

基于功率检测和选择性集成的无线MAC协议识别技术

杨俊安^{1,2}, 黄科举^{1,2}

(1. 国防科技大学电子对抗学院, 合肥, 230037; 2. 安徽省电子制约技术重点实验室, 合肥, 230037)

摘要: 通过检测无线网络的功率变化识别介质访问控制(Medium access control, MAC)协议在认知无线电和认知电子战领域具有重要意义。为提高识别能力, 本文在现有特征的基础上, 提出碰撞概率估计特征和Fisher统计量特征。针对识别目标的网络配置与训练样本不同时识别效果较差的问题, 提出利用基于Q学习的选择性集成方法, 从训练的多个基分类器中选择一部分组成分类系统, 提高算法的泛化能力。利用OPNET仿真软件采集4种MAC协议的功率变化数据进行验证。实验结果表明, 所提特征能够提升对不同MAC协议的区分度, 当目标网络与训练网络的配置不同时, 选择性集成方法的识别效果优于单分类器和全部集成方法。

关键词: 认知无线电; 认知电子战; 介质访问控制; 协议识别; 选择性集成

中图分类号: TP393 **文献标志码:** A

Wireless Medium Access Control Protocol Identification Based on Power Detection and Selective Ensemble

YANG Junan^{1,2}, HUANG Keju^{1,2}

(1. Electronic Countermeasure Institute, National University of Defense Technology, Hefei, 230037, China; 2. Key Laboratory of Electronic Restricting Technique of Anhui Province, Hefei, 230037, China)

Abstract: Wireless medium access control (MAC) protocol identification based on power detection plays an important role in cognitive radio and cognitive electronic warfare. To improve accuracy, features like collision proportion estimation and Fisher statistic are derived. Aiming at the problem of low accuracy when the target network is different from training samples, the selective ensemble method based on Q-learning is utilized. The identification system is constructed from a set of selected base classifiers to improve generation capacity. The proposed method is tested with samples of four MAC protocols by OPNET simulation. Experimental results show that the features proposed in this article increase discrimination ability of different MACs. Moreover, when the target network is different from training samples, the selective ensemble method has higher accuracy than the single classifier and the ensemble of all classifiers.

Key words: cognitive radio; cognitive electronic warfare; medium access control; protocol identification; selective ensemble

引 言

在认知无线电和认知电子战领域,识别目标网络的介质访问控制(Medium access control, MAC)协议具有重要意义。对于认知无线电,利用主用户的 MAC 协议信息,可以预测频谱空穴随时间的变化情况,在充分利用频谱资源的同时降低对主用户的干扰^[1]。例如,当主用户的 MAC 协议分别为载波侦听多路访问(Carrier sense multiple access, CSMA)和时分复用(Time division multiple access, TDMA)时,为了最大化频谱资源利用率,认知无线电用户应采用不同的接入机制^[2-3]。在认知电子战领域,干扰目标的 MAC 协议信息能够用于决定最佳干扰时机^[4],针对不同的 MAC 协议采取特定的干扰方式,有助于提高干扰效能^[5-7]。认知无线电和认知电子战的相关研究均假定已知目标的 MAC 协议和参数,事实上,由于次用户相对于主用户的透明特性、干扰方与被干扰方的非合作特性,系统无法直接获取目标的 MAC 协议信息,需要研究基于目标网络功率变化情况的 MAC 协议识别方法。

认知无线电技术赋予无线电台感知频谱环境以及推理学习的能力^[8],机器学习是实现认知无线电的学习、决策过程的重要途径^[9]。利用机器学习中的分类算法,根据目标的功率变化进行 MAC 协议分类具有一定的可行性。

Hu 等人^[10]通过提取目标无线网的信号功率、信道空闲时间和信道占用时间的统计特征,训练支持向量机(Support vector machine, SVM)分类器,实现对 TDMA、CSMA/CA、ALOHA、S-ALOHA 四种协议的分类。但该方法假定网络中不同节点的功率和数据包长度均相同,在实际应用中具有一定的局限性。Rajab 等人^[11]通过检测目标信道的能量变化,提取活动时间和非活动时间的分布特征,并分别采用 K 最近邻(K nearest neighbor, KNN)分类器和朴素贝叶斯(Naïve Bayes, NB)分类器,识别 802.11b/g/n 三种协议。上述研究均采用单一分类器,并假定训练样本的网络配置与待识别目标相同。但在实际中,由于网络配置环境的多样性,目标的网络配置一般与训练样本不同,如网络通信负载、节点数及节点发射功率等。单分类器方法容易过拟合,对不同于训练样本配置的目标识别效果较差。

