Received 14 March 2025, accepted 9 April 2025. Date of publication 00 xxxx 0000, date of current version 00 xxxx 0000.

Digital Object Identifier 10.1109/ACCESS.2025.3560213

Spectrum Hole Geolocation for Database-Driven IoT-Enabled Dynamic Spectrum Access

DAYAN ADIONEL GUIMARÃES
National Institute of Telecommunications-Inatel, Santa Rita do Sapucaí, Minas Gerais 37540-000, Brazil
e-mail: dayan@inatel.br

This work was supported in part by the Rede Nacional de Ensino e Pesquisa [National Education and Research Network (RNP)] with Resources from Ministério da Ciência, Tecnologia, Inovações e Comunicações [Ministry of Science, Technology, Innovations and Communications (MCTIC)] through the "Brazil 6G Project" of the Radio-Communication Reference Center of Inatel under Grant 01245.020548/2021-07; in part by the Empresa Brasileira de Pesquisa e Inovação Industrial [Brazilian Company for Industrial Research and Innovation (EMBRAPII)]-Inatel Competence Center on 5G and 6G Networks with Resources from Programa de Parcerias de Investimentos [Investment Partnerships Program (PPI)] Internet of Things (IoT)/Manufacture 4.0 of Ministério da Ciência, Tecnologia e Inovação [Ministry of Science, Technology and Innovation (MCTI)] and Fundação de Amparo à Pesquisa do Estado de Minas Gerais [Research Support Foundation of the State of Minas Gerais (FAPEMIG)], PPE-00124-23, Project XGM-AFCCT-2024-4-1-2, under Grant 052/2023; in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo [Research Support Foundation of the State of São Paulo (FAPESP)] under Grant 2022/09319-9; in part by FAPEMIG under Grant APQ-05305-23; and in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico [National Council for Scientific and Technological Development (CNPq)] under Grant 302589/2021-0.

ABSTRACT The effective realization of dynamic spectrum access (DSA) subsumes that, when a secondary terminal seeks a vacant band, the geolocation of the candidate spectrum hole coincides with the position of the terminal. Otherwise, stemming from a spectrum hole detection not corresponding to the terminal location, harmful interference to the primary network may occur. Non-cooperative spectrum sensing can inherently deliver accurate spectrum hole geolocation, as it corresponds to the position of the spectrum sensor. However, performance is degraded due to propagation-induced phenomena in the environment. While the performance of cooperative spectrum sensing is less prone to such phenomena, the accurate geolocation of the spectrum hole is compromised. This is because a vacant band is determined by a collective of spectrum sensors whose dispersed spatial distribution prevents from achieving high resolution concerning where in space the band is in fact vacant. This article delves into the problem of spectrum hole geolocation for DSA. As a solution, it is introduced a cooperative spectrum sensing scheme based on overlapped-clustering, integrated with a database-driven Internet of Things-enabled DSA framework. It is shown that the proposed solution achieves a high rate of correct estimations of spectrum hole geolocation, with small variance and under diverse system conditions.

INDEX TERMS Cognitive radio, dynamic spectrum access, dynamic spectrum sharing, spectrum sensing.

I. INTRODUCTION

The escalating demand for novel telecommunications services has emerged as the primary catalyst for more advanced technologies. This is evident in recent efforts, particularly in the context of the fifth generation (5G) communication networks, the Internet of things (IoT), and ongoing discussions and research on the sixth generation (6G) communication networks [1], [2].

To materialize many envisioned telecommunications services, particularly in the realm of wireless systems, it is imperative to overcome the bottleneck posed by the scarcity of radio-frequency (RF) spectrum. This scarcity is attributed to the prevailing fixed spectrum allocation policy, which confers to the incumbent (primary user, PU) network the exclusive right to use specific frequency portions. Consequently, as wireless communication systems proliferate, only a limited amount of free bands remain available.

There is a widely accepted belief that the fixed spectrum allocation policy may prove inadequate for accommodating further expansions in wireless communication systems. A paradigm shift is needed, wherein spectrum sharing becomes the novel approach. This concept is encapsulated by dynamic spectrum access (DSA) techniques. DSA capitalizes on the under-utilization of many spectral bands already allocated to PU networks, potentially enabling sharing with secondary user (SU) networks [3], [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Pietro Savazzi.

In a scenario of shared spectrum, SU transmissions can occur in two primary ways: simultaneously with PU transmissions, provided no harmful interference is inflicted upon the PU network (referred to as underlay sharing), or in a non-interfering basis, by opportunistically occupying unused licensed bands (known as interweave sharing). This article deals with the interweave spectrum sharing approach.

The cognitive radio (CR) concept [5] has emerged within this spectrum-sharing context. A CR transceiver possesses the capability to acquire cognition about the surrounding environment and the network, adapting operational parameters to optimize performance targets. Among various cognitive attributes, a CR can identify vacant bands for opportunistic access through a technique known as spectrum sensing [1], [6], [7], [8], [9]. In few words, spectrum sensing is the technique for monitoring the RF spectrum to identify vacant spectral bands, commonly referred to as spectrum holes or white-spaces.

Spectrum sensing can be made by each SU, independently of the other SUs, which characterizes a non-cooperative spectrum sensing (NCSS) approach. When NCSS is adopted, a given SU, even if it is reachable by the primary network signal in terms of the distance from the transmitter, it may be incapable of detecting the signal. This situation can happen, for instance, if the sensed signal is subjected to severe fading or is blocked by obstacles between the PU transmitters and the SU receivers. In such situation, an SU might initiate transmission on an occupied frequency band, resulting in interference with primary network devices. To address, or at least mitigate this issue and counteract the adverse effects of multi-path fading and shadowing, an alternative to NCSS is the adoption of cooperative (or collaborative) spectrum sensing (CSS).

As the nomenclature implies, CSS involves a collective effort from a group of SUs to determine the occupancy status of the sensed band, thereby enhancing PU signal detection capabilities. There are two primary forms of CSS: distributed detection with distributed decision, and distributed detection with centralized decision. In the case of distributed decision CSS, the SUs exchange the spectrum sensing information, which includes local decisions or received samples. Subsequently, a collective decision is made, often through consensus, regarding the occupancy status of the sensed frequency band. Conversely, in centralized decision CSS the spectrum sensing information is conveyed to a fusion center (FC), which can be the base station (BS) of the secondary network or a specialized SU. The FC processes this information to arrive at a global decision concerning the state of the sensed band. The resulting decision is then broadcast to the SUs via control channels. Access to the unoccupied band is made through a multiple access technique tailored to the secondary network.

Centralized decision CSS is categorized based on how the spectrum sensing information is transmitted to, and processed by the FC, a process known as fusion. Decision fusion occurs when local decisions on spectrum occupancy are transmitted to the FC. In this case, the logic decision rule k-out-of-m is often applied, declaring a busy frequency band if at least k out of the m cooperating SUs decide that the band is busy. On the other hand, data fusion takes place when the samples received by the SUs, or quantities derived from these samples are sent to the FC, where a test statistic is formed and compared with a decision threshold to yield the final decision [7].

A. PROBLEM DESCRIPTION

A spectrum hole or white-space, refers to a portion of the RF spectrum that is temporarily unused or unoccupied by any licensed or primary users within a specific geographic area and time period. In other words, it represents frequencies within the spectrum that are available for secondary or unlicensed users to utilize for communication purposes without causing harmful interference to the existing primary users.

The concept of spectrum holes arises due to the dynamic nature of radio frequency usage and signal propagation characteristics of the environment. While certain portions of the spectrum are allocated to licensed users, these users may not be continuously active in all allocated frequency bands at all times, and throughout the entire coverage area. As a result, there are often periods and locations in which portions of the spectrum become opportunities for secondary users to access for their own communication needs.

When a secondary terminal is seeking for a vacant band for DSA, the spectrum hole geolocation information must be in agreement with the position of that terminal. While the terminal stays at a given position, a spectrum hole that is detected is only useful if it refers to that position, not to any other. Stemming from a mismatch between spectrum hole detection and terminal location, detrimental interference with the primary network may occur.

It is worth emphasizing that the declaration of a spectrum hole not only depends on the state of the PU transmitter, but also is affected by the propagation characteristics of the coverage area. Even if a PU transmitter is active, it is possible for certain areas within the coverage region to experience weak signal strength or attenuation due to factors such as distance, terrain, and obstacles. Hence, in scenarios where the PU transmitter's signal strength diminishes significantly beyond a certain distance or due to obstructions, SUs may still be able to utilize the frequency spectrum within these areas without causing harmful interference to the primary communication system. This is because the PU network may not be capable of operating in these locations, creating opportunities for SUs to access the spectrum resources without causing disruption to authorized services.

Non-cooperative spectrum sensing inherently ensures precise spectrum hole geolocation, as it corresponds directly to the position of the spectrum sensor. Nevertheless, the spectrum sensing performance may suffer due to propagation-induced phenomena within the environment, like signal shadowing and multi-path fading. Conversely, cooperative spectrum sensing exhibits greater resilience to environmental phenomena but compromises the precision of spectrum hole geolocation, owed to the fact that a vacant band is identified collectively by a distributed set of spectrum sensors, whose dispersed spatial arrangement hinders achieving high resolution regarding the actual geolocation of the vacant band.

The problem described herein can be named spectrum hole localization in the spatial domain, since spectrum hole localization in the frequency and time domains is not a problem because frequency and time information is inherent to the spectrum sensing task. In other words, the instants and frequencies selected for spectrum monitoring are defined on-the-fly by the network, as a normal task in the operation of spectrum sensing. Thus, hereafter, the term spectrum hole geolocation (SHG) refers solely to the spatial information associated with the absence of the primary signal throughout the secondary network area. Roughly speaking, SHG can be linked with a primary signal coverage (from the SUs' perspective) analysis whose result is updated according to the pace of spectrum sensing.

To the best of the authors' knowledge, the real-time spectrum hole geolocation problem has not yet been addressed within the context of dynamic spectrum access. This omission is somewhat surprising and may stem from a common, yet not necessarily accurate, implicit assumption: that a band declared vacant by spectrum sensing is considered available across the entire secondary network. Such an unrealistic approach assumes that any secondary user can access the declared vacant band whenever the primary user's transmitter is inactive.

The DSA approach considered herein aims at a more realistic scenario: a given SU terminal needs to establish communication over a vacant PU network band that is unused either because all PU transmitters are *off*, or the location of the SU terminal corresponds to insufficient PU signal coverage due to the propagation characteristics of the environment. As an example, consider a primary outdoor wireless communication network. The outdoor-to-indoor path-loss is capable of producing spectrum holes inside a building, even if the PU transmitter is on. Similarly, an outdoor signal-shadowed region in the PU network coverage area can be regarded as spectrum holes, from the SU network perspective.

B. RELATED WORK

Solutions for shared spectrum access leveraging the propagation characteristics of the environment already exist. These approaches rely on spectrum availability databases, which secondary users query to identify vacant channels. Such solutions have been adopted by regulatory authorities worldwide and are incorporated into standards such as IEEE 802.22 [10], IEEE 802.11af [11], and the IETF protocol to access white-space (PAWS) [12].

In 2014, the Federal Communication Commission (FCC) designated several database administrators, among which five have been approved: Spectrum Bridge, Inc.; Iconectiv; Keybridge Global, LLC; Google, Inc.; LS Telcom, Inc. and RadioSoft, Inc. RadioSoft has been subsequently acquired by LS Telcom, Inc., but kept in operation [13], [14], [15]. Google has ended its television white-space (TVWS) project in 2018 to prioritize a solution targeted to the citizens broadband radio service (CBRS), managing a spectrum portion reserved to the US Federal Government to avoid interference with the US Navy radar systems and aircraft communications [15], [16], [17].

