



教育部高等学校电子信息类专业教学指导委员会规划教材
高等学校电子信息类专业系列教材

信息与通信工程

Fundamental Introduction to Information
Theory and Coding (2nd Edition)

信息论与编码 基础教程

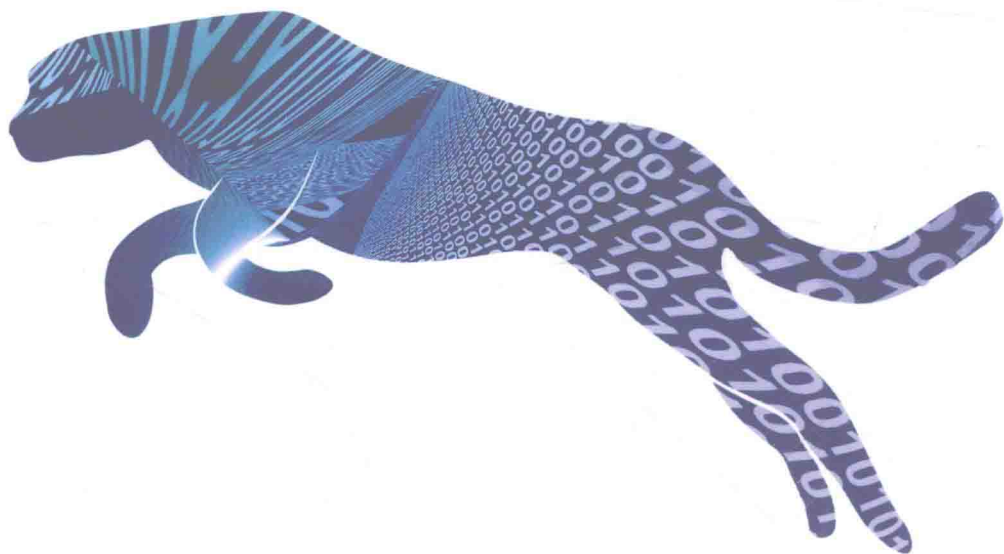
(第2版)

孙海欣, 张猛 张丽英 编著

Sun Haixin

Zhang Meng

Zhang Liying



清华大学出版社





教育部高等学校电子信息类专业教学指导委员会规划教材
高等学校电子信息类专业系列教材

Fundamental Introduction to Information Theory and Coding
(2nd Edition)

信息论与编码 基础教程

(第2版)

孙海欣 张猛 张丽英 编著

Sun Haixin

Zhang Meng

Zhang Liying

清华大学出版社
北京

内 容 简 介

本书以香农信息论为基础,对信息论基础和编码理论进行介绍。系统介绍了信息论的3个基本概念(信源熵、信道容量和信息率失真函数),以及无失真信源编码、限失真信源编码、信道编码和加密编码中的理论知识及其实现原理。为了便于教学和加深对概念的理解,每章的最后还附有小结和习题。此外,本书配有相应的教学课件。

本书概念清晰、重点突出,用通俗易懂的文字和图形图表阐述基本概念、基本理论及实现原理,既有必要的数学分析,又强调物理概念的理解。本书既可作为普通高等院校电子信息类、通信工程、信息与计算科学及相关专业的本科生教材,也可作为从事相关专业的科研和工程技术人员的参考用书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

信息论与编码基础教程/孙海欣,张猛,张丽英编著.—2版.—北京:清华大学出版社,2017
(高等学校电子信息类专业系列教材)
ISBN 978-7-302-46046-6

I. ①信… II. ①孙… ②张… ③张… III. ①信息论—高等学校—教材 ②信源编码—高等学校—教材 IV. ①TN911.2

中国版本图书馆CIP数据核字(2016)第306333号

责任编辑:梁颖 柴文强

封面设计:李召霞

责任校对:梁毅

责任印制:杨艳

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印装者:三河市吉祥印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:12.5 字 数:305千字

版 次:2010年4月第1版 2017年2月第2版 印 次:2017年2月第1次印刷

印 数:1~2000

定 价:29.00元

产品编号:060024-01

高等学校电子信息类专业系列教材

顾问委员会

谈振辉	北京交通大学 (教指委高级顾问)	郁道银	天津大学 (教指委高级顾问)
廖延彪	清华大学 (特约高级顾问)	胡广书	清华大学 (特约高级顾问)
华成英	清华大学 (国家级教学名师)	于洪珍	中国矿业大学 (国家级教学名师)
彭启琮	电子科技大学 (国家级教学名师)	孙肖子	西安电子科技大学 (国家级教学名师)
邹逢兴	国防科学技术大学 (国家级教学名师)	严国萍	华中科技大学 (国家级教学名师)

