

# 数字调制信号识别的特征参数优化方法

刘明骞 李兵兵 赵 雷

(西安电子科技大学综合业务网理论与关键技术国家重点实验室 西安 710071)

**摘 要** 针对数字调制信号识别中特征参数数目多和特征冗余的问题,提出一种数字调制信号识别的特征参数优化方法。首先利用正交实验对常用的 20 个特征参数进行优化选择,然后利用 RBF 神经网络识别 9 种数字调制信号,最后分别与主分量分析方法(PCA)和核主分量分析方法(KPCA)进行比较。仿真结果表明,该方法在高斯和多径信道下均能够有效地对特征参数进行优化选择,比 PCA 方法和 KPCA 方法有更好的优化能力。

**关键词** 调制识别,特征参数选择,正交实验,主分量分析,核主分量分析

中图分类号 TN391.4 文献标识码 A

## Feature Optimization for Digital Modulation Signals Recognition

LIU Ming-qian LI Bing-bing ZHAO Lei

(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

**Abstract** Aiming at the problem of feature parameter numerousness and feature redundancy in the recognition of digital modulation signals, a feature optimization method for digital modulation signals recognition was proposed in this paper. Firstly, the method optimized the feature parameters of twenty selected features by orthogonal experiment. And then it recognized nine kinds of digital modulation signals using RBF neural network. Finally, the method compared to PCA method and KPCA method. The simulation results show that the method is able to optimize feature parameters effectively in Gaussian and multipath channel, and has much better optimization ability than PCA and KPCA methods.

**Keywords** Modulation recognition, Feature parameter selection, Orthogonal design experiment, PCA, KPCA

## 1 引言

在现代数字通信中,数字信号的调制识别有着重要的作用,它是信号检测与解调的关键步骤,在军事、民用中都扮演着重要的角色。随着通信系统多样性和复杂性的增加,数字信号的调制识别领域也涌现出了许多新的理论和方法。在这些理论和方法中,使用了大量的特征参数。由于在实际应用时,数据集中有着大量无关或冗余的特征参量,会对模式识别产生不利影响,造成“维数灾难”等问题,因此需要根据识别任务来确定选用的特征参数。如何从这些特征中选出识别性能较好、稳健性较强的互补特征参数,对于理论研究和实际应用都有着十分重要的意义<sup>[1]</sup>。

目前选择特征参数的方法主要有两种:一类是根据实验将特征参数分为若干类,再通过对对比试验在各类特征参数中选出一类;另一种方法是对一类特征参数通过实验确定特征向量的阶数。这两种方法都是为了最终得到一类特征参数或简单地加上其他几个特征参数的组合,没有不同类型特征参数之间的组合。为此,正交实验的特征参数选择方法有效地解决了这些问题。近年来,利用正交实验方法研究了信号识别中的特征参数选择,主要集中在语音信号识别<sup>[2-4]</sup>和病态噪

声识别领域<sup>[5]</sup>,而在数字调制信号识别领域研究甚少。

针对数字调制信号识别中特征参数数目多和特征冗余的问题,本文首先利用正交实验的方法对数字调制信号识别中常用的 20 个特征参数进行优化选择,然后采用优化后的特征在高斯和多径信道下分别识别 9 种数字调制信号集合,最后与基于主分量分析法(PCA)的特征参数选择和基于核主分量分析法(KPCA)的特征参数选择进行对比。仿真结果表明,基于正交实验的特征参数优化选择是有效可行的,数字调制信号的识别性能不是随着特征参数的增加而单调增加的,并且该方法的优化效果优于 PCA 方法和 KPCA 方法。

## 2 实验数据与实验方法

### 2.1 数字调制信号识别的特征参数集

目前,在数字调制信号识别中,经常用到瞬时特征、小波变换特征、高阶累积量特征、高阶矩特征和分形理论特征等<sup>[6-9]</sup>。本研究从上述特征中选取 20 个常用的特征参数作为数字调制信号识别的特征参数集,包括归一化中心瞬时幅度的峰值 AMPLITUDE\_matrix;高阶累积量特征 ANG\_C40 和 ANG\_C41;6dB 带宽 B\_F\_6 和 26dB 带宽 B\_F\_26;高阶累积量与高阶矩组合特征 C41\_M21;四阶累积量 C<sub>42</sub>;六阶累积量

