

基于多重分形谱的物理层帧结构检测算法研究

李歆昊* 张旻 韩树楠

(国防科技大学电子对抗学院 合肥 230037)

(安徽省电子制约技术重点实验室 合肥 230037)

摘要: 针对解调数据有无物理层帧结构的检测问题, 该文提出了一种基于多重分形谱的物理层帧结构检测算法。首先, 在定义序列有偏性与 0, 1 出现概率关系的基础上, 根据编码序列、扰码序列和物理帧的生成原理, 通过计算相关序列中的 0, 1 出现概率推导出它们的有偏性关系, 即物理帧、编码序列和扰码序列的有偏性逐渐减小; 然后, 利用上述结论并结合多重分形谱具有描述有偏性分布的特点, 对待识别序列进行固定长度拆分并计算多重分形谱宽度值; 最后, 通过观察多重分形谱宽度值的分布实现解调数据有无物理层帧结构的检测。仿真试验验证了算法的有效性, 表明该算法具有一定的工程应用价值。

关键词: 物理层; 帧结构; 多重分形谱; 有偏性

中图分类号: TN911.22

文献标识码: A

文章编号: 1009-5896(2018)02-0400-08

DOI: 10.11999/JEIT170356

Study on Frame Structure Detection in Physical Layer Based on Multi-fractal Spectrum

LI Xinhao ZHANG Min HAN Shunan

(School of Electronic Countermeasures, National University of Defense Technology, Hefei 230037, China)

(Key Laboratory of Anhui Electronic Restricting Technique, Hefei 230037, China)

Abstract: In order to achieve the justification of whether a frame structure detection exists or not under the condition that protocol is unknown, an algorithm for frame structure detection in physical layer based on multi-fractal spectrum is proposed. Firstly, the relationship between bias and occurrence probability of 0 and 1 bit is defined. Through analyzing the generation principle of channel code, scrambler, frame sequence and calculating the occurrence probability of 0 and 1 bit, the conclusion is that the bias of synchronization word is larger than the others. Then, due to the above conclusion and the character that bias distribution can be described by multi-fractal spectrum, the width of multi-fractal spectrum of split sequences is calculated. At last, frame structure can be detected by observing distribution of the width of multi-fractal spectrum of split sequences. Simulation results show the proposed algorithm is effective, and it has value in engineering application.

Key words: Physical layer; Frame structure; Multi-fractal spectrum; Bias

1 引言

随着社会的飞速发展, 无线通信和智能通信在人们的生活中扮演着越来越重要的角色, 而通信协议分析是上述通信过程中不可或缺的重要组成部分^[1]。在非合作的军事对抗领域, 协议分析是获取敌方通信信息和进行有效干扰的前提^[2]。因此, 针对通信协议识别的研究受到了国内外学者的广泛关注。物理层主要负责通信双方之间的比特数据传输, 高层

协议的识别工作必须建立在物理层协议解析的基础之上^[2]。根据通信协议类型的不同, 数据在物理层调制发射前还会选择性插入信令和导频块, 形成物理层帧结构, 因此, 物理层帧结构的检测是未知数据通信协议识别的前提。

数据在物理层传输时会选择性成帧, 即在信道编码的基础上选择性插入由信令和导频块构成的同步字。根据协议类型的不同, 每帧的编码长度不定但都远大于同步字长度。因此, 非合作通信中的协议识别就面临物理帧的检测问题且难度较大。目前, 国内外类似研究主要集中于同步字的识别^[3-9]。相关算法如文献[5]提出的累积滤波算法、文献[6]提出的集中插入式算法、文献[7]提出的编码 QS 算法和文献[8]提出的容器算法。对于误码率高的情况, 文献[9]利用二次相关滤波、哈达玛变换和模糊匹配完

收稿日期: 2017-04-20; 改回日期: 2017-10-27; 网络出版: 2017-12-04

*通信作者: 李歆昊 lixinhao1989616@126.com

基金项目: 国家自然科学基金(61171170), 安徽省自然科学基金(1408085QF115)

Foundation Items: The National Natural Science Foundation of China (61171170), The Natural Science Foundation of Anhui Province (1408085QF115)

成同步字的识别，具有较好的鲁棒性。分析已有算法，它们大多是在确定物理帧存在的条件下完成同步字的识别。在非合作的军事对抗领域，由于通信协议类型未知，需要首先判断数据是否存在物理帧然后进行下一步识别，然而已有研究还未涉及物理帧检测。因此，研究一种有效的物理层帧结构检测算法具有重要的意义。

分形概念揭示了自然界中一大类无规则形体的内在规律——标度不变性，其中多重分形所描述的主要是某个参量的有偏性分布^[10,11]。由于同步字、物理帧、编码和扰码的有偏性逐渐减小，因此本文提出了一种基于多重分形谱的物理层帧结构检测算法。通过分析编码序列、扰码序列和物理帧的有偏

性差异，得出物理帧、编码序列和扰码序列的有偏性逐渐减小的结论；利用上述结论，对待识别序列进行拆分并计算每段拆分序列的多重分形谱宽度值，通过观察多重分形谱宽度值的分布，实现解调数据有无物理帧的判断。

2 问题描述

2.1 系统组成

已有的无线通信和智能通信的信道传输系统主要由发射端和接收端两部分组成，其中的发射端系统主要由模式调整、信号流调整、前向纠错和物理层成帧 4 部分组成，接收端包括了物理帧检测、去除信令和导频、解纠错编码、解交织和帧同步，结构如图 1 所示。

