研究与开发

基于区块链与深度学习的空间分集协作频谱感知系统

肖仲杉[1]，王春琦[2]，冯大权[1]
（1. 深圳大学，广东 深圳 518060；2. 国家无线电监测中心，北京 100037）

摘　要：协作频谱感知是认知无线电中的关键技术。针对协作频谱感知中存在的安全性、隐私性、激励性和隐藏终端等问题，提出了一种运行在智能合约上的数据驱动的智能化空间分集的协作频谱感知系统。具体地，利用区块链的去中心化、数据难以篡改等特性，设计了一种激励性的协作频谱感知系统，并采用深度学习的方法来识别系统中的恶意用户。此外，针对如何更为高效地在该系统中招募感知节点以达到较高的感知准确率，设计了基于性能权重和空间分集的硬判决协作频谱感知融合算法。实验结果表明，所提算法在安全性、隐私性、激励性、感知准确率方面优于传统的协作频谱感知算法。

关键词：协作频谱感知；认知无线电；区块链；智能合约；机器学习
中图分类号：TP393
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2023193

A spatial diversity cooperative spectrum sensing system based on blockchain and deep learning

XIAO Zhongshan[1], WANG Chunqi[2], FENG Daquan[1]
1. Shenzhen University, Shenzhen 518060, China
2. The State Radio Monitoring Center, Beijing 100037, China

Abstract: Cooperative spectrum sensing is a key technology in cognitive radio. A data-driven intelligent cooperative spectrum sensing system was proposed with spatial diversity running on a smart contract to address issues of security, privacy, incentive and hidden terminals in cooperative spectrum sensing. Specifically, a motivated spectrum sensing system was designed by taking advantage of the decentralization of blockchain technology and the immutability of data. Secondly, a deep learning-based approach was proposed to identify malicious users in the system. In addition, to achieve higher accuracy in recruiting sensing nodes more efficiently in the system, a hard decision cooperative spectrum sensing fusion algorithm based on performance weights and spatial diversity was designed. The experimental results indicate that the proposed solution outperforms traditional cooperative spectrum sensing algorithms in terms of security, privacy, motivation, and sensing accuracy.

Key words: cooperative spectrum sensing, cognitive radio, blockchain, smart contract, machine learning

收稿日期：2023−08−18；**修回日期：**2023−09−10
基金项目：国家重点研发计划项目（No.2020YFB1807600）
Foundation Item: The National Key Research and Development Program of China (No.2020YFB1807600)

0 引言

随着 5G 无线通信网络中高清视频传输、虚拟现实、增强现实和其他高带宽服务的兴起，通信流量呈爆炸式增长[1-4]，据统计，平均每月每部移动设备的流量消耗将从 2020 年的 5 GB 增长至 2030 年的 250 GB[5]，可用频谱资源变得愈发紧张。当前我国的频谱管理主要采用固定频谱分配策略，无线设备在不同的频率上“各行其道”。根据美国联邦通信委员会调查报告[6]中的统计数据，传统静态频谱分配方案的许可频谱利用率为 15%～85%，这无疑对不断增长的无线通信网络流量需求有很大的限制。因此，迫切需要寻找新的提高无线通信网络频谱利用率的解决方案。

认知无线电（cognitive radio，CR）的提出是缓解频谱资源短缺和未充分利用的可行解决方案，美国联邦通信委员会在 2003 年提出，只要具备自适应频谱感知能力的无线电都可称为 CR，并鼓励在不干扰主用户正常通信的前提下使用 CR。频谱感知是认知无线电的关键技术之一[7-10]。它通过感知无线环境中未被充分利用的频段，让认知无线电用户接入该频段以实现动态频谱共享。从参与感知的节点数量上看，频谱感知可以分为单用户频谱感知和协作频谱感知。由于单用户频谱感知容易受噪声干扰和障碍物阻挡等因素干扰而产生误差，因此协作频谱感知在研究中更加受关注。集中式协作频谱感知将来自多个次用户（secondary user，SU）的单个感测结果组合在一起以确定主用户（primary user，PU）的存在。但是由于信道条件复杂，很难得出最佳的协作频谱感知策略。例如，与远离 PU 的 SU 相比，靠近 PU 的 SU 可能更可靠地检测 PU，而由于无线信道的空间相关性，彼此接近的 SU 很可能报告相似的感知结果，但过多相似且出错的感知结果的上报会占用大量带宽资源和增加频谱感知融合的复杂度，严重时甚至导致频谱感知的准确率低下，此外，相邻频带间发射功率的泄露也会对频谱感知的结果造成影响。集中式协作频谱感知架构引入了融合中心来负责收集各个节点的感知结果，为了有效地共享频谱，PU 和 SU 都需要与融合中心交互私人信息，这种方式容易受单点攻击和隐私泄露的影响[11-15]。而且，可能存在别有用心的用户故意上传错误的信息企图破坏融合中心的判决结果，南京邮电大学[16]、国防科技大学[17]和陆军装甲兵学院[18]的学者针对该类情况展开过详细的研究和讨论。此外，对于感知节点来说，收集感知数据是需要成本的，尤其是对于资源有限的物联网设备来说。其成本通常包括能量、网络资源（如通信、缓存和计算）和时间，除了物理资源消耗，上文提到的安全威胁和隐私泄露都会影响用户的积极性。因此，传统的集中式协作频谱感知可能无法吸引足够的感知节点进行频谱感知。此外，协作频谱感知大多基于中心化的架构，难以在感知节点和感知需求者之间建立信任。因此，需要提出合适的激励机制来吸引足够多的感知节点参与到协作式频谱感知任务中[19]。

为了解决上述问题，本文提出了一种面向安全性、隐私性、激励性的数据驱动智能化空间分集协作频谱感知系统。主要贡献如下。

- 针对在协作频谱感知中的安全性、隐私保护和激励的问题，设计出了一种带有激励性的基于智能合约的协作频谱感知系统，并在 Remix 集成开发环境中对智能合约进行了测试。
- 针对在协作频谱感知系统中可能存在的上传虚假感知数据的现象，采用了深度学习的方法，对恶意用户及其行为进行识别。
- 针对隐藏终端、噪声干扰和硬件感知性能参差不齐带来的如何选择合适的协作频谱感知节点的问题，设计了基于性能权重和空间分集的硬判决协作频谱感知融合算法。