The main drawbacks of such database-assisted spectrum sharing approaches are: (i) the reliability of the spectrum availability information depends on the accuracy of the signal coverage prediction used to feed the database; (ii) the spectrum availability data does not necessarily correspond to up-to-date information, owed to the fact that it is not feasible to perform real-time coverage prediction [18], [19].

In [15], a comprehensive review of various DSA solutions based on spectrum sensing alone, database alone, and the combination of these two approaches are presented. A promising solution based on such combination is proposed as an attempt to circumvent the above-mentioned drawbacks. It makes use of a supporting IoT network in which some IoT devices are equipped with spectrum sensors, forming the so-called spectrum sensing IoT (SSIoT) devices. These SSIoTs are responsible for frequent spectrum sensing, and for updating a spectrum availability database that is then queried by the SUs when a vacant PU band is needed. Note that, as a drastic paradigm shift, the spectrum sensing task is transferred from the SUs to the supporting network of SSIoT devices, alleviating the complexity and energy consumption of the SU terminals. Moreover, the solution proposed in [15] allows real-time update of the spectrum availability information, which is crucial for implementing DSA when the primary network has a high rate of channel occupation and release. Additionally, thanks to the transfer of sensing tasks to the SSIoTs, the implementation of modern integrated sensing and communications technologies (ISAC) is alleviated [20].[1] Indeed, the same result obtained from ISAC is indirectly achieved by the reception of information on spectrum occupancy retrieved from the database within normal control communication.

The DSA approach proposed in [15] has been adopted as the enabling technique for the overlapped-clustering spectrum hole geolocation solution proposed in this study. Thus, this solution significantly diverges from previous studies, offering a novel perspective to enhance spectrum sharing capabilities in wireless communication systems.

C. CONTRIBUTION AND ORGANIZATION OF THE ARTICLE

From above, it can be concluded that the sole utilization of spectrum sensing, either in its NCSS or CSS form, is not enough to attain, simultaneously, high detection performance and accurate SHG. In the case of NCSS, performance is penalized whereas SHG may be accurate. In CSS, performance can be improved with respect to NCSS,

[1]ISAC enables the simultaneous use of wireless systems for both high-data-rate communication and environmental awareness tasks such as detecting, localizing and tracking objects, or even creating high-resolution maps of the environment. Thus, spectrum sensing can be included under the broader umbrella of ISAC principles, but it is not the primary focus of ISAC.

but SHG becomes inaccurate. Thus, the database-driven IoT-enabled DSA technique described in [15] arises as a promising supporting architecture to solve the SHG problem.

This article proposes an SHG strategy for CSS, in which the spectrum sensing task is performed by a support IoT network, possibly having fixed and mobile nodes, whose SSIoT devices are clustered, and CSS is performed in the cluster level. The details of this technique are given in Section III, but, in summary, the SSIoT devices are classified into overlapped clusters defined around the crossing points of a regular rectangular grid in the area of interest. A single cluster can be activated at a time for spectrum sensing, with sequential clusters' activation allowing for sequentially scanning the coverage area for constructing the spectrum availability map in the spatial domain. Alternatively, all clusters can be activated at the same time to speedup the construction of the map, at the cost of a large control overhead peak.

The map is built from the spectrum sensing decisions linked with each cluster around each of the grid crossing points. The SHG is given by the coordinates of the crossing points where the decision is in favor of a vacant band. The resolution of the map is governed primarily by the thickness of the grid, which is a design variable of the technique. On the other hand, the performance of spectrum sensing is mainly determined by the cluster size, which is also a design variable. By controlling these variables, a trade-off is established between CSS performance and accuracy of the SHG information.

In summary, this work proposes a novel overlapped-clustering strategy integrated with a support network of IoT-based spectrum sensors. The proposed framework employs data-fusion and decision-fusion cooperative spectrum sensing whose results feed a database of spectrum occupancy information in real-time, aiming at identifying propagation-induced and non-propagation-induced spectrum holes throughout the primary network coverage area for the purpose of driving dynamic spectrum access.

The remaining sections of this article are organized as follows. Section II briefly describes the database-driven IoT-enabled DSA framework. Section III is devoted to the overlapped-clustering technique. The sensing channel, signal, noise, and spectrum sensing models are described in Section IV. Numerical results are presented and discussed in Section V. The conclusions and opportunities for further research are addressed in Section VI.

II. DATABASE-DRIVEN IOT-ENABLED DSA

A. OVERVIEW

The well-known conventional CSS approach is the one that applies distributed detection with centralized decision (see Section I). In this approach, it is noteworthy that the SU terminals need to be equipped with spectrum sensing capability, which increases their complexity and energy consumption, possibly increasing either their cost or physical dimensions, and perhaps reducing portability. Moreover, all control tasks related with spectrum sensing must be implemented and managed by the secondary network.

The database-driven DSA solution proposed in [15] is illustrated in Fig. 1 in a simplified way. The spectrum sensing task is performed by SSIoT devices instead of SUs. An SSIoT is simply a device formed by connecting an ordinary IoT node to a spectrum sensing (SS) module, through a standard wired or wireless interface [9], [15]. Notice in Fig. 1 that not all IoT devices are SSIoTs.

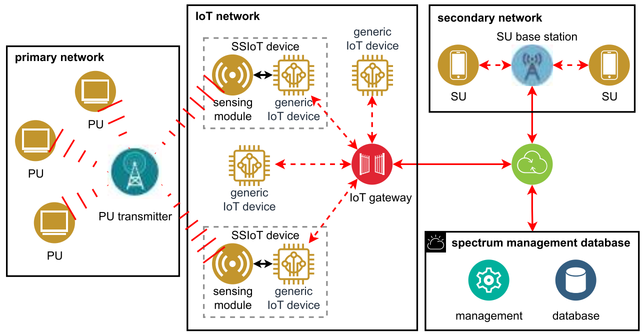


FIGURE 1. Architecture of the database-driven Internet of things-enabled dynamic spectrum access framework.

The IoT device and the SS module are equipped with their own antennas, distinguishing themselves through practical characteristics, primarily their bandwidth and central operating frequency. For instance, the IoT network may operate in a specific frequency range, while spectrum sensing is conducted within the secondary network's operating frequency range.

The SSIoT devices are responsible for scanning the RF spectrum and relaying the acquired sensing information to an IoT gateway. The gateway gathers IoT-related data from a group of closely positioned IoT devices, also acting as an aggregator that can offload specific responsibilities from the nodes, particularly those related with the security of both IoT and SSIoT devices. This approach simplifies the IoT devices and helps protecting against malicious attacks.

The gateway establishes communication with the IoT network, which may or may not be part of the Internet. Whenever requested, the spectrum occupancy database accesses this network to refresh the pool of accessible channels for DSA purposes. Additionally, the database has the capability of analyzing ongoing and historical activities within the primary network. This enables the provision of spectrum usage predictions and other pertinent information, enhancing the effectiveness of the quest for unoccupied frequency bands. The database management function is responsible for managing all tasks related to the spectrum market.

The acquisition of spectrum usage information is performed either on-demand or through continuous querying of the database, which may occur during the regular control communication between SU terminals and their base station.

The database-driven IoT-enabled DSA solution just described takes advantage of the high density and large coverage area of typical IoT networks, allowing for the construction of a fine-grid spectrum occupation database that can be updated in an approximate real-time fashion.

B. FIXED VERSUS MOBILE SSIOT NODES

An IoT network composed exclusively of fixed SSIoT devices can be employed in the current SHG strategy. However, incorporating mobile nodes can enhance the accuracy of the SHG process by refining the information on SHGs through temporal variations in the spatial distribution of the SSIoT devices. Furthermore, the mobility of SSIoTs offers adaptability to the dynamics of the primary network, particularly in scenarios involving mobile PU transmitters. Examples include temporary networks for emergency public safety communications, military tactical communications, temporary broadcast facilities, and wireless microphone transmissions during events. Examples of fixed IoT nodes are smart home appliances, industrial sensors, and infrastructure monitoring systems. Mobile IoT nodes include devices like connected vehicles, wearable technology, and mobile health monitors.

As of 2023, fixed IoT nodes constitute the majority of IoT deployments. This dominance is attributed to extensive applications in sectors like smart cities, industrial automation, and home automation, where stationary devices takes place. For instance, the consumer segment, which includes many fixed devices, accounted for approximately 60% of all IoT devices in 2023 [21].

Looking ahead, the prevalence of fixed IoT nodes is expected to persist. Projections indicate that the number of IoT devices will nearly double from 15.9 billion in 2023 to over 32.1 billion by 2030, with the consumer segment maintaining its significant share [21]. However, mobile IoT nodes are anticipated to experience substantial growth, driven by advancements in technologies such as 5G/6G and edge computing. These developments facilitate applications in autonomous vehicles, logistics, and mobile health monitoring. For example, the integration of 5G is expected to enhance the capabilities of mobile IoT devices, enabling faster data transfer and more reliable connections [22].

Thus, while fixed IoT nodes currently dominate the landscape and are projected to remain prevalent, the rapid expansion of mobile IoT applications suggests a trend toward a more balanced distribution between fixed and mobile nodes in the near future.

III. OVERLAPPED-CLUSTERING

Clustering is a well-known strategy in the context of cluster-based spectrum sensing [23], where the SUs are grouped to form clusters, and spectrum sensing is performed by all cluster members, which send the sensing information to a cluster head (CH) where the cluster-level decision is made. The CHs' decisions associated with all clusters are then forwarded to the secondary network FC, where the final decision upon the occupation state of the sensed band is arrived. In this approach, it is noteworthy that a single decision is made for the entire coverage area, meaning that SHG information is meaningless.

Fig. 2 shows a hypothetical square coverage area of a secondary network; the half-lengths of the sides are normalized to 1 for simplicity. A number of $N = 100$ SUs pertaining to the secondary network are shown as circular dots, which are uniformly distributed over the area. In this example, the SUs are grouped according to the k-means clustering algorithm [24], yielding $c = 9$ non-overlapped clusters with possibly different numbers of members. For a clear visualization, the clusters are bounded by their corresponding convex hulls.[2] This figure is intended to show the result of a typical clustering algorithm adopted in the context of cluster-based spectrum sensing.

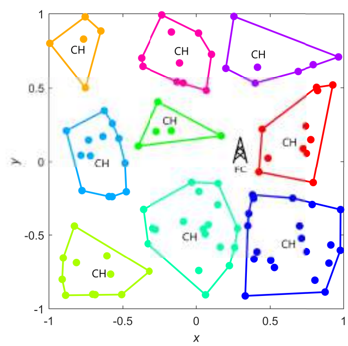


FIGURE 2. Conventional cluster-based CSS with k-means clustering applied to 100 SUs, yielding 9 disjoint clusters. Notably, the clusters may have different numbers of members.

In opposition to the common clustering strategy previously described, clustering is adopted here to create smaller sensing areas in order to allow for detecting possibly different channel occupation states throughout the coverage area. The database-driven IoT-enabled DSA solution explained in the previous section is adopted as the driving architecture.

The same normalized network coverage area and node positions depicted in Fig. 2 are once again shown in Fig. 3. However, the $N = 100$ nodes are now SSIoT devices pertaining to the supporting IoT network, instead of SUs, and the overlapped-clustering approach is being used to form the clusters, instead of the k-means algorithm. A regular square grid is shown in dashed lines, in this case having $G^2 = 9$ crossing points, where G is the number of horizontal or vertical grid lines, which is hereafter referred to as the grid tightening parameter. Each of the $c = G^2 = 9$ clusters is formed by the $m = 20$ SSIoT nodes closest to the corresponding grid crossing point. Due to intersections among clusters, it may happen that some SSIoT devices are not part of any cluster. These devices are more likely to be located near the borders of the coverage area, and their

[2] A convex hull is the smallest convex set that contains a given set of points in a Euclidean space. In simpler terms, it is the smallest convex polygon that encloses all the given points without any indentations.

number tends to increase as the proportion of SSIoT nodes with respect to the number of clusters is increased.