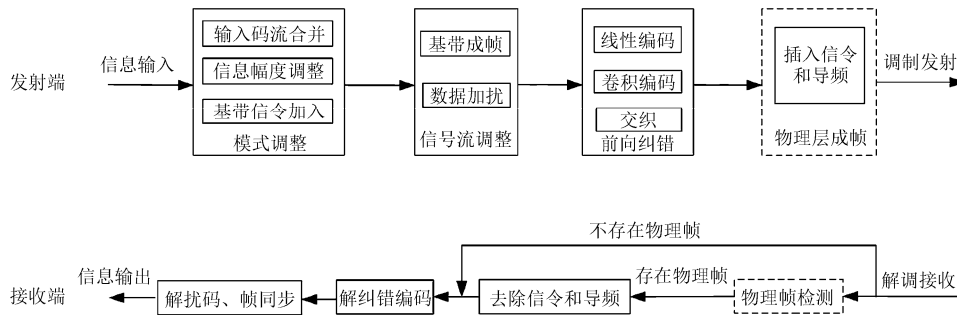


图 1 系统组成

发射端的模式调整模块主要对输入码流进行合并和信息幅度调整，加入基带信令从而告知该模块的内容；信号流调整模块将输入数据按照固定长度打包完成基带成帧并进行加扰处理；前向纠错模块对输入的基带帧数据进行纠错编码，并进行交织以便在接收端更好地完成纠错恢复；根据协议类型的不同，部分协议数据会进入如图 1 虚线框所示的物理层成帧模块，对前向纠错帧插入信令和导频，形成物理层帧结构。如空间数据系统咨询委员会 (CCSDS) 制定的遥控传输单元 (TC) 存在物理层帧结构，而遥测传输单元 (TM) 不存在物理层帧结构^[2]。

在非合作通信中，接收端要想准确恢复出信息需要首先对解调数据进行有无物理帧的判断，如图 1 中虚线框所示。若存在物理帧需要去除信令和导频，然后解纠错编码、解扰码和帧同步，最后才能提取出有用信息。

2.2 物理帧

根据协议类型的不同，部分数据传输到物理层会以图 2 所示的物理层帧结构出现。

物理层帧结构中，物理层信令和导频块构成同步字，用于确定每帧起始位；紧跟后面由编码或扰码序列构成一帧中传输的信息段，信息段中的数据

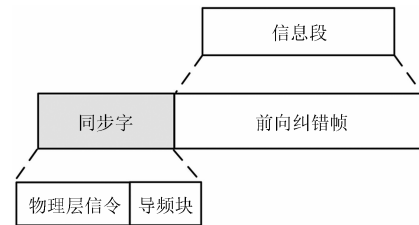


图 2 物理层帧结构

不能与同步字相同^[12]。在非合作的军事对抗领域，由于协议类型未知，接收端解调数据类型存在 3 种可能：编码序列、扰码序列和物理帧序列。

2.3 序列的多重分形谱

将长度为 L_i 的拆分序列 I_i 按分割长度 ξ 划分成 $\lfloor L_i/\xi \rfloor$ 段 ($\lfloor \cdot \rfloor$ 表示向下取整)，把各段转换成相应的十进制数并计算十进制数种类 ϕ (L 足够大时 $\phi = 2^\xi$)，统计各十进制数的出现概率 P_i 并对其进行分档，计算每一档下概率 P_i 的数目 N_j 。令分形尺度 $\varepsilon = 1/\phi$ ，则有式(1)和式(2)成立^[13]。

$$P_i(\varepsilon) \propto \varepsilon^\alpha \tag{1}$$

$$N_j(\varepsilon) \propto \varepsilon^{-f(\alpha)} \tag{2}$$

α 是奇异指数，反映了分形上各十进制数的奇

异程度; $f(\alpha)$ 表示相同 α 值的子集的分形维数, 一般将 $f(\alpha)$ 称为多重分形谱^[13], 多重分形谱的宽度能够反映序列的有偏性分布特征。

3 序列的有偏性

常用的纠错编码主要包括线性分组码与卷积码, 由于卷积码的产生使用移位寄存器, 导致伪随机性大于线性分组码^[14]; 自同步扰码因其具有较好的保密性而广泛应用^[15]。因此, 本文针对信息段是线性分组码和线性分组码的自同步加扰序列进行有偏性分析。

定义 1^[15] 若二进制序列中 1 的出现概率为 ς , 则序列的有偏性是 $1/2 - \varsigma$, 它反应了序列中 0, 1 bit 的分布情况。

3.1 线性分组码

信息位长度是 k 、编码长度是 n 的线性分组码 (n, k) , 它的 $k \times n$ 阶生成矩阵 \mathbf{G} 由式(3)定义。

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} = \begin{bmatrix} g_{0,0} & g_{0,1} & \cdots & g_{0,n-1} \\ g_{1,0} & g_{1,1} & \cdots & g_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,n-1} \end{bmatrix} \quad (3)$$

