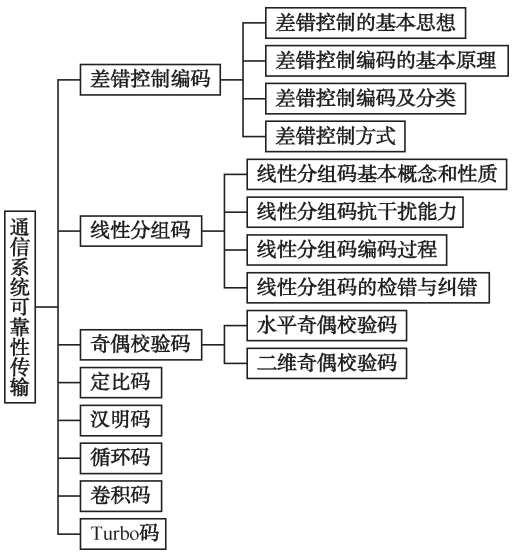
模块4

通信系统可靠性传输

知识分布网络



导入案例

传感器网络有着巨大的应用前景，被认为是将对 21 世纪产生巨大影响力的技术之一。已有和潜在的传感器应用领域包括军事侦察、环境监测、医疗、建筑物监测等。随着传感器技术、无线通信技术、计算技术的不断发展和完善，各种传感器网络将遍布我们生活环境，从而真正实现“无处不在的计算”。

传感器网络研究最早起源于军事领域，实验系统有海洋声呐监测的大规模传感器网络，也有监测地面物体的小型传感器网络。现代传感器网络应用中，通过飞机撒播、特种炮弹发射等手段，可以将大量便宜的传感器密集地撒布于人员不便于到达的观察区域如敌方阵地内，收集到有用的微观数据；在一部分传感器因为遭破坏等原因失效时，传感器网络作为整体传感器网络仍能完成观察任务。传感器网络的上述特点使得它具有重大军事价值。

无线信道环境是相当恶劣和复杂的。对于接收端的信号，不但存在由于地理环境引起的衰落和阴影，而且还要受到开放式信道结构带来的各种干扰和噪声的影响。这些衰落和干扰所造成的误码有随机差错和突发差错，通常以多径衰落和长突发差错为主，这将严重损害通信质量。因此，在无线通信这种变参的混合信道中，必须采用差错控制方案和其他抗衰落技术来提高信号的传输质量，保证信息可靠传输。

差错控制是通信网络中一个非常重要的错误处理机制。信源产生二进制符号信息，信道编码器将这些符号信息，按一定的规则加上冗余，从而产生更高比特率的编码数据。接收端的信道译码器利用这些冗余来判断发送端发送的比特信息是否正确。差错控制方案按进行纠错的工作方式，可以分为前向纠错、自动重发请求和混合模式。

在前向纠错控制的方案中，接收端不但能够利用所附加的冗余信息（监督码元）来检测接收到的信息是否有错误，并且由于冗余信息是按一定规则生成的，所以还能够纠正接收端的错误。前向纠错既能检测错误，也能纠正一定数量的错误，其优点是发送时不需存储，不要反馈信道；而缺点是译码设备复杂，纠错码与信道干扰情况相关。

自动重发请求中，信源产生的信息码元在编码器中被分组编码后，到达接收端的译码器。如果根据监督码元检测出有错，则进行请求重发。

混合模式是 FEC 及 ARQ 两种方式的混合。混合纠错的工作方式是：少量错误在接收端自动纠正，差错较严重，超出自行纠正能力时，就向发信端发出信号，要求重发。

差错控制主要利用检测码或纠错码进行检错或者纠错。所谓检错码是能够自动发现错误的编码；纠错码是能够发现错误且又能自动纠正错误的编码。

通信系统中，信源产生要发送的信息，经过信道编码，在无线信道中进行传输，在到达接收端之前，进行译码，做相应的差错处理。信道编码的过程，就是按照一定规则在信元上附加“冗余”信息的过程。在接收端进行译码的过程实质是根据冗余规则，进行错误的检测和处理。

思考：

差错控制方式有哪些？接收端怎样进行检错纠错？

学习目标

- ☞ 理解差错控制的基本思想和有关概念。
- ☞ 会计算编码的最小码距离，并能根据最小码距离计算该种编码的纠错检错位数。
- ☞ 理解奇偶校验码、汉明码、循环码、卷积码、Turbo 码编码的特点和构造思路。
- ☞ 掌握奇偶校验码、汉明码、循环码的编码方法。

4.1　差错控制编码

扫一扫看差错控制编码及线性分组码教学课件

4.1.1　差错控制的基本思想

信号在传输过程中不可避免地受到干扰，原因主要归结为两个方面：一是信道特性不理想造成的码间干扰；二是噪声对信号的干扰。信号到达接收端时，接收信号是信号与各种干扰的叠加，接收电路在取样时判断信号电平。如果干扰对信号叠加的结果在电平判断时出现错误，就会引起通信数据的错误，就出现了误码。数字通信系统中码元的错误有三种形式。

1．随机错误

错误的出现是随机的，一般而言错误出现的位置是随机分布的，即各个码元是否发生错误是互相独立的，通常不是成片地出现错误。这种情况一般是由信道的加性随机噪声引起的。因此，一般将具有此特性的信道称为随机信道。

2．突发错误

错误的的出现是一连串出现的。通常在一个突发错误持续时间内，开头和末尾的码元总是错的，中间的某些码元可能错也可能对，但错误的码元相对较多。错码出现时，在短时间内有一连串的错码，而该时间过后又有较长时间无错码。这种情况如移动通信中信号在某一段时间内发生衰落，造成一串差错；汽车发动时电火花干扰造成的错误；光盘上的一条划痕等。这样的信道称之为突发信道。

3．混合错误

既有突发错误又有随机差错的情况。这种信道称之为混合信道。移动通信的传输信道属于变参信道，它不仅会引起随机错误，而更重要的是造成突发错误。

差错控制是对传输差错采取的技术措施，目的是提高传输的可靠性。差错控制的基本思想是通过对信息序列做某种变换，使原来彼此独立的、没有相关性的信息码元序列，经过某种变换后，产生某种规律性，从而在接收端有可能根据这种规律来检查，进而纠正传输序列中的差错。变换的方法不同就构成不同的编码和不同的差错控制方式。差错控制的核心是抗干扰编码，即差错控制编码，又称纠错编码、可靠性编码，也称信道编码。不同的编码方法，有不同的检错或纠错能力，差错控制编码一般是在用户信息序列后（称为信息码元）插入一定数量的新码元（称为监督码元）。监督码元不受用户的控制，最终也不发送给接收用户，只是在系统传输过程中为了减少传输差错而采用的一种措施。如果信道传输速率一定，加入差错控制编码，就降低了用户输入的信息速率，新加入的码元越多，冗余度就越大，检错纠