As briefly described in Section I-C, spectrum sensing is made in a cluster level basis, meaning that a decision upon the occupation state of the sensed band is made by each cluster, independently of the others clusters' decisions. Specifically, the m SSIoTs belonging to a given cluster collect samples of the received PU signal in the band of interest. The signal samples acquired by the m SSIoTs in the case of data fusion CSS, or the SSIoTs' decisions in the case of decision fusion CSS are then sent to the FC (notice that there is no cluster head), through the IoT network. This FC can be located in any convenient place, as defined by the IoT network administrator. The FC subsequently makes a decision lined with each cluster, and stores the corresponding channel occupation state in the database, along with frequency range and geolocation information of the associated grid crossing point, which is an information known in advance.

The above-described process is made for all clusters, sequentially or simultaneously. The choice of sequential or simultaneous sensing depends on the trade-off between the needs in terms of spectrum sensing updating speed, and control traffic in the IoT network. At the end of a round of sensing events throughout the entire coverage area, a spectrum availability map in the spatial domain can be constructed. Thus, an estimated SHG is given by the coordinates of the crossing point where the decision is in favor of a vacant band.

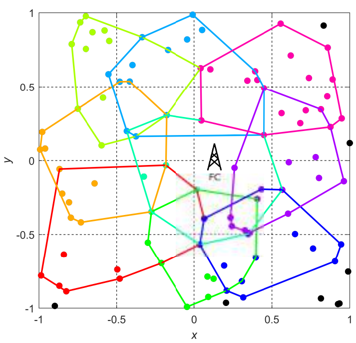


FIGURE 3. **Overlapped-clustering applied to $N = 100$ SSIoT nodes, yielding $c = G^2 = 9$ clusters with intersections. Each cluster has $m = 20$ members. Black dots are nodes that do not belong to any cluster. This figure is better viewed in color.**

Based on Fig. 3, one can infer that the resolution of the spectrum availability map is governed primarily by the grid tightening parameter, G. Since the number of crossing points is G^2, the map resolution increases as G increases. For a given value of G, larger clusters are formed if the number m of SSIoTs in cooperation is increased, which also increases the overlap among them. Nonetheless, it is known from the spectrum sensing literature that a larger number of nodes in cooperation improves the spectrum sensing performance, but in diminishing-return fashion, that is, the addition of more SSIoTs in cooperation yields progressively smaller increments in the spectrum sensing performance.

Thus, by controlling G and m, a trade-off is established between CSS performance and accuracy of the SHG information, as demonstrated in Section V.

In terms of scalability, one can see Fig. 4 as part of a wider or denser network of spectrum sensors. For example, a network having $N = 400$ SSIoT nodes can be thought as the concatenation of four figures similar to Fig. 3, one occupying each quadrant of a square. In this case, the new grid tightening parameter is $G_4 = 2G_1 + 1 = 7$, where the subscript 1 refers to the original basis network (i.e., with $G_1 = 3$) and the subscript 4 refers to the new network with 4 times more nodes (i.e., $N_4 = 4N_1$) and 4 times larger area, or with node density 4 times higher. In general, if $N_s = sN_1$, then

$$G_s = \sqrt{s}\,(G_1 + 1) - 1, \tag{1}$$

where the scaling factor s is a perfect square number, i.e. it is an integer that can be expressed as the square of another integer. Note in Fig. 4 that the number $m = 20$ of cluster members has been maintained with respect to the value considered in Fig. 3.

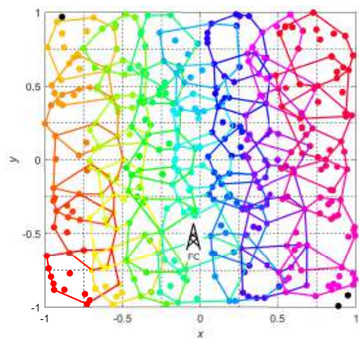


FIGURE 4. **Overlapped-clustering applied to $N = 400$ SSIoT nodes, yielding $c = G^2 = 49$ clusters with intersections. Each cluster has $m = 20$ members. Black dots are nodes that do not belong to any cluster. This figure is better viewed in color.**

The overlapped-clustering strategy can give rise to a potential problem whose consequences must be accounted for: inter-cluster decisions become more correlated as the cluster overlapping area increases. Nonetheless, it is important to note that the phenomenon of signal shadowing also produces spatial correlation, thereby hiding the potential negative impact of the correlation among inter-cluster decisions. Furthermore, it is well-known that such shadowing-induced spatial correlation can affect the performance of any spectrum sensing strategy, unless the spacing between spectrum sensors is sufficiently large to lower this correlation to a negligible value. The extent of this problem is also analyzed in Section V.

IV. SYSTEM MODELS

This section addresses the models for the transmitted and received signals, the noise, the sensing channel, and the detection technique. The models have been chosen to be realistic enough to confer reliability to the results and discussions presented in Section V.

Notably, the most realistic models used in the context of spectrum sensing are those considered in [23]. Herein, such models are adapted to the overlapped-clustering strategy, with an important increment in regard to the sensing channel: besides distance-dependent received signal levels, non-uniform noise power levels, and multi-path fading with environment-dependent random Rice factor, a log-normally distributed spatially correlated shadowing term is incorporated into the received signal modeling.

A. SIGNAL, NOISE, AND CHANNEL MODELS

The CSS employs m SSIoTs in cooperation in each cluster. Each of these SSIoTs is responsible for collecting n samples of the PU signal during a given sensing interval. The samples gathered by the SSIoTs belonging to the jth cluster, for $j = 1, \dots, c$, are grouped at the FC to form the matrix $\mathbf{Y}_j \in \mathbb{C}^{m\times n}$, which is given by

$$\mathbf{Y}_j = \mathbf{h}_j\mathbf{x}^{\mathrm{T}} + \mathbf{V}_j, \tag{2}$$

where $\mathbf{x} \in \mathbb{C}^{n\times 1}$ is the vector containing the samples of the transmitted PU signal. These samples are modeled as zero-mean complex Gaussian random variables whose variance is governed by the average SNR across all SSIoTs.

The sensing channel in (2) is modeled by the vector $\mathbf{h}_j \in \mathbb{C}^{m\times 1}$, whose elements $h_{i,j}$, $i = 1, \dots, m$, represent the channel gains between the PU transmitter and the ith SSIoT of the jth cluster. These gains are assumed to be constant during the sensing interval and independent and identically distributed (i.i.d.) across subsequent sensing events. The jth channel vector in (2) is defined according to

$$\mathbf{h}_j = \mathbf{G}_j\mathbf{a}_j, \tag{3}$$

where $\mathbf{a}_j \in \mathbb{C}^{m\times 1}$ is a vector of complex-Gaussian random variables $a_{i,j} \sim \mathcal{CN}[\sqrt{K_{i,j}/(2K_{i,j}+2)}, 1/(K_{i,j}+1)]$, and where $K_{i,j} = 10^{K_{i,j}^{\mathrm{dB}}/10}$, with $K_{i,j}^{\mathrm{dB}} = 10\log_{10}(K_{i,j})$ being the Rice factor, in dB, of the channel between the PU transmitter and the ith SSIoT of the jth cluster.

In [25], it has been found that $K_{i,j}$ can be modeled by a Gaussian random variable with mean μ_K and standard deviation σ_K, both in dB, determined according to the environment. It is considered an urban area herein, for which $\mu_K = 1.88$ dB and $\sigma_K = 4.13$ dB [25].

Non-uniform received signal power levels across the SSIoTs are assumed, owed to different distances between the PU transmitter and the SSIoTs, and due to signal shadowing. This condition is modeled by the diagonal gain matrix $\mathbf{G}_j \in \mathbb{R}^{m\times m}$ in (3), which is given by

$$\mathbf{G}_j = \mathrm{d}iag\left(\sqrt{\frac{\mathbf{p}_j}{P_{\mathrm{t}}}}\right), \tag{4}$$

where $\mathbf{p}_j = [P_{\mathrm{r}1,j}\, P_{\mathrm{r}2,j} \dots P_{\mathrm{r}m,j}]^{\mathrm{T}}$ is the vector containing the received PU signal powers across the m SSIoTs of the jth cluster, P_{t} is the transmitted PU signal power, $\mathrm{d}iag(\cdot)$ returns the diagonal matrix whose diagonal is the vector in the argument, and $[\cdot]^{\mathrm{T}}$ denotes matrix or vector transposition.

Following the log-distance path loss model described in [26], the local-mean signal power, in dBm, received by the ith SSIoT of the jth cluster is given by

$$P_{\mathrm{r}_{i,j}}^{\mathrm{dBm}} = 10\log_{10}\left[10^3 P_{\mathrm{t}}\left(\frac{d_0}{d_{i,j}}\right)^{\eta}\right] + S_{i,j}, \tag{5}$$

where d_0 is a reference distance in the far-field region of the PU transmit antenna, $d_{i,j}$ is the distance from the PU transmitter to the ith SSIoT of the jth cluster, η is the dimensionless, environment-dependent path-loss exponent, and $S_{i,j}$ models the log-normal signal shadowing [27] component affecting the ith SU of the jth cluster, being a zero-mean Gaussian random variable, in dB, whose standard deviation is σ_{s} dB. Obviously, the power given in (5), in watts, becomes

$$P_{\mathrm{r}i,j} = 10^{-3}10^{P_{\mathrm{r}_{i,j}}^{\mathrm{dBm}}/10}. \tag{6}$$

A specific value of $S_{i,j}$ in (5) can be viewed as a realization of a location-dependent random variable that corresponds to an element of a Gaussian random matrix $\mathbf{S}_{\mathrm{c}}$. Given the secondary network coverage area, the realizations of $S_{i,j}$ throughout the area form the whole matrix $\mathbf{S}_{\mathrm{c}}$, as described in the sequel.

The row-wise and column-wise correlations between the elements of $\mathbf{S}_{\mathrm{c}}$ are governed by a covariance matrix Σ that is formed according to the well-known negative-exponential correlation model [28]. This model is characterized by the correlation function $\rho(\delta) = \exp(-\delta/\lambda)$, where δ is the distance between two points and λ is the correlation length (also often referred to as the correlation distance). The correlation length is the distance for which the shadowing correlation coefficient falls to $1/e$. Then, the elements of Σ are $\Sigma_{z,u} = \exp(-\delta_{z,k}/\lambda)$, where $\delta_{z,k}$ is the distance between the points defined by the row (z) and column (k) indexes of Σ. Hence, λ is a distance between points in the square coverage area with sides of length equal to the number of rows or columns of Σ.

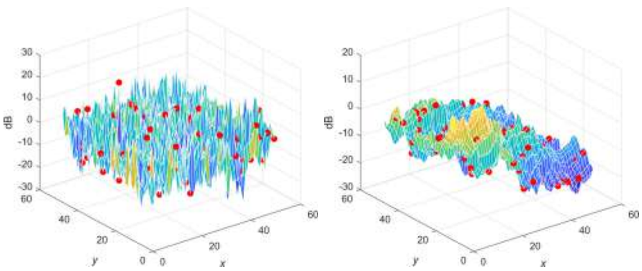
Given a square matrix of uncorrelated normal random variables, $\mathbf{S}_{\mathrm{u}} \in \mathbb{R}^{u\times u}$, the Gaussian matrix $\mathbf{S}_{\mathrm{c}} \in \mathbb{R}^{u\times u}$ with correlated values associated with the set of spatially-correlated points is formed according to the transformation

$$\mathbf{S}_{\mathrm{c}} = \sigma_{\mathrm{s}}\mathbf{L}\mathbf{S}_{\mathrm{u}}\mathbf{L}^{\mathrm{T}}, \tag{7}$$

where $\mathbf{L}$ is the lower-triangular matrix retrieved from the Cholesky decomposition of Σ.