到稿日期:2010-12-10 返修日期:2011-03-25 本文受国家“863”高技术研究发展计划项目(2007AA01Z288),国家自然科学基金项目(60772138),高等学校学科创新引智计划项目(B08038)资助。

刘明骞(1982-),男,博士生,主要研究方向为通信信号处理、模式识别,E-mail:mqliu@mail.xidian.edu.cn;李兵兵(1955-),男,博士,教授,主要研究方向为通信信号处理、数字通信;赵 雷(1985-),男,硕士生,主要研究方向为模式识别。

$C_{63}$ ; 方差系数特征 COV; 信号小波变换后的方差系数 CWT\_COV; 零中心归一化非弱信号段瞬时频率绝对值标准偏差 FREQUENCY\_matrix; 归一化中心瞬时功率谱密度幅度最大值 GAMMAX; 信号瞬时相位的高阶矩 K2 和 K3; 范数熵特征值 NOEN; 零中心归一化瞬时幅度绝对值标准偏差 SIGMA\_aa; 零中心归一化非弱信号段瞬时频率绝对值标准偏差 SIGMA\_af; 瞬时零中心非弱信号段瞬时相位非线性分量绝对值的标准偏差 SIGMA\_ap; 瞬时零中心非弱信号非线性分量标准偏差 SIGMA\_dp 和谱对称性 SYMMETRY\_spectrum。

## 2.2 正交实验方法

正交实验方法是从全面实验的样本点中挑选出部分有代表性的点来做实验, 这些点具有“均匀”和“整齐”的特点, 其作用是只用较少的实验次数就可以找出因素水平间的最优搭配或由实验结果推断出最优搭配。本次试验由于原始特征个数  $m=20$ , 根据  $2^{n-1} \leq m \leq 2^n$  需要做  $2^n$  (即 32) 次实验。

正交实验方法对数字调制信号识别的特征参数进行优化选择, 具体分为 3 个步骤。

(1) 构造正交表: 一般正交表的构造需要用到很艰深的数学理论, 但可以将特征的选择和不选择看成特征是二水平的 (分别用 1 和 0 表示), 这样正交表的构造就变得很容易。用文献[5]中介绍的 Hadamard 矩阵法构造正交表, 具体构造法如下:

$$\begin{aligned} \text{记 } H_2 &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \text{ 令 } H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}, \\ H_8 &= \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix}, H_{16} = \begin{bmatrix} H_8 & H_8 \\ H_8 & -H_8 \end{bmatrix}, \\ H_{32} &= \begin{bmatrix} H_{16} & H_{16} \\ H_{16} & -H_{16} \end{bmatrix} \end{aligned} \quad (1)$$

将矩阵中所有的 -1 元素改为 0 元素, 由于  $H_{32}$  的第 1 列全为 1, 不满足下述正交表特征中的第一个特征, 故去除第 1 列。由于用到的特征参数只有 20 个, 因此需要做 20 个因子的实验, 即只需 20 列, 故还需去掉 11 列。为简单起见, 去掉最后的 11 列, 这样就得到了一个  $32 \times 20$  的矩阵, 即为所要构造的正交表, 记为  $(a_{ij})_{32 \times 21}$ 。这样构造出来的正交表有下面两个特征: 一是表中代表因子水平的数字出现的次数是相等的; 二是在任意两列中, 代表两个因子的搭配方式的有序对出现的次数是相等的。

(2) 正交实验: 由于正交表具有以上两条性质, 因此用它来安排实验时, 各因子的各种水平搭配是均衡的。正交表中每行表示一个实验方案, 是各个因素所处状态的一种组合, 每列数字表示对应因子所处的状态, 按正交表对应的实验方案完成实验, 将结果填入正交表中。