错越强，效率越低。差错控制编码是通过增加冗余码来达到提高可靠性传输的目的的。

差错控制技术简单地说就是一种保证接收完整、准确数据的方法。例如，我们日常使用的电话线路是不稳定的，那么数据在传输过程中就会出现数据顺序的错乱和丢失。为了使这些错误能够得到及时地纠正，调制解调器在发送端必须对发送的数据进行信道编码，并将监督码元和信息码元同时发送，调制解调器在接收端对编码过的数据进行解码，也就是检验监督码元和信息码元是否符合该编码的规律。若不符合规律，则表明数据在传输过程中被破坏，接收端的调制解调器就会向发送端的调制解调器发送一个命令，要求数据重发。图 4.1 就是一种差错控制技术的机理图。

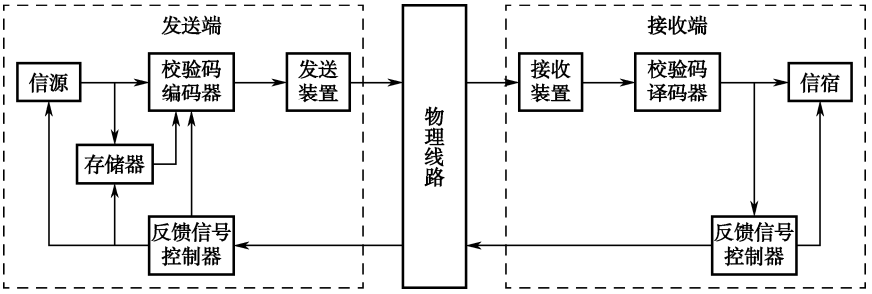


图 4.1　差错控制技术机理图

4.1.2　差错控制编码的基本原理

信息码序列中加入监督码元才能完成检错和纠错功能，其前提是监督码元和信息码元之间要有一种特殊的关系，即符合一定的规律。下面我们举例说明检错和纠错的基本原理。

假设要发送一组具有四个状态的数据信息，比如电压信号的四个值分别为 1 V、2 V、3 V 和 4 V。我们首先要用二进制码对数据信息进行编码，用 2 位二进制就可以完成。

假设不经信道编码，在信道中直接传输按表 4.1 中编码规则得到 0、1 数字序列，则在理想情况下，这样编码接收端接收没有问题。但是在实际通信中由于干扰的影响，会使信息码元发生错误，从而出现误码。比如码组 00 变成 01、10 或者 11。任何一组码不管是一位还是两位发生错误，都会使该码组变成另外一组信息码，从而引起信息传输错误，而且接收端无法判断是否有错误。因此，以这种编码形式得到的数字信号在传输过程中不具备检错和纠错的能力。

表 4.1　数据信息编码方案

数据信息	1 V	2 V	3 V	4 V
数据编码	00	01	10	11

为了使接收端能具有检错能力，我们在每组码后面再加 1 位码元，使监督码元和信息码元中 1 的个数为偶数，这样 2 位码组就变成了 3 位码组，如表 4.2 所示。这样，在 3 位码组的 8 种组合中只有 4 组（000、011、101 和 110）是按照编码规则允许使用的码字，称为许用码组，而其余 4 种（001、010、100 和 111）不符合编码规则的码字，被称为禁用码组。表 4.2 中每个码组右边加上的 1 位码元就是监督码元，加入监督码元的原则就是使监督码元和信息码元中 1 的个数为偶数。如果许用码组 000 在传输过程中出现一位误码，即变成了 001、

010 或者 100 三个码组中的一个，则不满足编码规则（信息码元和监督码元中 1 的个数为偶数），成为禁用码组。当接收端收到这三个禁用码组中的任何一个时，按照监督码元和信息码元的关系（1 的个数为偶数）判断出是误码。因此表 4.2 可以发现一位错误。但是当接收端收到一个误码 010 时，可能是 000、011、110 错一位得到，也可能是 101 错两位得到，没有办法判断是哪一位错误得到 010，因此没有办法对收到的错码 010 进行纠错。

表 4.2　信道编码方案 A

数据信息	1 V	2 V	3 V	4 V	×	×	×	×
数据编码	000	011	101	110	001	010	100	111

为了使接收端具有纠错能力，在表 4.1 数据信息编码后面增加 4 位监督位，如表 4.3 所示（由于禁用码太多，没有列出来）。如果接收端收到码字 000001，那么可能是 000000 错 1 位，011011、101101 和 110110 错 3 位得到的。因为传输中码字错的位数比多的位数少出现的概率小得多。因此，如果接收端收到码字 000001，那么接收端会认为是 000000 错 1 位得到的，接收端则直接把收到的码字判为 000000，这样就达到了纠正错码的目的。

表 4.3　信道编码方案 B

数据信息	1 V	2 V	3 V	4 V
数据编码	000000	011011	101101	110110

4.1.3　差错控制编码及分类

从不同的角度出发，信道编码有不同的分类方法，如图 4.2 所示。

1．按码组的功能分

按码组的功能分，有检错码和纠错码两类。一般来说，在译码器中能够检测出错码，但不知道错码的准确位置的码，称为检测码，它没有自动纠正错误的能力。若在译码器中不仅能发现错误，而且知道错码的位置，自动纠正错误的码，则称为纠错码。

2．按码组中的监督码元和信息码元之间的关系分

按码组中的监督码元和信息码元之间的关系分，有线性码和非线性码两类。线性码是指监督码元与信息码元之间呈线性关系，即可用一组线性代数方程联系起来；否则为非线性关系。

3．按照信息码元与监督码元的约束关系分

按照信息码元与监督码元的约束关系，又分为分组码和卷积码两类。所谓分组码就是将信息序列以每 k 个码元分组，通过编码器在每 k 个码元后按照一定的编码规则产生 r 个监督码元，组成长度为 $n=k+r$ 的码组，每一个码组中的 r 个监督码元仅监督本码组中的信息码元，而与别组无关。分组码一般用符号（n，k）表示。

在卷积码中，每组的监督码元不但与本组码的信息码元有关，而且还与前面若干组的信息码元有关，即不是分组监督，而是每个监督码元对它的前后码元都实行监督，前后相连，有时也称为连环码。