1 系统模型

协作频谱感知系统模型如图 1 所示。该系统

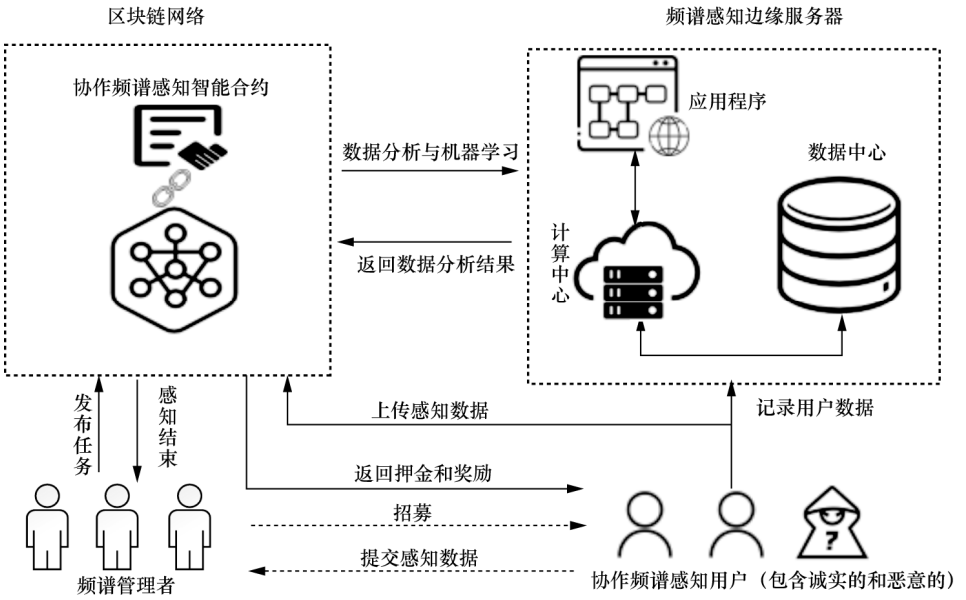


图 1　协作频谱感知系统模型

由 4 个主体组成：频谱管理者、运行在区块链上的智能合约、协作频谱感知用户、频谱感知边缘服务器。频谱管理者和协作频谱感知用户借助区块链网络达到信息交互的目的，各部分的主要职责如下。

频谱管理者：频谱管理者的目标是检测一定地理范围内的频谱环境，即频谱空闲和占用的情况，并通过在智能合约中发布协作频谱感知任务，设置一定的以太币奖励，达到招募用户参与协作频谱感知任务的目的。

协作频谱感知用户：协作频谱感知用户是申请参与协作频谱感知任务的感知节点，其通过在智能合约中查看频谱管理者上传的感知任务后选择是否参与。参与协作频谱感知的用户需要提交一定量的以太币作为押金，并及时上传其感知结果。

频谱感知边缘服务器：频谱感知边缘服务器会收集各个用户的感知数据与每轮频谱感知任务的融合结果，其带有一个能识别恶意用户的深度学习模型和一个基于性能权重和空间分集的硬判决协作频谱感知融合算法。在频谱感知任务开始后，频谱感知边缘服务器通过将各个协作频谱感知用户的数据输入模型中，预测出该用户是否为恶意用户，并将预测结果发送至智能合约；通过一个基于性能权重和空间分集的硬判决协作频谱感知融合算法，得出最佳的频谱感知融合策略，并将结果发送至智能合约，为智能合约最终的感知数据融合与发放以太币奖励提供数据化和智能化的支持。

智能合约：区块链智能合约技术作为支撑，在协作频谱感知系统中建立了信任机制、安全机制、激励机制。具体实现形式如下。

- 智能合约在频谱管理者与协作频谱感知用户之间建立了信任机制，在不需要第三方参与的情况下，系统会自动执行合约所规定的内容。
- 智能合约对协作频谱感知用户提交的数据进行融合处理，得出最终的感知结果。通过分布式算法和密码学技术，确保了区块链网络中的数据一致性和难以篡改性，避免了单点攻击导致数据丢失或被恶意篡改的风险，由此建立了安全机制。
- 智能合约是一个可编程的、按照代码逻辑

自动执行的合约，在每轮感知任务结束后，智能合约会将该轮感知结果作为依据，与用户上传的感知数据进行对比。基于提前预训练好的用于识别恶意用户及其行为的深度学习模型，将该用户的历史数据和当前感知轮次的数据作为神经网络的输入特征，对恶意用户进行识别。如果判定的结果是诚实的用户，系统将无须人工操作，自动退还押金并发放以太币奖励。否则，系统将没收该用户数的押金。随着参与协作频谱感知任务的用户数不断增加，系统能得到更多用于判定恶意用户的有利数据，使得恶意用户或感知性能较差的用户及其行为更容易被系统识别，实现良性循环。由此建立了激励机制。

2 协作频谱感知流程

协作频谱感知流程如图 2 所示，具体如下。

步骤 1 频谱管理者在智能合约中发布频谱感知任务，设置押金的金额，并存入一定数额的以太币用于任务结束的奖励发放。

步骤 2 协作频谱感知用户通过调用智能合约中的相关函数，获取到可参加的任务列表。

步骤 3 参加协作频谱感知任务的用户在本地进行频谱感知，并将本地的感知结果分别提交至区块链网络和频谱感知边缘服务器，实现本地数据与智能合约的交互，确保数据难以被篡改。

步骤 4 频谱感知边缘服务器在收集到所有的协作频谱感知用户提交的感知数据后，结合用户的历史数据，采用基于性能权重和空间分集的硬判决协作频谱感知融合算法进行节点招募分析，并将分析结果上传到区块链网络，为智能合约对数据融合提供指导。

步骤 5 协作频谱感知任务招募结束后，智能合约对收集到的数据进行融合处理，得到最终的感知结果，并发送至频谱管理者和频谱感知边缘服务器。

步骤 6 频谱感知边缘服务器收集该轮的感知数据，并进行训练，更新模型的权重，值得注意的是，该步骤只是在之前训练好的模型基础上进行训练，不需要从无到有重新训练，以节省大量的计算资源和时间，且并非每一轮感知任务结束后都执行该步骤，在具备较多数据集之后，可以每隔几轮执行一次该步骤。

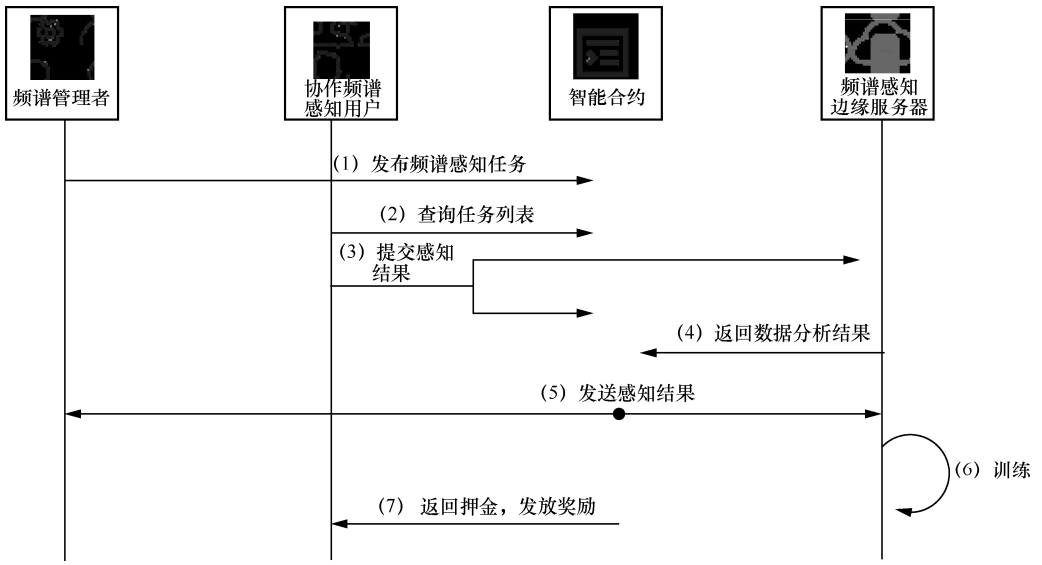


图 2 协作频谱感知流程

步骤 7　智能合约以频谱感知边缘服务器的数据分析结果为依据，对是否给协作频谱感知用户返回押金和发放奖励做出选择。

3　智能合约的设计及测试

针对前文所提出的基于区块链的智能化协作频谱感知系统，本章将详细阐述该系统智能合约中函数的组成与设计，并展示在 Remix 集成开发环境上的实验测试结果。

3.1　智能合约的设计

3.1.1　变量符号

本文采用 Solidity 语言编写智能合约，并在 Remix 上对智能合约进行了编译、部署和测试。智能合约变量说明见表 1。

表 1　智能合约变量说明

变量	名称	类型
addr	地址	address
taskId	任务 ID	uint
goalNum	目标招募人数	uint
currentNum	当前招募的人数	uint
deposit	押金	uint
channelNum	需检测的信道数	uint
award	奖励	uint
notice	公告	string
sensingData	感知数据	string
taskEnd	任务结束时间	uint
resultData	感知融合后的结果	string