编审委员会

主任	吕志伟	哈尔滨工业大学		
副主任	刘旭	浙江大学	王志军	北京大学
	隆克平	北京科技大学	葛宝臻	天津大学
	秦石乔	国防科学技术大学	何伟明	哈尔滨工业大学
	刘向东	浙江大学		
委员	王志华	清华大学	宋梅	北京邮电大学
	韩焱	中北大学	张雪英	太原理工大学
	殷福亮	大连理工大学	赵晓晖	吉林大学
	张朝柱	哈尔滨工程大学	刘兴钊	上海交通大学
	洪伟	东南大学	陈鹤鸣	南京邮电大学
	杨明武	合肥工业大学	袁东风	山东大学
	王忠勇	郑州大学	程文青	华中科技大学
	曾云	湖南大学	李思敏	桂林电子科技大学
	陈前斌	重庆邮电大学	张怀武	电子科技大学
	谢泉	贵州大学	卞树檀	第二炮兵工程大学
	吴瑛	解放军信息工程大学	刘纯亮	西安交通大学
	金伟其	北京理工大学	毕卫红	燕山大学
	胡秀珍	内蒙古工业大学	付跃刚	长春理工大学
	贾宏志	上海理工大学	顾济华	苏州大学
	李振华	南京理工大学	韩正甫	中国科学技术大学
	李晖	福建师范大学	何兴道	南昌航空大学
	何平安	武汉大学	张新亮	华中科技大学
	郭永彩	重庆大学	曹益平	四川大学
	刘缠牢	西安工业大学	李儒新	中科院上海光学精密机械研究所
	赵尚弘	空军工程大学	董友梅	京东方科技集团
	蒋晓瑜	装甲兵工程学院	蔡毅	中国兵器科学研究院
	仲顺安	北京理工大学	冯其波	北京交通大学
	黄翊东	清华大学	张有光	北京航空航天大学
	李勇朝	西安电子科技大学	江毅	北京理工大学
	章毓晋	清华大学	张伟刚	南开大学
	刘铁根	天津大学	宋峰	南开大学
	王艳芬	中国矿业大学	靳伟	香港理工大学
	苑立波	哈尔滨工程大学		
丛书责任编辑	盛东亮	清华大学出版社		

序

FOREWORD

我国电子信息产业销售收入总规模在 2013 年已经突破 12 万亿元,行业收入占工业总体比重已经超过 9%。电子信息产业在工业经济中的支撑作用凸显,更加促进了信息化和工业化的高层次深度融合。随着移动互联网、云计算、物联网、大数据和石墨烯等新兴产业的爆发式增长,电子信息产业的发展呈现了新的特点,电子信息产业的人才培养面临着新的挑战。

(1) 随着控制、通信、人机交互和网络互联等新兴电子信息技术不断发展,传统工业设备融合了大量最新的电子信息技术,它们一起构成了庞大而复杂的系统,派生出大量新兴的电子信息技术应用需求。这些“系统级”的应用需求,迫切要求具有系统级设计能力的电子信息技术人才。

(2) 电子信息系统设备的功能越来越复杂,系统的集成度越来越高。因此,要求未来的设计者应该具备更扎实的理论基础知识和更宽广的专业视野。未来电子信息系统的设计越来越要求软件和硬件的协同规划、协同设计和协同调试。

(3) 新兴电子信息技术的发展依赖于半导体产业的不断推动,半导体厂商为设计者提供了越来越丰富的生态资源,系统集成厂商的全方位配合又加速了这种生态资源的进一步完善。半导体厂商和系统集成厂商所建立的这种生态系统,为未来的设计者提供了更加便捷却又必须依赖的设计资源。

教育部 2012 年颁布了新版《高等学校本科专业目录》,将电子信息类专业进行了整合,为各高校建立系统化的人才培养体系,培养具有扎实理论基础和宽广专业技能的、兼顾“基础”和“系统”的高层次电子信息人才给出了指引。