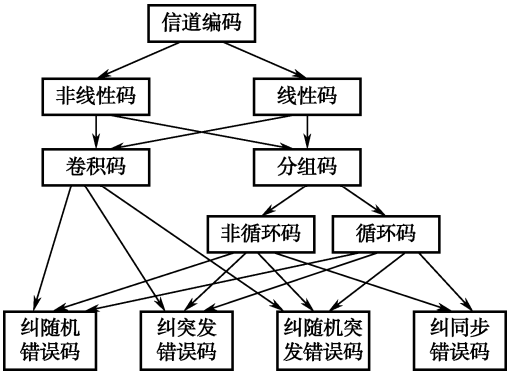


图 4.2　信道编码分类

4．按照信息码元在编码前后是否保持原来的形式不变分

按照信息码元在编码前后是否保持原来的形式不变，可划分为系统码和非系统码。系统码的信息码元和监督码元在分组内确定位置，而非系统码中信息码元则改变了原来的信号形式。

4.1.4　差错控制方式

在数字通信系统中，信道编码和差错控制方式是结合起来使用的，如图 4.3 所示。比如，前向纠错码和纠错编码结合起来使用，前向纠错码就不能和检错码结合起来使用。常用信道编码的差错控制的方式主要有前向纠错、自动重传请求和混合纠错三种。

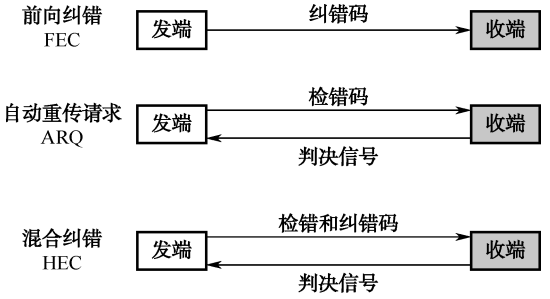


图 4.3　信道编码和差错控制方式

1．前向纠错（Forward Error-Correction，FEC）

FEC 的基本思想是利用纠错编码来控制传输差错，在发送端将信息按照一定规则附加冗余码元，使之具有一定的纠错能力；在接收端收到码元后，按预先规定的规则校验信息与冗余码元之间的关系，若发现错误则确定其出错位置并进行纠正。通过纠错编码可以降低误比特率，但如果差错超过了其纠错能力，那接收的码组将被错误地译码，并将错误码组传给用户。

其主要优点是：不需要反向信道就能进行一个用户对多个用户的同时通信（广播），特别适合于移动通信；而且系统的传输效率高；译码延迟固定，信息传输时延和时延抖动都比较小，实时性好，较适用于实时传输系统；控制电路比较简单。

然而 FEC 也存在一些缺点：纠错编码是以引入冗余比特，加大开销为代价，可能会导致不必要的浪费；当译码出现错误时，错误的信息会传递给用户，所以其可靠性较差；为了获得较高的可靠性，设计时必须使用长码和选用纠错能力强的码组，这会增加译码电路复杂度，提高成本；编解码使计算的开销和复杂性大大增加，在丢包率很高时，性能下降明显；只适合一次发送一个数据包的应用；其整体性能受丢包最严重的接收者制约等。另外，FEC 采用“事先避免”的策略，即使事后仍有丢包，也不再重传。因此，单纯的 FEC 技术并不能完全保证数据传输的正确性。

2．自动重传请求（Automatic Rpeat Request，ARQ）

ARQ 的基本思想是在发送端和接收端之间引入反向链路，发送端对信息进行编码，编码后的信息具有很强的检错能力，通过前向信道发送到接收端。在接收端进行检错译码，如果没有检出错误，则提交给用户（或存入缓冲寄存器备用），同时，通过反向信道向发送端返回一个确定应答（ACK），通知发送端此信息已经正确接收。如果检出错误，则通过反向信道返回一个否定应答（NAK），请求对方把刚才的信息重发一次，这样持续进行下去直到正确接收或达到最大重传次数为止。由此可知，应用 ARQ 方式必须存在一条反馈信道，并要求发送端信息的产生速率可以控制（或有大容量的信息发送缓冲存储器），整个通信系统的发送端和接收端必须密切协作，互相配合，因此 ARQ 方式的控制过程相对比较复杂。由于进行反馈重发的次数与信道情况有关，若信道情况较差，则系统经常处于反馈重发的状态，所以信息传输的实时性和连贯性较差。该方式的优点是编解码设备简单，尤其是解码设备，在冗余度一定的情况下其检错能力比纠错码的纠错能力要高很多，所以检错能力极强，因而整个差错控制系统的适应性很强，特别适用于干扰情况特别复杂的短波和散射等信道以及对误码率要求极低的场合。ARQ 系统与 FEC 系统相比，不仅设备简单，而且可靠性高。但它必须存在一个反向信道，并且当信道误码率太大时，系统会经常处于重传状态而使传输效率非常低。

ARQ 有三种基本的重传机制：停止等待 ARQ（见图 4.4）、连续重传 N-ARQ（见图 4.5）和选择重传 ARQ（见图 4.6）。后两种协议是滑动窗口技术与请求重传技术的结合，由于当窗口尺寸足够大时，帧在线路上可以连续地发送，因此又称其为连续 ARQ 协议。三者的区别在于对于出错数据块的处理机制不同。三种 ARQ 协议中，复杂性递增，效率也递增。

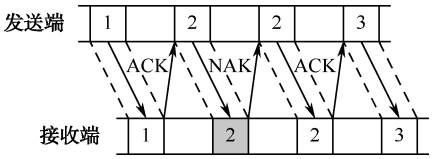


图 4.4　停止等待 ARQ

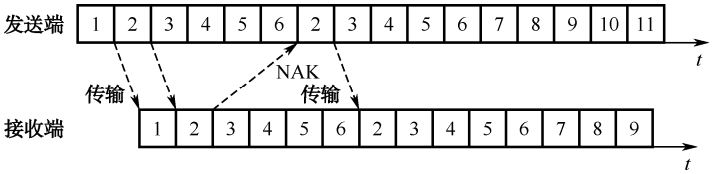


图 4.5　连续重传 ARQ

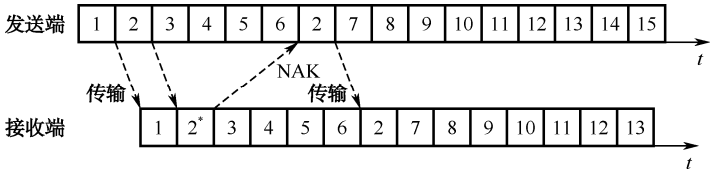


图 4.6　选择重传 ARQ

3．混合纠错（Hybrid Error-Correction，HEC）

FEC 和 ARQ 分别是利用纠错码和检错码实现差错控制的技术。ARQ 方式检错能力强，但需要一个反馈信道，并且实时性较差。相反，FEC 方式的通信实时性好，收发控制系统电路简单，但纠错码往往是以最坏信道条件来进行设计，因此编码的效率较低。混合纠错检错方式是前向纠错方式和检错重发方式的结合。在这种系统中，发送端发出同时具有检错和纠错能力的码，接收端收到码后，检查错误情况，如果错误少于纠错能力，则自行纠正；如果干扰严重，错误很多，超出纠正能力，但能检测出来，则经反向信道要求发端重发。混合纠错检错方式在实时性和译码复杂性方面是前向纠错和检错重发方式的折中。

