文章编号: 1000-6893(2001) 06-0567-03

用重要性抽样方法估计分组码同时纠错和检错时的性能

徐大专, 许宗泽
(南京航空航天大学 信息科学与技术学院, 江苏 南京　210016)

ESTIMATING PERFORMANCE OF BLOCK CODES FOR BOTH ERROR DETECTION AND CORRECTION USING IMPORTANCE SAMPLING

XU Da-zhuan, XU Zhong-ze
(Institute of Information Science and Technology,
Nanjing University of Aeronautics and Astronautics, Nanjing　210016, China)

摘　要: 研究用重要性抽样仿真方法估计线性分组码同时纠错和检错时的不可检错误概率。从理论上得到了最佳偏置转移概率和仿真效率的表达式。对(15, 7)BCH 码和(23, 12)Golay 码的仿真表明, 重要性抽样是一种非常有效的仿真方法。

关键词: 分组码; 不可检错误概率; 重要性抽样; 计算机仿真
中图分类号: TN911. 2　　文献标识码: A

Abstract: The undetected error probability of linear block codes for both error detection and correction is estimated with importance sampling. Formulas for optimum bias transition probability and simulation efficiency are obtained theoretically. Simulation results for the (15, 7) BCH code and the (23, 12) Golay code show that the importance sampling is a very efficient method.

Key words: block code; undetected-error probability; importance sampling; computer simulation

重要性抽样(IS) 方法是一种非常有效的仿真方法[1~3], 文献[1]用 IS 方法成功地估计通信系统的误码率。文献[4]用 IS 方法估计 Viterbi 译码器的误码率, 也获得了满意的结果。文献[5]用 IS 方法估计纯检错系统的性能, 本文将该结果推广到同时纠错和检错的情况, 从理论上得到了最佳偏置转移概率和仿真效率的表达式。对(15, 7) BCH 码和(23, 12)Golay 码的仿真结果表明, IS 方法的效率很高, 仿真值与理论值非常吻合。

1　仿真模型

设(n, k)线性分组码的最小距离为 d, 译码器对汉明重量小于等于 t 的错误图样进行纠正, 对汉明重量大于 t 的错误图样进行检测, 这里 $0\le t\le[(d-1)/2]$, $[x]$表示小于等于 x 的最大整数。显然, 纯检错是 $t=0$ 的特殊情况, 纯纠错是 $t=[(d-1)/2]$的特殊情况。

系统仿真模型如图 1 所示: C 表示发送码字; E 表示错误图样; $X=C+E$ 表示接收接收向量, 其分量等于 C 和 E 对应分量的模 2 和, 它们都是 n 维线性空间 V^n 中的向量。$y=g(X)$表示同时纠检错译码器的系统函数。令 $w(X, C)$表示接收向量与码字之间的汉明距离。所有满足 $w(X, C)\le t$ 的向量 X 的集合 $S(C)=\{X\in V^n | w(X, C)\le t\}$ 称为以 C 为中心, 半径为 t 的汉明球。那么有如下 3 种情况:

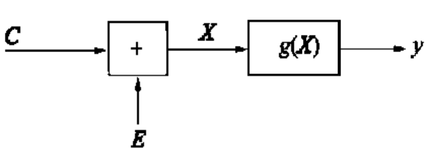

图 1　同时纠检错系统仿真模型

Fig. 1　Simulation model of correction/detection system

(1) 当 $X\in S(C)$ 时, 译码器能够正确译码, 用 $Y=0$ 表示。

(2) 当 X 不属于任何码字的汉明球时, 译码器能够检测出错误, 用 $Y=1$ 表示。

(3) 当 X 属于另一个码字的汉明球时, 出现了一个不可检测的错误, 用 $Y=2$ 表示。显然, 不可检错误的概率

$$p_{ud}=P\{Y=2\}=P\{g(X)=2\}$$

令 $8=\{X\in V^n | g(X)=2\}$表示全部不可检错误图样的集合, 设 X 的概率分布函数为 $f_X(x)$, 那么

$$p_{ud}=\sum_{X\in 8} f_X(x) \qquad (1)$$

收稿日期: 2000-11-13; 修订日期: 2001-04-25
基金项目: 航空基础科学基金(99F52043)资助项目
文章网址: http://www.hkxb.net.cn/hkxb/2001/06/0567/

引入开关函数

$$H[g(x)] = \begin{cases} 0, x \mid 8 \\ 1, x \in 8 \end{cases} \qquad (2)$$

那么

$$p_{ud} = \sum_{x \in V^n} H[g(x)] f_X(x) \qquad (3)$$

上式表明 p_{ud} 可以看成 $H[g(X)]$ 的数学期望。下面用MC 方法估计 p_{ud}, 由式(3) 得

$$\overset{d}{p}_{ud} = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} H[g(x_i)] \qquad (4)$$