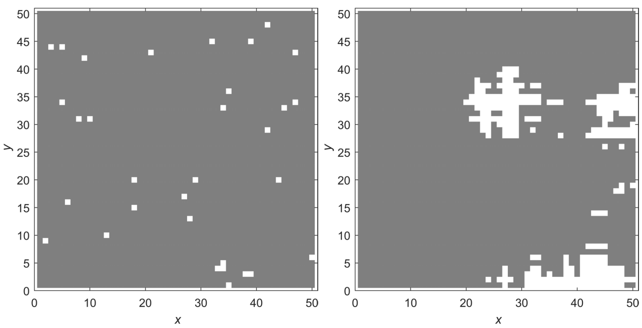
Fig. 5 shows the surface plot of $\mathbf{S}_{\mathrm{c}} \in \mathbb{R}^{50\times 50}$ in two situations in terms of the spatial correlation. Fig. 5a adopts $\lambda = 1$, meaning very small correlation between neighbor values, whereas Fig. 5b considers $\lambda = 40$, which corresponds to a high correlation. In both cases, $\sigma_{\mathrm{s}} = 7$ dB. This figure also shows 100 dots that represent possible positions of SSIoT nodes in the SU coverage area. The shadowing affecting these nodes is associated with the dots' vertical displacements.

Complementing Fig. 5, in Fig. 6 are shown the shadowed areas, in white, that could lead to spectrum holes. In these areas, the shadowing is less than or equal to -15 dB, this value being arbitrarily chosen just for illustration purpose. Fig. 6a corresponds to Fig. 5a, whereas Fig. 6b is associated with Fig. 5b. Fig. 6 clearly demonstrates that a spatially-correlated shadowing model is by far more realistic than an uncorrelated one. The realistic aspect of this model arises from the fact that obstacles located in-between a wireless transmitter and a receiver create valleys in the local-mean received signal power across an area, not point-wise as in Fig. 6a.



(a) (b)

FIGURE 5. Pictorial view of S_c for $\sigma_S = 7$ dB. Fig. 5a is for $\lambda = 1$; Fig. 5b considers $\lambda = 40$. The shadowing values affecting 100 SSIoTs are also shown as red dots. This figure is better viewed in color.



(a) (b)

FIGURE 6. Pictorial view of shadowed areas. Fig. 6a corresponds to Fig. 5a (uncorrelated shadowing), and Fig. 6b refers to Fig. 5b (correlated shadowing).

Returning to the model described in (2), it is well-established that the levels of thermal noise at the receivers' front-ends are typically unequal in practice. This disparity arises primarily from uncalibrated circuitry and nonuniform ambient temperature. Additionally, the non-uniform noise may exhibit time-varying characteristics, and undesired interfering signals entering the receivers can mimic changes in noise levels. To account for this non-uniformity, the elements in the ith row of $\mathbf{V}_j \in \mathbb{C}^{m\times n}$ are modeled as independent Gaussian random variables with zero mean and variance

$$\sigma_{i,j}^2 = (1+\zeta u)\bar{\sigma}^2, \tag{8}$$

where u is a realization of the random variable U that is uniformly distributed in $[-1, 1]$, $\bar{\sigma}^2$ is the noise variance averaged across the SSIoTs, and $0 \le \zeta < 1$ is the fractional variation of the noise power about the average $\bar{\sigma}^2$.

Let SNR be the signal-to-noise ratio averaged across all SSIoTs of the supporting IoT network, that is,

$$\text{SNR} = \mathbb{E}\left[\frac{P_{\text{r}i,j}}{\sigma_{i,j}^2}\right] = \mathbb{E}\left[P_{\text{r}i,j}\right]\mathbb{E}\left[\frac{1}{\sigma_{i,j}^2}\right], \tag{9}$$

where in the right-hand side it has been used the fact that the received signal power can be considered independent of the noise. Moreover, from the separation of the mechanisms that govern the distance-dependent area mean power and the local-mean shadowing, it follows that, in light of (5) and (6), $\mathbb{E}\left[P_{\text{r}i,j}\right]$ can be written as

$$\mathbb{E}\left[P_{\text{r}i,j}\right] = \mathbb{E}\left[P(d)\right]\mathbb{E}\left[10^{S/10}\right], \tag{10}$$

where $P(d) = P_{\text{t}}(d_0/d_{i,j})^{\eta}$ replaces the area-mean received power at a distance $d_{i,j} = d$ from the PU transmitter to an SSIoT, and S replaces $S_{i,j}$ to account for the same distribution of the shadowing over i and j. Furthermore, stemming from the fact that the distribution of the noise is the same over i and j, these indexes can be dropped from $\sigma_{i,j}^2$, and the SNR can be compactly written as

$$\text{SNR} = \mathbb{E}\left[P(d)\right]\mathbb{E}\left[10^{S/10}\right]\mathbb{E}\left[\frac{1}{\sigma^2}\right]. \tag{11}$$

The three expectations in (11) are derived in **Appendix A, B** and **C**, respectively yielding

$$\mathbb{E}[P(d)] = \frac{P_{\text{t}}}{L^2 d_0^{\eta}}\int_0^L\int_0^L\left[(x-x_{\text{t}})^2+(y-y_{\text{t}})^2\right]^{-\frac{\eta}{2}}dx\,dy, \tag{12}$$

$$\mathbb{E}[10^{S/10}] = \exp\left[\frac{\sigma^2\ln^2(10)}{200}\right], \tag{13}$$

and

$$\mathbb{E}\left[\frac{1}{\sigma^2}\right] = \begin{cases}\dfrac{1}{2\bar{\sigma}^2\zeta}\ln\left(\dfrac{1+\zeta}{1-\zeta}\right), & \text{for } 0<\zeta<1\\ \dfrac{1}{\bar{\sigma}^2}, & \text{for } \zeta = 0\end{cases} \tag{14}$$

From these expressions, $\bar{\sigma}^2$ is retrieved for a given SNR according to (11) and the given expectations. Then, $\bar{\sigma}^2$ is plugged into (8), along with the corresponding realizations of the random variable U, resulting in the variance $\sigma_{i,j}^2$ of the noise samples in the ith row of the matrix $\mathbf{V}_j$. In terms of computer simulations, new elements of the set $\{\sigma_{i,j}^2\}$ are computed for each sensing event, conferring the time-varying character to the noise power.

Returning to the spectrum sensing process, the matrix $\mathbf{Y}_j$ given in (2) is formed at the FC. Under the hypothesis $\mathcal{H}_0$ that there is a spectrum hole, then $\mathbf{Y}_j = \mathbf{V}_j$. Under $\mathcal{H}_1$, it follows that $\mathbf{Y}_j = \mathbf{h}_j\mathbf{x}^{\mathrm{T}} + \mathbf{V}_j$. The sample covariance matrix $\mathbf{R}_j$ of order $m \times m$ associated with the jth cluster is subsequently formed at the FC, yielding

$$\mathbf{R}_j = \frac{1}{n}\mathbf{Y}_j\mathbf{Y}_j^{\dagger}, \tag{15}$$

where † denotes complex conjugate and transposition. This matrix is subsequently processed to form the test statistic of any covariance-based detector.

B. DATA FUSION SPECTRUM SENSING MODEL

In each of the clusters throughout the SU coverage area, a distributed-detection centralized data fusion CSS is adopted to identify SHGs associated with the corresponding grid crossing points (see Section III). The Pietra-Ricci index detector (PRIDe) has been chosen as the test statistic, owed to its appeal in terms of state-of-the-art metrics in regard to performance, complexity and latency [29], [30], [31].

Given $\mathbf{R}_j$, referring to the CSS in the jth cluster, the PRIDe test statistic is computed at the FC as follows. Define $r_{j,z,k}$ as the element in the zth row and kth column of $\mathbf{R}_j$, for $j = 1, \ldots, c$ and $z, k = 1, \ldots, m$. The average of $r_{j,z,k}$ is

$$\bar{r}_j = \frac{1}{m^2}\sum_{z=1}^{m}\sum_{k=1}^{m} r_{j,z,k}. \tag{16}$$

The PRIDe test statistic for the jth cluster is defined as

$$T_{\mathrm{PRIDe}_j} = \frac{\sum_{z=1}^{m}\sum_{k=1}^{m}\left|r_{j,z,k}\right|}{\sum_{z=1}^{m}\sum_{k=1}^{m}\left|r_{j,z,k} - \bar{r}_j\right|}. \tag{17}$$

The determination of the occupancy state of the band sensed by the jth cluster is based on a comparison between $T_{\mathrm{PRIDe}j}$ and a decision threshold γ_j, predefined according to the target probability of false alarm, Pfa. If $T_{\mathrm{PRIDe}_j} > \gamma_j$, the jth cluster declares the presence of a primary signal. Otherwise, the band is assumed to be vacant, and its geolocation is determined based on the coordinates of the grid crossing points associated with the clusters that declared it vacant.

C. COMPUTATIONAL COMPLEXITY

The computational complexity of the PRIDe detector used for cluster-level spectrum occupancy decisions is dominated by the cost of calculating the sample covariance matrix of order $m \times m$ from n samples per sensor, which is $\mathcal{O}(nm^2)$ [29], [32]. Since the scalability of the overlapped clustering approach grows linearly with the number of clusters, c, the overall computational cost of the overlapped clustering method with mPRIDe detection in each cluster is $\mathcal{O}(cnm^2)$. Hence, although the complexity growth rate is mainly dictated by the number of SSIoTs per cluster, the overall computational cost can become a limiting factor in networks with a very large number of clusters.

V. NUMERICAL RESULTS

The network model utilized in this work is depicted in Fig. 7. In this model, multiple SSIoTs are randomly distributed across an area equal to or larger than the coverage region of the secondary network. Both the SSIoTs and the secondary network are situated outside the exclusion zone of the primary network. This arrangement exemplifies a typical spectrum-sharing scenario, carefully designed to prevent harmful interference with the primary network [33]. The existence of an exclusion zone is mimicked in the system model described in Section IV by means of a PU transmitter location distant some amount (a multiple of side length L of the SSIoT coverage) from the primary network coverage area.

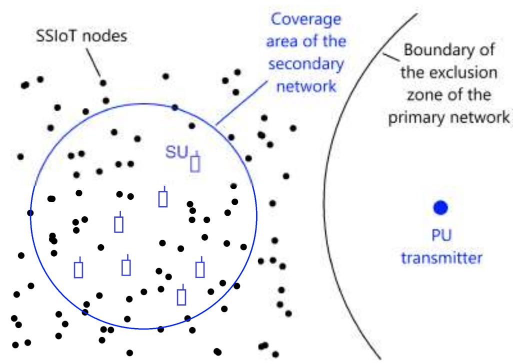


FIGURE 7. Network model.

The identification of a signal level below which a spectrum hole is declared depends on several factors, including regulatory requirements, specific communication system characteristics, and operational constraints. In this work, a minimum SNR threshold of $\mathrm{SNR}_{\mathrm{th}} = -15$ dB is adopted. A spectrum hole exists if the primary signal level falls below this threshold. This value reflects the minimum acceptable quality of service (QoS) required for primary communication systems operating within the spectrum band. For instance, in wireless communication systems, a specific minimum $\mathrm{SNR}_{\mathrm{th}}$ is necessary to ensure reliable data transmission with acceptable error rates. Consequently, frequency bands with signal levels below this threshold are deemed suitable for secondary usage as spectrum holes.