针对 MAC 协议识别的特征提取问题,本文提出根据功率检测序列提取碰撞概率估计特征和 Fisher 统计量特征。针对待识别目标与样本配置不同时识别效果较差的问题,本文提出一种将选择性集成应用于 MAC 协议识别的方法,以提升分类系统的泛化能力。利用 OPNET 软件仿真得到的 TDMA、CSMA/CA、ALOHA、S-ALOHA 协议数据进行协议识别实验。结果表明,本文所提特征对不同 MAC 协议具有较好的区分度,并且当目标与训练样本的网络配置不同时,选择性集成方法优于单分类器和全部集成方法。

1 系统模型

为简化问题,本文以单频点的无线 MAC 协议作为研究对象。单频点无线网络的 MAC 协议可分为无竞争和有竞争两类,其中有竞争 MAC 协议包括基于随机访问和基于预约机制^[12]。采用无竞争 MAC 协议的各节点按照事先确定的规则访问频谱资源,该规则保证各节点通信时不会发生碰撞,如 TDMA 协议。基于预约机制的 MAC 协议中节点在利用控制帧预约频谱资源后才能发送数据,如 CSMA/CA 协议。基于随机访问的 MAC 协议没有碰撞避免机制,节点在需要发送数据时立刻发送,如 ALOHA 协议和 S-ALOHA 协议。因此本文对 TDMA、CSMA/CA、ALOHA 和 S-ALOHA 四种 MAC 协议进行识别。

在进行 MAC 协议识别时,首先通过侦察获取无线网络的通信频率,采集该频点上信号的功率变化信息;其次根据功率变化信息提取用于 MAC 协议分类的特征;最后基于样本训练基分类器,并采用强化学习方法对基分类器进行选择集成,得到分类系统。在利用该分类系统进行 MAC 协议识别时,采用投票法根据基分类器的结果得到系统的分类结果。

1.1 数据采集

假设通过一个接收节点采集目标无线网络传输信道上的功率信息,无线信道环境为加性高斯白噪声环境,如图 1 所示,则接收机在时刻 t 接收到的信号功率为

$$p(t) = \sum_{i=1}^{N_T} p_i(t) + n(t) \quad (1)$$

式中: N_T 表示目标无线网络中的节点个数; $p_i(t)$ 表示接收机接收到网络中第 i 个节点的信号功率; $n(t)$ 表示背景噪声。设接收节点的采样时间间隔为 τ , 则目标网络的功率变化由序列 $p_m = p(t = m\tau)$, $m = 1, 2, \dots, M$ 表示。

图 2 为采样得到的功率序列 $\{p_m\}$ 示意图, 由于背景噪声, 序列 $\{p_m\}$ 存在一定的波动, 但仍能根据功率变化情况估计目标网络中节点的通信行为。当功率变化大于一定阈值 e 时表示网络中存在节点通信行为的变化, 例如图中 $m=6, 15, 25, 36$ 时 $p_m - p_{m-1} > e$, 表示有新的节点开始发送数据。

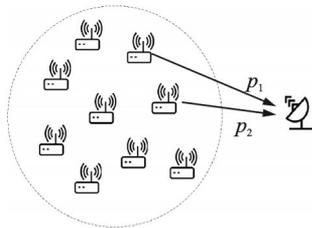


图1 信号接收模型

Fig.1 Signal receiving model

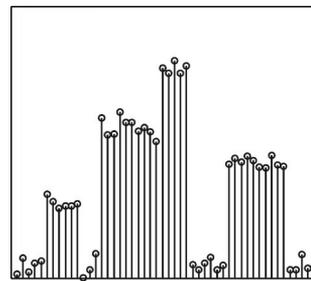


图2 功率序列

Fig.2 Power sequence

1.2 特征提取

为了区分 4 种 MAC 协议, 根据功率信息提取的特征, 需要体现两方面特性: 碰撞概率特性和发送时间周期性。如图 3 所示, TDMA 和 CSMA/CA 协议分别采用时隙分配方式和 RTS/CTS 预约机制以保证传输过程中不发生碰撞, ALOHA 和 S-ALOHA 协议采用随机接入的方式, 传输过程中容易发生碰撞; TDMA 和 S-ALOHA 协议采用分时机制, 其数据包发送时间存在一定的周期性, CSMA/CA 和 ALOHA 协议的数据包发送时间较为随机。根据碰撞概率和发送时间周期特性的不同, 可以有效区分 4 种协议。

1.2.1 碰撞概率特征

当目标网络中不同节点同时发送信号时出现数据包碰撞, 此时接收节点检测到的功率为多个节点所发送信号的功率之和。文献[10]假定所有节点的发送功率相同, 当碰撞概率较大时, 接收功率的变化范围也较大, 因此采用接收功率的均值和方差作为碰撞概率特征。为了更符合实际情况, 本文假定不同节点的发送功率不同, 并在仿真时为不同节点设置不同的信号发送功率。图 4 为仿真得到的目标网络分别为不同的 MAC 协议时, 对接收节点得到的功率序列 $\{p_m\}$ 归一化后的累积概率分布函数(Cumulative distribution function, CDF)。由于 TDMA 和 CSMA/CA 协议中不会发生数据包碰撞, 其 CDF 呈阶梯状, 在每个阶梯处的功率值对应目标网络中某些节点的信号功率; 由于 ALOHA 协议和 S-ALOHA 协

