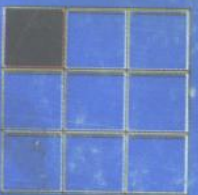


ARQ

编 码 通 信

王仲文 编著



机械工业出版社

ARQ 编码通信

王仲文 编著



机 械 工 业 出 版 社

本书专门研究数据 ARQ 通信，特别是计算机网络中的差错控制问题。其内容反映了最新科研成果。在叙述方法上，本着实用简洁的原则，尽量采用实例加以说明，使读者容易理解从而得到启发。

全书共分七章，前五章研究编码问题，后两章研究 ARQ 通信系统的有关问题。本书可供从事通信、遥控遥测、计算机通信的技术人员使用，特别适合有关专业的研究生和大专院校学生使用。

ARQ 编码通信

王仲文 编著

责任编辑：贡克勤 版式设计：胡金瑛
封面设计：李兴民 责任校对：张颖

机械工业出版社出版（北京阜成门外百万庄南街一号）

（北京市书刊出版业营业许可证出字第 117 号）

机械电子部自动化所印刷厂印刷

新华书店北京发行所发行·新华书店经售

开本 787×1092 1/16·印张 13¼·字数 315 千字
1991 年 7 月北京第一版·1991 年 7 月北京第一次印刷

印数 0,001—3500·定价：8.20 元

ISBN7-111-02814-7/TN·50 (x)

前 言

数据的传输已深入到人们生活的各个领域，电子计算机的普遍应用，要求数据传输提供前所未有的服务。可以说，凡是有计算机的地方，就有数据的传输服务，这也是计算机具有的通信能力所做出的贡献。从这个意义上说，一个计算机系统，就是一个数据传输系统。显然，用户对计算机在数据处理、交换和存贮过程中的数据传输的可靠性提出了越来越高的要求。因此，如何保证远程数据通信、卫星通信以及计算机网络通信中数据传输的可靠性就成为迫切需要解决的课题了。

所谓数据传输的可靠性，是指当数据在信道上传输时，对于噪声干扰所造成的错误，接收端能够发现或自动纠正的性能。这种能够发现错误甚至能纠正错误的数据传输系统被称为差错控制系统。完成差错控制的主要手段之一是，对所传信息（数据）进行编码，即利用代数方法，给信息附加一些保护数据，使这些保护数据和所传输数据之间建立起一些相互约束关系，从而完成编码。另一方面，为了更有效地进行差错控制，在纠错编码的基础上，利用正反两个方向信道，收端把信息收到正确与否的情况及时通知发端确定是否需要重发。这样就构成了收发双方的自动应答关系，这样的系统就称做 ARQ 差错控制系统。

近来，由于编码理论的进展，大规模集成电路和计算机的飞速发展，使得差错控制编码和 ARQ 差错控制技术在通信和计算机系统（特别是计算机网络系统）得到广泛应用。这是因为现今提供的差错控制系统，成本不再那么昂贵，技术的实现不是那么过于复杂，特别是许多环节都可以在计算机上或者是在网络的规程协议中用软件模块完成；从性能价格比上看，由于采用差错控制技术，使系统的可靠性和效率都有明显的改善，而这种改善就相当于提高发端的传输功率，所以是很有实际应用价值的。这就是说，这一领域是很有前途的。目前，西方发达国家不但在宇宙通信、卫星通信和计算机网络等高级通信系统使用了差错控制技术，而且在一般通信传输系统中也广泛采用差错控制技术，且许多好的差错控制方案都列为国家标准和国际标准而加以推广应用，例如大家熟知的 CCITT 建议在 HDLC 中应用 CRC-16 就是一例。

本书是根据作者多年从事这方面的教学和科学研究所积累的资料写成的，在《工程遥控遥测系统》（82 年第一版，曾获兵工系统优秀图书三等奖）一书第三篇基础上，重新改写而成的。其主要目的是向广大计算机和通信工作者介绍这门技术。其叙述方法力求实用简洁。

全书共分七章，第一章差错控制和编码的基本概念，主要给出本书的必要入门知识和概念；第二章线性分组码，是下一章的基础；第三章循环码，是本书的重点，因在实际中应用最广的就是循环码；第四章 BCH，是循环码中纠随机错误能力很强的一种码，本章给出 BCH 定义、性质和编译原理及方法；第五章卷积码，是另一类有别于分组码的纠随机错误和突发错误的码，它也是一种很有实用价值的码；第六章 ARQ 差错控制系统，是本书所要介绍的系统章节，首先给出三种基本型 ARQ 系统，接着又给出几种改进型 ARQ 系统，最后介绍了半速率可逆码和详细地分析了 Stop-and-Wait ARQ 系统性能；第七章介绍具有 DCM 码的混合记忆型 ARQ 系统，给出一种 DCM 码，详细地讨论了 DCM 码的性质，给出新的码组合方法，以及

