

网络编码研究

付琳 付志雄
(成都信息工程学院 成都通信勘察设计院)

摘要: 最大流最小割的理论决定了网络的最大吞吐量, 网络编码可以使这一理论在单源多播方式的网络环境下得以实现。本文就网络编码的概念进行了详细说明, 并介绍了这个新兴技术的优点和不足之处, 以及它未来的研究方向。

关键词: 网络编码 吞吐量 负载均衡

中图分类号: TN711

文献标识码: A

文章编号: 1672-3791(2007)03(a)-0094-02

1 引言

在现有的计算机通信网络中, 信息传输都是由源节点经过中间节点, 以存储转发的方式传送到目标节点的。除了数据复制以外, 一般来说在网络的中间节点并不需要做任何数据处理。在许多实际应用中, 人们为了信息分析、信息安全以及交换的目的, 总是要在中间节点进行某种形式的数据处理。人们普遍认为, 中间节点所进行的数据处理对数据传输过程本身并不会带来任何好处。

编码最初是在物理层, 在高层是否能带来好处一直以来是有争议的^[1]。而网络路由的传统操作是尽可能避免数据流的碰撞, 但最近研究显示在多播环境下与单一路由相比, 如果允许中间节点处理信息可明显提高

传输速率^[2]。网络编码的核心思想就是允许并提倡网络中间节点对信息进行融合。路由本身则被视为一种特殊的编码, 即节点的输出是输入的一组数字排列。

网络编码主要优势在于如果已知网络的链路性能, 就可以计算出源点到多个终点的最大信息吞吐量, 它可以使资源得到充分利用, 并且能使网络资源利用达到理论上限。这有赖于网络节点(比如路由器)在转发信息前对接收数据进行编码。具备了网络编码功能的节点从各个输入链路获得信息, 并进行编码, 然后把信息传送给所有输出链路^[3]。

另外, 编码还能用于网络边缘移动造成的永久性连接失败, 甚至是在各种失效环境下, 如果要建立一个多播连接, 可通过单个静

态编码来确保连接的健壮性^[4]。

网络编码受到很多关注, 吸引了众多研究兴趣, 它对网络的管理和设计有着巨大影响。网络编码必将更好地挖掘出诸如网络互连和无线带宽等共享资源, 而且从理论上讲, 它也是一门非常具有吸引力的交叉学科, 该领域涉及信息论、算法、代数、编码理论以及图论^[5]。

2 网络编码的概念

我们先定义一些术语和名词概念以便描述网络编码的问题, 然后解释一下用于表征已知网络允许码速的最大流最小割理论。

我们用有向图 $G=(V, E)$ 来表示点到点的通信网络, 其中 E 是边集, 如此一来, 信息可以从节点 i 无噪地传送到节点 j , 有 $(i, j) \in E$ (即一条边就代表了两点间的连接), 而 V 则是网络的点集。我们将关注一种称为单源问题的特殊情况, 也就是网络中只有一个单独的信息源。换句话说, 网络中仅有一个节点 s (称为源点) 传送信息到一组终点 t_1, t_2, \dots, t_L 。对于特定的 L , 问题就被当作是一个单源多终点的问题。边集 $(i, j) \in E$ 的容量用 R_{ij} 表示。

$F=[F_{ij} | (i, j) \in E]$ 是图 G 中由 s 到 t_L 的流量, 如果满足 $(i, j) \in E$, 则 $0 \leq F_{ij} \leq R_{ij}$ 。从而除开 s 和 t_L , 对于所有的 $i \in V$, 就有:

$$\sum_{i:(i,j) \in E} F_{ij} = \sum_{j:(j,i) \in E} F_{ji}$$

也就是说, 输入节点 i 的总流量等于输出节点 i 的总流量。 F_{ij} 被认为是 F 在边集 $(i, j) \in E$ 上的值。 F 的值被定义为:

$$\sum_{j:(j,i) \in E} F_{ji} - \sum_{i:(i,j) \in E} F_{ij}$$

它等价于:

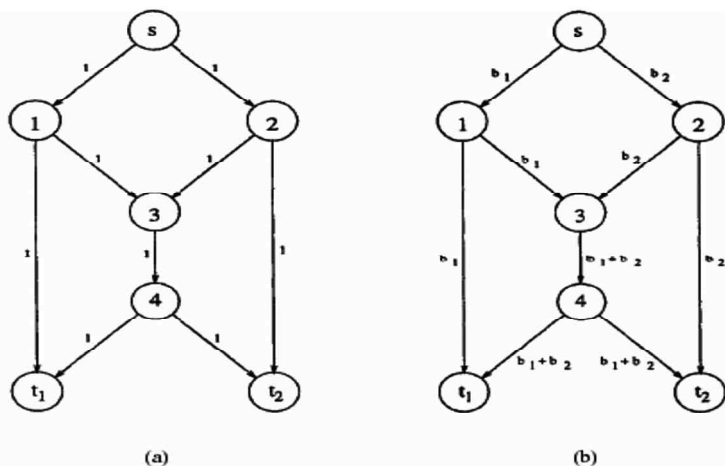


图1 单源两终点网络

上来看, 就需要一些更新、更好的算法, 现在很多科学家都在致力于这方面的研究, 但还不成熟。无论如何, 生物医学或 DNA 的数据挖掘对人类可谓受益非浅。

目前, 在很多领域都用到数据挖掘, 除了上面阐述的几个方面外, 数据挖掘还应用于科学、工业、体育、财务、制造厂、通讯、网络相关行业、零售商、制造业、医疗保健及制药业等等领域。总之, 数据挖掘技术应用非常广泛, 它就是面向应用而诞生的。

2 数据挖掘的未来展望

数据挖掘的信息来源于数据仓库, 而企业获取及利用信息的方式又决定着企业的竞争优势, 所以数据仓库将成为极重要的领

域, 同时市场对于能够将数据转变成可靠与可用信息的系统需求也会增加, 这样将会有一系列的 Data Warehousing (数据仓库化) 程序, 同时将会有更多的人才、数据挖掘工具和数据库以及 OLAP (在线分析处理) 技术更紧密的整合。

3 结语

现代的企业经常搜集市场信息、客户信息、供货商信息、竞争对手信息以及未来趋势等信息, 但是信息超载与无结构化, 使得企业决策单位无法有效利用现存的信息而做出决策。所以妥善的运用数据挖掘技术, 从海量数据库中, 发掘出有用的信息, 作为决策支持之用, 必能产生企业的竞争优势。这也是数据挖掘的应用精华所在。

参考文献

- [1] 范明, 孟小峰. 数据挖掘概念与技术[M]. 机械工业出版社, 2001.8.
- [2] 王哲, 高文, 王实. Web 数据挖掘[J]. 计算机科学, 2000, (4)
- [3] 胡国强. 数据挖掘在远程教育决策支持系统的运用[J]. 开放教育研究, 2003, (5) 44-45.

$$\sum_{i:(i,t_l) \in E} F_{i,t_l} - \sum_{j:(t_l,j) \in E} F_{t_l,j}$$

F 就是图 G 中从 s 到 t_L 的最大流, 它大于或等于从 s 到 t_L 的其它流量。对于单源单终点的图, 从源点到终点的最大流就称为图的容量[2]。

在多播通信中, 可能存在具有多个最大流的多个收端, 因为最大流定义为点到点的通信。为了估计多播网络的唯一容量, 我们定义点到多点通信的最大流为:

$$F_{multicast} = \min F(i), i = (t_1, t_2, \dots, t_L)$$

这里的 $F(i)$ 就是从 s 到 i 的最大流[3], 这个现象与图论里著名的最大流最小割理论很相似。

然而, 对于网络信息流, 最大流最小割理论描述为: 由图 $G=(V, E)$ 表示的网络多播容量 (即从源点到所有收端能同时传送的最大比特数) 等于所有收端的最大流的最小值。因此, 为了寻找网络的容量, 我们应先计算每个收端各自的最大流, 然后再对这些最大流取最小值, 即可得到网络的容量。

下面给出一些例子, 描述最大流最小割理论, 并演示网络编码是如何提高网络吞吐量的。图1演示了单源两终点的情况, 图1(a)给出了网络中每个边的容量。我们可以容易的得到从 s 到 t_L 的最大流为2, $L=1, 2$ 。由最大流最小割理论可知能同时发送2比特信息给 t_1 和 t_2 , 因为图的最大流为2。图1(b)则给出一种方案, 可以产生我们期望的速