除了上述三种主要的方式以外，还有所谓狭义信息反馈系统（Information Repeat Request，IRQ）和检错删除。狭义信息反馈是指接收端将收到的码元原封不动地通过反馈信道送回发送端，发送端比较发送的与反馈回来的消息，若发现错误，发送端把传错部分对应的原消息再次传送，最后达到使对方正确接收消息的目的。该方式的缺点是须采用双向信道，传输效率也很低。检错删除是指在接收端发现错码后，立即将其删除。适用在发送码元中有大量多余度，删除部分接收码元不影响应用之处。

4.2　线性分组码

分组码一般可用（n，k）表示，其结构如图 4.7 所示。其中，k 是每个码组二进制信息码元的数目，n 是编码组的码元总位数，又称为码组长度，简称码长。$n-k=r$ 为每个码组中的监督码元数目。简单地说，分组码是对每段 k 位长的信息组以一定的规则增加 r 个监督元，组成长为 n 的码字。在二进制情况下，共有 2^k 个不同的信息组，相应地可得到 2^k 个不同的码字，称为许用码组。其余 2^n-2^k 个码未被选用，称为禁用码组。

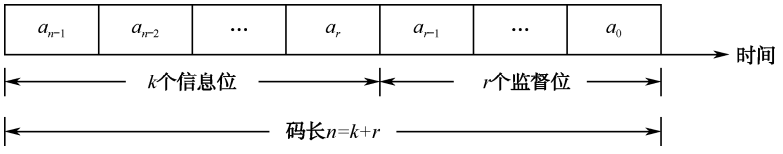


图 4.7　分组码的结构示意图

4.2.1　线性分组码基本概念和性质

线性分组码是所有纠错编码中最基本最容易研究的一类码，它概念清楚，易于理解，而且能方便地反映出各类编码中广为使用的一些基本参数和名称。因此，线性分组码就成了讨论其他各类码的基础。

在（n，k）分组码中，若每一个监督元都是码组中某些信息码元按模 2 和得到的，即监

督码元是信息码元按线性关系相加而得到的，则称线性分组码。或者说，可用线性方程组表述码规律性的分组码称为线性分组码。线性分组码是一类重要的纠错码，应用很广泛。

码字中码元的数目称为码长，如 001，码长为 3。码字中非 0 码元的个数称为该码字的码重，又称为汉明重量。如 001，码重为 1。两个等长码字之间对应位不同的个数称为两个码字之间的码距，又称为汉明距离。如 001 和 000 之间的码距为 1。在（n，k）线性分组中，任意两个不同码字之间的距离最小值称为该分组码的最小汉明距离，用 $d_{\min}$ 表示。如 000、011、101 和 110 两两之间的码距有 2 和 3，最小码距则是 2。

监督码元的引入，增加了原始信息码元的数目，这就引入了编码效率的概念。若码字中信息位为 k，监督位为 r，码长为 $n=k+r$，编码效率是指信息码元数与码长之比，通常用 R_c 来表示。

$$R_c=\frac{k}{n}=\frac{n-r}{n} \tag{4-2-1}$$

采用差错控制编码是为了提高通信系统的可靠性，但是它是以降低有效性为代价换来的。对信道编码的基本要求是：检错和纠错能力尽量强；编码效率尽量高；编码规律尽量简单。实际中要根据具体指标要求，保证有一定的纠、检错能力和编码效率，并且易于实现。

线性分组码具有以下两个性质。

（1）封闭性：任意两个许用码组相加（模 2 加）后，所得码组仍是许用码组。

（2）最小码距：等于除全“0”码组以外的最小码重。

4.2.2　线性分组码抗干扰能力

采用表 4.2 信道编码方案 A，信息码元后面增加了一位监督位，可以发现 1 位发生错误或者 3 位出现错误的码组，而无法检出 2 位错误。采用表 4.3 信道编码方案 B 增加了 3 位监督位，可以发现错误，并纠正 1 位错误。

那么能否得出这样的结论：增加监督码元的位数就能增加检错位数或实现纠错功能？将表 4.1 中的编码增加 2 位监督码元，采用重复编码，变成 4 位编码，观察情况如何，如表 4.4 所示。

表 4.4　信道编码方案 C

数据信息	1 V	2 V	3 V	4 V	×	×	×	×
数据编码	0000	0101	1010	1111	0001	0010	0100	1000
					0011	0111	1001	1011
					1100	1110	1101	0110

用这种编码方案可以发现 1 位错误，如 0000 错一位变成 0001、0010、0100、1000 四个禁用码组中的一个，由此可以判断出误码，但是无法判断出是哪一位错误。若 0000 错两位可能变成 0011、0101、0110、1001、1010、1100 中的任何一种，而 0101、1010 是许用码组，故如果 0000 变成了 0101、1010 则无法检测出错误，因此，这种编码方案只能检测 1 位误码，不能纠正 1 位误码，也不能检测 2 位误码。

由此可见，表 4.4 相对于表 4.2 增加了监督码元位数，并没有提高检错与纠错能力，那么检错与纠错能力究竟与什么有关呢？

一种编码方式的检错与纠错能力与许用码组中的最小码距有关。一般情况下，分组码的最小码距 $d_{\min}$ 和分组码的检错纠错能力存在如下关系：

（1）要检测 e 位误码，则要求

$$d_{\min} \geqslant e+1 \tag{4-2-2}$$

（2）要纠正 t 个错误，则要求

$$d_{\min} \geqslant 2t+1 \tag{4-2-3}$$

（3）要码字用于纠正 t 个错误，同时检测 e 个错误，则要求

$$d_{\min} \geqslant t+e+1 \tag{4-2-4}$$

显然，要提高编码的检错纠错能力，不能仅靠简单地增加监督码元位数（即冗余度），更重要的是要加大最小码距，而最小码距的大小与编码冗余度是有关的，最小码距增大，码元的冗余度就增大。

4.2.3 线性分组码编码过程

对于偶监督码，使用了一位监督位 a_0，设码字 $A=[a_{n-1},a_{n-2},\cdots,a_1,a_0]$，有

$$a_{n-1} \oplus a_{n-2} \oplus a_{n-3} \oplus \cdots \oplus a_1 \oplus a_0 = S \tag{4-2-5}$$