A spectrum hole can arise in two distinct but complementary scenarios: (i) when the PU transmitter is declared inactive, or (ii) when the PU transmitter is deemed active, but shadowed regions exist within the primary network's coverage area where SUs can operate. To align with these scenarios, the performance analysis in this section follows a structured procedure: first, the PU transmitter's state is determined using a global hard-decision rule at the FC, which aggregates local spectrum sensing decisions from all clusters. The FC's decisions are then used to estimate the global probabilities of false alarm (P_{fa}) and detection (P_{d}). If the PU transmitter is declared inactive, the entire coverage area is assumed to be a spectrum hole. Conversely, if the active state is declared, the estimated spectrum occupation map is generated based on the decisions made by individual clusters.

Spectrum holes are represented as square regions of side length L, centered at grid intersections (see Fig. 3). These squares correspond to clusters that have reported undetected PU signals. The square's side length matches the grid spacing, ensuring seamless connectivity between neighboring squares without overlaps or gaps.

A. PSEUDO-CODE

Pseudo-code 1 served as the basis for constructing the MATLAB codes included in the code pack provided in [34]. This pack comprises three main codes: the first, designed mainly for didactic purposes and preliminary system assessment, enables visual analysis of the actual and the estimated SHGs, the overlapped-clustering process, and the receiver operating characteristic (ROC) curves for both local (cluster-level) and global decisions. A second code, derived from the first, generates histograms of the spectrum hole geolocation detection rate (SHGDR) and provides numerical values for key local and global performance metrics. The third code provides the visualization of the SHGDR as a function of key system parameters. It also includes features for plotting error bars and summarizing statistics related to the SHGDR. Additionally, the pack contains MATLAB functions that support the operations of the main codes.

As outlined in **Pseudo-code 1**, to reflect the coexistence of fixed and mobile IoT nodes described in Section II-B, $100\alpha\%$ of the SSIoT devices are designated as fixed nodes, while $100(1-\alpha)\%$ of the SSIoTs are randomly positioned in each sensing event to simulate mobility.

Step 10 of **Pseudo-code 1** refers to the interpolation **Algorithm 1**, which is designed to take into account the shadowing correlation in the SHG estimation process, ensuring that clusters' decisions are not isolated when surrounded by a majority of opposing decisions. Specifically, if clusters' decisions are represented as 1 for an occupied band and 0 otherwise, the algorithm flips any 0 to 1 if it has τ_{01} or more neighboring 1s, and flips any 1 to 0 if it has τ_{10} or fewer neighboring 1s. To disable such interpolation, the setting $\tau_{01} = 9$ prevents any 0 from being flipped to 1 because no 0 can have at least 9 neighboring 1s, given that the maximum possible neighbors is 8. Similarly, setting $\tau_{10} = -1$ prevents any 1 from being flipped to 0, as having at most -1 neighboring 1s is an impossible condition. The impact of this algorithm on decision patterns is further analyzed in the next subsection.

Unless otherwise explicitly stated, the default system parameter are those listed in Table 1.

B. PRELIMINARY RESULTS

This subsection presents a first set of numerical results to validate the Monte Carlo simulation and the SHG estimation process, and to allow for a batter understanding of the whole system operation.

Pseudo-code 1

1) Define the values of the system parameters (see Table 1).
2) Place αN uniformly-distributed fixed SSIoT nodes on the SSIoT coverage area, $0 \le \alpha \le 1$.
3) **For** each sensing round, **Do**:
   - 2.1) Place $(1-\alpha)N$ uniformly-distributed mobile SSIoT nodes on the coverage area.
   - 2.2) Apply overlapped-clustering to generate G^2 clusters, each having m nodes.
   - 2.3) Compute the distances $d_{i,j}$ from the PU transmitter to the ith SU of the jth cluster, for $i = 1, \dots, m$, $j = 1, \dots, G^2$.
   - 2.4) Define a set of points in the coverage area, according to the dimensions of $\mathbf{S}_\text{c}$, and generate the correlated shadowing matrix for all points, each point linked with a row and column of $\mathbf{S}_\text{c}$.
   - 2.5) From $\mathbf{S}_\text{c}$, extract the shadowing random variable, $S_{i,j}$, according to the coordinates of the N SSIoTs.
   - 2.6) Compute $P_{\text{r}_{i,j}}$ via (5) and (6).
   - 2.7) Compute $\bar{\sigma}^2$ via (14) and related equations, then plug into (8) to yield $\sigma^2_{i,j}$, which is used to generate the noise matrix $\mathbf{V}_j$.
   - 2.8) Using $\mathbf{Y}_j = \mathbf{V}_j$ compute $\mathbf{R}_j$ and T_{PRIDe_j} under $\mathcal{H}_0$, and using $\mathbf{Y}_j = \mathbf{Y}_j = \mathbf{h}_j\mathbf{x}^\text{T} + \mathbf{V}_j$ compute $\mathbf{R}_j$ and T_{PRIDe_j} under $\mathcal{H}_1$. Store T_{PRIDe_j} under $\mathcal{H}_0$ and $\mathcal{H}_1$ for each sensing round.
   - 2.9) Using equations (5) and (6), compute local-mean received signal powers across all points in the coverage area (change i, j to the coordinates of each point). Define the local-mean SNR_lm in each point as the quotient between the local-mean power and the average noise variance $\bar{\sigma}^2$. Determine the actual spectrum holes in every point where $\text{SNR}_\text{lm} \le \text{SNR}_\text{th}$.

   End for (sensing rounds).
4) Estimate the empirical CDFs of T_{PRIDe_j} under $\mathcal{H}_0$ and $\mathcal{H}_1$, from where the clusters' $(P_{\text{fa}_j}, P_{\text{d}_j})$ are respectively calculated.
5) Compute the global $(P_\text{fa}, P_\text{d})$ by plugging the clusters' $(P_{\text{fa}_j}, P_{\text{d}_j})$ into Eqs. (12) and (13) of [23].
6) For debugging and cooperation gain analysis, plot the ROC curves related to the clusters' $(P_{\text{fa}_j}, P_{\text{d}_j})$, for $j = 1, \dots, G^2$, and to the global $(P_\text{fa}, P_\text{d})$.
7) Find the clusters' decision thresholds γ_j for the target local P_fa.
8) Compare T_{PRIDe_j} with γ_j under $\mathcal{H}_1$ (PU transmitter *on*) to obtain clusters' decisions for each sensing round.
9) Compare T_{PRIDe_j} with γ_j under $\mathcal{H}_0$ (PU transmitter *off*) to obtain clusters' false alarm decisions for each sensing round.
10) Apply the k-out-of-c global decision rule and store all clusters' and global decisions.
11) Form the binary decision matrix $\mathbf{D}$ containing the clusters' decisions in every simulation run. Then, expand and interpolate $\mathbf{D}$ according to **Algorithm 1**, yielding the matrices of estimated SHGs. If the PU transmitter is deemed *off*, the entire coverage area is considered free for secondary usage.
12) To assess the SHG estimations, do the logical complement of the exclusive OR (XNOR) operation between the matrix of estimated SHGs and the actual matrix of spectrum holes found in Step 2.9. The metric SHG detection rate (SHGDR) in each simulation run is defined as the average of all elements of the resultant matrix.

Fig. 8a provides a snapshot of the actual SHG, while Fig. 8b and Fig. 8c show the corresponding estimated SHGs. Fig. 8b incorporates **Algorithm 1** for interpolating cluster decisions, using parameters $\tau_{01} = 2$ and $\tau_{10} = 0$. In contrast, Fig. 8c omits this algorithm, setting $\tau_{01} = 9$ and $\tau_{10} = -1$. The strong resemblance between Fig. 8a and Fig. 8b highlights the effectiveness of the SHG estimation process and the interpolation provided by **Algorithm 1**. On the other hand, Fig. 8c reveals noticeable gaps and isolated spectrum holes where they should not exist due to the spatially-correlated shadowing, demonstrating the necessity of **Algorithm 1** in handling such shadowing effect.

Algorithm 1 Expansion and Interpolation of Matrix D

1) **Input**: Let $\mathbf{D}$ be the decision matrix of order $G \times G$, which contains the binary spectrum occupancy decisions made by the $c = G^2$ clusters.
2) Given the matrix $\mathbf{D}$, flip to 1 all zeros that have $\tau_{01} = 2$ or more neighbor ones, and flip to 0 all ones that have $\tau_{01} = 0$ neighbor 1. This step acts like an interpolation, turning isolated clusters' decisions 0 into 1, and turning isolated clusters' decisions 1 into 0.
3) Transform $\mathbf{D}$ into an expanded matrix $\mathbf{D}_\mathrm{e}$ having the same order of the matrix $\mathbf{S}_\mathrm{c}$ of shadowing values, i.e. $u \times u$. Each of the c bits in $\mathbf{D}$ turns into a multiplicity of $u/(G+1)$ equal bits in $\mathbf{D}_\mathrm{e}$, centered at the crossing points of G^2 grid lines. This step creates the square region associated to the spectrum hole detected by each cluster.
4) **Output**: The SHG if formed by the union of the square regions where clusters report an undetected PU signal, after interpolation, under the condition of global decision for an occupied band (PU transmitter deemed *on*).

TABLE 1. Main default system parameters.

Parameter	Value
Total number of SSIoTs in the basis network, N_1	400
Grid tightening parameter, G_1	20
Number of SSIoTs per cluster, m	15
Scale factor of the SSIoT network, s	1
Global decision combining rule	MAJ
Side length of the SSIoT coverage area, L	2 m
Fraction of fixed SSIoT nodes, α	0.5
Number of samples per SSIoT, n	100
Signal-to-noise ratio, SNR	-5 dB
Number of Monte Carlo simulation runs	10000
Path-loss exponent, η	2.5
Reference distance for path-loss calculation, d_0	$0.001L$ m
PU transmit power, P_t	5 W
PU transmitter location, $(x_\mathrm{t}, y_\mathrm{t})$	$(2L, 2L)$ m
Fraction of noise power variations about the mean, ζ	0.5
Mean of the Rice factor, μ_K	1.88 dB
Standard deviation of the Rice factor, σ_K	4.13 dB
Standard deviation of the shadowing, σ_s	7 dB
Threshold SNR for actual SHG, $\mathrm{SNR_{th}}$	-15 dB
Order of the matrices $\boldsymbol{\Sigma}$, $\mathbf{S}_\mathrm{u}$ and $\mathbf{S}_\mathrm{c}$	$u \times u$
Correlation length of the shadowing model, λ	$0.8u$
Reference local P_fa for threshold calculation	0.3
Interpolation thresholds, (τ_{01}, τ_{10})	$(2, 0)$

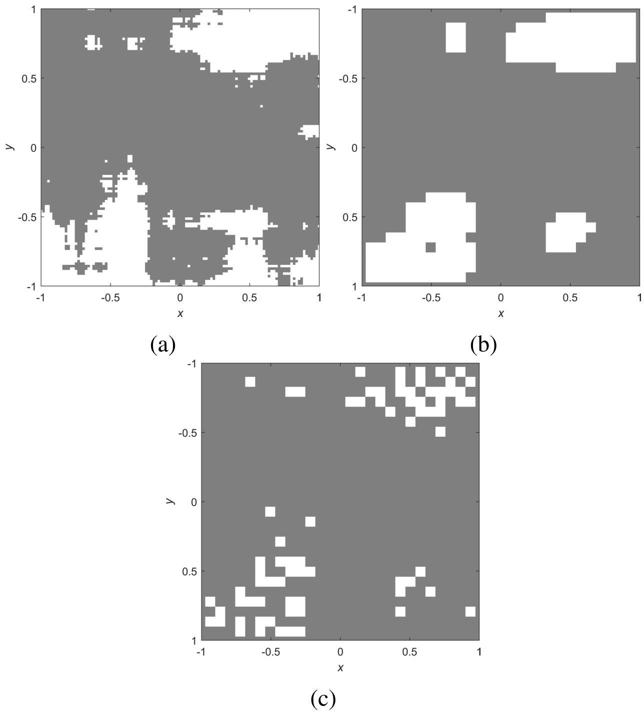


FIGURE 8. A single snapshot of actual SHG (a) and the corresponding estimated SHGs (b, c). The SHG in (b) adopts Step 2 of **Algorithm 1**, whereas (c) does not. The estimated SHGs resulted from 27 × 27 10-member clusters formed from 1600 SSIoTs.