其编译码方法。最后介绍了系统的组成和性能分析，证明该系统效率是目前最高的。

由于作者水平有限，这方面经验不足，不当之处在所难免，欢迎批评指正。

作者由衷地感谢北京理工大学俞宝传教授、北京航空航天大学范仁周教授、燕山大学郑绳榷副教授、北京市遥测技术研究所刘瑞复高级工程师。感谢他们详细审阅、修改了本书全稿，并提出许多宝贵意见。原子能科学研究院王海成高级工程师和解放军第 61 研究所于康友同志在 ARQ 通信系统性能分析章节中，方勇同志在编码算法章节中都做了大量工作。王东同志担负全书绘制图表和计算机程序编写、调试工作。在当前学术著作出版难的情况下，李兴民和贡克勤同志为本书创造了出版条件，在此向他们表示感谢。

作 者

1990.8

北 京

目 录

绪 论	1
第一章 差错控制和编码的基本概念	5
§ 1-1 数字传输系统的质量要求	5
§ 1-2 纠错和检错原理	8
一、基本概念	8
二、差错的检出及纠正原理	9
三、编码的几何表示	11
四、纠单一错误的编码	12
五、多重错误的纠正及检出原理	12
六、差错概率	13
七、最大后验概率判定法	14
八、最大似然判定法	16
§ 1-3 常用检错码举例	18
一、重复码	18
二、奇偶校验码	19
三、恒比码	21
四、方阵码	22
第二章 线性分组码	23
§ 2-1 代数基础知识	23
一、有限域(伽罗华域)	23
二、有限域(伽罗华域)运算	23
§ 2-2 线性分组码的概念	34
§ 2-3 伴随式(校验子)的概念	38
§ 2-4 线性码的纠错能力	40
§ 2-5 汉明码	43
第三章 循环码	49
§ 3-1 二元码的多项式表示	49
一、码多项式和生成多项式	49
二、循环码与线性码之间的关系	50
§ 3-2 循环码的定义及其性质	51
§ 3-3 循环码的编码方法	53
§ 3-4 循环码的硬件编码实现	57
一、编码和译码器的基本电路	57
二、循环码的 $n-k$ 级编码器	60

三、循环码的 k 级移位编码器	65
§ 3-5 循环码的软件编码实现	69
一、软件算法 I	69
二、软件算法 II	71
§ 3-6 循环码的译码	74
一、循环码的伴随式计算	74
二、循环码的捕错译码	77
三、改进的捕错译码	79
§ 3-7 循环码的大数逻辑译码	81
一、大数逻辑译码法的原理	82
二、大数逻辑译码器	90
三、一步大数逻辑可译码——最大长度码	97
四、差集码	99
五、 L 步大数逻辑译码	102
第四章 BCH 码	107
§ 4-1 简单的 BCH 码例	107
§ 4-2 关于 BCH 码	109
一、BCH 码的定义	109
二、BCH 码的最小距离	109
§ 4-3 BCH 码的译码	116
§ 4-4 非二元 BCH 码的译码	125
第五章 卷积码	130
§ 5-1 卷积码的结构特点	130
§ 5-2 卷积码的编码	135
§ 5-3 卷积码的大数逻辑译码	139
第六章 ARQ 差错控制系统	143
§ 6-1 概述	143
§ 6-2 三种基本型 ARQ 通信系统	144
一、Stop-and-Wait ARQ 通信系统	144
二、Go-back- n ARQ 通信系统	145
三、选择重传 (Selective-Retransmission) ARQ 通信系统	146
§ 6-3 具有有限接收存贮器的选择重传 ARQ 通信系统	147
§ 6-4 混合型 ARQ 通信系统	147
§ 6-5 半速率可逆码	149
§ 6-6 Stop-and-Wait ARQ 通信系统的性能分析	152
一、引言	152
二、Stop-and-Wait ARQ 系统的性能分析	152
第七章 混合记忆型 ARQ 通信系统	166
§ 7-1 DCM 码及其性质	166

§ 7-2 一种新型的码组合方法	173
§ 7-3 DCM 码的编译码	175
一、DCM 码的编译	175
二、DCM 码的译码	175
§ 7-4 具有 DCM 码的混合记忆型 ARQ 通信系统	181
一、混合记忆型 ARQ 通信系统的原理	181
二、混合记忆型 ARQ 通信系统的性能	183
附录 用独立多项表示码的 μd	189
参考文献	199

