文章编号:1001 - 9081(2006)07 - 1730 - 02

LDPC 比特翻转译码算法的分析与改进

张 谨,苏广川
(北京理工大学 电子工程系,北京 100081)
(yikangzhang@163.com)

摘 要:利用统计译码思想由 LDPC(Low Density Parity Check)码校验矩阵通过矢量的线性组合构造出一个新的低密度校验矢量集合,并结合 LDPC 码并行比特翻转译码算法的环检测等特点的分析,提出了一种新的硬判决译码方案。仿真结果表明:改进算法在译码性能上接近 BP 算法,又保持了并行比特翻转算法迭代次数少的优点。

关键词:LDPC 译码;比特翻转;统计译码
中图分类号: TP919.22 **文献标识码**:A

Analysis and improvement of the bit-flipping decoding algorithm for LDPC codes

ZHANG Jin, SU Guang-chuan
(*Department of Electronic Engineering, Beijing Institute of Technology, Beijing* 100081, *China*)

Abstract: According to the statistical decoding method, a new enlarged set of the low density check vectors was constructed by the linear combination among the vectors of the LDPC code's old check matrix. Combined with the analysis of the Parallel-Bit-Flipping(PBF) decoding algorithm, such as loop detection technique, a new hard-decision decoding algorithm was presented for LDPC codes. The simulation results show that our algorithm has small iterative numbers just like the PBF algorithm, and the decoding performance approaching the belief propagation algorithm.

Key words: LDPC code; bit flipping; statistical decoding

0 引言

Gallager 在 1962 年提出的 LDPC 码[1](Low Density Parity Check)是一类可以用非常稀疏的奇偶校验矩阵或二分图定义的线性分组纠错码,以其卓越的译码性能备受关注,是 4G 移动通信、深空通信等领域首选的编码方案。目前主要有两大类译码算法[1]:一类是基于概率的置信传播(Belief Propagation)迭代译码算法,简称 BP 算法。这类算法译码在码长较大时性能可逼近香农限,但实现复杂度较高。另一类是基于校验和统计迭代的比特翻转译码算法(Bit Flipping Algorithm),简称 BF 算法。BF 算法实现复杂度很低,目前提出了多种有效的改进方案,如加权的 BF 算法(WBF)及其改进形式[2,3]、结合符号可靠性的软判决 BF 算法[4,5]、引入"环检测"和比特翻转约束机制等[6,7],使其译码性能逐步接近 BP 算法,但同时也使算法复杂性增加不少。

统计译码算法是 A. Al Jabri 为了分析 McEliece 公钥加密体制提出的一种译码方法[8]。研究表明,可以借鉴统计译码思想改进现有 LDPC 的 BF 译码算法。现有各种 BF 算法都是在 LDPC 译码原有校验矩阵 $\boldsymbol{H}$ 的基础上,以逐比特串行翻转算法为基本形式改进而来的。如果利用 $\boldsymbol{H}$ 矩阵中各校验矢量的线性组合能够构造出新的更多低重量的校验矢量,得到扩充的校验矩阵,就可以把统计译码思想与现有 BF 算法相结合,从而提高译码性能。由于并行比特翻转算法是最简单的 BF 算法,与其他算法相比,该算法可以减少迭代次数,且具有很好的环检测能力。本文利用统计译码算法结合并行比特翻转设计译码方案。仿真结果表明:整个方案简洁有效,实现复杂度很低,译码性能较好。整个方案在性能和实现复杂性之间进行了有效的折中。

1 统计译码算法

约定符号:对于任一(n, k)分组码,C 是码字空间 $H = C^{\perp}$,是 C 的对偶空间。$\forall c \in C, h \in H, ch^{\mathrm{T}} = 0$。则 $rh^{\mathrm{T}} = (c \oplus e)h^{\mathrm{T}}$,其中 r 为接收码字,e 称为差错矢量,也叫差错图样。

由 Jabri 提出的统计译码算法,其核心思想是:通过在给定线性分组码的对偶码空间中选择具有相同或相近重量的码矢量集合(记为 $\boldsymbol{Hw}$),利用它对接收码字的校验作用,可以对各种错误图样提供有关错误位置的统计信息,从而实现有效译码。原算法描述[8] 如下:

1)从校验空间 $\boldsymbol{H}$ 中选择重量较大且接近的校验矢量组成集合 $\boldsymbol{Hw}$;

