

# Learning experience on applying car model into the course of information theory and coding

SONG Zhen-zhen, YU Lin-chen, ZHENG Chang, WANG Mao-cai

China University of Geosciences (Wuhan), Wuhan, 430074, China

linechenyu@yahoo.com.cn

**Abstract:** Taking Shannon's information theory as a representative, information theory plays an important role in the engineering of modern information transmission. However, it is too abstract for students to understand. The analogies between the process of information transmission and car's goods transmission are proposed. Based on this, it has brought up a car model. Car model can explain theories in the transmission of information clearly. It will show that the teaching quality is improved and the good teaching effect is implemented by car model.

**Keywords:** information theory and coding; information transmission model; car model

## 小车模型应用在信息论与编码课程学习中的一点体会

宋珍珍, 余林琛, 郑 倡, 王茂才

中国地质大学计算机学院 武汉 430074

linechenyu@yahoo.com.cn

**【摘要】**以香农三大定理为代表的信息论是现代信息传输工程的重要理论基础,但因为过于抽象,在平时的学习过程中非常难以理解。本文利用类比的方法,将信息传输的过程看作是小车运输货物的过程,通过小车模型来解释信息传输过程中的各种理论,极大地促进信息论课程的教学与学习效果。

**【关键词】**信息论与编码;信息传输模型;小车模型

### 1 引言

信息论与编码是信息安全专业的一门重要的学科基础课程,它围绕香农三大定理(无失真信源编码定理、有噪信道编码定理和保真度准则下的信源编码定理),研究和分析了信息从信源到信宿传输过程中的有效性、可靠性和安全性问题<sup>[1,2]</sup>。但课程中的一些概念较抽象,脱离本科生的实践经验,也不符合大学生的认知规律,使得学生在学习过程中无法形成正确的认识和理解。为了解决这个问题,笔者结合自己在信息论与编码课程教学与学习中的经验与体会,信息论与编码中的抽象概念实例化,构造一个小车运输货物的模型来对信息论中的各种理论进行形象化地辅助说明,提高学习效果。

本文受中国地质大学(武汉)2009年度教学研究项目(No. 2009B20)和2010年度教学研究项目(No. 2010B34)资助。

宋珍珍,郑倡,中国地质大学(武汉)计算机学院信息安全专业本科生。余林琛(1977-),本文通信作者,女,讲师,博士研究生,主要研究方向:网络与信息安全;王茂才(1974-),男,讲师,博士,主要研究方向:网络与信息安全、公钥密码理论等。

### 2 基于“小车模型”的信息论与编码相关概念

#### 2.1 “小车模型”与通信系统传输模型的对应关系

信息论的研究对象是如图1所示的通信系统模型,人们通过系统中消息的传输与处理过程来研究信息的传输和处理<sup>[3,4]</sup>。在该模型中,消息首先从信源输出,将信息带出信源,然后经过编码器上载到信号,再通过信号在信道中的传输将信息带到译码器。译码器接收到信号以后,从信号中提取出信源发送的消息,传给信宿。信宿接收到消息,如果信息足够,则信息传输成功。

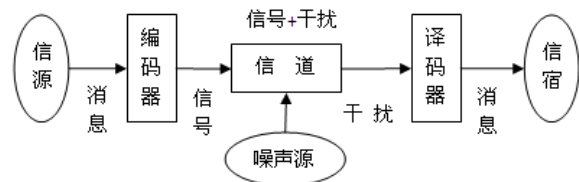


Figure 1. Communication system model

图1. 通信系统模型

将上述过程实例化为小车运输货物的过程，就构造出如图 2 所示的小车模型。在“小车模型”中，货箱首先从发货方出发，将货物带出发货方，然后经过中转站转载到小车上，再通过小车在公路上的运输将货物带到接收站。接收站接收到小车以后，从小车上卸载下发货方发送的货物，送给收货方。收货方接收到货箱，如果货物足够，则货物传输成功。