传统的电子信息学科专业课程体系呈现“自底向上”的特点,这种课程体系偏重对底层元器件的分析与设计,较少涉及系统级的集成与设计。近年来,国内很多高校对电子信息类专业课程体系进行了大力度的改革,这些改革顺应时代潮流,从系统集成的角度,更加科学合理地构建了课程体系。

为了进一步提高普通高校电子信息类专业教育与教学质量,贯彻落实《国家中长期教育改革和发展规划纲要(2010—2020 年)》和《教育部关于全面提高高等教育质量若干意见》(教高【2012】4 号)的精神,教育部高等学校电子信息类专业教学指导委员会开展了“高等学校电子信息类专业课程体系”的立项研究工作,并于 2014 年 5 月启动了《高等学校电子信息类专业系列教材》(教育部高等学校电子信息类专业教学指导委员会规划教材)的建设工作。其目的是为推进高等教育内涵式发展,提高教学水平,满足高等学校对电子信息类专业人才培养、教学改革与课程改革的需要。

本系列教材定位于高等学校电子信息类专业的专业课程,适用于电子信息类的电子信

息工程、电子科学与技术、通信工程、微电子科学与工程、光电信息科学与工程、信息工程及其相近专业。经过编审委员会与众多高校多次沟通,初步拟定分批次(2014—2017年)建设约100门课程教材。本系列教材将力求在保证基础的前提下,突出技术的先进性和科学的前沿性,体现创新教学和工程实践教学;将重视系统集成思想在教学中的体现,鼓励推陈出新,采用“自顶向下”的方法编写教材;将注重反映优秀的教学改革成果,推广优秀的教学经验与理念。

为了保证本系列教材的科学性、系统性及编写质量,本系列教材设立顾问委员会及编审委员会。顾问委员会由教指委高级顾问、特约高级顾问和国家级教学名师担任,编审委员会由教育部高等学校电子信息类专业教学指导委员会委员和一线教学名师组成。同时,清华大学出版社为本系列教材配置优秀的编辑团队,力求高水准出版。本系列教材的建设,不仅有众多高校教师参与,也有大量知名的电子信息类企业支持。在此,谨向参与本系列教材策划、组织、编写与出版的广大教师、企业代表及出版人员致以诚挚的感谢,并殷切希望本系列教材在我国高等学校电子信息类专业人才培养与课程体系建设中发挥切实的作用。

吕志伟 教授

前言

PREFACE

“信息论与编码”是一门理论性很强的课程,是电子信息类学科及相关专业的必修课程之一,其主要内容包括信息论基础和编码理论。“信息论与编码”指出了通信工程的一般性规律和理论极限,它对实际通信系统的设计产生了深刻的影响。由于其中涉及大量的数学分析、论证和建模,对于学生特别是本科生而言,是具有一定难度的。针对这种情况,我们编写了《信息论与编码基础教程》。根据学生对以往的信息论与编码类教材的反馈意见及教学中的体会及心得,本书对《信息论与编码基础教程》进行再版,遵照由浅入深、循序渐进的教学规律,系统地组织本书的内容。

本书概念清晰、重点突出,在编写过程中强调基本原理的理解,选材时充分考虑其实用性,把信息论涉及的数学知识限制在工科高等数学和工科数学的范畴内,尽量用通俗易懂的文字和图形图表阐述基本概念、基本理论及实现原理,既有必要的数学分析,又强调物理概念的理解。全书内容包括7章:

第1章绪论,介绍了信息论与编码的基本概念,通信系统的模型及信息论与编码的研究内容和意义;

第2章信源及信源熵,介绍信息论的一些基本概念,包括离散信源、连续信源、自信息量、信源熵、条件熵、联合熵、熵的基本性质、离散序列熵、连续信源熵及最大熵定理、冗余度等,对信源的信息测度给出定量的描述;

第3章信道与信道容量,介绍信道的基本概念,包括信道的分类、数学模型及参数、互信息、平均互信息、数据处理中信息的变化、信道容量的概念及计算、串联信道和并联信道的信道容量及信源与信道的匹配等;

第4章信息率失真函数,介绍失真函数和信息率失真函数的定义及性质,给出了在一定失真限度内信源能输出的最小传输速率;