3.1.2　用户注册及任务发布

频谱管理者将押金和以太币奖励的数额、需要招募的协作频谱感知用户的数量发布到链上，其功能的具体实现如函数 1 所示；在每轮协作频谱感知规定的任务时间内，如果没有达到频谱管理者规定的招募数额，协作频谱感知用户可以进行注册并将感知数据和规定的押金上传到链上，其功能的具体实现如函数 2 所示。

函数 1　频谱管理者任务发布函数

输入　Admin{taskId, goalNum, currentNum, channelNum, deposit, award, notice},

输出　successfully submit

```
function AdministratorSubmit(uint taskId,uint goalNum, uint currentNum, uint deposit, uint channelNum, uint award, string notice)
```

函数 2　协作频谱感知用户注册以及提交感知数据函数

输入　User{addr,taskId,sensingData, deposit}

输出　successfully submit

```
function UserSubmit(address addr, string sensingData,uint deposit)
    require now <= taskEnd
    require msg.value >= deposit
    require currentNum <= goalNum
```

3.1.3　频谱感知数据融合

在每轮协作频谱感知任务结束之后，频谱管理者将调用频谱感知融合函数，第 5.1～5.2 节将对本文采用的基于性能权重和空间分集的硬判决协作频谱感知融合算法展开详细的分析，此处本文采用传统的 *K*-*N* 硬判决融合准则在合约中进行测试，*K*-*N* 硬判决融合准则的含义是：如果在 *K* 个提交了数据的用户中有 *N* 个用户的数据是一致的，则将该 *N* 个用户的数据结果视为最终判决结果。在智能合约中，本文将 *N* 值设置为 *K*/2，其功能的具体实现如函数 3 所示。

函数 3　频谱感知数据融合函数

输入　dataFusion{taskId}

输出　successfully fusion

```
function fusion(uint resultData)
    if currenrNum > goalNum then
        for each channelNum do
            for each sensingData do
                resultDa-
```

```
ta[channelNum]+=sensingData[channelNum]
                    if resultData [chan-
nelNum] > N then
                        resultDa-
ta[channelNum]    = 1
                    else
            resultData[channelNum]= 0
```

3.1.4 价值评估

在当前轮次的频谱感知任务结束之后，系统将执行价值评估函数，该函数会对每一个协作频谱感知用户进行分析，判别其是否是诚实的用户，如果是，则返回押金并给与以太币奖励，否则不会退回押金，也不会获得任何以太币奖励。本文采用基于深度学习的方法对用户是否诚实进行判定，这将在第 4.2～4.3 节展开详细分析，此处的仿真仅采用一致判别法，即当用户提交的感知数据与最终融合后的数据一致才认为该协作频谱感知用户是诚实的。其功能的具体实现如函数 4 所示。

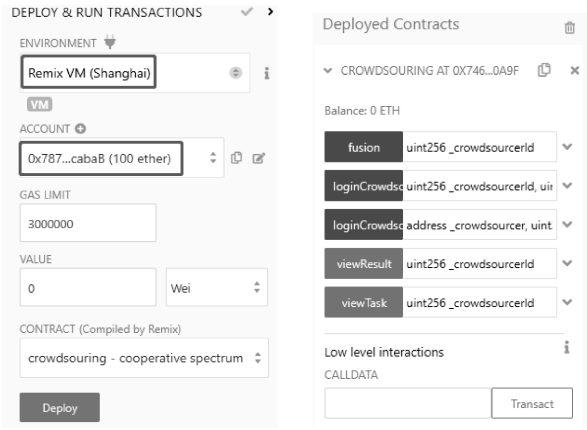
函数 4 价值评估函数

```
function evaluate()
if sensingData == resultData then
    addr.transfer(deposit + award)
```

3.2 智能合约的测试

本文在 Remix 上对智能合约进行了测试，通过 1 位频谱管理者、3 位协作频谱感知用户进行测试，智能合约测试数据见表 2。

智能合约界面如图 3 所示。首先进入智能合约的部署界面，如图 3（a）所示，选择合约部署的环境是 Remix VM(Shanghai)，在对合约进行部署之后，便可以调用函数接口与合约进行交互。图 3（b）是用户与合约的交互页面，用户在对应的函数框中输入数据，点击函数按钮后便可让合约执行该函数。该页面中的函数从上至下依次是协作频谱感知融合函数、协作频谱感知用户注册函数、频谱管理者注册函数、查看当前感知任务的感知结果函数、查看当前正在发布中的频谱感知任务公告函数。



(a) 智能合约的部署界面　　(b) 用户与合约的交互页面

图 3　智能合约界面

本文使用频谱管理者账户地址 0x78731…cabaB 进行频谱管理者的注册和任务发布，频谱管理者注册函数执行结果如图 4 所示，执行函数之后，频谱管理者发布的招募数量、招募公告、押金数量等信息被成功写入智能合约，并因此消耗了 264 443 gwei 汽油费，汽油费是指执行智能合约所需支付的费用，其单位通常用 gwei 表示，其与以太币的换算关系为 1 gwei 等于 10^{-9} 个以太币，汽油费的消耗量与智能合约代码的复杂度以及数据量的大小有关。

表 2　智能合约测试数据

用户	用户类别	地址	频段数	押金	奖励	感知数据	招募数量
Admin1	频谱管理者	0x78731…cabaB	1	2ether	6ether	\	3
User1	协作频谱感知用户	0x5B38D…eddC4	1	\	\	1	\
User2	协作频谱感知用户	0xAb848…35cb2	1	\	\	0	\
User3	协作频谱感知用户	0x4B209…C02db	1	\	\	1	\

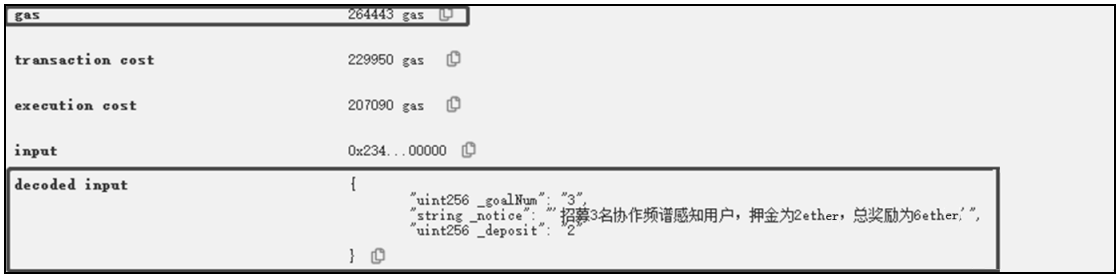


图 4　频谱管理者注册函数执行结果

随后，协作频谱感知用户便可以在合约中查询到当前发布中的协作频谱感知任务，调用频谱管理者注册函数的结果如图 5 所示，在输入框中输入 1 再点击 viewTask 按钮表示查询第一个发布协作频谱感知任务的频谱管理者的发布信息。下方出现的查询结果中第一行是发布的公告，第二行是当前已经招募到的协作频谱感知用户的人数，第三行是该任务总的以太币奖励，单位为 wei。

