IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 10, 2021, accepted December 5, 2021, date of publication December 27, 2021, date of current version January 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3138888

Performance of Machine Learning-Based Techniques for Spectrum Sensing in Mobile Cognitive Radio Networks

MURAD A. ABUSUBAIH[1] AND SUNDOUS KHAMAYSEH[2]
[1]Department of Electrical Engineering, Palestine Polytechnic University, Hebron 00970, Palestine
[2]Department of Computer Engineering and Information Technology, Palestine Polytechnic University, Hebron 00970, Palestine
Corresponding author: Murad A. Abusubaih (murads@ppu.edu)

ABSTRACT Communication technologies are evolving drastically in recent years. However, the scarcity of spectrum began to appear with the accelerating usage of various communication technologies, as well as the preservation of traditional channel access methods. Cognitive Radio (CR) is an innovative solution for spectrum scarcity. Spectrum sensing is a key task of the CR life-cycle that gains significance as spectrum holes can be detected during this task. This paper studies and compares the performance of the **KMeans-based spectrum sensing technique** with the **non-cooperative spectrum sensing technique**, the **And-based spectrum sensing technique**, and the **Or-based spectrum sensing technique in stationary and mobile CR networks (CRNs).** The effect of the fading channel type has also been considered. Small-scale CRNs were simulated using the third version of the network simulator. In this context, two models have been developed. The first was built based on the well-known $\kappa - \mu$ general fading model to simulate the fading effects. The latter is the noise model to simulate different noise conditions. The results reveal that spectrum sensing techniques provide better performance in stationary networks as compared to mobile networks. Further, our experimental results show that at least three secondary users and about 1500 samples are needed to reach acceptable performance. In addition, the results show that the **KMeans-based technique** slightly outperforms the **Or-based technique**, especially in highly noisy environments and under severe fading channels.

INDEX TERMS Additive white Gaussian noise (AWGN), cognitive radio (CR), KMeans clustering, network simulator 3 - NS3, propagation model, spectrum sensing techniques.

I. INTRODUCTION

Cognitive Radio (CR) is an intelligence system able to switch between radio access methods as well as transmitting in different portions of the radio spectrum, [1]. The reconfigurability of the CR passes through cognition tasks which are: sensing the spectrum, analyzing the spectrum, and making joint decisions on spectrum selections. Spectrum sensing (SS) is the first task of the CR life-cycle that gains significance since the spectrum holes can be detected during this task. Spectrum sensing task operates in a **non-cooperative (Non-CSS)** or a **cooperative (CSS)** modes, whereby the secondary users (SUs) cooperate to determine the channel state. There are a plethora of works of spectrum sensing techniques in cognitive radio networks (CRNs). Most of these types are classified as energy detection-based, cyclostationary matrix-based, and covariance-based techniques. Machine learning-based techniques are another modern type of innovative spectrum sensing technique [2]. In such methods, the sensing process in detecting the primary user's activities passes through two phases which are: the feature extraction phase and the decision-making phase.

Literature studies conclude that the performance of the CSS schemes can be affected by many factors. First, the number of collaborating SUs. Second, the PU transmit power. Increasing the number of collaborating SUs and the PU transmit power can improve the performance. Third, the number of active PUs. Increasing the number of active PUs can deteriorate the performance because of the high interference. Finally, the number of training samples. Increasing the number of training samples increases the classification time and the computational complexity. However, these studies

The associate editor coordinating the review of this manuscript and approving it for publication was Tawfik Al-Hadhrami.

are conducted for stationary CRNs. As nodes in wireless networks are normally mobile, studying the performance of CSS techniques in mobile CRNs is crucial.

The contribution of this paper can be summarized as: firstly we study and compare the performance of the **Non-CSS**, the **And-based**, the **Or-based** techniques, as well as the **KMeans-based ML technique** in stationary CRNs. In contrast to the majority of published research, we then examine the performance of that mentioned spectrum sensing techniques in mobile CRNs. Also, we try to grasp the effect of the fading channels on the sensing performance. Moreover, we try to find the optimal parameters that can improve the performance of various spectrum sensing techniques. Finally, we investigate the circumstances in which KMeans-based spectrum sensing techniques provides superior to traditional techniques.

The rest of this paper is organized as follows: Section II provides a summary of some previous related works. Section III describes the spectrum sensing problem, theories, and basic concepts. Section IV describes the system model, while section V demonstrates the experiment's setup. Section VI discusses the results before we conclude the paper in section VIII.

