的延迟 $\text{Delay}(P_{u,v}) = \sum_{e \in E(P_{u,v})} D(e)$ 。对任意监测节点 v_m 定义

节点监测数据被聚集的带宽要求函数 $B: V_m \rightarrow R^+ \cup \{0\}$ 。对任意链路 e , 定义提供给监测信息被聚集的最大可利用带宽函数 $U: E \rightarrow R^+ \cup \{0\}$ 。则带宽延迟约束的层次化网络监测模型为以下定义:

定义2 已知监测域的拓扑结构为 $G(V, E)$, 节点子集 V_m 表示监测域中的所有监测节点, 节点子集 V_a 是满足带宽延迟约束的聚集节点集合。对任意监测节点 v_m 存在唯一聚集节点 v_a 使得 $\lambda(v_m) = v_a$, 且满足:

(1) 对任意聚集节点 v_a 满足 $\forall v_m \in S(v_a) \rightarrow \text{Delay}(\text{Path}_{v_a, v_m}) \leq \delta$, 其中 δ 称为模型的延迟容忍系数。

(2) 对任意链路 e 满足流经链路 e 聚集的监测数据的流量小于链路可以提供给监测信息聚集的最大可利用带宽 $U(e)$ 。

由于到管理和协调聚集节点需要代价, 求解满足带宽延迟约束的层次化网络监测模型的问题就是求满足以上带宽和延迟约束的最小聚集节点集。下面给出问题的整数规划形式。

引入 0-1 变量 x_{ij} 表示是否监测节点 v_i 分配给聚集节点 v_j 。0-1 变量 b_e^{ij} 表示是否链路 e 在节点 v_i 和 v_j 之间的连接路径上。0-1 变量 y_j 表示顶点 $v_j (v_j \in V)$ 是否是聚集节点。

则寻找最小聚集节点集的整数规划形式如下:

$$\text{目标: Minimize } \sum_{j=1}^{|V|} y_j$$

约束条件为

$$\sum_{j=1}^{|V|} x_{ij} = 1 (\forall v_i \in V_m) \quad (1)$$

$$x_{ij} \leq y_j (\forall v_i \in V_m, \forall v_j \in V) \quad (2)$$

$$\sum_{e \in E} b_e^{ij} D(e) x_{ij} \leq \delta (\forall v_i \in V_m, \forall v_j \in V) \quad (3)$$

$$\sum_i \sum_j b_e^{ij} B_i x_{ij} \leq U(e) (e \in E) \quad (4)$$

$$x_{ij} \in \{0, 1\} (\forall v_i \in V_m, \forall v_j \in V) \quad (5)$$

$$y_j \in \{0, 1\} (\forall v_j \in V) \quad (6)$$

约束条件式(1)表示监测节点 v_i 恰好分配给一个聚集节点。约束条件式(2)保证如果某监测节点分配给顶点 v_j , 则 v_j 必须是聚集节点。约束条件式(3)是保证聚集节点与监测节点之间的监测信息聚集满足延迟约束。约束条件式(4)保证监测域中的流经链路 e 的监测数据的流量小于链路可以提供给监测信息聚集的最大可利用带宽。

3 求解策略

作者在以前的研究^[7]中已经证明了求解满足延迟约束的层次化网络监测模型的优化问题是 NP 难的, 所以求解满足带宽延迟约束的层次化网络监测模型的优化问题也是 NP 难的。在以上给出模型的整数规划形式的基础上, 可以采用分支定界的方法求解。众所周知, 分支定界方法最坏情况将有指数级的运行时间花费。由于寻找最小聚集节点集是 NP 难问题, 这就意味着至今尚无多项式时间算法来求解问题的解。实际应用的需要迫使我们退一步设法求出该问题的近似最优解。

实际中一般采取贪婪策略, 在随机或按照一定策略选择某节点作为聚集节点后, 能尽可能多地将来未分配的监测节点分配给该节点, 且监测信息的聚集过程满足以上带宽和延迟的约束。然而令人遗憾的是: 假设节点 v 作为聚集节点, 求解该节点最大监测节点集合且监测信息的数据聚集过程满足以上带宽和延迟约束也是 NP 难的。

然而在实际应用中可以将问题进一步简化, 只考虑节点间只有唯一路由的情况下(如路由预先给定, 以最短路径(Open Short Path First)作为节点间的路由)。下面分析节点间路径唯一情形下的一种求解算法。

已知监测域 $G(V, E)$ 和监测节点子集 V_m , 记 $C(v)$ 为节点 $v (v \in V)$ 作为聚集节点时, 监测信息的数据聚集过程满足以上带宽和延迟约束的监测节点集合。对任意顶点 v , 可以采用宽度优先算法计算其最大可能分配的监测节点集合。若从聚集节点 v 到其他监测节点只存在唯一的路由, 以下算法 ComputeMonitoringNodeSet 将计算集合 $C(v)$:

