

Received September 1, 2021, accepted September 27, 2021, date of publication October 13, 2021, date of current version October 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3119524

Artificial Intelligence for Radio Communication Context-Awareness

MAŁGORZATA WASILEWSKA¹, (Student Member, IEEE),
ADRIAN KLIKS¹, (Senior Member, IEEE), HANNA BOGUCA¹, (Senior Member, IEEE),
KRZYSZTOF CICHON¹, (Member, IEEE), JULIUS RUSECKAS², GEDIMINAS MOLIS²,
AUŠRA MACKUTĖ-VARONECKIENĖ², AND TOMAS KRILAVIČIUS²

¹Faculty of Computing and Telecommunications, Institute of Radiocommunications, Poznań University of Technology, 61-131 Poznań, Poland

²Baltic Institute of Advanced Technology, 01124 Vilnius, Lithuania

Corresponding author: Małgorzata Wasilewska (malgorzata.wasilewska@put.poznan.pl)

This work was supported in part by the Project CERTAIN of the Research Council of Lithuania under Grant S-LL-18-9, and in part by the National Science Centre of Poland through the Polish-Lithuanian Funding Initiative DAINA under Project 2017/27/L/ST7/03166.

ABSTRACT This paper surveys Artificial Intelligence (AI) methods for acquiring and managing context-of-operation awareness of radio communication nodes, links, and networks. The meaning and significance of context information and suitability of Machine Learning (ML) methods for the enrichment of context information is discussed. A number of context features are considered in this regard and thorough analysis on which ML methods are suitable to which part of context learning is provided. The added value of the paper is the presentation of a synthesized framework of context-information processing, sharing, and management in a radio communication network by delineating a network-embedded subsystem for this management. Recommendations for a future AI/ML-based radio communication system architectures are also provided.

INDEX TERMS Artificial intelligence, cognitive radio, context awareness, machine learning, radio context information, wireless communication.

I. INTRODUCTION

In the era of ubiquitous information access and pervasive communication networks, systems and nodes are needed to be aware of their context of operation, utilizing information on ambient networks, links, devices and applications. This context awareness will allow improvements in the efficiency of existing services, and provision of personalized services. For example, networks will need to be more aware of the application requirements, Quality of Experience (QoE) and Quality of Service (QoS) metrics, local (or more global) conditions of operation, and apply specific ways to adapt the application flows to meet users' needs under specific environmental conditions. The context-based adaptations of various transmission and network parameters will have to take into account the device-level, user-level, link-level, network- and application-level context. The context information itself consists of different parts/components, each of which affects the individual steps of the decision making process in a different way. More specifically, various parts

that constitute the context are related to the following levels: (i) the hardware platform, which poses specific hardware constraints and implementation issues, (ii) radio environment conditions in terms of location-specific parameters, wireless channel quality, spectrum availability, other-users characteristics and signal features, traffic patterns, interference levels, etc., (iii) required performance (QoS) parameters that can be identified in all layers of the system protocol stack, and are considered to be the basis for the evaluation of decisions made, (iv) network management policies as a set of rules used to control the behavior of nodes, manage available resources, regulate interference to other deployed systems, obtain identified trade-offs, etc. Note that in communication system management, (i) and (iv) are constraints, (ii) are changing radio-environment parameters and characteristics, and (iii) are the goals of optimisation. In this paper, we concentrate specifically on the radio context, i.e., on (ii).

Apart from and beyond gathering of context-information building data, radio-environment awareness can be enriched with reasonable overhead by using suitable machine learning (ML) methods and artificial intelligence (AI) embedded in communication networks, either at the network edge, or in

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco Rafael Marques Lima¹.

a central entity, or in a dedicated subsystem. In this paper, we survey such ML/AI methods, and discuss their suitability for particular radio-environment contexts and radio applications.

There are a few survey papers related to our topic. Let us now overview these published surveys, and compare them with the content of our work in the following aspects that we undertake: (a) the considered radio communication scenario and related meaning of context-awareness, (b) the completeness of radio-context information and context awareness, (c) considered methods for radio-context information acquisition and context awareness enrichment and (d) formal context-information processing and management framework embedded in a radio communication network.

In [1], the authors focus on context importance in the global computing network, particularly in the Internet of Things (IoT). They overview methods of context-aware *computing* in desktop, web, mobile, and sensor networks of IoT. A number of solutions are considered in terms of systems, middleware, applications, techniques, and context-aware computing models. Likewise, survey [2] focuses on mobile IoT, provides a comprehensive overview of context-aware middleware design, and categorizes context-aware applications, recognizing human-centric and community-based social activities in IoT as key future directions. The authors of [3] focus strictly on middleware, and how it handles context modelling, management, reasoning and provisioning of related functions. In [4], context-aware mobile networking integrated with (cloud) computing is considered. Here, the meaning of context awareness, its uncertainty levels, functionalities and classification are provided in the form of a proposed taxonomy scheme that is mapped to mobile cloud computing. The authors of [5] and [6] focus on ML methods for big data processing (computing), IoT and social networks. Thus, papers [1]–[6] differ significantly from ours in the considered scenarios, i.e. in aspect (a). Moreover, they do not address the practical design of context-aware radio networks (our aspect (d)), although [2] mentions this as one of the key challenges.

The authors of [7] surveyed the spectrum sensing methodologies for Cognitive Radio (CR). In [8], methods for obtaining energy-efficiency in cooperative spectrum sensing have been reviewed. Local sensing algorithms, selection of cooperating, sensing, reporting and relaying nodes, the fusion rule, and network organisation have been discussed for the energy-efficiency purpose. Likewise, paper [9] concentrates on the context-awareness subtopic, namely on localization methods. Neither in [7] and [8], nor in [9], have ML methods been considered to improve the awareness of spectrum availability, and no relative formal context-awareness management model has been proposed. Thus, these papers' scopes are narrowed with respect to our mentioned aspect (b) and not addressing aspects (c) and (d).

Paper [10] is not a survey, but a tutorial on Artificial Neural Networks (ANN) in wireless communication networks. It also presents some ANN applications, i.e., in unmanned

aerial vehicles communication, virtual reality applications, edge computing and caching. Thus, this paper does not fully cover our aspects (a) and (b), and presents a subset of ML methods, hence it is addressing aspect (c) only partially. No formal ML-based context-information management framework (aspect (d)) is considered.

Although not really a survey paper, [11] analyzes the role of cross-layer information exchanges in context-aware CR networks by adopting a generic layered model composed of physical, link, network and higher layers managed by a *cognitive engine*. The considered context-information refers to QoS parameters in the interference networks. Learning methods for communication traffic prediction are only briefly mentioned in this paper. Thus, this paper content is narrowed in aspect (b) and (c) when compared to our survey.

In [12], compressive sensing (sampling), also known as sparse sampling, is considered for a few elements of radio-context awareness, e.g. SNR or channel estimation. The paper strictly focuses on the theory that certain signals can be recovered from far fewer samples than required by traditional methods. Thus, this paper is quite disparate from ours, narrowing the considerations in aspects (a) and (b). It also does not touch upon aspect (c).

Finally, in paper [13], learning problems in CR systems have been considered, and the state-of-the-art of learning methods, mainly applied for detection of primary users and their transmission parameters, has been presented. The paper classifies applied learning methods, and focuses on primary-users' transmission-related parameters, not on the radio-context as a whole. Apart from a general CR cycle (or engine), it also does not consider any practical context-information processing subsystem embedded in a real-world network. Thus, [13] is different from our work in aspect (a), and in addressing aspects (b) and (d) only partially.

Having discussed the recent surveys on context-awareness, we claim that our paper is original in all considered aspects (a)–(d), presenting a complete overview of AI/ML methods applied for context-awareness in radio communication systems and discussing a practical framework for context information management. Distinctively, apart from the problem of context-awareness acquisition and enrichment (via ML techniques) in a radio network, we elaborate on a practical network-embedded subsystem to store, share, process and manage context information. This framework is our original contribution beyond the survey of the existing works. Based on that, we formulate recommendations for context-aware wireless network design and architectures, which also constitutes our contribution and the survey recapitulation.

The paper is organized as follows. In Section II, we review and set out the exact meaning and significance of context information in wireless communication systems. In Section III, we review and discuss the suitability of ML methods for the enrichment of context information. In Section IV, we concentrate on spectrum sensing with ML methods for spectrum availability detection and prediction. In Section V, we review ML methods for signal-features

detection as another important context-information category, while in Section VI, we focus on gathering the localization information with AI algorithms. We address ML-based traffic pattern recognition and prediction in Section VII. Section VIII synthesizes the preceding sections by looking at context-information processing, sharing and management as a whole, addressing a practical framework and delineating a network-embedded subsystem for this management. Finally in Section IX, we conclude our work and attempt to give recommendations for future AI/ML-based radio communication system architectures.

II. THE SIGNIFICANCE OF CONTEXT AWARENESS IN WIRELESS COMMUNICATIONS

Context-awareness (CA), being a basis for pervasive computing for more than thirty years [1], [14], has recently been considered as a highly promising, yet challenging concept in the domain of wireless communications [15]–[17]. Although the algorithms applied in contemporary wireless systems try to adjust the transmission profile to varying channel conditions (such as adaptive bit and power loading [18]–[21], advanced modulation schemes [22], [23], already standardized modulation and coding schemes (MCS) [24], [25], close- or open-loop power control [26], [27]), the application of AI/ML toolboxes fosters the implementation of situation-aware communication systems. The possibility of collecting vast and rich information about the surrounding (radio) environment delineates new communications paradigms, where user-centered transmission is adjusted to the current communication context. Practically, the data may be collected from dedicated sensors or directly from end-user terminals (as it is done currently in, e.g., minimization of drive test (MDT) schemes in cellular networks [28]), applying the concept of crowd-sensing or crowd-sourcing [29], [30]. Access to the gathered data allows us to construct the context of the considered scenario, which can be further used to adjust the communication link and network profile. However, as the amount of gathered data increases (due to a high number of data sources, increased frequency of data collection, or increased amount of observed data per single read), context-aware communications have to deal with problems typical for *big data processing* [31]–[34]. It is clear that the lack of efficient big-data processing makes the collected information simply useless. *Raw data*, generated directly by sensors, need to be processed, managed, checked for consistency, and complemented to create the complete context information.

A. RADIO COMMUNICATION AWARENESS

Various definitions of the term *context* can be found in the literature. The Merriam-Webster's Dictionary provides a general definition of it - a context refers to *the interrelated conditions in which something exists or occurs: environment, settings*" [35]. Similarly, the historically important definitions related to computing and communications, such as in [36], [37], define the term *context* in a specific situation

or refer to some certain use case. In particular, the authors of [36] describe context as specific locations, identities of close objects and creatures, and changes to them. However, in a broader sense, context can be specified operationally as proposed in [38], and for the sake of clarity, we quote this definition verbatim below:

Context is any information that can be used to characterize the situation of entities (i.e., whether a person, place, or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves. Context is typically the location, identity, and state of people, groups and computational and physical objects [38].

Finally, in [39], context is presented as a set of interrelated events between which logical and timing relations can be identified. The events are classified into discrete (such as starting a call) and continuous ones (executing the call). Assuming that there is a set of interrelated events that specify a given context, the logical relations are defined as the Boolean formula of the appearance of these individual events. For example, context is said to be a unit context, when all of the constituent events have to be understood as *true*.

In view of the above discussion, for completeness, it is worth analyzing the meaning of *context awareness*. One may say that a system or algorithm will be aware of an existing context if it uses the context (and all data related to it) to provide detailed information to the end-user. In this sense, a wireless system that is context-aware will support various features, and possess specific attributes, such as the ability to observe the surrounding environment, sense it, perform data acquisition and processing, and finally react.

When it comes to radio communications, the terms context and context awareness should be adjusted accordingly. In particular, radio communication context will be a set of information and data that characterize the communication-related situation of the network or of the entire wireless system. Thus, radio communication context will be constituted by such descriptors as geographical location, identity and state of wireless nodes (persons or things, base stations, transmission points etc.), status of wireless channels, transmission requirements and system performance. Again, the radio context may also be presented as a set of interrelated events describing the functioning of the wireless transceivers and whole network. Thus, radio communication context awareness will be the ability (of a device, a network or a system of networks) to observe the surrounding radio and geographical environment, sense it, perform data acquisition and processing, and react accordingly. It is evident that the ability to be radio-communication context-aware is an inherent feature of the cognitive radio and cognitive networks [40]. Access to rich context information about the surrounding radio environment leverages the application of more tailored communication schemes. Contemporary wireless communication systems utilize such a kind of data to some extent. One may think of the adaptation of transmit power (by means of so-called open or close control loops) or selection

of the best modulation-and-coding scheme depending on the instantaneous channel conditions. However, in the context of cognitive radio and cross-layer cooperation, much more advanced strategies have been considered, i.e. where advanced information exchange across the protocol stack has been proposed [11], [41], [42].

Various kinds of radio context information can be identified. First, following [1], it may be divided into two key classes - primary and secondary. As the former one is defined as the set of information retrieved without any data fusion operation (one may think of directly accessible information such as received signal strength), the latter refers to any information that can be derived based on the primary context data (such as the location of a user computed based on any triangulation scheme). Furthermore, various smaller classes of radio context information can be specified, i.e., information related to location, time, identity and activity. Regardless of the exact classification of context data, it is worth summarizing jointly the examples of radio context information available in contemporary wireless systems (mainly cellular networks, but also wireless local and personal area networks). They are typically used for describing the observed or predicted quality of signal, assessing the measured signal power, defining the best method of adaptive signal processing or just for describing the generic system setup. These are presented in Tab. 1. One must notice that this table is not complete, and many other parameters can be specified. However, considering them jointly, and particularly with association to a certain location and time, detailed context awareness can be achieved by contemporary systems. At the same time, further exploration and processing of various types of information may increase the overall context-awareness, and in consequence, may lead to better exploitation of available resources at the expense of data processing (hardware resources and related energy consumption). We believe that AI/ML are excellent tools for exploiting acquired data, extracting useful information contained in these data and enriching the context awareness of a considered system.

B. CONTEXT INFORMATION LIFE CYCLE

It is evident that the context awareness may have various time-scales. Some data will be valid for a short time, whereas other will represent long-term trends. This phenomenon is often referred to as information ageing [43], and recently the age-of-information metric has been considered as a tool for measuring the applicability of certain data while stating their validity. However, the age of information entails the need for updating such a kind of data, creating a so-called context life cycle [1]. In the domain of computer science, two terms are typically discussed, mainly data lifecycle management and information lifecycle management. In the case of radio communication awareness, such a lifecycle may define how radio data pass from one stage to another, whereas a stage represents the validity of the data and its ageing. For example, four stages of such a lifecycle can be identified, namely: radio context acquisition - modelling - reasoning - dissemination,

as presented in, e.g., [1]. Another approach is to identify these phases as: data collection - classification - processing and storage - sharing and dissemination, as we present in Fig. 1. However, much more advanced schemes are possible. For example, [44] proposed a more advanced scheme, where such phases as data collection, classification, handling and storage, release and backup are identified, as graphically presented in Fig. 1. In any way, there is a need for updating the collected information (permanently, periodically or on-request). As we discuss further in this paper, in an advanced wireless communication system, the process of data collection and processing for increasing local and global context awareness should be steered by appropriate AI/ML tools.

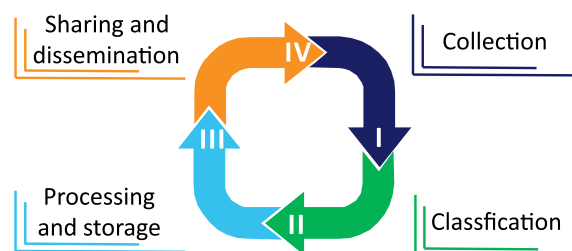


FIGURE 1. Generic cycle for radio context information management.

C. VARIOUS DOMAINS FOR CONTEXT INFORMATION GATHERING

Table 1 gathers commonly used metrics or parameters for defining the instantaneous radio communication context. They may be broadly classified as information related to power management (such as received signal strength, RSS, signal to noise ratio, SNR), channel quality measurements (such as channel quality indicator, CQI, channel state information, CSI), network configuration (such as absolute radio frequency channel number, ARFCN), selected signal processing schemes (such as rank indicator, RI or MCS) or traffic characterization (e.g., maximum bit rate, MBR, QoS class identifier, QCI). However, the list is obviously not complete, and the number of parameters used in contemporary wireless networks is continuously growing. From the point of view of creating rich radio context information, it is worth investigating new domains of network functioning, which can be used for context information gathering and enrichment.

First, following the findings made in the cognitive radio domain, much information about the presence of other transmissions can be gained through single-node or cooperative spectrum sensing. Although the ultimate aim of traditional spectrum sensing is to detect the presence or absence of primary users (i.e., the users of wireless systems licensed to occupy a specific frequency band at a certain location and time), this scheme can be extended to the detection of any existing transmission in the vicinity, or even more - to detect specific features of these transmissions (such as the type of signal, applied modulation scheme, etc.). The process

TABLE 1. Examples of radio context information utilized in various contemporary wireless networks.

Name	Description	Class	Similar Descriptors
ARP - Allocation and Retention Priority	Specifies relative importance compared to other bearers for the allocation and retention of a new bearer	Secondary	QCI (QoS Class Identifier), GBR (Guaranteed Bit Rate), MBR (Maximum Bit Rate), AMBR (Aggregate Maximum Bit Rate), AC (Access Categories)
ARFCN - Absolute radio-frequency channel number	Allows for the identification of the communication channel number and center frequency	Secondary	5G NR ARFCN (variant for New Radio), EARFCN (E-UTRA Absolute Radio Frequency Channel Number), UARFCN (UTRA Absolute Radio Frequency Channel Number),
CQI - Channel Quality Indicator	Indicates the quality of the wireless channel	Secondary	CSI (Channel State Information)
CSO (Cell Selection Offset)	Defined in various ways, used for cell range extension in heterogeneous networks	Primary	cellIndividualOffset, q-offsetCell
MCS - Modulation and Coding Scheme	Specifies used modulation and coding configuration	Secondary	MCS (Modulation and Coding Set - in IEEE 802.11 networks)
RI - Rank Indicator	Indicates the number of layers and signal streams transmitted in the downlink (LTE)	Secondary	PMI (Precoding Matrix Index), CQI (Channel Quality Indicator)
RSS - Received Signal Strength	Measured power of the received signal	Primary	RSSI (Received Signal Strength Indicator), RSRP (Reference Signal Receive Power), RSRQ (Reference Signal Received Quality), RCPI (Received Channel Power Indicator), Signal Strength, SS-RSRP (Synchronization Signal RSRP), OSTP (OFDM Symbol Transmit Power), RX (Received Power, e.g. in Bluetooth) etc.
SNR - Signal to Noise Ratio	Describes the ratio between the power of the wanted signal within certain band and the noise power observed in this band	Primary	SINR (Signal-to-Interference-plus-Noise Ratio)
QCI - QoS Class Identifier	Defines the quality of packet communication provided by cellular networks (LTE)	Secondary	5QI (5G QoS Identifier)
Generic physical parameters defining system	Fundamental information such as center frequency, bandwidth, start and stop frequency, maximum allowed power	Any	Various modifications of these parameters

of spectrum sensing may be realized by each device independently or cooperatively between communication-network nodes, and it may also be assumed that dedicated sensing nodes (deployed and devoted only for this purpose) are applied to improve data collection. In consequence, using various techniques for prospective big-data processing, rich radio-context information may be inferred. Spectrum sensing seems to be then one of the important new domains of radio context exploration that could be inherently integrated into the future wireless networks. In Sec. IV, we analyse key findings in the application of AI/ML tools for data gathering through spectrum sensing.

As already mentioned, besides the detection of the presence or absence of other ongoing transmissions, there is a possibility to enrich context information by the recognition of various signal features. This could include modulation recognition (e.g., if it is single carrier or multi-carrier) or identification of types of signals (if it is, e.g., a signal of a 3G, 4G or 5G network or other type of a distinctive local or wider area network). Exploiting such context information may serve multiple communication tasks, e.g., being able to transmit using OFDM (orthogonal-frequency division multiplexing) subcarriers orthogonal to the detected ones or spread-spectrum techniques not affecting the detected narrow-band signal with a detected modulation type.

Radio context information is highly dependent on (geographical) localization. Thus, the incorporation of user localization techniques are of paramount importance in wireless communication systems. As in the previous cases of

discussed context information domains, also here AI/ML tools can be applied to improve user localization.

Finally, the utilization of available radio resources (such as time-frequency chunks, allowable power, etc.) may be improved when the transmission patterns of other existing transmissions are known. These can be used for the specification of the types of other parallel signals and in consequence, for better adjustment of planned transmission to such a radio communication context. The above-identified domains are graphically presented in Fig. 2.

It should be observed that the increase of the amount of collected context information from new domains increases the overall processing complexity of system management (see Fig. 3). On the one hand, the richer the radio context information, the better the adjustments of the system setup, on the other, there is a significant increase in the complexity of such a system. Thus, there is a need, first, for accurate acquisition of the context information, and second, for advanced (big) data processing. In both cases, the application of AI/ML tools may provide reliable solutions.

D. APPLICATION EXAMPLE

As an example of context information importance, the case of Long Term Evolution (LTE) users moving on high speed trains can be analyzed. In this case, the context information is the speed and location of a train. In some real-world scenarios, users speeds can reach 500 km/h, and when using the OFDM modulation, inter-carrier interference (ICI) may have a large impact on the system performance due

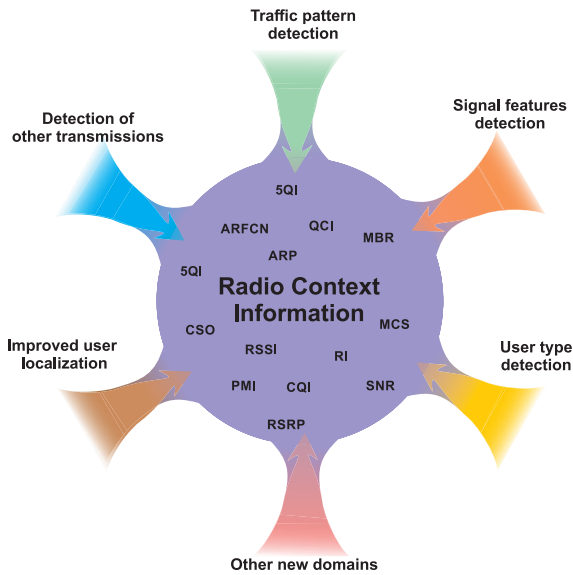


FIGURE 2. Enrichment of radio context information by adding new context information domains.