II. RELATED WORK

In machine learning-based (ML-based) techniques, the classifier can apply soft or hard combining schemes for decision making after extracting features in the CSS mode. The determined type depends on the nature of the feature vectors. SUs in hard combining schemes implement a mechanism to digitize their local observation, while they explicitly exchange their local decisions in soft combining schemes. And/Or-based CSS techniques are the most common examples of hard combining schemes. In the And-based CSS, the channel is considered occupied if and only if all SUs determine that the primary user (PU) is active. Unlike, the Or-based CSS considers the channel to be busy if at least one SU determines that the PU is active, [3]. Classification of ML-based CSS schemes, based on the type of ML algorithm, include: (I) unsupervised-based CSS, (II) supervised-based CSS, and (III) reinforcement-based CSS, [4]

In the unsupervised-based CSS, the features are fed into the classifier without declaring their distinct labels. The k-means algorithm and its updated version, fuzzy c-mean, as ML classifier with feature extraction that based on energy detection (ED) were adopted in [5], [6]. In the ED-based scheme, the Normalized energy determines the channel state at a predetermined threshold. Another scheme based on the geodesic distance calculation was proposed in [7]. In this scheme, the geodesic distance is used as feature vectors. These features are derived from the sensing matrices and their Riemann means. Reference [8], [9] proposed eigenvalue/eigenvector-based schemes via applying the linear transformation on the sensing matrix. The same scheme with the Gaussian Mixture model was adopted in [10]. [11] introduces a signal processing scheme that combines the empirical mode decomposition algorithm and the wavelet threshold algorithm in order to remove the noise components and thus reduce the noise effects.

In the supervised-based, the features are filled along with their labels within the classifier to construct the final decision. Huge research has been conducted based on the support vector machine (SVM) [12], [13]. A probability vector with linear and polynomial kernel was proposed in [12]. This study aims to reduce the multidimensional energy vectors to two-dimensional probability vectors thus less training and testing time. Multi-class SVM for model large-scale CRN with ED-based scheme was adopted in [14], [15], whereas the supervised beamformer-based technique was proposed in [16]. In order to implement Multi-class SVM, an approach of classes coding is needed. For example, the one versus rest approach characterizes a particular class as a positive class and the rest as negative classes. Last but not least, [17], [18] studies applying deep learning in the sensing process.

Reinforcement-based techniques address specific problems such as power consumption level, throughput, energy efficiency, etc. The classifier is given rewards when evaluating its behavior. These awards are evaluated based on the problems intended to be solved. Reinforcement-based CSS techniques are not thoroughly studied in the literature. Sensing policy based on ϵ-*greedy* policy was proposed in [19]. The reward in this study represents the immediate throughput corresponding to the selected sub-band. Finally, an efficient sub-band selection policy based on replicated Q-learning was proposed in [20]. This technique introduces a partially observable Markov decision process that awards high rewards for a large number of idle channels while otherwise being awarded smaller rewards.

Though a lot of research has been invested in addressing the efficiency of ML-based spectrum sensing techniques, very few has considered the mobility effects [21] as well as different types of fading channels [5], which we believe would highly impact the results and thus constitute an important gap to be filled out.

III. PROBLEM STATEMENT

During the spectrum sensing task, SUs learn and perceive the surrounding environment, then detect the spectrum holes. Each SU that has the right access senses the spectrum, then, determines the PU signal existence. The channel is considered **idle** under the null hypothesis (H_0) condition and **busy** under the alternative hypothesis (H_1) condition [1]. The challenging problem is to find methods that accurately detect the existence of PUs transmissions.

A. ED-BASED SCHEMES

Let the received signals samples received by the SU denoted by $Z_n(k)$, where $Z(k)$ is the energy value corresponding sample k. Therefor, the total energy (E) is estimated by the n^{th} SU as:

$$E_n = \frac{1}{K}\sum_{k=1}^{K} |Z_n(k)|^2 \tag{1}$$

K is the number of samples. SUs observe the communication channel in the sensing duration and then estimate the value of E. The column vector from energy values $\{E_1, E_2, \ldots, E_n, \ldots, E_N\}^T$ is then used to determine the channel state.

B. KMEANS-BASED SCHEMES

The K-means algorithm maps the feature vectors to non-overlapping clusters, at the nearest Cartesian distance. Each possible cluster represents a set Ψ and is indexed by j. So, the KMeans-based methods try to find the Js of clusters corresponding to various channel states. Each cluster has its own centroid C_j that represents the cluster arithmetic mean. Therefore, the objective function of the KMeans-based CSS technique (i.e., the *distortion function,* Θ) is to find the minimum squared distance of overall clusters from their corresponding centroid as shown in the following equation

$$\Theta_{(\{\Psi_j\},\{C_j\})} = argmin \sum_{j=1}^{J} \sum_{l \in \Psi_j} \eta_{lj} ||l - C_j||^2 \tag{2}$$

where l is a feature vector and $\{.\}$ reflects the cardinality of Ψ_j and C_j sets. $||.||$ is the ℓ^2-norm, η_{lj} takes 1 if l is belong to Ψ_j and 0 otherwise. After training the clusters, the items and the centroid of each cluster become known.