ComputeMonitoringNodeSet (v)
{

```

 $C(v) = \Phi$  /*初始化候选监测节点集合为空*/
 $Q_W = \{v\}$  /*初始化待处理节点队列*/
 $Q_D = \Phi$  /*初始化已处理节点队列为空*/
Delay(Pathv,v) = 0
While  $Q_W$  not empty
{
    Select and remove first node from
 $Q_W$  which is named  $u$ .
    Repeat select all nodes adjacent to  $u$ 
named  $w$  where  $w \notin Q_D$  And Let Delay(Pathu,v) =
Delay(Pathu,v) + Delay( $u, w$ )
    {
        If Delay(Pathu,v)  $\leq \delta$ 
        Then  $Q_W = Q_W \cup \{w\}$ 
        If Delay(Pathu,v)  $\leq \delta$  And  $w \in S_m$ 
then  $C(v) = C(v) \cup \{w\}$ 
    } /*Repeat 循环结束, 计算出满足延迟约束
的监测节点集合*/
    }
    Sort  $C(v)$  increasingly by bandwidth request /*
满足延迟约束的监测节点集合  $C(v)$  按带宽要求从小到大排
序*/
    For (count = 0 to  $|C(v)|$ )
    {
        Select and remove first node from
 $C(v)$  which is named  $u$ .
        If  $P_{vu}$  violating the bandwidth constraint
then remove  $u$ 
        Else Re-Calculate the residual bandwidth
in Edges relative to path  $P_{vu}$ .
    } /*For 循环结束, 计算出满足延迟和带宽约束的
监测节点集合*/
    Return  $C(v)$ 
} /*结束*/

```

在节点间只有唯一路由的情况下, 算法 ComputeMonitoringNodeSet 可以精确求出某节点作为聚集节点时, 监测信息的数据聚集过程满足延迟带宽约束的最大监测节点集, 同时算法在有穷步内结束。所以节点间只有唯一路由的情况下, 求解某节点作为聚集节点对应的满足延迟带宽约束的最大监测节点集问题是属于 P 类的。

然后采取贪婪策略, 每步选择最大可能分配的监测节点集合 $C(v)$ 。

Procedure ComputeAggregatingNodeSet() {

(1) $S \leftarrow V_m$ /*初始化未分配监测节点集合为整个监测节点集合*/

(2) $U \leftarrow \Phi$ /*初始化聚集节点集合为空*/

(3) While ($S \neq \Phi$) {

(4) Pick a node $v(v \in V, v \notin U)$ satisfying that maximize $C(v) \cap S$

(5) $U \leftarrow U \cup \{v\}$

(6) $S \leftarrow S / C(v)$

}

(7) return U } /*结束*/

令 $v_1^m, v_2^m, \dots, v_n^m$ 是以上算法顺次选择分配的监测节点。

对于任意 $v_k^m (v_k^m \in V_m)$ 假设在算法第 4 步被选择的聚集节点 v_i^m 的监测节点集合 T_i 覆盖。不失一般性选择一个聚集节点的代价是 1, 并定义监测节点 v_k^m 代价 $\text{cost}(v_k^m)$ 为 $1/|T_i \cap S|$ 。令算法计算出的聚集节点集合为 U , 最优(小)聚集节点集为 U^* 。有如下引理:

引理 1 对任意 $i(1 \leq i \leq n)$, $\text{cost}(v_i) \leq |U^*| / (n - i + 1)$

证明 假设最优聚集节点集 U^* 的子集 U_i^* 中的节点对应的监测节点集合覆盖了 $v_i^m, v_{i+1}^m, \dots, v_n^m$, 且 $C_i = \{C(v) | v \in U_i^*\} = \{T_{i1}, T_{i2}, \dots, T_{ix}\}$ 。令 $D = \{v_1^m, v_2^m, \dots, v_{i-1}^m\}$, $D' = \{v_i^m, v_{i+1}^m, \dots, v_n^m\}$ 。考虑顶点 $v_l^m \in D'$, 设 T_{lj} 是集合 C_i 中第一个覆盖 v_l^m 的监测节点集合。定义 $f(v_l^m) = 1/|T_{lj} - D - T_{i1} - T_{i2} - \dots - T_{i_{j-1}}|$ 。注意到 $\sum_{l=i}^n f(v_l^m) = |C_i| \leq |U^*|$, 所以对于 $v_l^m \in D'$ 满足 $f(v_l^m) \leq |U^*| / (n - i + 1)$ 。