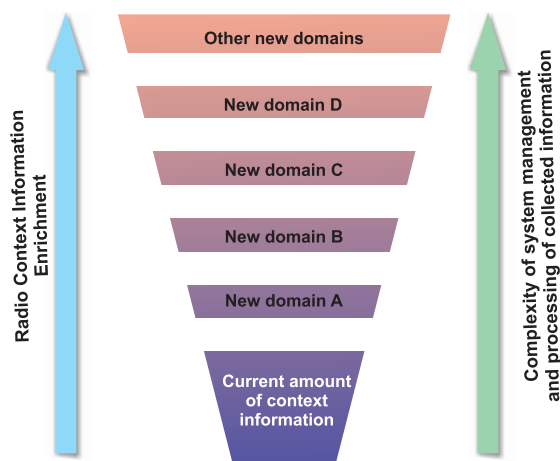


FIGURE 3. Improved radio context information leads to increased complexity of system management.

to the Doppler effect. Robust vehicle-to-vehicle communication techniques require the perception of the surrounding environment, prediction and compensation of the Doppler shift, so methods like car-embedded-sensor aided [45] and location-aware (in the context of frequency domain) Doppler distortion estimation and compensation [46] are proposed. This helps to improve bit error rate (BER) up to two orders of magnitude [45]. In the case of 5G for high velocity trains, the ICI compensation method helps to achieve the targeted 10% block error rate (BLER) when the train is moving at speeds up to 250 km/h, and a high modulation coding scheme like 256-QAM (quadrature amplitude modulation) with a coding rate of 3/4 is used [46]. Moreover, in the

latter paper, it was shown that with sub-meter positioning, a high accuracy Doppler frequency estimate can be obtained for Doppler frequency pre-compensation allowing significantly improved demodulation performance in the train relay operating at 30 GHz. Also in [47], it was shown that at the base station side, Doppler frequency pre-compensation (in the down-link) allows legacy devices to perform well even at 350 km/h velocity. If a line-of-sight (LOS) Doppler component is pre-compensated using accurate information on the localization and speed of a train, it is possible to achieve bit-rates up to 10 Gbit/s, when using high MCS schemes like 256-QAM and the train is moving at 500 km/h speed [48]. Using a link level simulator, it was shown that the LOS component is 45 dB stronger than the second strongest component (non-LOS, NLOS), and it is enough to pre-compensate only the LOS signal component to achieve these bit-rates. Thus, it is evident that context awareness, in the case of users moving of high speeds, is essential to achieve low-latency and high-throughput goals declared in the 5G road map.

Another example of context information importance is mmWave micro-cell discovery procedures in future 5G network architectures. A significant drawback of high frequency (mmWave) radiation is a high path-loss and complete blocking of the radiation in NLOS scenarios by common obstacles such as trees, buildings, etc. Due to the high path-loss, beam-forming technique is extensively used to increase SNR. A variety of the beamforming techniques can be used, and when seeking to increase the user discovery distance, a narrow beam must be formed, thus increasing the number of possible beam orientations and consequently, the time needed for discovery, if the random scanning approach is used [49]. Hence, the closest-cell discovery is a task where context information provided by a base station using legacy networks and exploiting lower frequencies, becomes important to minimize the initial cell discovery time [50], [51]. Furthermore, as the network density increases by introducing mmWave picocells and femtocells, the problem of determining which radio access technologies (RAT) a user should use for cell discovery at a given time, becomes more complex. New methods are proposed, such as ones based on the context-aware radio access technology (CRAT) [52]. A mathematical model of CRAT, considering the user and network context is derived, adopting an analytical hierarchical process (AHP) for weighting the importance of the selection criteria and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [53] for ranking the available RATs. The simulation results obtained using the NS3 simulation environment, show that this approach outperforms the conventional A2A4-RSRQ approach, used in LTE, in terms of the number of handovers, average network delay, throughput, and packet delivery ratio by 20-100 %.

III. MACHINE LEARNING METHODS FOR CONTEXT AWARENESS

Machine learning methods can be broadly organized into three groups: supervised learning, unsupervised learning and

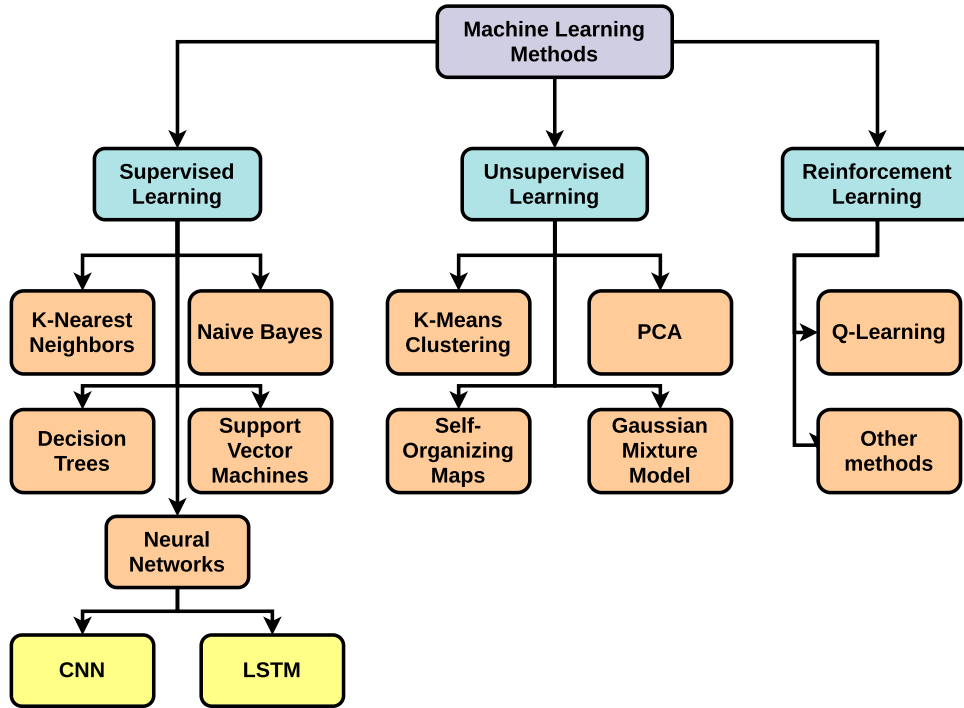


FIGURE 4. Most popular machine learning methods used in communication networks.

TABLE 2. Application of ML methods for context awareness.

Category	Method	Examples of applications
Supervised learning	K-nearest neighbors	Single node spectrum sensing [54], cooperative spectrum sensing [55]–[57], modulation recognition using feature set based on higher-order statistics [58], indoor user localization [59]
	Naive Bayes	Single node spectrum sensing [60]
	Decision Trees	Modulation recognition [61], [62]
	Support vector machines	Single node spectrum sensing [63], cooperative spectrum sensing [64], [65], cooperative decision scheme for spectrum sensing in vehicular communication environment [66], user classification [67]
	Neural Networks	Single node spectrum sensing [68], modulation recognition [69], multi-carrier modulation recognition [70]
	Neural Networks: CNN	Single node spectrum sensing [71], combining sensing results in cooperative spectrum sensing [72], modulation recognition [73], traffic pattern recognition [74]
	Neural Networks: LSTM	Prediction of channel state [75], modulation recognition [76], prediction of PU’s next spectrum state [77]
Unsupervised learning	K-means clustering	Single node spectrum sensing [78], cooperative spectrum sensing [64]
	PCA	Modulation recognition [79], multi-carrier modulation recognition [80]
Reinforcement learning	Q-learning and others	Learning spectrum sensing policy [81], cooperative sensing [82]

reinforcement learning. A schematic grouping of various machine learning methods is shown in Figure 4. Below, we outline the basics of these methods and their application to build and enrich context awareness in radio communication systems. The possible applications of different methods for context awareness are displayed in Table 2.

In supervised learning, a predictive model is constructed by using the training data that consists of inputs together with corresponding output values. The goal of the model is to minimize the difference between the model output and the actual values. There are many supervised learning methods, which we shortly outline below.

k-nearest neighbors (kNN) is a non-linear method where the predicted output is the average of the values of k nearest

neighbours of the input. For the distance metric in the input space, the Euclidean distance is commonly used. The kNN models are easy to interpret, fast in training, and have a small number of parameters to tune. However, the accuracy of prediction is generally limited. In the domain of radio context awareness, the kNN method has been employed, among other things, for spectrum sensing [55]–[57] and modulation recognition [58]. However, the limited accuracy means that kNN is most useful in resource-constrained environments.

Naive Bayes (NB) method is based on the Bayes theorem for calculating probabilities using prior probabilities. A Naive Bayes classifier assumes that all features are conditionally independent. It requires a small amount of training data and is recommended when the dimensionality of the input is high.

Decision trees (DTs) is a flow-chart model in which each internal node represents a test on an attribute. Each leaf node represents a response, and a branch represents the outcome of the test. DTs have parameters such as desired depth and the number of leaves in the tree. They do not require any prior knowledge of data, and are robust against outliers or label noise in data [83]. The complexity-cost of using a tree is logarithmic in the number of data points provided for training. Decision trees may be biased if some classes dominate in the training dataset, therefore, a balanced dataset is required prior to fitting. Unlike other methods, decision trees can process categorical and numerical data even without data normalization. Decision trees have been applied for spectrum sensing [84], and modulation recognition [61], [62].

Support vector machines (SVMs) use training data to come up with a hyperplane that separates classes with the largest margin. The sample points that form the margin are called support vectors and establish the final model. When a good linear separator cannot be found, kernel techniques are used to project data points into a higher dimensional space where they can become linearly separable. Thus, the correct choice of kernel parameters is crucial for obtaining good results. This method, in general, shows high accuracy in prediction, and it can also behave very well with non-linear problems when using appropriate kernel methods. An exhaustive search must be conducted on the parameter space, thus complicating the task. Since the optimal problem solution by SVMs is convex, SVMs deliver a unique solution, in contrast to Neural Networks which provide multiple solutions associated with local minima. SVMs may provide high performance for small problems, however, their computation and storage requirements increase rapidly with the number of training vectors. SVMs are not scale invariant, therefore the data need to be re-scaled before being fed as input into SVM. An SVM classifier can be used for spectrum sensing and decision-making [63]–[67].

An artificial neural network (ANN) is a statistical learning model consisting of interconnected nodes, also called neurons [85]. A neuron gets information from all neighboring neurons and gives an output depending on its activation functions. The connection strengths between neurons are represented by adaptive weights. During the learning process, the weights are adjusted until the output of the network is approximately equal to the desired output. As a special type of neural networks, convolutional neural networks (CNNs) use convolution operations with a set of kernels (filters) instead of employing full connections between layers of neurons [85]. Since convolution operations are invariant in terms of translation, CNNs are useful for analysing spatial data. Another type of neural networks, recurrent neural networks (RNNs) are designed for modeling sequential data, where sequential correlations exist between samples. RNNs use recurrent connections from a neuron in one layer to neurons in previous layers. In the training of traditional RNNs, vanishing or exploding gradient problems frequently occur, which make them hard to train. The long short term memory (LSTM) is a special

kind of RNN that mitigates these issues by introducing a set of gates [86]. The connection pattern between different layers of neurons, the learning process for updating the weights of interconnections, and the activation function that converts a neuron's weighted input to its output activation are the most important parameters to be trained. Neural networks may face slow training depending on the network size. They provide multiple solutions associated with local minima, and for this reason may not be robust over different samples. ANNs can be used at the secondary User (SU) end in cognitive radios to act dynamically whenever there is a primary user's (PU's) activity detected over the channel to avoid collision. ANNs can be used in spectrum sensing and adapting radio parameters in cognitive radio [68]–[70]. RNN and LSTM have the capability to find embedded characteristics and leverage long-time dependency in a sequence.

In unsupervised learning, only unlabeled data is provided, and the goal of a model is to find a pattern in data. The most common applications of unsupervised learning methods are clustering, dimensionality reduction and anomaly detection.

The purpose of clustering is to identify groups of data and build a representation of the input. Clustering methods can be classified as non-overlapping, hierarchical and overlapping. Among non-overlapping methods, K-means clustering and self-organizing maps (SOMs) are most popular. K-means clustering aims to partition observations into clusters, so that each observation belongs to a cluster with the nearest mean and within-cluster variance is minimized. The k-means applications for context awareness improvement mainly perform sensing [64], [78]. The most common algorithm uses an iterative refinement technique: k clusters are created by associating every observation with the nearest mean, and then the centroid of each of the k clusters becomes the new mean. In SOM, unlabeled data are fed into a neural network to produce a low-dimensional, discretized representation of the input space of training samples, called a map. In overlapping clustering an observation can exist in more than one cluster simultaneously. Gaussian mixture models belong to this class of methods.

Dimensionality reduction methods produce lower dimensional models of high-dimensional datasets. Principal component analysis (PCA) does this by creating new combinations of features, which project data onto a lower dimensional subspace by identifying correlated features in data distribution. The principal components (PCs) with the greatest variance are retained, and all others are discarded to preserve maximum information and retain minimal redundancy. PCA can be useful in modulation recognition [79], [80].

The purpose of reinforcement learning (RL) is to learn through dedicated agent an optimal (or sub-optimal) policy maximizing so-called reward function based on the observed immediate rewards. The RL may be model-based (where the dedicated model is generated) or model free [87]. A basic reinforcement learning model consists of environment states, possible actions, rules for transition between states, rewards of transitions, and rules for observation. The learner, called

an agent, interacts continuously with the environment by selecting actions. The environment changes by responding to these actions and the agent receives numerical rewards as the environment response. In reinforcement learning, the agent tries to maximize rewards over time. Learning can be centralized in a single agent or distributed across multiple agents. Reinforcement learning is useful in sequential decision and control problems where it is not possible to provide explicit supervision and only a reward function can be given. Using RL, each SU can sense the spectrum, perceive its current transmission parameters and take necessary actions when a PU appears.

There are many RL algorithms and the review of all of them is out of scope of this paper. One of the popular methods is Q-learning, used for example for discovering of the optimal spectrum sensing policy [81] or interference control [88]. In Q-learning the algorithm computes the expected rewards (Q) of an action taken in a given state, independent of the policy being followed [87]. Then the next action is chosen using a policy derived from value Q and the observed rewards are used to update Q with the weighted average of the old value and new information.

Of all machine learning methods, supervised learning algorithms are most developed and are most frequently used in applications of machine learning for radio-context awareness considered in the literature. Supervised learning can be applied, for example, in spectrum sensing, modulation recognition or user classification. A possible drawback of supervised learning is the requirement for labeled data, which can be time-consuming to acquire. Unsupervised learning, employed not as commonly as supervised learning, is used to automatically find patterns in a large amount of data. Reinforcement learning is mostly suitable for complex network-control problems.

IV. SPECTRUM SENSING AND DECISION MAKING: A BASIS FOR INCREASING CONTEXT AWARENESS

While discussing context awareness and the process of collecting information about the environment, a natural and logical association could be to the contributions in the cognitive radio (CR) domain. In general, cognitive radio and cognitive radio networks (CRN) are assumed to follow the well-known cognitive cycle in order to optimize the overall functioning of a network and improve its performance. Within this cycle, CR and CRN tend to observe the environment, learn and perform decisions based on collected data. It is evident that observation of the environment, also through spectrum sensing or access to some databases, is a rudimentary requirement of the CR technology [40], [89]. As CR has already been investigated for more than two decades, let us start our analysis on AI-based context-awareness by reviewing recent findings in this domain.

Spectrum sensing (SS) is in principle the process of observing a given spectrum portion and making a decision about the presence/absence of a licensed signal in the observed band at the given location. A great number of schemes have

been proposed in the last twenty years for efficient spectrum sensing in the literature. These methods may either be blind or use some a-priori knowledge about PU signals and noise variance [7], [90]. Well-known representatives of the first group of SS algorithms are eigenvalue-based detection, higher-order-statistics based detection, or a solution focusing on the symmetry property of the cyclic autocorrelation function. Next, the well-known semi-blind technique requiring just the knowledge on noise variance is energy detection, i.e., a simple scheme yet highly inaccurate in low SNR regions. On the other hand, a classic non-blind method is matched-filtering [91]. Irrespective of the selected sensing technique, a sensing entity can make correct or incorrect sensing decisions and the spectrum band can be occupied or vacant, resulting in four cases constituting a well-known confusion matrix. When a node detects a signal and it is indeed present, one refers to the probability of detection P_d . In the case of signal absence and correct decision, the correct-negative scheme is considered. Next, two types of errors are known - false positive described by the probability of false alarm P_{fa} (the probability that the signal is absent but it is decided to be present), and false negative measured typically as the probability of misdetection P_{md} (signal is present but it is not detected). The decision process based on these estimated probabilities is referred to as the double-hypothesis statistic test [8]. Typically, the performance of any receiver is expressed by means of the so-called receiver operating characteristics (ROC), which is the plot of a true positive rate (TPR) (which is the estimation of P_d) against a false positive rate (FPR) (which is the estimation of P_{fa}) for various test-function threshold settings. The area under the ROC curve is widely used for sensing performance evaluation, and is denoted as AUROC, i.e., the area under the receiver operating characteristics.

Machine Learning is becoming a more and more popular tool for improving or even replacing traditional spectrum sensing methods. The use of artificial intelligence algorithms in spectrum sensing allows for not only deciding on the PU's transmission state [63], [75], [92]–[95] but also enables the estimation of the number or localization of active PUs [96]–[99]. ML can also be useful in SS with feature recognition which can be used in PU and SU signal differentiation [67] or PU behavior recognition. Since the best chance to find an idle spectrum band is to search in a wide frequency range, ML has also been applied for wideband sensing in sparse signals [100]–[105] which is another utilization of ML in SS. Also, as an indirect way of using ML for SS, one can distinguish applying ML algorithms for SS threshold adaptation and also using ML in the fusion center (FC) in Cooperative Sensing as a decision method instead of typical OR, AND or majority rules [64], [106]–[110]. Last but not least, ML can be used for future channel state prediction based on SS data. In the following, however, we split the discussion into three dominant categories: A. Single node sensing enhancements, B. Wideband spectrum sensing, and C. Cooperative sensing improvements, including advanced data fusion techniques. Within each part, we try to arrange

the analysis of the existing solutions in two perspectives: the pure sensing process, and the resultant decision making.

A. SINGLE NODE SENSING IMPROVEMENT

Various AI/ML techniques have been proposed for the improvement of single node spectrum sensing, focusing on specific aspects of this procedure. A comprehensive comparison of numerous techniques has been presented in [92]. Mainly, the authors have verified the performance of 13 detection methods (including 11 ML methods) for SPN-43 radar detection. Actual radar transmission data in the form of spectrograms were used as input data for ML. Three classical ML algorithms have been chosen, namely SVM, kNN, and a Gaussian mixture model (GMM). The remaining eight applied ML algorithms are deep learning methods, and more specifically CNN methods: VGG-16, VGG-19, ResNet-18, ResNet-50, the Inception-V1 network, DenseNet-121, and two algorithms designed by the authors; one CNN algorithm, and LSTM RNN. All of the ML algorithm performances were compared with classical SS methods: energy detection and sweep-integrated energy detection. To compare the performance of the used algorithms, ROC curves have been drawn, also the speed of algorithms performances has been compared. The authors claimed that based on the evaluations of real-world data the superiority of ML-based detection has been proved when compared to the energy-based scheme. A similar comparison of the performance of ML-based solutions and traditional SS schemes has been presented in [93]. The authors raise an important question: when is it better to use ML techniques, and when is it more beneficial to rely on classical statistical signal processing methods. To investigate this issue, ML and signal processing methods were proposed for multiple transmitter detection and automatic modulation classification. For multiple transmitter transmission detection, two ML algorithms were proposed: TxMiner based on a Rayleigh-Gaussian mixture model and Log-Rayleigh mixture model. The performances of both of those algorithms have been compared with the signal processing method: multiple hypothesis testing based on normalized threshold binning. For automatic modulation classification, one ML algorithm, namely kNN, and one signal processing method (maximum likelihood) have been used. However, the authors claimed that there exists a significant trade-off between accuracy and computation/implementation complexity. In particular, it has been shown that ML-based solutions offer better accuracy at the expense of significantly higher complexity.

1) APPLICATION OF CLASSIFIERS

As spectrum sensing may easily be represented as some sort of classification, various AI-based classification tools have been considered in numerous papers. For example, in [63] the SVM algorithm has been proposed to classify spectrum into free or occupied classes. The category of the free spectrum is further classified into a few subcategories that indicate what power SU can use to transmit. This approach is supposed to minimize the interference

level in the case of misdetection. Other papers that consider the application of SVM-based solutions for single node sensing are, e.g., [111]–[113]. Also, [114] applies SVM to achieve better sensing performance while combining it with genetic algorithms and self-organizing maps. In [115], eigenvalue-based spectrum sensing with SVM in a multi-antenna cognitive radio was investigated. It has been further evaluated in [116], where SVM has been adopted for temporal, as well as joint spatio-temporal sensing, together with beamformer-aided feature extraction for enhancing the capability of SVM. The method allows for determining the actual number of active PUs and their locations in the network during the sensing interval. Least square regression, SVM, and manifold learning have been applied to classify features extracted by the energy detector, waveform-based sensing, and cyclostationarity-based sensing in [117].

Zhang *et al.* in [78] proposed a machine learning-based spectrum sensing framework for a scenario where PU operates on more than one transmit power level. The method does not require prior information of PU or the environment. Before sensing, SU undergoes a learning phase, where K-means clustering is applied to discover PU's transmission patterns as well as its statistics. Then, SVM is implemented to train SU to distinguish PU's status based on energy feature vectors. A similar approach is proposed in [118] for spectrum sensing using a sample covariance matrix of the received signal vector from multiple antennas. K-means clustering is performed to discover the primary user's transmission patterns, and afterwards, a decision is made using SVM. The feature vector used in learning is extracted from the covariance matrix and consists of the ratio of maximum and minimum eigenvalues and the ratio between the absolute sum of all matrix elements and diagonal elements.

In [84] the authors applied several ML methods for spectrum sensing to data from the GSM 850 MHz band on one day of March 2016. Records of the power of a radio channel have been obtained in 290 ms intervals throughout 15 h per day. kNN, SVM, logistic regression (LR), and DT classifiers have been investigated. It was found the best results were obtained using the DT classifier. However, SVM, kNN, and LR classifiers take much less time than the decision trees.

Although sub-Nyquist spectrum sensing in the narrowband case was researched in [103], the authors claim the proposed algorithm could be used in wideband as well, after some modifications. The algorithm uses a low sampling rate and a learned dictionary to recover the sampled signal. In the end, ML classifiers are used to enhance detection. Two ML algorithms are tested: support-vector machine and deep neural network. As feature vectors, absolute gradients are used. The ML algorithm improved detection performance significantly, which is shown using the probability of detection and probability of false alarm for different SNR values. The algorithm has also been tested over lab measurements.