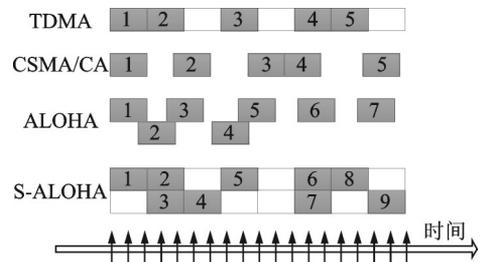


图3 不同MAC协议数据包传输情况示意

Fig.3 Packet transmission in different MAC protocols

议中数据包碰撞概率较大,接收的功率可能为多个信号功率之和,其可能的取值更多,CDF 曲线较为平滑,且相对集中分布在较小的功率值处。因此功率的统计特性能够一定程度反映目标网络中数据包碰撞的情况,本文采用序列 $\{p_m\}$ 的平均值 E_p , 方差 V_p , 偏度 S_p , 峰度 K_p 作为特征。

除功率统计特性外,本文提出根据序列 $\{p_m\}$ 直接估计目标网络中发生碰撞的数据包比例 R_C 。记 $R_C = N_C/N$, 其中 N 为 $\{p_m\}$ 中的数据包个数估计, N_C 为 $\{p_m\}$ 中发生碰撞的数据包个数估计。按照如下方式统计 N 和 N_C : $p_m - p_{m-1} > e$ 时记为一次数据包传输,如图 2 中 $m=6, 15, 25, 36$ 时满足条件,此时 $N=4$; $p_m - p_{m-1} > e$ 且 $p_{m-1} > e$ 时记为一次数据包碰撞,如图 2 中 $m=25$ 时满足条件,此时 $N_C=1$,得到图 2 所示 $\{p_m\}$ 的 $R_C = N_C/N = 0.25$ 。

1.2.2 发送时间周期性特征

对于采用分时机制的 MAC 协议,虽然节点发起通信的行为是随机的,但节点进行通信的起始时刻是一个时隙的起始时刻,因此发送时间周期性可以作为区分 MAC 协议的特征之一。文献[10]假定所有数据包的长度相同,则采用分时机制协议的空闲时间和通信时间均为数据包传输时间的整数倍,统计空闲时间和通信时间的最小值、最大值和中位数作为特征。上述特征不适用于数据包长度变化的情况,且容易受到通信负载的影响。因此本文假定数据包长度随机变化,通过统计发送起始时刻的自相关函数最大值 a 和 Fisher 统计量 g ,用于估计通信节点发送行为的周期性。

首先利用序列 $\{p_m\}$ 得到目标网络的发送起始时刻标记序列 $\{f_m\}$, 即当 m 为发送起始时刻时 $f_m = 1$, 否则令 $f_m = 0$ 。图 2 所示的序列 $\{p_m\}$ 对应的序列 $\{f_m\}$ 如图 5 所示。序列 $\{f_m\}$ 的自相关函数能够体现其周期性,按照如下方式计算自相关函数的最大值 a 为

$$a = \max_{k=1}^{m-1} \left(\frac{c_k}{c_0} \right) \quad (2)$$

$$c_k = \frac{1}{M} \sum_{m=1}^{M-k} (f_m - \bar{f})(f_{m+k} - \bar{f}) \quad (3)$$

式中: \bar{f} 为序列 $\{f_m\}$ 的平均值; c_k 表示序列 $\{f_m\}$ 延时为 k 的自相关系数。

Fisher 周期性检验中的统计量 $g^{[13]}$ 可以体现时间序列的周期性, g 越大表明序列的周期性越强,其计算方式如下

$$g = \frac{\max_{k=1}^{M/2} F_k}{\sum_{k=1}^{M/2} F_k} \quad (4)$$

$$F_k = \left| \sum_{m=1}^M f_m \cdot e^{-i2\pi \frac{km}{M}} \right|^2 \quad (5)$$