其中, $g_{i,j} = (0 \text{ or } 1)$, $\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{k-1}$ 是线性独立的 n 阶向量。假设 k 阶信息向量 $\mathbf{x}_i = (x_{i,0}, x_{i,1}, \dots, x_{i,k-1})$, $i = 1, 2, \dots$, 经过式(4)编码后, 得到码字 $\mathbf{c}_i (i = 1, 2, \dots)$ 。

$$\mathbf{c}_i = \mathbf{x}_i \cdot \mathbf{G} = (x_{i,0}, x_{i,1}, \dots, x_{i,k-1}) \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} \quad (4)$$

输出的编码序列 $\mathbf{c}_i = (c_{i,0}, c_{i,1}, \dots, c_{i,n-1})$, $i = 1, 2, \dots$ 。

假设 (n, k) 线性分组码 $\mathbf{c}_i = (c_{i,0}, c_{i,1}, \dots, c_{i,n-1})$ 的信息序列 $\mathbf{x}_i = (x_{i,0}, x_{i,1}, \dots, x_{i,k-1})$ 中 0, 1 分布有偏性是 φ , 生成矩阵 $\mathbf{G} = (\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{k-1})'$ 的每一列码重是 $W_i, i = 0, 1, \dots, n-1$, 信息序列中 1 出现的次数用 v 表示、出现概率可表示成 $\Pr(x_{i,j} = 1) = (1/2) - \varphi, j = 0, 1, \dots, k-1$, 则 $c_{i,j} = 1, j = 0, 1, \dots, k-1$ 出现的概率可以表示为^[15]

$$\begin{aligned} \Pr(c_{i,j} = 1) &= \sum_{v=1,3,\dots}^{W_i} \left(\frac{W_i}{v} \right) \left(\frac{1}{2} - \varphi \right)^v \left(\frac{1}{2} + \varphi \right)^{W_i-v} \\ &= \frac{1}{2} [1 - (2\varphi)^{W_i}] \end{aligned} \quad (5)$$

编码码字 $\mathbf{c}_{i,j}$ 的有偏性可以表示为

$$\varphi_{c_{i,j}} = \frac{1}{2} - \Pr(c_{i,j} = 1) = \frac{1}{2} (2\varphi)^{W_i} \quad (6)$$

由于常用编码都是系统码, 其生成矩阵 \mathbf{G} 的前

$k-1$ 列是单位矩阵, 因此 \mathbf{G} 的前 $k-1$ 列的码重取值为 $W_i = 1, i = 0, 1, \dots, k-1$, 后 $n-k$ 列的码重取值 $W_i > 1, i = k, k+1, \dots, n-1$ 。由于 $\varphi < 0.5$, 所以 $\varphi_{c_{i,j}} < \varphi$, 整个编码序列 \mathbf{c}_i 的有偏性 φ_{c_i} 可以表示成式(7)的形式^[16]。

$$\varphi_{c_i} = \frac{1}{n} \sum_{j=0}^{n-1} \varphi_{c_{i,j}} \leq \frac{1}{n} \sum_{j=0}^{n-1} \varphi_j = \varphi \quad (7)$$

编码序列的有偏性可进一步表示为

$$\begin{aligned} \varphi_{c_i} &= \frac{1}{n} \sum_{j=0}^{n-1} \varphi_{c_{i,j}} = \frac{1}{n} \sum_{j=0}^{k-1} \frac{1}{2} (2\varphi_j) + \frac{1}{n} \sum_{j=k}^{n-1} \frac{1}{2} (2\varphi_j)^{W_j} \\ &= \frac{k}{n} \varphi + \frac{1}{n} \sum_{j=k}^{n-1} \frac{1}{2} (2\varphi)^{W_j} \approx \frac{k}{n} \varphi \end{aligned} \quad (8)$$

由式(8)可知, 线性分组码序列中 0 和 1 分布的有偏性小于原有信息序列^[16]。

3.2 扰码序列

假设自同步扰码器以 $f(x) = 1 + b_1x + b_2x^2 + \dots + b_Lx^L$ 作为本原多项式, 输入线性分组码 $\mathbf{c} = (c_L, c_{L+1}, \dots, c_{2L-1})$, 得到扰码序列 $\mathbf{a} = (a_L, a_{L+1}, \dots, a_{2L-1})$, 则扰码器加扰过程如式(9)^[16]

$$a_i = c_i \oplus \sum_{j=1}^L b_j a_{i-j}, \quad i = L, L+1, \dots, 2L-1 \quad (9)$$

将式(9)表示成矩阵形式:

$$\begin{pmatrix} a_L \\ a_{L+1} \\ \vdots \\ a_{2L-1} \end{pmatrix} = \begin{pmatrix} c_L \\ c_{L+1} \\ \vdots \\ c_{2L-1} \end{pmatrix} + \begin{pmatrix} a_{L-1} & a_{L-2} & \cdots & a_0 \\ a_L & a_{L-1} & \cdots & a_1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{2L-2} & a_{2L-3} & \cdots & a_{L-1} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{pmatrix} \quad (10)$$

对有偏性是 ψ 的线性分组码 \mathbf{c} 进行自同步加扰, 假设得到有偏性是 δ 的线性分组码加扰序列 $\mathbf{a} = (a_L, a_{L+1}, \dots, a_{2L-1})$, 此时扰码序列的有偏性 $\delta = 0$, 下面给出证明。