绪 论

一提到通信, 自然地被理解为传统的远程数据通信。其实除了传统的通信外, 由于计算机具有很强的数据处理能力, 数据在计算机各部分的传递本身就是通信。无论远程数据通信、传真通信, 还是计算机网络通信, 都对数据传输的可靠性提出更高的要求, 例如计算机局域网要求误码率低于 10^{-9} 。

提高数据传输可靠性, 有两种途径: 或者增加发送信号的功率, 提高接收端的信号噪声比; 或者采用编码方法对信道差错进行控制。前者常常受条件限制, 不是所有情况都能采用。后者是建立在 Shannon 理论基础上的, 近几十年发展起来差错控制编码技术。编码和提高信号发送功率都能在接收端有效地抑制噪声信号, 而有效地在噪声中恢复出有用信号。从这个意义上说, 差错控制编码中的编码和提高发送信号功率是等效的。

所谓差错控制编码, 就是对所传输的信号加以改造, 使信号内部结构具有更强的规律性和相互关联性, 以致在信道上当噪声破坏了信号的部分结构时, 仍能根据信号原有规律性和相互关系来发现错误, 纠正错误, 恢复原有信息。

编码一般方法是按某种规律, 对原来的信息组添加一些新码元, 也叫监督元。监督码元的作用是监督该码组中传输过程中是否发生了错误, 并指出是哪位或哪些位码元发生了错误, 以便纠正。附加监督码元的方法(规律或规则)不同, 就形成不同的编码方法。无论按什么规则来编码, 总要在原码组中加进若干监督码元。这就说明, 纠错编码码组抗干扰能力的提高是以牺牲部分传输效率为代价的。Shannon 编码定理指出: 当码组足够长时总能找到一种编码方法, 按照这种编码方法, 监督码元数与总码元数之比趋于任意小。这是理想的编码结果, 可惜迄今所有已经找到的编码方法距理想编码结果还相差甚远, 因此差错控制编码还在不断发展。大体上, 现在的差错控制编码都属于代数编码, 就是说, 码组中的信息码元和监督码元之间按一定的代数关系相互约束。从结构上讲主要有分组码和卷积码两大类。

把编码技术应用到通信中去, 对信道中噪声产生的差错进行控制, 组成一个差错控制通信系统。目前利用的都是 ARQ 技术(自动重复请求技术), FEC 技术(前向纠错技术)虽然效率较高, 但它的可靠性远不如 ARQ 技术。此外, FEC 技术对信道的适应能力较差, 无论信道优劣, 它的效率都是一个恒定值。FEC 技术只适用于没有反向信道或不适合建立反向信道的系统。ARQ 技术的应用范围极为广泛, 目前它已应用在电报系统、电传系统、传真通信系统、计算机通信网、卫星通信、计算机存贮系统、遥测遥控系统等。ARQ 技术之所以得到广泛的应用是由于它有一系列优点。首先它实现简单、可靠性高, 对信道具有很广的适应能力。ARQ 系统的主要缺点是在信道干扰较强时效率降低。所以对 ARQ 技术的研究主要集中在如何提高效率上。近十几年人们提出了许多方案, 着眼点都在于提高它的效率。下面简单介绍几种有代表性的方案。

1975 年 Sastry^[1]提出了一种方案, 它对基本的 Go-back-N 方案做了一些修正, 使浪费在重传码组上的时间得以减少, 其效率如下:

$$\eta'_c = \eta \frac{1}{1+p(E+K(N-1))}$$

式中 η_c 为码速率; P 为码组错误概率; K 为方案中定义的参数; N 为往返延迟; E 为正确接受一个码组所需的平均传送次数, $E=1/(1-p)$ 。

Go-back-N 方案的效率为:

$$\eta_c = \eta \frac{1}{1+NpE}$$

所以, 当 $1+Np(1-p)^{-1} > 1+p(1-p)^{-1}+Kp(N-1)$, 也即 $p > 1-1/K = \frac{1}{2}$ ($K=2$) 时, Sastry 方案要比 Go-back-N 方案的效率高。

一般的信道在绝大多数情况下 $p < 1/2$, 所以这时 Sastry 方案要比 Go-back-N 效率低。只有当 $p > 1/2$ 时 (即信道很差时), Sastry 方案才优于 Go-back-N, 说明 Sastry 方案有一定的局限性。Morris^[12] 对 Sastry 方案做了一些修正, 使得它更加减少了浪费在多余重传码组中的时间, 效率有了进一步提高 (在 p 较大情况下), 在 p 较小时, 它与 Sastry 方案基本是相同的。