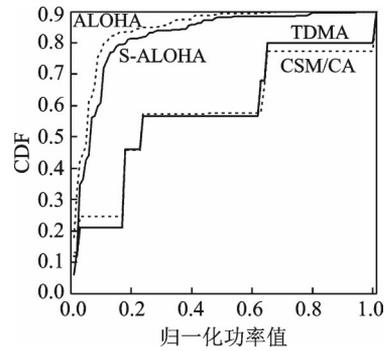


图 4 归一化后功率的累积概率分布函数及概率密度函数

Fig.4 Cumulative probability and probability density of normalized power values

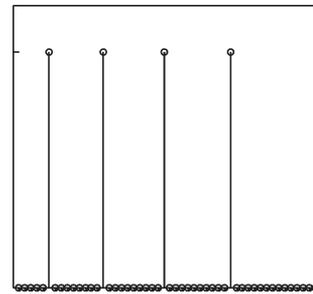


图 5 发送起始时刻标记序列

Fig.5 Identity sequence of transmitting start time

式中 $\{F_k\}$ 为序列 $\{f_m\}$ 的周期图。

从序列 $\{p_m\}$ 提取 $E_p, V_p, S_p, K_p, R_c, a, g$ 共7维特征用于MAC协议分类,其中 E_p, V_p, S_p, K_p 以及 R_c 体现协议的碰撞概率特性, a 和 g 体现协议的发送行为周期性。

1.3 基于强化学习的选择性集成算法

泛化能力是指分类算法对新鲜样本的适应能力。泛化能力强的算法能够学习到隐含在数据背后的规律,对具有相同规律的训练集以外的数据,算法也能保持较好的分类效果^[14]。因此,提升分类算法的泛化能力,有助于提升对不同于训练集配置网络的协议识别效果。

集成学习针对同一问题训练多个基分类器,并将基分类器集成为一个分类系统,该系统的分类准确率优于单个分类器,且具有较强的泛化能力。相比于将所有基分类器全部集成到系统中,从中选取较优的部分基分类器构建的集成系统可能提高分类效果,并且所需的预测时间更短^[15]。选择部分基分类器进行集成的方法称为选择性集成(Selective ensemble, SE),或者称为集成剪枝^[16]。

在构建集成系统时,首先利用训练集训练多个基分类器,其次利用验证集,按照特定算法选择最优的基分类器子集组成集成系统,如图6所示。基分类器的选择是选择性集成的关键。理论上可以通过枚举法,枚举所有可能的基分类器集合,并衡量每个基分类器集合的预测效果选择最优的集合,但这一方法的计算复杂度过高,实际中并不可行。采用贪婪方法直接选择分类效果最优的数个基分类器,得到的基分类器集合不一定为最优。

强化学习作为一种基于试错的行为学习方法^[17],能够用于基分类器的选择。文献[18]采用强化学习方法,将基分类器的选择问题建模为马尔可夫决策过程,以分类准确率作为优化目标,在决策过程中对所有可能的基分类器集合的分类效果进行估计,并选择最优的集合,最终收敛到最优的基分类器集合。该方法比枚举法更加高效,并且能够一定程度上避免贪婪算法容易找到局部最优解的问题。

记基分类器集合为 $T=\{C_1, C_2, C_3, \dots, C_T\}$,强化学习中的状态 s_i 由已选择的基分类器和当前考虑的基分类器组成,行为 a_i 包括是否集成当前考虑的基分类器,如图7所示。在每次状态转移时决定是否包含特定基分类器到集成系统中,在每个状态 s_i ,以 s_i 的基分类器集合对验证集的分类效果作为该次决策的回报值。采用强化学习中的Q-学习^[19]算法,将所选基分类器的分类效果作为回报值优化基分类器选择策略,最终收敛到最优的基分类器选择策略,其算法流程如图8所示,其中参数 ϵ, α, γ 分别表示随机选择行为的概率、学习速率和未来回报值的折扣比例,事先设置为固定值。