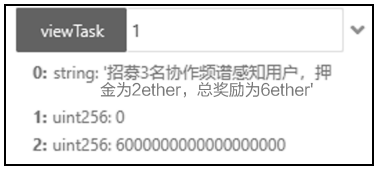


图 5　调用频谱管理者注册函数的结果

在频谱感知招募阶段结束以后，本文在频谱管理者的账户中调用 fusion 函数，合约执行频谱感知融合，随后协作频谱感知用户 User1 和 User3 均收到退回的 2 个以太币押金和 3 个以太币奖励，协作频谱感知用户 User2 因被判决为非诚实的用户，押金没有退还。感知数据融合结果如图 6 所示，是频谱管理者调用 viewResult 函数后得到的协作频谱感知的最终结果，在指定频段上的感知结果为 1，被占用。任务结束后账户以太币的变化如图 7 所示，是任务结束后各账户的以太币数额（初始状态每个账户各有 100ether）。

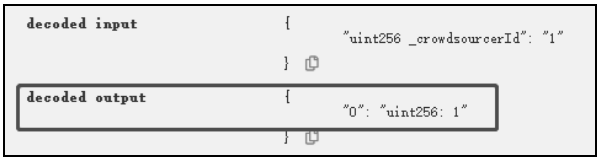


图 6　感知数据融合结果

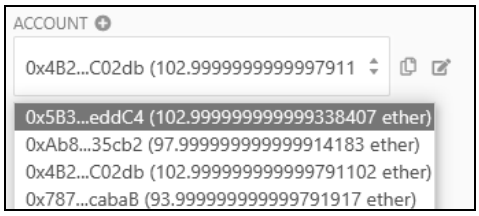


图 7　任务结束后账户以太币的变化

4　基于深度学习的恶意用户识别方法

针对上文提出的协作频谱感知系统，本节将详细阐述系统中可能存在的恶意用户及其行为，提出采用深度学习的方式对系统中的恶意用户进行检测，以保持系统的纯洁性、公平性。

4.1　协作频谱感知中存在的安全威胁

在本文所提出的协作频谱感知系统中，按照攻击目的划分，大致存在以下类型的恶意用户。

（1）企图破坏融合中心判决结果的用户

其发动的攻击称为伪造频谱感知数据（spectrum sensing data falsification，SSDF）攻击[20]。从恶意用户发动攻击的概率角度上看，主要分为“Always”型 SSDF 攻击和概率型 SSDF 攻击。概率型 SSDF 的攻击者会按一定的概率选择是否发动攻击，因此攻击者更难以被识别，目前的研究大部分关注于“Always”型攻击，其主要分为以下 3 种类型：

- “Always Yes”攻击，即只发送表示各个频段主用户存在的感知数据，其目的是阻止认知用户接入空闲频段；
- “Always No”攻击，即只发送表示各个频段主用户不存在的感知数据，其目的是干

扰主用户的通信；

- “Always False”攻击，即发送和本地感知结果相反的感知数据，其目的是破坏融合中心的判决结果。

理论上来说，“Always”型 SSDF 是一种特殊的概率型 SSDF，即，概率为 1 的概率型 SSDF。

（2）企图骗取以太币的用户

该类用户不会在本地进行频谱感知，而是发送一段随机生成的二进制感知数据到智能合约，以骗取以太币奖励的用户，本文将该恶意行为称为 EF（Ether fraud），同理，也分为“Always”型 EF 和概率型 EF，概率型 EF 因其有时候会发送真实的感知数据而较难被识别。

4.2 问题定义与数据生成

该部分将介绍基于深度学习的恶意用户识别的优化问题与数据生成，本文使用 Python 编程语言模拟了用户参加协作频谱感知任务的过程，并记录下整个过程中产生的数据，随后对数据进行清洗、处理，以形成能够作为神经网络的输入数据。对于每个参与协作频谱感知任务的用户，先在本地进行单用户频谱感知，单用户频谱感知的结果可以表示为一个二元假设检验问题，分别用 H_0 和 H_1 表示目标频段空闲和占用，用在每轮频谱感知任务中，第 k 个协作频谱感知用户在第 n 个频段上采样得到的信号可以表示为：

$$x_k(n)=\begin{cases} w_k(n), & H_0 \\ h_k(n)s(n)+w_k(n), & H_1 \end{cases} n=1,2,\cdots,N \quad (1)$$

其中，N 为感知频段的总数，$w_k(n)$ 表示第 k 个协作频谱感知用户在第 n 个频段上的附加噪声，$s(n)$ 表示在第 n 个频段上的采样值，$h_k(n)$ 表示在第 n 个频段上的信道增益。智能合约收集到的所有招募的协作频谱感知用户提交的数据可表示为：

$$C=\{c_1,c_2,\cdots,c_K\} \quad (2)$$

其中，每个协作频谱感知用户提交的感知数据表示为：

$$c_k=\left(c_{k,1},c_{k,2},\cdots,c_{k,n}\right) \quad (3)$$

其中，$c_{k,n}$ 表示第 k 个用户在第 n 个频段上的感知结果，$c_{k,n}\in\{0,1\}$，$c_{k,n}=0$ 表示空闲，$c_{k,n}=1$ 表示占用。为了看出不同用户感知数据之间的差异程度，本文以 $d_{i,j}$ 表示用户 i 与用户 j 感知数据之间的感知距离，计算式如下：

$$d_{i,j}=\sum_{n=1}^{N}\left|c_{i,n}-c_{j,n}\right| \quad (4)$$

在进行数据分析时，考虑用户的历史数据，本文实际使用的特征是用户 i 的平均感知距离：

$$\bar{d}_i=\frac{1}{K}\sum_{k=1}^{K}d_{i,k} \quad (5)$$

感知距离示意图如图 8 所示，当系统中的恶意用户占比较少时，诚实用户的平均感知距离要小于恶意用户。这是本文将该数据作为神经网络输入特征的原因。同理，本文还将用户提交的数据与融合中心的判决结果之间的感知距离也作为一种判别恶意用户的特征数据，并称之为感知偏差。