1979 年, Towsley^[13] 提出了一个 “Idealized ARQ” 作为一种参考标准, 利用排队论分析系统, 提出了排队长度和等待时间两项指标来描述系统性能。这种描述和用效率描述实质是一样的, 排队时间长, 排队长度大, 相当于效率低; 反之, 效率高。Towsley 还提出了一种 “Stutter ARQ” 方案, 它实际上也是对 Go-back-N 方案的一种修正。对于中、低误码率情况下, “Stutter ARQ” 方案和 “Idealized ARQ” 方案几乎是一样的。在高误码率及往返延迟大的情况下, Stutter ARQ 方案比标准的 Go-back-N 方案要平均排队长度与平均延迟时间 (等待) 时间上有所改善。

随着大规模集成电路及计算机应用的不断发展, 为智能化 ARQ 方案的实现提供了可能。1982 年 Weldon^[14] 提出了一种改进的 “Sw-ARQ” 与 “Go-back-N ARQ” 方案, 它在效率上明显地要比 “Sw-ARQ” 与 “Go-back-N ARQ” 方案效率高。主要思想是使系统对信道有自适应能力, 信道好时, 发送的次数少 (速度高), 信道差时, 发送次数多 (速度低)。Sw-ARQ 方案工作过程分 q 步进行。第 0 步, 发端收到正确接收认可信号 ACK, 继续发送新码组; 第 1 步, 如果发端收到不能正确接收的否定信号 NAK, 那么将相应的码组重传 n_1 次; 第 2 步, 如果第一步重传的码组全部出错, 那么将相应的码组再重传 n_2 次; 这样一直进行下去……; 第 q 步, 如果第 $q-1$ 次重传的 $q-1$ 个码组合全部出差错, 那么缓存器发生溢出, 码组再被重传 n_q 次; 第 $q+1$ 步, 如果第 q 步重传的 n_q 个码组都出错, 那么缓存器溢出, 码组被再次重传 n_q 次, 一直保持这个数字重传下去, 直到正确接收为止。分析表明, 该方案效率有所提高, 且它更具有一般性。当 $q=0$, $n_1=1$ 时, 该方案就退化为 Go-back-N 方案。事实上, 该方案效率的提高是有限的, 且由于每次重发码组数是变化的, 故它对控制设备提出更高的要求。

怎样进一步提高效率呢? 我们知道 FEC 之所以效率高是因为纠错靠纠错编码, 而不进行重传。如果把 FEC 技术和 ARQ 技术结合起来, 效率会进一步提高。1982 年, Shu Lin 与 Philip S. Yu 首次将 FEC 和 ARQ 技术结合起来, 提出了一种混合型 ARQ 方案^[15]。它的主要设计思想是, 用于纠错的一致校验位只在需要纠错时才发送出去。一般情况下, 只发送具有一定检错能力的码组。当发送的一个信息组 D 被检出错时, 重传不是 D 的简单重复, 而是 D 的一致校验位 $p(D)$ 与 D 交替地重复。 $p(D)$ 是基于 D 及一个能纠正不多于 t 个错误同时能检出不多于 d ($d > t$) 个错误的半速率可逆码而形成的。收端收到 $p(D)$ 后, 利用可逆性质从 $p(D)$

中直接恢复出 D 或者利用 $(D, p(D))$ 进行纠错而恢复 D 。该方案的效率明显优于上述方案,原因是 FEC 技术和 ARQ 技术相结合,相应地克服了各自缺点,而保留了各自的优点。

此方案被称做混合 I 型方案,虽然它的效率较高,但在利用码组方面不够充分,即没有充分利用半速率可逆码的纠错能力。另外,在组码方式上不够灵活。针对上述缺点,1984 年 Yu-ming-Wang 与 Shu Lin^[16]对混合 I 型方案又进行了修正。修正方案与原始方案相比,在传输性能方面略有改善,然而在利用码的纠错能力方面则更灵活了。修正方案更具有—般性,文献^[16]中证明了纯选择重传 ARQ 方案只是修正方案的一个特例。

在一些对可靠性要求较高的系统中(如武器控制、火箭控制、卫星控制),对效率的要求变为次要的指标,而对可靠性要求变成重要的指标。1986 年 Tadao Kasami, Tohiro Fujiwara 和 Shu Lin 提出了一种级联码 ARQ 方案^[17],它的可靠性极高,已经被推荐用于美国 NASA 遥控系统中。1987 年 Robert 等人对级联码 ARQ 方案做了详尽的性能分析^[18]。