Fig. 8 considers a base network configuration with $N_1 = 100$ SSIoTs, grid tightening parameter $G_1 = 6$, and a network scale factor $s = 16$. This scaling results in a network of $N_{16} = 16N_1 = 1600$ SSIoTs, with $G_{16} = \sqrt{16}(6+1) - 1 = 27$, yielding $c = G_{16}^2 = 729$ overlapping clusters, each comprising $m = 10$ members. The remaining system parameters are specified in Table 1.

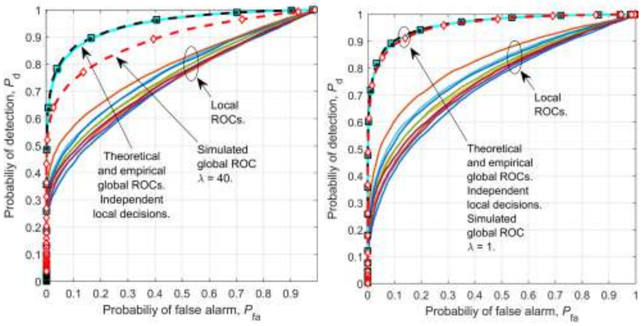
Fig. 9 presents the local (at cluster levels) and global performance in terms of ROC curves for SNR $= -10$ dB. The analysis is based on a network configuration with $N_1 = 100$, $G_1 = 3$ and $s = 1$, resulting in an SSIoT network with $N = N_1 = 100$ SSIoTs and $c = G_1^2 = 9$ overlapping clusters, each containing $m = 10$ members. The remaining system parameters are detailed in Table 1. Local decision thresholds were configured to achieve the local false alarm probability of $P_\mathrm{fa} = 0.2$ for all clusters, leading to local detection probabilities of approximately 0.54, 0.53, 0.59, 0.54, 0.58, 0.65, 0.62, 0.65, and 0.68 for clusters 1 to 9, respectively. The corresponding global metrics were $P_\mathrm{fa} \approx 0.02$ and $P_\mathrm{d} \approx 0.65$.

The primary aim of Fig. 9 is to illustrate the local performances, demonstrate the cooperation gain, and evaluate the global performance both from simulation and theory. The intermediate global ROC curve in Fig. 9a represents the actual global performance, which deviates noticeably from the upper ROC curve. This discrepancy arises from the fact that the theoretical expressions for global probabilities of detection and false alarm from [23] assume uncorrelated decisions under the k-out-of-c combining rule. Thus, incorporating correlated shadowing into the channel model introduces a significant deviation from the theoretical predictions.

Unfortunately, the derivation of theoretical expressions for these probabilities under correlated shadowing is highly challenging, if not impractical. This difficulty arises from the high complexity of modeling the correlation between clusters' decisions within the current network framework: since each hard decision results from multiple sensors that are randomly distributed within a cluster, it becomes intricate to generalize a model that captures the dependencies accurately.

Fig. 9b illustrates the scenario where $\lambda = 1$, representing nearly uncorrelated shadowing. As a result, the clusters' decisions are approximately uncorrelated. In this case, the theoretical and empirical ROC curves derived under the assumption of independent decisions align closely with the simulated ROC, confirming the validity of the theoretical model for uncorrelated scenarios.

Another important aspect to emphasize is the relationship between the local and global probabilities of false alarm and the corresponding probabilities of detection. The local



(a) (b)

FIGURE 9. Local and global performances of the overlapped-clustering strategy.

decision thresholds adopted in Fig. 9 were determined to achieve a local $P_{fa} = 0.2$ for all clusters, which led to a global $P_{fa} \approx 0.02$. This result highlights that the false alarm rate at the cluster level must be sufficiently high to avoid a near-zero global P_{fa}. Unless the global ROC curve is ideal ($P_{fa} = 0$ and $P_d = 1$), a near-zero global P_{fa} will possibly result in a low global P_d, because the operating point will fall away from the knee of the ROC curve, in the left direction.

In practical terms, to ensure that the global P_{fa} and P_d are close to the knee of the ROC curve, where the trade-off between detection and false alarms is typically most favorable, the local P_{fa} must be set significantly higher than the expected global P_{fa}.

C. STATISTICAL ANALYSIS

Before proceeding, it is important to further clarify the definition of the spectrum hole geolocation detection rate. An SHGDR of 0.8, for example, should not be interpreted as an estimate of the probability of detecting a spectrum hole. Rather, it means that 80% of the instances of spectrum hole presence and absence over the whole coverage area have been correctly identified. For further illustration, consider Fig. 8b and Fig. 8c. An SHGDR of 0.8 implies that the gray and white regions in Fig. 8b align with their counterparts in Fig. 8c with 80% accuracy.

That said, Fig. 10 gives the histogram of the SHGDR for 10000 sensing events. The main system parameters are shown in the figure, and the remaining ones are listed in Table 1. The mean SHGDR is ≈ 0.85, with a variance of ≈ 0.005. Notice the local $P_{fa} = 0.3$, yielding near-perfect global decisions.

The remaining figures of this section show the mean, maximum, minimum and the variance of the SHGDR when important system parameters are varied. For the sake of conciseness, the results presented in this section are restricted to the effect of those parameters that are directly related to the overlapped-clustering approach. Nonetheless, the MATLAB code [34] can be easily adapted to allow for the analysis of other parameters as well.

Fig. 11 presents the SHGDR as a function of the SNR. The SHGDR consistently improves as the SNR increases,

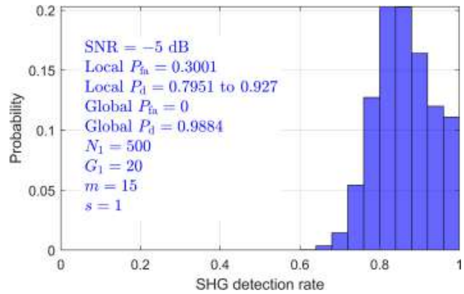


FIGURE 10. Histogram of the spectrum hole geolocation detection rate.

which is an expected outcome. The variance of the SHGDR is notably low across all SNR values, indicating robust SHG detection performance. For SNR values above −10 dB, near-perfect global detection probability ($P_d \approx 1$) and negligible global false alarm probability ($P_{fa} \approx 0$) have been achieved, which suggests that the system is highly effective at distinguishing between the states of the PU transmitter even under challenging noise conditions. These near-perfect global probabilities also bring certainty to a vacant spectrum being caused by shadowed regions, not by the *off* state of the PU transmitter.

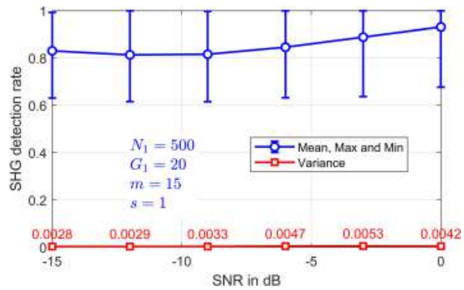


FIGURE 11. Statistics of spectrum hole geolocation detection rate versus SNR.

Fig. 12 shows the SHGDR as a function of the shadowing standard deviation. The SHGDR decreases as σ_s increases, with a particularly pronounced drop at higher values, reflecting the increased uncertainty introduced by significant shadowing effects. Despite this, the system maintains reasonable performance, with low variance, global $P_d > 0.96$ and near-zero P_{fa} for $\sigma_s < 8$ dB. This behavior underscores the robustness of the system in handling moderate shadowing, though extreme shadowing impacts accuracy due to the performance loss of local spectrum sensing.

Fig. 13 illustrates the SHGDR as a function of the total number of SSIoTs. The SHGDR improves slightly and stabilizes as N_1 increases. The initial improvement is achieved until N_1 reaches a value beyond which clustering starts to leave SSIoTs outside any cluster (see Fig. 3). This is because the chance of having SSIoTs not belonging to any cluster increases as N_1 increases and the number of clusters is maintained. The mean SHGDR remains consistently high, with minimal variance, indicating reliable performance

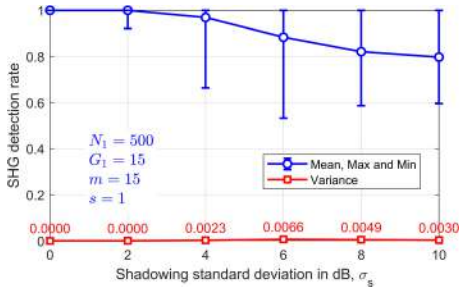


FIGURE 12. Statistics of spectrum hole geolocation detection rate versus standard deviation of the shadowing, σ_s.

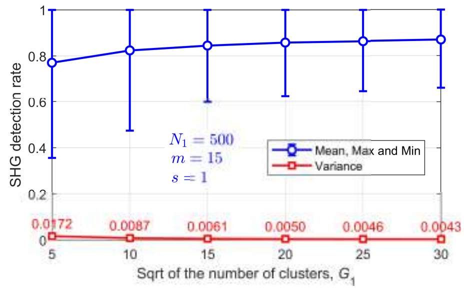


FIGURE 14. Statistics of spectrum hole geolocation detection rate versus the square root of the number of clusters, G_1.

across network sizes. The near-zero global P_{fa} and a global $P_{\text{d}} > 0.98$ for $N_1 > 400$ further affirm the system's effectiveness.

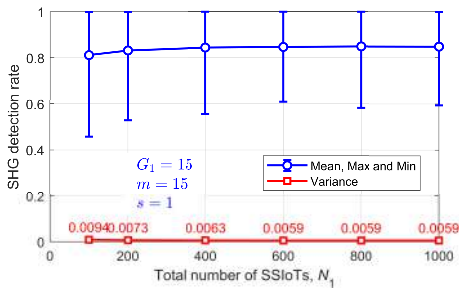


FIGURE 13. Statistics of spectrum hole geolocation detection rate versus the total number of SSIoTs, N_1.

Fig. 14 depicts the SHGDR as a function of the grid tightening parameter, G_1, which is the square root of the number of clusters, c, when $s = 1$. The SHGDR increases as G_1 grows, reaching a plateau at higher values. This trend reflects the benefits of increasing the spatial resolution of clustering, which improves the system's ability to detect spectrum holes accurately. The variance remains low, and the global probabilities indicate that the cooperative decision-making process remains effective. For example, near-zero global P_{fa} and a global $P_{\text{d}} > 0.98$ has been achieved for $G_1 > 15$. The observed stability in the SHGDR suggests that beyond a certain cluster density, further increasing G_1 results in diminishing returns.

Since diminishing returns are attained by increasing the cluster density, one should opt for the smallest number of clusters that yields a desired SHGDR, since the control overhead required to transmit the clusters' decisions to the fusion center grows linearly with the number of clusters.