In the testing phase, the classifier becomes able to make a suitable decision about the channel state, i.e., the channel is available or not. Let l' denotes the test vector, then the decision making can be defined as:

$$\frac{||l' - C_i||}{min_{j=1,\ldots,J} ||l' - C_j||} \geq \zeta \tag{3}$$

(here, $C_i < min(C_j)$ usually represents the noise cluster). The test vector l' is classified to the cluster C_i if (3) is satisfied. Otherwise, it is classified to the cluster C_j.

In the next sections we are interested in evaluating the performance of ML-based spectrum sensing techniques in mobile CRNs.

IV. SYSTEM MODELING

In this paper, we consider a small-scale CRN that consists of a PU network placed in the center and N multiple SUs around it. Two mobility scenarios are considered, which are the stationary and the mobile CRNs (they are respectively abbreviated StaCR and MobCR). The PU network in the two scenarios is the same and consists of a fixed PU transmitter (PU-Tx) and a fixed PU receiver (PU-Rx) placed at 15 meters far from the PU-Tx. SUs in the first scenario are also fixed and positioned 120 meters away from the PU network. In the second scenario, SUs are mobile and move randomly in the area (i.e., move in random direction and velocity). However, both scenarios start with the initial configuration as shown in figure 1.

Different levels of cooperation are considered. The non-cooperative mode is firstly considered, in which only one node participates in the sensing process. Then, different cooperative SUs (i.e., 2SUs, 5SUs, and 10SUs) are participating in the sensing process. In addition, the effect of various fading channels is examined. We adopt the general $\kappa - \mu$ fading model. The general $\kappa - \mu$ fading model was first proposed and examined in [22]. What makes this model preferable is that completely different fading channels can be controlled and modeled by two parameters of the distribution, the κ and the μ.

The general $\kappa - \mu$ fading model confirms that Rayleigh fading combines the Rician fading set and the Nakagami fading set. Rayleigh fading can be simulated as κ approaches 0 and μ equals 1. Different types of Rician fading can be modeled by fixing the μ parameter and tuning the κ. Also, different types of Nakagami fading can be modeled by fixing the κ parameter and tuning the μ parameter. **The experiments for other types of fading channels were carried out for the $\kappa \to 0, \mu = 3.5$ (Nakagami) and the $\kappa = 2.65, \mu = 1$ (Rician).**

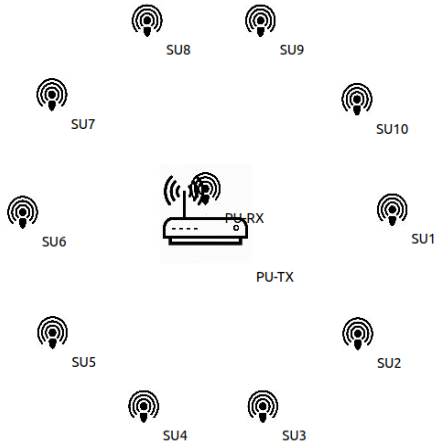


FIGURE 1. System modeling.

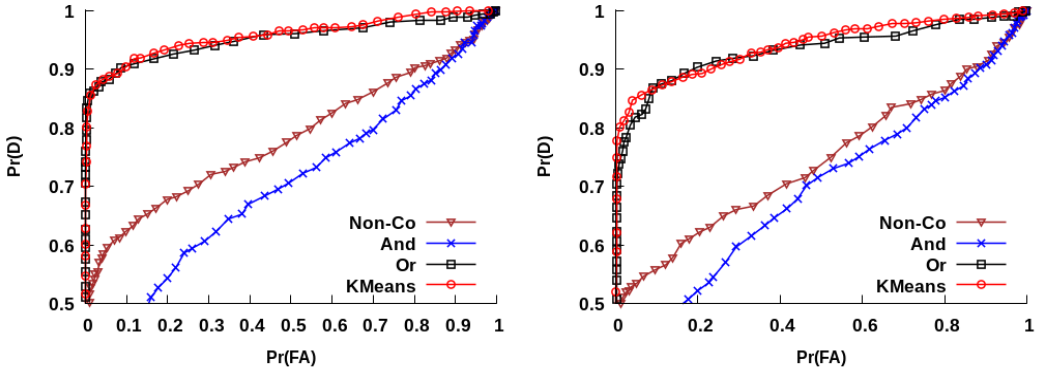
V. EXPERIENTIAL SETUP

We used the well-known third version of the discrete-event network simulator (ns3.30) to model a small-scale CRN and to generate datasets. In this model, we assumed that the PU network always operates at channel 36 of the IEEE802.11n-5GHz wireless technology. The Wi-Fi mode for unicast data frames is indexed to the 'HtMcs6' value which is a metric to several parameters of the Wi-Fi connection such as the 64-QAM modulation type, the 3/4 coding rate, and one spatial stream.