1986 年, G. Benelli 研究了一种在收端附加缓存器的 Go-back-N 方案^[26],还研究了一种利用软判决译码的记忆型 ARQ 方案^[33]以及与调制解调方式相结合的 ARQ 方案^[34]。1985 年, B. Arazi 和 J. Hsieh 利用循环码的捕错译码(Error-trapping decoding)特点提出了一种收敛 ARQ 方案^[27]。1977 年, P. S. Sindhu 研究了一般的记忆型重传差错控制方案。1983 年, J. J. Metzner 和 D. Chase 还提出一种适于高噪声信道的 SR-ARQ 方案^[38]。1968 年, R. L. Kashyap 研究了反向信道有噪声干扰时的反馈编码方案^[42]。1979 年, J. M. Morris 研究了 ARQ 差错控制编码方案的最佳分组长度^[53]。

在研究 ARQ 方案的同时,理论工作者对各种 ARQ 方案性能也做了详细的分析。这些性能主要集中在两个方面:①效率和可靠性分析。②利用排队论对待发信息的排队长度等待时间的分析。前一种主要用来评价各种 ARQ 的优劣,后一种主要用于在设计 ARQ 系统时对系统的缓存器容量以及信息传播延迟时间做一估价。1979 年, Don Towsley 与 Jack K. Wolf^[19]对利用 ARQ 重传方案的统计多工器的排队长度及等待时间做了统计分析,得出了两者平均值及各自的生成函数。1980 年, Alen G. Konheim^[20]对分组 ARQ (block ARQ) 与选择 ARQ (Select ARQ) 做了排队分析。另外, E. Gelenbe 和 C. Poulain^[21]以及 M. E. Anagnostou 和 E. N. Protonotarios^[22]对几种 ARQ 方案也做了排队分析, S. Tasaka 和 K. Lshida^[23]对在 ALOHA 网的 SURC 协议中利用 Go-back-N ARQ 方案在卫星信道条件下的性能做了理论分析,分析表明,在 SURC 协议中利用 Go-back-N 方案比利用 Stop-and-Wait ARQ 方案的性能有很大改善。

D. Raychaudhuri^[24]对在 Stotted ALOHA 系统中多信息包传输条件时利用 ARQ 重传技术解决了当信息包发生冲突时系统性能下降的问题。文中分析了利用 Go-back-N ARQ 与 SR-ARQ 两种不同的 ARQ 方案的性能。结果表明,当适当选择参数时,利用 SR-SRQ 方案比不用该方案的 ALOHA 系统的效率至少高出 25%~40%。

1985 年, David Chase^[25]从信息论的基本定理出发,概括地阐述了码组合方法在可靠通信中的应用。它的依据是:当信道上传输的数据的速率低于信道容量时无差错通信是可能的。利用差错控制编码技术可以满足用户提出的可靠性要求。然而,设计一个可靠通信系统的一个基本问题是如何选择码速率。目前业已存在的各种纠错编码的码速率远远低于理论最佳值,有相当多的校检位是多余的,码组合技术则提供了一种匹配码速率与信道条件的方法。这种技术是将待传信息以包的格式发送。每个包都按较高的码速率编成一个特定的码组,通过利用重传技术可以获得可靠通信。当在一个包中的校验位不足以克服信道干扰时,接收端则组合

各次发送过来的受噪声干扰而不能恢复正确信息包（码组合）而得到一个码速率足够低的码组，使它足以抵抗信道的干扰。文献〔25〕中讨论了几种码组合方法，并给出了极大似然译码方法。

总之，由于对通信系统性能要求日益提高，人们在不断地寻求各种有效的方法，差错控制编码技术是其中的一个方向。随着各种有效的纠错编码的不断出现，大规模集成电路功能不断提高以及计算机的大量应用，给差错控制编码技术的应用开辟了广阔的前景。

第一章 差错控制和编码的基本概念

§ 1-1 数字传输系统的质量要求

对数字传输系统的技术要求是信息传输的速率、差错率（误码率）、可靠性、适应性和经济性等。我们总是希望信息传输的速率尽可能的快，误码率尽可能的低，可靠性尽可能的高，维护方便，适应性强，经济耐用。其中信息的传输速率和误码率是衡量数字传输系统的两项主要技术指标。