第5章信源编码,介绍信源编码的基本概念、码的分类、无失真信源编码定理、常用的无失真信源编码方法,包括香农码、费诺码、哈夫曼码、游程编码、算术编码等,以及限失真信源编码定理及方法;

第6章信道编码,介绍信道编码的基本概念与分类、信道编码的基本参数、纠错编码的基本原理、线性分组码的基本概念,包括生成矩阵和一致校验矩阵、伴随式与标准阵列译码、线性分组码的纠错能力、汉明码等,循环码的基本概念,包括循环码的多项式描述及构造方法、生成矩阵和校验矩阵、编译码电路及常用的循环码,卷积码的基本概念和描述方法、维特比译码算法及卷积码的性能限。

第7章加密编码,介绍加密编码的基本概念及典型的加密算法,还引入信息安全性的概念及相关技术。

本书由孙海欣主编,其中第1章由张丽英编写,第2章、第4章、第5章和第6章由孙海欣编写,第3章和第7章由张猛编写。全书由孙海欣统稿。在编写过程中,本书的插图得到了王守利、王浩的大力帮助,在此表示衷心的感谢。

由于编者的学术水平有限,书中难免有不妥或谬误之处,殷切希望读者批评指正。

编 者

2017年元月

目录

CONTENTS

第 1 章 绪论	1
1.1 信息论与编码的基本概念	1
1.1.1 信息的一般概念	1
1.1.2 信息论与编码的发展历程	3
1.2 通信系统的模型	4
1.3 信息论与编码的研究内容与意义	5
1.3.1 信息论研究的主要内容	5
1.3.2 香农信息论对信道编码的指导意义	6
1.3.3 香农信息论对信源编码的指导意义	6
习题	7
第 2 章 信源及信源熵	8
2.1 信源的数学模型及分类	8
2.1.1 离散无记忆信源	9
2.1.2 离散有记忆信源	10
2.1.3 马尔可夫信源	11
2.1.4 连续信源	14
2.2 离散信源熵	15
2.2.1 自信息量	15
2.2.2 信源熵	16
2.2.3 条件熵和联合熵	18
2.2.4 熵的基本性质	20
2.3 离散序列信源熵	21
2.3.1 离散无记忆信源的序列熵	21
2.3.2 离散有记忆信源的序列熵	22
2.3.3 马尔可夫信源的序列熵	23
2.4 连续信源熵	24
2.4.1 幅度连续的单个符号信源熵	25
2.4.2 波形信源熵	26
2.4.3 最大熵定理	27
2.5 冗余度	27
本章小结	29
习题	30

第 3 章 信道与信道容量	32
3.1 信道的基本概念.....	32
3.1.1 信道的分类	32
3.1.2 信道的数学模型及参数	33
3.2 互信息.....	36
3.2.1 互信息与平均互信息	36
3.2.2 平均互信息的性质	38
3.2.3 数据处理中信息的变化	39
3.3 离散符号信道容量.....	40
3.3.1 信息传输率与信道容量	40
3.3.2 无干扰离散信道	41
3.3.3 对称 DMC 信道	42
3.3.4 准对称 DMC 信道	43
3.3.5 一般 DMC 信道	45
3.4 离散序列信道及容量.....	46
3.5 串联信道和并联信道的信道容量.....	47
3.5.1 串联信道	47
3.5.2 并联信道	49
3.6 连续信道及其容量.....	49
3.6.1 连续单符号加性信道	50
3.6.2 多维无记忆加性连续信道	51
3.6.3 加性高斯白噪声波形信道	54
3.7 信源与信道的匹配.....	56
本章小结	58
习题	59
第 4 章 信息率失真函数	61
4.1 失真函数.....	61
4.1.1 失真函数(失真度)	61
4.1.2 平均失真	63
4.2 信息率失真函数及其性质.....	64
4.2.1 信息率失真函数 $R(D)$	64
4.2.2 信息率失真函数的性质	64
4.2.3 信息率失真函数与信道容量的比较	67
4.3 无记忆信源信息率失真函数的计算.....	67
4.3.1 离散无记忆信源的 $R(D)$ 计算	67
4.3.2 连续信源的 $R(D)$ 计算	70
本章小结	70
习题	71
第 5 章 信源编码	73
5.1 编码的基本概念.....	74
5.1.1 码的定义及分类	74
5.1.2 码树	75