在贪婪算法选择的聚集节点对应的监测节点集合覆盖顶点 v_i^m 以前, 由于 D' 中节点没有被覆盖, 则监测节点集合 $T_{i1}, T_{i2}, \dots, T_{ix}$ 对应的任意聚集节点不会被算法选择。如果某聚集节点被算法选择, 且对应的监测节点集合是 T_{ij} , 满足 T_{ij} 覆盖 v_i^m , 顶点 v_i^m 将分配代价: $\text{cost}(v_i^m) = 1/|T_{ij} - D| \leq f(v_i^m) \leq |U^*| / (n - i + 1)$ 。证毕

利用引理1, 将证明以下定理。

定理 1 $|U| \leq H(n) |U^*|$, 其中 $H(x) = \sum_{i=1}^x 1/i$, $n = |V_m|$ 。

证明 $|U| = \sum_{i=1}^n \text{cost}(v_i) \leq |U^*| \sum_{i=1}^n 1/(n - i + 1) = H(n) |U^*|$ 。

进一步由于 $H(n) \leq \int_1^n (1/x) dx + 1 = \ln n + 1$, 故 $|U| \leq H(n) |U^*| \leq (\ln n + 1) |U^*|$ 。证毕

需要说明的是, 当某节点 v_m 可以成为多个聚集节点的成员时, 可以按不同的策略进行指派。如:

(1) 先到先分配 即节点 v_m 若先被聚集节点 v_a 对应的监测节点集合 $C(v_a)$ 命中, 则将 v_m 指派给 v_a 做成员。

(2) 延迟最小优先 若从节点 v_m 到多个聚集节点的路径都满足延迟要求, 且节点 v_m 到聚集节点 v_a 的路径延迟最小,

则将 v_m 指派给 v_a 做成员。

(3) 随机选择 若从节点 v_m 到多个聚集节点的路径都满足延迟要求，则将 v_m 指派给任意一个满足延迟要求的聚集节点 v_a 做成员。

4 实验比较

4.1 聚集节点规模的比较

本小节实验基于节点间最短路由作为唯一路径的情况下，考察算法求解的聚集节点规模。测试中采用 Waxman 模型^[10]作为网络拓扑结构。在 Waxman 模型中网络节点随机分布在笛卡儿空间，考虑节点所有可能连接，网络链路以一定概率方式加入图中。其中概率计算方法为 $P_{e(u,v)} = \beta \exp(-d(u,v)/\alpha L)$ 。其中 L 是网络任意节点间最大距离， $d(u,v)$ 是节点间的距离，参数 $0 \leq \alpha, \beta \leq 1$ ，调整连接链路的数目和连接节点距离特征。固定参数 $\alpha = 0.2$ ，通过变化 β 来产生不同的网络拓扑。任意链路分配一定比率的带宽资源给监测信息的数据聚集，在测试中假定分别为 5% 和 10%。链路的延迟值在 $[0.1, 2.0]$ 范围内变化，而延迟容忍系数 $\delta = 10$ 。测试比较的基准为问题的整数规划的分支定界的方法求解。

表 1 节点数为 200 时聚集节点规模的比较

	带宽资源分配比率 (%)	β 参数 (%)			
		0.05	0.10	0.15	0.20
线性规划	5	11.973	7.457	5.893	4.067
	10	13.287	10.018	72.856	5.427
文中算法	5	13.265	9.364	7.327	5.985
	10	15.284	14.136	10.956	8.049

表 2 节点数为 400 时聚集节点规模的比较

	带宽资源分配比率 (%)	β 参数 (%)			
		0.05	0.10	0.15	0.20
线性规划	5	15.278	11.487	8.975	7.013
	10	17.345	15.857	10.475	8.976
文中算法	5	-	-	-	-
	10	-	-	-	-

从以上实验结果表明，节点间路径唯一情形下，贪婪算法求出的聚集节点集合的规模比较接近精确解。

4.2 负载均衡的比较

本小节实验考察当某节点 v_m 可以成为多个聚集节点的成员时，不同的指派策略的性能。对算法的评价一方面希望聚集节点数量尽量少，另外一方面聚集节点之间负载均衡也是考虑的指标。若按带宽延迟约束的要求把集合 V_m 分成

S_1, S_2, \dots, S_k 互不相交的子集，引入不平衡因子 λ ， λ 满足： $\max(|S_i| - n/k) \leq \lambda (i = 1, 2, \dots, k)$ ，该参数依赖于聚集节点之间的实际负载能力差异，是评测算法效能的一个重要指标。指派成员的策略，如上所述包括：先到先分配、延迟最小优先、随机选择 3 种。本小节实验对应的节点数为 600, 800, 1000。每组节点数随机产生 10 组拓扑。比较的参数包括：孤立点比例和相对不平衡因子 $\lambda/(n/k)$ ，结果如表 3。