信息传输速率（简称传信率），是指数字传输系统每秒钟所传送的二进制码元数，单位是比特每秒（bit/s），它不应当包括信道编码器所附加的监督码元在内。它代表数字传输系统的一个本质指标，表示系统每秒所传输的信息量。注意传信率与信号速率（以波特为单位）是截然不同的两个概念，信号速率是指每秒钟所传输的数字波形数（例如脉冲波、余弦波、升余弦波等）。传信率 bit/s 和信号速率波特之间有一定的数量关系，这决定于数字传输系统传输信息采用的是几元制。结 N 元制来说，每个数字波形携带 $\log_2 N$ (bit) 信息量，故若它的波形速度为 M 波特，那么它的信息传输速率为 $M \log_2 N$ bit/s。对二元制来说，波特和比特在数值上是相同的。以图 1-1 为例来说明上述概念，图中给出一个四元制波形图，以四种不同电平组成四元制脉冲序列。这个波形在数字信道上传输必须把每个码元以两个二进制码元来表示。如果信号速率是 800bit，则传信率应为 1600bit/s。在信号速率（波特）中，已包含“每秒钟”的意思在内，而在传信率（bit）中，不包含“每秒”的意思，所以必须计为若干 bit/s。

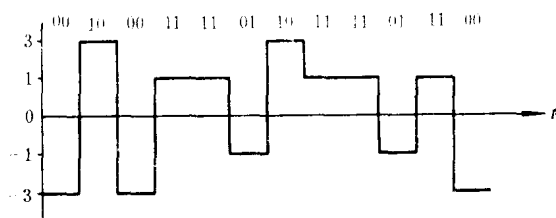


图 1-1 说明信号速率与波形速率不同的四元制波形

衡量数字传输系统的另外一个指标是误码率。它是数字传输系统正常工作时传输系统传输信息可靠程度的一个重要指标，是指所传信息总量中发生错误的信息量所占的比例，由下式确定：

$$p_e = \frac{\text{发生错误的码元数}}{\text{接收的总码元数}} \quad (1-1)$$

这是一个统计平均数，例如说误码率 $p_e = 10^{-4}$ 就是指平均每传输 10000 bit 信息，平均错了 1 bit。根据多年的统计，在一般调制制度和中等传信率（600~4800bit/s）情况下，普通电报电

话线的 $p_e = 10^{-4} \sim 10^{-6}$ 。而高频无线电信道或散射信道的 $p_e = 10^{-2} \sim 10^{-4}$ 。误码率所依赖的因素很多，诸如信道类型、调制种类和噪声等等。

传信率和误码率二者是互相制约的，例如利用纠错编码可以减少误码率，提高信息传输的可靠性。但是，这就降低了传信率，为了合理的设计一个数字传输系统，必须根据使用场合和技术要求综合考虑通信的效率和可靠性这两个主要问题以及二者之间的关系。

除了改善信道和合理地选择调制制度以外，采用纠错编码技术可大大降低误码率。但是传信率也降低了。可见提高信息传输的可靠性是用降低传信率换取的。所谓纠错编码就是在信源端将已变换成二进制的信息按某种规律再附加一些监督码元编成抗干扰码。在接收端则按同一规则还原（解码）成二进制信息。

山农（Shannon）曾经证明过一个定理。定理指出：**对于有干扰的信道，存在一定的信道容量 C ，只要信息传输速度不超过信道容量 C ，就一定存在一种编码方法，使得译码错误概率（即误码率）为**

$$p_e \leq e^{-nE(R)} \quad (1-2)$$

式中， R 是信息传输速率； n 是码长； $E(R)$ 是 R 是函数，当 $R < C$ 时 $E(R) > 0$ 。对于白高斯噪声，信道容量 C 可从下式求出：

$$C = W \log_2 \left(1 + \frac{S}{N_0 W} \right) \quad (1-3)$$

式中的 W 是信道的带宽； N_0 是噪声的功率谱密度，所以 $N_0 W$ 是噪声功率； S 是信号功率。由此可见，只要信息传输速率 $R < C$ ，就有一种编码存在，使得误码率满足式（1-3）。如果选择 n 足够大，则误码率 p_e 可以任意小。山农给出了信道容量的公式并没有给出具体的编码方法，但是指出了用编码来解决这个问题的途径。虽然到目前为止，还没有找到能符合山农编码定理的极限的编码方法，但是，要使误码率降低，必须增加码长，这一点是肯定的。