(3) 正交实验分析: 对特征进行正交试验时, 其目的就是比较每个实验方案识别率, 以分析各个特征参数对识别率的影响, 并得出理论上较优的特征参数组合。其方法是通过计算实验的  $K$ 、 $S$  参数来完成特征的优化选择, 其计算方法如下:

$$K(1, j) = \frac{1}{16} \sum_{a_{ij}=1} r(i), j=1, 2, \dots, 20 \quad (2)$$

$$K(0, j) = \frac{1}{16} \sum_{a_{ij}=0} r(i), j=1, 2, \dots, 20 \quad (3)$$

$$S(j) = |K(1, j) - K(0, j)|, j=1, 2, \dots, 20 \quad (4)$$

式中, 第  $i$  号实验方案的识别率为  $r(i)$ ;  $K(1, j)$  表示选用第  $j$

个特征参数时的平均识别率;  $K(0, j)$  表示不选用第  $j$  个特征参数时的识别率;  $S(j)$  表示第  $j$  个特征参数选用与否对识别结果的影响程度, 称为极差。极差越大, 则说明它的选用与否对识别结果的影响越大。再由此确定理论较优方案, 如果  $K(1, j) \geq K(0, j)$ , 说明选用该特征时识别率较高。否则, 说明不选用该特征时识别率较高。

## 2.3 PCA 和 KPCA 实验方法

主分量分析 (PCA) 方法<sup>[10]</sup> 是将原来的输入特征参数重新组合成一种新的互不相关的几个综合特征参数来代替原来的特征参数, 同时根据具体的要求, 从中选取几个综合特征参数尽可能多地反映原来特征参数的信息。为了尽可能多地反映原来特征的信息, 常用方法是根据新的特征参数的方差来衡量其信息量, 将方差贡献率最大的前  $p$  个综合特征参数作为其主成分分量, 即新的综合特征。这样, 就完成了新的特征参数的提取。

假定第  $k$  个特征参数为  $X_k = (x_1, x_2, \dots, x_n)^T$ ,  $n$  是特征参数的维数。首先, 求出  $N$  个特征参数的协方差矩阵  $S$ :

$$S = \frac{1}{N} \sum_{k=1}^N (X_k - \bar{X})(X_k - \bar{X})^T \quad (5)$$

式中,  $\bar{X} = \frac{1}{N} \sum_{k=1}^N X_k$ , 表示全体特征参数的平均值。不难发现,  $S$  为对称矩阵。然后计算  $S$  矩阵的特征值  $\lambda_1, \lambda_2, \dots, \lambda_n$  和对应的特征向量  $u_1, u_2, \dots, u_n$ 。对特征值从大到小进行排列:  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , 则特征值越大, 对应的特征向量的方差贡献率越大。定义第  $i$  个主分量  $y_i$  的方差贡献率<sup>[11]</sup> 为:

$$\phi_i = \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} \quad (6)$$

则前  $m$  个主分量的累计方差贡献率为:

$$\phi(m) = \frac{\sum_{i=1}^m \lambda_i}{\sum_{k=1}^n \lambda_k} \quad (7)$$

本文选择累计方差贡献率达到 85% 的前  $m$  个主分量作为提取后的特征参数。前  $m$  个主分量的计算公式为:

$$Y = U^T X_k \quad (8)$$

式中,  $U = (u_1, u_2, \dots, u_m)$ ,  $Y = (y_1, y_2, \dots, y_m)$ 。

核主分量分析 (KPCA) 方法<sup>[12]</sup> 实际上是将核方法应用到 PCA 方法中, 通过使用“核技巧”将输入特征参数映射到高维空间, 使输入特征参数具有更好的空间可分性的一种方法。其基本思路是将原始输入特征参数  $x$  利用核函数映射到高维特征空间  $\Phi(x)$ , 然后在高维特征空间中对  $\Phi(x)$  进行线性 PCA 计算, 即相当于在输入空间中对输入特征参数  $x$  的非线性 PCA 计算。本次试验选取的核函数为高斯核函数, 其表达式为:

$$k(x, y) = \exp \left\{ -\frac{\|x - y\|^2}{2\sigma^2} \right\} \quad (9)$$

## 3 仿真结果及分析

### 3.1 正交实验结果及分析

为了验证正交实验的有效性和可行性, 采用上述的 20 个特征参数在高斯信道下通过正交设计对信号源 16QAM、2ASK、4ASK、BPSK、QPSK、2FSK、4FSK、MSK 以及 OFDM 这 9 种信号, 利用 RBF 神经网络分类器进行了识别。表 1 是在高斯信道下信噪比为 0dB 时通过正交实验对 20 个特征参