证明

$a_i = 1$ 出现的概率可以表示为

$$\Pr(a_i = 1) = 1/2 - \delta \quad (11)$$

当 $c_i = 0, (a_{i-1} \ a_{i-2} \ \cdots \ a_{i-L})(b_1 \ b_2 \ \cdots \ b_L)' = 1$ 时, 编码序列中 0 出现的概率:

$$\Pr(c_i = 0) = 1/2 + \psi \quad (12)$$

扰码序列中 1 出现的概率:

$$\begin{aligned} \Pr \left((a_{i-1} \ \cdots \ a_{i-L}) \begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix} = 1 \right) \\ = \sum_{v=1,3,\dots}^W \left(\frac{W}{v} \right) \left(\frac{1}{2} - \delta \right)^v \left(\frac{1}{2} + \delta \right)^{W-v} \end{aligned} \quad (13)$$

其中, v 表示1出现的次数, W 是序列 $(b_1 b_2 \cdots b_L)$ 的码重。当 $c_i = 1, (a_{i-1} a_{i-2} \cdots a_{i-L})(b_1 b_2 \cdots b_L)' = 0$ 时, 编码序列中1出现的概率:

$$\Pr(c_i = 1) = \frac{1}{2} - \psi \quad (14)$$

扰码序列中0出现的概率:

$$\Pr \left(\begin{matrix} (a_{i-1} \cdots a_{i-L}) \\ \vdots \\ (b_1 \cdots b_L) \end{matrix} = 0 \right) = \sum_{v=2,4,\dots}^W \left(\frac{W}{v} \right) \left(\frac{1}{2} - \delta \right)^v \left(\frac{1}{2} + \delta \right)^{W-v} \quad (15)$$

因此, 线性分组码加扰序列中1出现的概率可表示为

$$\begin{aligned} \Pr(a_i = 1) &= \Pr \left(\begin{matrix} (a_{i-1} \cdots a_{i-L}) \\ \vdots \\ (b_1 \cdots b_L) \end{matrix} = 1 \right) \Pr(c_i = 0) \\ &\quad + \Pr \left(\begin{matrix} (a_{i-1} \cdots a_{i-L}) \\ \vdots \\ (b_1 \cdots b_L) \end{matrix} = 0 \right) \Pr(c_i = 1), \\ \frac{1}{2} - \delta &= \left[\sum_{v=1,3,\dots}^W \left(\frac{W}{v} \right) \left(\frac{1}{2} - \delta \right)^v \left(\frac{1}{2} + \delta \right)^{W-v} \right] \left(\frac{1}{2} + \psi \right) \\ &\quad + \left[\sum_{v=2,4,\dots}^W \left(\frac{W}{v} \right) \left(\frac{1}{2} - \delta \right)^v \left(\frac{1}{2} + \delta \right)^{W-v} \right] \left(\frac{1}{2} - \psi \right), \\ \frac{1}{2} - \delta &= \frac{1}{2} [1 - (2\delta)^W] \left(\frac{1}{2} + \psi \right) + \frac{1}{2} [1 + (2\delta)^W] \left(\frac{1}{2} - \psi \right) \end{aligned} \quad (16)$$

由于扰码序列中0和1的出现概率基本相同, 因此加扰序列的有偏性 $\delta < 0.5$, 式(16)可进一步化简为

$$\frac{1}{2} - \delta \approx \frac{1}{2} \left(\frac{1}{2} + \psi \right) + \frac{1}{2} \left(\frac{1}{2} - \psi \right) \quad (17)$$

即 $\delta = 0$ 。

证毕

3.3 物理帧

假设一帧数据由同步字长度是 l_1 、信息段长度是 l_2 的数据构成, 同步字中0和1分布的有偏性是 λ , 信息段中0和1分布的有偏性是 η , 如图3所示。

计算图3所示的物理帧中0和1分布的有偏性, 假设输出的物理帧是 $d = (d_1, d_2, \dots, d_{l_1+l_2})$ 。由上述分析



图3 物理帧信息

和图3可知, 同步字中1的出现概率如式(18)所示。

$$\Pr(d_i = 1) = \frac{1}{2} - \lambda, \quad i = 1, 2, \dots, l_1 \quad (18)$$

信息段中1的出现概率如式(19)所示。

$$\Pr(d_i = 1) = \frac{1}{2} - \eta, \quad i = 1, 2, \dots, l_2 \quad (19)$$

因此, 一帧中1出现的概率可由式(20)表示。

$$\begin{aligned} \Pr(d_i = 1) &= \frac{(1/2 - \lambda) \cdot l_1 + (1/2 - \eta) \cdot l_2}{l_1 + l_2} \\ &= \frac{1}{2} - \frac{\lambda \cdot l_1 + \eta \cdot l_2}{l_1 + l_2}, \quad i = 1, 2, \dots, l_1 + l_2 \end{aligned} \quad (20)$$

由式(20)可知, 一帧的有偏性 $\gamma = (\lambda \cdot l_1 + \eta \cdot l_2) / (l_1 + l_2)$, 可化简成式(21)和式(22)的形式。

$$\gamma = \eta + \frac{\lambda - \eta}{l_1 + l_2} \cdot l_1 \quad (21)$$

$$\gamma = \lambda + \frac{\eta - \lambda}{l_1 + l_2} \cdot l_2 \quad (22)$$

(1)信息段是线性分组码: 通过式(8)可知, 编码后序列的有偏性降低。因此, 由固定字符串构成的同步字的有偏性大于线性分组码, 此时同步字的有偏性 λ 大于信息段的有偏性 η 。