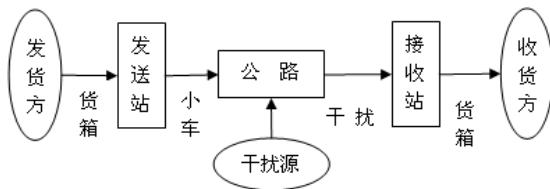


Figure 2. Car model  
图 2. 小车模型

小车模型，有助于讲解信息论中的“信源编码”、“信道容量”和“率失真函数”等方面的知识，增强学生对相关概念的理解。

### 2.2 结合“小车模型”讲述信源编码

信源编码就是编码器在信源符号从信源输出以后对其所载荷信息进行重编码，目的是减少信源输出符号序列中的冗余度，提高通信的有效性。在学习中，可以结合“小车模型”来讲述信息论中的有关知识。

①信息熵  $H(S)$ ：信息熵是信源  $S$  中每个符号的平均信息量，在“小车模型”中表示每一个货箱的平均载货量。

②极大值  $\log r$ ： $r$  个码元符号在等分布时，平均每个码符号能携带的信息量获得极大值  $\log r$ ，在“小车模型”中表示小车的最大载货量。

③最佳码平均码长  $\bar{L}$ ：

$$H(S)/\log r \leq \bar{L} < 1 + H(S)/\log r$$

在“小车模型”中就可以实例化为每运输一个货箱的货物所需要的最少小车数。

④平均每个码元所携带的信息量  $H(X)$ ，即编码后信道的信息传输率，在“小车模型”中表示每一辆小车的实际载货量。

⑤冗余度  $\gamma$ ： $\gamma = 1 - H(X)/\log r$  表示某一符号集内各符号的载荷信息能力的相对剩余量，在“小车模型”中表示小车装载货物的相对剩余空间。

假设每箱货物的重量均为  $H(S) = 1.5$  吨/箱，现用最大载重量  $\log r = 1$  吨/车的小车来运输。此时小车运输货物的过程就相当于对熵  $H(S) = 1.5 \text{bit/符号}$  的信源进行二进制编码。如果编码对象为  $N = 1$  的单符号，则相当于每个货箱分开运输，则每个货箱至少需要  $H(S)/\log r = 2$  辆小车，此时，小车运载能力的冗余度

$\gamma$  为  $1 - H(X)/\log r = 0.25$ 。但如果编码对象为  $N = 2$  的符号序列，我们将两批货物放在一起重新组装运输，此时则每两个货箱至少需要  $NH(S)/\log r = 3$  辆小车，每辆小车的实际载货为  $H(X) = NH(S)/LN = 1$  吨/辆，小车运载能力的冗余度  $\gamma$  为  $1 - H(X)/\log r = 0$ 。

从上述过程可以得到结论：小车在运输货物的过程中，首要前提是不能超载，即：

$$H(X) = NH(S)/\bar{L}_N \leq \log r$$

则：每  $N$  个货箱至少需要的小车数  $\bar{L}_N$  应满足下式：

$$\bar{L}_N \geq NH(S)/\log r \quad (1)$$

其次是让需要的小车数尽量的小，即：

$$NH(S)/\log r + 1 > \bar{L}_N \quad (2)$$

由(1)式与(2)式得到，要想货物没有损失的运输，则每个货箱需要的最少小车数需满足：

$$H(S)/\log r + 1/N > \bar{L}_N/N \geq H(S)/\log r$$

这正是香农第一定理。

从上述过程中可知，信源编码实质就是一种对符号信息进行重整重载的过程，在这个过程中，编码器将信源输出符号载荷的信息提取出来，用一种编码效率更高的方法来对这些信息进行重编码，已达到提高每一个符号平均信息量的目的。而在这个信息转移的过程中，其核心是实际平均信息量  $H(X)$  不能大于对应的极大值  $\log r$ ，在小车模型中就是小车的实际载货量不能大于小车的最大载货量。