率, 很明显每个收端都能正确地接收到数据推断出所发送的2比特信息。事实上, 对于这种特定的结构, 如果没有使用网络编码, 是不可能以2比特速率传送的。

图2给出了网络编码与IP多播等其它多播技术的比较, 由图2(a)和最大流最小割理论可知, 图中容量为4, 而仅仅只有使用了网络编码的图2(d)能以这个速率传送。

3 网络编码的优势

- (1) 在多播环境下可以达到最大传输速率
- (2) 提高频谱利用率
- (3) 提供负载均衡
- (4) 恢复链路失效

4 网络编码的缺点

- (1) 网络中一个分组的丢失, 会在接收端造成多个分组丢失。
- (2) 要接收完所有必要信息, 才能恢复数据, 有延迟。
- (3) 实时系统的同步问题。
- (4) 需要改进网络拓扑结构。
- (5) 可以设计编码来克服链路失败, 但译码时必须知道失败的原因, 即需要链路信息反馈。

5 网络编码的未来研究工作

- (1) 网络中出现多个源点同时多播时, 如何编码, 即多源网络编码问题。

(2) 联合信源信道的网络编码问题。

(3) 改进网络编码的生成算法, 目前是线性的, 因为编/译码实现起来简单。

(4) 对有环网络的编码问题。

(5) 网络安全与网络管理的应用问题。

6 结论

网络编码的理论创新具有普遍意义, 应用前景十分广阔。因而近年来, 网络编码的理论及应用在信息论、编码理论、网络交换、无线通信、计算机科学、运筹学、矩阵理论以及许多其他学科领域, 都受到人们的普遍关注, 网络编码的新应用也不断涌现。

参考文献

- [1] D. Tuninetti, C. Fragouli, "Processing along the way: Forwarding vs. Coding," ISITA, Parma 2004
- [2] R. Ahlswede, N. Cai, S.-Y. R. Li and R. W. Yeung, "Network information flow," IEEE Trans. on Information Theory, vol. 46, pp. 1204-1216, 2000
- [3] Taku Noguchi, Takahiro Matsuda, Miki Yamamoto, "Performance Evaluation of New Multicast Architecture with Network Coding", IEICE Trans. Comm. June, 2003
- [4] R. Koetter, M. Medard, "Beyond Routing: An Algebraic Approach to Network Coding", INFOCOM, 2002
- [5] C. Fragouli, E. Soljanin, A. Shokrollahi, "Network coding as a coloring problem", Processing of CISS 2004

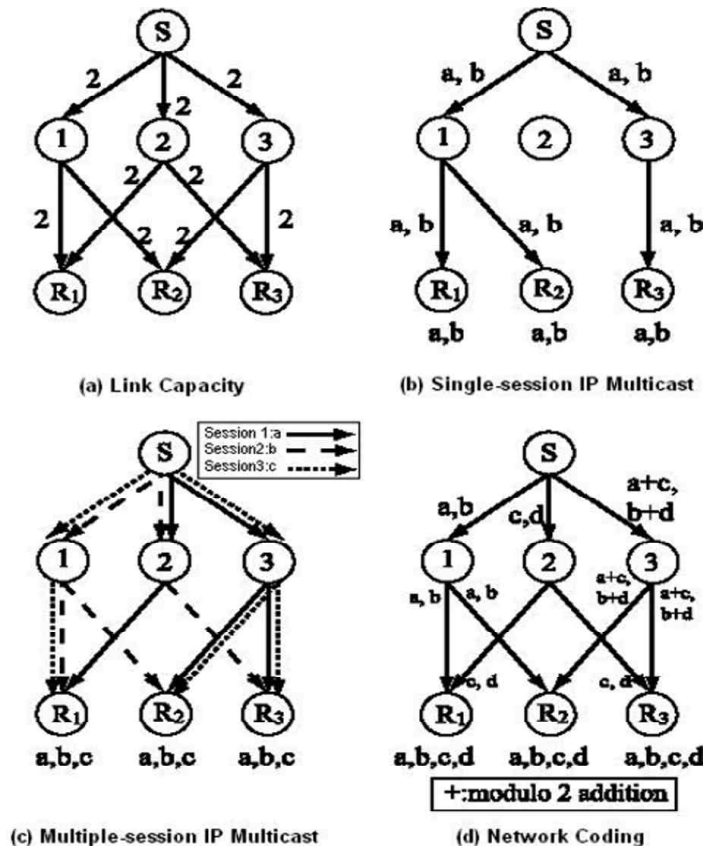


图2 网络编码与已有方案的比较