(2)信息段是扰码: 通过式(17)可知, 加扰序列的有偏性为0。由于同步字的有偏性大于0, 此时同步字的有偏性 λ 大于信息段的有偏性 η 。

根据上述分析可知, 同步字的有偏性大于信息段的有偏性, 即 $\lambda > \eta$ 。因此, 由式(21)可知, 信息段是线性分组码和扰码的物理帧的有偏性大于信息段, 即 $\gamma > \eta$; 由式(22)可知, 同步字的有偏性大于物理帧的有偏性, 即 $\lambda > \gamma$ 。因此可以得出结论, 同步字、物理帧、线性分组码和扰码的有偏性逐渐减小。

4 多重分形谱的计算

多重分形谱地计算包括解析方法和统计物理方法^[13], 由于后者具有更广的应用范围, 因此本文选用统计物理方法进行多重分形谱地计算。

首先定义一个配分函数 $\chi_q(\varepsilon)$, 对各十进制数的出现概率 $P_i(\varepsilon)$ 用 q 次方进行加权求和, 其数学表达式为

$$\chi_q(\varepsilon) \equiv \sum P_i(\varepsilon)^q = \varepsilon^{\tau(q)} \quad (23)$$

如果式(23)后面的等式成立, 即配分函数 $\chi_q(\varepsilon)$ 和分形尺度 ε 有幂函数关系, 则可以从 $\ln \chi_q \sim \ln \varepsilon$ 曲线的斜率得到

$$\tau(q) = \frac{\ln \chi_q(\varepsilon)}{\ln \varepsilon}, \quad \varepsilon \rightarrow 0 \quad (24)$$

一般把 $\tau(q)$ 称为质量指数^[13]。

按概率 $P_i(\varepsilon)$ 的大小进行分档后求和, 式(23)可表示成式(25)的形式:

$$\chi_q(\varepsilon) \equiv \sum P_i(\varepsilon)^q = \sum N(P)P^q \quad (25)$$

式中的 $N(P)$ 是概率为某一档 P 的数目。由式(1)和式(2), 式(25)可以表示为

$$\chi_q(\varepsilon) = \sum \varepsilon^{-f(\alpha)} \varepsilon^{\alpha q} = \sum \varepsilon^{\alpha q - f(\alpha)} = \varepsilon^{\tau(q)} \quad (26)$$

最后的等式在系统属于多重分形时成立, 它可以改写为

$$\sum \varepsilon^{\alpha q - f(\alpha) - \tau(q)} = 1 \quad (27)$$

当 $\varepsilon \rightarrow 0$ 时, 求和号中 $\alpha q - f(\alpha) - \tau(q) > 0$ 的项趋于0, $\alpha q - f(\alpha) - \tau(q) > 0$ 的项不应该出现, 因为此时将出现无限大, 这样一来只有 $\alpha q - f(\alpha) - \tau(q) = 0$ 的项目保留下来, 即

$$\tau(q) = \alpha q - f(\alpha) \quad (28)$$

而 α 可以从以下的微商得出, 即^[13]

$$\alpha = \frac{d\tau(q)}{dq} \quad (29)$$

由式(23)、式(24)和式(29)可得

$$\begin{aligned} \alpha &= \frac{d\tau(q)}{dq} = \frac{d}{dq} \left(\frac{\ln \chi_q(\varepsilon)}{\ln \varepsilon} \right) = \frac{d}{dq} \left(\frac{\ln \sum P_i(\varepsilon)^q}{\ln \varepsilon} \right) \\ &= \frac{\sum [P_i(\varepsilon)^q \cdot \ln P_i(\varepsilon)]}{\ln \varepsilon \cdot \sum P_i(\varepsilon)^q} \end{aligned} \quad (30)$$

将式(24)和式(30)代入式(28)得

$$f(\alpha) = \frac{q \cdot \sum [P_i(\varepsilon)^q \cdot \ln P_i(\varepsilon)]}{\ln \varepsilon \cdot \sum P_i(\varepsilon)^q} - \frac{\ln \sum P_i(\varepsilon)^q}{\ln \varepsilon} \quad (31)$$

由奇异指数 α 和多重分形谱 $f(\alpha)$ 构成的图形就是所求的多重分形谱, 多重分形谱的宽度越宽表示对应序列的有偏性越大^[13]。由于物理帧、线性分组码和扰码序列的有偏性存在差异, 通过计算拆分序

列多重分形谱宽度值的分布, 可以准确判断解调数据中有无物理帧。

5 算法步骤

选取序列拆分长度 $\varphi = 96$ bit、拆分序列分割长度 $\xi = 8$ bit 进行识别, 具体步骤如下。

步骤 1 将长度为 L 的待识别序列 I , 按照每段拆分长度 φ 划分成 $\lfloor L/\varphi \rfloor$ 段 ($\lfloor \cdot \rfloor$ 表示向下取整);

步骤 2 按照拆分序列分割长度 ξ 计算各段的多重分形谱宽度值 $\alpha_1 \alpha_2 \dots \alpha_{\lfloor L/\varphi \rfloor}$ 并绘制分布图;

步骤 3 若分布图存在周期性峰值, 则解调数据中存在物理帧, 反之不存在。

6 仿真实验

实验选取的无物理帧数据分别由(7,4)分组码、(15,11)分组码和以 $f(x) = 1 + x^{18} + x^{23}$ 作为本原多项式, (7,4)分组码作为输入信息数据的自同步扰码序列构成, 在上述数据前添加表1中的同步字, 构成实验所需的物理帧数据。