Finally, it is worth mentioning the Kalman filter-based channel estimation technique for tracking a temporally correlated slow fading channel, as presented in [119]. This

technique adapts parametric classifiers to changing channel conditions. Moreover, [60] proposed spectrum sensing in an OFDM system using an NB classifier. A class reduction-assisted NB method was used to train the model and reduce spectrum sensing time. The method has been tested on a simulation of a second-generation terrestrial digital video broadcasting (DVB-T2) system. It has been shown that compared with non-ML methods, the proposed method achieves higher spectrum sensing accuracy, particularly in at critical areas of low SNRs.

2) APPLICATION OF NEURAL NETWORKS

Another big class of AI-tools applied for spectrum sensing are neural-network-based solutions. In [68], ANN has been applied to predict the state of a radio channel using as input of the network the results of energy detection and cyclic spectrum feature detection. This allowed for improving the detection of PU at low SNRs. ANN was also proposed in [120] where as input features, energy (from energy detection and Likelihood Ratio Test statistic) was used. The method has taken shown to outperform the classical energy detection method.

In [71] the authors employed CNN for spectrum sensing in design, experimental assessment and software defined radio (SDR) implementation of the SU link. One-dimensional (1D) CNN has also been applied for spectrum sensing in [121]. As an input to CNN, a matrix composed of energy and cyclic spectrum features has been used; similar features have been employed in [68], where a back propagation neural network (BPNN) has been applied. Simulations show that the proposed algorithm has a higher detection probability than cyclostationary feature detection. Also in [122], the authors proposed a method based on CNN for spectrum sensing of the OFDM signal. According to this method, a covariance matrix (CM) is normalized and transformed into a gray-level representation which is classified using CNN. To examine the effectiveness of the algorithm, simulation experiments are performed. In the simulation platform, the transmitter uses packaged OFDM signal frame data based on the 802.11a protocol and the data stream is subject to Rayleigh fading and Gaussian white noise. It has been shown that the algorithm has good performance in low SNR environments and can be rapidly trained.

Next, a mixture of CNN and LSTM was considered in [123], where a deep learning model for spectrum sensing was investigated. The dataset for the experiments in this paper was obtained from a radio frequency signal sampled from digital radio. The experiments have shown that for in-band SNR in the range from -9 dB to -5 dB, the proposed model can achieve a 25 – 38% performance improvement over the energy detection method, and this performance improvement does not require the introduction of any prior information on the signal of interest.

Finally, various papers investigated the usage of autoencoders. The so-called autoencoder-based spectrum sensing (SAE-SS) has been proposed in [99] and the

stacked autoencoder-based spectrum sensing method with time-frequency domain signals (SAE-TF) has been used to detect the activity states of PU in OFDM signals. The methods allowed the authors to detect PU activity solely based on the received signals and proved robust to noise uncertainty, timing delay, and carrier frequency offset.

3) APPLICATION OF Q-LEARNING

Another big AI-class of solutions proposed for signal detection was Q-learning and Gaussian mixtures. In particular, in [88] real-time multi-agent reinforcement learning, known as decentralized Q-learning, has been proposed to manage the aggregated interference generated by multiple SUs. Similarly, in [81] a reinforcement learning algorithm that allows each autonomous SU to distributively learn its own spectrum sensing policy has been developed, assuming that PU channel occupancy follows a Markovian evolution (mainly, a Q-learning algorithm with a decentralized, partially observable Markov decision process, DEC-POMDP).

4) PREDICTION OF CHANNEL STATE

As AI tools are widely applied for future prediction, such an application has also been evaluated in the context of spectrum sensing and prospective spectrum occupancy identification. The prediction of signal presence exploits the temporal correlation of the collected historical signal. This approach can be very beneficial in terms of time and energy consumption.

In particular, in [75] the application of LSTM for spectrum state prediction was considered. As input data for experiments, two real spectrum datasets were used: GSM1800 downlink and satellite signals. The temporal correlation of collected data was utilized to make a future spectrum state prediction. The Taguchi method has been introduced to determine the network's architecture. Next, in [94] an LSTM network was used and also for the purpose of making a prediction on future channel state. As the signal to be detected the frequency hopping signal was used. As historical data used for predicting the future spectrum state, SS results for a specific frequency and timeslot were used. The authors in [95] employed two neural network (NN) algorithms for spectrum occupancy prediction, mainly, the multilayer perceptron and recurrent neural network, and also two SVM algorithms: SVM with Linear Kernel and SVM with Gaussian Kernel. Also, three different network traffic models were analyzed, namely, Poisson, interrupted Poisson, and self-similar traffic.

5) OTHER AI TOOLS

The variety of existing AI-tools is very high and can be classified in various ways. Thus, let us provide just a small glimpse of other AI-tools, which we have not assigned to the previous key groups. For example, in [97] the proposed method to improve SS performance for low SNR is a modified cyclostationarity feature detection algorithm based on softmax regression. As ML features, characteristic cyclic values are extracted from the spectrum when the signal is present and when it is not. The softmax regression is trained using

this feature dataset. The Hidden Markov model has been also considered in [124], where an algorithm for estimating channel parameters, based on expectation maximization (EM) is proposed. Similarly, the authors in [125] investigated an EM-based spectrum sensing algorithm. The number of active users in a given frequency band, the power received from each user, the occupied time slots, and the noise floor was estimated. The received estimated power was modeled as a Gaussian mixture, the Gaussian with the lowest mean is associated with the noise floor and used to estimate an adaptive threshold. The method has been validated in a Wi-Fi experimental setup, where real-world data have been acquired with a software-defined radio.

In order to better classify various single-node spectrum sensing algorithms, we have created a dedicated Tab. 3 which concludes all ML for SS improvement papers.

6) OBSERVED TRENDS

Based on the analysis of Tab. 3, one can conclude that obviously, the ultimate goal of the application of AI/ML tools for single-node spectrum sensing is to improve the accuracy of PU detection. Numerous approaches have been considered, spreading all classes of AI/ML algorithms: supervised, unsupervised, and reinforced. However, from the point of view of the gained context information, other aspects of the observed spectrum can be identified beside the accurate knowledge of PU presence. For example, by applying advanced ML methods, the knowledge of the PU traffic pattern or types of signal classes (in terms of detected modulation type, prospective SNR, etc.) can be obtained. Moreover, reliable prediction of PU behavior can also be guaranteed. Thus, the application of AI/ML tools leads not only to improving the performance of single node spectrum sensing but can be the source of some additional context information.

B. WIDEBAND SENSING IMPROVEMENTS

Cognitive radio should be able to adjust its behavior to the current channel conditions fast and dynamically. The ideal CR has equipment versatile enough to be able to sense and transmit on any frequency that is idle at the moment. This causes a strong need of having a wide range of frequencies that CR can sense, which is very difficult to obtain because of hardware limitations. On the other hand, a high sampling rate is needed. Using high sampling frequency would cause serious time delays which are unacceptable in SS, where the sensing time should be as short as possible to leave enough time for transmission [126].

In compressive sampling it is claimed that some signals can be recovered using sub-Nyquist sampling rate. The signals that can be used in Compressive Sampling must be sparse in some domain, for example in frequency domain obtained by using Fourier transform, or on the wavelet basis, [127]. Compressive Sensing can be used not only in Spectrum Sensing itself but also in signal parameter estimation and the number

of active PUs, PU's location and transmit power estimation as well. These applications are described in more detail in [12].

Compressive sensing can be used along with ML in order to further improve its performance. Papers like [100]–[104] concern compressive sensing and ML algorithms.

The authors of [100] present the use of PU probability prediction of transmission in a given channel for reducing the recovery time in Bayesian compressive sensing (BCS). Here, the relevance vector machine (RVM) is used to accelerate BCS Performance. In the paper, recovery time is compared between only BCS and BCS with RVM. The proposed algorithm works faster while maintaining the level of reconstruction error for different SNR values. Since the PU probability prediction-based BCS is used for spectrum sensing, the probability of detection depending on the SNR value is also presented. The modified BCS algorithm performance is superior to the basic BCS, and the results are even better for a higher number of PU predicted.

Next, in [101] the Bayesian compressive sensing was used in Cooperative Sensing with an adaptive threshold. BCS is used to sense a wideband spectrum. Then, the results of the sensing are sent to the fusion center, where RVM with an adaptive threshold is used in order to correct the recovery of sensed signals. In the end, the restricted Boltzmann machine (RBM) is used in the fusion center to decide about spectrum occupancy based on the recovered signals. In the proposed application, RBM learns to obtain static relation weights among SUs, which enables better decisions on the channel state. To obtain an adaptive threshold, an SVM algorithm was used. The results of adaptive threshold Bayesian compressive sensing (ABCS) were compared with BCS, basis pursuit method, and block orthogonal matching pursuit, and it was shown that ABCS has better signal recovery accuracy, as well as shorter recovery time. The RBM-based cooperative Bayesian Compressive Spectrum Sensing shows improvement in terms of detection probability comparable only to ABCS and basis pursuit adaptive threshold Bayesian compressive sensing. Compressive Sensing needs some prior information on the spectrum sparsity level. Usually, the sparsity level is assumed to be constant in time and its mean value is considered. Due to the fact that spectrum occupancy changes in time, so does the sparsity level. In [102], it is proposed to use an ML supervised algorithm to provide an estimated sparsity level, which should improve Compressive Sensing and minimize energy consumption, as well as the probability of false alarm. The sparsity level is being estimated separately in real time in every spectrum block by FC. The main algorithm of cooperative sensing can be explained as follows: FC calculates the mean value of occupancy and sends it to SUs. If a SU wants to transmit, it senses the spectrum, and sends the results to FC, where the actual sparsity is predicted. The new sparsity level is sent back to the neighbors of SU that want to transmit to perform spectrum measurements, which are used by FC to perform spectrum occupancy recovery. In spectrum occupancy prediction two ML algorithms are used: Linear regression using a gradient

TABLE 3. ML for single node SS improvement (selected papers).

Paper	Number of PUs and SUs	Signal type	Channel type	Feature set	Blind sensing	ML method	ML used for:	Gained context information
[60]	1 PU, 1 SU	OFDM (DVB-T2)	6-tap typical urban (TU6)	SNR samples	Yes	NB with class reduction	SNR are first mapped to classes, NB with class reduction is used for spectrum sensing	Knowledge of PU presence, SNR classes
[63]	1 PU, 1 SU	-	AWGN	Three statistical features based on signal samples	-	SVM	SS; when spectrum is classified as free; the received signal is classified into several sub-classes	Knowledge of PU presence, signal classes
[68]	1 PU, 1 SU	AM	AWGN	Four features are detected: Energy, three features for cyclostationarity	Partially (ANN is adjusted to AM cyclostationarity features)	BPNN	SS	Knowledge of PU presence; knowledge on AM signal presence
[75]	Not applicable (measurements)	GSM1800 downlink, satellite signals (measured)	Real measurements	Signal's power spectral density samples	Yes	LSTM network	Spectrum state prediction, estimation channel quality	Channel occupancy prediction with dedicated metrics (root mean square and classification accuracy)
[71]	1 PU, 1 SU	OFDM (4 × 2.5 MHz channels)	AWGN (not well specified)	IQ samples, spectrograms	Yes	Deep CNN	Detection of channel utilization pattern	Presence of PU, channel utilization pattern
[78]	1 PU, 1 SU	Gaussian	AWGN	Energy feature vectors	Partially	K-means; SVM	K-means used for discovery transmission patterns and PU statistics, SVM for PU status based on energy feature set	PU presence, PU statistics, PU transmission patterns
[81]	Multiple SUs, multiple PU channels sensed	not specified	Binary Symmetric Channel	-	-	DEC-POMDP	signal presence sensing, improvement of medium access scheme	PU's behaviour learnt in distributed way by each SU
[92]	Not applicable (measurements)	Radar signal - SPN-43 (measured)	Real measurements	spectrograms data	Yes	SVM, kNN, GMM, 7 different CNNs, LSTM RNN	SS	Knowledge of PU presence; knowledge of spectrum occupancy statistics and power distributions
[93]	N PUs, 1 SU	Not specified	Rayleigh fading channel, AWGN	Spectral values	Yes	TxMiner, Log-Rayleigh Mixture Model	multiple PUs detection	Knowledge of PU presence; modulation classification
[94]	1 PU, 1 SU	Frequency hopping	Real measurements	Windowed historical SS results	Yes	LSTM, back propagation network	Predicting feature spectrum state	Knowledge of PU presence
[95]	1 PU, 1 SU	-	-	Traffic patterns	Partially	Multilayer perceptron, RNN, LSVM, GSVM	Predicting feature spectrum state	Prediction of PU absence periods
[97]	1 PU, 1 SU	OFDM	AWGN	Features derived from cyclostationarity	Yes	Softmax regression	Improving cyclostationarity SS	Knowledge of PU presence
[112]	1 PU, 1 SU	BPSK	AWGN	Signal samples	Partially	SVM	Detection of PU presence	Knowledge of PU presence

Continued on next page

TABLE 3. (Continued.) ML for single node SS improvement (selected papers).

Paper	Number of PUs and SUs	Signal type	Channel type	Feature set	Blind sensing	ML method	ML used for:	Gained context information
[113]	1 PU, 1 SU	-	AWGN	differential entropy vector	Yes	SVM, kNN, Random Forest, logistic regression	Detection of PU presence	Knowledge of PU presence
[114]	1 PU, 1 SU	BPSK	AWGN	Signal samples	partially	SOM, genetic algorithms (GA), SVM	Detection of PU presence	Knowledge of PU presence
[115]	1 PU, 1/many SUs	BPSK	AWGN	Signal samples	Partially	SVM	Detection of PU presence	Knowledge of PU presence
[116]	Many PUs, many SUs	BPSK	AWGN	Signal samples	partially	SVM, multi-class SVM, beamforming-aided SVM	Presence of PU signal; prediction of PU status	Knowledge of the presence of multiple PUs; Prediction of PU behaviour
[118]	1 PU, 1 SU	Gaussian	AWGN and some fading	Signal samples	Yes	K-means, SVM	Detection of PU presence	Knowledge of PU presence; knowledge of PU traffic patterns
[84]	Not applicable (measurements)	GSM 850 band measurements	Real measurements	Spectral data	Yes	kNN, SVM, DT, LR	Detection of PU presence	Knowledge of PU presence
[119]	Many PUs, many SUs	BPSK	AWGN and Rayleigh fading	Signal samples	Yes	K-means supported by Kalman filter tracking	Detection of PU presence, channel estimation	Knowledge of PU presence; enhanced channel estimation
[120]	Not applicable (measurements)	FM band, EGSM band, DCS band, UHF TV band	Real measurements	Signal samples	Yes	BPNN	Detection of spectrum occupancy	Knowledge of PU presence
[121]	1 PU, many SUs	BPSK	AWGN	Signal samples; cyclostationarity features	Yes	CNN	Detection of PU presence	Knowledge of PU presence
[123]	1 PU, 1 SU	Artificially generated burstly transmission	Measured data	Signal samples	Yes	CNN, LSTM, fully connected NN (FCNN)	Detection of PU presence	Detailed knowledge of PU presence

descent, and support vector regression (SVR). To evaluate the results, the predicted occupancy has been compared with actual occupancy focusing on P_d and P_{fa} .

In contrast to the previously mentioned articles, papers [104], [105] do not apply compressive sensing. In [104] sparse Bayesian learning (SBL) is used. Based on this algorithm, the Expectation and Maximization algorithm is proposed. Sparse Bayesian learning is compared with two conventional Compressive Sensing algorithms, namely, the Basis Pursuit algorithm and Orthogonal Matching algorithm by examining ROC curves and reconstruction error. In [105] multicoset sampling is used, which enables an analysis of the wideband spectrum in multiple narrow bands separately. Then, SBL is used for learning to detect occupied channels. The proposed method is compared with multi-orthogonal

matching pursuit (M-OMP) and multiple signal classification (MUSIC) by comparing P_d .

In order to compare solutions presented in the papers considering wideband sensing, Table 4 is proposed.

Observed trends As in the previous case, where traditional narrowband single node spectrum sensing was considered, similar conclusions can be drawn for wideband spectrum sensing. As the AI/ML tools improve the performance of the compressed sensing procedure, also other kinds of information can be potentially fetched from the observed signal samples, such as the prediction of PU behavior. Please note that wideband spectrum sensing may also be treated as a first step in gathering context information - once the activity of PUs is detected, a more detailed narrowband, AI/ML-based spectrum sensing algorithm can be applied.

TABLE 4. ML for wideband sensing.

Paper	Number of PUs and SUs	Signal type	Channel type	Spectrum sensing approach	Blind sensing	ML method	ML used for:	Gained context information
[100]	M PUs, N SUs	-	AWGN	Compressive sensing	Yes	RVM	Accelerating performance of Bayesian compressive sensing	Knowledge of PU presence
[101]	M PUs, N SUs	-	Fading, AWGN	Compressive sensing	Yes	RBM, RVM, SVM	Cooperative Bayesian compressive sensing, threshold adaptation	Knowledge of PU presence
[102]	M PUs, N SUs	-	Path loss, AWGN	Compressive sensing	Yes	GD, SVR	Spectrum occupancy prediction	Knowledge of PU presence; occupancy prediction
[103]	-	-	AWGN	Using dictionary	Yes	SVM, DNN	Enhancing spectrum sensing performance	Knowledge of PU presence
[104]	Not applicable (measurement)	Signal acquired from RTL-SDR (25 MHz - 1.7 GHz)	Measured data	Compressed sensing and SBL	Yes	SBL, EM	Recovering sparse signal	Knowledge of PU presence
[105]	M PUs (at different channels), 1 SU	Multi band, multi channel sinc-based signals	AWGN	SBL	Yes	SBL	Learning detecting occupied subchannels	Knowledge of PU presence

C. COOPERATIVE SPECTRUM SENSING IMPROVEMENTS

In cooperative spectrum sensing (CSS), the final decision on the global state of the spectrum is made in a fusion center (FC), which collects data from collaborating SUs present in the network. In general, SU nodes may either deliver raw sensed data (e.g., the value of measured energy) or some local decisions, and the role of FC is to process the delivered data in order to make reliable decisions. In the final step, the decision on spectrum occupancy may be sent back to the interested SUs. In such an approach, the decision-making process is more robust against the negative impact of a wireless channel on the sensing process. Traditionally, FC uses AND, OR, or k-over-n rules to make a decision whether the spectrum is occupied or free. In cooperative spectrum sensing ML may be used both in each node independently to improve sensing or local-decision making process, or in the FC to enhance the performance of the whole system. As the single node case has been discussed in the previous subsection, we will now focus on AI/ML application at the FC node.

1) CLASSIFICATION METHODS APPLIED AT FC

Similar to the single node sensing improvements, also in the collaborative case various AI-based classification methods have been proposed in the vast literature. To improve the performance of FC, many papers proposed the use of ML algorithms. Let us start with [109], where the focus was on reducing the cooperative overhead, such as overlong sensing time and energy consumption, by introducing SUs grouping algorithms. To achieve this goal, the SVM algorithm is implemented. The proposed ML framework consists of

four modules: an SVM training module, an SVM classification module, a user grouping module, and a group scheduling module. The ML algorithm is trained and tested based on the input energy vectors feature dataset. The training module is responsible for training the ML algorithm. The classifier can determine whether a given energy vector implies that a spectrum is occupied or free. The user grouping module groups users into different subsets depending on the usefulness of information received from users, for example, redundant SUs, SUs that suffer from severe fading, malfunctions, etc., are not included in sensing. A similar approach has been investigated in [106], where the authors proposed two-hybrid adaptive boosting (AdaBoost) algorithms, i.e., the algorithms where the so-called weak learners deliver their outputs to one entity who combines them into a weighted sum and produces boosted classifier. The first method is a decision stump-based AdaBoost, whereas the second is an SVM-based AdaBoost algorithm. The results presented in the paper were compared with SVM, kNN, K-means, and OR and AND rules.

SVM classifier is also considered in [65], where the aim was to alleviate the noise uncertainty effect by applying a novel ML algorithm called Fuzzy SVM with a nonparallel hyperplane (NP-FSVM). The authors especially emphasize the important fact that the noise level is usually unknown to SU. The proposed algorithm reduces the effect of noise on feature data by introducing the probability of each data and double hyperplanes for representing value deviations.

A two-stage cooperative SS is presented in [128]. In the first stage, offline training is performed. In the second, online classification takes place. The main classification algorithm

is K-means clustering that groups feature data into two categories: occupied and free channel. The features are extracted by principle component analysis (PCA). Another approach has been verified in [110], where a learning-based NB classifier is used to tell whether the channel is occupied or free.

The performance of three kinds of classifiers in the context of cooperative sensing is provided in [129], namely SVM, kNN, and NB. As feature vectors, energy levels are used. The ML task is to determine whether a given feature vector means that the spectrum is available or not. However, the final decision on channel availability is made by weighted voting using the results of all three classifiers. Combining results from many classifiers can compensate for the differences of performance of each classifier, as each of them concentrates on different aspects of data. The proposed weighted voting method is a particle swarm optimization method (PSO). The PSO method determines the weights for each ML classification decision and combines the weighted decisions linearly. As a verification, the results of the classification error rates of the proposed method and of separate ML algorithms are compared.

Another comprehensive comparison is presented in [130], where various unsupervised and supervised ML techniques in CSS are evaluated for a fixed received SNR. As previously, the vector of energy levels estimated by the devices is treated as a feature vector and supplied as input to the classifier which determines whether the channel is occupied or not. As unsupervised methods, the authors considered K-means clustering and GMM, whereas support vector machines and weighted kNN have been employed as supervised methods. The performance of each classification technique has been quantified in terms of the average training time, sample classification delay, and ROC curve. The authors found that spectrum sensing methods based on kNN and SVM are more adaptive to changing signal environments; SVM performed better than the kNN method, whereas K-means clustering performed better than the Gaussian mixture model. In [131], kNN, SVM, NB, and DT classifiers are trained over a set containing energy test statistics of PU channel frames. The simulation results show that the ML classifier-based fusion algorithm has the same accuracy as the conventional fusion rules with shorter sensing time, overheads, and extra operations.

kNN for cooperative spectrum sensing has also been employed in [55] as a counting mechanism. A global energy detection threshold for different rules of decision combinations in FC is proposed, which does not take into consideration the weight of individual SUs and their performance history. In [56], kNN is used in building a TV white space database for the reconstruction of the missing spectrum sensing points. kNN determines a label based on the majority labels of the neighboring data points. In [57], a CSS scheme based on kNN is proposed. In its training phase, each SU produces a sensing report, and local decisions are combined by majority voting at FC. At each SU the global decision is compared to the actual PU activity, which is ascertained

by an acknowledgment signal. In the classification phase, the sensing reports are sorted to sensing classes using kNN. To accurately calculate the distance between the current sensing report and existing members of the sensing classes, Smith-Waterman algorithm is used. Each SU is assigned a weight based on its effectiveness. The scheme has good performance even at low values of SNR in a fading environment.