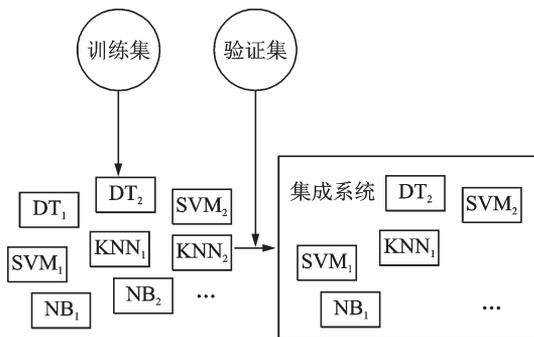


图6 选择性集成系统构建过程

Fig.6 Construction of selective ensemble system

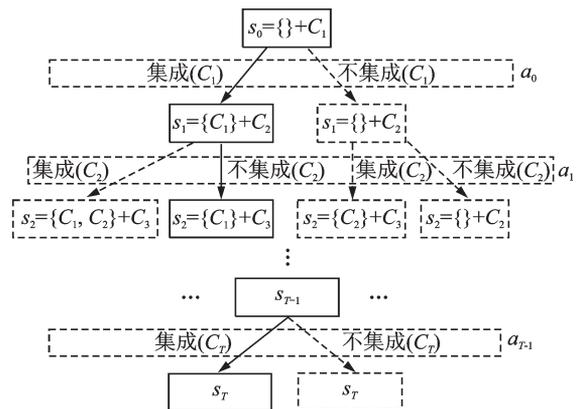


图7 选择性集成的马尔可夫决策过程

Fig.7 Markov decision process of selective ensemble

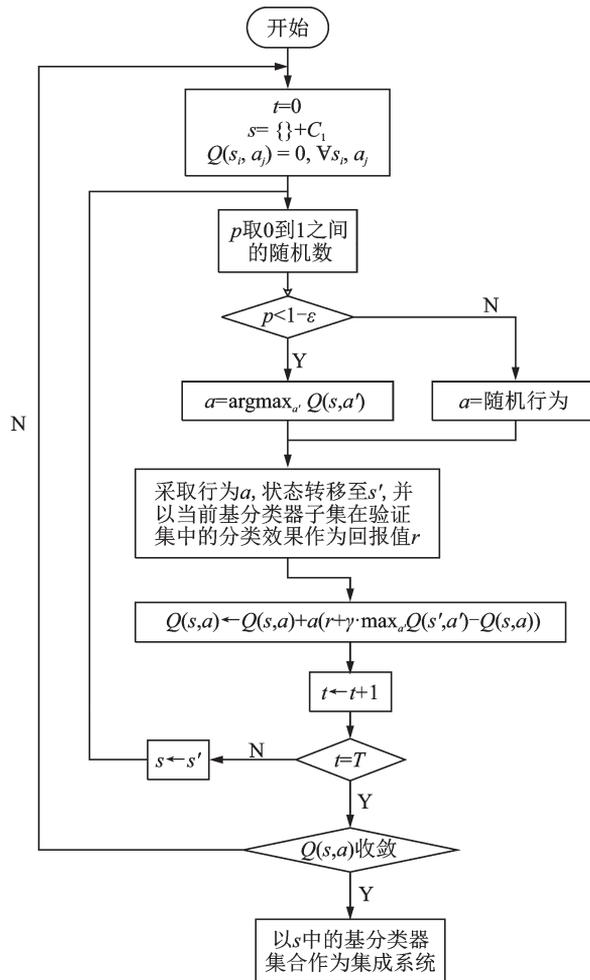


图8 基于 Q-学习的选择性集成算法流程

Fig.8 Process of Q-learning based selective ensemble algorithm

2 实验结果和分析

2.1 数据集及算法参数

本文通过 OPNET 软件进行仿真,其无线网络环境配置如图 9 所示,目标无线网络包含 9 个节点,通过一个接收节点获取目标网络通信信道上的功率变化情况,设置接收节点的采样速率为 10^4 次/s,每次采样 10 s,即 10^5 个采样点作为一个接收功率序列 $\{p_m\}$,并根据该序列提取特征得到一个样本。本文实验分别仿真 TDMA、CSMA、ALOHA、S-ALOHA 四种 MAC 协议。为了体现无线网络配置场景的多样性,在仿真时随机设置通信节点的发射功率和数据包长度,并设置通信负载 G 为 0.1, 1 或 0.1~10 之间的随机数,分别得到 500, 500, 1 500 个样本,总样本数为 10 000 个,样本数据分布情况如表 1 所示。

本文采用 DT、SVM、KNN、NB 四种基分类器,基分类器由基于 Python 语言的 Scikit-Learn 库实现,其中 DT、KNN、NB 分类器均采用默认设置,SVM 分类器设置核函数为径向基函数,惩罚因子 $C=500$,核参数 $\gamma=20$ 。用于选择性集成的 Q-学习算法相关参数设置为 $\epsilon=0.1, \alpha=0.1, \gamma=0.9$ 。

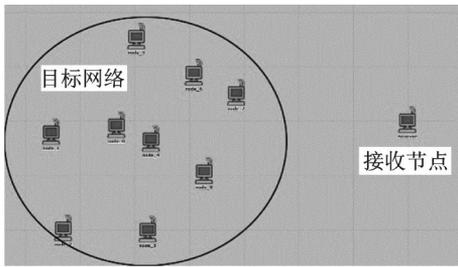


图9 OPNET软件仿真场景

Fig.9 Simulation scenario in OPNET

表1 样本数据分布情况

Table 1 Distribution of sample data

协议	$G=0.1$	$G=1.0$	$G=\text{random}$ (0.1, 10.0)	合计
TDMA	500	500	1 500	2 500
CSMA/CA	500	500	1 500	2 500
ALOHA	500	500	1 500	2 500
S-ALOHA	500	500	1 500	2 500

2.2 评价指标

分类系统对待分类样本进行分类后,对每种协议,统计如下3个统计量:TP(True positive)表示被分类系统识别为该协议且确实属于该协议的样本数;FP(False positive)表示被分类系统识别为该协议但实际不属于该协议的样本数;FN(False negative)表示被分类系统识别为不属于该协议但实际属于该协议的样本数。分类系统的准确率 P 和召回率 R 定义如下