实验 1 无误码识别

(1)(7,4)分组码构成信息段: 在无误码的条件下, 各选取 2×10^5 bit 由(7,4)分组码构成的物理帧 1、物理帧 2 和无物理帧数据, 其中物理帧每帧中的分组码长度为 2800 bit。选取序列拆分长度 $\varphi = 96$ bit、拆分序列分割长度 $\xi = 8$ bit, 对上述 3 种数据进行多重分形谱宽度的计算, 结果如图 4 所示。

由图 4 可知, 物理帧 1 与物理帧 2 拆分序列的多重分形谱宽度分布存在周期性峰值, 序列中包含有物理帧; 而仅由(7,4)分组码构成的序列的多重分形谱宽度分布不存在周期性峰值, 各段拆分序列的多重分形谱宽度基本相同。表 2 对应的是 3 种序列

表 1 物理帧特征统计

物理帧	同步字	长度(bit)
物理帧 1	11000101110001011100010111000101110001011100010111000101011110011110101110010000	80
物理帧 2	1011	64

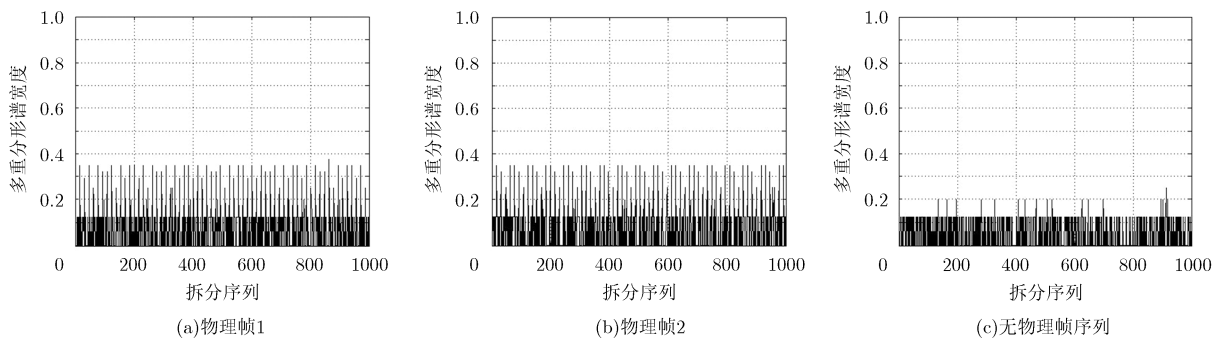


图 4 (7,4)分组码构成信息段, 多重分形谱宽度分布

表 2 (7,4)分组码构成信息段，多重分形谱宽度值分布的方差

序列	物理帧 1	物理帧 2	无物理帧序列
方差	0.0084	0.0086	0.0041

表 3 (15,11)分组码构成信息段，多重分形谱宽度值分布的方差

序列	物理帧 1	物理帧 2	无物理帧序列
方差	0.0057	0.0057	0.0029

多重分形谱宽度值分布的方差取值。

由表 2 可知，物理帧 1 与物理帧 2 的方差接近，大于无物理帧序列的方差，表明物理帧 1 与物理帧 2 拆分序列的多重分形谱宽度取值分布比无物理帧序列的离散程度高，即物理帧 1 与物理帧 2 各拆分序列之间的多重分形谱宽度值差别较大。

(2)(15,11)分组码构成信息段：在无误码的条件下，各选取 2×10^5 bit 由(15,11)分组码构成的物理帧 1、物理帧 2 和无物理帧数据，其中物理帧每帧中的分组码长度为 3000 bit。选取序列拆分长度 $\varphi = 96$ bit、拆分序列分割长度 $\xi = 8$ bit，对上述 3 种数据进行多重分形谱宽度的计算，结果如图 5 所示。

由图 5 可知，物理帧 1 与物理帧 2 拆分序列的多重分形谱宽度分布存在周期性峰值，序列中包含有物理帧；而仅由(15,11)分组码构成的序列的多重分形谱宽度分布不存在周期性峰值，各段拆分序列的多重分形谱宽度基本相同。表 3 对应的是 3 种序列多重分形谱宽度值分布的方差取值。

由表 3 可知，物理帧 1 与物理帧 2 的方差相同，大于无物理帧序列的方差，表明物理帧 1 与物理帧

2 拆分序列的多重分形谱宽度取值分布比无物理帧序列的离散程度高，即物理帧 1 与物理帧 2 各拆分序列之间的多重分形谱宽度值差别较大。

(3)扰码数据构成信息段：在无误码的条件下，各选取 2×10^5 bit 以 $f(x) = 1 + x^{18} + x^{23}$ 作为本原多项式，(7,4)分组码作为输入信息数据的自同步扰码序列构成的物理帧 1、物理帧 2 和无物理帧数据，其中物理帧每帧中的扰码序列长度为 3000 bit。选取序列拆分长度 $\varphi = 96$ bit、拆分序列分割长度 $\xi = 8$ bit，对上述 3 种数据进行多重分形谱宽度的计算，结果如图 6 所示。

由图 6 可知，物理帧 1 与物理帧 2 拆分序列的多重分形谱宽度分布存在周期性峰值，序列中包含有物理帧；而仅由扰码数据构成的序列的多重分形谱宽度分布不存在周期性峰值，各段拆分序列的多重分形谱宽度基本相同。表 4 对应的是 3 种序列多重分形谱宽度值分布的方差取值。

