邮电高等学校教材

纠错编码技术

陈宗杰 左孝彪 编

人民邮电出版社

邮电高等学校教材

纠错编码技术

陈宗杰
左孝彪 编

人民邮电出版社

内 容 提 要

本书是邮电高校无线电通信专业统编教材。全书共十章，内容包括：纠错编码的基本概念；近世代数基础；线性分组码；循环码；汉明码和戈莱码；BCH码；纠突发错误码；卷积码及其概率译码；以及差错与控制系统的应用与设计等。

本书兼顾理论性和实用性，内容自成体系，由浅入深。可作高等院校电子与通信类专业高年级教材，也可供从事数字通信、数据传输、计算机网络等科研和工程技术人员参考。

邮电高等学校教材

纠错编码技术

陈宗杰
左孝彪 编

人民邮电出版社出版
北京东长安街27号
广益印刷厂印刷
新华书店北京发行所发行
各地新华书店经售

开本：850×1168 1/32 1987年11月 第一版
印张：$13^{8}/_{32}$ 页数：212 1987年11月北京第1次印刷
字数：348千字 印数：1—3,500册

ISBN 7115-03494-X/Z

定价：2.55元

前　言

本书系邮电高等院校无线电通信专业统编教材。学习本书要求的先修课有：线性代数，脉冲与数字电路，通信原理等有关课程。

本书共十章，第二章为学习本课程提供必要的数学基础；第一章概述了纠错码的基本知识；第三、四、五、六章讨论了最基本的纠随机错的线性分组码及其子类循环码、汉明码、戈莱码以及 BCH 码等；第七章介绍纠突发错误码；第八、九两章讨论了卷积码及其概率译码；第十章简要讨论差错控制系统在工程中的应用与设计问题。

本书经邮电高校教材无线电技术编审委员会讨论通过，推荐作为统编教材出版。

本书是在原纠错编码技术讲义基础上重新改写而成的。原讲义第一至六章由左孝彪编写，七至十章及前面部分内容由陈宗杰编写。本书由陈宗杰统编，并对全书作了增删和润色，左孝彪提供了许多修改意见。本书各章所配习题由左孝彪、陈宗杰分别进行演算，并给出了部分习题参考答案。

本书适用 54 至 60 学时，书中凡有“＊＊”处可不作基本教学内容；第二章也可根据教学时数作弹性处理；当课时在 40 学时以内时，可讲更基本的内容，有“＊”处可不讲。

参加本书编审活动的院校有：清华大学、北京邮电学院、南京邮电学院、重庆邮电学院、西安邮电学院、长春邮电学院等，在此，我们一併表示衷心的谢意。

由于编者水平有限，加之时间仓促，书中定有欠妥之处，热忱欢迎读者给予批评指正。

编者 1986 年 4 月

目　录

第一章 绪 论

引 言

数字通信、数据传输、图象传输、计算机网等数字信息交换和传输中所遇到的最主要的问题是可靠性问题，也就是数字信号在交换和传输过程中出现差错的问题。不同用途的用户对可靠性的要求是很不相同的。例如，对于普通的电报，差错概率在 10^{-3} 时是可以接受的；而对导弹运行轨道数据的传输，过高的差错率将使导弹偏离预定的轨道，这显然是不允许的。使数字信号在传输过程中产生不同差错率的主要原因，是不同传输系统的性能及其在传输过程中受到的不同干扰。因此，要从多种途径来研究提高系统可靠性的方法。首先，要合理地选择系统和调制解调方式，这些内容不属本书研究范围。在上述条件已确定的情况下如何提高可靠性，降低错误概率属于信道编码的范围，也就是本书所要研究和解决的问题。所谓信道编码，概括地说就是在一个具有确定长度的数字信息序列 m 后面，人为地按一定规则加进非信息数字序列，从而构成了一个一一个码字 C（信道编码），然后将此码字 C 经调制器变换为适合信道传输的信号。经信道传输后，在接收端经解调器判决输出的数字序列称为接收序列 R。由于信道中干扰的存在，使接收序列 R 可能与所发码字 C 不相同。二者都是有序数列，序列中每一位数字都是一一对应的。如果信道平静（无干扰），R 与 C 中每一个对应位上的数字都是相同的；在有干扰时，有的对应位上的数字是相同的，有的则

不相同。例如，发送码字为 $C=(1011100101)$，接收序列为 $R=(1010100111)$，则在第 4 位和第 9 位发生错误。接收序列 R 经信道译码器译码后输出信息序列为 $\bar{m}$，此 $\bar{m}$ 为发送信息序列 m 的估值。由于编码时所加的多余度只是供译码器纠（检）错用的，它并不是用户需要的，所以一般不再输出。信源译码器将 $\bar{m}$ 变换成用户需要的信息形式 $\bar{S}$ 并送至用户。$\bar{S}$ 为信源输出的 S 的估值。这一数字信号传输过程可用图 1-1 来表示。

