

线性网络编码及其在 P2P 网络中的应用研究

张旋,姬建新,王颖
(西安理工大学,陕西 西安 710048)

摘要: 通过对网络编码理论和现有 P2P 文件共享系统的深入研究,设计了一种基于线性网络编码的 P2P 文件共享系统。该系统的优点是解决了现有 P2P 文件共享系统中存在的不能充分利用网络资源、“种子”节点突然退出造成文件下载不完等问题。实验结果表明,该系统克服了现有系统中存在的问题,同时也提高了整个系统的吞吐量并增强了系统的稳定性。

关键词: 对等网络;网络编码;比特洪流

中图分类号: TP393.0

文献标识码: A

文章编号: 1674-7720(2013)01-0048-03

Linear network coding and its application research in P2P network

Zhang Xuan, Ji Jianxin, Wang Ying
(Xi'an University of Technology, Xi'an 710048, China)

Abstract: By studying up on network coding theory and the existed P2P file-sharing system, this paper proposes a feasible scheme for P2P file-sharing system based on linear network coding. The advantages of this system are that it can solve the problems of the existed P2P file-sharing system that network resources can not be sufficiently utilized and the source file can not be downloaded completely after the “seeder” node suddenly quit the system. Experiment shows that this system could correctly solve these problems, improve the throughput of P2P network and enhance the robustness of system.

Key words: peer-to-peer network; network coding; bittorrent

对等网络 P2P (Peer-to-Peer Network) 是分布式系统与计算机结合的产物,是采用对等模式工作的计算机网络,它开发了网络中每个节点的网络能力。基于 P2P 的文件共享系统摆脱了原有基于客户/服务器(C/S)体系结构集中式访问模式的束缚,通过汇集网络边缘可用资源提供服务,现已成为 Internet 上下载大文件的主流模型。但是现有的 P2P 文件共享系统也存在一些问题:

(1) 如果“种子”节点突然离开 P2P 网络,可能造成网络中其余计算机不能完整地下载源文件。

(2) 一个好的数据块调度算法时间复杂度高,执行效率低,这将影响到客户端的下载时间。

(3) 在现有的 P2P 通信网络中,信息的传输都是从源节点出发,经过中间节点的存储转发到目的节点。在这个过程中,中间节点起着中继作用,其并未对收到的信息做任何处理,这种传统的通信模式很难达到网络的最大吞吐量。

本文在对网络编码和现有的 P2P 文件共享系统工

作原理研究的基础上,提出了一种基于线性网络编码的文件共享系统实现方案。利用网络编码的优势,来解决现有 P2P 文件共享系统中存在的问题,完善和增强系统的性能。

1 线性网络编码

1.1 网络编码原理

2000年,AHLWEDE R^[1]提出了网络编码。图1以著名的蝶形网络图^[1](Butterfly Network)为模型,阐述网络编码的基本原理。图1中, s 是信源, x 和 y 是信宿,各边的带宽均为1 bit/s,现有2 bit数据 a 和 b 同时从 s 传送到 t_1 、 t_2 。如图1(a)所示,由于两条路径间存在共有的链路 $\langle 3, 4 \rangle$,因此 a 、 b 不能同时在边 $\langle 3, 4 \rangle$ 上传输,则 s 到 t_1 、 t_2 的最大信息流速为1.5 bit/s。若采用网络编码方法,如图1(b)所示,在节点上对 a 、 b 进行异或操作并转发,则节点 t_1 可以通过 $a \oplus b \oplus a$ 的计算解出 b ,同理 t_2 也可以解出 a ,从而使 s 到 t_1 、 t_2 的信息流速达到2 bit/s,该网络的信息吞吐量提高了1/3。由此可见,网络编码可

网络与通信 Network and Communication

以节省带宽,提高信息吞吐量,可以充分利用网络资源,这就是网络编码的优势。

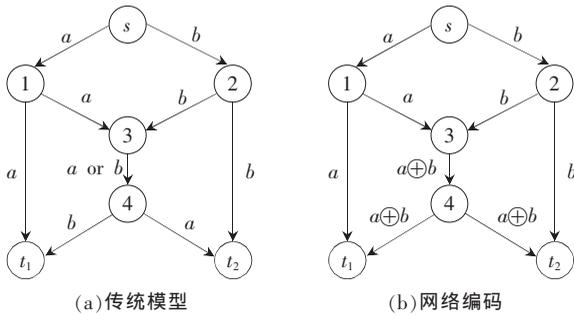


图1 蝶形网络图

1.2 线性网络编码

网络中所有节点对其输入的数据包进行线性操作称为线性网络编码,否则为非线性网络编码。在网络编码概念提出之后,LI S Y R 等人证明了使用线性网络编码已经能够达到网络多播容量^[2]。