在接收端解码时，实际上就是计算式（4-2-5）中 S 的结果，若结果为 1，则认为有错，结果为 0，则认为无错。式中，S 称为校正子，取值只有两种，故只能代表有错和无错两种信息，若增加一位监督位，则能增加一个类似于上式的监督关系式。若有两个校正子，它们有 4 种可能值组合，故能表示 4 种不同信息，则除了表示有无错信息外，还能有其余可能值来表示错误的位置信息。同理，r 个监督位能表示 2^r-1 个可能错误的位置。

现以（7，4）分组码为例来说明线性分组码的特点。设其码字为 $A=(a_6,a_5,a_4,a_3,a_2,a_1,a_0)$，其中前 4 位是信息码元，后 3 位是监督码元，可用下列线性方程组来描述该分组码，产生监督码元。

$$\begin{cases} a_2 = a_6 + a_5 + a_4 \\ a_1 = a_6 + a_5 + a_3 \\ a_0 = a_6 + a_4 + a_3 \end{cases} \tag{4-2-6}$$

显然，这三个方程是线性无关的。经计算可得（7,4）码的全部码字，如表 4.5 所示。

表 4.5 （7,4）线性分组码码字

	码　字			码　字	
序　号	信息码元	监督码元	序　号	信息码元	监督码元
0	0000	000	8	1000	111
1	0001	011	9	1001	100
2	0010	101	10	1010	010
3	0011	110	11	1011	001
4	0100	110	12	1100	001
5	0101	101	13	1101	010
6	0110	011	14	1110	100
7	0111	000	15	1111	111

根据线性分组码的性质和最小码距与分组码的抗干扰能力的关系，不难看出，上述（7,4）线性分组码的最小码距 $d_{\min}=3$，它能纠正一个错误或检测两个错误。

将式（4-2-6）改写成如下形式：

$$\begin{cases}1\cdot a_6+1\cdot a_5+1\cdot a_4+1\cdot a_3+1\cdot a_2+0\cdot a_1+0\cdot a_0=0\\1\cdot a_6+1\cdot a_5+0\cdot a_4+1\cdot a_3+0\cdot a_2+1\cdot a_1+0\cdot a_0=0\\1\cdot a_6+0\cdot a_5+1\cdot a_4+1\cdot a_3+0\cdot a_2+0\cdot a_1+1\cdot a_0=0\end{cases}\tag{4-2-7}$$

这组线性方程可用矩阵形式表示为

$$\begin{bmatrix}1&1&1&0&1&0&0\\1&1&0&1&0&1&0\\1&0&1&1&0&0&1\end{bmatrix}\left[a_6\,a_5\,a_4\,a_3\,a_2\,a_1\,a_0\right]^T=\begin{bmatrix}0\\0\\0\end{bmatrix}\tag{4-2-8}$$

式（4-2-8）简记为 $\boldsymbol{H}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{O}^{\mathrm{T}}$，或 $\boldsymbol{A}\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{O}$，其中 $\boldsymbol{A}^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 转置矩阵，$\boldsymbol{O}^{\mathrm{T}}$ 是 $\boldsymbol{O}=[0\ 0\ 0]$ 的转置。

$$\boldsymbol{H}=\begin{bmatrix}1&1&1&0&1&0&0\\1&1&0&1&0&1&0\\1&0&1&1&0&0&1\end{bmatrix}=[\boldsymbol{P}\ \boldsymbol{I}_r]\tag{4-2-9}$$

$\boldsymbol{H}$ 称为监督矩阵，一旦 $\boldsymbol{H}$ 给定，信息位和监督位之间的关系也就确定了，$\boldsymbol{H}$ 为 $r\times n$ 阶矩阵，$\boldsymbol{H}$ 矩阵每行之间是彼此线性无关的。式（4-2-9）中的 H 矩阵可以分成矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{I}_r$ 两部分，其中 $\boldsymbol{P}$ 为 $r\times k$ 阶矩阵，$\boldsymbol{I}_r$ 为 $r\times r$ 阶单位矩阵，我们将具有 $\boldsymbol{H}=[\boldsymbol{P}\ \boldsymbol{I}_r]$ 形式的监督矩阵称为典型监督矩阵。一般形式的 $\boldsymbol{H}$ 矩阵可以通过行的初等变换将其化为典型形式。$\boldsymbol{H}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{O}^{\mathrm{T}}$，说明 $\boldsymbol{H}$ 矩阵与码字的转置乘积必须为零，可以用来作为判断接收码字 $\boldsymbol{A}$ 是否出错的依据。

将式（4-2-6）补充为下列方程

$$\begin{cases}a_6=a_6\\a_5=a_5\\a_4=a_4\\a_2=a_6+a_5+a_4\\a_1=a_6+a_5+a_3\\a_0=a_6+a_4+a_3\end{cases}\tag{4-2-10}$$

并改写为矩阵形式

$$\begin{bmatrix}a_6\\a_5\\a_4\\a_3\\a_2\\a_1\\a_0\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&1&1&0\\1&1&0&1\\1&0&1&1\end{bmatrix}\begin{bmatrix}a_6\\a_5\\a_4\\a_3\end{bmatrix}\tag{4-2-11}$$

两边求转置，得

$$\boldsymbol{A}=[a_6\ a_5\ a_4\ a_3\ a_2\ a_1\ a_0]=[a_6\ a_5\ a_4\ a_3]\boldsymbol{G} \tag{4-2-12}$$

其中

$$\boldsymbol{G}=\begin{bmatrix}1&0&0&0&1&1&1\\0&1&0&0&1&1&0\\0&0&1&0&1&0&1\\0&0&0&1&0&1&1\end{bmatrix}=[\boldsymbol{I}_k\ \ \boldsymbol{Q}] \tag{4-2-13}$$

$$\boldsymbol{Q}=\boldsymbol{P}^{\mathrm{T}} \tag{4-2-14}$$

$\boldsymbol{G}$ 称为生成矩阵，由 $\boldsymbol{G}$ 和信息码组$[a_6\,a_5\,a_4\,a_3]$就可以产生全部码字。$\boldsymbol{G}$ 为 $k\times r$ 阶矩阵，各行也是线性无关的。生成矩阵也可以分成两部分 $\boldsymbol{I}_k=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$ 和 $\boldsymbol{Q}=\begin{bmatrix}1&1&1\\1&1&0\\1&0&1\\0&1&1\end{bmatrix}$，$\boldsymbol{I}_k$ 为 k 阶单位矩阵，$\boldsymbol{Q}$ 为 $k\times r$ 阶矩阵，可以写成式（4-2-13）形式的 $\boldsymbol{G}$ 矩阵，称为典型生成矩阵。非典型形式的生成矩阵经过简单的行运算也一定可以化成典型生成矩阵形式。任意码组 $\boldsymbol{A}$ 都是 $\boldsymbol{G}$ 的各行的线性组合。实际上，$\boldsymbol{G}$ 的各行本身就是一个许用码组。