### 2.3 结合“小车模型”讲述信道容量和有噪信道编码定理

信道的任务是以信号方式传输信息和存贮信息，研究重点是研究信道中能够传送或存储的最大信息量，即信道容量问题。信号在信道中传输会引入噪声或干扰，它使信号通过信道后产生错误和失真，因此信道的输入和输出之间一般不是确定的函数关系，而是统计依赖关系。为了保证信息传输的可靠性，就必须在信号进入信道前，向对应的码字中加入一定量的校验信息，以在信号从信道输出后完成码字的校验，减少错码率。这就是信道编码。在学习中，可以结合“小车模型”来讲述信息论中的有关知识。

①先验熵  $H(X)$ ：先验熵是信道输入信源  $X$  的熵，是在接收到输出  $Y$  以前，关于输入变量  $X$  的先验不确定性；在“小车模型”中表示进入公路的每一辆小车的平均载货量。

②后验熵  $H(X|b_j)$ ：后验熵是接受到  $b_j$  后关于  $X$  的不确定性，是关于输入符号的信息测度；在“小车模型”中表示一车货物在公路运输过程中的损耗量。

③信道疑义度  $H(X|Y)$ ：信道疑义度是输出端收到

全部输出符号  $Y$  后, 对输入  $X$  所存在的平均不确定程度, 这种不确定程度是由传输过程中信道的干扰引起的; 在“小车模型”中表示每辆小车在公路上运输货物过程中货物的平均损耗量。

④平均互信息  $I(X;Y)$ :  $I(X;Y)=H(X)-H(X|Y)$

平均互信息是接收到符号集  $Y$  后平均每个符号获得的关于  $X$  的信息量, 是输入与输出两个随机变量之间的统计约束程度; 在“小车模型”中表示每辆小车经过在公路上的运输以后货物的平均剩余量。

⑤信道容量  $C$ :  $C=\max\{I(X;Y)\}$ , 表示信道中平均每个符号所能传送的最大信息量, 是信道传输信息的最大能力; 在“小车模型”中表示公路允许小车的最大载货量。

在无损信道中, 信道疑义度等于 0, 信道的容量等于输入符号信息的最大载荷量, 即分布等概时的极大值  $\log r$ ; 在有损信道中, 信道疑义度大于 0, 信道容量等于符号的信息最大载荷量减去传输过程中平均的损失量。

⑥冗余度与有效性和可靠性的关系: 冗余度越小, 信息传输的有效性越好; 冗余度越大, 信息传输的可靠性越好。信道编码在码字中加入校验信息增大了信源编码形成的平均码长  $\bar{L}$ , 减小了信源编码的效率, 但同时减少了信息在信道传输过程中的损耗, 在“小车模型”中就是增大了每一个货箱重新装车以后需要的小车数。

从上述过程可以得到结论: 小车在公路上运输货物过程中, 不仅不能超载, 而且载货量还不能大于公路的最大容量, 即:  $H(X) \leq \log r$ , 且  $H(X) \leq C$  这正是香农第二定理。

从上述过程可知, 信道编码的目的就是为了保证信息传输的可靠性, 为了达到这个目的每一个信道传输符号的平均信息量都不能大于信道的容量, 否则一定会有信息的损失。在“小车模型”中就是小车的实际载货量不能大于公路的最大容量, 只要不大于这个容量, 就一定存在可以无损传输的方法。

## 2.4 结合“小车模型”讲述率失真函数

率失真函数表示在允许一定失真度  $D$  的情况下, 信源输出的信息传输率可压缩到  $R(D)$  值。在“小车模型”中表示在允许一定损耗的情况下, 发送站需要从每个货箱中取出进行装车的货物的最小量。

在信息传输过程中, 不可能保证信息总是能够无失真的到达信宿, 传输符号在信道中传输时, 因为噪

声的影响, 最后能传输到信宿的信息不等于原始的信源熵, 而等于信源熵减去由噪声引起的损失熵。但只要这种损失维持在一定程度内, 还是可以达到信息传输的要求, 这就是失真信源编码的理论基础。在“小车模型”中只要能达到要求, 发送站是不需要把货箱中的货物都重新装车的。