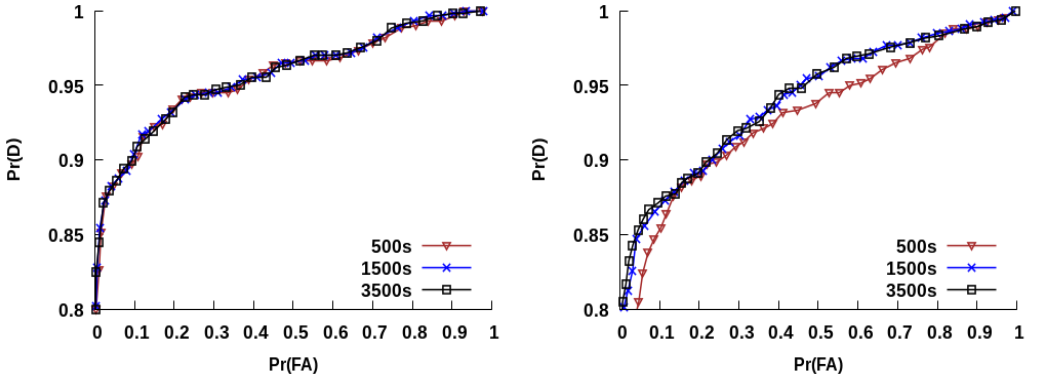
During the simulation time, the PU-Tx randomly broadcasts 1500 byte-length UDP packets to the PU-RX with a data rate of 5Mbps. SUs must observe and estimate the instantaneous signal-to-noise ratio (SNR) of each packet for overall decision making. In the simulation experiments, SUs listen to channel 36 for 5 ms per second, then estimate the normalized SNR over the entire simulation time. In our experiments, we assume that the probability of PU-Tx activity is 0.5.

We aim at investigating the performance of the KMeans-based technique as well as several other CSS technologies in the general $\kappa - \mu$ fading channel. Unfortunately, the ns3 package has no such model of this type of fading. Therefore, we developed our $\kappa - \mu$ fading model for the ns3 simulator. Here, we employed the well-known rejection sampling method to directly sample random variables from the $\kappa - \mu$ distribution.

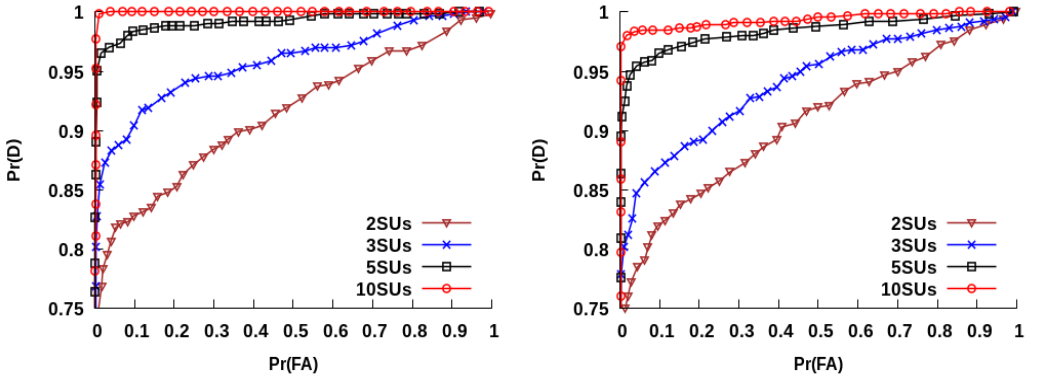
As aforementioned, ns3 is in nature a discrete-event simulator. Thus, we can only track and extract the data when the PU becomes active (i.e. during the ON-intervals). To overcome this problem, we developed a noise model using the python language. The noise model follows the Gaussian distribution and can be controlled using the mean of the



(a) ROC of different CSS techniques, [Rayleigh, N=3SUs, M=1500s]



(b) KMeans-based CSS with different training samples, [Rayleigh, N=3SUs]



(c) KMeans-based CSS with different number of collaborative SUs, [Rayleigh, M=1500s]

FIGURE 2. StaCR (left) vs. MobCR (right) of Rayleigh fading.

distribution (μ) as well as the standard deviation (σ). Large values for μ and σ indicate a very noisy environment while lower values indicate a low-noise environment. The resulting data was then combined with that of ns3 to obtain the complete dataset for experiments.