5.2	无失真信源编码定理	76
5.2.1	定长编码定理	77
5.2.2	变长编码定理	79
5.3	无失真信源编码的常用方法	81
5.3.1	香农码	81
5.3.2	费诺码	82
5.3.3	哈夫曼码	83
5.3.4	游程编码	86
5.3.5	算术编码	88
5.4	限失真信源编码定理及方法	92
5.4.1	保真度准则下信源编码定理	93
5.4.2	矢量量化	93
5.4.3	预测编码	94
5.4.4	变换编码	95
	本章小结	98
	习题	99
第6章	信道编码	102
6.1	信道编码概述	102
6.1.1	信道编码的基本概念与分类	103
6.1.2	信道编码的基本参数	106
6.1.3	纠错编码的基本原理	106
6.1.4	最优译码与最大似然译码	108
6.2	线性分组码	110
6.2.1	基本概念	110
6.2.2	生成矩阵和一致校验矩阵	112
6.2.3	伴随式与标准阵列译码	116
6.2.4	线性分组码的纠错能力	119
6.2.5	汉明码	121
6.3	循环码	122
6.3.1	循环码的多项式描述及构造方法	122
6.3.2	循环码的生成矩阵和校验矩阵	125
6.3.3	循环码的编码电路	127
6.3.4	循环码的译码电路	130
6.3.5	常用的循环码	133
6.4	卷积码	137
6.4.1	卷积码的基本概念和描述方法	138
6.4.2	卷积码的最大似然译码——维特比算法	144
6.4.3	卷积码的性能限	149
	本章小结	151
	习题	152
第7章	加密编码	155
7.1	加密编码概述	155
7.1.1	加密编码的基本概念	155

7.1.2	加密编码算法分类	158
7.1.3	密码系统的安全性及其分类	158
7.1.4	加密编码中的熵概念	160
7.2	数据加密标准	161
7.2.1	换位和替代密码	162
7.2.2	DES 密码算法	163
7.2.3	DES 密码的改进	167
7.3	公开密钥加密法	169
7.3.1	公开密钥密码体制	169
7.3.2	RSA 密码体制	170
7.3.3	报文摘要	173
7.3.4	公开密码体制的优缺点	176
7.4	信息安全和确认技术	176
7.4.1	信息安全的基本概念	177
7.4.2	数字签名	177
7.4.3	防火墙	181
7.4.4	常用的信息安全技术应用实例	182
	本章小结	185
	习题	186
	参考文献	187

科学技术的发展使人类进入了高速发展的信息化时代,特别是 20 世纪后半叶,计算机技术、微电子技术、通信技术的迅猛发展及相关设备的迅速更新换代,极大地提高了人们处理信息、存储信息、控制和管理信息的能力,信息的重要性不言而喻。作为现代科学技术基础理论之一的信息论在各个领域的应用和推广使许多经典的概念有了全新的解释,使过去曾经不确切的描述有了精确的定量分析方法。

信息论是人们在长期通信工程的实践中,由通信技术、概率论、随机过程和数理统计等学科相结合而逐步发展起来的一门学科。它主要研究信息、信息熵、通信系统、数据传输、密码学、数据压缩等问题。信息论是信息科学的主要理论基础之一。它主要研究可能性和存在性问题,为具体实现提供理论依据。

信息论在学术界引起了巨大的反响,在香农信息论的指导下,为提高通信系统信息传输的有效性和可靠性,人们在信源编码和信道编码两个领域进行了卓有成效的研究,取得了丰硕的成果。随着信息理论的迅猛发展和信息概念的不断深化,信息论所涉及的内容早已超越了通信工程的范畴,进入了信息科学这一更广、更新的领域,并渗透到许多学科,得到了多个领域的重视。

本章主要讨论以下内容:

- 信息的定义及特征,信息论与编码理论的发展历史;
- 通信系统的模型及其各部分的作用;
- 信息论的主要研究内容;
- 信息论对信源编码和信道编码研究的指导意义。

1.1 信息论与编码的基本概念

1.1.1 信息的一般概念

当今社会,人们在各种生产、科学研究和社会活动中,无处不涉及信息的交换和利用。可以说,在我们周围充满了信息,我们正处于“信息社会”中。通过电话、电报、传真和电子邮件,人们可以自由地交流信息;通过报纸、书刊、电子出版物和因特网等媒介,人们可以有选择地获取大量信息;通过电台、电视台等视听媒体,人们可以“身临其境”地感受最新信息。但以上所述还远不能概括信息的全部含义:四季交替透露的是自然界的消息,而牛顿定律提示的是物体运动内在规律的信息。信息含义之广几乎可以涵盖整个宇宙,且内容庞杂,层