其中: $\overset{d}{p}_{ud}$ 是对 p_{ud} 的估计值; x_i 是随机向量 x 的样本; N_{MC} 是样本总数。MC 方法产生的方差为[2]

$$R^2_{MC} = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} H[g(x_i)](1 - p_{ud}) f_X(x_i) \qquad (5)$$

通常 $p_{ud} n 1$, 因此

$$R^2_{MC} = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} H[g(x_i)] f_X(x_i) \qquad (6)$$

为了应用重要性抽样方法, 引入 X 的一个新的概率分布函数 $f^*_X(x)$, 称为偏置概率分布。把 p_{ud} 重写为

$$p_{ud} = \sum_{x \in V^n} H[g(x)] w_X(x) f^*_X(x) \qquad (7)$$

其中:

$$w_X(x) = f_X(x) / f^*_X(x) \qquad (8)$$

称 X 为关于样本 x 的权函数。用 IS 方法对 p_{ud} 重新进行估计

$$p_{ud} = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} H[g(x_i)] w_X(x_i) \qquad (9)$$

式中: x_i 是用偏置概率分布 $f^*_X(x)$ 产生的样本; N_{IS} 是用 IS 方法仿真的样本总数。IS 方法的方差为[2]

$$R^2_{MC} = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} H[g(x_i)] \,õ$$

$$(w_X(x_i) - p_{ud}] f_X(x) \qquad (10)$$

通常 $w_X(x) m p_{ud}$, 所以

$$R^2_{MC} = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} H[g(x_i)] w_X(x_i) f_X(x_i) \qquad (11)$$

比较式(6)和式(11)可以看出, 在样本总数一定时, 只要保证当 $x \in 8$ 时, $w_X(x) < 1$ 就有 $R^2_{MC} > R^2_{IS}$。为了便于比较, 定义

$$r = [N_{MC}/N_{IS}]\, û_{R^2_{MC} = R^2_{IS}} = \frac{\sum_{x \in V^n} H[g(x)] f_X(x)}{\sum_{x \in V^n} H[g(x)] w_X(x) f_X(x)} \qquad (12)$$

r 称为 IS 方法的效率, 它表示在相同的精度下, IS 方法比 MC 方法节约样本的倍数。

重要性抽样方法的本质是用偏置概率分布代替原来的概率分布, 使错误译码的数量大量增加, 而权函数的作用就是把增加的部分抵消掉。如何选择偏置概率分布是重要性抽样方法成功的关键。

2 最佳偏置概率分布问题

设(n, k, d)线性分组码 C 的重量分布为 A_i, $0 \leqslant i \leqslant n$, 这里 A_i 表示重量为 i 的向量的数目。令 $N_t(i, j)$ 表示重量为 j 的码字的数目, 它们与重量为 i 码字的距离小于等于 t, 显然 $i - t \leqslant j \leqslant i + t$。考虑转移概率为 p 的二元对称信道(BSC)。那么, 某个特定的重量为 j 的错误图样 X 出现的概率为

$$f_X(x) = p^j (1 - p)^{n-j} \qquad (13)$$

因此, 不可检错误的概率为

$$p_{ud} = \sum_{i=d}^{n} A_i \sum_{j=i-t}^{i+t} N_t(i, j) p^j (1 - p)^{n-j} \qquad (14)$$

通常 $p n 1$, 因此, 上式简化为

$$p_{ud} = \sum_{i=d}^{n} A_i N_t(i, i - t) p^{i-t} (1 - p)^{n-(i-t)} \approx$$

$$A_d N_t(d, d - t) p^{d-t} (1 - p)^{n-(d-t)} \qquad (15)$$

式中: $A_d N_t(d, d - t)$ 表示那些重量为 $d - t$ 错误图样的数目, 它们与全部重量为 d(最小重量) 的码字的距离小于等于 t。

在 IS 方法中, 选取如下的偏置概率分布

$$f^*_X(x) = p^j_{IS} (1 - p_{IS})^{n-j} \qquad (16)$$

这里 p_{IS} 称为偏置转移概率, j 是 x 中"1"的数目, $p < p_{IS}$。这时

$$w_X(x) = \left(\frac{p}{p_{IS}}\right)^j \left(\frac{1 - p}{1 - p_{IS}}\right)^{n-j} \qquad (17)$$

注意到式(12)的分子就是不可检错误概率, 而式(12)的分母

$$\sum_{X \in V^n} H[g(x)] w_X(x) f_X(x) = \sum_{x=8} w_X(x) f_X(x) = p_{ud} \left(\frac{p}{p_{IS}}\right)^j \left(\frac{1 - p}{1 - p_{IS}}\right)^{n-j} \qquad (18)$$

把式(15)、式(18)代入式(12)得

$$r = \left(\frac{p_{IS}}{p}\right)^{d-t} \left(\frac{1 - p_{IS}}{1 - p}\right)^{n-(d-t)} \qquad (19)$$

