基于信息增益的无线通信信号指纹构建及 识别机制研究

黄伟庆^{1,2},杨召阳^{1,2},魏 冬^{1,2},张 萌^{1,2},王 文^{1,2},叶 彬^{1,2}

¹中国科学院信息工程研究所 北京 中国 100093 ²中国科学院大学 北京 中国 100049

摘要 为实现对特定区域电磁空间的有效监管,本文建立针对无线通信信号的唯一识别及标定机制,利用无线通信信号调制方式、发射机器件一致性及通信协议等多种因素造成的信号差异,构建表征无线通信信号完备差异信息的指纹集,并引入信息增益评价机制评估指纹集中各特征对差异信息的贡献度,建立精简有效的电磁指纹。在此基础上,设计基于支持向量机的指纹匹配算法,实现无线通信信号身份的有效识别。仿真实验结果表明,本文方法在接收信噪比大于4dB时可实现特定区域电磁空间内无线通信信号身份的有效识别。

关键词 无线通信信号;指纹;信息增益;身份识别 中图法分类号 TN918.91 DOI号 10.19363/J.cnki.cn10-1380/tn.2020.11.02

Fingerprint construction and identity recognition of Wireless signal based on information gain

HUANG Weiqing^{1,2}, YANG Zhaoyang^{1,2}, WEI Dong^{1,2}, ZHANG Meng^{1,2}, WANG Wen^{1,2}, YE Bin^{1,2}

¹Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China ²University of Chinese Academy of Sciences, Beijing 100049, China

Abstract With the question of supervising of the special electromagnetic space. This paper aims to construct the unique identity of the wireless signal. Through fusing the wireless signal difference features caused by modulation type, device uniformity and communication protocol, the fingerprint set of wireless signal is constructed. Meanwhile, aiming at evaluating the fingerprint contribution degree to identity recognition of signal and constructing refined and valid fingerprint set, this paper introduces into the information gain to evaluate every fingerprint's degree to identity recognition and support vector machine (SVM) method to match refined fingerprint subset. Simulation and experiment results show that, this paper's fingerprint set method can effectively distinguish the identity of different wireless signal in special electromagnetic space under the condition that the signal-noise ratio is larger than 4 dB.

Key words wireless signal; fingerprint; information gain; identity recognition

1 引言

随着无线通信技术的发展及广泛应用,电磁空 间中无线信号呈现复杂化及多样化的特点,如何对 电磁空间中的无线信号进行唯一识别是认知无线 电、军事通信对抗及信息安全等领域重点关注的问 题之一。

当前,针对无线通信信号识别的研究主要包括 调制方式识别及辐射源识别等。其中,调制方式识别 是利用不同调制技术对信号特征改变的差异性来进 行识别^[1-14],主要利用瞬时特征^[1-2]、累积量特征^[3-9]、 循环统计量特征^[10-14]、变换域特征^[15-16]及星座结构 特征^[17]等,目前已实现了DSB、FM、MPSK、MFSK、 MQAM、MASK、OFDM 等调制方式的识别。在辐 射源识别方面,主要基于射频器件的不一致性对信 号及其辐射源进行识别,如文献[18-32]利用辐射源 发射调制参数稳定性差异、非线性特征差异等,并利 用希尔伯特-黄变换^[18]、贝叶斯方法^[19]、能量包络的

通讯作者:魏冬,博士,副研究员,Email:weidong@iie.ac.cn。

本课题得到国家自然科学基金项目(No. 61501458)资助,中国科学院战略性先导 C 类 XDC02000000 专项支持。 收稿日期: 2018-12-27;修改日期: 2019-03-13;定稿日期: 2020-09-22

多项式拟合^[20]、双谱分析^[21]、短时傅里叶变换^[22]、 维格纳分布^[23]、小波变换^[24-25]、累积量^[26]、二次型 时频分析^[27-28]、固有时间尺度分解(ITD)^[29]及变分模 态分解(VMD)^[30]等理论对差异进行表征,进而实现 辐射源的识别。基于上述分析可知,当前信号识别的 研究主要基于调制方式或射频器件一致性,并未对 信号的差异进行全面利用,如通信协议对发射参数 上造成的差异也可以作为识别信号的特征,包括突 发间隔、突发持续时间、跳频间隔等。事实上,若能 基于信号全部差异特征,一方面可增加用于识别信 号的判决依据,提升算法识别的准确率及鲁棒性; 另一方面也可对信号的差异性进行全面描述,建立 针对特定信号的唯一表征,实现对特定电磁空间内 无线通信信号的唯一识别。

为达到上述目的,本文提出一种基于信息增益 的无线通信信号指纹构建算法。该算法融合信号调 制差异、辐射源差异及通信协议差异,建立针对特定 电磁空间内无线通信信号的指纹集;然后,为提取 贡献度大的指纹特征并祛除指纹集中冗余特征信息, 联合余弦相似度及信息增益评价机制建立对无线通 信信号指纹特征的评价机制,祛除各类信号差异特 征间的冗余信息,构建表征指纹的最佳差异特征组 合方式。最后,基于所构建的指纹,设计基于支持向 量机的无线通信信号指纹分类算法,对特定电磁空 间内无线通信信号进行唯一标定及识别。

2 无线通信信号接收模型

在莱斯信道下,无线通信信号非合作接收模型 可表示为:

$$s(t) = \alpha \sum_{k} \sum_{l} r(t) g\left(t - k\left(T_{s} + \delta_{T_{s}}\right)\right)$$

$$e^{j\left(2\pi\left(f_{c} + \delta_{f_{c}}\right)t + \theta\right)} e^{j2\pi l\Delta f_{c}} + \lambda_{3}\left(A(t)\right)^{3} + \lambda_{5}\left(A(t)\right)^{5} + n$$

$$\exists t \to \cdot$$
(1)

 $\alpha = \sqrt{(x+\beta)^2 + y^2}$ 为信道衰减系数, x,y为均 值为 0, 方差为 σ^2 的高斯变量, β 为直射分量。r(t)为信号基带调制码元序列。 $g(t-k(T_s+\delta_{T_s}))$ 为信号 基带码元波形, T_s 为码元持续时间, δ_{T_s} 为码元时 间误差。 f_c 为信号载波频率, δ_{f_c} 为载波误差, θ 为 信号初始相位。l表示子载波个数。 Δf_c 为载波间隔。 λ_3 , λ_5 表示信号中的非线性成分,由非线性器件产 生;由于只有奇数阶非线性量能落在有用信号的带 宽内,并且影响较大的主要为三阶,五阶非线性分,因此这里只考虑三阶和五阶非线性量。*n*为高斯白噪声。