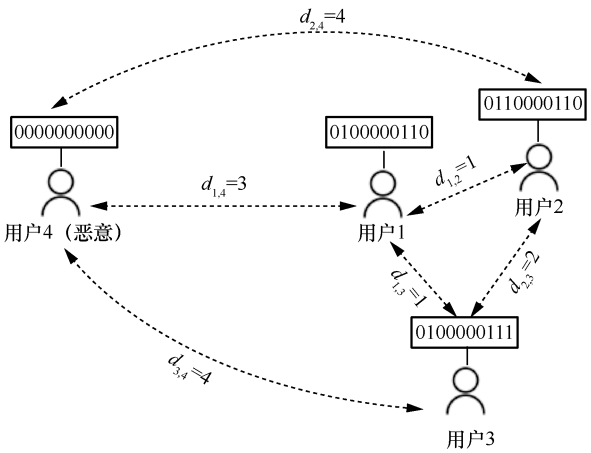
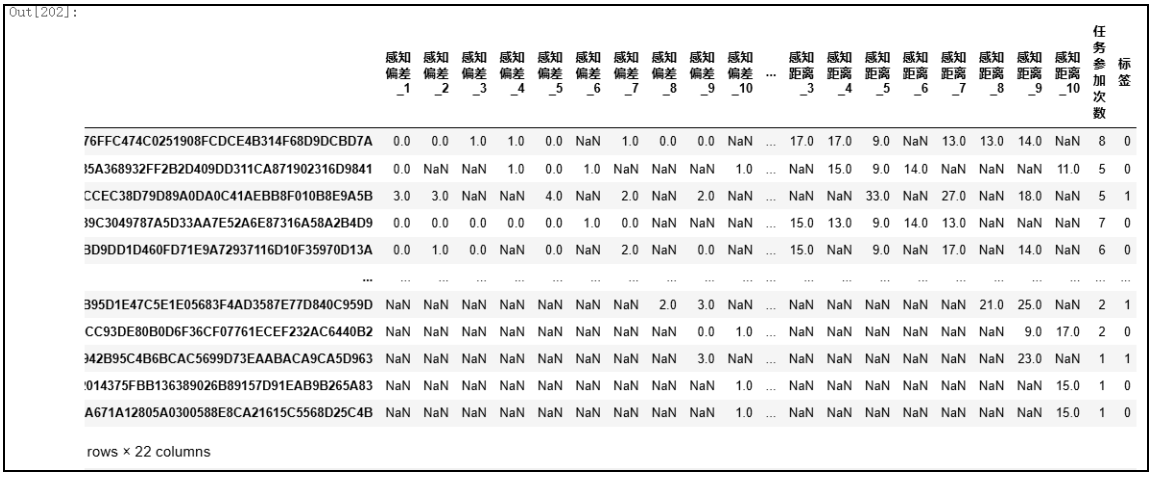


图 8 感知距离示意图

本文使用 Python 对系统的运行过程进行了仿真测试，并采用蒙特卡洛实验，在此过程中收集到了足够多的可用于训练的数据。通过仿真程序，最终生成得到的未处理过的数据如图 9 所示，数据在 Python 中以 DataFrame 的数据类型保存下来。



Out[202]:

	感知偏差_1	感知偏差_2	感知偏差_3	感知偏差_4	感知偏差_5	感知偏差_6	感知偏差_7	感知偏差_8	感知偏差_9	感知偏差_10	...	感知距离_3	感知距离_4	感知距离_5	感知距离_6	感知距离_7	感知距离_8	感知距离_9	感知距离_10	任务参加次数	标签
76FFC474C0251908FCDCE4B314F68D9DCBD7A	0.0	0.0	1.0	1.0	0.0	NaN	1.0	0.0	0.0	NaN	...	17.0	17.0	9.0	NaN	13.0	13.0	14.0	NaN	8	0
5A368932FF2B2D409DD311CA871902316D9841	0.0	NaN	NaN	1.0	0.0	1.0	NaN	NaN	NaN	1.0	...	NaN	15.0	9.0	14.0	NaN	NaN	NaN	11.0	5	0
CCEC38D79D89A0DA0C41AEBB8F010B8E9A5B	3.0	3.0	NaN	NaN	4.0	NaN	2.0	NaN	2.0	NaN	...	NaN	NaN	33.0	NaN	27.0	NaN	18.0	NaN	5	1
9C3049787A5D33AA7E52A6E87316A58A2B4D9	0.0	0.0	0.0	0.0	0.0	1.0	0.0	NaN	NaN	NaN	...	15.0	13.0	9.0	14.0	13.0	NaN	NaN	NaN	7	0
3D9DD1D460FD71E9A72937116D10F35970D13A	0.0	1.0	0.0	NaN	0.0	NaN	2.0	NaN	0.0	NaN	...	15.0	NaN	9.0	NaN	17.0	NaN	14.0	NaN	6	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
395D1E47C5E1E05683F4AD3587E77D840C959D	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.0	3.0	NaN	...	NaN	NaN	NaN	NaN	NaN	21.0	25.0	NaN	2	1
CC93DE80B0D6F36CF07761ECEF232AC6440B2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.0	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	9.0	17.0	2	0
42B95C4B6BCAC5699D73EAABACA9CA5D963	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	3.0	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	23.0	NaN	1	1
014375FBB136389026B89157D91EAB9B265A83	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15.0	1	0
A671A12805A0300588E8CA21615C5568D25C4B	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15.0	1	0

rows × 22 columns

图 9　生成得到的未处理过的数据

由于用户并非参与每一轮的协作频谱感知任务，因此可以看到数据是高度稀疏的，数据量的缺乏会影响神经网络的模型效果，本文需要对数据进行预处理。

数据处理的步骤如图 10 所示。

协作频谱感知用户的真实标签构成的集合为 $Y=\{y_1,y_2,\cdots,y_n\}$，将处理后的数据与真实标签对应后输入模型进行训练，最终的预测值为：$Y=\{\hat{y}_1,\hat{y}_2,\cdots,\hat{y}_n\}$。损失函数采用交叉熵损失：

$$L(y_i,\hat{y}_i)=-\sum_{i=1}^{N}\left[y_i\log\hat{y}_i+(1-y_i)\log(1-\hat{y}_i)\right] \quad (6)$$

恶意用户识别可以表示为以下优化问题：

$$\min L(y_i,\hat{y}_i)=-\frac{1}{N}\sum_{i=1}^{N}\left[y_i\log\hat{y}_i+(1-y_i)\log(1-\hat{y}_i)\right] \quad (7)$$

4.3　实验结果及分析

在这一部分，本文利用 Python 语言对基于深度学习的恶意用户识别方法进行了测试与评估。在本文的引言部分谈到，已经有部分国内学者对协作频谱感知中的恶意用户攻击防御问题展开了大量研究工作，但是据笔者所知，这是第一个将深度学习应用于抵御骗取以太币行为的工作。

本文所使用的性能指标有 AUC、Accuracy、F1-score，AUC 是 ROC 曲线与坐标轴所围成的面积，ROC 曲线的横坐标是假阳率（false positive rate，FPR），纵坐标是真阳率（true positive rate，TPR），该曲线能反映出分类器在不同阈值下的性能表现，AUC 的值为 0.5～1，在本文的恶意用户检测算法中，其表示模型能正确预测某一用户是否为恶意用户的性能，值越大则模型的预测性能越好。

准确率 Accuracy 表示分类器将所有样本正确分类的比例，即分类器正确分类的样本数与总样本数之比，值越大表示分类性能越好。准确率可以定义为：

$$\text{Accuracy}=\frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \quad (8)$$

其中，TP（true positive）表示真正例，TN（true negative）表示真负例，FP（false positive）表示假正例，FN（false negative）表示假负例。