表 3 负载均衡的比较

节点规模	孤立点比例 (%)		相对不平衡因子			
	先到先分配	延迟最小优先	先到先分配	延迟最小优先	先到先分配	延迟最小优先
	600	11.75	0	8.39	0.8336	0.4496
800	5.31	0	3.37	0.7742	0.4175	0.5561
1000	6.64	0	2.29	0.6816	0.3849	0.6692

从实验的测试数据来看，采用贪婪方法选择聚集节点，延迟最小优先方法分配成员孤立点基本没有，而相对不平衡因子的性能也更好，所以可认为延迟最小优先的成员分配策略是相对合适的。在算法中，实际是通过比较到各个聚集节点的传输延迟来决定自己的被指派关系，可以有效地消除孤立点，减少相对不平衡因子系数。

5 结束语

本文详细描述了带宽延迟约束的层次化网络监测模型。由于该问题是个 NP 难的，在节点之间的数据传输路径是唯一的情况下，本文给出了一个基于贪婪策略近似算法，证明了算法的近似比不超过 $1 + \ln d$ ，其中 d 是网络中监测节点的数目，并测试算法的性能。网络环境中节点之间的数据传输路径唯一只是一种简单情况，下一步研究将考虑数据传输多路径情况。

考虑到以上网络监测模型的研究都是针对已知的网络拓扑。然而大型网络的拓扑结构是演化、渐变的，具有一定的动态性。所以动态拓扑结构的层次化监测模型的研究也是进一步研究的方向。

参考文献

[1] Asgari A, Trimintzios P, Irons M, Pavlou G, Egan R, and den Berghe S V. A scalable real-time monitoring system for supporting traffic engineering. In: Proc. of the IEEE Workshop on IP Operations and Management, Dallas, 2002: 202-297.

[2] Breitbart Y, Chan CY, Garofalakis M, Rastogi R, and Silberschatz A. Efficiently monitoring bandwidth and latency in IP networks. In: Proc. of the IEEE INFOCOM, Anchorage, 2001: 933-942.

- [3] Liotta A, Pavlou G, and Knight G J. Active distributed monitoring for dynamic large-scale networks. Proceedings of the IEEE International Conference on Communications (ICC'01), Helsinki, Finland, June 2001, Vol. 5: 1544~1550.
- [4] Liotta A, Pavlou G, and Knight G. Exploiting agent mobility for large scale network monitoring [J]. *IEEE Network, special issue on Applicability of Mobile Agents to Telecommunications*, 2002, 16(3): 7-15.
- [5] Liu X H, Yin J P, Tang L L, and Zhao J M. Analysis of efficient monitoring method for the network flow [J]. *Journal of Software*, 2003, 14(2): 300-304.(in Chinese with English abstract).
- [6] Li L, Thottan M, Yao B, and Paul S. Distributed network monitoring with bounded link utilization in IP networks. In: Proc. of the IEEE INFOCOM, San Francisco, 2003: 1189-1198.
- [7] Liu X H, Yin J P, Lu X C, Cai Z P, and Z J M. The distributed monitoring model with bounded delay constraints [J]. *Wuhan University Journal of Natural Science*, 2004, 9(4): 429-434.
- [8] Thottan M, Li L, Yao B, Mirrokni V S, and Paul S. Distributed network monitoring for evolving IP networks. In: Proc. of the 24th Int'l Conf. on Distributed Computing Systems, Hachioji: 2004: 712-719.
- [9] Cai Z P, Yin J P, Liu F, and Liu X H. Distributed monitoring model with bounded delay. *Journal of Software*, 2006, 17(1): 117-123. (in Chinese with English abstract).
- [10] Waxman B M. Routing of multipoint connections [J]. *IEEE Journal on Selected Areas in Communications*, 1988, 6(9): 1617-1622.
- 刘湘辉: 男, 1973 生, 博士, 博士后, 主要研究领域为计算机网络测量、网络安全、网络算法设计与分析.
- 景 宁: 男, 1963 生, 博士, 教授, 博士生导师, 主要研究领域为分布式数据库、地理信息系统.
- 殷建平: 男, 1963 生, 博士, 教授, 博士生导师, 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.
- 李 军: 男, 1972 生, 博士, 副教授, 主要研究领域为分布式数据库、计算机应用技术.
- 靳肖闪: 男, 1978 生, 博士生, 主要研究领域为算法设计与分析、计算机应用技术.