An interesting approach was presented in [99], for the CSS framework with mobile SUs based on non-parametric Bayesian machine learning. There, the beta process sticky hidden Markov model (BP-SHMM), is introduced for capturing the spatial-temporal correlation in the data collected at different times and locations by various SUs. Bayesian inference is then carried out to group sensing data into different classes in an unsupervised manner, where the spectrum data in each class share a common spectrum state. Based on the classification results the locations of PUs together with their transmission ranges are inferred by the Levenberg-Marquardt algorithm.

Finally, let us mention the application of an SVM-based cooperative decision scheme for spectrum sensing in vehicular communication environment to mitigate shadowing and multipath fading, as discussed in [66]. In the proposed scheme, individual vehicles perform sensing using energy detection and the local results are sent to a central node which constructs vectors of energy levels for classification. The proposed SVM-based sensing performs better than the hard fusion combining rule in a low SNR regime.

2) APPLICATION OF ARTIFICIAL NEURAL NETWORKS AT FC

Apart from traditional classification methods, various types of artificial neural networks have been considered as promising tools for cooperative spectrum sensing improvements. In [107], three fusion methods were considered: conventional CSS model with hard fusion rules, ML-based fusion, and a cluster-based model. In the conventional fusion model, all SUs collect energy detection results, and send them to FC to evaluate global results using one of the rules: AND rule, OR rule, or majority rule. In ML-based fusion, ANN is proposed as a decision-making algorithm. As input features, energy detection decisions and SU locations are used. In the clustering model, two different fusion models are considered. In the first model called OR-OR fusion, OR decisions are made in a given cluster head, and then, globally in FC. Similarly, the second model employs ANN at both cluster-fusion levels.

The extreme learning machine (ELM), being a sort of ANN devoted to high learning speed applications, has been discussed in [108]. The authors considered a CR network with multiple PUs, where each PU transmits in a separate channel. FC in the proposed system receives energy vectors of length N , consisting of energies calculated by N SUs. The FC's task is to match those vectors with sets of output values that give information which channels are occupied by PUs. ELM was compared with traditional SVM results. Next, in [132] the ensemble learning (EL) framework for CSS has been

adopted in an OFDM signal-based cognitive radio system. The spectral coherence density is provided as input and is classified by CNN locally at each SU. The SUs are considered as weak learners and the stacking strategy in EL is adopted for FC to integrate their results. For this task, another deep neural network learner is used in FC. Finally, [72] proposed a deep CNN for combining sensing results in cooperative spectrum sensing. The strategy for combining single-node sensing results of SUs is learned autonomously with CNN using training sensing samples. Single-node sensing results from different bands and SUs constitute two-dimensional input data for CNN, thus both spectral and spatial correlation of single-node sensing outcomes are taken into account. The proposed scheme can achieve higher sensing accuracy than the K-out-of-N scheme or a scheme based on SVM.

3) REDUCTION OF METHOD COMPLEXITY

As the complexity of various AI/ML methods may be very high, especially with the great number of data entries, there is a need of finding ways of its optimization. The authors of [64] investigated a new method of using a low-dimensional probability vector in ML classification instead of an N-dimensional feature vector. This method is supposed to shorten the training and classification time. The probability vector is represented as a vector of two values of the probability density function of an energy vector under the condition of PU's signal present and not present, accordingly. Those feature vectors are used in two ML algorithms to classify the spectrum as occupied or free, namely, K-Means clustering and SVM algorithms. Another method which achieves a substantial training time reduction is presented in [133], where SVM-based FC soft decision algorithms are proposed. One of them keeps a constant P_{fa} , the other enables adapting the value of P_{fa} . Both of them focus on redefining the problem of finding a decision boundary in the SVM algorithm in order to make the training process faster. The results are compared with traditional SVM algorithm's performance.

Next, the authors in [82] proposed a reinforcement learning-based cooperative sensing (RLCS) method to address the cooperation overhead problem and to improve cooperative gain in cognitive radio networks. In RLCS, SU acting as the fusion center is represented as a decision-making agent that interacts with the environment consisting of cooperating neighbors and their observations of PU activity. The authors utilize temporal-difference (TD) learning to address cooperation overhead issues and show that the optimal solution obtained by the RLCS approach improves the detection performance under correlated shadowing while minimizing the control channel bandwidth requirement. RLCS converges asymptotically with the option of optimal stopping for fast response in a dynamic environment, mitigates the impact of control channel fading, improves the reliability of user and sensing data selection, and adapts to PU activity changes and the movement of SUs.

4) OTHER AI/ML TOOLS CONSIDERED

Again, the variety of AI/ML-based schemes applied to CSS is very high. For example, in [134] the authors investigated distributed algorithms using no-regret methods to detect malicious and incapable secondary users in collaborative spectrum sensing. In [135], a linear fusion rule for CSS is developed, and to obtain linear coefficients, the Fisher linear discriminant analysis has been used. In [136] the authors proposed a distributed multi-agent, multiband reinforcement learning-based sensing policy. The sensing policy employs SU collaboration with neighbor SUs through local interactions. Whereas in [137], EM algorithm for the detection of PU in multi-antenna cognitive radio networks was investigated. The PU signal is detected and the unknown channel frequency responses and noise variances over multiple subbands are jointly estimated iteratively. A distributed implementation of the proposed scheme to reduce communication overhead is researched.

In order to concisely summarize the above-discussed papers, a dedicated table has been created - see Table 5 which contains the summary of papers regarding decision making in FC.

5) OBSERVED TRENDS

Being an extension of single-node spectrum sensing, cooperative spectrum sensing benefits in a similar way from the application of AI/ML tools. Mainly, the ultimate goal of most of the considered solutions is to improve the performance of PU presence or absence detection. However, again, as in previous cases, besides the knowledge of the PU activity, also other kinds of information can be fetched. In particular, the presence of multiple sensing nodes allows for the deduction of the PU signal source location as well as the number of PU signals. Moreover, the quality of the observed signal can be estimated in a much better way, when AI/ML tools are applied for data collected from many sources.

V. DETECTION OF SIGNAL FEATURES

Aside from signal detection, it is beneficial to detect features of received signals to gain further context information. Different technologies of transmission can be recognized and used, for example, in cognitive radio to establish whether the received signal belongs to PU or SU. Knowing the difference between PU and SU signal allows for protecting PU transmission. Another example of the application of feature detection are military systems for recognizing friendly and hostile transmission.

In this section, we will focus on the intelligent distinction of different wireless systems which is achieved using modulation recognition and different transmission technologies recognition.

A. MODULATION RECOGNITION

Methods for automatic modulation recognition (AMR) can be divided into two major groups: likelihood-based (LB)

TABLE 5. ML for decision making in fusion centers in cooperative spectrum sensing.

Paper	Number of PUs and SUs	Sensed signal type	Channel type	Blind sensing	ML input data	ML method	ML used for:	Gained context information
[55]	1 PU, N SUs	PU's signal modeled as Gaussian process	AWGN	No	Energy values used (input features), energy detection decisions (labels during training)	kNN, SVM, NB, DT	Decision making in FC - PU presence detection	spectrum state
[56]	M PUs, N SUs	TV broadcasting	AWGN	Yes	Collected signal power values in different locations	kNN	Completing the missing spectrum points	Spatial power map, spectrum white space database
[57]	1 PU, N SUs	-	AWGN	Yes	Energy detection decisions	kNN	Local classification of sensing data into PU absent or PU present classes	Spectrum state
[64]	1 PU, N SUs	-	AWGN	No	Probability vector (probabilities of signal presence/absence)	K-means, SVM	Decision making on spectrum occupancy in FC	Spectrum state
[65]	M PUs, N SUs	-	Path loss, AWGN	Yes	Energy detection test statistics	NP-FSVM (fuzzy SVM with nonparallel hyperplane), for comparison: SVM and K-means clustering	Alleviating noise uncertainty effect, decision making on spectrum occupancy in FC	Spectrum state
[66]	1 PU, N SUs	TV	AWGN, fading	No	Energy detection decisions collected from SUs	SVM (RBF-SVM, polynomial-SVM, linear-SVM)	Decision making in FC - PU presence detection	Spectrum state
[72]	1 PU, N SUs	-	AWGN, path loss, fading, shadowing	Yes	Energy detection decisions collected from SUs	CNN	Decision making on spectrum occupancy in FC	Spectrum state
[82]	1 PU, N SUs	-	AWGN, fading, shadowing	Yes	Local sensing decisions collected from SUs	Proposed reinforcement learning algorithm (Q-learning, Sarsa, Action-Critic-based)	Decision making on spectrum occupancy in FC	Spectrum state
[99]	M PUs, N SUs	-	AWGN	Yes	Energy values vectors collected from SUs (energy values in different time moments)	Proposed beta-process sticky hidden Markov model (BP-SHMM)	Defining number and type of different spectrum states	Spectrum state and information on PUs' locations
[106]	1 PU, N SUs	-	Path loss, AWGN	Yes	Energy values measured by SUs	Hybrid SVM based AdaBoost and decision stumps based AdaBoost	Decision making on spectrum occupancy in FC	Spectrum state
[107]	1 PU, N SUs	PSK	AWGN	No	Energy detection decisions collected from SUs	ANN	Decision making on spectrum occupancy in FC and in cluster head	Spectrum state
[108]	M PUs, N SUs	-	AWGN, fading	Yes	Energy values collected from SUs	ELM	Decision making on spectrum occupancy in FC	Spectrum state
[109]	1 PUs, N SUs	-	AWGN	Yes	Energy values collected from SUs	SVM	Grouping SUs in order to make cooperation more efficient, decision making on spectrum occupancy	Spectrum state, the quality of a given SU's data
[110]	1 PU, N SUs	-	AWGN, path loss	No	Normalized energy values collected from SUs	NB	Estimating channel occupancy probability	Spectrum state
[129]	M PUs, N SUs	-	AWGN, path loss	No	Normalized energy values collected from SUs	SVM, kNN, NB	Decision making on spectrum occupancy in FC	Spectrum state

Continued on next page

TABLE 5. (Continued.) ML for decision making in fusion centers in cooperative spectrum sensing.

Paper	Number of PUs and SUs	Sensed signal type	Channel type	Blind sensing	ML input data	ML method	ML used for:	Gained context information
[128]	1 PU, N SUs	-	AWGN	Yes	Data extracted by applying PCA	K-means	Decision making on spectrum occupancy in FC	Spectrum state
[130]	M PUs, N SUs	-	AWGN, fading, shadowing	No	Normalized energy values collected from SUs	kNN, SVM, K-means, GMM	Decision making on spectrum occupancy in FC	Spectrum state
[131]	1 PU, N SUs	-	AWGN	No	Energy detection decisions collected from SUs	kNN, SVM, NB, DT	Decision making on spectrum occupancy in FC	Spectrum state
[133]	1 PU, N SUs	-	AWGN, fading	Yes	Power Spectral Density	two SVM-based algorithms	Decision making on spectrum occupancy in FC	Spectrum state
[134]	M PUs, N SUs	-	-	-	Sensing decisions collected from SUs	No-regret learning	detecting malicious SUs	Knowledge of the presence and number of malicious users
[135]	M PUs, N SUs	-	AWGN, fading, shadowing	No	Energy detection decisions collected from SUs	Fisher linear discriminant analysis	Determining linear coefficients for all SUs sensing decisions in cooperative sensing	Spectrum state and PUs' locations
[136]	M PUs, N SUs	-	AWGN	No	Energy detection with soft combining decisions collected from SUs	Proposed multiagent reinforcement learning	Determining the free frequency bands	Spectrum state

methods and feature-based (FB) methods. LB methods suffer from high complexity due to the number of unknowns that must be integrated into the likelihood function [58]; they typically require buffering a large number of samples [138]. The loss in performance of such methods may be impacted by phase/frequency offset, residual channel effect, and timing errors [58]. In recent years most of the research has been focused on feature-based methods. This is influenced by the ease of application within the AMR domain, unnecessary prior knowledge about the received signal, and lower computational complexity [58].

The main problem in AMR is defining an ML input feature set that accurately reflects the characteristics of different types of modulation. This feature set should be resistant to the changing radio channel and noise. A combination of feature subsets based on different feature extraction methods is used and a broad list of different machine learning methods is employed in feature-based AMR methods.

A popular type of input dataset is higher-order statistics (HOS). A feature set based on two types of fourth-order and eighth-order HOSs has been employed in [58]. Here, a simple kNN algorithm has been used to classify detected signals according to their modulations (phase shift keying, PSK, or QAM). The second, fourth, sixth, and eighth-order HOSs have been employed with three types of DTs by Subbarao and Samundiswary [62]. The different decision trees used are fine tree (FT), medium tree (MT), and coarse tree (CT). FT and MT have been proven to achieve better results than CT. In addition to HOS, Hazar *et al.* [139] proposed to

use spectral features. This decision has been motivated by the instantaneous amplitude and phase of the signal containing information about modulation type. The author used their dataset on many types of ML, e.g. ANN, SVM, RF, kNN, Hoefding tree, logistic regression, NB, gradient boosted regression trees (GBRT). Another addition to HOS values has been proposed by Xiong *et al.* [140]. The authors proposed to extract local patterns from IQ data by employing a Fisher Kernel framework which can capture non-linearity in underlying data. An SVM algorithm was used to classify modulations.

Modulation information can also be inferred from cyclostationary properties of signals. This approach has been presented in [141] where cyclostationary features based on the spectral coherence function (SCF) and cyclic domain profile (CDP) were used as ML input data. Many different types of modulations have been classified using a feed-forward back propagation neural network. Another example of employing cyclostationarity into AMR has been presented by Li [142]. In the proposed AMR method the cyclic spectra of modulated signals are calculated and then denoised. A deep CNN model is trained with the denoised spectrum images to learn features automatically and identify modulation types.

Another interesting approach to AMR using signal features has been presented by Yang [143] who proposed the expanded neural networks (ENN), which are established through the energy natural logarithm model and are using amplitude, phase, and frequency sub-networks.

A more direct use of collected samples can also be useful in AMR. Fu *et al.* [144] proposed to use normalized frequency spectrum vectors, time-domain normalized vectors, and normalized higher-order spectral vectors. The DL algorithm with the Restricted Boltzmann machine for pre-training has been proposed as an AMR ML method. It has been presented in the paper that to recognize modulations of the FSK type, the spectrum information is sufficient. The modulations ASK and QAM also need time-domain amplitude data to be categorized correctly, and modulations of the PSK group need all three types of information.

Popular ML algorithms used for AMR utilize CNNs. There are some visible trends of using collected samples directly as input of CNN. Zhou *et al.* [73] proposed an algorithm where CNN is applied to explore features in collected signal samples. The CNN output is treated as a new collection of data that can be used as input features for the classification algorithm. The authors use SVM as an algorithm that classifies the data created by CNN into different modulation categories. The results have been compared with a classification performed directly by CNN with softmax layer as the last layer.

As CNN algorithms are usually used on 2-dimensional data, there is a need to create two-dimensional datasets. Cheong *et al.* [69] used raw IQ samples as CNN input. To achieve 2D datasets, the in-phase part and quadrature part of the samples create an extra dimension. A similar approach has been presented in [145], where IQ samples are both obtained via simulation and measurements. The classification is performed by CNN and a residual network. Using the IQ samples directly as ML input data can achieve good results, however, the errors introduced during signal capturing, e.g. frequency offset, sampling rate offset, can deteriorate AMR. In response to this issue Li *et al.* [146] proposed an architecture that reduces the impact of the previously mentioned offsets before modulation classification. It is done by creating a sequential model consisting of two parts. The first part, a network-based signal spatial transformer module adaptively corrects estimation errors. The second part consists of CNN.

Another way of feeding the CNN's input with image-like data has been proposed by Peng *et al.* [147]. The idea is to use collected IQ samples to create constellation diagrams and use them as ML input data. The authors noticed that the proposed CNN-based approach may not always outperform existing methods, due to information loss in the limited resolution of images obtained using a data conversion procedure from complex samples to images. The benefits of modulation constellation have been exploited in [148] where three input datasets have been tested. The first one consists of IQ constellation points, the second one consists of centroids of those constellation points obtained by the C-means algorithm, and the last one is HOSs of received samples. The proposed ML method consists of two independent layers of autoencoder (AE) based DNN.

A kind of variation of popular CNN algorithms is a convolutional long short-term deep neural network (CLDNN).

Liu *et al.* [149] compared a few deep learning networks, including residual networks (ResNet), densely connected networks (DenseNet), and CLDNN, and concluded that CLDNN outperforms other network architectures and the research on network depth shows that the number of convolutional layers does not improve AMR. CLDNN performance has also been compared to ResNet and LSTM by Ramjee *et al.* [79]. The results indicated a better performance of CLDNN and ResNet deep neural network architectures at low SNR, and LSTM and ResNet architectures at high SNR. The authors focused on reducing the time of training by minimizing the size of the training dataset, reducing ML input data dimensions through PCA and subsampling.

In addition to single-carrier systems to which the above articles relate, AMR methods can also be used in multi-carrier systems, such as OFDM systems. The aim of AMR in multi-carrier systems can be the identification of combined modulation parameters of the carrier number and modulation type. Liu *et al.* [70] proposed the DNN model, which is divided into two sub-networks, one for the detection of the number of active carriers and one for the detection of used modulation. Another ML input feature set was proposed by Keshk *et al.* [150]. The authors performed AMR in the OFDM system, using discrete transforms and cepstral coefficients (Mel-frequency cepstral coefficients (MFCCs)). The Modulation order has been determined by extracting cepstral features from the signal after the application of transformers, such as discrete wavelet transform (DW), discrete cosine transform (DCT), and triangular transform with common properties with the DCT (DST). A classification of the defined features using the SVM algorithm for different modulation types has been performed. Other problems in multi-carrier modulation systems can be distinguishing different multi-carrier waveforms along with the used modulation or distinguishing single-carrier signals from multi-carrier. Duan *et al.* [80] focused on detecting signals that use OFDM, universal-filtered multi-carrier (UFMC), and filterbank-based multi-carrier (FBMC). Instead of using IQ samples as input data, the amplitude values extracted from PCA output have been employed as input for the CNN algorithm. The authors showed that this dataset gives better results than just raw IQ data. Norouzi *et al.* [151] on the other hand, focused on differentiating OFDM signals from single-carrier signals. The key features used in the proposed ML algorithm (modified K-means algorithm) are two statistics of the amplitude of the received signal that is calculated at the output of a quadrature mixer.

The comparison of all of the above-mentioned papers is presented in Tab. 6.

In the above-mentioned papers, there are some trends visible regarding data used as ML input. The main categories of input data that can be observed are higher-order statistics, cyclostationary properties of received signals, measured signal information as frequency spectrum obtained via FFT, direct use of collected IQ samples (or their slight modification, e.g., in the form of 2D images), and constellation

TABLE 6. ML for automatic modulation recognition.

Paper	Modulation type	Channel type	Feature set	ML method	Signal information
[58]	(2, 4, 8)-PSK, (16, 64, 256)-QAM	AWGN, fading	HOS (fourth-order, eighth-order)	kNN	Simulated
[62]	(2, 4, 8)-PSK, (4, 16, 64)-QAM	AWGN	HOS (second, fourth, sixth and eighth-order)	DT algorithms: FT, MT, CT	Simulated
[69]	(2, 4, 8)-PSK, (16, 64)-QAM	AWGN	IQ samples	CNN	Simulated - IEEE 802.11ac radio signals
[70]	(2, 4, 8)-PSK, 16-QAM	AWGN, fading	IQ samples	DNN	Simulated - multi-carrier signal (OFDM)
[73]	(2, 4, 8)-ASK, (2, 4, 8)-FSK, (2, 4, 8)-PSK, (4, 16, 64)-QAM, OFDM, LFM, MSK	AWGN	Collected signal samples	CNN and CNN-SVM	Simulated
[79]	(2, 4, 8)-PSK, (16, 64)-QAM, (2, CP)-FSK, 4-PAM, WBFM, AM-DSB	AWGN, fading	IQ samples	CNN, DenseNet, CLDNN, LSTM, ResNet	Simulated - RadioML2016.10b
[80]	OFDM-QAM, UFMC, FBMC-OQAM	AWGN	Amplitudes extracted from PCA output	CNN, for comparison: expert-based SVM and expert-based kNN	Simulated - multi-carrier signal (OFDM, UFMC, FBMC)
[139]	(2, 4, 8)-ASK, (2, 4, 8)-PSK, (16, 64)-QAM	AWGN, fading	HOS and some spectral features	ANN, SVM, RF, kNN, Hoefding tree, logistic regression, NB, GBRT	Simulated
[140]	(4, 8)-PSK, (8, 16, 64)-QAM	AWGN	HOS, local patterns based on the Fisher kernel	SVM	Simulated and measured (USRP testbed)
[141]	(2, 4, 8, 16, 32, 64)-PSK, (2, 4, 8, 16, 32, 64)-QAM, (2, 4, 8, 16, 32, 64)-FSK, (2, 4, 8, 16, 32, 64)-PAM	AWGN, fading	Cyclostationary features based on SCF and CDP	Feed-forward back propagation neural network (FF-BPNN)	Simulated
[142]	(2, 4)-PSK, (2, 4)-FSK, MSK, AM, FM	AWGN	Denoised cyclic spectrum images	AN-SF-CNN - antinoise processing and deep sparse-filtering convolutional neural network	Simulated
[143]	(2, 4, 8)-PSK, 16-QAM	AWGN, fading	IQ samples	ENN	Simulated
[144]	(2, 4, 8)-ASK, (2, 4, 8)-PSK, (2, 4, 8)-FSK, (16, 64)-QAM	AWGN	Frequency spectrum data, time-domain amplitudes, and normalized higher-order spectral data	DL with Restricted Boltzmann Machine for pre-training	Simulated
[145]	OOK, (4, 8)-ASK, (2, 4, 8, 16, 32)-PSK, (6, 32, 64, 128)-APSK, (16, 32, 64, 128, 256)-QAM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC, FM, GMSK, OQPSK	AWGN, fading	IQ samples	CNN, ResNet	Simulated - RadioML, measured-B210 SDR
[146]	(2, 4, 8)-PSK, (16, 64)-QAM, (G, CP)-FSK, 4-PAM	AWGN, fading	IQ samples	Spatial transform method (STM) + CNN model	Simulated - RadioML2016.04c
[147]	(4, 8)-PSK, (16, 64)-QAM	AWGN	Sample points converted into a constellation diagram	CNN (AlexNet model)	Simulated
[148]	(2, 4, 8)-PAM, (2, 4, 8)-PSK, (16, 64, 256)-QAM	AWGN, flat-fading	IQ constellation points / centroids of the constellation points / HOS	DNN with two independent AEs	Simulated
[149]	(2, 4, 8)-PSK, (16, 64)-QAM, (B, CP)-FSK, 4-PAM, WBFM, AM-DSB	AWGN, fading	IQ samples	CNN, ResNet, DenseNet, CLDNN	Simulated - RadioML2016.10b
[150]	(2, 4, 8, 16, 64)-PSK	AWGN, fading	Cepstral features	SVM	Simulated - multi-carrier signal (OFDM)
[151]	4-FSK, 4-PSK, 16-QAM, OFDM	AWGN	Kurtosis, amplitude	K-means	Simulated - multi-carrier signal (OFDM) and single-carrier signals

diagrams and information regarding them. The summary of those popular input datasets along with their corresponding and most popular ML algorithms is presented in Fig 5. It can be observed that the DL method, especially CNN, is very popular. The more complex the ML algorithm, the simpler input datasets can be used, and the simpler

the ML method, the more processed the input dataset has to be.