由于数字传输系统和电子计算机都广泛采用二进制，所以下面的讨论仅限于二进制。

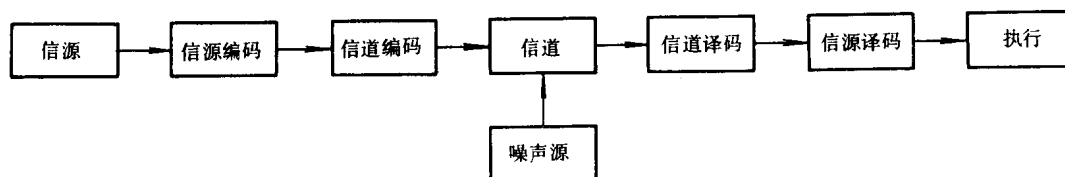


图 1-2 数字传输系统模型

根据以上讨论，可以把具有纠错编码的传输系统归结为图 1-2 所示的模型。这里信源是指任意消息源（被测量或遥控命令），通常被测量都是连续信源，遥控大都是离散信源。信源编码器把消息转换为二进制序列以便于下一环节进行传输，如果消息本来就是离散二进制序列（例如遥控信源），信源编码器可以省去。这里的信道是广义的信道，包括调制解调器和实际信道。

数字信号在信道中传输，不可避免地要发生错误。对于二元制的情况，错误可以有两种

形式：随机错误和突发错误。随机错误常常发生于恒参的相加性平稳随机干扰信道中，它的特点是，码元之间的错误是互相独立的。换句话说，每个码元的错误概率与它前面的码元和后面的码元是否有错是无关的。突发性错误则不然，一个码元的错误往往影响前后码元的错误概率。或者说，一个码元产生错误，则后面几个码元都可能产生错误。在磁性存贮系统中，磁带或磁性记录材料的缺陷以及读写磁头接触不良等，都会产生突发性错误。举例来说，假定发送的码序列为 0000000...，由于干扰而变为 0110010...，则称收发序列之间的差别 11001 为错误图样，头一个错误与最后一个错误的距离（码元数）称为突发长度 b ，上例中 $b=5$ 。在实际信道中，上述两种错误形式往往兼而有之。产生随机错误的信道称为随机信道或无记忆信道。产生突发错误的信道称为突发信道或有记忆信道。两者兼而有之的称为复合信道。严格地说，复合信道是比较普遍的，不过如果随机错误比突发错误常见的多，也可以认为是随机信道。对于某一具体的抗干扰编码和译码方法，对抗随机错误和突发错误的能力是不一样的。因此，设计抗干扰编码时，必须首先确定错误的形式，然后选择适当的编码方法，以达到给定的技术要求。

从每个码元的错误性质看，信道又可以分为三种：对称信道、不对称信道和删除信道。在信道中传输由“0”和“1”构成的二元序列时，产生错误不外乎“0”变成“1”，或“1”变成“0”。若“1”变成“0”或“0”变成“1”的概率相等，或二者是独立的，则称为二元对称信道（简称 BEC），可以用图 1-3 来表示。图中 p_0 称为信道转移概率。若“1”变成“0”和“0”变成“1”的概率不相等，则称为不对称二元信道。所谓删除信道（简称 BSC）是指二元码在信道传输时，“0”和“1”被正确接收时的概率均为 q ，而“0”和“1”被删除的概率为 $1-q$ ，“0”被删除后并不变为“1”，“1”被删除后并不变为“0”。“0”或“1”被删除后均可被发现，只是不知道原来究竟是“0”还是“1”而已。显然，这种信道的纠错要比 BSC 容易。删除信道可以用图 1-4 来描述。

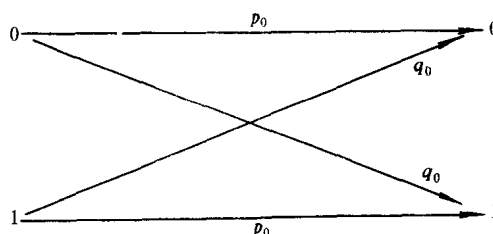


图 1-3 BSC 信道转移概率

编码理论工作者利用信道转移概率模型找出许多有效的码字，随着大规模集成电路的出现，使得这些码的编码器和译码器得以硬件实现。这可从下面事实看出：设 $c=c_{n-1}, c_{n-2}, \dots, c_2, c_1, c_0$ 为发送码字， $r=r_{n-1}, r_{n-2}, \dots, r_2, r_1, r_0$ 为接收码字，由于噪声存在使得接收序列 r 可能与发送序列 c 不同。如果序列 r 与序列 c 在长度 n 内有 i 个不同值，那么接收的条件概率为：