数所建立的正交表,其中1表示选用该特征,0表示丢弃该特征,R表示识别率。表2是在不同信噪比下应用正交实验分

析得到的优化特征参数组合进行识别的识别率,R表示识别率。表3是在不同信噪比下各个特征的极差。

表1 2水平因子正交表

实验方案编号	特征参数编号																				R
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.6556
2	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0.6321
3	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	0.6451
4	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0.6383
5	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0.5605
6	0	1	0	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0	0	0.6951
7	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0.7494
8	0	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0	0	1	0	0.5840
9	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	0.5617
10	0	1	0	1	0	1	0	0	1	0	1	0	1	0	1	1	0	1	0	1	0.7611
11	1	0	0	1	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	1	0.8593
12	0	0	1	1	0	0	1	0	1	1	0	0	1	1	0	1	0	0	1	1	0.7760
13	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0.7407
14	0	1	0	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0	0	0.6642
15	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0	0	0.6833
16	0	0	1	0	1	1	0	0	1	1	0	1	0	0	1	1	0	0	1	0	0.7333
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0.6543
18	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	0.6086
19	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	0	0	1	1	0	0.6981
20	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	1	0	0	0.6364
21	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0.5642
22	0	1	0	0	1	0	1	1	0	1	0	0	1	0	1	0	1	0	1	1	0.6932
23	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	0.7642
24	0	0	1	0	1	1	0	1	0	0	1	0	1	1	0	0	1	1	0	1	0.6840
25	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3593
26	0	1	0	1	0	1	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0.7630
27	1	0	0	1	1	0	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0.8648
28	0	0	1	1	0	0	1	0	1	0	0	1	1	0	0	1	1	0	0	0	0.7858
29	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	1	0.7414
30	0	1	0	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0	1	1	0.6846
31	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	1	1	1	0.8086
32	0	0	1	0	1	1	0	0	1	1	0	1	0	0	1	0	1	1	0	1	0.7432

表2 优化特征的识别率

特征参数编号	信噪比																				R
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	1	1	1	0.8698
5	0	0	0	1	0	0	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0.9642
10	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0	1	1	1	0.9883
15	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	1	1	1	1	1	0.9895
20	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	1	1	1	1	1	0.9895

表3 特征的极差分析

特征参数编号	信噪比				
	0	5	10	15	20
1	0.0108	0.0473	0.0127	0.0147	0.0193
2	0.0821	0.0885	0.0754	0.0270	0.0296
3	0.0723	0.0709	0.0789	0.0560	0.0535
4	0.0122	0.0052	0.0215	0.0269	0.0329
5	0.0271	0.0287	0.0204	0.0282	0.0263
6	0.0314	0.0111	0.0164	0.0202	0.0230
7	0.0233	0.0057	0.0044	0.0015	0.0015
8	0.0667	0.0531	0.0485	0.0665	0.0803
9	0.0257	0.0088	0.0130	0.0022	0.0041
10	0.0402	0.0421	0.0170	0.0197	0.0238
11	0.0087	0.0166	0.0679	0.0450	0.0114
12	0.0036	0.0291	0.0276	0.0235	0.0355
13	0.0466	0.0203	0.0073	0.0022	0.0040
14	0.0565	0.1050	0.0940	0.0513	0.0359
15	0.0871	0.0005	0.0582	0.0309	0.0468
16	0.0071	0.0107	0.0001	0.0030	0.0031
17	0.0072	0.0068	0.0039	0.0059	0.0037
18	0.0324	0.0345	0.0208	0.0077	0.0027
19	0.0173	0.0139	0.0021	0.0003	0.0005
20	0.0270	0.0125	0.0027	0.0035	0.0021