基于上述模型,可对特定场所内无线通信信号 差异进行归类:

1) 信号发射参数差异

该差异体现在码元持续时间 T_s ,成型滤波器方 式 $g(t-k(T_s+\delta_{T_s}))$ 。

2) 调制方式差异

该差异体现在信号瞬时幅值、瞬时相位及瞬时 频率的统计变化。

3) 调制参数稳定性差异

该差异体现在码元持续时间 T_s 及载波 f_c 稳定性上,即 δ_T 和 δ_f 。

4) 非线性差异

该差异体现在信号非线性成分 $\lambda_3(A(t))^3$ 和 $\lambda_5(A(t))^5$ 。

5) 通信协议差异

该差异体现在信号的同步序列、突发时间间隔、 突发持续时间、载频分布及跳频间隔等特征。

3 无线通信信号指纹构建

基于上述差异性,从信号中提取含有差异性的 特征并构建指纹集:{ $\Gamma_1, \Gamma_2, \Gamma_3$ },该指纹集共包含 以下三类特征,其中 Γ_1 为信号瞬时特征:

$$\Gamma_{1} = \left\{ \gamma_{\max} \sigma_{ap} \sigma_{dp} \sigma_{aa} \sigma_{af} \right\}$$
(2)

 Γ_2 为信号小波特征:

$$\Gamma_2 = \left\{ V _ WT_{a=5} \quad V _ WT_{a=9} \right\}$$
(3)

 Γ_3 为信号统计特征:

$$\Gamma_{3} = \begin{cases} c_{2,0} & c_{2,1} & c_{4,0} \\ c_{4,1} & c_{4,2} & c_{6,0} \\ c_{6,1} & c_{6,3} & c_{8,0} \\ c_{r2,0} & c_{r2,1} & c_{r4,0} \\ c_{r4,1} & c_{r4,2} & c_{r6,0} \\ c_{r6,1} & c_{r6,3} & c_{r8,0} \end{cases}$$
(4)

上述 25 个特征构建无线通信信号身份指纹集, 具体每个指纹特征含义如下所述。

1) 瞬时特征

γmax 为零中心归一化瞬时幅度最大值:

$$\gamma_{\max} = \max \left| DFT(a_{cn}(i)) \right|^2 / N_s \tag{5}$$

其中N_s为样本数据长度, a_{cn}为零中心归一化瞬时

幅度, 具体参见公式(6)

$$a_{cn}(i) = \frac{a(i) - \frac{1}{N_s} \sum_{i=1}^{N_s} a(i)}{\frac{1}{N_s} \sum_{i=1}^{N_s} a(i)}$$
(6)

a(i)为信号瞬时幅度。

σ_{av} 为零中心归一化瞬时相位绝对值标准差:

$$\sigma_{ap} = \sqrt{\frac{1}{N_s} \left(\sum \phi_{NL}^2(i) \right) - \left(\frac{1}{N_s} \sum \left| \phi_{NL}(i) \right| \right)^2}$$
(7)

其中 $\phi_{NL}(i)$ 为零中心归一化瞬时相位。

 σ_{dn} 为零中心归一化瞬时相位标准差:

$$\sigma_{dp} = \sqrt{\frac{1}{N_s} \left(\sum \phi_{NL}^2(i) \right) - \left(\frac{1}{N_s} \sum \left(\phi_{NL}(i) \right) \right)^2} \qquad (8)$$

 σ_{aa} 为零中心归一化幅度绝对值标准差:

$$\sigma_{aa} = \sqrt{\frac{1}{N_s} \left(\sum a_{cn}^2 \left(i \right) \right) - \left(\frac{1}{N_s} \sum \left| a_{cn} \left(i \right) \right| \right)^2} \qquad (9)$$

 σ_{af} 定义为零中心归一化瞬时频率绝对值标准差:

$$\sigma_{af} = \sqrt{\frac{1}{N_s} \left(\sum f_N^2(i) \right) - \left(\frac{1}{N_s} \sum \left| f_N(i) \right| \right)^2} \qquad (10)$$

其中 $f_N(i)$ 为零中心归一化瞬时频率,具体参见 公式(11)

$$f_N(i) = \frac{f(i) - \frac{1}{N_s} \sum_{i=1}^{N_s} f(i)}{r_s}$$
(11)

f(i)为信号瞬时频率, r_s 为信号码率。

以上瞬时特征基于信号瞬时幅度、相位及频率 信息进行计算,可直接反映信号调制方式、非线性及 通信协议所产生的差异性。

2) 小波特征

本文采用的小波特征 *V_WT* 用于体现非线性特征,在一个符号内对接收信号进行 Haar 小波变换:

$$cwt(a,\tau) = \int s(t)\psi_{a}^{*}(t)d(t)$$

$$= \frac{1}{\sqrt{a}}\int s(t)\psi^{*}\left(\frac{t-\tau}{a}\right)d(t)$$

$$= \frac{4s_{i}}{j\sqrt{a}\omega_{c}}\sin^{2}\left(\omega_{c}\frac{a}{4}\right)e^{j(\omega_{c}\tau+\theta+\phi_{i})}$$

$$+ \frac{4\lambda_{3}s_{i}^{3}}{j\sqrt{a}\omega_{c}}\sin^{2}\left(\omega_{c}\frac{a}{4}\right)e^{j(\omega_{c}\tau+\theta+3\phi_{i})}$$
(12)

其中 $s_{i} = \sqrt{a_{i}^{2} + b_{i}^{2}}, \quad \phi_{i} = \tan^{-1}(b_{i}/a_{i}), \quad a_{i}, \quad b_{i}$
为信号

调制码元序列, $\omega_c = 2\pi (f_c + \delta_{f_c})$ 。

由公式(12)可以看出信号的 Haar 小波变换 $|cwt(a,\tau)|$ 在一个符号内具有固定值,其同信号的调 制序列 s_i 、频率 f_c 、载波误差 δ_{f_c} 及设备非线性系数 λ_3 有关。为祛除随机性对小波特征造成的影响,本文 将特征 V_WT 定义为小波幅度 $|cwt(a,\tau)|$ 的方差,并 选 取 两 种 小 波 尺 度 作 为 指 纹 特 征 : $\{V_WT_{a=5}, V_WT_{a=9}\}$ 。