(1) 编码。当编码节点有 n 个数据包 (M_1, M_2, \dots, M_n) , 线性编码系数 (g_1, g_2, \dots, g_n) , 其中编码系数都是有限域中的元素。建立方程: $X = \sum_{i=1}^n g_i M_i$, 其中 $g = (g_1, g_2, \dots, g_n)$ 称为编码向量, X 称为信息矢量。编码完成后, 节点将编码向量和信息向量 (g, X) 同时发送出去。将编码向量也发送出去的目的是接收节点对信息向量进行解码, 恢复原始信息。网络中间节点可以对收到的已编码过的信息向量进行再次编码, 然后再转发出去。假定中间节点收到了 m 个已经编码过的数据包 $(g^1, X^1), (g^2, X^2), \dots, (g^m, X^m)$, 它对 m 个信息向量再次编码, 生成新的信息向量 (g', X') , $g' = (g'_1, g'_2, \dots, g'_n)$ 。中间节点再次编码的编码系数是 $h = (h_1, h_2, \dots, h_n)$ 。 X' 表示为:

$$X' = \sum_{i=1}^m h_i X^i = \sum_{i=1}^m h_i \left(\sum_{j=1}^n g_j^i M_j \right) = \sum_{j=1}^n \left(\sum_{i=1}^m h_i g_j^i \right) M_j = \sum_{j=1}^n g'_j M_j \quad (1)$$

(2) 译码。当节点收到 m 个编码后的数据包 $(g^1, X^1), (g^2, X^2), \dots, (g^m, X^m)$ 后, 为了恢复出 n 个原始数据包 M_1, M_2, \dots, M_n , 只需求解方程组 (2), 它是一个含有 n 个未知数和 m 个方程的线性方程组。只有当线性无关方程个数大于或等于未知数个数 (即 $m \geq n$) 时线性方程组才有唯一解。也就是当接收节点至少收到 n 个线性无关的数据包就可以将原始数据解出来, 即:

$$\begin{cases} X^1 = g_1^1 M_1 + g_2^1 M_2 + \dots + g_n^1 M_n \\ X^2 = g_1^2 M_1 + g_2^2 M_2 + \dots + g_n^2 M_n \\ \vdots \\ X^m = g_1^m M_1 + g_2^m M_2 + \dots + g_n^m M_n \end{cases} \quad (2)$$

编码向量之间的线性相关程度取决于所选取的有限域的大小。在实际应用环境中, 符号域的大小为 $q=2^8$ 有限域, 编码向量之间出现线性相关的概率就非常小^[3]。

1.3 线性网络编译码器实现算法说明及性能分析

在实际工作中, P2P 网络的拓扑结构经常会随着节点的加入或退出发生变化, 所以网络编码采用随机线性网络编码^[4], 即网络节点对编码系数是随机选取的, 并且对输入数据包进行线性操作, 其具有良好的拓扑适应性。但从网络编码原理分析可知, 网络编码算法时间复杂度较高, 因此, 本文在编译码器的实现程序中采取了一些优化策略, 以尽可能降低计算量, 提高编译码的速度。策略如下:

(1) 有限域运算。编译码过程中需要有大量的有限域乘除运算, 本程序采用离散对数方法来减少运算量。利用有限域中特殊元素生成元, 将有限域中任何非零元素唯一的表示为生成元的指数形式, 因此有限域元素相乘除的运算都可转化为指数运算。

(2) 对源文件采用“代”(Generation)划分^[3]。P2P 网络中等待下载的文件通常都在百兆以上, 数量庞大的源文件分组使相应的编码矩阵和解码矩阵维数很大, 加剧了编译码运算过程中的运算量。本文对源文件采用“代”划分的方法, 以达到有效降低编译码矩阵维数、简化运算量的目的。

(3) 稀疏矩阵。网络编码的编码系数是随机选择的, 随机选择的编码系数是均匀分布的, 编码系数中零的个数很少, 所以程序的编码矩阵选择使用稀疏矩阵^[5], 这样可以有效降低编译码计算量。

基于随机线性网络编码的原理和上述算法优化策略, 本文采用 C++ 在 Linux 环境下实现了传统的线性网络编码编译码器和随机线性网络编码编译码器 (采用上述优化策略)。用 100 MB 的文件在两种编译码器上分别测试了分组大小从 64 KB 到 2 MB 的编译码运行速率。通过分析图 2 的编译码器性能图可以得到, 经过优化策略的编译码器性能明显优于传统网络编码的编译码器。实验结果证明, 将上述策略应用于网络编码的编译码器是有效的, 所以选择合适大小的分组可以充分发挥编译码器的性能。