To sum up, AMR is a very popular topic with multiple different approaches examined and available to be compared. One can observe that single-carrier signals are dominant and there is much less research on multi-carrier signals. The

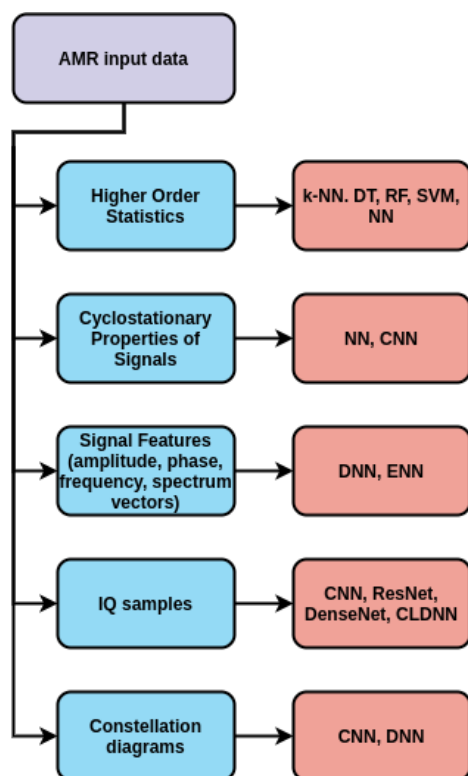


FIGURE 5. AMR typical input datasets with their corresponding most popular ML algorithms.

more complex ML algorithms, such as DL algorithms are becoming more popular, as their input data requires less preprocessing. The AMR methods can be a good choice for PU/SU differentiation, when different modulations are used by those users. If the modulation information is not enough to recognize PU and SU signals, there is an option to analyze received signals even deeper and try to identify different transmission methods. This approach is presented and compared in the next part.

B. SIGNAL IDENTIFICATION

Although most of the feature detection papers concern AMR, some examples focus mainly on finding features in signals that can clearly specify the type of the telecommunication signal. This topic is much less examined in papers, however, it is a great tool to differentiate PU and SU signals, as well as perform simple signal detection.

Subekti *et al.* [67] proposed an algorithm for SS and detected signal classification to determine whether the present signal belongs to PU or another SU. Two ML algorithms have been used, namely, the deep autoencoder NN algorithm for signal feature learning, and the SVM algorithm for classifying autoencoder output into PU and SU signal categories. As ML input dataset, two sets of images were used: spectrograms and images created from amplitude and phase difference. The PU signal is a radar signal, while the SU one is an LTE transmission. Another issue of signal differentiation

has been addressed in [152]. Here, the authors (Alhazmi *et al.*) emphasized the need for the coexistence of three cellular systems, namely 3G, 4G, and 5G, while the gradual transfer to the 5G system progresses. On this basis, the new system joining the previous systems can be considered as secondary. As a solution for signal (UMTS, LTE, and 5G) differentiation, a CNN algorithm has been implemented.

A CNN algorithm has also been used in [153] for the identification of radar signals when they can overlap with LTE signals and WLAN. Here, the radar is a PU and the WLAN and LTE are SUs. The collected data used for training is processed, so the amplitude and phase shift properties of radar signals are emphasized. This approach allows the CNN algorithm to perform better.

Tekbiyik *et al.* [154] focused on using the SVM algorithm to identify interferences coming from cellular signals (GSM, wideband code division multiple access, WCDMA, and LTE). As the input data for SVM classification, FFT of the signal, auto-correlation function, power spectral density, and spectral correlation function are utilized. The results of SVM classification are compared to DL neural network results, namely, CLDNN and LSTM.

The last two papers concern the identification of IEEE 802.11 and IEEE 802.15 signals. Schmidt *et al.* [155] applied a deep CNN to distinguish between 15 signal classes, which represent packet transmissions of IEEE 802.11 b/g, IEEE 802.15.4, and IEEE 802.15.1 signals. On the other hand, Rajab *et al.* [156] proposed using two simpler ML methods, namely, kNN and NB for 802.11 b/g/n classifications. Here, the energy detection method is implemented to collect channel activity/inactivity characteristics by constructing histograms. Knowing the temporal distributions, a combination of features is derived that can characterize the wireless technologies. The features include some statistical information on the distributions.

Table 7 contains a summary of the considered papers. By analyzing the table, it can be concluded that for signal identification, the SVM and CNN algorithms are especially popular. CNN is able to analyze spectrograms and find complex features that are characteristic of a given signal. On the other hand, SVM is a simple algorithm that works on less complex datasets and is particularly useful if some input data preprocessing is performed, like for example FFT or correlation functions calculation. It can be observed that different signal identification is much less popular than AMR methods.

VI. USER LOCALIZATION

The Knowledge of user's self-location and positions of other transmitting devices can be utilized in context-awareness in many ways. A user's location can be vital to determining a spectrum occupancy state, either by spectrum sensing or by using a spectrum database. Some locations are more probable to be crowded with transmitters, and some might be known for their smaller user density, which is relevant information for spectrum reuse. The information on the position might also be important in channel propagation conditions

TABLE 7. ML for users/signals classification.

Paper	Signal type	Channel type	Feature set	ML method	ML used for:
[67]	Radar signal (PU), LTE (SU) - measured	-	Spectrogram / image created from amplitude and phase difference	Deep autoencoder NN, SVM	Signal classification into two categories: PU's signal and SU's signal
[152]	UMTS, LTE, 5G - simulated	AWGN, fading	Signal samples in form of 2D images	CNN	Signal classification
[153]	Radar, WLAN, LTE - measured	AWGN, fading	Spectrograms, amplitude and phase difference representations	CNN	Detection of radar signals (PU)
[154]	GSM, WCDMA, LTE - measured	AWGN, fading	Outputs of SCF, FFT, autocorrelation function, power spectral density of the received signal	SVM, DLNN (CLDNN, LSTM)	Signal classification and spectrum sensing
[155]	IEEE 802.11 b/g, IEEE 802.15.4, IEEE 802.15.1 - measured	AWGN, fading	IQ samples	CNN	Signal classification
[156]	IEEE 802.11 b/g/n - measured	AWGN	A set of statistical features derived from activity distribution	kNN, NB	Signal classification

estimation. When it comes to locating other users, it is important to have a general overview of other SUs and most importantly PUs to protect their transmissions and minimize the level of interference.

A. TRADITIONAL METHODS OF LOCALIZATION

The most popular positioning techniques are based on signal measurements and can be classified into the following categories: trilateration, triangulation, proximity, scene analysis, and hybrid methods [9]. The most popular positioning systems - global navigation satellite systems (GNSS) use the trilateration (multilateration) method, namely the time of arrival (ToA) for positioning. Other examples of this method are the time difference of arrival (TDoA) and RSS measurements.

In cellular communication systems, a cell ID is a popular method of user localization. It is an example of the proximity method, where the location is estimated as the position of a known transmitter - in the case of a cellular network, the known transmitter is a base station.

Databases are widely used for scene analysis location methods. To obtain knowledge on location, measurements of the channel and received signals are compared with data stored in the database of *fingerprints*. Each fingerprint characterizes a given location.

In triangulation, the angle of arrival (AoA) of the received signal is used. Last but not least, in hybrid methods, a combination of any two or more localization methods can be used to achieve expected results.

The above-mentioned methods are useful mostly for self-localization, although the information on single-node user location can be stored and shared using databases. Other methods of self-positioning and other users' location acquisition can be based on machine learning algorithms which use collected signals and channel data to predict the distribution of transmitting devices in space, frequency, and time.

B. MACHINE LEARNING-BASED SELF-LOCALIZATION

Self-localization can be carried out using ML techniques either directly on the available data or by improving existing localization methods. Self-positioning can be divided into outdoor localization and indoor localization.

Indoor positioning is a subject that is becoming more and more popular. The existing methods usually employ RSS and fingerprinting. AlHarji *et al.* [59] proposed an algorithm for indoor localization by employing gathering data by the sensor network. The surrounding environment is identified by applying the ML method (kNN) to collected data, mainly on RSS samples. Knowing the type of surrounding environment makes it possible to choose the most appropriate selection of data that yields the highest localization accuracy. Another example of using sensor networks and RSS-based fingerprinting has been presented in [157], where an autoencoder-based deep extreme learning machine has been employed to extract features from input data collected from sensors and establish localization inside a building.

The fingerprinting-based localization methods usually employ deep learning as a good method of intricate pattern detection. There is a need to collect large amounts of data in order to train the DL algorithm. For localization purposes, crowdsourcing is a good way of gathering large amounts of data from different users. This means that some sort of a centralized server would have to collect the data from the users which poses a threat to user privacy. Ciftler *et al.* [158] proposed to use federated learning, which enables training multiple deep learning models in a decentralized manner, and then gathers the obtained models and creates one global DL model, which enables both taking advantage of many input data sources, as well as keeping user privacy.

Outdoor positioning methods employing machine learning can be considered as less popular, as usually GPS or GNSS positioning is available. But in some cases, satellite positioning system signals can be blocked or cannot be received.

TABLE 8. ML for users localization.

Category	Paper	Sensed signal type	cooperative or single-node sensing	ML algorithm	ML used for	ML input data
self-localization	[59]	(not specified) - Measured	Single-node	kNN	Identifying the type of surrounding environment, choosing the best combination of RF features for indoor localization	Received signal strength, channel transfer function, frequency coherence function
	[157]	Wireless sensor network signals - simulated	Single-node	Autoencoder based deep extreme learning machine (ADELM)	Indoor localization	RSS values
	[158]	IEEE 802.11 - simulated	Cooperative	DL (federated learning)	Indoor localization	RSS values
	[159]	OFDM - measured	Single-node	NN	Outdoor positioning improvement	Channel impulse response estimation
other users localization	[96]	Wireless microphones, TV transmitters - simulated	Cooperative	RVM, Bayesian Sparse Learning	PU signal strength, location, number of PUs	Power values [dBm] calculated by SUs
	[99]	Random signal (follows Gaussian distribution) - simulated	Cooperative	Non-parametric unsupervised Bayesian learning (named beta process hidden Markov model)	Clustering sensing data - localizing PUs and their transmission ranges	Normalized energy values calculated by SUs
	[135]	(no signal considered - only power values in space)- simulated	Cooperative	Fisher Linear Discriminant Analysis (tool used in ML)	Determining linear coefficients for all SUs sensing decisions in cooperative sensing	Energy detection sensing decisions calculated by SUs
	[160]	Simulated	Cooperative	SOM	PU transmission range estimation	Power values calculated by SUs

In [159] a case like this was considered, where a device that happens to be in need of determining its position is also in a multipath environment with no direct Line of Sight from the transmitter. The multipath components of the channel are actually used to estimate a radio map using the ray-tracing method. The proposed method combines all estimated channel impulse response, simulated ray-tracing, and machine learning (NN) which enables localization by matching the amplitude and delay information to a specific position.

The papers described above are summarized in Table 8. Usually for self-localization RSS is employed for individual localization estimation. Like in other applications, more complex ML algorithms will work well with simpler input datasets. Indoor positioning is the most popular, as it is not covered by satellite navigation systems.

C. MACHINE LEARNING-BASED OTHER USERS LOCALIZATION

Localizing other transmitters nearby can be very helpful and can provide an important piece of context information. For example localizing PUs is an important step in protecting their transmission. SU that is aware of other transmitting devices' location can determine in which areas it is safe to transmit and what power to use in order not to cause unnecessary interferences. Locating other users in spectrum awareness can therefore be defined as delimiting areas where those users' transmitting power meets some predefined conditions, such as a power level high enough to be received by the target device, a signal level exceeding some predefined threshold, etc.

Determining PU location and signal range can be achieved in multiple ways. In the simplest case, it can be carried out by sensing the PU's signal and determining its presence in a certain detection area, which gives a general idea of the PU's location. Choi *et al.* [135] presented a method of detecting PU by applying cooperative sensing. Each SU tries to detect PU signal within a detection area around itself by using gathered sensed information from other nearby SUs. In the proposed method, the SUs must evaluate received sensing results from other SUs before incorporating them into the final detection decision. This is because SUs located too far away are more likely to be less reliable in their sensing results. To properly evaluate the received sensing information, and combine in a linear fusion, the Fisher linear discriminant analysis is applied, which enables SUs to learn coefficients for the fusion.

More precise location information acquiring methods have been presented in [96]. The authors proposed a method based on sparse Bayesian learning which is able to reconstruct the power propagation map. SUs work cooperatively by gathering received power values and sending that information to a fusion center, which creates a power map. Having a piece of information like that makes it possible to not only determine the ranges of PU transmission and the power level but also to determine their quite exact location and also the number of PUs. The accuracy of the designated power map depends largely on SUs' number and their right distribution in space. The solution to the insufficient number of sensing devices may be to take advantage of the fact that SUs are mobile. This approach, presented in [99] allows finding locations of PUs by collecting data in different locations in space by moving

SUs. In the paper, many SUs collect spectrum data as time series. A global spatial-temporal correlation in the collected data is captured by a proposed algorithm based on Bayesian learning. The collected data is stored in a cluster head, which categorizes given areas in space into categories related to the presence of a signal, the number of PUs, etc.

The positioning of PUs and their transmission range can also be performed in a more indirect way. Wang and Yao [160] considered a sensor network in which each sensor (SU) senses a PU signal. In the paper, it has been shown that it is possible to determine PU's transmission range by estimating each SU's location and their sensed PU status. Sensor locations are estimated using an unsupervised machine learning algorithm called self-organizing maps [161]. Knowing the SU's location and received PU power, the PU's signal coverage boundary is derived by applying the SVM algorithm.

The summary of other users' positioning is presented in Table 8. The topic of other users localization is less discussed in the literature than self-localization. Localizing other users requires a lot of sensing data, hence all of the proposed approaches require SUs' cooperation. Popular input datasets for ML algorithms include energy or power values collected by SUs located in different places in space, sometimes also changing their position which can be beneficial for localization, as it enables collecting data from different points in space.

VII. TRAFFIC PATTERN RECOGNITION

Traffic pattern recognition is a way of recognizing statistical properties or time, frequency, and dependencies in the received signal. By detecting those dependencies it is easier to estimate the current spectrum state or predict spectrum state in the near future. Predicting the future spectrum state translates into a more efficient reuse of spectral resources, better spectrum management, shorter sensing time, and therefore also lower energy consumption. Other users' traffic patterns depend on the statistical distribution of PU activity, sensing area, time of day, telecommunication system, etc.

The other users' activity statistics are not always known, the transmitted signal is complex, and subtle transmission space-time-frequency dependencies and correlations are hard to see. ML algorithms prove to be very useful for this application.

The main methods of pattern recognition employing ML methods can be categorized into the following groups: time patterns recognition, frequency patterns recognition or spatial patterns recognition, or a combination of any of those. The combined time and frequency patterns can be seen in Figure 6. This figure shows two scenarios: one outdoors and one indoors. One can observe that the location of measurements has a great impact on the results. Last but not least, the spatial pattern in the form of different SNR values in different locations in space is shown in Figure 7. The most common causes of traffic pattern occurrence are summarized in Figure 8. Usually, ML is not used to find the patterns

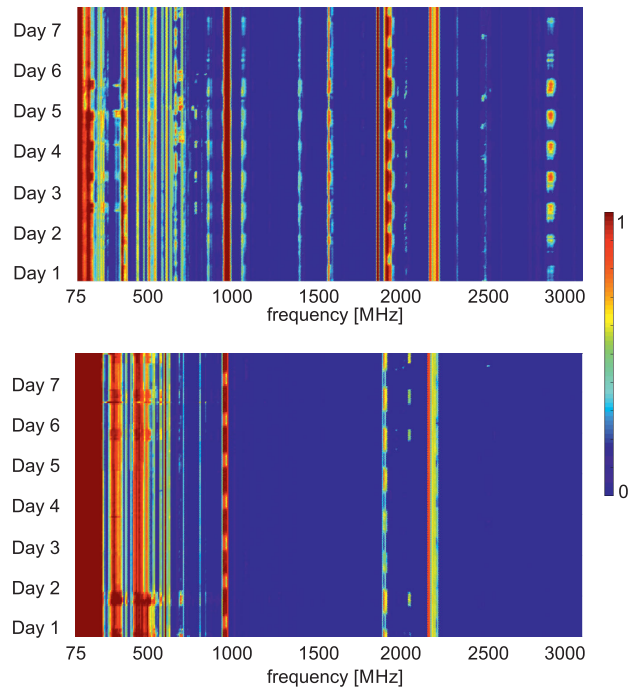


FIGURE 6. Examples of measured real-data time-frequency patterns of wireless signals: top figure - the measurements carried out outdoors (rooftop); bottom figure - the measurements carried out indoors. Measurements done by some of the authors of this paper.

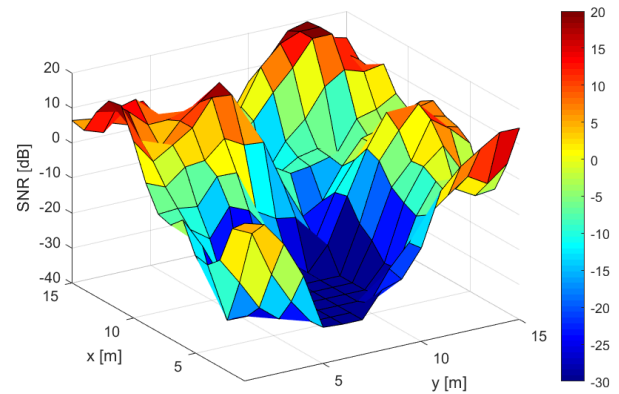


FIGURE 7. SNR values varying in space. Here, the AWGN and fading effect are causing variety of SNR values. The path loss is not taken into account.

explicitly, but rather to predict the next signal occurrence, or probability of its occurrence in the future.

All of the below mentioned papers are summarized in Table 9.

A. SENSING/PREDICTION OF SIGNALS WITH TIME DEPENDENCIES

The typical approach to finding patterns in sensed data is to look for temporal dependencies and correlations. Knowing the history of the signal occurrence in time, it is possible to improve its sensing or predict its occurrence in the future.

TABLE 9. ML for traffic pattern recognition.

paper	pattern type [T - time, F - frequency, S - spatial]	Sensed signal/ system type	Data measured / simulated	ML algorithm	ML used for	ML input data
[54]	T + F (separately for different locations in space)	LTE	simulated	kNN, Random Forest	spectrum sensing	Energy values/energy detection decision per a given resource blocks, information on energies of adjacent resource blocks, time and frequency information
[74]	T + F	IEEE 802.11	simulated	CNN	Transport protocol detection, traffic pattern, transmission rate	2D images - IQ samples in time, Short-Time Fourier Transformation of IQ samples in frequency
[75]	T (for different frequencies)	GSM1800 downlink, satellite signal	Measured	RNN	Prediction of PU's next spectrum state	Power spectral density values
[77]	T + F	3 MHz - 5.4 MHz band	measured	LSTM-based architecture	Prediction of PU's next spectrum state	Availability status of current and previous spectrum states
[162]	T	450 MHz-800 MHz band of land mobile radio	Measured	RNN, proposed ConvLSTM algorithm	Prediction of PU's next spectrum state	-
[163]	T	Generated random QPSK signal	Measured	RNN	Prediction of PU's next spectrum state	IQ samples
[164]	T	Signal generated according to Poisson distribution	Simulated	SVR-based online learning	Prediction of PU's next spectrum state	Received signal power values
[165]	T + F + S	big data in 5G	- (theoretical paper)	K-means clustering for feature extraction, other ML algorithms for spectrum prediction	Spectrum data feature extraction, spectrum prediction	(not specified)
[166]	T + F + S	LTE	Simulated	kNN, Random Forest	Spectrum sensing	Energy values/energy detection decision per a given resource blocks, information on energies of adjacent resource blocks, time and frequency information
[167]	T + F	LTE/5G, sensor network signals	Simulated	NN, RNN, CNN	Next time slot occupancy prediction	For CNN: 2D images - energy values per resource block, time and frequency information
[168]	T + F	IoT signals	Measured	DBSCAN	Spectrum sensing	Energy detection decisions
[169]	T + S	MIMO-OFDM signals	Simulated	Delayed feedback reservoir (RNN)	Spectrum sensing	Signal samples

One of the most popular ML algorithms for finding sequence dependencies is RNN. Rutagemwa *et al.* [162] employed RNN to learn the time-varying probability distribution of received power samples. RNN predicts the next samples and makes it possible to establish the suitability of sharing the channel with other users. Roy *et al.* [163] also proposed to use RNN to take advantage of the temporal statistical distribution of received signals. The received data samples are in the form of in-phase and quadrature-phase components and for each of the time intervals, several samples are collected. Therefore the collected data is in the form of two-dimensional temporal data, where the first dimension is the time dimension and the second dimension is the multiple samples per time interval. This gives it an opportunity to use CNN to take advantage of the two-dimensional form of the data. The authors propose a hybrid of RNN and CNN algorithms: ConvLSTM algorithm, which further improves prediction results. Another example of RNN usage for detection of signals correlated in time has been presented by Hamedani *et al.* [169], where a new

class of RNN, namely, a delayed feedback reservoir has been applied.

In older papers, one can find other, less complex ML algorithms employed for prediction. For example, Zhang *et al.* [164] proposed SVR-based online learning, where the past signal power measurements for a given frequency are used to establish the probability of the next spectrum occupancy state. The probabilities are then used by SU to decide which channel to select for transmission.