从表1和表2可以看出,不同的特征组合的识别率有所不同,经过正交实验分析之后优化的特征参数组合的识别率显然高于原始特征的识别率。但是,分析后的识别率不一定高于正交实验表中的最大识别率,这是因为在正交实验分析时,把各个特征割裂开来进行分析,并没有分析其组合之后的识别特性。由表3可知,可以根据原始特征的极差特性来分析各个特征对识别率的影响情况。例如,在0dB情况下,第2、3、15个特征的极差比较大,而由表2可以看出,其第2、3个特征应该不选用,第15个特征应该选用以达到较高的识别率,由表1可以得到该分析是正确的。

### 3.2 对比实验结果及分析

正交试验方法分别与PCA方法、KPCA方法和利用原始特征的方法在高斯信道下和多径信道下的识别性能对比如图1和图2所示。

从图1可以看出,这3种方法在高斯信道下都能够较好地实现优化数字调制信号的特征参数。由于原始的20个特征的相关性太强,经过PCA方法降维之后,得到的新特征只有一个。尽管这一个特征对于这9个源信号的识别有些少,

但其识别率接近于原始的 20 个特征的识别率。原始 20 个特征经过 KPCA 降维之后,得到了 17 个特征,虽然识别率相对于原始 20 个特征的识别率有所提高,但提高的程度不大,说明这 20 个原始特征的非线性不是很强。正交试验方法的识别率均高于 PCA 方法和 KPCA 方法的识别率。由图 2 可知,正交试验在多径环境下仍能保持较高的识别率。

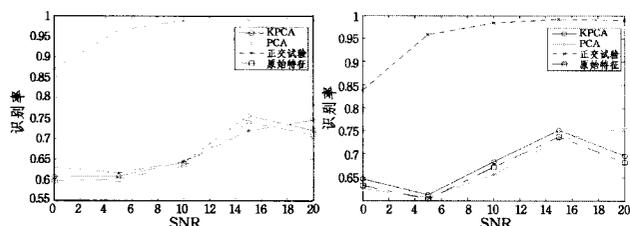


图 1 高斯信道下识别性能对比图 图 2 多径信道下识别性能对比图

这 3 种优化特征参数方法的共同点在于减少无用特征,保留有用特征,用较少的特征达到更高的识别率。3 种方法的不同点在于:PCA 和 KPCA 方法都是利用方差贡献率来选取较优特征,PCA 方法的较优特征是由原始 20 个特征经过线性变换得到的新特征;而 KPCA 方法的较优特征是由原始的 20 个特征先经过非线性变换升维之后,再经过线性变换降维得到的新特征;正交试验法是根据构造的正交表中的特征组合进行试验分析得到较优特征,较优特征只是从原始特征中选取出来的特征,不经过任何变换,不是新特征。选用特征参数的相关性太强,非线性不强,导致 PCA 以及 KPCA 方法的优点没有凸显出来,尽管达到了一定的优化特征参数,而相对于正交试验方法识别率明显较低;在正交试验方法的分析中,没有考虑各个特征参数的相关性,即其组合特征的识别率,故其分析之后的较优特征有可能不是最优特征组合。