The influence of the number of SSIoTs per cluster in the SHGDR is demonstrated in Fig. 15. The SHGDR stabilizes after an initial improvement as m increases. This trend reflects the role of intra-cluster cooperation, where larger cluster sizes enhance local decision reliability. The system demonstrates low variance in SHGDR and $P_{\text{d}} \approx 0.99$ around $m = 15$, with negligible false alarms. The results suggest that beyond a certain point, increasing m yields diminishing returns, as the benefits of additional sensors per cluster are limited by the overall system configuration.

Moreover, one must recall from Section IV-C that the computational complexity of the PRIDe detector grows as $\mathcal{O}(nm^2)$, meaning that large numbers of SSIoTs per cluster are not attractive also from the computational complexity perspective. Owed to the fact that computational complexity is often direct proportional to latency, increasing m may delay the computation of the decision on the spectrum occupancy state. The increase of m for a fixed number of clusters also increases the amount of control overhead to send sensing information to the fusion center, which may be undesirable.

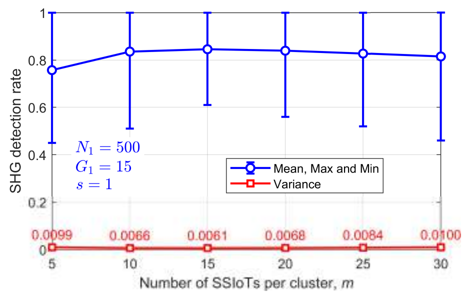


FIGURE 15. Statistics of spectrum hole geolocation detection rate versus the number of SSIoTs per clusters, m.

Fig. 16 shows the SHGDR as a function of the number of samples per SSIoT, n. It can be noticed that when n is around 150, the maximum value of the mean SHGDR is attained. Surprisingly, larger values of n do not bring performance improvements to the mean SHGDR, although the probability of detection in each cluster consistently increases with n, as expected. This confirms that increasing n improves local detection, though not necessarily global geolocation accuracy. The variance of the SHGDR also attains its smaller value around $n = 150$. Zero global P_{fa} and a global $P_{\text{d}} > 0.93$ have been achieved over all values of n.

The influence of the shadowing correlation length, λ, on the SHGDR is shown in Fig. 17. Small values of λ produces small SHGDR, with this metric monotonically

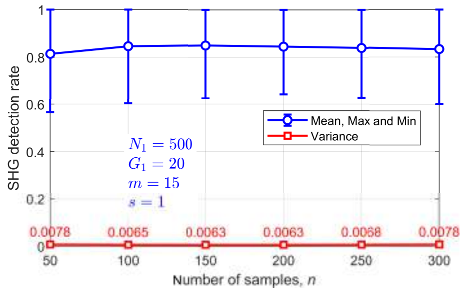


FIGURE 16. Statistics of spectrum hole geolocation detection rate versus the number of samples per SSIoT, n.

increasing as λ increases. The same pattern is observed in regard to the dispersion of the SHGDR.

The behavior observed in this figure can be interpreted in light of Fig. 6, where it can be seen that smaller shadowing spatial correlations produce point-wise spectrum holes, which are intentionally undetectable due to the action of **Algorithm 1**. As λ increases, isolated point-wise spectrum holes tend to become rarer, and **Algorithm 1** starts acting as intended, increasing the SHGDR.

The local probability of detection was around 0.91 for $\lambda = 0.01$ (which is the smallest λ in Fig. 17), going to around 0.96 for the other values of λ. Zero global P_{fa} and a global $P_{\text{d}} > 0.99$ have been achieved over all values of λ.

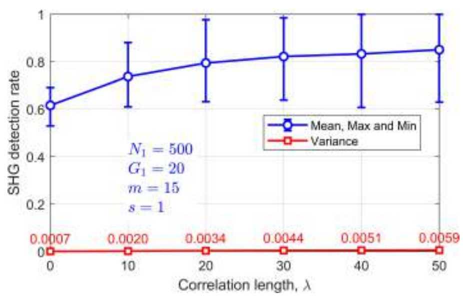


FIGURE 17. Statistics of spectrum hole geolocation detection rate versus the shadowing correlation length, λ.

The results across all figures highlight the robustness and reliability of the proposed system under varying conditions. The low variance in SHGDR in all scenarios underscores consistent detection performance, while the near-zero global false alarm probabilities and high detection probabilities demonstrate the effectiveness of cooperative sensing in determining the state of the PU transmitter with accuracy. However, the system exhibits high sensitivity to the SNR and σ_s.

Lastly, the determination of an optimum set of system parameters is not straightforward from the results presented in this section, mainly because the influence of a given parameter is dependent of the other parameters. Nonetheless, an attractive set of parameters could be: SNR $= -5$ dB, $N = N_1 = 500$, $G = G_1 = 20$ and $m = 15$, which were those adopted to generate Fig. 10.

D. COMPLEMENTARY DISCUSSIONS

1) COMPARISON WITH EXISTING METHODS

To the best of the author's knowledge, only one other spectrum hole geolocation method exists besides the proposed overlapped clustering approach, which is based on propagation (or coverage) prediction. Comparing the present method with a propagation prediction approach would require real or hypothetical digital terrain data (e.g., digital elevation maps) and commercial coverage prediction software. However, this would hinder statistical comparisons unless numerous real scenarios were analyzed, making such a comparison practically unfeasible. Additionally, since coverage prediction methods do not produce real-time results, the comparison would be inherently unfair to the proposed method.

2) THE CHOICE OF CLUSTER-LEVEL DETECTORS

Any other detector suitable for CSS with data fusion can replace the PRIDe without requiring modifications to the overlapped clustering approach. The impact on spectrum hole geolocation performance will depend on how the chosen detector compares to PRIDe in terms of spectrum sensing performance, potentially leading to either improvement or degradation, according to the improvement or degradation of the alternative detector with respect to PRIDe.

3) PRACTICAL FEASIBILITY

The practical feasibility of the overlapped clustering strategy for spectrum hole geolocation assumes that a dense IoT network (or a similar network) is already deployed and that its administrator opts to leverage spectrum trading revenues by adapting the network to accommodate the largest possible number of SSIoTs. This presents the most significant limitation for implementation, primarily due to hardware constraints in converting a standard IoT device into an SSIoT. Some directions towards the implementation of spectrum sensors can be found in [31] and references therein.

Once the SSIoT network is deployed, the challenges associated with implementing the spectrum sensing algorithm become considerably smaller, primarily dictated by the complexity of cluster formation and the fusion and processing of spectrum sensing data. The next major hurdle is the establishment of the spectrum market; potential approaches for achieving this are explored in the seminal work [15].

4) THE CHOICE OF CLUSTERING ALGORITHM

The overlapped clustering strategy was designed with four main objectives: 1) to provide sufficient spatial resolution for accurate spectrum hole geolocation; 2) to achieve low complexity, thus minimizing the computational cost of the clustering process; 3) to facilitate the straightforward definition of cluster locations; and 4) to ensure uniform cluster sizes, enabling the reuse of the same spectrum sensing algorithm across all clusters, which helps reduce complexity. It is important to highlight that adopting alternative clustering methods may not necessarily align with the proposed spectrum hole geolocation approach, potentially increasing complexity and requiring modifications to the overall geolocation strategy to accommodate the clustering method rather than the other way around.

VI. CONCLUSION

This work proposed a novel approach to address the challenge of spectrum hole geolocation in dynamic spectrum access systems by leveraging an overlapped-clustering strategy within a database-driven IoT-enabled framework. The results demonstrated that the integration of cooperative spectrum sensing with overlapped clusters provides robust performance in detecting spectrum holes while maintaining spatial resolution for accurate geolocation. The approach takes advantage of the Pietra-Ricci index detector, whose low complexity and high performance are favorable to the demands of lowering energy consumption and processing burden of the spectrum sensing information coming from a high number of sensors.

The simulation results validated the effectiveness of the proposed solution under varying system parameters, including SNR, shadowing standard deviation, cluster density, and SSIoT network size. The SHG detection rate exhibited high mean values with low variance across all tested scenarios, confirming the robustness and reliability of the proposed technique. The interplay between cluster overlap and shadowing effects was shown to have minimal adverse impact, likely due to the highly-deleterious influence of the shadowing standard deviation. Moreover, the results highlighted the necessity of balancing local and global decision thresholds to optimize detection performance.

While the proposed strategy demonstrated accuracy in regard to the spectrum hole geolocation estimation, opportunities for further research remain. The following initiatives could be explored:

- Assess the impact of interference on primary users when secondary networks rely on spectrum hole geolocation information derived from the proposed technique.
- Perform comparative studies with standalone (non-cooperative) sensors, which could provide additional insights into performance trade-offs, mainly related to the total number of SSIoT nodes needed.
- Explore the use of unequal cluster sizes to address inter-cluster performance disparities.
- Integrate artificial intelligence or machine learning for adaptive parameter tuning, spectrum sensing optimization or adaptive clustering, aiming at enhancing the overall system performance.
- Evaluate the spectrum hole geolocation detection rate as a function of other parameters, such as the flip control parameters τ_{01} and τ_{10} in **Algorithm 1**, the target local P_{fa}, the hard-decision combining parameter k, the path-loss exponent η, and the PU transmitter location $(x_{\text{t}}, y_{\text{t}})$.
- Perform additional statistical analyses, for example the analysis of variance (ANOVA), aiming at assessing the interaction among system parameters.

In conclusion, the proposed overlapped-clustering strategy represents a promising advancement for real-time spectrum hole geolocation for dynamic spectrum access in current and future wireless communication networks, providing a novel framework to enable radio-frequency spectrum sharing.

APPENDIX A
EXPECTED VALUE OF $P(D)$

The received power at a distance d from the transmitter is

$$P(d) = P_{\text{t}}\left(\frac{d}{d_0}\right)^{-\eta}, \tag{18}$$

where P_{t} is the transmit power, d_0 is a reference distance, and η is the path loss exponent.

The distance between the PU transmitter located at $(x_{\text{t}}, y_{\text{t}})$ and an SSIoT located at (x, y) is given by

$$d = \sqrt{(x - x_{\text{t}})^2 + (y - y_{\text{t}})^2}. \tag{19}$$

Since the SSIoTs are uniformly distributed over a square area with side length L, the expected value $\mathbb{E}[P(d)]$ is obtained by integrating $P(d)$ over the square area and normalizing by the area, that is,

$$\mathbb{E}[P(d)] = \frac{1}{L^2}\int_0^L \int_0^L P_{\text{t}}\left(\frac{d}{d_0}\right)^{-\eta} dx\, dy. \tag{20}$$

Substituting the expression for d into the integral, then

$$\mathbb{E}[P(d)] = \frac{P_{\text{t}}}{L^2 d_0^{-\eta}}\int_0^L \int_0^L \left[(x - x_{\text{t}})^2 + (y - y_{\text{t}})^2\right]^{-\frac{\eta}{2}} dx\, dy. \tag{21}$$

The double integral in the given expression, though expressed in closed form, is analytically complex and often requires numerical methods for evaluation. However, numerical problems may arise, particularly when the PU transmitter is located within the coverage area of the SSIoT network. This issue stems from the term $d^{-\eta}$, which increases sharply as $d \rightarrow 0$, especially for small values of η. Such behavior results in large outliers that influence the calculation, yielding instability. To address this numerical issue, a Monte Carlo integration can be adopted [34], although instabilities may still occur for d in the vicinity of zero.