3) 高阶统计量特征

高阶统计量具有抗高斯噪声的特点,本文利用 高阶累积量和循环累积量作为指纹特征。

高阶累积量可有效反映平稳信号的变化特点, 本文主要利用以下9种累量特征:

$$\left\{c_{2,0}, c_{2,1}, c_{4,0}, c_{4,1}, c_{4,2}, c_{6,0}, c_{6,1}, c_{6,3}, c_{8,0}\right\}$$

上述各累积量可基于信号累积矩求得^[31]:

$$c_{k,n}(x) = \sum_{U_{p=1}^{q} I_{p} = I} (-1)^{q-1} (q-1)! \prod_{p=1}^{q} m_{k_{p},n}(x)$$
(13)

其中 $\sum_{U_{p=I}^{q} I_{p}=I} (\cdot)$ 表示空间 I所有无交连非空集合的分

割, k_p 代表分割 I_p 中元素个数。

其中累积矩定义为:

$$m_{k,n}\left(x\right) = E\left(x^{k-n}\left(x^*\right)^n\right) \tag{14}$$

联合式(14),(15)得:

$$c_{2,0} = m_{2,0} \tag{15}$$

$$c_{2,1} = m_{2,1} \tag{16}$$

$$c_{4,0} = m_{4,0} - 3m_{2,0}^2 \tag{17}$$

$$c_{4,1} = m_{4,1} - 3m_{2,1}m_{2,0} \tag{18}$$

$$c_{4,2} = m_{4,2} - m_{2,0}^2 - 2m_{2,1}^2 \tag{19}$$

$$c_{6,0} = m_{6,0} - 15m_{4,0}m_{2,0} + 30m_{2,0}^3$$
(20)

$$c_{6,1} = m_{6,1} - 5m_{4,0}m_{2,1} - 10m_{4,1}m_{2,0}$$
(21)

$$+30m_{21}m_{20}^2$$

$$c_{6,3} = m_{6,3} - 9m_{4,2}m_{2,1} - 6m_{4,1}m_{2,0} + 18m_{2,1}m_{2,0}^2 + 12m_{2,1}^3$$
(22)

$$c_{8,0} = m_{8,0} - 28m_{6,0}m_{2,0} - 35m_{4,0}^2$$

$$+420m_{4,0}m_{2,0}^2 - 630m_{2,0}^4$$
(23)

基于上述公式,在符号等概率、零均值及平均功 率归一化为 E 假设下,针对 4G 移动通信系统调制 信号以及集群对讲机系统调制信号,计算了相应的 信号累积量;其中 4G 移动通信系统采用 OFDM 制式,每个子载波的调制类型包括 MPSK 或 MQAM;

集群对讲机系统参考 DMR、PDT 等制式,调制类型 为 *M*FSK,具体如表 1 所示:

				表1	累积量				
				Table 1	Cumulant				
	<i>c</i> _{2,0}	<i>c</i> _{2,1}	$c_{4,0}$	$c_{4,1}$	<i>c</i> _{4,2}	<i>c</i> _{6,0}	<i>c</i> _{6,1}	<i>c</i> _{6,3}	<i>c</i> _{8,0}
MFSK	0	Ε	0	0	$-E^2$	0	0	$4E^{3}$	0
MPSK	0	Ε	$-E^2$	0	$-E^2$	0	$4E^3$	$4E^{3}$	$-34E^{4}$
16QAM	0	Ε	$-0.68E^2$	0	$-0.68E^{2}$	0	$2.08E^{3}$	$2.08E^{3}$	$-13.981E^4$
64QAM	0	Ε	$-0.619E^2$	0	$-0.619E^2$	0	$1.797E^{3}$	$1.797E^{3}$	$-11.502E^4$
256QAM	0	Ε	$-0.604E^2$	0	$-0.604E^2$	0	$1.734E^{3}$	$1.734E^{3}$	$-10.97E^4$
OFDM	-	-	0	0	0	0	0	0	0

循环累积量可有效反映非平稳信号的差异,本 文选取的循环累积量特征包括以下9种:

 $\{c_{r_{2,0}}, c_{r_{2,1}}, c_{r_{4,0}}, c_{r_{4,1}}, c_{r_{4,2}}, c_{r_{6,0}}, c_{r_{6,1}}, c_{r_{6,3}}, c_{r_{8,0}}\}$ 其中

$$c_{rk,x}^{\alpha} = \lim_{T \to \infty} \frac{1}{T} \sum c_{k,x}(t,\tau) e^{-j\alpha t}$$

=
$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{U_{p=1}^{q} I_{p}=I} (-1)^{q-1} (q-1)! \prod_{p=1}^{q} m_{k_{p},x}(t,\tau_{p}) e^{-j\alpha t}$$

=
$$\sum_{U_{p=1}^{q} I_{p}=I} (-1)^{q-1} (q-1)! \sum_{\alpha_{1}+\cdots+\alpha_{q}=\alpha} \prod_{p=1}^{q} m_{k_{p},x}^{\alpha_{p}}(\tau_{p})$$

(24)

 $m_{k,x}^{\alpha}(\tau)$ 为信号循环累积矩, α 为循环频率, 可基于如下公式进行计算^[32]:

 $m_{k,x}^{\alpha}(\tau) = \frac{1}{T} \sum_{i=0}^{N-1} x(i) x(i+\tau_1) \cdots x(i+\tau_{k-1}) e^{-j2k\pi\alpha t} \quad (25)$

综上所述,本文构建的无线通信信号指纹集覆 盖了调制方式、辐射源器件一致性及通信协议等差 异。上述指纹集中各特征对信号差异的贡献度是不 同的,并且各特征间存在信息的冗余。

4 基于有效信息增益的指纹特征评价算法

针对指纹特征集中特征贡献度不同以及特征间 冗余的问题,本节提出一种基于有效信息增益的指 纹特征贡献度评价方法,对指纹特征贡献度进行定 量计算,以提取有效指纹特征。

信息增益用以度量信息熵的差值,对随机变量 x(i),其概率分布为 p_i ,则其信息熵定义为:

$$H(p_i) = -\sum_i p_i \log_2(p_i)$$
(26)