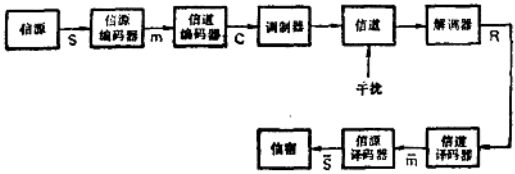

图 1-1　数字信息传输系统方框图

从广义的信息传输来说，信源可以是语言、图象或各种数据的电信号 S，从波形上看可以是连续波形，也可以是离散波形，经过信源编码器输出为二进（0 或 1）的信息序列 m，对于信道编码器来说，输入的是二进信息序列 m，输出为二进的码字 C；信道是传输经过调制器处理后的含有有用信息的信号的媒质，它可以是电话线、电缆线、各类高频无线电线路，以及计算机的存贮系统和磁带、磁盘装置等。信道常受各种自然的或人为的干扰，例如，在各类无线电线路上，干扰可能有热噪声、闪电干扰、以及各种工业用电干扰，而磁带或磁盘的损伤对所存贮的信息来说将造成差错，所以也可视为干扰。调制器是将包含信息 m 的码字 C 变换成适合信道传输的各类信号，由于这不属于本书所研究的范围，其内容不再叙述。

由上述可知，利用信道编码——译码器，可以显著地改善信息在传输过程中的错误概率指标，有效地提高系统抗干扰能力，也就是提高了系统可靠性，这是数字信号传输中的最重要的课题之一。

在编码过程中，每个码字所加进的非信息序列是供在译码器中检出或纠正错误用的，常称为监督元。它们本身不是信息，单纯从信息传输的角度来说是多余的，这种多余度降低了信息传输的效率。但是从另一角度来说，所增加的多余度换来的是使码字获得了纠错能力，提高了信息传输的可靠性。一般地说，码字中多余度越高，纠错能力越强，可靠性越高。由此可见，可靠性是以降低效率、即有效性为代价换来的。编码的问题就是在一定的抗干扰能力要求下，合理设计和选择多余度最小的码字问题，也就是如何使可靠性和有效性二者能够得以合理兼顾的问题。

从理论上说，纠错编码是建立在近世代数学——近世代数基础上的，它又是信息论的一个重要分支，从五十年代开始至今发展很快，许多内容已经建立起完整的、严密的理论体系。随着数字技术的迅速发展，大规模、超大规模集成电路的不断出现，为纠错编码技术的应用开辟了无限广阔的前景；目前，适应不同需要的多种类型的码字不断出现，使这门学科呈现了无限生机。

本书为电子与通信系统专业本科高年级学生教学用书，因此不可能全面地介绍编码理论和技术，在有限的内容中，我们只打算给读者提供有关纠错编码的基本理论和技术，并且选择近年来常见的几种码型进行分析，以使本书内容兼顾理论性和实用性。

本书主要讨论实践中最常用的二元编码。为此，首先要了解二元信道对信号传输的影响，而二元对称信道是实际中常用信道的一种较简单模型，因此我们先讨论这一信道。

§ 1-1 二元对称信道

上面已经对二元数字信号在系统中的传输过程作了描述，前面提及的信道可以称为狭义信道。本书研究的对象是系统中的编码器和译码器，而调制器的输入是二元序列（码字C），解调器的输出也