APPENDIX B
EXPECTED VALUE OF $10^{S/10}$

Given that S is a Gaussian random variable with zero mean and standard deviation σ_{s}, its probability density function (PDF) is

$$f_S(s) = \frac{1}{\sqrt{2\pi\sigma_{\text{s}}^2}} \exp\left[-\frac{s^2}{2\sigma_{\text{s}}^2}\right]. \tag{22}$$

The expected value $\mathbb{E}[10^{S/10}]$ can be calculated using the definition of the expectation of a function of a random variable, that is,

$$\mathbb{E}[10^{S/10}] = \int_{-\infty}^{\infty} 10^{s/10} f_S(s)\, ds. \tag{23}$$

Recognizing that $10^{s/10} = e^{s \ln(10)/10}$, and rearranging to form a completing-the-square term in the exponent, then it follows that

$$\mathbb{E}[10^{S/10}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{1}{2\sigma_s^2}\left(s^2 - \frac{s\sigma_s^2 \ln(10)}{5}\right)\right] ds. \tag{24}$$

Since

$$s^2 - \frac{s\sigma_s^2 \ln(10)}{5} = \left(s - \frac{\sigma_s^2 \ln(10)}{10}\right)^2 - \left(\frac{\sigma_s^2 \ln(10)}{10}\right)^2, \tag{25}$$

after some manipulations the integral can be written as

$$\mathbb{E}[10^{S/10}] = e^{\frac{\left(\frac{\sigma_s^2 \ln(10)}{10}\right)^2}{2\sigma_s^2}} \int_{-\infty}^{\infty} \frac{e^{-\frac{\left(s-\frac{\sigma_s^2 \ln(10)}{10}\right)^2}{2\sigma_s^2}}}{\sqrt{2\pi\sigma_s^2}}\, ds. \tag{26}$$

The integral now involves a shifted Gaussian PDF, which integrates to 1. Then, it finally follows that

$$\mathbb{E}[10^{S/10}] = \exp\left[\frac{\sigma^2 \ln^2(10)}{200}\right]. \tag{27}$$

APPENDIX C
EXPECTED VALUE OF $1/\sigma^2$

The expected value of $1/\sigma^2 = 1/[(1 + \zeta U)\bar{\sigma}^2]$ is the expected value of a function of the uniform random variable $U \sim \mathcal{U}[-1, 1]$, which is

$$\mathbb{E}\left[\frac{1}{\sigma^2}\right] = \frac{1}{2\bar{\sigma}^2} \int_{-1}^{1} \frac{1}{1 + \zeta u}\, du. \tag{28}$$

Performing the substitution $t = 1 + \zeta u$, then $dt = \zeta\, du$ and $du = dt/\zeta$. When $u = -1$, $t = 1 - \zeta$, and when $u = 1$, $t = 1 + \zeta$. Thus, it follows that

$$\mathbb{E}\left[\frac{1}{\sigma^2}\right] = \frac{1}{2\bar{\sigma}^2\zeta} \int_{1-\zeta}^{1+\zeta} \frac{1}{t} dt. \tag{29}$$

Since the integral of $1/t$ is $\ln|t|$, then, after some simple manipulations, it is found that

$$\mathbb{E}\left[\frac{1}{\sigma^2}\right] = \frac{1}{2\bar{\sigma}^2\zeta} \ln\left(\frac{1+\zeta}{1-\zeta}\right) \tag{30}$$

and

$$\mathbb{E}\left[\frac{1}{\sigma^2}\right] = \frac{1}{\bar{\sigma}^2} \tag{31}$$

for $0 < \zeta < 1$ and $\zeta = 0$, respectively.

REFERENCES

[1] A. Nasser, H. Al Haj Hassan, J. Abou Chaaya, A. Mansour, and K.-C. Yao, "Spectrum sensing for cognitive radio: Recent advances and future challenge," *Sensors*, vol. 21, no. 7, p. 2408, Mar. 2021, doi: 10.3390/s21072408.

[2] W. Dias, A. Ferreira, R. Kagami, J. S. Ferreira, D. Silva, and L. Mendes, "5G-RANGE: A transceiver for remote areas based on software-defined radio," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2020, pp. 100–104, doi: 10.1109/EuCNC48522.2020.9200925.

[3] F. Communications Commission. (2002). *Spectrum Policy Task Force Report*. [Online]. Available: https://books.google.com.br/books?id=p9MQtwAACAAJ

[4] D. Das and S. Das, "A survey on spectrum occupancy measurement for cognitive radio," *Wireless Pers. Commun.*, vol. 85, no. 4, pp. 2581–2598, Dec. 2015, doi: 10.1007/s11277-015-2921-1.

[5] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999, doi: 10.1109/98.788210.

[6] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, Dec. 2010, Art. no. 381465, doi: 10.1155/2010/381465.

[7] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, no. 1, pp. 40–62, Mar. 2011, doi: 10.1016/j.phycom.2010.12.003.

[8] Y. Arjoune and N. Kaabouch, "A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions," *Sensors*, vol. 19, no. 1, p. 126, Jan. 2019, doi: 10.3390/s19010126.

[9] D. Guimarães, "Spectrum sensing: A tutorial," *J. Commun. Inf. Syst.*, vol. 37, no. 1, pp. 10–29, Feb. 2022, doi: 10.14209/jcis.2022.2.

[10] *IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands*, Jul. 2011. [Online]. Available: https://standards.ieee.org/ieee/802.22/4742/

[11] *IEEE Standard for Information Technology-telecommunications and Information Exchange Between Systems-local and Metropolitan Area Networks-specific Requirements-part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Television White Spaces (TVWS) Operation*, Feb. 2014.

[12] L. Zhu, V. Chen, J. Malyar, S. Das, and P. McCann, *Protocol to Access White-Space (PAWS) Databases*, document RFC 7545, Internet Eng. Task Force, Fremont, CA, USA, 2015.

[13] F. Communications Commission, *Amendment Part 15 Commission's Rules for Unlicensed Operations Television Bands, Repurposed 600 MHz Band, 600 MHz Guard Bands and Duplex Gap, and Channel 37: Notice of Proposed Rulemaking*, document FCC: Notice of Proposed Rulemaking, ET Docket No. 14-144, 2014.

[14] *Amendment Part 15 Commission's Rules for Unlicensed Operations Telev. Bands, Repurposed 600 MHz Band, 600 MHz Guard Bands and Duplex Gap, and Channel 37: Report and Order*, document FCC: Notice of Proposed Rulemaking, ET Docket No. 14-144, 2015. [Online]. Available: https://www.fcc.gov/document/amendment-part-15-commissions-rules-unlicensed-operations-television-bands

[15] D. A. Guimarães, E. J. T. Pereira, A. M. Alberti, and J. V. Moreira, "Design guidelines for database-driven Internet of Things-enabled dynamic spectrum access," *Sensors*, vol. 21, no. 9, p. 3194, May 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/3194

[16] F. Communications Commission, *Wireless Telecommunications Bureau and Office of Engineering and Technology Approve Four Spectrum Access System Administrators for Full-scale Commercial Deployment in the 3.5 GHz Band and Emphasize Licensee Compliance Obligations in the 3650-3700 MHz Band Under Part 96*, document GN Docket No. 15-319, Federal Communications Commission, 2020. [Online]. Available: https://docs.fcc.gov/public/attachments/DA-20-1102A1.pdf

[17] F. C. Commission. (Apr. 2020). *3.5 GHz Band Overview*. [Online]. Available: https://www.fcc.gov/wireless/bureau-divisions/mobility-division/35-ghz-band/35-ghz-band-overview

[18] Z. Qin, Y. Gao, and C. G. Parini, "Data-assisted low complexity compressive spectrum sensing on real-time signals under sub-nyquist rate," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1174–1185, Feb. 2016.

[19] Y. Luo, L. Gao, and J. Huang, *Economics of Database-Assisted Spectrum Sharing*. Cham, Switzerland: Springer, 2016.

[20] J. Wang, N. Varshney, C. Gentile, S. Blandino, J. Chuang, and N. Golmie, ''Integrated sensing and communication: Enabling techniques, applications, tools and data sets, standardization, and future directions,'' *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23416–23440, Dec. 2022, doi: 10.1109/JIOT.2022.3190845.

[21] S. Inc. (2025). *Number of Internet of Things (IoT) Connections Worldwide From 2022 to 2023, With Forecasts From 2024 to 2033*. Accessed: Apr. 9, 2025. [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

[22] J. Doe and J. Smith, ''The role of 5G in advancing mobile IoT devices,'' in *Advances in Emerging Technologies for IoT*, A. Brown, Ed., Cham, Switzerland: Springer, 2023, ch. 1, pp. 201–220, doi: 10.1007/978-981-99-3668-7_11.

[23] D. A. Guimarães, ''Hybrid fusion of Pietra–Ricci index detector information for cooperative spectrum sensing,'' *Ad Hoc Netw.*, vol. 150, Nov. 2023, Art. no. 103265, doi: 10.1016/j.adhoc.2023.103265. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870523001853

[24] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, ''An efficient k-means clustering algorithm: Analysis and implementation,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.

[25] S. Zhu, Tahereh. S. Ghazaany, S. M. R. Jones, R. A. Abd-Alhameed, J. M. Noras, T. Van Buren, J. Wilson, T. Suggett, and S. Marker, ''Probability distribution of Rician K-factor in urban, suburban and rural areas using real-world captured data,'' *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3835–3839, Jul. 2014, doi: 10.1109/TAP.2014.2318072.

[26] T. S. Rappaport, *Wireless Communications: Principles and Practice*. London, U.K.: Pearson, 2002.

[27] J. Vallet García, ''Characterization of the log-normal model for received signal strength measurements in real wireless sensor networks,'' *J. Sensor Actuator Netw.*, vol. 9, no. 1, p. 12, Feb. 2020. [Online]. Available: https://www.mdpi.com/2224-2708/9/1/12

[28] M. Gudmundson, ''Correlation model for shadow fading in mobile radio systems,'' *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[29] D. A. Guimarães, ''Pietra-ricci index detector for centralized data fusion cooperative spectrum sensing,'' *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12354–12358, Oct. 2020, doi: 10.1109/TVT.2020.3009440.

[30] D. A. Guimarães, E. J. T. Pereira, and R. Shrestha, ''Resource-efficient low-latency modified Pietra–Ricci index detector for spectrum sensing in cognitive radio networks,'' *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 11898–11912, Sep. 2023.

[31] E. J. T. Pereira, D. A. Guimarães, and R. Shrestha, ''VLSI architectures and hardware implementation of ultra low-latency and area-efficient pietra-ricci index detector for spectrum sensing,'' *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 71, no. 5, pp. 2348–2361, May 2024.

[32] M. Lin, W. Wang, X. Hong, and W. Zhang, ''GLRT approach for Multi-Antenna-Based spectrum sensing under interference,'' *IEEE Commun. Lett.*, vol. 24, no. 7, pp. 1524–1528, Jul. 2020.

[33] The Institute of Electrical and Electronic Engineers, IEEE (2011). *IEEE 802—Part 22: Cognitive wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands*. [Online]. Available: http://standards.ieee.org/getieee802/ download/802.22-2011.pdf

[34] D. A. Guimarães. (Dec. 2024). *MATLAB Code for Performance Analysis of Overlapped Clustering for Spectrum Hole Geolocation*. [Online]. Available: https://www.dropbox.com/scl/fi/u3fw0pr5vcz4pj2utjrjy/Code_pack.zip?rlkey=bbfzlxd0or63enc9glmbt8e51&dl=0

DAYAN ADIONEL GUIMARÃES received the M.Sc. and Ph.D. degrees in electrical engineering from the State University of Campinas (Unicamp), Brazil, in 1998 and 2003, respectively. He has been affiliated with the National Institute of Telecommunications-Inatel, Brazil, for more than 30 years, where he is currently a Researcher and a Senior Lecturer. His research interests include wireless communications, with a focus on radio propagation, digital transmission, spectrum sensing, dynamic spectrum access, and random signal processing.