其表示随机变量的所包含的信息量。若在条件 D

下随机变量x(i)的条件概率为 q_i/D ,则其信息熵为:

$$H(q_i/D) = -\sum q_i/D\log_2(q_i/D)$$
(27)

其表示在条件 D 下随机变量 x(i)的信息量,则 在条件 D 下随机变量的信息增益为:

$$G(x(i),D) = H(p_i) - H(q_i/D)$$
(28)

该式反映的是在条件 D 下随机变量 x(i)的信息 量或熵的变化。

为此, 单个指纹特征向量 **P**_{*i*,M} 的信息增益可表 示为:

$$G(Q_{M,N}, P_{i,M}) = H(Q_{M,N}) - H(Q_{M,N}/P_{i,M})$$
(29)
其中 $H(Q_{M,N})$ 定义为本文样本 $Q_{M,N}$ 的信息熵:

$$H(Q_{M,N}) = -\sum_{k=1}^{K} \frac{C_k}{Q_{M,N}} \log_2\left(\frac{C_k}{Q_{M,N}}\right)$$
(30)

*M*为样本个数,*N*为指纹集中特征个数,*K*为样本中类别数,*C*_k表示类别*k*中的样本个数。

 $H(Q_{M,N}/P_{i,M})$ 定义在特征向量 $P_{i,M}$ 分割下样本 $Q_{M,N}$ 的信息熵:

$$H\left(Q_{M,N}/\boldsymbol{P}_{i,M}\right) = -\sum_{k=1}^{K} \frac{C_k/\boldsymbol{P}_{i,M}}{Q_{M,N}/\boldsymbol{P}_{i,M}} \log_2\left(\frac{C_k/\boldsymbol{P}_{i,M}}{Q_{M,N}/\boldsymbol{P}_{i,M}}\right)$$
(31)

基于公式(29)可以发现,所定义的指纹特征向量 信息增益尽管能够表示某个特征向量对样本分类的 贡献,但是信息增益倾向于选择取值较多的特征。当 特征 $P_{i,M}$ 取值多时,则每个 $P_{i,M}$ 值下的 C_k 的种类数 k 变少,其值 $C_k/P_{i,M}$ 和 $Q_{M,N}/P_{i,M}$ 会越接近,即 $\left|\log_2 \frac{C_k/P_{i,M}}{Q_{M,N}/P_{i,M}}\right|$ 的取值将变小,因此信息熵 $H(Q_{M,N}/P_{i,M})$ 的值则趋向变小。极端情况下, $P_{i,M}$ 的取值个数等于样本个数时, $H(Q_{M,N}/P_{i,M})=0$, 此特征条件下的信息增益 $G(Q_{M,N},P_{i,M})$ 则会取得最 大值。然而,取值多的特征有时候与分类相关性小。 为解决上述问题,本文引入余弦相似度作为另一个 特征排序指标,定义为每个特征向量 $P_{i,M}$ 和标签向 量 B_{MK} 的余弦相似度

$$C(\mathcal{Q}_{M,N}, \boldsymbol{P}_{i,M}) = \frac{p_{i,M} \cdot \boldsymbol{B}_{M,K}}{\|\boldsymbol{p}_{i,M}\| * \|\boldsymbol{B}_{M,K}\|}$$
(32)

其表示特征向量 $P_{i,M}$ 和标签向量 $B_{M,K}$ 的相似程度,因此可作为对信息增益偏大却与分类相关性小的指纹特征进一步筛选。

综上,本文所设计的基于信息增益的特征评价 如图1所示:





首先,根据公式(29)计算单个指纹特征向量 $P_{i,M}$ 的信息增益, $G(Q_{M,N}, P_{i,M}) = H(Q_{M,N}) - H(Q_{M,N}/P_{i,M})$ 然后,结合信息增益和余弦相似度计算有效信息增益:

 $C_{p} = G(Q_{M,N}, P_{i,M}) * \tilde{C}(Q_{M,N}, P_{i,M})$ (33) $\pm \Psi \tilde{C}(Q_{M,N}, P_{i,M}) \rtimes C(Q_{M,N}, P_{i,M})$ 的归一化数据。

最 后 按 照 特 征 有 效 信 息 增 益 指 数 $C_p = G(Q_{M,N}, P_{i,M}) * \tilde{C}(Q_{M,N}, P_{i,M})$ 对特征 $P_{N,M}$ 进行排序 获得指纹特征排序集 P, 然后通过分类器匹配出合 适的指纹集。

5 基于支持向量机的信号指纹匹配算法

基于上述建立的指纹特征,利用支持向量机 (SVM)^[33]建立相应的指纹匹配算法,以获取合适的 指纹集。如图 2 所示,该算法流程如下:



图 2 SVM 指纹匹配算法流程 Figure 2 SVM fingerprint matching algorithm

1) 基于信息增益的指纹特征评价方法获得指纹 特征排序集 P。

2) 从指纹特征排序集 *P* 中提取第一个特征,并利用训练样本对支持向量机进行训练。

3) 利用训练过的 SVM 对测试样本进行测试, 若测试结果达到分类效果或者达到所有指纹特征分 类时的效果时,则停止进行下一步,此时训练用的 指纹特征就是所获得的匹配指纹子集,若不满足则 进行第4步。

4) 根据指纹特征排序集 P, 按照顺序增加一个 指纹特征重复第2步和第3步中样本训练 SVM 和测 试分类的过程。若测试结果达到所需分类效果, 则停 止。此时训练用的指纹特征集就是所获得的匹配指 纹子集, 若不满足则重复第4 步中上述过程:按顺序 增加一个指纹特征、样本训练和测试分类直到达到 所需分类效果或 P 中的所有指纹特征。

5) 最后, 根据第 4 步的结果即可获得匹配的指 纹特征子集。

6 仿真分析

仿真利用实验室多种设备获取无线通信信号 时域样本,并利用 Matlab 仿真平台对本文所提出 的无线通信信号身份识别算法进行测试和分析。仿 真采用的无线通信设备包括: AM 电台一部, GSM 手机三部, TETRA 手台二部, IDEN 手台一部, 4G 手机一部。

6.1 指纹特征有效性仿真分析

本部分对无线通信信号指纹集中特征进行仿真验证,目的如下:

1)验证指纹特征对信号身份区分的有效性及不同指纹特征间的贡献度差异;

2) 对不同特征在不同信噪比情况下贡献度的变 化进行分析;

3)分析不同特征间对信号身份区分贡献度的关系,评估特征间信息冗余性。

为此,本节采集上述设备信号,在不同接收信

噪比(SNR)的情况下对指纹特征进行仿真分析。考虑 到实验中三台 GSM 设备之间差异相对较小,在全部 设备特征图中差异不易看出,所以单独列出其图。特 征计算数据长度按 2000 个采样点计算。

6.1.1 信号瞬时特征仿真分析

根据采集的无线通信信号时域样本,在不同信 噪比下仿真了瞬时相位,瞬时幅度,瞬时频率。

通过图 3 可以发现, AM 设备发射信号与 IDEN 设备发射信号在特征 σ_{ap} 上反映出的差异较大, 但受 噪声影响 σ_{ap} 值的波动范围也较大。基于图 4 可以发 现, 瞬时幅度特征 σ_{aa} 的差异贡献度很小, 无法用于 区分各类设备的发射信号。由图 5 看出, 瞬时频率特征 σ_{af} 在高信噪比时可用于区分 AM、IDEN、LTE、

GSM、TETRA 设备间的发射信号,但随着信噪比的恶化,该特征的贡献度也在降低,当信噪比小于4dB时,无法区分AM、GSM_A、GSM_B、GSM_C设备的发射信号。



图 3 零中心归一化瞬时相位绝对值标准差 Figure 3 Normalization instantaneous phase







图 5 零中心归一化瞬时频率绝对值标准差 Figure 5 Normalization instantaneous frequency

6.1.2 信号统计量特征仿真分析

在不同信噪比下采集信号时域样本,并计算累积量特征 $\{c_{2,0}, c_{4,0}, c_{4,2}, c_{6,0}, c_{6,1}, c_{6,3}\}$ 进行对比分析。

从图 6 可知, 特征 $c_{2,0}$ 在高信噪比下能有效区分 AM、LTE 和 TETRA 设备的发射信号, 但当信噪比 低于 5 dB 时, 区分度下降明显; 当信噪比大于 4 dB 时, 基于 $c_{2,0}$ 特征值可有效区分 GSM_A、GSM_B、 GSM_C 设备间的发射信号。图 7 显示特征 $c_{4,0}$ 能够 有效区分 IDEN、LTE、GSM_B、GSM_C、和 TETRA_B 设备的发射信号,并且上述设备发射信号的 $c_{4,0}$ 特征 随噪声强度增加没有明显变化。图 8 显示了特征 $c_{4,2}$ 对 TETRA、GSM、IDEN 设备发射信号有明显的区 分度,但对 GSM_A、GSM_B、GSM_C、AM 及 LTE 设备发射信号则无法区分。





图 7 累积量 c_{4,0} Figure 7 Cumulants c_{4.0}





由图 9 可以看出, TETRA 设备发射信号与其他 设备发射信号间的 *c*_{6,0}特征值差异明显, AM、GSM、 IDEN 及 LTE 设备发射信号间的 *c*_{6,0}特征值差异则相 对较小, GSM_A、GSM_B、 GSM_C 三类设备间发 射信号的 *c*_{6,0}特征已无法支撑区分。通过图 10 可知, 特征 *c*_{6,1}无法区分 GSM_A 设备发射信号与 IDEN 设 备发射信号,以及 GSM_C 设备发射信号与 AM 设备 发射信号;此外,随噪声强度变化,各类设备信号的 $c_{6,0}$ 特征值没有明显波动。图 11 显示了 $c_{6,3}$ 特征值无 法对 GSM_A、GSM_B、 GSM_C、AM 及 LTE 设备 发射信号进行区分;此外,随噪声强度变化,各设备 信号的 $c_{6,3}$ 特征值也没有明显变化。

循环累积量提取六类特征: $\{c_{r_{2,0}}, c_{r_{4,0}}, c_{r_{4,2}}, c_{r_{6,0}}, c_{r_{6,1}}, c_{r_{6,3}}\}$ 进行仿真对比。



Figure 12 Cyclic cumulants c_{r2.0}





图 12 显示特征 $c_{r2,0}$ 可明显区分 AM、TETRA 和 LTE 设备间的发射信号,而对 GSM 设备发射信号和 IDEN 设备发射信号无法区分,对三类 GSM 设备及两类 TETRA 设备间的发射信号也不能区分。此外,各类设备 信号的 $c_{r2,0}$ 特征值随噪声增加没有明显变化。由图 13 可知特征 $c_{r4,0}$ 能够明显区分 AM、TETRA 和 LTE 设备 间的发射信号,并且三者的 $c_{r4,0}$ 特征值随噪声没有明显 变化; GSM 设备发射信号和 IDEN 设备发射信号间的 $c_{r4,0}$ 特征值在信噪比大于 4 dB 时能够实现区分; GSM_A、GSM_B、GSM_C 设备发射信号间的 $c_{r4,0}$ 特 征值不具备有效的区分度。此外,从图 14 可以看出,各设 备发射信号的 $c_{r4,2}$ 特征值同 $c_{r4,0}$ 特征值结果基本一致。









由图 15 可以看出 TETRA 设备发射信号与其他 设备发射信号的 c_{r6.0} 特征值差异明显, AM、GSM、 IDEN 及 LTE 设备发射信号间的 c_{r6.0} 特征值差异则 相对较小, GSM_A、GSM_B、GSM_C 设备发射信号 间的 c_{r6.0} 特征值在信噪比大于 9 dB 时能够实现区分; TETRA_A、TETRA_B 设备发射信号之间的 c_{r6.0} 特 征值在信噪比大于 2dB 时能够明显区分。此外,从图 16 和图 17 可以看出,不同设备间信号的 c_{r6,1} 特征值 和 c_{r6,3} 特征值与 c_{r6,0} 有相似的区分效果。

6.1.3 信号小波特征仿真分析

图 18 显示了小波特征 *v*_*wT*_{*a*=5} 的结果,可以看出 AM、TETRA 和 LTE 设备发射信号间的 *v*_*wT*_{*a*=5} 特征