**结束语** 本文用正交试验方法、PCA 方法和 KPCA 方法对 9 种数字调制信号的 20 个特征参数进行了优化选择,然后分别在高斯和多径信道下利用优化特征对数字调制信号进行

了识别。结果表明,基于正交实验的数字调制信号特征参数优化的效果均优于 PCA 和 KPAC 方法。

## 参考文献

- [1] 吕铁军,王河,肖先赐.新特征选择方法下的信号调制识别[J].电子与信息学报,2002,24(5):661-666
- [2] 韩志艳,王健,王旭.基于正交实验设计的语音识别特征参数优化[J].计算机科学,2010,37(1):214-216
- [3] 杨大利,徐明星,吴文虎.语音识别中一种新的特征参数选择方法[J].清华大学学报:自然科学版,2003,43(1):79-82
- [4] 杨大利,徐明星,吴文虎.语音识别特征参数选择方法研究[J].计算机研究与发展,2003,40(7):963-969
- [5] 彭策,熊屹,陈文西.病态噪声识别特征参数的优化选择[J].中国生物医学工程学报,2007,26(5):675-679
- [6] 李佩,孙进平,成功,等.一种基于支持向量机的数字调制识别方法[J].电路与系统学报,2010,15(3):7-12
- [7] 张弛,吴瑛,周欣.基于高阶累积量的数字调制信号识别[J].数字采集与处理,2010,25(5):575-579
- [8] 李楠,曲长文,平殿发,等.基于分形理论的辐射源识别算法[J].航天电子对抗,2010,26(2):62-64
- [9] Zhang Jing-jing, Li Bing-bing. A new modulation identification scheme for OFDM in multipath rayleigh fading channel[C]//International Symposium on Computer Science and Computational Technology. Shanghai, China, IEEE, 2008:793-796
- [10] Subasi A, Gursoy M I. EEG signal classification using PCA, ICA, LDA and support vector machines[J]. Expert Systems with Applications, 2010,37(12):8659-8666
- [11] 任若恩,王惠文.多元统计数据分析[M].北京:国防工业出版社,1997:92-109
- [12] Xu Y, Lin C, Zhao W. Producing computationally efficient KPCA-based feature extraction for classification problems[J]. Electronics Letters, 2010,46(6):452-453
- [13] Wu H C, Huang S H S. Masquerade Detection Using Command Prediction and Association Rules Mining[C]//2009 International Conference on Advanced Information Networking and Applications. Aina, IEEE, 2009:552-559
- [14] 田新广,段冰毅,程学旗.基于 Shell 命令和多重行为模式挖掘的用户伪装攻击检测[J].计算机学报,2010,33(4):697-705
- [15] Coull S E, Branch J W, Szymanski B K, et al. Sequence Alignment for Masquerade Detection[J]. Computational Statistics & Data Analysis, 2008,52(8):4116-4131
- [16] 田新广,段冰毅,孙春来,等.采用 shell 命令和隐 markov 模型进行网络用户行为异常检测[J].应用科学学报,2008,26(02):175-181
- [17] Tian Xin-guang, Duan Mi-yi, Li Wen-fa, et al. Anomaly Detection of User Behavior Based on Shell Commands and Homogeneous Markov Chains[J]. Chinese Journal of Electronics, 2008,17(2):231-236
- [18] Lane T. Machine Learning Techniques for the Computer Security Domain of Anomaly Detection[D]. West Lafayette, Indiana: Purdue University, 2000
- [19] Maxion R A, Townsend T N. Masquerade Detection Using Truncated Command Lines[C]//International Conference on Dependable Systems & Networks. Washington, DC, USA, 2002:219-228
- [20] Lane T, Brodley C E. An Empirical Study of Two Approaches to Sequence Learning for Anomaly Detection[J]. Machine learning, 2003,51(1):73-107
- [21] 孙宏伟,田新广,李学春,等.一种改进的 IDS 异常检测模型[J].计算机学报,2003,26(11):1450-1455
- [22] Schonlau M, DuMouchel W, Ju W H, et al. Computer Intrusion: Detecting Masquerades[J]. Statistical Science, 2001,16(1):58-74
- [23] Szymanski B K, Zhang Y Q. Recursive Data Mining for Masquerade Detection and Author Identification[C]//Proceedings of the 5th IEEE System, Man and Cybernetics Information Assurance Workshop. West Point, NY, USA, 2004:424-431
- [24] Kim H S, Cha S D. Empirical Evaluation of SVM-based Masquerade Detection Using Unix Commands[J]. Computers & Security, 2005,24(2):160-168
- [25] Wang K, Stolfo S J. One Class Training for Masquerade Detection[C]//ICDM Workshop on Data Mining for Computer Security(DMSEC 03). Citeseer, 2003
- [26] Karlin S, Taylor H M. A First Course in Stochastic Processes [M]. Second Edition, Beijing: Post & Telecom Press, 2007
- [27] 田新广.基于主机的入侵检测方法研究[D].长沙:国防科学技术大学, 2005