While the modulation and coding index (MCS Index) was set to 'HtMcs6'. The packet size and the application data rate were also set to 1500-byte and 5Mb/s respectively. We found the best configuration of the noise model to generate 2-samples per time key (time key indicates the second's index, i.e., 2-samples/s). The best number of samples taken during the sensing duration can be calculated by dividing the sensing interval by the packet size and the application data rate. This suits the data that was produced from the ns3 and reflects the probability of the PU-Tx being active. Therefore, the accurate selection of the MCS Index, the packet size, and the application data rate have a direct effect on the performance measurement.

VI. EXPERIMENTAL RESULTS & DISCUSSION

We started the experiments by assuming that the parametric noise model has -89.75 dBm and 1.0 dBm values for the μ and the σ parameters. The low-noise environment, abbreviated as Env1/1. Then, we raised these values to model

different channel conditions. Table 1 above depicts the various noise conditions that were considered.

TABLE 1. The different noise conditions.

**	low-noise environment		High-noise environment	
μ (dBm)	1.0		1.5	
σ (dBm)	-89.75	-89.25	-88.25	-87.5
Abbrev.	Env1/1	Env1/2	Env2/1	Env2/2

A. STATIONARY CRN VS. MOBILE CRN

The experiments were initially carried out for the Rayleigh fading (i.e., $\kappa \rightarrow 0$ & $\mu = 1$) and the low-noise environment Env1/1. Figure 2 and figure 3 show that the performance is generally better for the StaCR scenario as compared the MobCR scenario. Figure 2a-left shows that the And-based technique in the stationary mode and under Rayleigh fading provides the worst-case that is close to the Non-CSS performance. This is because when a SU wrongly detects the presence of the PU, it affects the global decision made by all SUs.

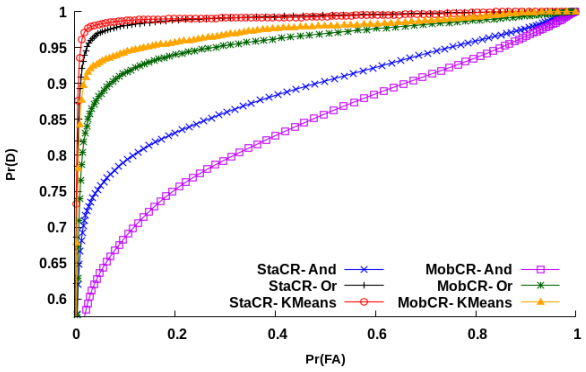
Figure 2a-left also shows that the Or-based and the KMeans-based techniques have comparable detection performance. In mobile CRN under Rayleigh fading, the KMeans-based technique is slightly superior to the Or-based technique, as clearly appears in figure 2a-right. However, the Or-based technique gives a good performance while the Non-CSS and the And-based techniques give a degraded performance.

Figure 2b and figure 2c show that the number of samples, M, has no clear effect on the performance, while the number of SUs, N, have an obvious effect on the performance of the KMeans-based technique regardless the mobility nature of the SUs nodes. Figure 2b and figure 2c generally show that we need about 1500 samples and at least 3 SUs to reach acceptable performance.

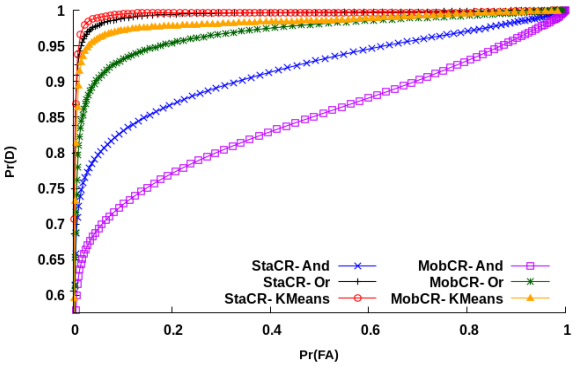
Figure 2 confirms that employing more than 5SUs and that above of 1500 samples for the sensing system does not improve the efficiency or accuracy of the system. The results also show that all techniques provide better performance with StaCR as compared to MobCR. This is due to the more dynamic nature of mobile channels, which introduces difficulties in identifying the presence of PU. This leads to a higher probability of false alarm, Pr(FA).