图 17 循环累积量 cr6.3

Figure 17 Cyclic cumulants c_{r6.3}



图 18 小波特征 V_WT_{a=5}







值具有明显的区分度,并且三者的 $v_w T_{a=5}$ 特征值随 噪声没有明显变化;同时 GSM_A、GSM_、GSM_C 设备发射信号间的 $v_w T_{a=5}$ 特征值也能够实现区分; 但 GSM_A 设备发射信号与 IDEN 设备发射信号的 $v_w T_{a=5}$ 特征值接近,在信噪比大于 6 dB 时两者才 能够区分;TETRA A、TETRA B 设备发射信号之间 的 $v_w T_{a=5}$ 特征值在信噪比大于 2 dB 时能够明显区 分。此外,由图 19 可发现 AM、TETRA_A、TETRA_B 和 LTE 设备发射信号间的 $v_w T_{a=9}$ 特征值与 $v_w T_{a=5}$ 特征值具有相似的区分度。

基于 6.1.1、6.1.2、6.1.3 仿真分析, 可得到以下 结论: 首先,指纹特征仿真结果同第三章理论分析结 果一致,能够覆盖无线通信信号间身份差异,实现 无线通信信号的的区分:如瞬时特征在信噪比大于 6 dB 时能够实现 AM、IDEN、LTE、GSM、TETRA 设备发射信号的区分;四阶累积量能够实现 GSM、 TETRA、IDEN、LTE 设备发射信号的区分,同时对 三类 GSM 设备发射信号以及两类 TETRA 设备发射 信号也有较好的区分能力;六阶循环累积能够购实 现 AM、GSM、TETRA、IDEN、LTE 设备发射信号 的区别。同时也可以看到不同指纹特征对信号的区 分度有明显差异,因此需要对各指纹特征的贡献度 进行定量评价,以便进行筛选。

其次,功率衰减造成的信噪比变化会对各指纹 特征贡献度有不同的影响:比如当信噪比在 4~6 dB 区间时,小波特征 v_wT_{a=5}和 v_wT_{a=9} 针对 GSM_A、 GSM_B、GSM_C、IDEN 等设备发射信号间的区分 度下降明显;当信噪比在 2~4 dB 区间时,瞬时相位 特征 σ_{ap} 针对 IDEN 设备发射信号与其他设备发射信 号的区分度下降较快;当信噪比在 0~2 dB 区间时,6 阶循环累积量针对两类TETRA 设备发射信号区分度 明显接近;特征 $c_{2,0}$ 各设备发射信号区分度则随信噪 比下降一直成下降趋势;而特征 $c_{6,3}$ 和 $c_{r2,0}$ 则表现出 了较好抗噪性。可知信号陡降造成的信噪比下降会 降低小波特征和瞬时特征等的区分度。由于指纹集 中高阶统计量能够一定程度消除高斯噪声影响,因此 信号陡降造成的信噪比下降对四阶和六阶统计量特 征区分度的整体影响相对较小。

最后,根据前面的分析,不同特征间存在信息 差异的冗余:比如 $c_{r4,0}$ 、 $c_{r4,2}$ 两者区分效果相似; $c_{r6,1}$ 、 $c_{r6,3}$ 与 $c_{r6,0}$ 有相似的区分效果; σ_{af} 、 $c_{4,0}$ 、 $c_{4,2}$ 和 $v_w T_{a=5}$ 都能用于 IDEN 设备发射信号、TETRA 设备发射信号间的区分等。因此需要在对特征贡献 度定量评价分析的基础上,对指纹集进行匹配,剔 除冗余特征,构建精简的指纹集。

6.2 特征贡献度仿真分析

根据 6.1 仿真的指纹特征数据,同样按照 2000 个样本长度计算特征,并选取 1600 个设备特征样 本。为了对本文特征贡献度评价方法进行合理性分 析,选取基于距离模型的 Fisher 分离度特征评价方 法进行了对比仿真,结果如下:

图 20 显示了 25 个特征的信息增益,可以看出: 特征指纹的信息增益能够表征不同指纹对分类贡献 度的不同:比如由图 3~图 5 可知 σ_{af} 相比 σ_{aa} 和 σ_{ap}



图 20 每个指纹特征的信息增益值 G(Q_{M,N}, P_{i,m}) Figure 20 Information entropy of fingerprint feature G(Q_{M,N}, P_{i,m})



图 21 每个指纹特征的有效信息增益值 C_p

Figure 21 Availability information entropy of fingerprint feature C_p



图 22 每个指纹特征的 Fisher 分离度 D Figure 22 Fisher separation degree of fingerprint feature D

特征区分度有明显的增加,同其信息增益的值相似。 其次,可以看出特征 c_{2,0} 受噪声影响很大,特征 c_{4,2}, c_{6,3}和 c_{r2,0} 受噪声影响较小,同特征仿真结果类似。 图 21 显示了 25 个特征的有效信息增益指数, 其结果表明有效信息增益指数相比信息增益对特征 贡献度进一步筛选:比如特征 *v_wT_{a=5}、v_wT_{a=9}*和 *c*_{2,0}在有效信息增益指数下,其在在指纹集中贡献度 排序有明显下降。

图 22 显示了 25 个特征的 Fisher 可分离度。其 定义为:

$$D = \frac{\sum_{l=i,j} (mean(k_l) - mean(mean(k_l)))^2}{\sum_{l=i,j} \operatorname{var}(k_l)}$$

其中 $mean(k_l)$ 和 $var(k_l)$ 代表第l类的均值和方差。 $mean(mean(k_l))$ 代表了所有类的均值。i和 j代表不同的类。

6.3 特征评价算法效果分析

基于对信息增益和有效信息增益指数的分析, 有效信息增益相比信息增益有更好的特征评价效果, 同时考虑与 Fisher 分离度特征评价方法的对比,本 仿真根据信息增益,有效信息增益和 Fisher 可分离 度对指纹特征的定量评价结果,按照评价指数 $G(Q_{M,N}, P_{i,M})$ 、 C_p 和D的特征排序,利用本文提出 SVM 指纹匹配算法对识别效果进行仿真:



Fisher 分离度 D 识别效果对比 Figure 23 recognition results between $G(Q_{M,N}, P_{i,m})$, C_n and D