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

F 值是准确率和召回率的调和平均值,其定义如下

$$F = 2 \times \frac{P \times R}{P + R}$$

由于 F 值综合考虑分类系统的准确率和召回率,本文采用 F 值作为分类效果的评价指标。

2.3 实验结果

2.3.1 特征提取方法比较

采用选择性集成方法作为分类算法,比较不同特征对MAC协议识别效果的影响。在10 000个样本中,从4种MAC协议中各随机选取500个样本,其中1 000个样本作为训练样本集,另外1 000个样本作为验证样本集,其余8 000个样本作为测试样本集,分别采用文献[10]提取的特征和本文所提特征进行测试,得到对不同协议的分类 F 值。将上述实验重复10次后取平均得到的结果如表2所示。实验结果表明采用本文所提特征能够取得更好的分类效果,尤其对ALOHA协议和S-ALOHA协议的分类效果有明显提升。

表2 不同特征提取方法的分类效果

Table 2 Classification result of different feature extraction methods

提取方法	TDMA	CSMA/CA	ALOHA	S-ALOHA	平均
文献[10]所提特征	0.991 5	0.974 9	0.654 3	0.631 9	0.813 2
本文所提特征	0.991 7	0.979 6	0.839 8	0.794 7	0.901 5

图10所示为相同样本分别采用文献[10]和本文特征时的分布情况,图10(a)采用文献[10]中的功率方差和通信时间中位数特征,图10(b)采用本文中的数据包碰撞比例 R_c 和Fisher周期性检验统计量 g 。由于本文假定数据包发送时间长度为随机值,通信时间中位数无法准确反映网络中通信行为的周期性,因此图10(a)中的ALOHA协议和S-ALOHA协议样本点不易区分。而图10(b)中除 R_c 接近0时ALOHA协议和S-ALOHA协议不易区分外,4种MAC协议在特征 R_c 和 g 上具有一定的可区分性。

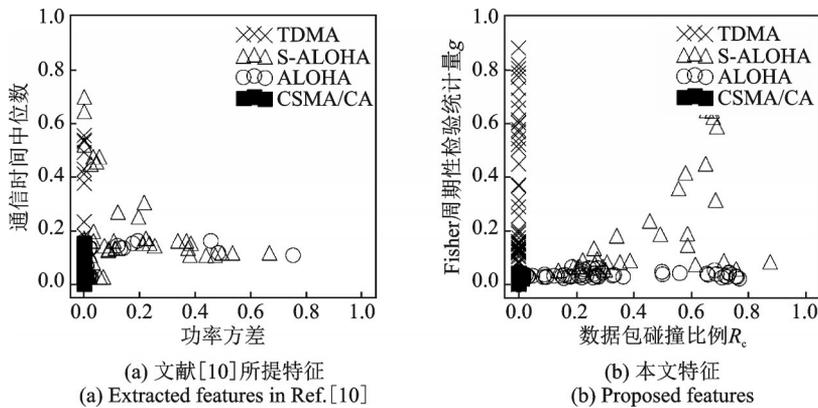


图 10 相同样本提取的文献[10]特征和本文特征分布情况
 Fig.10 Distribution of features of same samples in Ref.[10] and this paper

2.3.2 分类算法比较

在单分类器和全部集成方法的训练过程中,仅需要训练集样本用于分类器训练。在选择性集成方法的学习过程中,除需要训练集样本用于基分类器训练外,还需要验证集样本用于基分类器的选择。因此在将这些方法进行对比时,为保证其在训练时所用样本数相等,需要保证单分类器或全部集成方法的训练集样本数等于选择性集成的训练集与验证集样本数之和。

为验证训练样本对识别效果的影响,在实验 1 中随机选取训练样本,在实验 2 中从 $G=1.0$ 的数据中选取训练样本,在实验 3 中从 $G=0.1$ 的数据中选取训练样本。在实验 1 中,从 4 种 MAC 协议中各随机选取 500 个样本组成集合 Y ,作为单分类器和全部集成方法的训练集。从 Y 中随机选取 1 000 个样本作为选择性集成方法的训练集,剩余 1 000 个样本作为其验证集。将所有样本中除 Y 以外的样本作为测试集,得到不同方法对不同 MAC 协议的分类 F 值。将上述实验重复 100 次后取平均得到结果如表 3 所示。在实验 2 中,从 4 种 MAC 协议中各选取 $G=0.1$ 的 500 个样本组成集合 Y ,其他步骤与实验 1 相同,得到结果如表 4 所示。在实验 3 中,从 4 种 MAC 协议中各选取 $G=1.0$ 的 500 个样本组成集合 Y ,其他步骤与实验 1 相同,得到结果如表 5 所示。