次混叠,不易理清。现代的信息技术使得我们能够快速、有效地获取有价值的信息,而信息的价值则体现在社会生活的方方面面,可以给我们带来无尽的益处。因此,迅速获取信息,正确处理信息,充分利用信息,既能促进科学技术和国民经济的飞跃发展,又能在各种形式的竞争中占得先机。

如今有关信息的新名词、新术语层出不穷,信息产业在社会经济中所占的份额也越来越大,信息基础设施建设速度之快成了当今社会的重要特征之一,物质、能源、信息构成了现代社会生存发展的三大基本支柱。

信息的价值在于它为人们能动地改造外部世界提供了可能,信息所揭示的事物运动规律为人们应用这些规律提供了可能,而信息所描述的事物状态也为人们推动事物向有利的方向发展提供了可能。掌握的资源和能量越多,面对同样的信息时人们能用以改造世界的可能性也越大。今天我们所掌握的物质力量比过去增大了不知多少倍,因此,信息对于当今社会的发展和人们生活的重要性较之几百年前、几十年前甚至十几年前都有很大的提高。这是信息社会的一个重要特征。

信息的重要性不言而喻,那么,如此神通广大、无处不在而又无所不能的信息究竟是什么?

信息是信息论中最基本、最重要的概念,既抽象又复杂。关于信息的科学定义,到目前为止,国内外已提出近百种,它们从不同的侧面和不同的层次来揭示信息的本质。从本质的意义上说,信息是人类社会活动所产生的各种状态和消息的总称,信息是人们对客观事物运动规律及其存在状态的认识。

在信息论和通信理论中会经常遇到信息、消息和信号这3个既有联系又有区别的名词,下面对它们进行定义和比较。

信息是指各个事物运动的状态及状态变化的方式。人们从对周围世界的观察中获取信息。信息是抽象的意识或知识,它是看不见、摸不着的,而且信息仅仅与随机事件的发生相关,非随机事件的发生不包含任何信息。也就是说,信息量的大小与随机事件发生的概率有直接的关系,概率越小的随机事件一旦发生,它所包含的信息量就越大;而概率越大的随机事件一旦发生,它所包含的信息量就越小。

消息是信息的载体。它是指包含信息的语音、文字、数字和语言等。例如我们每天从广播节目、报纸和电视节目中获得的各种新闻及其他消息。在通信中,消息是指担负着传送信息任务的单个符号或符号序列。可见,消息是具体的,它载荷信息,但它不是物理性的。信息只与随机事件的发生有关。每时每刻在世界上的每个地方,都会有各种事件发生,这些事件的发生绝大多数是随机的,即这些随机事件的消息中含有信息;如果事件的发生是确定的,那么该消息中就不含信息,该消息的传输也就失去了意义。

信号是消息的物理体现。为了在信道上传输消息,就必须把消息加载(调制)到具有某种物理特征的信号上去。信号是信息的载体,是具有物理性的,如电信号、光信号、声信号等。

通信系统传送的本质内容是信息,发送端需将信息表示成具体的消息,再将消息载至信号上,才能在实际的通信系统中传输。信号到了接收端(信息论里称为信宿)经过处理变成文字、语音或图像等形式的消息,人们再从中得到有用的信息。在接收端将含有噪声的信号经过各种处理和变换,从而获取有用信息的过程就是信息提取,提取有用信息的过程或方法

主要有检测和估计两类。载有信息的可观测、可传输、可存储及可处理的信号,均称为数据。作为系统设计人员,我们所接触的只是信号,而这种信号最终要变成消息的形式才能被大众所接受。

信息的基本概念在于它的不确定性,任何已确定的事物都不含有信息。信息具有以下特征:

- (1) 接收者在收到信息之前,对其内容是未知的,所以信息是新知识、新内容。
- (2) 信息是能使认识主体对某一事物的未知性或不确定性减少的有用知识。
- (3) 信息可以产生,也可以消失,同时信息可以被携带、存储及处理。
- (4) 信息是可以被度量的。