2)对于接收码字 r,计算错误位置统计矢量 $\boldsymbol{v} = \sum_{h \in \boldsymbol{Hw}} (rh^{\mathrm{T}})h$。把矢量 $\boldsymbol{v}$ 中统计数值表现异常(偏离平均值)的位置分量标记为不可靠位置,其他则为可靠位置。

3)根据接收码字中可靠位置的比特信息,利用信息集合译码算法对整个接收码字进行纠错。

上述算法的关键在于合理选取校验子集 $\boldsymbol{Hw}$。文献[2]提出选择重量大的码元组成 $\boldsymbol{Hw}$,此时若接收码元中误码个数为偶数,则 $\boldsymbol{v}$ 在码字相应位置上的统计表现为峰值;反之若误码个数为奇数,则相应的统计位置表现为谷值。由于事先并不知道误码的奇偶性,必须分别假设这两种情况,舍弃不可靠位置的比特符号,选择可靠位置的比特符号,利用信息集合译码算法估计出信息矢量。从统计的角度来看,确定 $\boldsymbol{Hw}$ 的原则是:

1)一致性。$\boldsymbol{Hw}$ 所含矢量的重量尽可能相同或相近,使得各种错误图样尽可能具有相似的统计表现,便于区分。

2)均匀性。$\boldsymbol{Hw}$ 所含矢量在各分量位置上的 0、1 分布,总

收稿日期:2006 - 01 - 04;**修订日期**:2006 - 04 - 07 **基金项目**:国家 863 计划项目(2003AA146010)

作者简介:张谨(1968 -),男,安徽萧县人,副研究员,博士研究生,主要研究方向:编码、信息安全; 苏广川(1941 -),男,江苏扬州人,教授,博士生导师,主要研究方向:通信安全技术.

体上尽可能相同或相近,便于对各分量位置上比特符号的可靠性做出正确判断。

3) 必要性。选择的 $\boldsymbol{Hw}$ 应能达到纠错能力的设计要求。

4) 小容量。为降低算法的总计算量 $\boldsymbol{Hw}$ 所含矢量不宜过多。

事实上还可以选取小重量的校验矢量组成校验矩阵 $\boldsymbol{Hw}$。而且此时无论接收码元中误码个数为偶数或奇数,错误比特在码字相应位置上的统计表现均表现为峰值。由于 LDPC 码(特别是规则 LDPC 码)的校验矩阵中的各矢量正好由重量基本相同的低重量矢量组成,符合统计译码在小重量情况下构造 $\boldsymbol{Hw}$ 的条件,因此统计译码算法自然地可以推广用于 LDPC 译码。但是受到具体编码方法的制约,构造出的 LDPC 码的矢量个数有限,分布也不够均匀。采用一步实现的统计译码算法效果不够理想。因此 LDPC 码现有各种 BF 译码算法均采用迭代形式,都可以理解为迭代的统计译码算法。其共同点是:都在 LDPC 码原有校验矩阵的基础上,根据校验和结果有效估计各分量位置上比特符号的可靠程度,每次对最不可靠的比特进行翻转;该过程迭代进行,直到算法收敛到一个码字矢量或超过规定的最大迭代次数。

下面以(20,15,3,4)规则 LDPC 码为例说明:在给定 LDPC 码校验矩阵以后,可以通过已有校验矢量的线性组合构造更多的校验矢量,组成新的校验矩阵 $\boldsymbol{Hw}$。该码校验矩阵行重量为 4,列重量为 3。现有校验矩阵给出了 15 个重量为 4 的校验矢量(组成 $\boldsymbol{H}_{w1}$)。

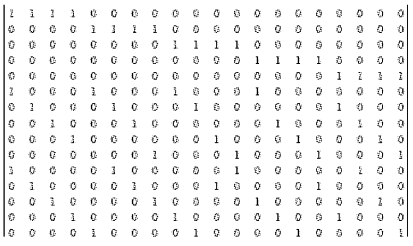

图 1　(20,15,3,4)LDPC 码的校验矩阵

利用有一处"1"位置相同的两个校验矢量对应分量模二加,可以得到 60 个重量为 6 的校验矢量(组成 $\boldsymbol{H}_{w2}$)。如:

$$(111100000000000000) \oplus (100010001000100000) = (011110001000100000)$$

表 1 给出三种规则 LDPC 码在扩展校验矩阵时得到的有关参数。值得注意的是尽管 $\boldsymbol{H}_{w2}$ 比 $\boldsymbol{H}_{w1}$ 的数量增加很多,但由于 $\boldsymbol{H}_{w2}$ 仍保持低密度校验矩阵的特点,因此在译码过程中总的计算量增加并不多。而且实验表明 $\boldsymbol{H}_{w2}$,$\boldsymbol{H}_{w1}$ 分开使用要比合并在一起使用效果更好。