在实验 1 中,单个 DCT 分类器的平均 F 值最大,对每种 MAC 协议的分类效果最优的方法均为单分类器。由于实验 1 中的训练样本包含了测试样本中所有的通信负载情况,因此单个分类器便能够取得较高的 F 值。在此实验条件下,选择性集成方法对泛化能力提升的效果不明显,同时其训练样本数少于单分类器和全部集成方法,因此选择性集成方法的识别效果不如单分类器和全部集成方法。在训练样本较为充分时,可能出现选择性集成方法不如最优的单分类器的情况^[20]。

在实验 2 和实验 3 中,对大部分 MAC 协议,选择性集成方法的分类 F 值最大,且选择性集成方法的平

表 3 实验 1 中不同分类方法的 F 值

Table 3 F value of different classification methods in Experiment 1

提取方法	DCT	SVM	KNN	NB	全部集成	选择性集成
TDMA	0.992 0	0.992 8	0.996 8	0.987 1	0.993 0	0.991 7
CSMA/CA	0.983 5	0.981 4	0.966 4	0.917 9	0.925 5	0.979 6
ALOHA	0.849 9	0.808 6	0.717 5	0.740 7	0.811 5	0.839 8
S-ALOHA	0.802 4	0.812 7	0.720 2	0.675 6	0.779 5	0.794 7
平均	0.907 0	0.898 9	0.850 2	0.830 3	0.877 4	0.901 5

表4 实验2中不同分类方法的F值

Table 4 F value of different classification methods in Experiment 2

提取方法	DCT	SVM	KNN	NB	全部集成	选择性集成
TDMA	0.734 4	0.668 4	0.975 0	0.795 4	0.935 3	0.974 5
CSMA/CA	0.798 1	0.463 2	0.460 5	0.500 8	0.660 2	0.930 9
ALOHA	0.584 7	0.407 6	0.579 0	0.565 7	0.599 3	0.629 4
S-ALOHA	0.474 3	0.487 9	0.299 4	0.330 2	0.458 9	0.600 4
平均	0.647 9	0.506 8	0.578 5	0.548 0	0.663 4	0.783 8

表5 实验3中不同分类方法的F值

Table 5 F value of different classification methods in Experiment 3

提取方法	DCT	SVM	KNN	NB	全部集成	选择性集成
TDMA	0.789 6	0.707 8	0.960 4	0.860 9	0.935 0	0.939 8
CSMA/CA	0.851 5	0.737 4	0.767 6	0.714 1	0.812 9	0.906 0
ALOHA	0.691 3	0.531 6	0.513 6	0.560 4	0.670 1	0.776 0
S-ALOHA	0.530 4	0.638 8	0.638 8	0.522 9	0.511 4	0.764 6
平均	0.715 7	0.592 9	0.720 1	0.664 6	0.732 4	0.846 6

均F值最大,识别效果提升明显。在实验2和实验3中用于训练的样本与测试样本的通信负载不同,单分类器由于泛化能力较差,对于和训练场景不同的样本识别效果较差,而全部集成方法和选择性集成方法通过提升分类系统的泛化能力,提高对未训练场景样本的识别效果,选择性集成方法的提升尤其明显。

在上述3个实验中,不同分类方法总体上对TDMA协议和CSMA/CA协议的分类F值更大,而对ALOHA协议和S-ALOHA协议的分类F值较小,这与2.2节中提取的特征造成ALOHA协议和S-ALOHA协议不容易区分有关。另外,实验3的识别效果总体上好于实验2,这是由于实验3中所用训练样本的通信负载 $G=1.0$,而实验2中所用训练样本的通信负载 $G=0.1$,通信负载较高时相同时间内的通信行为次数更多,MAC协议的特征更明显,因此训练得到的分类系统的识别效果更好。同时,在3个实验中,KNN分类器对TDMA协议的识别效果最好,但对ALOHA协议和S-ALOHA协议的识别效果较差,这与不同协议特征的分布情况和分类器的特性有关。

算法的识别时间对于实现实时分类具有重要影响。多次实验统计的结果表明,单分类器算法的单次识别时间为0.002 ms左右,全部集成方法的单次识别时间为0.05 ms左右,选择性集成方法的单次识别时间为0.01 ms左右。选择性集成方法的识别时间长于单分类器方法,但比全部集成方法更短,其识别时间满足实时性要求。

3 结束语

通过检测无线网络的功率变化识别其MAC协议对于认知无线电和认知电子战具有重要意义。由于不同无线网络配置不同,特征提取和分类算法的泛化能力对识别效果具有重要影响。本文根据典型MAC协议的特性,提出从碰撞比例特性和发送行为周期性两方面提取特征,所提特征适用于网络节点功率不同和数据包长度不同的情况,其分类效果优于原有特征。本文通过选择性集成学习提升分类系统的泛化能力,实验结果表明,当识别目标与训练样本的网络配置不同时,选择性集成方法的协议识别效果较单分类器和全部集成方法有明显提升。