在 p 给定时, r 是 p_{IS} 的函数, 并且当

$$p_{IS} = (d - t) / n \qquad (20)$$

时达到极大值。上式表明, p_{IS} 只与选用的码型有关, 而与信道转移概率 p 无关。把 p_{IS} 代入式(19)得

$$r_{max} = \left(\frac{d-t}{np}\right)^{(d-t)}\left[\frac{n-(d-t)}{n(1-p)}\right]^{n-(d-t)} \tag{21}$$

3　(15, 7) BCH 码和(23, 12) Golay 码的计算机仿真

IS 方法的仿真结果如图 2 和图 3 所示。图 2 给出了(15, 7) BCH 码不可检错误概率的仿真结果。当 $t=0$ 时, 最佳偏置概率 $p_{IS}=0.2$, 样本总数 $N_{IS}=2000$; 当 $t=1$ 时, $p_{IS}=0.2667$, $N_{IS}=2000$; 。当 $t=2$ 时, $p_{IS}=0.3333$, $N_{IS}=30000$。图 3 给出了(23, 12) Golay 码不可检错误概率的仿真曲线。当 $t=3$ 时, $p_{IS}=0.1796$, $N_{IS}=2000$; 当 $t=2$ 时, $p_{IS}=0.2174$, $N_{IS}=3000$; 当 $t=1$ 时, $p_{IS}=0.2609$, $N_{IS}=4000$; 当 $t=0$ 时, $p_{IS}=0.3043$, $N_{IS}=30000$。图 2 和图 3 给出的仿真结果与理论计算非常吻合。

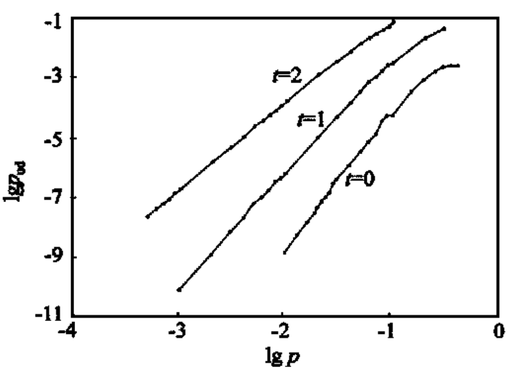

图 2　(15, 7)BCH 码仿真结果

Fig. 2　The undetected error probability of the (15, 7) BCH code

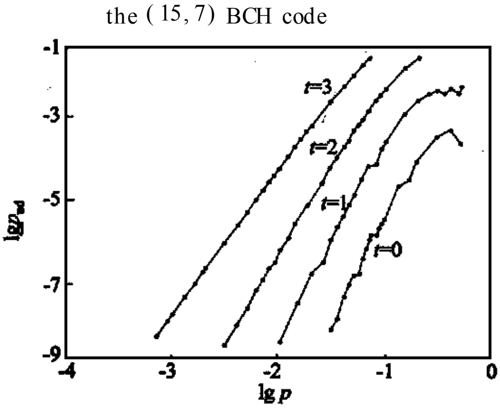

图 3　(23, 12)Golay 码仿真结果

Fig. 3　The undetected error probability of the (23, 12) Golay code

4　结　论

用重要性抽样方法估计线性码在同时纠错和检错时的不可检错误概率, 从理论上得到了最佳偏置概率分布和仿真效率的分析表达式。对(15, 7) BCH 码和(23, 12) Golay 码的计算机仿真表明, 仿真结果和实际情况非常吻合。本文的研究结果对重要性抽样方法在差错控制系统仿真中的应有重要意义。

参　考　文　献

[1]　Shammugan K S, Balaban P. A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems[J], IEEE Trans Commun, 1980, 28(11), 1916- 1924.

[2]　Hahn P M, Jeruchim M C. Developments in the theory and application of importance sampling[J]. IEEE Trans. Commun, 1987, 35(7), 706- 714.

[3]　Jeruchim M C. Techniques for estimating the bit error rate in the simulation of digital communications[J]. IEEE Journal on Selected Areas in Communications, 1984, 2(1): 153 - 170.

[4]　Herro M A, Nowack J M. Simulated Viterbi digital decoding using importance sampling[J]. IEE Proceedings F, 1998, 135(2): 133- 142.

[5]　徐大专. 用重要性抽样方法估计分组码的未检概率[J]. 通信学报, 1992, 13(1): 22- 27.

作者简介:

徐大专(1963-)　男, 江苏滨海人, 南京航空航天大学信息科学与技术学院教授, 博士, 博士生导师, 主要从事信息与编码理论、软件无线电和虚拟仪器方面的教学与科研工作。Email: xdzh@jlonline.com

许宗泽(1940-)　男, 江苏常州人, 南京航空航天大学信息科学与技术学院教授, 博士生导师, 主要从事扩频通信和移动通信系统方面的教学和科研工作。

(责任编辑: 俞　敏)