B. THE GENERAL $\kappa - \mu$ FADING SETUP

The experiments for other fading channels type were carried out for $\kappa \rightarrow 0$, $\mu = 3.5$ (Nakagami) and the $\kappa = 2.65$, $\mu = 1$ (Rician). In Nakagami fading channels, figure 3a, and Rician fading channels, figure 3b, it is clear that the KMean-based technique and the Or-based technique almost provide comparable performance that is better than the And-based technique in stationary CRN. On the other hand, the KMean-based technique outperforms other techniques in mobile CRN. In fact, when the κ and μ parameters increase, the performance of CSS techniques is consequently improving. Thus, the characteristic of the fading environment is highly affecting the



(a) Nakagami, [N=3SUs, M=3500s]



(b) Rician, [N=3SUs, M=3500s]

FIGURE 3. The performance of the various technique under the effect of the different fading channels.

performance of all CSS techniques as shown in the figures, figure 2a and figure 3.

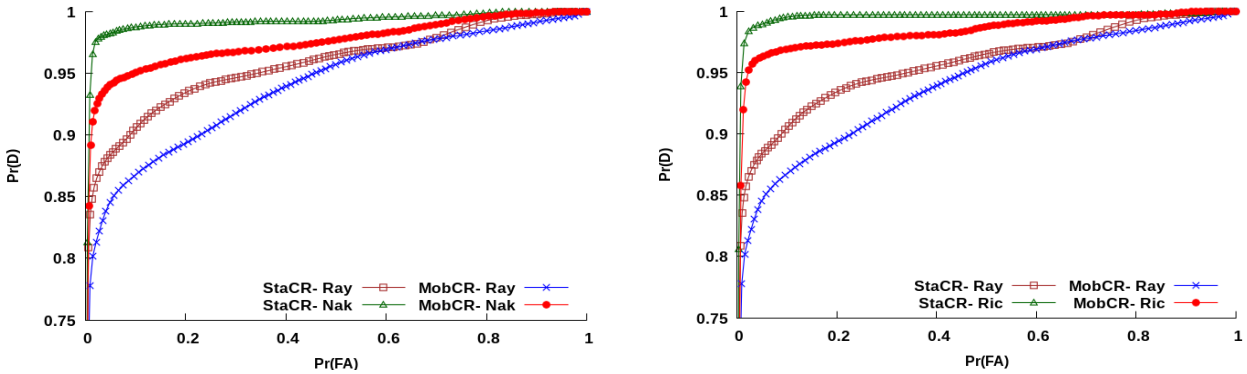
Figure 4 illustrates the performance of the KMean-based technique in Rayleigh fading compared to other fading distributions. Obviously, the results show that the performance of the KMean is better in the Nakagami and Rician fading channels. This is because when the value of κ or μ increases, the dispersion of the faded signal decreases. As a result, the distance between the clusters' centroids increases.

Figure 4 compares the performance of the KMean-based technique for StaCR and MobCR for those types of fading channels as well. We found that mobility causes performance degradation. Increasing the parameters κ and μ leads to a decrease in the dispersion of the fading signal. In contrast, the mobility increases the dispersion of the fading signal due to the spatial diversity effects of mobile nodes. Thus, the signal dispersion has a significant impact on the sensing performance.

C. NOISY ENVIRONMENT

In noisy environment, it becomes difficult for SUs to truly detect the presence of the PU signal. This is because they become unable to determine the nature of the captured signal. Figure 5 compares different CSS techniques in two different noisy environments, Env1/1 ($\mu = -89.75$, $\sigma = 1.0$) and Env2/1 ($\mu = -88.25, \sigma = 1.5$). It is clear that the performance of CSS deteriorates in a high-noise environment, especially for the And-based technique.

In StaCR, figure 5a, the performance of the KMeans-based and the Or-based technique are comparable in a low-noise



(a) Rayleigh vs. Nakagami, [N=3 SUs, M=1500]] (b) Rayleigh vs. Rician, [N=3 SUs, M=1500]