F1-score 是一种兼顾了精确率 Precision 和召

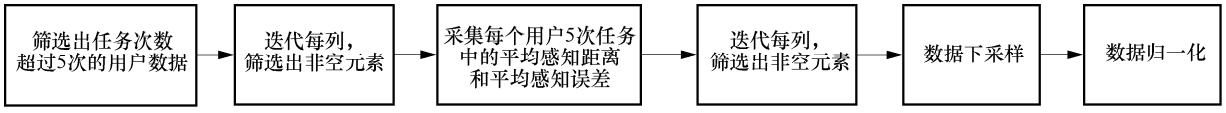


图 10　数据处理的步骤

回率 Recall 的性能指标，值越大表示分类性能越好，F1-score 可以定义为：

$$\text{F1-score} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

召回率定义为：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

精确率定义为：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

其中，精确率指的是模型预测为正样本中真正是正样本的比例，召回率指的是数据集中的正样本中被模型预测为正样本的比例。

笔者将本节所提出的算法与其他算法进行比较，具体如下。

（1）随机森林（random forest）：是一种集成学习算法，它的目的是在原始数据集上建立多个决策树，并取其平均值以提高准确性和避免过拟合的问题。

（2）逻辑回归（logistic regression）：是一种用于类别推断和概率估计的统计学方法，通过将输入特征与输出的对数概率之间的线性关系建模来进行分类，逻辑回归采用 sigmoid 函数来产生对数概率输出，将线性预测值映射到 0 和 1 之间的实数值。

（3）支持向量机（support vector machine，SVM）：SVM 方法的核心是在特征空间中寻找最优的分类超平面，使得正例和负例之间的间隔最大化，在处理非线性问题时，可以通过将数据映射到高维空间来实现线性分割。

（4）LightGBM（light gradient boosting machine）：一种基于梯度提升决策树（gradient boosting decision tree，GBDT）的预测模型，是一种集成学习方法，它采用了类似于 GBDT 的增量的提升策略来训练决策树，使得每一个新的决策树可以更好地拟合残差。与传统的 GBDT 相比，它采用了一些新的技巧来提高模型训练速度和准确率，是目前最先进的 GBDT 框架之一。

（5）XGBoost（eXtreme gradient boosting）是基于 gradient boosting 框架的提升树模型库，是一种集成学习方法，该算法的主要贡献来自它对决策树进行优化的方法，包括目标函数、正则化项、分裂点选择。

将上述所有算法在数据集上进行训练和预测，最后获得的实验结果，即在仿真数据集中各模型的性能对比见表 3。

本文采用的 DNN 模型在 AUC、准确率和 F1-score 性能指标上都优于其他机器学习算法，结果表明本模型能提供较优的恶意用户检测性能。

30%的概率型 EF 下的 ROC 曲线如图 11 所示，其描述了当发动恶意攻击概率为 30%的概率型 EF 行为存在于系统中时，所提出的基于深度学习的恶意用户检测算法与其他机器学习算法分别在 ROC 曲线上的性能表现进行了比较。

表 3　在仿真数据集中各模型的性能对比

对比项	随机森林	逻辑回归	SVM	LightGBM	XGBoost	DNN
准确率	0.777	0.645	0.639	0.760	0.731	0.787
AUC	0.835	0.713	0.703	0.825	0.800	0.860
F1-score	0.759	0.652	0.651	0.750	0.722	0.765
召回率	0.698	0.662	0.671	0.715	0.700	0.690

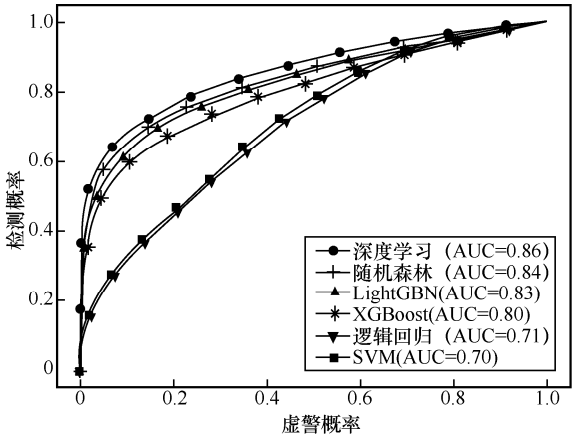


图 11　30%的概率型 EF 下的 ROC 曲线

在系统实际的运行过程中，恶意用户识别的效果将比仿真得到的性能更好，其原因是在生成数据的仿真程序中，笔者采用 *K*-*N* 硬判决法则对所有用户提交的感知数据进行融合，在融合之前并未使用训练好的模型对恶意用户进行排除，也未采用第 5 节提到的更优的协作频谱感知融合算法。而在实际设想中的每轮频谱感知任务中，首先通过事先训练好的模型识别排除恶意用户，然后利用基于性能权重和空间分集的协作频谱感知融合算法进行感知数据融合，从而有助于进一步提高感知和恶意用户识别的准确率。

5　基于性能权重和空间分集的硬判决协作频谱感知融合算法设计

传统的 *K*-*N* 硬判决频谱感知融合算法没有考虑不同的感知设备之间的性能差异，而且无法解决隐藏终端问题，使得处于同一干扰环境中的用户同时被招募作为协作频谱感知节点，这样不仅会增加频谱管理者的以太币奖励成本，而且会导致感知准确率低下。为了更好地解决上述问题，提出了一种基于性能权重和空间分集的硬判决协作频谱感知融合算法。首先，在频谱感知边缘服务器端通过分析用户的感知数据和当前用户的位置信息来判断用户的受干扰情况，本文采用 *K*-means 聚类算法，把在相同频段受到干扰的用户聚为一类；然后在不同聚类的簇中，依据用户的历史数据对不同用户的感知性能做一个排序，并依次从每个簇中挑选出感知性能最佳的用户进行 *K*-*N* 硬判决融合。

5.1　问题定义

假设有 K 个用户，每个用户都对 N 个频段进行检测，频谱感知边缘服务器端收到的样本数据集为：

$$S=\{\boldsymbol{s}_1,\boldsymbol{s}_2,\cdots,\boldsymbol{s}_K\} \tag{12}$$

其中，每个用户提交的数据 $\boldsymbol{s}_k(k=1,2,\cdots,K)$ 为一个 N+2 维的向量，包含该用户在 N 个频段上的感知结果和二维的地理位置坐标信息。接下来，本文提出一种改进的 *K*-means 算法对受到相似干扰的用户进行聚类。

传统的 *K*-means 算法会采用欧氏距离的距离计算方式去衡量每个样本与聚类中心的距离，从而量化出不同数据之间的相似度。显然，感知结果是由二进制字符组成，本文不能直接使用衡量两点之间距离的计算方式。其次，地理位置数据和二进制感知数据在数值大小、数据量上差异较大、在重要性上可能也有所不同，并且二进制感知数据包含了 N 个二进制字符，这 N 个二进制字符理应整体作为一个特征，因此本文不能直接对数据进行归一化。

考虑上述情况，本文所采用的方法是分别单独计算二进制感知数据和地理位置的距离矩阵，随后按照一定的权重比例合并成一个混合距离矩阵。二进制感知数据距离矩阵和地理位置距离矩阵分别如下：

$$\text{Sensing\_matrix}=\begin{bmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,K} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ d_{K,1} & d_{K,2} & \cdots & d_{K,K} \end{bmatrix} \tag{13}$$

其中，元素 $d_{i,j}$ 表示第 i 个用户与第 j 个用户之间的感知距离。地理位置距离矩阵如下：

$$\text{Euclidean\_matrix}=\begin{bmatrix} E_{1,1} & E_{1,2} & \cdots & E_{1,K} \\ E_{2,1} & E_{2,2} & \cdots & E_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ E_{K,1} & E_{K,2} & \cdots & E_{K,K} \end{bmatrix} \tag{14}$$