由图 23 可知: 基于本文的 SVM 指纹匹配算法, 在有效信息增益指纹评价基础上匹配 9 个特征时识 别效果整体达到全部指纹特征的识别效果。在信息 增益指纹评价基础上匹配 9 个特征时识别效果在 *SNR*<7dB 时整体达到全部指纹特征的识别效果, *SNR*>7dB 时接近全部指纹特征的识别效果。因此,可

以验证指纹特征集含有冗余信息,信息增益和有效 信息增益都能够提取贡献度高的指纹信息、同之前 算法分析一致。其次, 在同样个数的指纹特征下, 按 照有效信息增益选取的特征比按照信息增益选取的 特征能够获得更好的识别效果。验证了本文提出的 有效信息增益相比信息增益能更好评价指纹特征对 无线通信信号身份识别的贡献度大小。同时从图 23 可以得出按照 Fisher 分离度特征评价结果, 匹配 9 个 特征的识别效果在 SNR<7dB 时能够整体达到全部指 纹特征的识别效果, SNR>7dB 时识别效果介于信息 增益和有效信息增益之间。在匹配 6 个特征时效果 同信息增益整体一致。在匹配 3 个特征时效果低于 信息增益算法和有效信息增益算法。因此验证了基 于距离模型的 Fisher 分离度特征评价方法也可以提 取贡献度大的指纹特征,而本文提出的有效信息增 益特征评价法相比 Fisher 分离度能更好评价指纹特 征贡献度。最后,由上述分析可知基于有效信息增益 的指纹特征评价提取指纹特征排序P中的9个特征 整体上达到了全部特征的识别效果,即可以获得精 简的指纹集。

6.4 信号身份识别效果仿真分析

首先, 仿真了不同信噪比下对采集的各种信号 进行识别, 按照每组 2000 个采样点计算特征, 共采 集 1600 组数据进行识别, 结果如下





根据图 24 显示的仿真结果:第一,根据本文构 建无线通信信号特征指纹集可以对不同无线通信信 号身份进行分类识别。第二,针对本文所选取的不同 设备,在接收信噪比大于4 dB时识别效果达到93% 以上,当接收信噪比大于10 dB时识别效果达到99% 以上,验证了本文方法在对无线通信信号身份识别 上的有效性。 其次,计算指纹特征的样本数据长度对特征影响比较大。本文计算了不同数据长度对识别效果的 影响,根据采集数据按照每组 500,800,1200,1600 和 2000 的样本长度对识别效果进行了仿真:





图 25 的仿真结果显示,计算指纹特征的样本数 据长度越小识别效果越差,随着噪声的增加差别会 增加。同时随着计算指纹特征样本数据长度的增长, 识别效果越来越靠近,因此实际应用中可以选择合 适的样本长度计算特征。

7 结论

本文构建了表征无线通信信号差异信息的特征 指纹集,建立了对无线通信信号的唯一标识;同时 针对指纹集中指纹特征贡献度的差异以及可能存在 冗余信息,基于信息增益的特征评价方法对指纹特 征进行定量评价;并基于 SVM 分类方法设计指纹集 匹配方法,实现对特定空间无线通信信号的身份标 识和识别。仿真和实验结果显示本文的方法能够有 效识别广播电台、集群对讲机、2G 移动终端以及 4G 移动终端等设备间的发射信号,对通信场景具有普 遍的适应性。未来将基于 Lasso 回归等对特征评价机 制进一步优化,并建立针对 5G 移动通信系统的指纹 特征。

参考文献

- Sun P F, Zheng Z W, Li M. Recognition of digital modulation signals based on statistical parameters[C]. *International Conference on Transportation, Mechanical and Electrical Engineering*, 2012: 743-753.
- [2] Song J, Liu Y, Wang X D. The Recognition and Parameter Estimation of Hybrid Modulation Signal Combined with FSK and

BPSK[J]. Journal of Electronics & Information Technology, 2013, 35(12): 2868-2873.

(宋军, 刘渝, 王旭东. FSK/BPSK 复合调制信号识别与参数估 计[J]. *电子与信息学报*, 2013, 35(12): 2868-2873.)

- [3] Abdelbar M, Tranter W H, Bose T. Cooperative Cumulants-Based Modulation Classification in Distributed Networks[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2018, 4(3): 446-461.
- [4] Chang D C, Shih P K. Cumulants-based Modulation Classification Technique in Multipath Fading Channels[J]. *IET Communications*, 2015, 9(6): 828-835.
- [5] Tan X H, Chu G X, Zhang X J, et al. Modulation recognition algorithm based on higher-order cumulants and wavelet transform[J]. *Systems Engineering and Electronics*, 2018, 40(1): 171-177.
 (谭晓衡, 褚国星, 张雪静, 等. 基于高阶累积量和小波变换的 调制识别算法[J]. *系统工程与电子技术*, 2018, 40(1): 171-177.)
- [6] Zhao X W, Guo C X, Li J C. Mixed Recognition Algorithm for Signal Modulation Schemes by High-order Cumulants and Cyclic Spectrum[J]. *Journal of Electronics & Information Technology*, 2016, 38(3): 674-680.
 (赵雄文,郭春霞,李景春. 基于高阶累积量和循环谱的信号调

制方式混合识别算法[J]. *电子与信息学报*, 2016, 38(3): 674-680.)

- [7] Aslam M W, Zhu Z C, Nandi A K. Autmomatic Modulation Classification Using Combination of Genetic Programming and KNN[J]. *IEEE Transactions on Wireless Communications*, 2012,11(8): 2742-2750.
- [8] Gorcin A, Arslan H. Identification of OFDM signals under multipath fading channels[C]. *IEEE Military Communications Conference*, 2012: 1-7.
- [9] Kharbech S, Dayoub I, Zwingelstein-Colin M, et al. Blind digital modulation identification for time-selective MIMO channels[J]. *IEEE Wireless Commun. Lett.*, 2014, 3(4): 373-376.
- [10] Dobre O A, Bar-Ness Y, Su W. Robust QAM modulation classification algorithm using cyclic cumulants[C]. *IEEE Wireless Communication and Network Conference*, 2004: 745-748.
- [11] Spooner C M. On the Utility of Sixth-order Cyclic Cumulants for RF Signal Classification[C]. Conference Record of the Asilomar Conference on Signals, Systems and Computers, 2001: 890-897.
- [12] Xie L J, Wan Q. Cyclic Feature-Based Modulation Recognition Using Compressive Sensing[J]. *IEEE Wireless Communications Letters*, 2017, 6(3): 402-405.
- [13] Dobre O A, Bar-Ness Y, Su W. Highe order cyclic cumulants for high order modulation classification[C]. *IEEE Military Communications Conference*, 2003: 112-117.
- [14] Spooner C M, Mody A N, Chuang J, et al. Modulation recognition using second-and higher-order cyclostationarity[C]. *IEEE Interna-*

tional Symposium on Dynamic Spectrum Access Networks, 2017: 1–3.