4.2.4 线性分组码的检错与纠错

线性分组码的监督矩阵 $\boldsymbol{H}$ 和生成矩阵是密切联系在一起的。由生成矩阵 $\boldsymbol{G}$ 生成的 (n,k) 线性分组码发送后，接收端可以用监督矩阵 $\boldsymbol{H}$ 来检验收到的码字是否满足监督方程，即是否有错。设发送码组 $\boldsymbol{A}=[a_{n-1},a_{n-2},\cdots,a_1,a_0]$，在传输过程中可能发生误码。接收码组 $B=[b_{n-1},b_{n-2},\cdots,b_1,b_0]$，则收发码组之差定义为错误图样 $\boldsymbol{E}$，也称为误差矢量，即

$$\boldsymbol{E}=\boldsymbol{B}-\boldsymbol{A} \tag{4-2-15}$$

式中 $\boldsymbol{E}=[e_{n-1},e_{n-2},\cdots,e_1,e_0]$，且

$$e_i=\begin{cases}0 & \text{当} b_i=a_i\\ 1 & \text{当} b_i\neq a_i\end{cases}\qquad (i=0,1,\cdots,n-1) \tag{4-2-16}$$

上式也可以写成

$$\boldsymbol{B}=\boldsymbol{A}+\boldsymbol{E} \tag{4-2-17}$$

令 $\boldsymbol{S}=\boldsymbol{B}\boldsymbol{H}^{\mathrm{T}}$，$\boldsymbol{S}$ 称为伴随式或校正子，利用 $\boldsymbol{A}\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{O}$，得

$$\boldsymbol{S}=\boldsymbol{B}\boldsymbol{H}^{\mathrm{T}}=(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{E}\boldsymbol{H}^{\mathrm{T}} \tag{4-2-18}$$

由式（4-2-18）可见，校正子与错误图样 $\boldsymbol{E}$ 之间有确定的线性变换关系，校正子 $\boldsymbol{S}$ 只与错误图样 $\boldsymbol{E}$ 有关，可以用校正子 $\boldsymbol{S}$ 作判别错误的参量，如果 $\boldsymbol{S}$=0，则接收到的是正确码字；若 $\boldsymbol{S}\neq 0$，则说明 $\boldsymbol{B}$ 中存在着差错，接收译码器从校正子确定错误图样，然后从接收到的码字中减去错误图样，得到纠正后的码字。校正子 $\boldsymbol{S}$ 是一个 $1\times r$ 阶矩阵，也就是说校正子 $\boldsymbol{S}$ 的位数与监督码元个数 r 相等。(7,4) 码校正子 $\boldsymbol{S}$ 与错误图样 $\boldsymbol{E}$ 的对应关系如表 4.6 所示。

表 4.6 （7,4）码校正子 $\boldsymbol{S}$ 与错误图样 E 的对应关系

序号	错误码位	E	S
		$e_6\ e_5\ e_4\ e_3\ e_2\ e_1\ e_0$	$s_2\ s_1\ s_0$
0		0 0 0 0 0 0 0	0 0 0
1	b_0	0 0 0 0 0 0 1	0 0 1
2	b_1	0 0 0 0 0 1 0	0 1 0
3	b_2	0 0 0 0 1 0 0	1 0 0
4	b_3	0 0 0 1 0 0 0	0 1 1
5	b_4	0 0 1 0 0 0 0	1 0 1
6	b_5	0 1 0 0 0 0 0	1 1 0
7	b_6	1 0 0 0 0 0 0	1 1 1

实例 4.1 （7,4）分组码其监督方程为式（4-2-6），若接收端收到码字为 1100111，请分析是否有错，若有错，请纠正。

解：根据题意，接收码组 $\boldsymbol{B}=[1100111]$，其监督矩阵为式（4-2-9），即

$$\boldsymbol{H}=\begin{bmatrix}1&1&1&0&1&0&0\\1&1&0&1&0&1&0\\1&0&1&1&0&0&1\end{bmatrix}$$

通过式（4-2-18） $\boldsymbol{S}=\boldsymbol{B}\boldsymbol{H}^{\mathrm{T}}$，若接收码字 1100111 无错，那么计算结果 $\boldsymbol{S}=0$。

$$\boldsymbol{S}=[1100111]\begin{bmatrix}1&1&1&0&1&0&0\\1&1&0&1&0&1&0\\1&0&1&1&0&0&1\end{bmatrix}^{\mathrm{T}}=[1100111]\begin{bmatrix}1&1&1\\1&1&0\\1&0&1\\0&1&1\\1&0&0\\0&1&0\\0&0&1\end{bmatrix}=[110]=[s_2s_1s_0]$$

根据表 4.6 所示，得到错误码位置是 b_5，当接收码字为 1100111 时，纠正为 1000111。

4.3 奇偶校验码

扫一扫看奇偶校验码、定比码及汉明码教学课件

这是一种最简单的检错码，又称奇偶监督码，是奇校验码和偶校验码的统称，在计算机数据传输中应用广泛。

在发送端，奇（偶）监督码编码规则是先将所要传输的数据码元（信息码元）分组，在分组后的信息码元后加上一位监督码元，使信息码元和监督码元中 1 的个数为奇数（偶数），如表 4.7 所示。

表 4.7　奇偶校验码

原编码	奇校验码	偶校验码
0000	00001	00000
0010	00100	00101
1100	11001	11000
1010	10101	10100

$$a_{n-1} \oplus a_{n-2} \oplus \cdots \oplus a_1 \oplus a_0 = 0 \qquad (4\text{-}3\text{-}1)$$

$$a_{n-1} \oplus a_{n-2} \oplus \cdots \oplus a_1 \oplus a_0 = 1 \qquad (4\text{-}3\text{-}2)$$

在接收端用检查码组中 1 的个数是否符合编码规律来判断是否出错。设码组长度为 n，表示为$[a_{n-1}, a_{n-2}, \cdots, a_1, a_0]$，其中前 n–1 位为信息码元，第 n 位 a_0 为监督码元。对于偶监督码，要使码组中“1”的个数为偶数，其监督方程如式（4–3–1）所示。对于奇监督码，要使码组中“1”的个数为奇数，其监督方程如式（4–3–2）所示。如果发生奇数个错误，就会破坏上述方程式。这种奇偶校验码检测奇数个错误，不能检出偶数个错误，但是错一位的概率比错两位的概率大得多，错三位码的概率比错四位码的概率大得多。因此，绝大多数随机错误都能用简单奇偶校验查出，这正是这种方法被广泛用于以随机错误为主的通信系统的原因。但是这种方法难以应付突发错误，所以突发错误很多的信道中不能单独使用。