B. SENSING/PREDICTION OF SIGNALS WITH TIME AND FREQUENCY DEPENDENCIES

Another approach to predicting future spectrum states is to combine received signals both in time and frequency to create two-dimensional data matrices or spectrograms and use those to find combined temporal and frequency dependencies. The ML algorithm which is usually applied for this application is CNN. CNN deep learning algorithms are usually applied to analyze images, so their application in any two-dimensional datasets seems to be appropriate. Camelo *et al.* [74] used

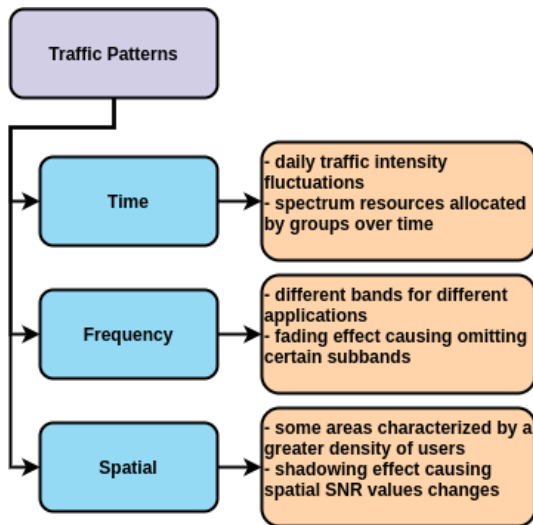


FIGURE 8. Types of traffic patterns occurring in telecommunication signals.

CNN on spectrum samples that are combined into a spectrogram images. The proposed algorithm is able to predict the next signals but also detect transport protocols and transmission rates. Wasilewska *et al.* [167] used the temporal and frequency dependencies to predict the state of a few next LTE/5G time slots. The LTE resources are allocated in bunches, so the probability of adjacent resource blocks being of the same allocated/idle state is high. Three deep learning algorithms are employed, namely deep NN, RNN, and CNN. RNN is trained to recognize the time dependencies for each of the frequencies separately, and CNN is trained to treat the spectrum energy data as two-dimensional images.

Yu *et al.* [75] used the RNN algorithm for each of the frequencies separately, to find idle spectrum in real measured GSM and satellite data. In this paper, the main focus is finding the best architecture of the DL RNN by employing the Taguchi method. Compared to [75], where RNN works separately for different frequency channels, in [77] the proposed deep learning algorithm based on LSTM can work jointly for multiple channels at the same time and predict the next time step spectrum occupancy states for those channels. Joint time/frequency sensing can be applied in IoT systems as well. As shown in [168], the transmitted IoT data is transmitted in the form of frames that create rectangular shapes in the time and frequency domains. The clustering algorithm is applied to find data points closely located in the time and frequency domains as a transmitted signal, while discarding the scattered points as falsely detected noise.

C. SENSING/PREDICTION OF SIGNALS WITH TIME, FREQUENCY, AND SPATIAL DEPENDENCIES

Last but not least, the most comprehensive solution is to combine spectrum data into three-dimensional datasets that contain information on signals in the time, frequency, and space domains. This approach enables obtaining the largest amount

of comprehensive context information, but also requires a lot of processing and calculations.

Liu *et al.* [165] proposed theoretical solutions for cooperative spectrum sensing where broadband big spectrum data are collected by many users and analyzed in Fusion Center servers. In this paper, ML algorithms are proposed to be used in two stages: in front-end ML and back-end ML. To extract relevant spectrum features, front-end ML is used, and after preprocessing the obtained feature data in order to generate time-frequency data maps, a back-end ML model obtains spectrum prediction information. In the front-end ML model, K-means clustering is proposed, which groups received data into categories of different communication systems, then for each of the categories, different back-end ML can be used according to the classified data characteristics.

Although the most intricate deep learning algorithms are popular in processing data of this complexity, simple algorithms can be employed in finding space/time/frequency dependencies as well. Wasilewska *et al.* [54], [166] focused on signal sensing for which there occur strong correlations in time, frequency and space. In [54] sensing is performed separately for different locations in space, but in [166] the learning is performed including localization information as ML input. In both of the papers, simple ML algorithms have been employed, namely, kNN and random forest.

As Table 9 shows, ML for traffic pattern recognition usually employs those detected patterns to predict the future spectrum state. The input datasets vary from simple IQ samples, to energy values, to spectrum sensing decisions. As the prediction of future spectrum state requires analyzing signals as time sequences, DL methods, especially RNN algorithms, are very popular. If the patterns are both in time and frequency, CNN algorithms are usually employed.

VIII. RADIO CONTEXT INFORMATION PROCESSING SUBSYSTEM

In the prior sections, we have discussed how AI/ML tools can be used for enriching context information in various specific domains, i.e., sensing of presence or absence of other simultaneous transmissions, detection of signal features, traffic patterns, as well as localization of other users. Of course, beside these four areas, other domains can also be identified. However, although the application of AI/ML schemes separately in each domain will benefit in better context information gathering, the true gain may be achieved by exchanging the data for their joint processing. It is evident that more coordinated processing may lead to a significant system improvement at the expense of increased control traffic and increased computational burden. Furthermore, more data to process entails the need for greater storage capacities of selected wireless nodes. Thus, one can conclude that there is a trade-off between improved information processing and increased complexity. However, access to various kinds of context information helps create a broader view, and in consequence, better adjustment of the wireless node (or the whole communication system) to the observed situation. All

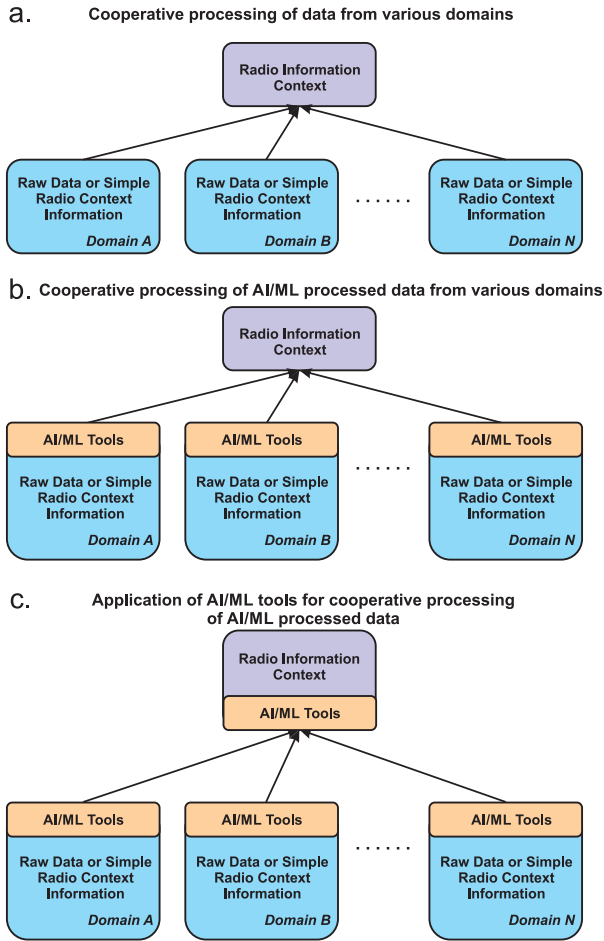


FIGURE 9. Three levels of data processing for enriching the radio context information.

sources of accurate and AI/ML improved context information can lead to the creation of an enriched radio information context. Thus, one can identify three levels of cooperative (i.e., originated from various sources and domains) data processing, as shown in Fig. 9. In the traditional way, no advanced AI/ML tools are used for the creation of the radio information context. Next, selected AI/ML mechanisms can be used to improve the quality and accuracy of data originated from a specific domain. In this paper, we have summarized the solutions applied in this approach. However, while implementing AI/ML algorithms for advanced processing of already AI/ML processed data, new possibilities are created for enriching radio context information.

Next, one can observe that various applications may benefit from accessing enriched context information, and the way they improve their functioning depends strongly on their characteristics. For example, traditional radio resource management may perform better if the knowledge of prospective traffic patterns and signal features of other users is provided in advance to advanced localization information. Next, user-association-to-base-station algorithms, typically based

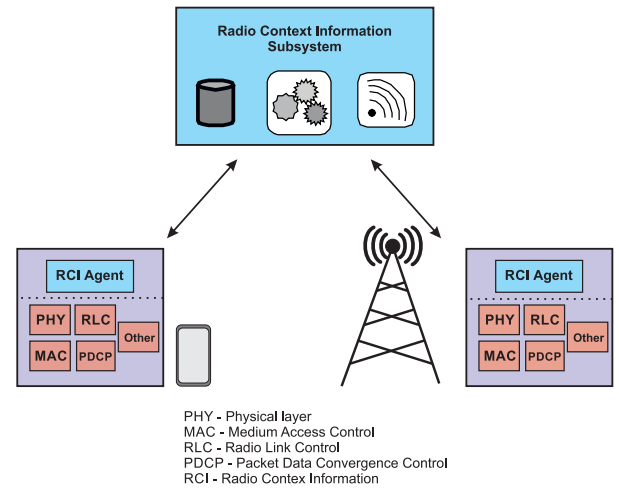


FIGURE 10. The generic idea of the radio information context subsystem.

on the strongest received signal power, may perform better if detailed localization is known, however, the knowledge of the detected modulation scheme is not that profitable here. Based on this observation we claim that there could be a dedicated radio-context-information sublayer (or subsystem) or radio context information function in the architecture of future wireless networks. This idea is visualized in Fig. 10. The role of such a subsystem or function would be to create a space for gathering, exchange and processing various kinds of context information from different places or nodes of the network. It is evident that depending on specific circumstances, different kinds of information will be accessible or may be delivered by different nodes. For example, it has been proved that in cooperative sensing the hidden-node problem may be mitigated, as with a sufficient number of sensing nodes the impact of being shadowed by obstacles is minimized. Thus, some nodes will be able to detect the presence of other ongoing transmissions, whereas others will not. Moreover, localization details about a certain, observed node (or user) may be deduced accurately based on delivered information from specific, well-located nodes, i.e., the nodes that can participate directly in the signal processing stage. The other nodes may benefit from accessing such localization information, but may not be able to contribute to the localization process. These two simple examples illustrate that cooperation and information exchange between neighboring nodes may be profitable to the system. Thus, again, cooperative information exchange and advanced processing may lead to the creation of the radio information context.

The role of such a radio context information subsystem or function may be broad. In the simplest case, such a subsystem (function) will only deliver some specific information per request. For example, a specific system entity (regardless of its placement in the network layered model) may benefit from accessing some specific kind of information. Thus, when necessary, it may request such information (or an update

of already possessed information) from the radio context information subsystem (function). If such information (or its update) is available, the requesting entity will utilize it accordingly. In the case when the subsystem (function) does not contain the requested data, it may perform some steps towards their collection from other connected nodes. Furthermore, in the case when the accuracy and quality of data possessed by the subsystem do not fit to the request, the context information subsystem may apply specific AI/ML tools (as those described in previous sections of this paper) in order to improve the data quality. Thus, the context information subsystem (function) may be reactive and apply necessary steps towards the delivery of requested data. This approach allows for such an implementation of various network functions that their behaviour could be related to the quality and amount of possessed context information. In such an approach, the context information subsystem would be equipped with a set of AI/ML supported solutions, tailored to specific application domains (e.g., improvement of localization of selected users, advanced traffic pattern detection schemes). One of the important aspects related strongly to any context-information processing is prospective data sensitivity. Detailed data management plan for end-user privacy and data security has to be applied, and much effort must be made to follow, e.g., General Data Protection Regulations addressing these aspects. Although in the context of better functioning of wireless systems, most of the context-information sources will be not sensitive from the perspective of the end-user (such as distributions of SINR values or historical frequency-channel occupancy at a given location), then data directly related to, e.g., end-user habits, have to be processed appropriately to not violate user rights, legal agreements, and societal rules. One of the popular solutions nowadays is to preserve privacy through data anonymization. From the perspective of the context information subsystem, a dedicated entity devoted to data management towards protecting end-user privacy and cyber-security has to be included. However, the context information subsystem may be created in a more advanced way, instead of being reactive it could also work in a preventing fashion. Besides providing replies to various queries from connected nodes (or some specific entities), it may proactively trace the behaviour of the nodes and prepare potentially useful information in advance, before some specific action has happened. Such an approach may be of great interest in situations where latency plays a crucial role. For example, knowing the periodic (e.g. daily) traffic patterns of some transeiving nodes in the system, the context information subsystem (function) may proactively deliver such information to the prospectively interested nodes. Advanced radio resource management functions may then utilize such information to better adjust resource allocation among users. Another example could be to compute or derive some information in a more dynamic scenario, where the nodes' mobility is considered. In particular, by observing the trajectories of moving users, the radio context information subsystem can predict their potential routes and perform necessary computations and

information processing, to deliver useful data to the nodes affected by these moving users.

Finally, in the most advanced, yet complicated scheme, the radio context information subsystem (function) could apply dedicated AI/ML tools to process, in an intelligent way, the information delivered from various specific domains, where other AI/ML tools have already been applied. It means that the heart of such a radio context information subsystem would be an advanced processing engine, able to derive advanced information as a result of efficient management of big data collected from various sources. And what is important, these kinds of data would be available on request, in a preventing fashion, but also may be processed proactively. This third operating mode can be treated as an extension of the preventing mode just described by adding additional intelligence to predict and foresee, some situations even in a longer time scale. Thus, we can call it predictive. For example, such an engine may merge information about the observed traffic patterns and localization, and deliver reliable suggestions towards better resource allocation by certain nodes. However, one can observe that in order to realize such a vision, there should be a tight cooperation between various nodes of wireless communication systems. It means that such approaches, like federated learning, would be particularly applicable here. It means that the computation load would be distributed among numerous nodes to achieve reliable results in an acceptable time. Moreover, although the intelligent engine could have a centralized view of the whole system, it is impossible to have one physically centralized entity. Following the solutions known from virtualized networks, the intelligent engine able to apply advanced AI/ML tools may be physically distributed but it will be logically centralized.

The above three levels of advancement of the radio context information subsystem are illustrated in Fig. 11 and Fig. 12.

Finally, the radio context information is not only related to computation (processing) resources. It is highly important to analyze also the issues related to storage - the way the radio context information may be stored, and how these repositories

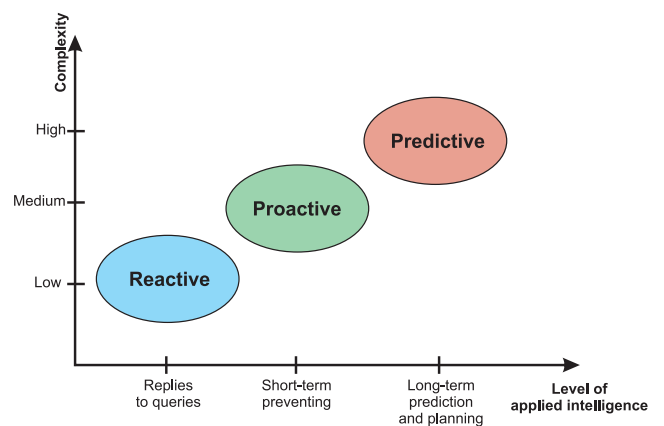


FIGURE 11. Three levels of advancement of the radio context information subsystem.

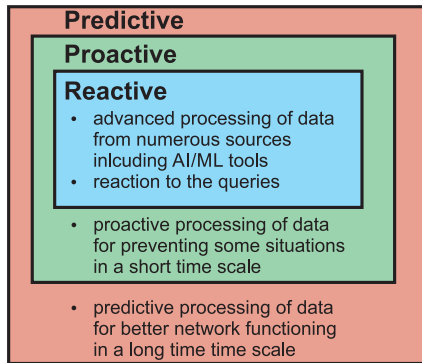


FIGURE 12. Characteristics of the three levels of advancement of the radio context information subsystem.

may be realized in practice. As we are not concentrating now on the specific way of data representation (i.e., if the radio context information is stored in the form of, e.g., key-value pairs, or in the form of dictionaries, or maps), our focus is on the databases themselves. As various kinds of databases are possible (e.g., relational, graphs etc.), they may either be distributed, centralized or hierarchical. In the first case, the information is distributed among numerous nodes of equal priority, and each node typically stores only information related to a specific geographical region. By exchanging data, neighboring nodes may enrich their local context information. In the second extreme, a fully centralized database is in the possession of all data from a wider region and delivers it - when necessary - to interested nodes. Though theoretically such an approach could be highly promising, it suffers from an extreme computational and storage burden. Thus, one of the prospective solutions is to apply the hierarchical approach, where the fully distributed and fully centralized approaches are mixed. Locally deployed repositories are in possession of locally-related (and typically of short time validity) data. These nodes create the lowest tier of nodes, the lowest tier of the radio context information subsystem. The higher tiers (and corresponding nodes located in these tiers) have an increasingly wider view on the system, and the top (highest) layer is treated as logically centralized tier. The higher the tier, the more generic the data, i.e., the data valid for a longer time-scale or for a wider area. Such a scheme is illustrated in Fig. 13.

IX. CONTEXT-AWARENESS DESIGN TRADE-OFFS, RECOMMENDATIONS, AND CONCLUSION

As discussed in the previous sections, the role of context information for the expected performance of future radio communication systems cannot be overestimated, and has been emphasized in a number of recent papers. Thus, the definition of a context-information framework for its acquisition, representation and distribution, as well as the definition of the suitable architecture, either centralized, distributed or mixed are of particular importance for future radio communications' broadly-understood efficiency. Let us now summarize

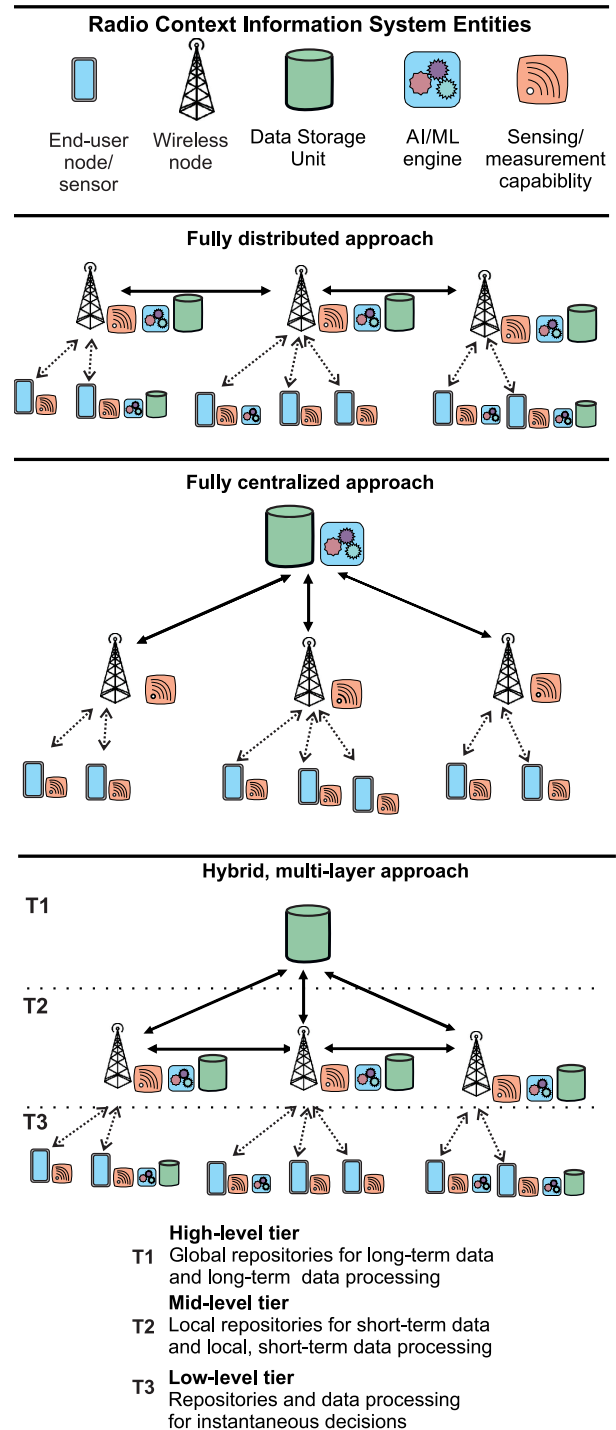


FIGURE 13. Distributed, centralized and hybrid architecture of the radio context information subsystem.

the design trade-offs and recommendations for such an architecture.

A. SIGNALLING OVERHEAD VS. RELIABILITY

Signaling overhead and information reliability are directly related to the key performance metrics of a radio network. Both of them reflect the methods and places of the

acquisition, representation, modification, and dissemination of context information. Signalling overhead (the cost) must be balanced with the performance improvement (the profit) that comes with exploiting context information. This is one of the main challenges of the deployment of distributed databases for environmental information. By designing an architecture that carefully considers the type and amount of information exchanged between different network layers and entities, the signaling overhead can be significantly reduced. Signaling overhead can also be reduced by a suitable choice of the dissemination strategy. An on-demand model is most appropriate when information is needed only rarely. A proactive model, on the other hand, may result in better performance for commonly needed and dynamically changing data. The AI/ML algorithms residing at the appropriate network point, e.g., at the network edge or dedicated subsystem, can reduce the signalling overhead by avoiding the transmission of data that can be retrieved by learning. The required information accuracy and reliability depends on the objectives of context-information usage and timescale. For instance, fast power control requires high precision, while dynamic spectrum allocation performs well with approximate or statistical information. The reliability also depends on time-dynamics and regional characteristics of information aggregation. The choice of large regions over a long period of time to reduce signaling overhead would come at the expense of the accuracy of the model or statistical characterization. Thus, whenever AI/ML methods are applied to enrich context awareness, they must converge at a required pace responding to the system dynamics, and with accuracy tailored to the application.

B. CONTEXT-INFORMATION ACQUISITION, STORAGE, AND DISTRIBUTION VS. POWER CONSUMPTION

In Section VIII, we discussed distributed, centralized and hybrid architectures of the RCI subsystem. Recall that a popular concept of providing context information is to have it stored, dynamically updated and made available in a centralized database (a storage unit) with an accompanying AI/ML engine to capture the dynamics and hidden dependencies of the information arriving from agents. The opposite approach to acquisition, storage and distribution of context information is to rely on a fully distributed architecture and edge intelligence (AI in the end-devices). Distributed architecture and operation of RCI subsystem requires significant additional traffic between the network devices, which in turn results in energy consumption. The *Smart Dust* project implemented in the University of Berkeley explored the limits on size and power consumption in autonomous sensor nodes [170]. This concept can be understood as decentralized acquisition and distribution of pieces of information. It incorporates the requisite sensing, communication, and computing hardware, along with a power supply, in the volume of a few cubic millimeters, while still achieving required performance in terms of sensor functionality and communication capability. The networking nodes consume extremely low power, communicate at bit rates measured in kbits per second and

potentially operate in high volumetric densities. However, originally, they do not possess any kind of AI. Considering future networks with various degrees of node mobility and density, it is possible that the idea of a network of simple low-power sensors (or information points) exchanging pieces of information, but applying edge AI/ML algorithms, could augment the concept of highly-reliable (but costly) centralised databases at low energy-cost. A major challenge is to define and incorporate the required functionalities of sensors/nodes for context-information acquisition, storage and distribution, while maintaining low power consumption.