1.1.2 信息论与编码的发展历程

信息论从诞生至今已经历半个多世纪,目前已成为一门独立的学科。而编码理论与技术研究也有半个世纪的历史,并从刚开始作为信息论的一个组成部分逐步发展成为较完善的独立体系。回顾它们的发展历史,我们可以清楚地看到该理论是如何在实践中经过抽象、概括、提高而逐步形成和发展的。

信息论与编码理论是在长期的通信工程实践和理论研究的基础上发展起来的。最早对信息进行科学定义的是哈特莱(L. V. R. Hartley),他在1928年发表的《信息传输》一文中,首先提出“信息”这一概念。他认为,发信者所发出的信息,就是他在通信符号表中选择符号的具体方式,并主张用所选择的自由度来测度信息。哈特莱的这种理解在一定程度上能够解释通信工程中的一些信息问题,但它存在没有考虑各种可能选择方法的统计特性的局限,正是这些限制了它的适应范围。

1948年,香农在发表的著名论文 *A mathematical theory of communication* (通信的数学理论)中提出了信息量的概念和信息熵的计算方法。他用概率测试和数理统计的方法系统地讨论了通信的基本问题,得出了几个重要而带有普遍意义的结论。该论文是一篇代表现代信息论的开创性的权威论文,为现代信息论奠定了基础,香农也因此成为信息论的奠基人。

在香农等人的研究基础上,通信技术界的科学工作者的主要精力转到信源编码和信道编码的实现技术上,并在这方面取得了稳步的发展。

20世纪60年代信道编码技术有了较大进展,成为信息论的又一重要分支。汉明首先做出了贡献,提出了一种重要的线性分组码——汉明码,此后人们把代数方法引入到纠错码的研究,形成了代数编码理论,并使分组码技术的发展达到了高峰,找到了大量可纠正多个错误的码,而且提出了可实现的译码方法。20世纪70年代,卷积码和概率译码有了重大突破,提出了序列译码和Viterbi译码方法,并被美国卫星通信系统采用,这使香农理论成为真正具有实用意义的科学理论。1982年G. Underboeck提出了将信道编码和调制结合在一起的网格编码调制方法,这种方法无需增大带宽和功率,以增加设备的复杂度换取编码增益,受到了广泛关注,在目前的通信系统中占据统治地位。

信源编码的研究落后于信道编码。香农首先提出了无失真信源编码定理,也给出了简单的编码方法。1952年,费诺和哈夫曼分别提出了各自的编码方法,并证明其方法都是最佳编码法。1959年香农在 *Coding theorems for a discrete source with a fidelity criterion*

(保真度准则下的离散信源编码定理)中系统地提出了信息率失真理论和限失真信源编码定理。这两个理论是数据压缩的数学基础,为各种信源编码的研究奠定了基础。20世纪60年代到70年代,人们开始将各种正交变换用于信源压缩编码,先后提出DFT、DCT、WHT、ST、KLT等多种变换,综合性能最好的是DCT变换,目前已被多种图像压缩国际标准用作主要压缩手段,得到了极为广泛的应用。除了经典的信源压缩编码方法的研究外,20世纪90年代初开始,针对图像类信源的特点,人们提出了多种压缩编码方法,包括小波变换编码、分形编码、模型编码等。这些方法可以有效地消除图像信源的各种冗余,并且还有很大的发展空间,有关其实际应用问题,还在继续探讨之中。

1.2 通信系统的模型

典型的通信系统物理模型如图1-1所示。下面介绍模型中各个部分的作用及需要研究的核心问题。

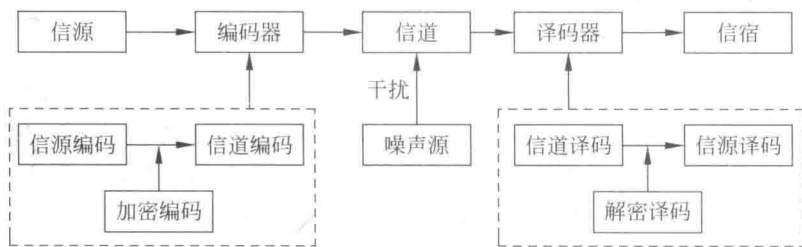


图 1-1 通信系统模型

1. 信源

信源是产生消息或消息序列的源头。信源本身十分复杂,它可以是人、生物或机器,在信息论中我们仅对信源的输出进行研究。信源输出的是以符号形式出现的具体消息,通常归纳为两类:离散消息,例如由字母、文字、数字等符号组成的符号序列,或者单个符号;连续消息,例如语音、图像以及在时间上连续变化的电参数等。