是二元序列（接收序列R），因此，调制器——狭义信道——解调器可以看成一个整体，称为广义信道，本书在无特别必要的情况下，统称信道，此信道输入是码字C，输出为接收序列R。同样，此信道由于干扰的存在，可能使传输码字C的不同位置（即码字中的元素——码元）发生差错，导致R与C可能不同。干扰的形式是各式各样的，且是随机的。但从其造成的效果来看，无非是使在信道中传输的码元1变为0，或0变为1。为了简化分析，我们可以将干扰在信道中的作用，同样用二元数字序列来表示，称为错型或错误图样E，如码字C中包含n个码元，即$C=(C_{n-1}, C_{n-2}, \cdots C_2, C_1, C_0)$；接收序列$R$一定也含有$n$个元素，同样可写成$R=(r_{n-1}, r_{n-2}, \cdots, r_2, r_1, r_0)$；相应地，错型$E$也包含$n$个元素，即$E=(e_{n-1}, e_{n-2}, \cdots, e_2, e_1, e_0)$。错型$E$对码字$C$的作用是相同位上的元素按模2相加（见第二章），即$1+0=1$，$0+1=1$，$0+0=0$，$1+1=0$。$C$、$R$、$E$三者关系可以写成$R=C+E$；或$(r_{n-1}, r_{n-2}, \cdots, r_2, r_1, r_0)=(C_{n-1}+e_{n-1}, C_{n-2}+e_{n-2}\cdots, C_2+e_2, C_1+e_1, C_0+e_0)$。在传输中没有错误的码元$C_i$所对应的$i$位上的错型元素$e_i=0$，即$r_i=C_i+e_i=C_i+0=C_i$；如果第$j$位的码元$C_j$传输后错了，则第$j$位上的$e_j=1$，则有$r_j=C_j+e_j=C_j+1=\bar{C}_j$。我们还是用上面的例子来说明。$C=(1011100101)$，$R=(1010100111)$，在第4位和第9位发生错误，此时的错型为$E=(0001000010)$。它们的关系为：$R=(1010100111)=(1+0, 0+0, 1+0, 1+1, 1+0, 0+0, 0+0, 1+0, 0+1, 1+0,)=(10\bar{1}1001\bar{0}1)=(1010100111)$。根据以上分析，数字信号传输系统可简化为图1-2所示。

具有下述性质的信道称为二元对称信道，即在编码器输出的码元为“0”时，经此信道传至译码器的输入端为“0”的概率为q，为“1”的概率为p；显然$q+p=1$；当所发为“1”时，译码器输入端为“1”的概率这时也为q，为“0”的概率也为p。

假设信道对每个码元在传输过程中的影响是独立的，上述性质

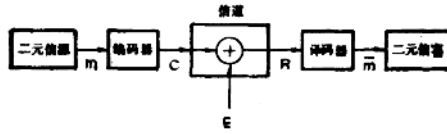

图 1-2　简化的传输系统

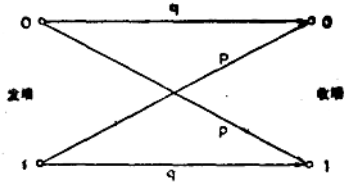

图 1-3　二元对称信道

可用图 1-3 表示。

在一般情况下，$q>\frac{1}{2}$。p 是描写传输发生错误的概率，称为转移概率。转移概率是一个条件概率。码字 C 第 i 个码元 C_i 取值为 0 或 1，接收序列 R 中第 i 个元素的取值也为 0 或 1。且有 $0\leqslant i\leqslant n-1$。此时的条件概率为：

$$\left.\begin{aligned}&P(r_i=0/C_i=0)=q\\&P(r_i=1/C_i=0)=p\\&P(r_i=1/C_i=1)=q\\&P(r_i=0/C_i=1)=p\end{aligned}\right\}\qquad(1\text{-}1)$$

所以有

$$P(r_i=0/C_i=0)=P(r_i=1/C_i=1)=q$$
$$P(r_i=1/C_i=0)=P(r_i=0/C_i=1)=p$$

式 (1-1)描述了二元对称信道对一单独码元的影响，其影响与所传码元的内容无关，且这种影响对各个码元是独立的。

对信道的实际传输效果我们经常关心的是码字的效果，即一个

码字在传输中出错的概率描述。用条件联合概率就可以描述码字在信道传输的情况。

若码字C中有n个码元，则接收端收到某一序列R的概率为：

$$P(R/C)=\prod_{i=0}^{n-1}P(r_i/C_i) \tag{1-2}$$

式中C表示发端所发的一个特定码字，而接收端收到序列具有多种可能性。其中收到某一特定序列R的概率为$P(R/C)$。式(1-2)描述了二元对称信道的性质。

在由该R与所发的C比较后知道R中含有d_e个错误码元，则式(1-2)右边的乘积中有d_e个p和$(n-d_e)$个q，即：

$$P(R/C)=q^{n-d_e}\cdot p^{d_e}$$

考虑到$p+q=1$，

$$P(R/C)=q^n\left(\frac{1}{q}-1\right)^{d_e} \tag{1-3}$$

当$q>\frac{1}{2}$时，有$0\leqslant\left(\frac{1}{q}-1\right)<1$。式(1-3)所表示的特定接收序列$R$出现的概率与错误码元数$d_e$的增加呈单调下降的关系。