4.3.1　水平奇偶校验码

为了提高奇偶校验检测突发错误的能力，引入了水平奇偶校验码。其构成思路是：将信息序列按行排成方阵，每行后加一个奇或偶校验码，发送时采用交织的方法，即按列的顺序进行发送。接收端排列成与发送端相同的方阵，然后按行进行奇偶校验。如图 4.8 所示，有信息序列 0101011 1000001 1110001 01101100 要进行水平奇偶校验，先将信息序列分组，假设 7 位为一组，排列成四行方阵，然后在水平方向上进行偶校验。发送端发送时按照方阵列的顺序发送，即发送顺序为 0110 1101 1100 0001 等。接收端先将码元排列成方阵，然后按照行进行偶校验。

发送顺序	信息码元							偶校验码
0	1	0	1	0	1	1	0	
1	0	0	0	0	0	0	1	
1	1	1	0	0	0	0	1	
0	1	1	0	1	1	0	0	

图 4.8　水平奇偶校验码

水平奇偶校验不但可以检测出各段同一位上的奇数位错,而且还能检测出突发长度≤p（p 是交织的行数）的所有突发错误，突发长度≤p 的突发错误必然分布在不同的行中，且每行一位，所以可以检查出差错。但是实现水平奇偶校验码时，不论是采用硬件还是软件方法，都不能在发送过程中产生奇偶校验冗余位边插入发送，而必须等待要发送的全部信息块到齐后，才能计算冗余位，也就是一定要使用数据缓冲器，因此它的编码和检测实现起来都要复杂一些。

4.3.2　二维奇偶校验码

二维奇偶校验码是在水平奇偶校验的基础之上增加了垂直奇偶校验。也就是在发送端将信息码元进行方阵排列后按行进行奇偶校验后，还增加了按列进行奇偶校验。发送时按行或列进行发送。接收端重新将码元排列成方阵，然后按行和列分别进行奇偶校验。

水平垂直奇偶校验（见图 4.9）能检测出所有 3 位或 3 位以下的错误（因为此时至少在某一行或某列上有一位错）、奇数位错、突发长度≤p+1 的突发错以及很大一部分偶数位错。这种方式的编码可使误码率降至原误码率的百分之一到万分之一。

	信息码元							偶校验码
发送顺序	0	1	0	1	0	1	1	0
	1	0	0	0	0	0	0	1
	1	1	1	0	0	0	0	1
	0	1	1	0	1	1	0	0
偶校验码	0	1	0	1	1	0	1	0

图 4.9　水平垂直奇偶校验码

水平垂直奇偶校验不仅可检错,还可用来纠正部分差错。例如数据块中仅存在 1 位错时,便能确定错码的位置就在某行和某列的交叉处，从而可以纠正它。

二维奇偶校验码检错能力强，又有一定的纠错能力，且实现容易，因而得到了广泛应用。

4.4　定比码

定比码的码字中 1 的数目与 0 的数目保持恒定比例，也称为恒比码。由于恒比码中，每个码组均含有相同数目的 1 和 0，因此恒比码又称等重码、定 1 码。这种码在检测时，只要计算接收码元中 1 的数目是否正确，就可判断有无差错。

我国电传通信用五位电码表示一位阿拉伯数字，再用四位表示一个汉字。电传通信普遍采用 3∶2 码，又称“5 中取 3”的定比码，即每个码组的长度为 5，其中 3 个“1”。这时可能编成的不同码组数目等于从 5 中取 3 的组合数 10，这 10 个许用码组恰好可表示 10 个阿拉伯数字。国际通用的 ARQ 电报通信系统采用“7 中取 3”的定比码。“7 中取 3”码可以检出所有的单比特差错和奇数个差错，但只能检出部分偶数位差错。

定比码比较简单，应用于电报、数据通信、计算机中，适合用在传输电传机或其他键盘设备产生的数字、字母和符号。

4.5　汉明码

汉明码是一种能够纠正单个随机错误的线性分组码，它是 1950 年由贝尔实验室的 R.W.Hamming 发明的。因其编译码器结构简单，故得到了广泛应用。

汉明码的特点：

（1）最小码距 $d_{\min}=3$，可以纠正一位错误。

（2）监督位数 $r=n-k$。

（3）信息位数 $k=2^r-r-1$。r 位监督位可以指示 2^r-1 个错误码元位置（当码元全部正确的时候用到一种情况）。对于码组长度为 n、信息码元为 k 位、监督码元为 $r=n-k$ 位的分组码，如果希望用 r 个监督位构造出 r 个监督关系来指示一个错码的 n 种可能，则要求 $2^r-1\geqslant n$ 或 $2^r\geqslant k+r+1$。

汉明码的检错、纠错基本思想是将有效信息按某种规律分成若干组，每组安排一个校验位进行奇偶性测试，然后产生多位检测信息，并从中得出具体的出错位置，最后通过对错误位取反（原来是 1 就变成 0，原来是 0 就变成 1）来将其纠正。

要采用汉明码纠错，须要按以下步骤来进行：确定校验位数→确定校验码位置→确定校验码→实现校验和纠错。下面来具体介绍这几个步骤。

（1）确定校验位数（监督位）。根据公式 $2^r-1\geqslant n$ 或 $2^r\geqslant k+r+1$ 来计算得出校验位数。

（2）确定校验码（监督位）的位置。将监督码元和信息码元的位置从左到右进行编号，1,2,3,4,5…其中 $2n$ 的位置就是校验码所在位置，其他位置就是信息码元的位置，可将信息码元从左到右依次填进去。如表 4.8 第一行第二行所示，第二行中 $d_1,d_2,d_3,\cdots,d_i,\cdots$ 就是信息码元，$p_1,p_2,p_4,p_8,p_{16}\cdots$ 就是监督码元。

（3）确定校验码（监督位）p_1,p_2,p_4,p_8,p_{16} 等。下列表格中 X 表示需要校验的位置，比如 p_1 是对第 3,5,7…等位置进行奇偶校验的监督码，p_2 是对 3,6,7,10,11 等进行偶校验的监督码，p_i 对从第 i 位开始校验 i 位，跳过 i 位校验 i 位…得到的偶校验码，如表 4.8 所示。