表 1　三种 LDPC 码扩展校验矩阵的有关参数

参数	码 1	码 2	码 3	参数	码 1	码 2	码 3
n	20	504	1 008	$W1$	4	6	6
k	15	252	504	H_{w1}	15	252	504
列重	3	3	3	$W2$	6	10	10
				H_{w2}	60	1 512	3 024

2　两种 BF 算法的性能比较

与 BP 算法类似,环也是影响 BF 算法译码性能的主要因素。这里的环是指在 BF 迭代译码过程中,由于重复翻转同一位置的比特符号,导致译码陷于死循环状态,直至达到最大迭代次数。如果迭代过程用函数 f 来表示,v 表示当前状态,$f^n(v)$ 表示迭代 n 次后的状态。若满足 $f^n(v)=v$,则称 v 是 f 的 n 阶环。由于环的存在使得迭代过程不及时收敛,造成了计算资源的无谓消耗。通过引入环的检测机制,可以及时跳出"死循环",提前终止算法,减少总迭代次数。若再能进一步有针对性地采取各种有效的"消环"措施,就可以大大提高译码性能。

环的存在本质上是由 LDPC 码的结构决定的,客观上受到迭代方式的影响。根据迭代方式的不同,目前硬判决 BF 算法主要包括并行比特翻转译码算法(Parallel Bit Flipping, PBF)和串行比特翻转译码算法(Sequential Bit Flipping, SBF)两类,其他算法都是在这两种算法的基础上加以改进而形成的[4]。PBF 算法步骤如下:

1) 给定最大迭代次数 n,接收码字 r,校验矩阵 $\boldsymbol{H}$,计算统计量 $\boldsymbol{v}=\sum_{h\in H}(rh^{\mathrm{T}})h$;若 $\boldsymbol{v}$ 为 0 矢量,则 r 是码字,算法终止,否则转 2)。

2) 标记 $\boldsymbol{v}$ 中统计数值最大的位置为不可靠位置。

3) 翻转 r 中所有处在不可靠位置上的比特符号。

4) 判断迭代终止条件:迭代次数若小于 n,转 1);否则终止。

SBF 算法与 PBF 基本相同,只是步骤 3) 改为:在 r 中所有处在不可靠位置上的比特符号中随机选择一个翻转。以(1008,504,3,6)LDPC 码为例,下面从环检测能力、总迭代次数、译码性能三个方面比较 PBF 和 SBF 两种硬判决 BF 算法:

1) SBF 的环检测算法比较复杂[7]。但 PBF 的环检测算法要简单得多,只要检测当前不可靠位置与上次迭代的不可靠位置是否完全相同:全同就进入环状态;否则不存在环。

2) PBF 的总迭代次数明显少于 SBF,参见图 2。这是因为在收敛情况下,由于串行迭代每次只改变一个比特,因此无论是固定形式还是随机选取形式,其平均迭代次数都随差错个数至少呈线性增长;而并行迭代的平均迭代次数随差错个数而缓慢增长。在不收敛的情况下,并行迭代能够快速准确检测到环的存在,及时终止迭代。

3) 在低信噪比(Signal Noise Ratio, SNR)(<4.8dB)时 SBF 算法性能优于 PBF 算法;在高信噪比(>4.8dB)时 PBF 算法性能优于 SBF 算法;二者总体上译码性能比较接近,参见图 1。

结论:与 SBF 算法相比,PBF 算法能够以最少的迭代次数和最少的计算量获得较好的译码性能;且能快速检测环。从总体上看 PBF 算法在效率和译码性能上实现了有效均衡,优于 SBF 算法。

3　改进方案与仿真结果

目前通过改进迭代方式达到"消环"目的的主要措施有:通过引入约束长度(设置禁翻)改进迭代方式;采用"翻转概率"代替硬性翻转;结合符号可靠性信息形成软判决 BF 算法等[2~7]。这些方法在提高译码性能的同时,也增加了计算复杂性。由于 PBF 算法有很多优点,若采用级联的译码方案会取得较好效果。整个译码过程可分为 PBF 译码、环检测、环处理三个部分。首先采用 PBF 算法译码,这样能够以较少的迭代次数获得较好的译码性能;再结合环检测,可以及时判断译码过程是否收敛。若不收敛,则利用其他更好但复杂性稍高的算法作进一步处理。下面利用统计译码得到的扩充校验矢量集合 $\boldsymbol{H}_{w2}$,结合 PBF 算法给出一种新的改进译码方案:

1) 首先利用 $\boldsymbol{H}_{w1}$,采用 PBF 算法进行译码。2) 若收敛,且满足效验矩阵则为码字,终止算法。3) 若译码过程不收敛,检查当前不可靠位置与上次迭代的不可靠位置是否完全相同。若相同则存在环,转 4);否则转 1)。4) 利用 $\boldsymbol{H}_{w2}$,仍采用 PBF 算法对接收码矢量重新进行迭代译码。5) 利用环检测机制,再次判断译码过程是否收敛。若收敛,且满足效验矩阵则为码字,终止算法;否则译码失败。

采用(1008,504,3,6)规则 LDPC 码对 SBF,PBF,BP 和改进算法共四种译码算法分别进行仿真比较。各算法的最大迭代次数均为200。采用BPSK调制方式,信道噪声假设为均值

(下转第 1734 页)

发的程序。测试是在两台通过 IP 交换机连接起来的 PC 机上进行的,其中一台是客户机,另一台是服务器。两机的配置大致相同:Intel Celeron 1.7GHz CPU;256M RAM;硬盘:Maxtor 60GB ATA Disk,最大数传率 40MB/s;100M 网卡;操作系统:RedHat 9.0。为进行对比,在此平台上分别测试了 SAMBA 服务器和 NRS 服务器。

5.2 测试数据及分析

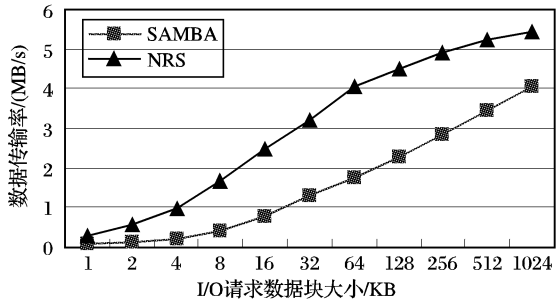

图 4 数据传输率曲线

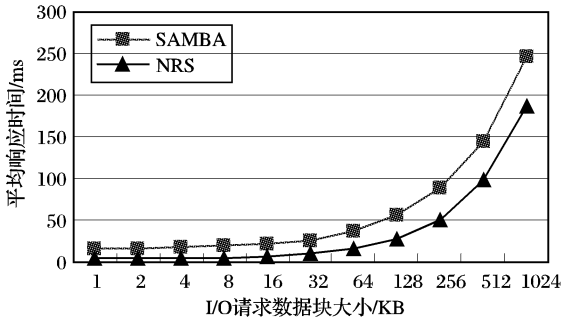

图 5 平均响应时间曲线

图 4,图 5 分别为 NRS 服务器和 SAMBA 服务器的吞吐率曲线和平均响应时间曲线。测试结果显示:随着 I/O 请求块大小的增加,两种服务器的数据传输率也都随之增大。总体上 NRS 服务器的数据传输率比 SAMBA 高 30% 左右,而 NRS 服务器平均响应时间则明显低于SAMBA。NRS服务器能够达到如此好的性能,主要是 NRS 系统采用 iSCSI 协议,iSCSI 协议采用数据块级的网络共享,要比 SAMBA 的文件级共享性能高得多,而且 NRS 系统采用磁盘文件作为缓存,采用后台刻录技术,在用户提交完成后进行后台刻录,这样既提高了共享的方便程度,也缩短了用户的刻录响应时间。这种共享方式同共享网络打印机方式一致,大大方便了用户的使用。

参考文献:

[1] GIBSON GA. Network attached storage architecture[J]. Communications of the ACM, 2000,43(11): 37 - 45.

[2] NAGLE D, GANGER G, BUTLER J, *et al.* Network Support for Network-Attached Storage[A]. Hot Interconnects'1999[C], 1999.

[3] LEE L - W, SCHEUERMANN P. File Assignment in Parallel I / O Systems with Minimal Variance of Service Time[J]. IEEE Computer, 2000,49(2):127 - 140.