其中，元素 $E_{i,j}$ 表示第 i 个用户与第 j 个用户地理位置之间的欧氏距离。笔者把得到的两个距离矩阵进行归一化，把特征缩小到[0,1]区间，再按照 0.5:0.5 的权重比拼接到一起成为融合后的混合距离矩阵。

K-means 聚类算法中簇的数目需要在算法运行之前设置，考虑在实际场景中本文并不清楚具体有多少不同类别的相似干扰用户。针对这个问题，笔者采用了手肘法和轮廓系数法对簇的数目进行判定，手肘法是计算每个点到最近簇的平方距离的总和，绘制误差平方和（sum of squared error，SSE）与 k 值的图形，然后把 k 值逐渐增加，选择在 k 值增加时 SSE 值变化下降最快的点作为最佳簇数目。手肘法曲线如图 12 所示，即采用手肘法绘制的曲线，可以看出，当前最佳簇的数目 I^* 是 5。考虑有些时候采用手肘法最佳簇的数目不好判断，本文又结合轮廓系数法进行判断，对于每个数据点，计算它与同一簇中所有其他点之间的平均距离和最近簇中所有点之间的平均距离，然后计算该点的轮廓系数（sihouette coefficient）。对于整个数据集的平均轮廓系数，选择其值最大的 k 作为最佳簇数目。轮廓法曲线如图 13 所示，即采用轮廓系数法绘制的曲线，可以看出，当前最佳簇的数目同样是 5。

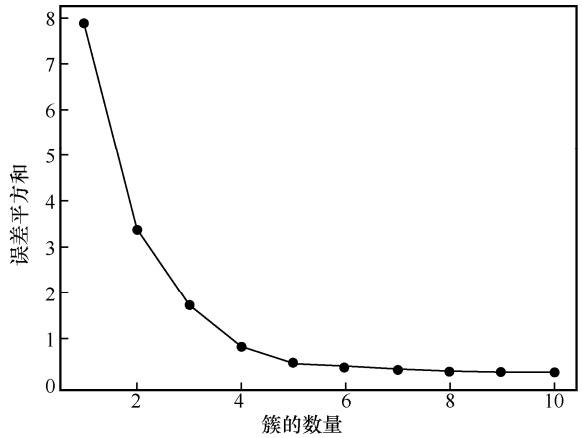


图 12　手肘法曲线

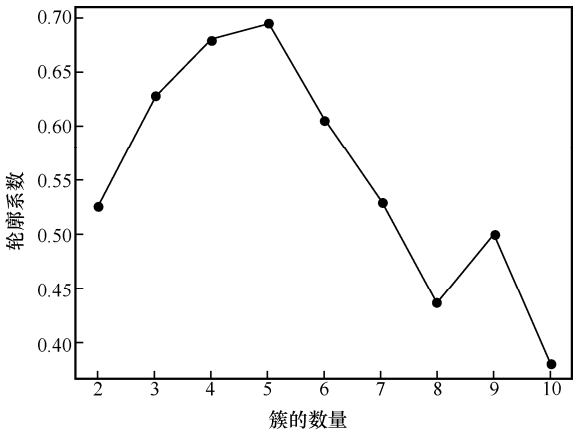


图 13　轮廓法曲线

在确定好簇的数目之后，本文便采用 *K*-means 算法来对融合后的距离矩阵进行聚类，最终将处于同一干扰环境中的用户聚类到相同的簇中。随后，在不同的簇中，选择感知性能最佳的用户，直到选择的用户数达到招募需求。整个过程在算法 1 中描述，具体如下。

算法 1　基于性能权重和空间分集的硬判决频谱感知融合算法

输入　$S=\{\boldsymbol{s}_1,\boldsymbol{s}_2,\cdots,\boldsymbol{s}_k\},K$

输出　sensing_result

采用手肘法和轮廓法判定出最佳簇的个数 I^*；

创建 I^* 个点作为起始质心；

当任意一个点的簇分配结果改变时：

对数据集中的每个数据点；

对每个质心；

计算质心与数据点之间的距离；

将数据点分配到距其最近的簇；

对每一个簇，计算簇中所有点的均值并将均值作为质心；

对每一个簇，将簇中的所有用户按照平均感知偏差从小到大进行排序；

当没有达到招募的目标人数 K 时：

对每一个簇；

招募该簇中第一个用户，随后将该用户从簇的集合中剔除；

K += 1

综合所有的用户数据，计算每个频段被感知为 1 的次数和 0 的次数；

如果该频段被感知为 1 的次数多于 0，则记录该频段为 1，否则为 0；

结束；

返回最终的全频段感知融合结果 sensing_result

5.2　实验结果及结果分析

本节将通过仿真实验来评估所提方案的性能。根据上文的手肘法和轮廓系数法，本文得到了合适的聚类簇数，将受到相似干扰的用户进行聚类，聚类效果如图 14 所示。图 14（a）是协作感知用户的分布场景，同时展示了在地理位置维度上的聚类效果，可以看到地理位置距离较近的用户受到的干扰也相似。为了在感知数据上实现聚类效果的可视化，笔者对其进行 PCA 降维处理，将 n 维的二进制感知数据映射为二维数据，其数据的聚类效果如图 14（b）所示。

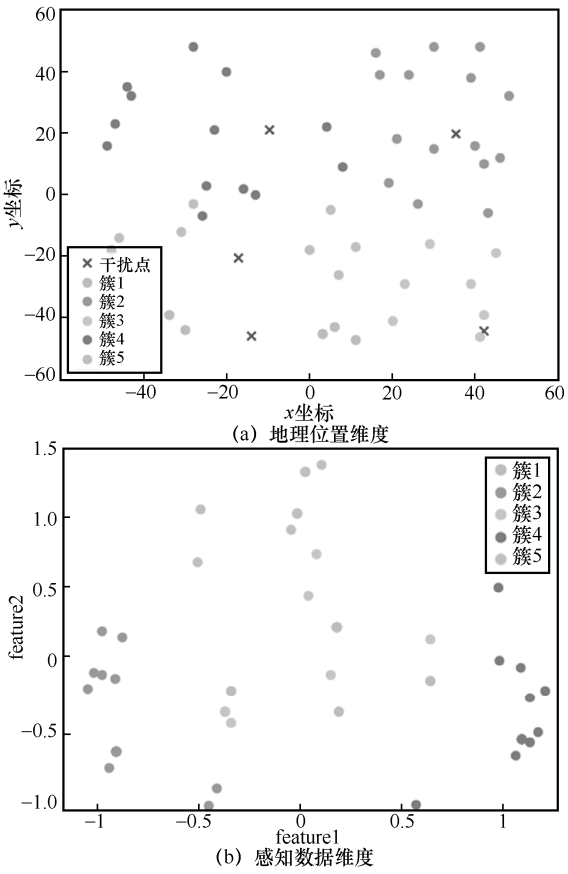


图 14　聚类效果

完成聚类相当于找出了在相同干扰环境中的用户，随后在不同的簇中，依据历史数据对用户的感知性能做一个排序，依次从不同簇中选择历史感知性能最佳的用户作为协作频谱感知节点再进行 *K-N* 硬判决融合法则。采用 100 次蒙特卡洛实验，把本文提出的基于性能权重的空间分集算法与传统的 *K-N* 硬判决算法在感知准确率上进行对比，同时，将所提算法拆分，把只采用性能权重的算法和只采用空间分集的算法也进行了实验分析，所提算法在实验中的性能如图 15 所示。