$$p[r/c] = \sum_{i=0}^{n-1} p[r_i/c_i] = p^i q^{n-i} \quad (1-4)$$

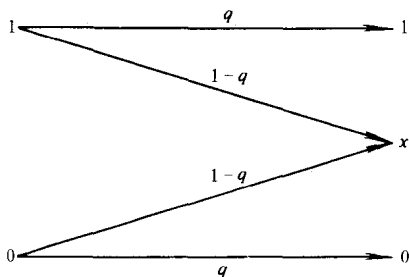


图 1-4 删除信道转移概率

因为 $q > p$, i 数目越大, $p[r/c]$ 越小, i 数目越小, $p[r/c]$ 越大。因此, 纠错译码的问题便是判断与 r 序列各对应码元位差值数目 i 最小的序列 c , 也即计算条件概率 $p[r/c]$ 最大的问题。

§ 1-2 纠错和检错原理

一、基本概念

设信源有 A、B、C 三个消息需要传送, 并将其变换为长度为 3 的二元符号序列作为传输的信息。显然, 长为 3 的二元符号序列共有 8 种组合:

$$\begin{array}{l}
 0 \ 0 \ 0 \\
 0 \ 0 \ 1 \\
 0 \ 1 \ 0 \\
 0 \ 1 \ 1 \\
 1 \ 0 \ 0 \\
 1 \ 0 \ 1 \\
 1 \ 1 \ 0 \\
 1 \ 1 \ 1
 \end{array} \tag{1-5}$$

这里, 挑选其中 3 个序列分别对应 A、B、C, 如下表所示:

$$\begin{array}{ll}
 \text{A} \longrightarrow & 001 \\
 \text{B} \longrightarrow & 111 \\
 \text{C} \longrightarrow & 110
 \end{array} \tag{1-6}$$

该表就称作编码表, 所对应的全体称之为编码, 所用到的符号 0、1 称为编码符号, 被传输的序列 001、111、110 分别称之为码字。

在传输这些码字途中将要发生差错。所谓差错也就是“0”变成“1”或“1”变成“0”。为方便问题的说明, 假定在所考虑的信道上, 仅第 1 和第 2 比特容易产生差错, 而没有两个比特同时出现错误的情况。如果把所产生的错误分别写成如

$$\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \quad (1-7)$$

的二元系列的形式是很方便的。其中 100 表示第一比特出现了错误, 010 表示第二比特出现的错误, 001 表示第三比特出现的错误。上面的序列统称为错误图样 (**error pattern**), 若把完全没有发生错误的系列的错误图样也计算在内, 则该信道中将有 000、100、010、001 四种错误图样。

如果将码字和错误图样都写成上面所示的行矢量, 并分别用 $\mathbf{w}$ 和 $\mathbf{e}$ 来表示, 则接收序列 $\mathbf{w}'$, 可写成:

$$\mathbf{w}' = \mathbf{w} + \mathbf{e} \quad (\text{mod } 2) \quad (1-8)$$

这里矢量相加就是其对应分量相加, $\text{mod } 2$ 表示以 2 为模的加法, 它所得到的结果偶数为 0, 奇数为 1。

例如在传输上述码字 B 即 111 时, 出现了 100 的错误, 此时有 $\mathbf{w} = 111$ $\mathbf{e} = 100$ 则

$$\mathbf{w}' = \mathbf{w} + \mathbf{e} = 211 = 011 \quad (\text{mod } 2) \quad (1-9)$$

同样, 用上例方法在接收端得到所有序列:

$$\begin{array}{lll} \text{A 所对应的码字} & 001 & 101 & 011 \\ \text{B 所对应的码字} & 111 & 011 & 101 \\ \text{C 所对应的码字} & 110 & 010 & 100 \end{array} \quad (1-10)$$

二、差错的检出及纠正原理

若码字和错误图样用式 (1-6)、(1-7) 给出, 则信源消息 A、B、C 和长度为 3 的 8 种二进制序列具有如图 1-5 所示的对应关系。

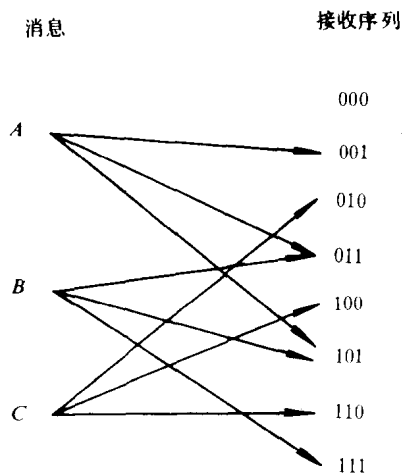


图 1-5 差错的检出及纠正原理