C. REDUCED VS. INCOMPLETE INFORMATION

The required reliability and the amount of context information in a network can be generalized by models of complete information vs. incomplete information, and full-information vs. reduced information. In game theory, a metric describing the cost of not having complete information, called the *Price of Ignorance* is defined as a relative loss of common welfare (e.g., network performance) that results from not having complete information. Network performance can be defined in a number of ways, e.g., as the total network energy saving or its spectral efficiency (sum throughput over available bandwidth) or sum-throughput net. Ignorance can be understood as either uncertainty of information or possessing complete (certain) information, which has a reduced representation of the information describing the players' environmental conditions and options in detail. As an example, we can consider channel state information required for optimal resource allocation in a network. Providing complete information on all link qualities of all players to all other players in the considered network is associated with a huge communication cost. Thus, complete information-based optimal resource allocation is not practically possible. Applying the Bayesian game models to the considered problem is even more impractical, because the fading statistics of all channels for all players are required to consider every player behavior with a given probability. In a dynamic radio environment these statistics change with time. Moreover, it is impractical to consider the channel gains probability density functions with high granularity because it exponentially increases the computational complexity of calculating the equilibrium point. This example shows that there is a fundamental trade-off between information availability and compactness, and network performance. Here again, AI/ML methods have a huge potential to uncover hidden elements of context information, transforming incomplete information into complete, although it is rather not possible to infer full context information once it has been reduced.

D. MACHINE LEARNING ALGORITHMS DESIGN VS. QUALITY DATASETS

Although a number of scientific papers, describing the application of machine learning for context awareness is constantly increasing, the majority of authors do not publish the datasets they used to generate results. This leads to a lack

TABLE 10. List of abbreviations used in the paper.

Abbreviation	Description
5QI	5G QoS Identifier
ABCS	Adaptive threshold BCS
AC	Access Categories
AE	Autoencoder
AHP	Analytical Hierarchical Process
AI	Artificial Intelligence
AM	Amplitude Modulation
AMBR	Aggregate Maximum Bit Rate
AMR	Automatic Modulation Recognition
ANN	Artificial Neural Network
AoA	Angle of Arrival
ARFCN	Absolute Radio-Frequency Channel Number
ARP	Allocation and Retention Priority
AUROC	Area Under the Receiver Operating Characteristics
AWGN	Additive White Gaussian Noise
BCS	Bayesian Compressive Sensing
BER	Bit Error Rate
BLER	Block Error Rate
BPNN	Back Propagation Neural Network
BPSK	Binary PSK
CA	Context Awareness
CDP	Cyclic Domain Profile
CLDNN	Convolutional Long short-term Deep Neural Network
CNN	Convolutional Neural Networks
CQI	Channel Quality Indicator
CR	Cognitive Radio
CRAT	Context-aware Radio Access Technology
CRN	Cognitive Radio Networks
CS	Compressive Sensing
CSI	Channel State Information
CSO	Cell Selection Offset
CSS	Cooperative SS
CT	Coarse Tree
DCS	Digital Communication System
DCT	Discrete Cosine Transform
DEC-POMDP	Decentralized POMDP
DT	Decision Tree
DNN	Deep Neural Network
DVB-T2	Digital Video Broadcasting - Terrestrial 2
DWT	Discrete Wavelet Transform
EARFCN	E-UTRA Absolute Radio Frequency Channel Number
EGSM	Extended GSM
ELM	Extreme Learning Machine
EM	Expectation Maximization
ENN	Expanded Neural Networks
FB	Feature Based
FBMC	FilterBank-based Multi-Carrier
FC	Fusion Center
FCNN	Fully Connected Neural Network
FM	Frequency Modulation
FPR	False Positive Rate
FSK	Frequency Shift Keying
FT	Fine Tree
GA	Genetic Algorithm
GBR	Guaranteed Bit Rate
GBRT	Gradient Boosted Regression Trees
GNSS	Global Navigation Satellite Systems
GMM	Gaussian Mixture Model
GMSK	Gaussian MSK
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HBB	Hidden Markov Model
HOS	Higher Order Statistics
ICI	Inter-Carrier Interference
IoT	Internet of Things
kNN	k-Nearest Neighbors
LB	Likelihood-based
LOS	Line of Sight
LR	Logistic Regression
LSTM	Long Short Term Memory
LTE	Long Term Evolution
MAC	Medium Access Control layer
MBR	Maximum Bit Rate
MCS	Modulation and Coding Scheme
MFCC	Mel-Frequency Cepstral Coefficients
MDT	Minimization of Drive Test
ML	Machine Learning
MSK	Minimum Shift Keying

TABLE 10. (Continued.) List of abbreviations used in the paper.

Abbreviation	Description
MT	Medium Tree
NB	Naive Bayes
NLOS	Non LOS
NN	Neural Network
OFDM	Orthogonal Frequency Division Multiplexing
OQAM	Offset QAM
OQPSK	Offset Quadrature Phase Shift Keying
OSTP	OFDM Symbol Transmit Power
PAM	Pulse Amplitude Modulation
PC	Principal Component
PCA	Principal Component Analysis
PDCP	Packet Data Convergence Protocol
PHY	Physical layer
PMI	Precoding Matrix Index
POMDP	Partially Observable Markov Decision Process
PSK	Phase Shift Keying
PSO	Particle Swarm Optimization
PU	Primary User
QAM	Quadrature Amplitude Modulation
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
RAT	Radio Access Technologies
RCI	Radio Context Information
RCPI	Received Channel Power Indicator
RI	Rank Indicator
RL	Reinforcement Learning
RLC	Radio Link Control
RNN	Recurrent Neural Networks
ROC	Receiver Operating Characteristics
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RSRP	Reference Signal Receive Power
RSRQ	Reference Signal Received Quality
RVM	Relevance Vector Machine
RX	Receiver
SCF	Spectral Coherence Function
SDR	Software Defined Radio
SINR	Signal to Interference to Noise Ratio
SNR	Signal to Noise Ratio
SOM	Self-organizing Map
SS	Spectrum Sensing
SS-RSRP	Synchronization Signal RSRP
SU	Secondary User
SVM	Support Vector Machines
ToA	Time of Arrival
TDoA	Time Difference of Arrival
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TPR	True Positive Rate
TX	Transmitter
UARFCN	UTRA Absolute Radio Frequency Channel Number
UFMC	Universal-Filtered Multi-Carrier
UHF	Ultra High Frequency
UTRA	Universal Terrestrial Radio Access
WCDMA	Wideband Code Division Multiple Access

of possibilities to make an objective comparison between machine learning methods and architectures. The successful application of ML models requires high quality datasets. Especially for larger neural networks that have a large set of parameters, sufficient training data volume is very important. However, mobile network datasets are scarce. Mobile data collected by sensors or network equipment is frequently affected by loss, redundancy, and mislabeling, thus requiring

cleaning before application for model training. In addition, mobile service providers and operators keep the collected data confidential, and are reluctant to share them for research purposes.

Moreover, the absence of public mobile networks datasets leads to another problem: many investigations are performed on private data. Without a comparison of the performance of various models on the same data it is hard to design and select an approach which works best, and to decide in what aspect it could be improved.

E. FINAL REMARKS

Above, we have surveyed the existing literature to address the following issues: (i) What is the role of context information, its availability and representation in contemporary and future radio communication networks? (ii) What are the suitable AI/ML methods to enrich context awareness in these networks? (iii) What kind of architectural framework utilizing an AI/ML engine is practical for context information acquisition, storage and distribution among nodes and networks in the considered scenarios? (iv) What are the design trade-offs and recommendations for intelligent context-aware radio communication? We believe that answering these questions is of particular relevance for future ubiquitous radio communication and its broadly-understood efficiency.

X. ABBREVIATIONS

Abbreviations used in this paper are listed in Table 10.

REFERENCES

- [1] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the Internet of Things: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 414–454, 1st Quart., 2014.
- [2] O. Yürür, C. Liu, Z. Sheng, V. Leung, W. Moreno, and K. Leung, "Context-awareness for mobile sensing: A survey and future directions," *IEEE Commun. Surveys Tuts.*, vol. 18, pp. 68–93, 1st Quart., 2016.
- [3] M. Knappmeyer, S. L. Kiani, E. S. Retz, N. Baker, and R. Tonjes, "Survey of context provisioning middleware," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1492–1519, 3rd Quart., 2013.
- [4] P. Makris, D. N. Skoutas, and C. Skianis, "A survey on context-aware mobile and wireless networking: On networking and computing environments' integration," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 362–386, 1st Quart., 2013.
- [5] Y. Liu, S. Bi, Z. Shi, and L. Hanzo, "When machine learning meets big data: A wireless communication perspective," *IEEE Veh. Technol. Mag.*, vol. 15, no. 1, pp. 63–72, Mar. 2020.
- [6] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-aware computing, learning, and big data in Internet of Things: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 1–27, Feb. 2018.
- [7] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.
- [8] K. Cicho, A. Kliks, and H. Bogucka, "Energy-efficient cooperative spectrum sensing: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1861–1886, 3rd Quart. 2016.
- [9] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1124–1148, 2nd Quart., 2017.
- [10] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.
- [11] A. Kliks, D. Triantafylloupolou, L. De Nardis, O. Holland, L. Gavrilovska, and A. Bantouna, "Cross-layer analysis in cognitive radio—Context identification and decision making aspects," *IEEE Trans. Cognit. Commun. Netw.*, vol. 1, no. 4, pp. 450–463, Dec. 2015.
- [12] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1838–1860, 3rd Quart., 2016.
- [13] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, 3rd Quart., 2013.
- [14] M. Weiser, "The computer for the 21st century," *SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 3, no. 3, pp. 3–11, Jul. 1999.
- [15] M. A. Munoz, M. Rodriguez, J. Favela, A. I. Martinez-Garcia, and V. M. Gonzalez, "Context-aware mobile communication in hospitals," *Computer*, vol. 36, no. 9, pp. 38–46, Sep. 2003.
- [16] J. Liu, J. Wan, D. Jia, B. Zeng, D. Li, C. Hsu, and H. Chen, "High-efficiency urban traffic management in context-aware computing and 5G communication," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 34–40, Jan. 2017.
- [17] Z. Zhou, G. Ma, M. Dong, K. Ota, C. Xu, and Y. Jia, "Iterative energy-efficient stable matching approach for context-aware resource allocation in D2D communications," *IEEE Access*, vol. 4, pp. 6181–6196, 2016.
- [18] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [19] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.
- [20] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.
- [21] E. Bedeer, O. A. Dobre, M. H. Ahmed, and K. E. Baddour, "Joint optimization of bit and power loading for multicarrier systems," *IEEE Wireless Commun. Lett.*, vol. 2, no. 4, pp. 447–450, Aug. 2013.
- [22] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart., 2015.
- [23] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, "Modulation and multiple access for 5G networks," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 629–646, 1st Quart., 2018.
- [24] T. Y. Arif, R. Munadi, and Fardian, "Energy efficiency opportunity at same data rate and different MCS in IEEE 802.11n," in *Proc. 9th Asia Modeling Symp. (AMS)*, Sep. 2015, pp. 142–147.
- [25] S. Kim, H. Kim, and D. Hong, "Joint power allocation and MCS selection in downlink NOMA system," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1–4.
- [26] H.-S. Jo, C. Mun, J. Moon, and J.-G. Yook, "Interference mitigation using uplink power control for two-tier femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4906–4910, Oct. 2009.
- [27] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2801–2806, Jun. 2009.
- [28] H. N. Qureshi and A. Imran, "Optimal bin width for autonomous coverage estimation using MDT reports in the presence of user positioning error," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 716–719, Apr. 2019.
- [29] V. Kumar, H. Li, J.-M.-J. Park, and K. Bian, "Crowd-sourced authentication for enforcement in dynamic spectrum sharing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 3, pp. 625–636, Sep. 2019.
- [30] J. Liu, W. C. Tang, Y. Chen, M. Li, and M. Guizani, "A novel crowd-sourcing inference method," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2019, pp. 55–60.
- [31] G. Xu, S. Gao, M. Daneshmand, C. Wang, and Y. Liu, "A survey for mobility big data analytics for geolocation prediction," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 111–119, Feb. 2017.
- [32] X. Cheng, L. Fang, L. Yang, and S. Cui, "Mobile big data: The fuel for data-driven wireless," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1489–1516, Oct. 2017.
- [33] S. Boubiche, D. E. Boubiche, A. Bilami, and H. Toral-Cruz, "Big data challenges and data aggregation strategies in wireless sensor networks," *IEEE Access*, vol. 6, pp. 20558–20571, 2018.
- [34] X. Cao, L. Liu, Y. Cheng, and X. S. Shen, "Towards energy-efficient wireless networking in the big data era: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 303–332, 1st Quart., 2018.

- [35] *Mw Online Dictionary*. Accessed: Jul. 29, 2020. [Online]. Available: <https://www.merriam-webster.com/dictionary/context>
- [36] B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *IEEE Netw.*, vol. 8, no. 5, pp. 22–32, Sep. 1994.
- [37] P. J. Brown, J. D. Bovey, and X. Chen, "Context-aware applications: From the laboratory to the marketplace," *IEEE Pers. Commun.*, vol. 4, no. 5, pp. 58–64, Oct. 1997.
- [38] A. K. Dey, G. D. Abowd, and D. Salber, "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications," *Hum.-Comput. Interact.*, vol. 16, nos. 2–4, pp. 97–166, 2001.
- [39] S. Ahn and D. Kim, "Proactive context-aware sensor networks," in *Wireless Sensor Networks*, K. Römer, H. Karl, and F. Mattern, Eds. Berlin, Germany: Springer, 2006, pp. 38–53.
- [40] J. Mitola and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Apr. 1999.
- [41] C. She, C. Yang, and T. Q. S. Quek, "Cross-layer optimization for ultra-reliable and low-latency radio access networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 127–141, Jan. 2018.
- [42] Q. Medhat Salih, M. A. Rahman, F. Al-Turjman, and Z. R. M. Azmi, "Smart routing management framework exploiting dynamic data resources of cross-layer design and machine learning approaches for mobile cognitive radio networks: A survey," *IEEE Access*, vol. 8, pp. 67835–67867, 2020.
- [43] A. Kosta, N. Pappas, and V. Angelakis, *Age Information: A New Concept, Metric, Tool*. Norwell, MA, USA: Now Publishers, 2017.
- [44] E. Hayden, "Data lifecycle management model shows risks and integrated data flow," *Inf. Secur. Mag.*, Jul. 2008.
- [45] R. Alieiev, T. Hehn, A. Kwoczek, and T. Kürner, "Sensor-based communication prediction for dynamic Doppler-shift compensation," in *Proc. 15th Int. Conf. ITS Telecommun. (ITST)*, May 2017, pp. 1–7.
- [46] T. Levanen, J. Talvitie, R. Wichman, V. Syrjala, M. Renfors, and M. Valkama, "Location-aware 5G communications and Doppler compensation for high-speed train networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–6.
- [47] *Study on Performance Enhancements for High Speed Scenario in LTE*, Standard 3GPP TR 36.878 V13.0.0, Rel. 13, Tech. Spec. Group Radio Access Network, Jan. 2016.
- [48] P. Luoto, K. Rikkinen, P. Kinnunen, J. Karjalainen, K. Pajukoski, J. Hulkkonen, and M. Latva-Aho, "Configurable 5G air interface for high speed scenario," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–5.
- [49] E. Ali, M. Ismail, R. Nordin, and N. F. Abdulah, "Beamforming techniques for massive MIMO systems in 5G: Overview, classification, and trends for future research," *Frontiers Inf. Technol. Electron. Eng.*, vol. 18, no. 6, pp. 753–772, 2017.
- [50] W. B. Abbas and M. Zorzi, "Context information based initial cell search for millimeter wave 5G cellular networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2016, pp. 111–116.
- [51] I. Filippini, V. Sciancalepore, F. Devoti, and A. Capone, "Fast cell discovery in mm-wave 5G networks with context information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 7, pp. 1538–1552, Jul. 2018.
- [52] A. Habbal, S. I. Goudar, and S. Hassan, "A context-aware radio access technology selection mechanism in 5G mobile network for smart city applications," *J. Netw. Comput. Appl.*, vol. 135, pp. 97–107, Jun. 2019.
- [53] C.-L. Hwang and K. Yoon, *Multiple Attribute Decision Making, Methods and Applications A State-of-the-Art Survey*. Berlin, Germany: Springer, 1981.
- [54] M. Wasilewska and H. Bogucka, "Machine learning for LTE energy detection performance improvement," *Sensors*, vol. 19, no. 19, p. 4348, Oct. 2019.
- [55] A. M. Mikaeil, B. Guo, and Z. Wang, "Machine learning to datafusion approach for cooperative spectrum sensing," in *Proc. 6th Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, Shanghai, China, 2014, pp. 429–434.
- [56] M. Tang, Z. Zheng, G. Ding, and Z. Xue, "Efficient TV white space database construction via spectrum sensing and spatial inference," in *Proc. 34th IEEE Int. Perform. Comput. Commun. Conf. (IPCCC)*, Nanjing, China, Dec. 2015, pp. 1–5.
- [57] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Commun. Mobile Comput.*, vol. 2018, Sep. 2018, Art. no. 5906097.
- [58] F. Adly, S. Muhaidat, and P. D. Yoo, "Cumulant-based automatic modulation classification over frequency-selective channels," in *Proc. IEEE Int. Conf. Cyber. Phys. Social Comput. (CPSCom)*, Jul. 2018, pp. 1–5.
- [59] M. AlHajri, N. Ali, and R. Shubair, "Indoor localization for IoT using adaptive feature selection: A cascaded machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, pp. 2306–2310, 2019.
- [60] J. Tian, P. Cheng, Z. Chen, M. Li, H. Hu, Y. Li, and B. Vucetic, "A machine learning-enabled spectrum sensing method for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11374–11378, Nov. 2019.
- [61] B. Xu, X. Wen, and X. Wang, "Research on modulation recognition technology based on machine learning," *Int. J. Comput. Appl. Technol. Res.*, vol. 8, no. 4, pp. 102–106, Apr. 2019.
- [62] M. Subbarao and P. Samundiswary, "Automatic modulation recognition in cognitive radio receivers using multi-order cumulants and decision trees," *Int. J. Recent Technol. Eng. (IJRTE)*, vol. 7, pp. 61–67, 11 2018.
- [63] S. U. Jan, V. H. Vu, and I. S. Koo, "Performance analysis of support vector machine-based classifier for spectrum sensing in cognitive radio networks," in *Proc. Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, Oct. 2018, pp. 385–3854.
- [64] Y. Lu, P. Zhu, D. Wang, and M. Fattouche, "Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Doha, Qatar, Apr. 2016, pp. 1–6.
- [65] Y.-D. Huang, Y.-C. Liang, and G. Yang, "A fuzzy support vector machine algorithm for cooperative spectrum sensing with noise uncertainty," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [66] C. Chembe, I. Ahmedy, R. Md Noor, D. Kunda, M. Oche, and A. B. Tambawal, "Cooperative spectrum decision in cognitive vehicular network based on support vector machine," *Malaysian J. Comput. Sci.*, vol. 32, no. 2, pp. 83–96, Apr. 2019.
- [67] A. Subekti, H. F. Pardede, R. Sustika, and Suyoto, "Spectrum sensing for cognitive radio using deep autoencoder neural network and SVM," in *Proc. Int. Conf. Radar, Antenna, Microw., Electron., Telecommun. (ICRAMET)*, Nov. 2018, pp. 81–85.
- [68] Y. Tang, Q. Zhang, and W. Lin, "Artificial neural network based spectrum sensing method for cognitive radio," in *Proc. 6th Int. Conf. Wireless Commun. Netw. Mobile Comput.*, Sep. 2010, pp. 1–4.
- [69] P. S. Cheong, M. Camelo, and S. Latré, "Evaluating deep neural networks to classify modulated and coded radio signals," in *Cognitive Radio Oriented Wireless Networks*, I. Moerman, J. Marquez-Barja, A. Shahid, W. Liu, S. Giannoulis, and X. Jiao, Eds. Cham, Switzerland: Springer, 2019, pp. 177–188.
- [70] F. Liu, Y. Zhou, and Y. Liu, "A deep neural network method for automatic modulation recognition in OFDM with index modulation," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–5.
- [71] F. Paisana, A. Selim, M. Kist, P. Alvarez, J. Tallon, C. Bluemm, A. Puschmann, and L. DaSilva, "Context-aware cognitive radio using deep learning," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Piscataway, NJ, USA, Mar. 2017, pp. 1–2.
- [72] W. Lee, M. Kim, and D. Cho, "Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3005–3009, Mar. 2019.
- [73] S. Zhou, Z. Yin, Z. Wu, Y. Chen, N. Zhao, and Z. Yang, "A robust modulation classification method using convolutional neural networks," *EURASIP J. Adv. Signal Process.*, vol. 2019, no. 1, p. 21, 2019.
- [74] M. Camelo, T. D. Schepper, P. Soto, J. Marquez-Barja, J. Famaey, and S. Latré, "Detection of traffic patterns in the radio spectrum for cognitive wireless network management," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [75] L. Yu, J. Chen, G. Ding, Y. Tu, J. Yang, and J. Sun, "Spectrum prediction based on Taguchi method in deep learning with long short-term memory," *IEEE Access*, vol. 6, pp. 45923–45933, 2018.
- [76] M. Zhang, Y. Zeng, Z. Han, and Y. Gong, "Automatic modulation recognition using deep learning architectures," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
- [77] L. Yu, Q. Wang, Y. Guo, and P. Li, "Spectrum availability prediction in cognitive aerospace communications: A deep learning perspective," in *Proc. Cognit. Commun. Aerosp. Appl. Workshop (CCAA)*, Jun. 2017, pp. 1–4.
- [78] K. Zhang, J. Li, and F. Gao, "Machine learning techniques for spectrum sensing when primary user has multiple transmit powers," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Macau, China, Nov. 2014, pp. 137–141.