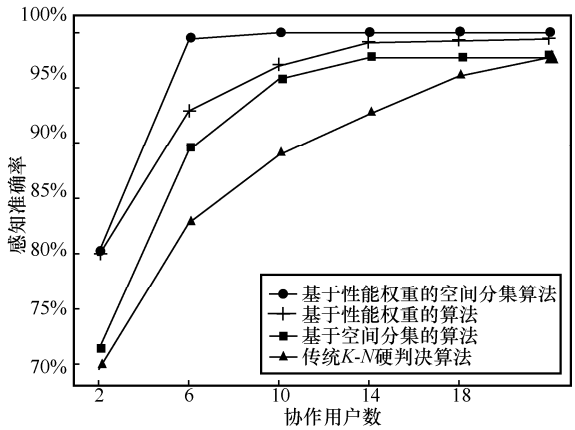


图 15　基于性能权重和空间分集的硬判决协作频谱感知融合算法性能

实验结果显示，对于不同的协作用户招募数量，本文所提的基于性能权重和空间分集的硬判决协作频谱感知融合算法在感知准确率方面优于传统的 *K-N* 硬判决算法，此外，本文将所提算法拆分后进行实验，只采用性能权重算法与只采用空间分集算法也均优于传统的 *K-N* 硬判决算法。当招募了 6 个协作用户时，所提算法的感知准确率接近 100%，对比传统的 *K-N* 硬判决算法，显然更能保障在保持较高的感知准确率的同时付出更少的协作用户招募成本。

6　结束语

本文研究了目前协作感知存在的安全性、激励性、隐藏终端的问题，为了避免集中式协

作频谱感知所带来的安全隐患，设计了一种运行在区块链智能合约上的协作频谱感知系统，并通过建立奖惩机制，结合人工智能自动识别恶意用户、奖励诚实且感知性能较高的用户，很好地解决了激励性和信任的问题。然后，提出了一种基于性能权重和空间分集的硬判决协作频谱感知融合算法，解决了隐藏终端导致的感知性能不佳的问题，且基于数据分析的方法，能够让系统在频谱感知数据融合阶段做出最佳选择。仿真结果表明，本文方案优于现有的集中式协作频谱感知。

参考文献：

[1] Cisco. Cisco visual networking index: global mobile data traffic forecast update, 2017–2022[R]. 20197.

[2] GAO A, DU C Y, NG S X, et al. A cooperative spectrum sensing with multi-agent reinforcement learning approach in cognitive radio networks[J]. IEEE Communications Letters, 2021, 25(8): 2604-2608.

[3] SHI Z J, GAO W, ZHANG S W, et al. Machine learning-enabled cooperative spectrum sensing for non-orthogonal multiple access[J]. IEEE Transactions on Wireless Communications, 2020, 19(9): 5692-5702.

[4] JANU D, SINGH K, KUMAR S. Machine learning for cooperative spectrum sensing and sharing: a survey[J]. Transactions on Emerging Telecommunications Technologies, 2022, 33(1): 43-52.

[5] 王晴天, 刘洋, 刘海涛, 等. 面向 6G 的网络智能化研究[J]. 电信科学, 2022, 38(9): 151-160.
WANG Q T, LIU Y, LIU H T, et al. Research on network intelligence for 6G[J]. Telecommunications Science, 2022, 38(9): 151-160.

[6] Federal Communication Commission. Spectrum policy task force report[R]. 2002.

[7] PANDIT S, SINGH G. An overview of spectrum sharing techniques in cognitive radio communication system[J].Wireless Networks, 2017, 23(2): 497-518.

[8] DENG Q Y, LI Z T, CHEN J B, et al. Dynamic spectrum sharing for hybrid access in OFDMA-based cognitive femtocell networks[J]. IEEE Transactions on Vehicular Technology, 2018, 67(11): 10830-10840.

[9] SARIKHANI R, KEYNIA F. Cooperative spectrum sensing meets machine learning: deep reinforcement learning approach[J]. IEEE Communications Letters, 2020, 24(7): 1459-1462.

[10] JALIL S Q, CHALUP S, REHMANI M H. Cognitive radio spectrum sensing and prediction using deep reinforcement learning[C]//Proceedings of 2021 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE Press, 2021: 1-8.

[11] LIU C H, TSAI H C. Traffic management for heterogeneous networks with opportunistic unlicensed spectrum sharing[J]. IEEE Transactions on Wireless Communications, 2017, 16(9): 5717-5731.

[12] ZHANG Y W, ZHANG S C, WANG Y H, et al. Riemannian mean shift-based data fusion scheme for multi-antenna cooperative spectrum sensing[J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(1): 47-56.

[13] CHEN Z B, XU Y Q, WANG H B, et al. Deep STFT-CNN for spectrum sensing in cognitive radio[J]. IEEE Communications Letters, 2021, 25(3): 864-868.

[14] LEE W, KIM M, CHO D H. Deep cooperative sensing: cooperative spectrum sensing based on convolutional neural networks[J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 3005-3009.

[15] LIU C, WANG J, LIU X M, et al. Deep CM-CNN for spectrum sensing in cognitive radio[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(10): 2306-2321.

[16] 季薇, 李炳星, 郑宝玉. 基于信誉与共识的分布式智能入侵防御方案[J]. 系统工程与电子技术, 2018, 40(3): 665-670.
JI W, LI B X, ZHENG B Y. Distributed intelligent intrusion prevention scheme based on reputation and consensus[J]. Systems Engineering and Electronics, 2018, 40(3): 665-670.

[17] 张孟伯, 王伦文, 冯彦卿. 基于强化学习和共识融合的分布式协作频谱感知方法[J]. 系统工程与电子技术, 2019, 41(3): 486-492.
ZHANG M B, WANG L W, FENG Y Q. Distributed cooperative spectrum sensing method based on reinforcement learning and consensus fusion[J]. Systems Engineering and Electronics, 2019, 41(3): 486-492.

[18] 苗成林, 李彤, 吕军, 等. 基于 Dempster-Shafer 证据理论与抗频谱感知数据篡改攻击的协作式频谱检测算法[J]. 兵工学报, 2017, 38(12): 2406-2413.
MIAO C L, LI T, LYU J, et al. Cooperative spectrum sensing algorithm against spectrum sensing data falsification based on dempster-shafer evidence theory[J]. Acta Armamentarii, 2017, 38(12): 2406-2413.

[19] FENG D Q, ZHANG L, ZHANG S L, et al. Blockchain-based secure crowdsourcing in wireless IoT[J]. Journal of Communi-

cations and Information Networks, 2022, 7(1): 23-36.

[20] PATNAIK M, PRABHU G, REBEIRO C, et al. ProBLeSS: a proactive blockchain based spectrum sharing protocol against SSDF attacks in cognitive radio IoBT networks[J]. IEEE Networking Letters, 2020, 2(2): 67-70.

[作者简介]

肖仲杉（1999− ），男，深圳大学硕士生，主要研究方向为频谱共享共存。

王春琦（1991− ），男，国家无线电监测中心工程师，主要研究方向为联盟链监管和频谱共享共存。

冯大权（1986− ），男，博士，深圳大学副教授，主要研究方向为 D2D 通信、频谱共享共存、车联网等。