由表 4 可知，物理帧 1 与物理帧 2 的方差接近，大于无物理帧序列的方差，表明物理帧 1 与物理帧 2 拆分序列的多重分形谱宽度取值分布比无物理帧序列的离散程度高，即物理帧 1 与物理帧 2 各拆分

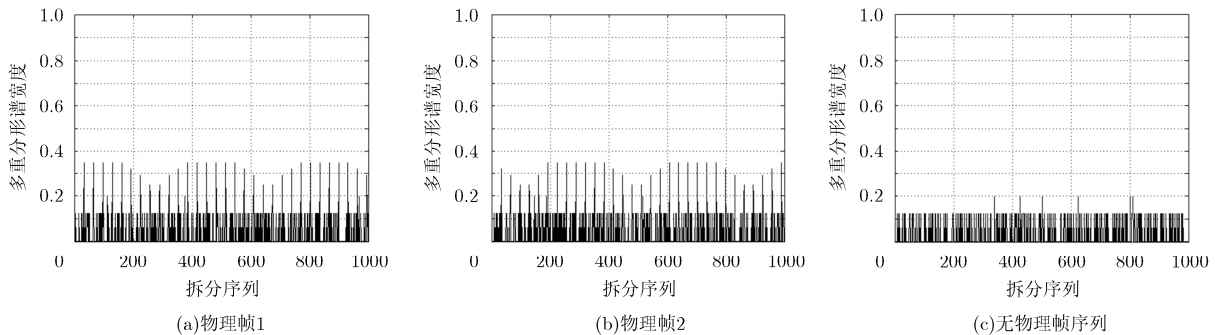


图 5 (15,11)分组码构成信息段，多重分形谱宽度分布

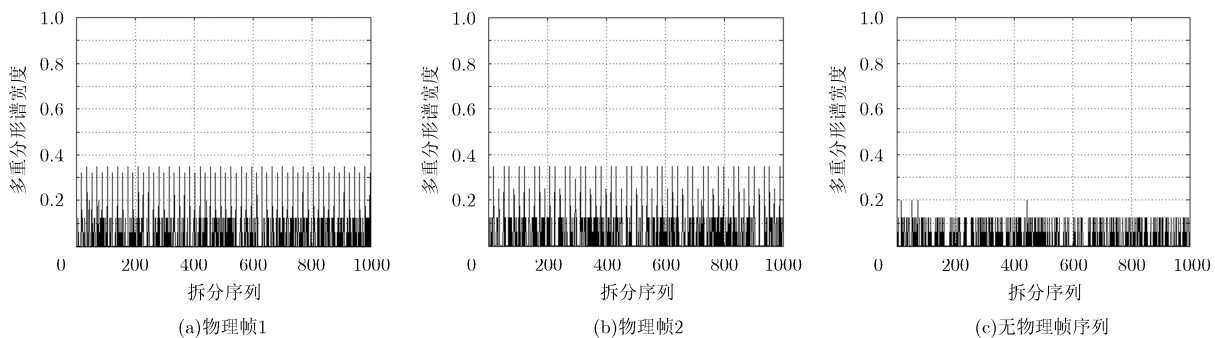


图 6 扰码数据构成信息段，多重分形谱宽度分布

表4 扰码数据构成信息段, 多重分形谱宽度值分布的方差

序列	物理帧 1	物理帧 2	无物理帧序列
方差	0.0079	0.0078	0.0029

序列之间的多重分形谱宽度值差别较大。

实验2 正确检测概率分析

选取 2×10^5 bit 由(15,11)分组码构成的物理帧1, 每帧中的分组码长度为 3000 bit。在不同误码率条件下各进行 100 次的蒙特卡洛仿真实验, 检验本文算法在不同误码率条件下正确检测出物理帧的概率, 结果如图7所示。

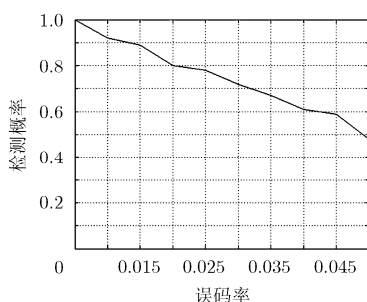


图7 正确检测概率分布图

当误码率小于 2% 时, 误码对物理帧中同步字和信息段的有偏性影响较小, 此时正确检测概率保持在 80% 以上; 随着误码率的升高, 物理帧中同步字和信息段的随机性增强, 两者的有偏性都接近于随机序列, 因此当误码率升高到 5% 时, 正确检测概率小于 50%。

7 结束语

本文提出基于多重分形谱的物理层帧结构检测算法, 解决了协议类型未知条件下解调数据有无物理帧的检测问题。算法利用同步字、编码、扰码和物理帧有偏性的差异, 在计算拆分待识别序列的多重分形谱宽度值的基础上, 通过观察取值的分布, 实现了解调数据有无物理帧的检测。由于误码对序列中 0 和 1 分布有偏性影响较小, 因此本文算法具有较好的抗误码性能。在检测出物理帧后, 同步字的识别是下一步协议破析的关键, 对于同步字的识别我们将在后续展开研究。