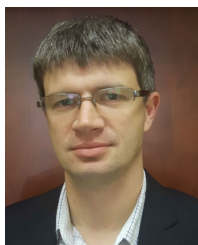
- [79] S. Ramjee, S. Ju, D. Yang, X. Liu, A. El Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," 2019, *arXiv:1901.05850*.
- [80] S. Duan, K. Chen, X. Yu, and M. Qian, "Automatic multicarrier waveform classification via PCA and convolutional neural networks," *IEEE Access*, vol. 6, pp. 51365–51373, 2018.
- [81] M. Bkassiny, S. K. Jayaweera, and K. A. Avery, "Distributed reinforcement learning based MAC protocols for autonomous cognitive secondary users," in *Proc. 20th Annu. Wireless Opt. Commun. Conf. (WOCC)*, Apr. 2011, pp. 1–6.
- [82] B. F. Lo and I. F. Akyildiz, "Reinforcement learning for cooperative sensing gain in cognitive radio ad hoc networks," *Wireless Netw.*, vol. 19, no. 6, pp. 1237–1250, Aug. 2013.
- [83] A. Ghosh, N. Manwani, and P. S. Sastry, "On the robustness of decision tree learning under label noise," in *Advances in Knowledge Discovery and Data Mining*, J. Kim, K. Shim, L. Cao, J.-G. Lee, X. Lin, and Y.-S. Moon, Eds. Cham, Switzerland: Springer, 2017, pp. 685–697.
- [84] D. D. Z. Soto, O. J. S. Parra, and D. A. L. Sarmiento, "Detection of the primary user's behavior for the intervention of the secondary user using machine learning," in *Future Data and Security Engineering*. Cham, Switzerland: Springer, 2018, pp. 200–213.
- [85] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [86] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," in *Proc. 9th Int. Conf. Artif. Neural Netw. (ICANN)*, vol. 2, Sep. 1999, pp. 850–855.
- [87] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.
- [88] A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for aggregated interference control in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1823–1834, May 2010.
- [89] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.
- [90] L. Pucker, *Review of Contemporary Spectrum Sensing Technologies*, Standard IEEE-SA P1900.6 Standards Group, 2010.
- [91] U. Salama, P. L. Sarker, and A. Chakrabarty, "Enhanced energy detection using matched filter for spectrum sensing in cognitive radio networks," in *Proc. Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV) 2nd Int. Conf. Imag., Vis. Pattern Recognit. (icVPR)*, Jun. 2018, pp. 185–190.
- [92] W. M. Lees, A. Wunderlich, P. J. Jeavons, P. D. Hale, and M. R. Souryal, "Deep learning classification of 3.5-GHz band spectrograms with applications to spectrum sensing," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 2, pp. 224–236, Jun. 2019.
- [93] A. Tiwari, H. Chenji, and V. Devabhaktuni, "Comparison of statistical signal processing and machine learning algorithms for spectrum sensing," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
- [94] L. Yu, J. Chen, and G. Ding, "Spectrum prediction via long short term memory," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 643–647.
- [95] A. Agarwal, S. Dubey, M. A. Khan, R. Gangopadhyay, and S. Debnath, "Learning based primary user activity prediction in cognitive radio networks for efficient dynamic spectrum access," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Bangalore, India, Jun. 2016, pp. 1–5.
- [96] D.-H.-T. Huang, S.-H. Wu, and P.-H. Wang, "Cooperative spectrum sensing and locationing: A sparse Bayesian learning approach," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–5.
- [97] L. Zhang, H. Huang, and X. Jing, "A modified cyclostationary spectrum sensing based on softmax regression model," in *Proc. 16th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2016, pp. 620–623.
- [98] Y. Aoki, T. Fujii, and T. Ide, "Time series analysis of multiple primary user environment using HMM-based spectrum sensing," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–5.
- [99] Y. Xu, P. Cheng, Z. Chen, Y. Li, and B. Vucetic, "Mobile collaborative spectrum sensing for heterogeneous networks: A Bayesian machine learning approach," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5634–5647, Nov. 2018.
- [100] N. Zhang, X. Sun, C. Guo, and L. Gao, "PU probability prediction based Bayesian compressive spectrum sensing," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Nov. 2015, pp. 1–5.
- [101] X. Sun, L. Gao, X. Luo, and K. Su, "RBM based cooperative Bayesian compressive spectrum sensing with adaptive threshold," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Jul. 2016, pp. 1–6.
- [102] B. Khalfi, A. Zaid, and B. Hamdaoui, "When machine learning meets compressive sampling for wideband spectrum sensing," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Valencia, Spain, Jun. 2017, pp. 1120–1125.
- [103] M. Nazzal, A. R. Ektf, A. Görçin, and H. Arslan, "Exploiting sparsity recovery for compressive spectrum sensing: A machine learning approach," *IEEE Access*, vol. 7, pp. 126098–126110, 2019.
- [104] K. V. Aiswarya, G. N. R. M. Naidu, K. Kalaiarasi, D. Kavya, and S. Kirthiga, "Spectrum sensing using sparse Bayesian learning," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2019, pp. 582–586.
- [105] P. Cheng, Y. Li, Z. Chen, and B. Vucetic, "High-resolution wideband spectrum sensing based on sparse Bayesian learning," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–5.
- [106] S. Chen, B. Shen, X. Wang, and H. Wu, "SVM and decision stumps based hybrid AdaBoost classification algorithm for cognitive radios," in *Proc. 21st Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2019, pp. 492–497.
- [107] R. Mustafa, R. R. Jaglan, and S. Agrawal, "Decision-fusion-based reliable CSS scheme in CR networks," *IET Commun.*, vol. 13, no. 7, pp. 947–953, Apr. 2019.
- [108] X. Ma, S. Ning, X. Liu, H. Kuang, and Y. Hong, "Cooperative spectrum sensing using extreme learning machine for cognitive radio networks with multiple primary users," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 536–540.
- [109] Z. Li, W. Wu, X. Liu, and P. Qi, "Improved cooperative spectrum sensing model based on machine learning for cognitive radio networks," *IET Commun.*, vol. 12, no. 19, pp. 2485–2492, Dec. 2018.
- [110] C. H. A. Tavares and T. Abrao, "Bayesian estimators for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE URUCON*, Oct. 2017, pp. 1–4.
- [111] A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales-Tirado, and J. J. Neel, "A survey of artificial intelligence for cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1578–1592, May 2010.
- [112] D. Zhang and X. Zhai, "SVM-based spectrum sensing in cognitive radio," in *Proc. 7th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2011, pp. 1–4.
- [113] P. Saravanan, S. S. Chandra, A. Upadhye, and S. Gurugopinath, "A supervised learning approach for differential entropy feature-based spectrum sensing," in *Proc. 6th Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2021, pp. 395–399.
- [114] H. Yang, X. Xie, and R. Wang, "SOM-GA-SVM detection based spectrum sensing in cognitive radio," in *Proc. 8th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2012, pp. 1–7.
- [115] O. P. Awe, Z. Zhu, and S. Lambodharan, "Eigenvalue and support vector machine techniques for spectrum sensing in cognitive radio networks," in *Proc. Conf. Technol. Appl. Artif. Intell.*, Dec. 2013, pp. 223–227.
- [116] O. P. Awe, A. Deligiannis, and S. Lambodharan, "Spatio-temporal spectrum sensing in cognitive radio networks using beamformer-aided SVM algorithms," *IEEE Access*, vol. 6, pp. 25377–25388, 2018.
- [117] J. Lu, L. Li, G. Chen, D. Shen, K. Pham, and E. Blasch, "Machine learning based intelligent cognitive network using fog computing," *Proc. SPIE*, vol. 10196, pp. 149–157, May 2017.
- [118] H. Xue and F. Gao, "A machine learning based spectrum-sensing algorithm using sample covariance matrix," in *Proc. 10th Int. Conf. Commun. Netw. China (ChinaCom)*, Aug. 2015, pp. 476–480.
- [119] O. P. Awe, S. M. Naqvi, and S. Lambodharan, "Kalman filter enhanced parametric classifiers for spectrum sensing under flat fading channels," in *Cognitive Radio Oriented Wireless Networks*. Cham, Switzerland: Springer, 2015, pp. 235–247.
- [120] M. R. Vyas, D. Patel, and M. Lopez-Benitez, "Artificial neural network based hybrid spectrum sensing scheme for cognitive radio," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7.
- [121] D. Han, G. C. Sobabe, C. Zhang, X. Bai, Z. Wang, S. Liu, and B. Guo, "Spectrum sensing for cognitive radio based on convolution neural network," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2017, pp. 1–6.
- [122] M. Zhang, L. Wang, Y. Feng, and H. Yin, "A spectrum sensing algorithm for OFDM signal based on deep learning and covariance matrix graph," *IEICE Trans. Commun.*, vol. E101.B, no. 12, pp. 2435–2444, 2018.
- [123] K. Yang, Z. Huang, X. Wang, and X. Li, "A blind spectrum sensing method based on deep learning," *Sensors*, vol. 19, p. 2270, 2019.

- [124] K. W. Choi and E. Hossain, "Estimation of primary user parameters in cognitive radio systems via hidden Markov model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 782–795, Feb. 2013.
- [125] D. Malafaia, J. Vieira, and A. Tomé, "Adaptive threshold spectrum sensing based on expectation maximization algorithm," *Phys. Commun.*, vol. 21, pp. 60–69, Dec. 2016.
- [126] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2007, pp. IV-1357–IV-1360.
- [127] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [128] X. Chen, F. Hou, H. Huang, and X. Jing, "Principle component analysis based cooperative spectrum sensing in cognitive radio," in *Proc. 16th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2016, pp. 602–605.
- [129] E. Ghazizadeh, B. Nikpour, D. A. Moghadam, and H. Nezamabadi-Pour, "A PSO-based weighting method to enhance machine learning techniques for cooperative spectrum sensing in CR networks," in *Proc. 1st Conf. Swarm Intell. Evol. Comput. (CSIEC)*, Mar. 2016, pp. 113–118.
- [130] K. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2209–2221, Nov. 2013.
- [131] A. M. Mikaeil, "Machine learning approaches for spectrum management in cognitive radio networks," in *Machine Learning*, H. Farhadi, Ed. Rijeka, Croatia: IntechOpen, 2018, ch. 6.
- [132] H. Liu, X. Zhu, and T. Fujii, "Ensemble deep learning based cooperative spectrum sensing with stacking fusion center," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2018, pp. 1841–1846.
- [133] M. Golvaei and M. Fakhrazadeh, "A fast soft decision algorithm for cooperative spectrum sensing," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 1, pp. 241–245, Jan. 2021.
- [134] Q. Zhu, Z. Han, and T. Basar, "No-regret learning in collaborative spectrum sensing with malicious nodes," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–6.
- [135] K. W. Choi, E. Hossain, and D. I. Kim, "Cooperative spectrum sensing under a random geometric primary user network model," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1932–1944, Jun. 2011.
- [136] J. Lunden, V. Koivunen, S. R. Kulkarni, and H. V. Poor, "Reinforcement learning based distributed multiagent sensing policy for cognitive radio networks," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, May 2011, pp. 642–646.
- [137] A. Assra, J. Yang, and B. Champagne, "An EM approach for cooperative spectrum sensing in multi-antenna CR networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1229–1243, Mar. 2016.
- [138] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.
- [139] M. A. Hazar, N. Odabasioglu, T. Ensari, Y. Kavurucu, and O. F. Sayan, "Performance analysis and improvement of machine learning algorithms for automatic modulation recognition over Rayleigh fading channels," *Neural Comput. Appl.*, vol. 29, no. 9, pp. 351–360, May 2018.
- [140] W. Xiong, P. Bogdanov, and M. Zheleva, "Robust and efficient modulation recognition based on local sequential IQ features," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2019, pp. 1612–1620.
- [141] S. A. Ghauri, I. M. Qureshi, I. Shah, and N. Khan, "Modulation classification using cyclostationary features on fading channels," *Res. J. Appl. Sci., Eng. Technol.*, vol. 7, no. 24, pp. 5331–5339, Jun. 2014.
- [142] R. Li, L. Li, S. Yang, and S. Li, "Robust automated VHF modulation recognition based on deep convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 946–949, May 2018.
- [143] G. Q. Yang, "Modulation classification based on extensible neural networks," *Math. Problems Eng.*, vol. 2017, pp. 1–10, Oct. 2017.
- [144] J. Fu, C. Zhao, B. Li, and X. Peng, *Deep Learning Based Digital Signal Modulation Recognition*, vol. 322. Cham, Switzerland: Springer, Jan. 2015, pp. 955–964.
- [145] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [146] M. Li, O. Li, G. Liu, and C. Zhang, "An automatic modulation recognition method with low parameter estimation dependence based on spatial transformer networks," *Appl. Sci.*, vol. 9, no. 5, p. 1010, Mar. 2019.
- [147] S. Peng, H. Jiang, H. Wang, H. Alwageed, and Y.-D. Yao, "Modulation classification using convolutional neural network based deep learning model," in *Proc. 26th Wireless Opt. Commun. Conf. (WOCC)*, Apr. 2017, pp. 1–5.
- [148] A. Ali, F. Yangyu, and S. Liu, "Automatic modulation classification of digital modulation signals with stacked autoencoders," *Digit. Signal Process.*, vol. 71, pp. 108–116, Dec. 2017.
- [149] X. Liu, D. Yang, and A. E. Gamal, "Deep neural network architectures for modulation classification," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Oct. 2017, pp. 915–919.
- [150] M. E.-H. Keshk, M. Abd-Elnaby, E.-S. El-Rabaie, M. I. Dessouky, and F. A. El-Samie, "Digital modulation recognition in OFDM systems using support vector machine classifier," *CiiT Int. J. Digit. Signal Process.*, vol. 5, pp. 516–525, Jan. 2013.
- [151] S. Norouzi, A. Jamshidi, and A. R. Zolghadrasli, "Adaptive modulation recognition based on the evolutionary algorithms," *Appl. Soft Comput.*, vol. 43, pp. 312–319, Jun. 2016.
- [152] M. H. Alhazmi, M. Alymani, H. Alhazmi, A. Almarhabi, A. Samarkandi, and Y.-D. Yao, "5G signal identification using deep learning," in *Proc. 29th Wireless Opt. Commun. Conf. (WOCC)*, May 2020, pp. 1–5.
- [153] A. Selim, F. Paisana, J. A. Arokkiar, Y. Zhang, L. Doyle, and L. A. DaSilva, "Spectrum monitoring for radar bands using deep convolutional neural networks," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.
- [154] K. Tekbiyik, O. Akbunar, A. R. Ekti, A. Gorcini, and G. K. Kurt, "Multi-dimensional wireless signal identification based on support vector machines," *IEEE Access*, vol. 7, pp. 138890–138903, 2019.
- [155] M. Schmidt, D. Block, and U. Meier, "Wireless interference identification with convolutional neural networks," in *Proc. IEEE 15th Int. Conf. Ind. Informat. (INDIN)*, Jul. 2017, pp. 180–185.
- [156] S. A. Rajab, W. Balid, M. O. Al Kalaa, and H. H. Refai, "Energy detection and machine learning for the identification of wireless MAC technologies," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Aug. 2015, pp. 1440–1446.
- [157] Z. E. Khatib, A. Hajihoseini, and S. A. Ghorashi, "A fingerprint method for indoor localization using autoencoder based deep extreme learning machine," *IEEE Sensors Lett.*, vol. 2, no. 1, pp. 1–4, Mar. 2018.
- [158] B. S. Ciftler, A. Albaseer, N. Lasla, and M. Abdallah, "Federated learning for RSS fingerprint-based localization: A privacy-preserving crowdsourcing method," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 2112–2117.
- [159] M. N. De Sousa and R. S. Thoma, "Mobile station localization emitter in urban NLoS using multipath ray tracing fingerprints and machine learning," in *Proc. 8th Int. Conf. Localization GNSS (ICL-GNSS)*, Jun. 2018, pp. 1–6.
- [160] H. Wang and Y.-D. Yao, "Primary user boundary detection in cognitive radio networks: Estimated secondary user locations and impact of malicious secondary users," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4577–4588, May 2018.
- [161] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.
- [162] H. Rutagemwa, A. Ghasemi, and S. Liu, "Dynamic spectrum assignment for land mobile radio with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.
- [163] D. Roy, T. Mukherjee, M. Chatterjee, and E. Pasilio, "Primary user activity prediction in DSA networks using recurrent structures," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Nov. 2019, pp. 1–10.
- [164] Z. Zhang, K. Zhang, F. Gao, and S. Zhang, "Spectrum prediction and channel selection for sensing-based spectrum sharing scheme using online learning techniques," in *Proc. IEEE 26th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Aug. 2015, pp. 355–359.
- [165] X. Liu, Q. Sun, W. Lu, C. Wu, and H. Ding, "Big-data-based intelligent spectrum sensing for heterogeneous spectrum communications in 5G," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 67–73, Oct. 2020.
- [166] M. Wasilewska and H. Bogucka, "Space-time-frequency machine learning for improved 4G/5G energy detection," *Int. J. Electron. Telecommun.*, vol. 66, no. 1, pp. 217–223, 2020.
- [167] M. Wasilewska, H. Bogucka, and A. Kliks, "Spectrum sensing and prediction for 5G radio," in *Proc. Int. Conf. Big Data Technol. Appl.* Cham, Switzerland: Springer, Dec. 2020, pp. 176–194.
- [168] B. A. Homssi, A. Al-Hourani, Z. Krusevac, and W. S. T. Rowe, "Machine learning framework for sensing and modeling interference in IoT frequency bands," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4461–4471, Mar. 2021.

- [169] K. Hamedani, L. Liu, and Y. Yi, "Energy efficient MIMO-OFDM spectrum sensing using deep stacked spiking delayed feedback reservoir computing," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 1, pp. 484–496, Mar. 2021.
- [170] J. M. Kahn, R. H. Katz, and K. S. Pister, "Emerging challenges: Mobile networking for 'smart dust,'" *J. Commun. Netw.*, vol. 2, no. 3, pp. 188–196, 2000.



ded intelligence, radio communications, spectrum sensing techniques, and intelligent radio resources management. She is the author of six research papers and the winner of Best Paper Award at 13th EAI International Conference on Wireless Internet Conference, in 2020.



ADRIAN KLIKS (Senior Member, IEEE) received the Postdoctoral degree in technical sciences, discipline: technical computer science and telecommunications, in February 2019. He works as a University Professor at the Institute of Radiocommunications, Poznań University of Technology. He took part in numerous international research projects: URANUS, NEWCOM++, ACROPOLIS, COGEU, NEWCOM#, COHERENT, in COST IC0902 and COST-Terra (IC 0905) and in national projects EcoNets, Bionets, OPUS project manager on V2X communication, and he managed numerous industrial and commissioned projects. He is a member of the IEEE Broadcasting Society, IEEE Communication Society, and IEEE Standard Association. He participated in the years 2012–2017 in the work of the IEEE P1900.x standardization group, as a member with voting rights, and also served as a Secretary. He is also a member of the groups: Radio Communications Committee and Research Group on Software Defined and Virtualized Wireless Access. From 2014 to 2016, he was the Membership Development/Web Visibility Chair in the IEEE for the EMEA area. Since 2019, he has been the Editor-in-Chief of the *Journal of Telecommunications and Information Technology* of the Institute of Communications.



HANNA BOGUCKA (Senior Member, IEEE) received the Ph.D. and the Doctor Habilitus degree in telecommunications from the Poznań University of Technology (PUT), Poznań, Poland, in 1995 and 2006, respectively. She is currently a Full Professor and the Director of the Institute of Radiocommunications, PUT. Moreover, she is the Co-Founder and the Head of Cooperation of RIMEDO Labs, the spin-off from PUT. She is involved in the research activities in the area of wireless communications: radio resource management, cognitive radio, and green-communication. She has been involved in multiple European 5th–7th Framework Programme and Horizon 2020 projects, the National Science Centre projects, and industry cooperation. She is the author of 200 research articles, three handbooks in the area of radio communications, and three scientific monographs on flexible and cognitive radio. She has been appointed as the IEEE Communications Society Director of the EMEA Region (Europe, Africa, and Middle East) and elected as the IEEE Radio Communications Committee Chair (2015–2016). She is currently the IEEE ComSoc Fog/Edge Industry Community Regional Chair in Europe and a member of the Polish Academy of Sciences.



KRZYSZTOF CICHON (Member, IEEE) received the Ph.D. degree from the Faculty of Electronics and Telecommunications, Poznań University of Technology (PUT), in 2017. He is currently employed at PUT, as an Assistant Professor. He has published over 20 papers in international and national journals and in international conference proceedings. He is active in IEEE DySpan-SC 1900.6 Group, as a Voting Member, since November 2014. He also participated in EU 7th Framework Programme projects, as well as in projects founded by the Polish National Science Centre. In 2017, he was awarded by Poznań's Chapter of the Polish Academy of Sciences for the best article published, in 2016, by the doctorate student.



JULIUS RUSECKAS received the Ph.D. degree in physics from Vilnius University, Vilnius, Lithuania, in 2006. From 2007 to 2020, he worked as a Research Professor with the Department of Physics, Vilnius University. Since 2019, he has been working with the Baltic Institute of Advanced Technology (BPTI), Vilnius, as a Researcher. He is a coauthor of more than 80 articles in peer-reviewed journals, mostly in the field of quantum optics. His current research interests include deep learning, computer vision, and neuromorphic computing.



GEDIMINAS MOLIS received the Ph.D. degree in physics, optoelectronics, from Vilnius University and the Center for Physical Sciences and Technology, Vilnius, Lithuania, in 2010. He worked at the Center for Physical Sciences and Technology, from 2002 to 2016, as a Researcher, and also in a period from 2008 to 2016, he worked at Teravil Ltd., a company producing THz range spectrometers, first as an Engineer and then as the Director. Since 2017, he has been working with the Baltic Institute of Advanced Technology, as a Researcher and the Program Manager. He is a coauthor of more than 30 articles in peer-reviewed journals, mostly in the field of optoelectronics. His current research interest includes the application of AI/ML methods in high frequency electronics systems.



AUŠRA MACKUTĖ-VARONECKIENĖ received the Ph.D. degree in informatics from Vytautas Magnus University, Kaunas, Lithuania, in 2007. She is currently a Researcher and a Lecturer with Vytautas Magnus University. Her Ph.D. dissertation topic was "Optimization methods for design of separation processes with possible irregularities of objective functions and feasible regions." Her current research interests include text mining, feature extraction and selection, signal and image analysis, text document classification/clustering, global optimization, multi-objective optimization, and multidimensional data visualization.



TOMAS KRILAVIČIUS is currently a Full Professor and the Dean of the Faculty of Informatics at Vytautas Magnus University, Kaunas, Lithuania. His scientific interests include the applications of artificial intelligence and language technologies, and large scientific infrastructures. He represents Lithuania in Information Systems Panel at NATO Science and Technology Organization. He has published a number of reviewed papers on the applications of AI and language technologies.

...