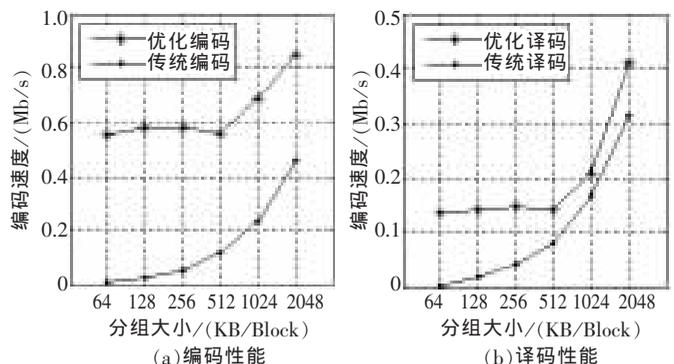


图2 编译码器性能比较图

2 线性网络编码在 P2P 文件共享系统中的应用

当前, BitTorrent 是使用最为广泛的 P2P 文件共享系统之一^[6]。通过对 BitTorrent 系统工作原理、线性网络编

网络与通信 Network and Communication

码理论的研究和分析,将本文提出的线性网络编译器应用到 BitTorrent 客户端系统中,实现基于线性网络编码的 P2P 文件共享系统。图 3 是基于线性网络编码的 BitTorrent 系统框架图。该系统主要由 5 个模块组成:(1)传输机制模块:提供 Socket 的通信,收发消息。(2)编码器模块:对数据进行编码。如果当前的节点是种子节点,就对原始数据进行编码;如果是非种子节点,就对当前收到的同一“代”的编码过的数据进行再次编码。(3)线性检测模块:对收到数据包的编码向量进行线性相关性的检测。如果线性无关,将该数据包存放在临时文件中;如果线性相关,将该数据包抛弃;如果临时文件存放的线性无关的数据包达到一定数量,就可以进行译码。(4)译码器模块:对收到编码过的数据包进行译码,恢复源文件。(5)成员管理模块:让当前客户端能够与系统中一定数目的其他客户端建立连接,并在运行过程中实现对其连接的其他客户端进行淘汰和更新。

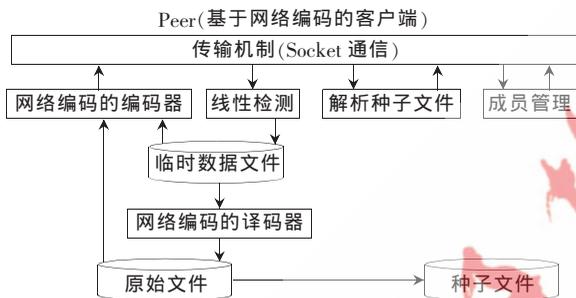


图 3 基于线性网络编码的 BitTorrent 系统框架图

将基于线性网络编码的 BitTorrent 文件共享系统运行在实验室实际的网络环境中(12 台 PC)。在实验过程中,将种子节点中途退出。实验最终结果表明,其余客户端都可以正确地下载完源文件。通过对实验结果进一步分析得到一些结论:(1)该系统可以解决 P2P 网络中“种子”突然退出可能造成的文件下载不完的问题;(2)基于网络编码的 BitTorrent 并没有比普通的 BitTorrent 性能快很多,这是由于编码和译码时间复杂度比较高,从而影

响系统的整体性能。

P2P 文件共享系统已经成为了 Internet 的主要应用之一。本文实现了一种基于线性网络编码的 P2P 文件共享系统。该系统可以提高网络的吞吐量、增强系统的可靠性并提高下载成功率。由于网络编码存在着计算复杂度高的问题,虽然在编译码器的实现方法中结合了一些有效策略,可以降低部分计算量,但是编译码器的计算量开销依然很大,因此研究降低网络编码的复杂性,实现最小代价的网络编码,具有重要的理论意义和使用应用价值。

参考文献

- [1] AHLWEDE R, CAI N, LI S Y R, et al. Network information flow [J]. IEEE Transactions on Information Theory, 2000, 46(4):1204-1216.
- [2] LI S Y R, YEUNG R W, Ning Cai. Linear network coding[J]. IEEE Transactions on Information Theory, 2003, 49(2):371-381.
- [3] CHOU P A, WU Y N, JAIN K. Practical network coding[C]. The 41st Allerton Conference on Communication, Control and Computing Monticello, Kluwer, 2003.
- [4] HO T, MEDARD M, KOETTER R, et al. A random linear network coding approach to multicast [J]. IEEE Transactions on Information Theory, 2006, 52(10):4413-4430.
- [5] MA G, XU Y, LIN M, et al. A content distribution system based on sparse linear network coding[C]. Proceedings of the 3rd Workshop on Network Coding, Theory, and Applications (NETCOD 2007), 2007.
- [6] GKANTSIDIS C, RODRIGUE Z. Network coding for large scale content distribution [C]. INFOCOM 2005, 2005 (4): 2235-2245.

(收稿日期:2012-04-10)

作者简介:

张旋,男,1980 年生,在职博士研究生,讲师,主要研究方向:网络编码和闪存编码。