参考文献

- [1] 张永光, 娄才义. 信道编码及其识别分析[M]. 北京: 电子工业出版社, 2010: 1-2.
ZHANG Yongguang and LOU Caiyi. Channel Coding and Recognition Analysis[M]. Beijing: Publishing House of Electronics Industry, 2010: 1-2.
- [2] 李相迎. CCSDS 数据链路层协议识别关键技术研究[D]. [博士学位论文], 中国科学院研究生院, 2011.
LI Xiangying. Key technologies of protocol identification for CCSDS data link layer[D]. [Ph.D.dissertation], Center for Space Science and Applied Research Chinese Academy of Sciences, 2011.
- [3] QIN J, HUANG Z, LIU C, *et al.* Novel blind recognition algorithm of frame synchronization words based on soft-decision in digital communication systems[J]. *Plos One*, 2015, 10(7): 135-144. doi: 10.1371/journal.pone.0132114.
- [4] LIANG Y, RAJAN D, and ELIEZER O. Sequential frame synchronization based on hypothesis testing with unknown channel state information[J]. *IEEE Transactions on Communications*, 2015, 63(8): 2972-2984. doi: 10.1109/TCOMM.2015.2438895.
- [5] 郭凯丰, 王萌. 基于等帧长信号的帧头检测方法研究[J]. 信号与信息处理, 2014, 44(6): 33-36.
GUO Kaifeng and WANG Meng. Research on detection of frame head on fixed frame length[J]. *Journal of Signal and Information Processing*, 2014, 44(6): 33-36.
- [6] 张玉, 杨晓静. 集中插入式帧同步识别方法[J]. 兵工学报, 2013, 34(5): 554-560.
ZHANG Yu and YANG Xiaojing. Recognition method of concentratively inserted frame synchronization[J]. *Acta Armamentarii*, 2013, 34(5): 554-560.
- [7] 王和洲. 面向比特流的链路协议识别与分析技术[D]. [硕士学位论文], 中国科学技术大学, 2014.
WANG Hezhou. Research on bit-stream oriented link protocol identification and analysis techniques[D]. [Master dissertation], University of Science and Technology of China, 2014.
- [8] 张一嘉. 局域网链路层数据帧识别算法的设计与实现[J]. 通信对抗, 2007, 99(4): 41-44.
ZHANG Yijia. Design and implementation of algorithm for LAN data frame recognition[J]. *Communication Countermeasures*, 2007, 99(4): 41-44.
- [9] 白彧, 杨晓静, 王懋. 基于相关滤波和哈达玛变换的帧同步码识别[J]. 探测与控制学报, 2011, 33(3): 69-72.
BAI Yu, YANG Xiaojing, and WANG Mao. Recognition method of frame synchronization codes based on relativity filter and hadamard transformation algorithm[J]. *Journal of Detection and Control*, 2011, 33(3): 69-72.
- [10] YADAV R P, KUMAR M, MITTAL A, *et al.* Fractal and multifractal characteristics of swift heavy ion induced self-affine nanostructured BaF₂ thin film surfaces[J]. *Chaos An Interdisciplinary Journal of Nonlinear Science*, 2015, 25(8): 083115. doi: 10.1063/1.4928695.
- [11] OROZCODUQUE A, NOVAK D, KREMEN V, *et al.* Multifractal analysis for grading complex fractionated

- electrograms in atrial fibrillation[J]. *Physiological Measurement*, 2015, 36(11): 2269–2284. doi: 10.1088/0967-3334/36/11/2269.
- [12] 张玉龙. DVB-S2 解调器关键算法及 VLSI 实现研究[D]. [硕士学位论文], 复旦大学, 2010.
- ZHANG Yulong. Research on VLSI implementation and key algorithm of DVB-S2 demodulator[D]. [Master dissertation], Fudan University, 2010.
- [13] 孙霞, 吴自勤, 黄响. 分形原理及其应用[M]. 合肥: 中国科学技术大学出版社, 2006: 1–2.
- SUN Xia, WU Ziqin, and HUANG Jun. Fractal Principle and Application[M]. Hefei: Press of University of Science and Technology of China, 2006: 1–2.
- [14] 李歆昊, 张旻, 史英春, 等. 基于游程特征的线性分组码与卷积码类型识别[J]. *数据采集与处理*, 2015, 30(6): 1205–1214.
- LI Xinhao, ZHANG Min, SHI Yingchun, *et al.* Blind identifying of type of linear block code and convolutional code based on run feature[J]. *Journal of Data Acquisition and Processing*, 2015, 30(6): 1205–1214.
- [15] 吕全通, 张旻, 李歆昊, 等. 基于码重分布距离的自同步扰码识别方法[J]. *探测与控制学报*, 2015, 37(5): 7–13.
- LÜ Quantong, ZHANG Min, LI Xinhao, *et al.* Self-synchronized scrambler recognition based on code weight distribution distance[J]. *Journal of Detection and Control*, 2015, 37(5): 7–13.
- [16] LIU X B, KOH S N, CHUI C C, *et al.* A study on reconstruction of linear scrambler using dual words of channel encoder[J]. *IEEE Transactions on Information Forensics and Security*, 2013, 8(3): 542–552.
- 李歆昊: 男, 1989 年生, 博士, 主要研究方向为信道编码识别.
- 张旻: 男, 1966 年生, 教授, 博士, 主要研究方向为通信信号处理、智能计算.
- 韩树楠: 男, 1989 年生, 博士, 主要研究方向为信道编码识别.