- [15] Hassanpour S, Pezeshk A M, Behnia F. A robust algorithm based on wavelet transform for recognition of binary digital modulations[C]. *International Conference on Telecommunications and Signal Processing*, 2015:. 508-512.
- [16] Yu Z, Shi Y Q, Su W. M-ary frequency shift keying signal classification based on discrete Fourier transform[C]. *IEEE Military Communications Conference*, 2003: 1167-1172.
- [17] Wood S L, Treichler J R. Performance of the Radon transform method for constellation identification[C]. ASILOMAR Conference on Signal, System, and Computers, 1988: 119-123.
- [18] Yuan Y J, Wu H, Wang X, et al. Specific Emitter Identification Based on Hilbert–Huang Transform-based Time–Frequency–Energy Distribution Features[J]. *IET Communications*, 2014, 8(13): 2404-2412.
- [19] Ureten O, Serinken N. Bayesian Detection of Wi-Fi Transmitter RF Fingerprints[J]. *Electronics Letters*, 2005, 41(6): 373.
- [20] Zhao C D, Huang L F, Hu L T, et al. Transient fingerprint feature extraction for WLAN cards based on polynomial fitting[C]. *The Sixth International Conference on Computer Science & Education*, 2011: 1099-1102.
- [21] Zhang X D, Shi Y, Bao Z. A New Feature Vector Using Selected Bispectra for Signal Classification with Application in Radar Target Recognition[J]. *IEEE Transactions on Signal Processing*, 2001, 49(9): 1875-1885.
- [22] Lopez-Risueno G, Grajal J, Sanz-Osorio A. Digital Channelized Receiver Based on Time-frequency Analysis for Signal Interception[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2005, 41(3): 879-898.
- [23] Lunden J, Koivunen V. Automatic radar waveform recognition[J]. IEEE Journal Seected. Topics Signal Processing, 2007, 1(1): 124-136.
- [24] Bertoncini C, Rudd K, Nousain B, et al. Wavelet Fingerprinting of



黄伟庆 现任中国科学院信息工程研究所 第四研究室主任、中国计算机学会信息保密 专委会秘书长。研究方向为:无线通信安 全、电磁信号处理、网络安全保密技术、物 联网安全、云计算安全技术等。Email: huangweiqing@iie.ac.cn



魏冬 于 2013 年获得北京邮电大学通信与 信息系统专业博士学位。现任中国科学院信 息工程研究所第四研究室副研究员。研究领 域为无线通信物理层安全、调制识别、信号 处理。Email: weidong@iie.ac.cn

Radio-Frequency Identification (RFID) Tags[J]. *IEEE Transactions* on *Industrial Electronics*, 2012, 59(12): 4843-4850.

- [25] Klein R W, Temple M A, Mendenhall M J. Application of Wavelet-based RF Fingerprinting to Enhance Wireless Network Security[J]. Journal of Communications and Networks, 2009, 11(6): 544-555.
- [26] Aubry A, Bazzoni A, Carotenuto V, et al. Cumulants-based radar specific emitter identification[C]. *IEEE International Workshop Information Forensics Security*, 2011: 1-6.
- [27] Li L, Ji H B, Jiang L. Quadratic Time–Frequency Analysis and Sequential Recognition for Specific Emitter Identification[J]. *IET Signal Processing*, 2011, 5(6): 568-574.
- [28] Gillespie B W, Atlas L E. Optimizing time-frequency kernels for classification[J]. *IEEE Transactions on Signal Processing*, 2011, 49(3): 485-496.
- [29] Ren D F, Zhang T, Han J. Approach of specific communication emitter identification combing ITD and nonlinear analysis[J]. *Journal of signal processing*, 2018, 34(3): 331-339.
 (任东方,张涛,韩洁.结合 ITD 与非线性分析的通信辐射源个体识别方法[J]. *信号处理*, 2018, 34(3): 331-339.)
- [30] Satiju V, Triredi N, Biswal G, et al. Specific emitter identification based on variational mode decomposition and spectral features in single Hop and relay scenarios[J]. *IEEE Transactions on information forensics and security*, 2019, 14(3): 581-591.
- [31] 张贤达 时序信号分析[M]. 北京:清华大学出版社 1996:15.
- [32] Zheng W X, Zhao G Q, Lou M. Blind estimation of OFDM sub-carrier frequencies based on the higher-order cyclic cumulants[J]. *Journal of Electronics and information technology*, 2008, 30(2): 346-349.

(郑文秀, 赵国庆, 罗明. 基于高阶循环累积量和 OFDM 子载波 盲估计[J]. *电子与信息学报*, 2008, 30(2): 346-349.)

[33] Bernhard S, Christopher J C Burges, Alexander J Smola. Advances in Kernel Methods: Support Vector Machines[M]. Cambridge: MIT Press, 1999: 1.



杨召阳 于 2011 年在同济大学物理专业获 得硕士学位。现在中国科学院信息工程研究 所计算机专业攻读博士学位。研究领域为无 线通信物理层安全、调制识别、信号处理。 Email: yzyphy@163.com



张萌 于 2013 年获得北京航空航天大学通 信与信息系统硕士学位,现任中国科学院 信息工程研究所第四研究室高级工程师, 研究方向包括:电磁泄漏发射检测防护技 术、无线通信安全技术。Email: zhangmeng@iie.ac. cn



主文 于 2015 年获得北京理工大学电子科 学与技术专业博士学位。现任中国科学院信 息工程研究所第四研究室高级工程师。研究 领域为电磁知识库构建与特征识别、卫星频 率兼容性分析与干扰机理研究。Email: wwen@iie.ac.cn



叶彬 于 2018 年获得桂林电子科技大学通 信专业硕士学位。现在中国科学院信息工程 研究所计算机专业攻读博士学位。研究领域 为信息隐写、信号处理等。Email: yebin@iie.ac.cn