本文提出的MAC协议特征对ALOHA协议和S-ALOHA协议的区分能力较差,还需进一步改进。由于选择性集成方法在学习过程中除需训练基分类器外,还需对基分类器进行选择,因此训练时间较长是该方法的不足。另外,对多频点MAC协议的特征提取方法和识别算法是今后的研究方向。

参考文献:

- [1] HU S, YAO Y D, YANG Z. MAC protocol identification approach for implement smart cognitive radio[C]// Proceedings of IEEE International Conference on Communications (ICC). Ottawa, Canada: IEEE, 2012: 5608-5612.
- [2] CHEN R R, LIU X. Coexisting with CSMA-based reactive primary users[C]// Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum. Aachen, Germany: IEEE, 2010: 1-7.
- [3] HU S, YAO Y D, YANG Z. Cognitive medium access control protocols for secondary users sharing a common channel with time division multiple access primary users[J]. Wireless Communications and Mobile Computing, 2014, 14(2): 284-296.
- [4] CHEN J, SEN S, CHIANG M, et al. A framework for energy-efficient adaptive jamming of adversarial communications[C]// Proceedings of 47th Annual Conference on Information Sciences and Systems (CISS). Baltimore, USA: IEEE, 2013: 1-6.
- [5] HAMZA T, KADDOUM G, MEDDEB A, et al. A survey on intelligent MAC layer jamming attacks and countermeasures in WSNs[C]// Proceedings of IEEE 84th Vehicular Technology Conference. Montréal, Canada: IEEE, 2016: 1-5.
- [6] LICHTMAN M, POSTON J D, AMURU S D, et al. A communications jamming taxonomy[J]. IEEE Security & Privacy, 2016, 14(1): 47-54.
- [7] TILOCA M, DE GUGLIELMO D, DINI G, et al. JAMMY: A distributed and dynamic solution to selective jamming attack in TDMA WSNs[J]. IEEE Transactions on Dependable and Secure Computing, 2017, 14(4): 392-405.
- [8] HAYKIN S. Cognitive radio: Brain-empowered wireless communications[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(2): 201-220.
- [9] BKASSINY M, LI Y, JAYAWEERA S K. A survey on machine-learning techniques in cognitive radios[J]. IEEE Communications Surveys & Tutorials, 2013, 15(3): 1136-1159.
- [10] HU S, YAO Y, YANG Z. MAC protocol identification using support vector machines for cognitive radio networks[J]. IEEE Wireless Communications, 2014, 21(1): 52-60.
- [11] RAJAB S A, BALID W, AL KALAA M O, et al. Energy detection and machine learning for the identification of wireless MAC technologies[C]// Proceedings of International Wireless Communications and Mobile Computing Conference (IWCMC). Dubrovnik, Croatia: IEEE, 2015: 1440-1446.
- [12] KUMAR S, RAGHAVAN V S, DENG J. Medium access control protocols for ad hoc wireless networks: A survey[J]. Ad Hoc Networks, 2006, 4(3): 326-358.
- [13] QUINN B G, HANNAN E J. The estimation and tracking of frequency[M]. Cambridge: Cambridge University Press, 2001: 14-15.
- [14] ABU-MOSTAFA Y S, MAGDON-ISMAIL M, LIN H T. Learning from data[M]. New York, USA: AML Book, 2012: 39-40.
- [15] ZHOU Zhihua. Ensemble methods: Foundations and algorithms[M]. Boca Raton, USA: CRC Press, 2012: 119-123.
- [16] DAI Qun, YE Rui, LIU Zhuan. Considering diversity and accuracy simultaneously for ensemble pruning[J]. Applied Soft Computing, 2017, 58: 75-91.
- [17] SUTTON R S, BARTO A G. Reinforcement learning: An introduction[M]. Cambridge: MIT Press, 1998: 3-6.
- [18] PARTALAS I, TSOUMAKAS G, VLAHAVAS I. Pruning an ensemble of classifiers via reinforcement learning[J]. Neurocomputing, 2009, 72(7): 1900-1909.
- [19] POURPANAH F, TAN C J, LIM C P, et al. A Q-learning-based multi-agent system for data classification[J]. Applied Soft Computing, 2017, 52: 519-531.
- [20] CAVALCANTI G D C, OLIVEIRA L S, MOURA T J M, et al. Combining diversity measures for ensemble pruning[J]. Pattern Recognition Letters, 2016, 74: 38-45.

作者简介:



杨俊安(1965-),男,教授,博士生导师,研究方向:信号处理、智能计算等。



黄科举(1994-),男,硕士研究生,研究方向:通信认知电子战,E-mail: huangkeju@163.com。