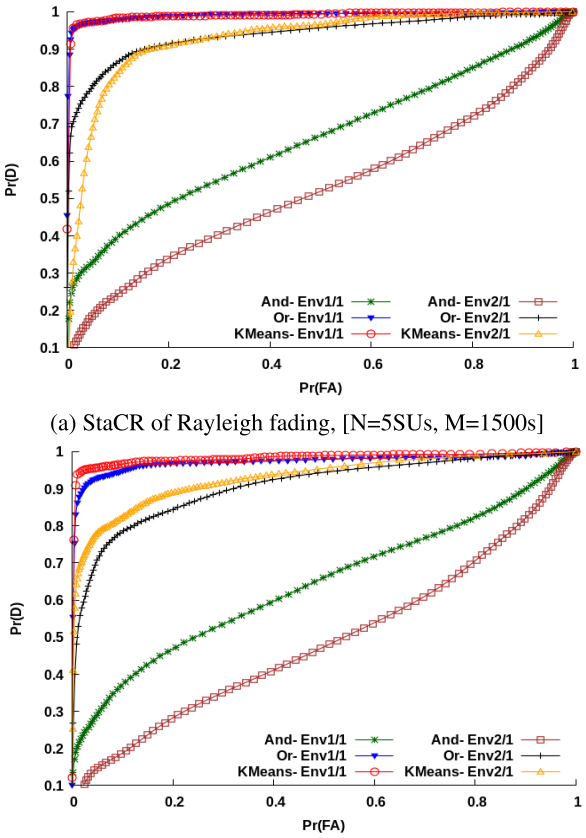
FIGURE 4. Comparison the effect of different fading channels on the performance of the KMeans-based technique.

environment. However, their performance is degraded in a high-noise environment. The performance of the Or-based techniques is slightly superior compared to the performance of the KMeans-based technique. But, the KMeans-based technique is obviously superior and more stable as compared to other techniques in the MobCR, figure 5b. Further, the results show that the performance of the And-based technique is degraded.

Figure 6 depicts the KMeans-based CSS performance of Rician fading. The data is depicted for two noisy environments, Env1/2 and Env2/1, which are characterized according to table 1. Figure 6b and figure 6d show that the mobility clearly affects the sensing data. The data is more dispersed in the mobility scenario, leading to altering the place of the centroids (i.e. the Euclidean distance in MobCR > the Euclidean distance in StaCR). This means that mobility is inevitably introduces low performance. Figure 6c and figure 6d show the effects of the noise level on the sensing data. Clearly, in the high-noise environment, the clusters become more condensed. Thus, clusters' centroids become more close, causing difficulties for the ML techniques to accurately classify the sensed data.

Table 2 summaries the numerical results for the noisy environments, Env 1/1 and Env2/1, under the effect of different fading channels. The table depicts the probability of detection, Pr(D), verses different reference points of probability of false alarm, Pr(FA), $r1 = 5\%$, $r2 = 10\%$, and $r3 = 15\%$. By comparing the measurements, it is clear that the Pr(D) decreases when the noise level is high, thus the performance in a high-noise environment (Env2/1, the second part of the table) is degraded. The second rows of various fading channels shows that the performance of the And-based technique is the lowest. The performance of this technique is far from reaching the value of 90% for Pr(D) versus the value of 10% for Pr(FA). For example, while the Pr(D) of stationary Rician fading was approximately equal 80% versus 10% of Pr(FA), it decreases too much less in most experiments (i.e., The Pr(D) is even lower 20% with the MobCR).

The performance of the Non-CSS techniques is low as shown in table 2. The Pr(D) was around forty percent versus 10% of Pr(FA) of the Rayleigh fading with Env2/1. table 2 also shows great convergence in the performance of the Or-based and the KMeans-based techniques. They achieve



(a) StaCR of Rayleigh fading, [N=5SUs, M=1500s]

(b) MobCR of Rayleigh fading, [N=5SUs, M=1500s]

FIGURE 5. The performance of the various technique under the effect of different noisy environments.

high performance that reaches 100 percent versus 10% of Pr(FA) in the KMeans-based techniques under stationary Rician fading and Env1/1.

Finally, figure 7 illustrates the mobility and noise effects on the performance of the KMeans-based technique for Rayleigh fading. As seen, the best performance corresponds to the lowest-noise environment, while the worst performance corresponds to the highest-noise environment. An exception here to the mobility conditions, the MobCR introduces better