根据上述理论, 可以假设出一个简单的信息失真传输模型来, 这个模型只有信源、信宿和试验信道, 见图 3。当信宿允许一定失真度  $D$  时, 信源输出的符号只要保证能传输过来的信息大于  $R(D)$  值, 就可以完成信息的传输。 $R(D)$  值等于信息传输符号所含平均互信息的最小值, 即传输的最小限度。把这个模型放在信源编码中, 就可以求出每一个信道传输符号的最小平均信息量, 即:

$$C = \min\{I(X;Y)\}$$

这正是香农第三定理。

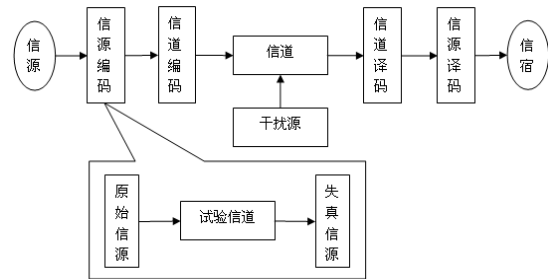


Figure 3. Test channel

图 3 试验信道

从上述过程可知: 信息率失真函数使用的试验信道不是信源与信宿间真正的信息传输信道, 而是一种失真算法。表示在允许一定失真度  $D$  的情况下, 将信源的熵  $H_1(X)$  压缩到  $H_2(X)$  后, 将一个原始信源转化成一个失真信源的过程。在“小车模型”中就是发送站需要从每个货箱中取出进行装车的货物的最小量。信息率失真函数对信息传输速率的影响, 主要在于经过试验信道的压缩以后, 失真信源的熵  $H_2(X)$  比原始信源的熵  $H_1(X)$  要小, 因为单位时间内输出的符号数  $N_1$  不变, 所以单位时间内需要传输的信息量就减小了。通过失真处理以后的信息量再经过无失真信源编码, 信息熵  $H_3(X)$  就会比  $H_2(X)$  小, 传输速率就会提高。该过程的数学表达式为:

原始信源:  $H_1(X)$ ,  $N_2$ ; 失真信源:  $H_2(X)$ ,  $N_2$ ;

信源编码:  $H_3(X)$ ,  $N_3$ ;

因为:  $H_2(X) \geq H_2(X)$ ,  $N_1 = N_2$ , 所以:  $H_1(X) * N_1 \geq H_2(X) * N_2$ ;

因为:  $H_2(X) * N_2 = H_3(X) * N_3$ ,  $H_2(X) \leq H_3(X)$ , 所以:  $N_2 \geq N_3$ 。

### 3 小结

本文从笔者自身的学习与实践出发,结合自己在信息论与编码课程教学与学习过程中的经验与体会,尝试在香农信息论的基础上构建一个实例,将抽象的概念与生活实例结合到一起,辅助教学,加深理解,取得了较好的效果。同时也需要指出,现有的“小车模型”还存在着不足,比如通信系统中的加密、解密部分没能在模型中体现,也希望能够与老师同学们展开更加深入的讨论,进一步提高信息论与编码的授课效果。

### References (参考文献)

- [1] Feng Gui, Lin Qiwei, Chen Huadong, Information Theory and Coding Theory[M]. BeiJing: Tsinghua University Press,2007.  
冯桂, 林其伟, 陈东华. 信息论与编码技术[M]. 北京: 清华大学出版社, 2007.
- [2] Thomas M. Cover, Joy A. Thomas, Elements of Information Theory [M]. 2ed Edition, John Wiley & Sons, Inc. , Publication.
- [3] C E Shannon. A Mathematical Theory of communications [J]. Bell system Technical Journal, 1948(3): 379-423.
- [4] DENG Jiaxian, Teaching Reform on Information Theory and Coding Theory[J], *Journal of EEE*, 2007,29(2),113-114.  
邓家先. 信息论与编码课程改革教学探讨. 电气电子教学学报, 2007, 29(2): 113-114.