信源发出的消息是随机的、不确定的,但又有一定的规律性。因此,信源的核心问题是它包含的信息到底有多少,怎样将信息定量地表示出来,即如何确定信息量。

2. 编码器

一般地说,通信系统的性能指标主要是有效性、可靠性、安全性和经济性。通信系统优化就是使这些指标达到最佳。除了经济性外,以上指标都是信息论的研究对象,可以通过各种编码处理使通信系统的性能最优化。根据信息论的各种编码定理和上述通信系统的指标,编码器可分类3类:信源编码、信道编码和加密编码。

(1) 信源编码。信源编码器的作用有两个:①把信源发出的消息变换成由二进制(或多进制)码元组成的码组,即基带信号;②通过信源编码压缩信源的冗余度,从而提高通信系统传输消息的效率。信源编码可分为无失真信源编码和限失真信源编码。前者适用于离散信源或数字信号;后者主要用于连续信源或模拟信号。从提高通信系统的有效性意义上说,信源编码器的主要指标是它的编码效率,即理论上所需的码率与实际达到的码率之比。一般来说,效率越高,编码器的代价也越大。

(2) 信道编码。信道编码器的作用是在信源编码器输出的码组上有目的地增加监督码元,使之具有检错或纠错的能力。信道编码包括调制和纠错检错编码。信道中的干扰常使通信质量下降,对于模拟信号,表现在收到的信号的信噪比下降;对于数字信号,就是误码率增大。信道编码的主要方法是增大码率或频带,即增大所需的信道容量。

(3) 加密编码。加密编码是研究如何隐蔽消息中的信息内容,以便在传输过程中不被窃听,提高通信系统的安全性。

在实际问题中,上述3类编码应统一考虑,以提高通信系统的性能。这些编码的目标往往是互相矛盾的。提高有效性必须去除信源符号中的冗余部分,此时信道误码会使接收端不能恢复原来的信息,这就需提高传送的可靠性,不然会使通信质量下降;反之,为了提高可靠性而采用信道编码,往往需增加码长,也就降低了有效性。安全性也有类似情况。从理论上说,若能把3种编码合并成一种编码来编译,即同时考虑有效性、可靠性和安全性,可使编译码器更理想化,在经济上也能更优越。但从理论上和技术上的复杂性看,要取得有用的结果,还是相当困难的。

3. 信道和噪声源

信道是传递消息的通道,又是传送物理信号的设施。信道可以是导线、同轴电缆、传输电磁波的空间、光导纤维等传输信号的介质。信息在信道中传输时,不可避免受到噪声和干扰的影响。对于任何通信系统,干扰的性质和大小是影响系统性能的重要因素。信道的研究问题是它能够传送多少信息,即信道容量的大小。

4. 译码器

译码是把信道输出的编码信号进行反变换,以获得解码的消息。与编码器相对应,译码器也可分成信源译码、解密译码和信道译码。

5. 信宿

信宿是消息传递的对象,即接收消息的人或机器。信宿接收的消息与信源发出的消息可以相同,也可以不同。因此,信宿需要研究的问题是能收到或提取多少信息。

1.3 信息论与编码的研究内容与意义

1.3.1 信息论研究的主要内容

目前,关于信息论研究的内容一般有以下3种理解:

1. 狭义信息论

狭义信息论是在信息可以量度的基础上,对如何有效、可靠地传递信息进行研究的科学,它涉及信息的度量、信息特性、信息传输速率、信道容量以及信源和信道编码理论等问题。这部分内容是信息论的基础理论,又称为香农信息论或经典信息论。

2. 一般信息论

一般信息论主要研究信息传输和处理问题。除了香农信息论以外,还包括噪声理论、信号滤波和预测理论、统计检测与估计理论、调制理论、信息处理理论和保密理论等。

3. 广义信息论

广义信息论不仅包括上述两方面的内容,而且包括所有与信息相关的自然和社会领域,如模式识别、计算机处理理论、心理学、遗传学、神经生理学、语言学、语义学,甚至包括社会