表 4.8　汉明码监督位与信息位关系

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
编码后数据位置		p_1	p_2	d_1	p_4	d_2	d_3	d_4	p_8	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	p_{16}	d_{12}	d_{13}	d_{14}	d_{15}
奇偶校验位覆盖率	p_1	X		X		X		X		X		X		X		X		X		X	
	p_2		X	X			X	X			X	X			X	X			X	X	
	p_4				X	X	X	X					X	X	X	X					X
	p_8								X	X	X	X	X	X	X	X					
	p_{16}																X	X	X	X	X

实例 4.2　对 1100 进行汉明编码，求编码后的码字。

解：计算监督位的位数 $k=2^r-r-1$，得出 r=3，码长 $n=k+r=7$

设编码后的码字为 $\boldsymbol{A}=(a_6,a_5,a_4,a_3,a_2,a_1,a_0)$，按照表 4.8 所示，$a_6,a_5,a_3$ 为监督位（采用偶校验），监督关系式为

$$\begin{cases}a_6+a_4+a_2+a_0=0\\a_5+a_4+a_1+a_0=0\\a_3+a_2+a_1+a_0=0\end{cases}$$

补充为下列方程：

$$\begin{cases} a_6 = a_4 + a_2 + a_0 \\ a_5 = a_4 + a_1 + a_0 \\ a_4 = a_4 \\ a_3 = a_2 + a_1 + a_0 \\ a_2 = a_2 \\ a_1 = a_1 \\ a_0 = a_0 \end{cases}$$

改写为矩阵形式：

$$\begin{bmatrix} a_6 \\ a_5 \\ a_4 \\ a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_4 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix}$$

根据上面的式子，将信息码 1100 带入上式 $\begin{bmatrix} a_4 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix}$，得到码字

$$\boldsymbol{A} = \begin{bmatrix} a_6 \\ a_5 \\ a_4 \\ a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix}^{\mathrm{T}} = [0\ 1\ 1\ 1\ 1\ 0\ 0]$$

因此，汉明编码后的码字为：0111100。

在接收端收到汉明码后，则把每个汉明码各校验码对它所校验的位组进行“异或运算”，即

$$G_1 = p_1 + d_1 + d_2 + d_4 + d_5 + \cdots$$

$$G_2 = p_2 + d_1 + d_3 + d_4 + d_6 + d_7 + d_{10} + d_{11} + \cdots$$

$$G_4 = p_4 + d_2 + d_3 + d_4 + d_8 + d_9 + d_{10} + d_{11} + \cdots$$

$$G_8 = p_8 + d_5 + d_6 + d_7 + d_8 + d_9 + d_{10} + d_{11} + \cdots$$

若各校验码采用偶（奇）校验，如果结果为 0 就是正确（1 则正确），为 1 则说明当前汉明码所对应的数据位中有错误（0 则错误），此时再通过其他校验位各自的运算来确定具体是哪个位出了问题。

假设接收端接收到汉明码为 0101100，G_1 = 0+0+1+0=1，G_2=1+0+0+0=1，G_4=1+1+1+0=0，从表 4.8 中可以看到，只有第三位即 d_1 出错，才会造成 G_1 和 G_2 同时出错。所以，正确的汉

明码为 0111100。

4.6 循环码

扫一扫看循环码、卷积码及 Turbo 码教学课件

循环码是线性分组码的一个重要子类，具有严密的代数学理论。循环码“线性”是指任意两个循环码模 2 相加所得的新码仍为循环码。循环码具有线性码的一般性质（即封闭性，指一种线性分组码的任意两个码组之和仍是该分组码的另一个码组）外，还具有循环性，即循环码中任意码组循环一位（将最右端码元移至左端，或反之）以后，仍为该码组中的一个码组。（n，k）循环码表示其中信息位为 k，监督位为 $n-k$ 位。

为了利用代数理论研究循环码，可以将码组用代数多项式来表示，这个多项式被称为码多项式，对于许用循环码 $\boldsymbol{A}=[a_{n-1},a_{n-2},\cdots,a_1,a_0]$，可以将它的码多项式表示为

$$A(x)=a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+\cdots+a_1x+a_0 \qquad (4\text{-}6\text{-}1)$$

对于二进制码组，多项式的每个系数不是 0 就是 1，x 仅是码元位置的标志。因此，这里并不关心 x 的取值。例如码组 101101 可以用码多项式 $x^5+x^3+x^2+1$ 来表示。

1．编码过程

在编码时，首先须要根据给定循环码的参数确定生成多项式 $g(x)$，也就是从 x^k+1 的因子中选一个$(n-k)$次多项式作为 $g(x)$；然后，利用循环码的编码特点，即所有循环码多项式 $A(x)$ 都可以被 $g(x)$整除，来定义生成多项式 $g(x)$。

根据上述原理可以得到一个较简单的系统：设要产生(n,k)循环码，$m(x)$表示信息多项式，循环码编码方法则其次数必小于 k，而 $x^{n-k}\cdot m(x)$ 的次数必小于 n，用 $x^{n-k}\cdot m(x)$ 除以 $g(x)$，可得余数 $r(x)$，$r(x)$的次数必小于$(n-k)$，将 $r(x)$加到信息位后作监督位，就得到了系统循环码。下面就将以上各步处理加以解释。

（1）用 x^{n-k} 乘 $m(x)$。这一运算实际上是把信息码后附加上$(n-k)$个“0”。例如，信息码为 110，它相当于 $m(x)=x^2+x$。当 $n-k=7-3=4$ 时，$x^{n-k}\cdot m(x)=x^6+x^5$，它相当于 1100000。而希望得到的系统循环码多项式应当是 $A(x)=x^{n-k}\cdot m(x)+r(x)$。

（2）求 $r(x)$。由于循环码多项式 $A(x)$都可以被 $g(x)$整除，也就是

$$\frac{A(x)}{g(x)}=\frac{x^{n-k}\cdot m(x)+r(x)}{g(x)}=\frac{x^{n-k}\cdot m(x)}{g(x)}+\frac{r(x)}{g(x)}$$

（3）因此，用 $x^{n-k}\cdot m(x)$ 除以 $g(x)$，就得到商 $Q(x)$和余式 $r(x)$，即

$$\frac{x^{n-k}\cdot m(x)}{g(x)}=Q(x)+\frac{r(x)}{g(x)}$$

这样就得到了 $r(x)$。

（4）编码输出系统循环码多项式 $A(x)$为

$$A(x)=x^{n-k}\cdot m(x)+r(x)$$

实例 4.3 已知循环码的生成多项式为 $G(x)=x^4+x^3+1$。若信息位为 11011110 时，写出它的监督码和码组。

解：$g(x)$的最高次幂是 $n-k=4$，$m(x)$的最高次幂是 7。

（1）用 x^{n-k} 乘以信息码多项式 $m(x)$得到