从信道传输的角度来说，人们所关心的是所传输码长为n(含有n个码元)的码字，其中含有d_e个错误码元的接收序列R的概率。这样的R概率为

$$P_{d_e}(R/C)=\binom{n}{d_e}q^{n-d_e}\cdot p^{d_e} \tag{1-4}$$

我们应当注意式(1-4)与式(1-3)的区别。式(1-3)中的R是接收端含有d_e个错误码元的某一特定序列，而式(1-4)中的R是指接收端含有d_e个码元错误的全部接收序列，在n个码元中，尽管有错的码元数是d_e，但其分布各不相同，一共有$\binom{n}{d_e}$种。也就是这样的接收序列R有$\binom{n}{d_e}$个。由于信道对码元干扰的独立性，故信道对各码字的影响也是独立的。加之，又由于信道对码元的影响呈对称性(错码率与码元内容无关)，可以得出这样的结论，即$P_{d_e}(R/C)$与C的内容无关。码字经过信道传输，含有d_e个错误

的接收序列R的概率就反映了这个信道的性质。对二元对称信道来说，其概率为：

$$P_{d_e}(R)=\sum_{L=1}^{M}P(C_L)P_{d_e}(R/C_L)$$

$$=P_{d_e}(R/C_L)\sum_{L=1}^{M}P(C_L)=P_{d_e}(R/C_L)$$

$$=\binom{n}{d_e}q^{n-d_e}\cdot p^{d_e} \tag{1-5}$$

后面将会知道，一种形式的编码所包含长度相同（n）的码字数是一定的（暂设为M个）。这M个码字集合记为$[C]$。C_L为集合$[C]$中的一个码字。可以认为，$[C]$中各个码字出现是等概率的，则码字C_L的概率$P(C_L)=\frac{1}{M}$，因而有

$$\sum_{L=1}^{M}P(C_L)=1$$

式（1-5）所描述的$P_{d_e}(R)$对d_e分布曲线如图1-4所示。

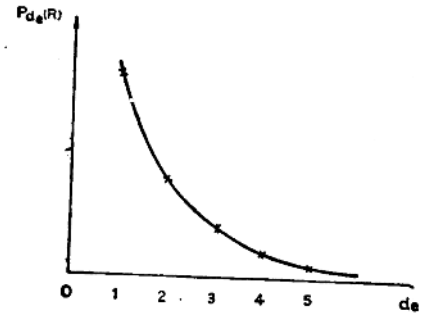

图 1-4 $P_{d_e}(R)$的分布曲线

由图可见，$R_{d_e}(R)$随d_e的增加而递减。这里我们还要说一下信道误码率P_e。设发送码字C，接收序列R。R与C相对照，当

$d_e=0$，表示无错；而 d_e 由 1 至 n 之间为任一值时，都说明码字传输时发生错误。因此，码字在信道中传输的错误概率，即误码率为：

$$P_e=\sum_{d_e=1}^{n}P_{d_e}(R)=\sum_{d_e=1}^{n}\binom{n}{d_e}q^{n-d_e}pd_e$$
$$=(q+p)^n-q^n=1-q^n \qquad (1\text{-}6)$$

在一般情况下，p 是很小的数，$d_e=1$ 和 $d_e=2$ 的概率 $P_1(R)$ 和 $P_2(R)$ 将占 P_e 中的很大部分。因此，降低信道误码率 P_e 首先关心的是对 d_e 较小数的纠正。

例如：一信道的 $q=0.999$，$p=0.001$。所传输的码字长度为 $n=15$，当 $d_e=1$ 时，$P_{d_e}(R)=P_1(R)=14.8\times10^{-3}$；当 $d_e=2$ 时，$P_2(R)=103.6\times10^{-6}$；而 $P_e=1-q^n=1-(0.999)^{15}\approx14.895\times10^{-3}$。

由此可见，总误码率 P_e 与 $d_e=1$ 的接收序列概率十分接近。

§ 1-2 信道编码的概念

设信源编码器输出的二元数字信息序列为(001010110001…)，序列中每一个数字都是一个信息元素。为了适应信道的最佳传输和进行编码，首先需要对信息序列进行分组。一般是以截取同等长度方式分组，每组长度为 k，即含有 k 个信息元。例如，对上面的信息序列以 $k=2$ 分组为：(0 0)、(1 0)、(1 0)、(1 1)、(00)、(01)、…。如果将这样的信息组直接送入信道传输，它们是没有抗干扰性能的。因为任一信息组中任一元素出错，都会成为另一个信息组。例如信息组(0 0)有一位错，变为(1 0)或(0 1)，而整个信息组也含有这两个。因此，在接收端无法判别传输是否有错。上述情况与分组长度 k 的大小无关。这就是说，不管 k 的大小如何，所分出的信息组是无抗干扰能力的，这是因为信源序列通常是足够