[4] BODEN NJ, COHEN D, FELDERMAN RE, *et al.* Myrinet: A Gigabit-per-Second Local Area Network[J]. IEEE Micro, 1995,15(1):29 - 36.

[5] GIBSON G. Cost-effective high-bandwidth storage architecture[A]. Proceedings of ACM ASPLOS[C], 1998.

[6] SACHS M, LEFF A, SEVIGNY D. LAN and I/O Convergence: A Survey of the Issues[J]. IEEE Computer, 1994,27(12):24 - 33.

[7] ORENSTEIN G, SHURTLEFF G. Integration Scenarios for iSCSI and Fibre Channel[Z]. SNIA IP Storage Forum, 2002.

[8] Intel iSCSI protect[EB/OL]. https://sourceforge.net/projects/intel-iscsi, 2002.

(上接第 1731 页)

等于零的高斯白噪声。图 2 和图 3 分别给出了四种算法的译码性能、迭代次数比较曲线。仿真结果表明:改进算法比其他硬判决算法的性能有较大提高,明显接近 BP 算法;同时改进算法的迭代次数明显低于 BP 算法,且没有浮点和乘法运算。

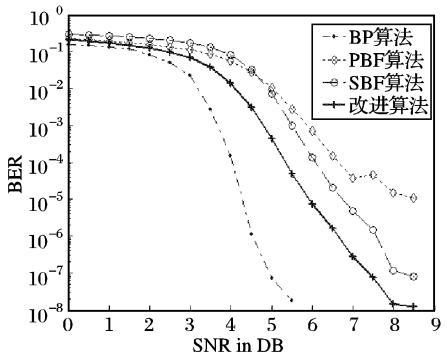

图2 (1008,504,3,6)码不同译码算法下的误码率(Bit Error Rate, BER)比较

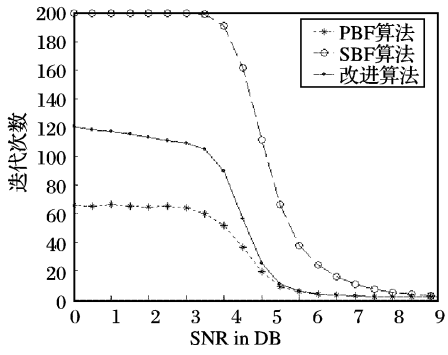

图 3 (1008,504,3,6)码不同译码算法下的迭代次数的比较

4 结语

本文结合统计译码思想,基于并行 BF 算法提出了一种新的硬判决译码的实现方案:在明显改进译码性能的同时,平均迭代次数仍保持较低水平,实现了译码性能和算法复杂性的有效均衡。由于统计译码仅仅是对 LDPC 码校验矩阵的扩充,因此其他各种改进的 BF 算法都可以与之结合,用以研究更为高效的译码方案。而且统计译码中构造校验子集的方法,既可推广用于非规则 LDPC 码,也可作为进一步优化 LDPC 编码方案的参考依据。

参考文献:

[1] GALLAGER RG. Low-Density Parity-Check Codes[D]. Cambridge, MA: MIT Press, 1963.

[2] MIADINOVIC N, FOSSORIER MPC. Improved Bit-Flipping Decoding of Low-Density Parity-Check Codes[J]. IEEE Transactions on Information Theory, 2005,51(4):1594 - 1606.

[3] ZHANG J, FOSSORIER MPC. A modified weighted bit-flipping decoding of low-density parity check code[J]. IEEE Communnication Letters, 2004,8(3):165 - 167.

[4] CHAN AM, KSCHISCHANG FR. A Simple Taboo based Soft Decision Decoding Algorithm for Expander Codes[J]. IEEE Communnication Letters, 1998,2 (7):183 - 185.

[5] KOU Y, LIN S, FOSSORIER MPC. Low density parity check codes based on finite geometries: A rediscovery and new results[J]. IEEE Transactions on Information Theory, 2001,47(7):2711 - 2736.

[6] NOUH A, BANIHASHEMI A. Bootstrap decoding of low-density parity check codes[J]. IEEE Communnication Letters,2002,6(9): 391 - 393.

[7] LIU Z, PADOS DA. A decoding algorithm for finite-geometry LDPC codes [J]. IEEE Transactions on Communications, 2005,53(3):415 - 420.

[8] Al JA. A new Class of Attacks On McEliece Public-Key and Related Cryptosystems[A]. The 2001 Canadian Workshop On Information Theory[C],2001.