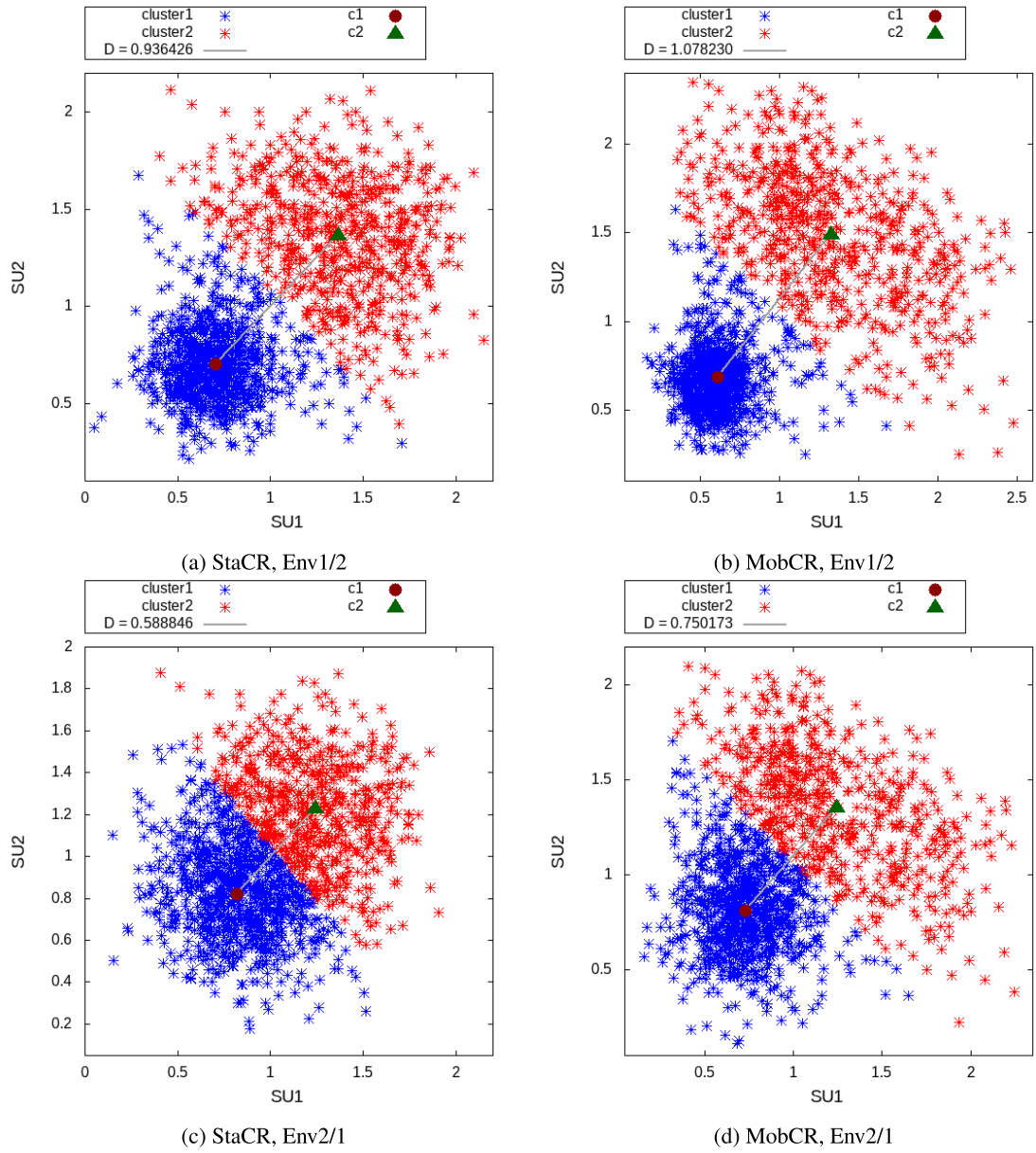


FIGURE 6. Data clustering for different environments, [Rician, M = 1500s].

performance compared to StaCR in high-noise environments. Actually, a deep thinking reveals that good performance comes from balancing the parameters of the Gaussian distribution and the parameters of the $\kappa - \mu$ distribution which is achieved in the high-noise environment.

VII. LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

We finish by the mention of some of the limitations and directions for future research.

A. LIMITATIONS RELATED TO THE DERIVED FEATURE VECTORS AND THE TYPE OF THE SPECTRUM SENSING TECHNIQUES

In mobile networks, spatial diversity that results from the different SUs' positions and their movement may provide useful information that can be utilized to accurately estimate the channel states. The accurate estimate of the channel states inevitably leads to massive improvement in the sensing performance. Several ideas can be addressed under this direction. Further research can be devoted to other ML-based spectrum sensing techniques, such as the GMM.

B. LIMITATIONS RELATED TO THE TYPE OF NETWORK TECHNOLOGIES, PROTOCOLS, AND STANDARDS

Correct tuning of sensing parameters, such as sensing duration and periods, probability of primary user activity, number of samples taken per sensing period, etc., has a direct impact on sensing performance. The selection of the exact values of these parameters is mainly modified based on the type of simulated networks and their protocols. Here, some limitations regarding network technologies and protocols are clearly visual and are worth considering as directions

TABLE 2. Summary of numerical results for the noisy environments Env 1/1 and Env2/1, [N = 5SUs, M = 1500s].

Low-noise environment, Env1/1

Row's order	Ch. type	Spectrum sensing techniques	StaCR $r1 = 0.05$	StaCR $r2 = 0.1$	StaCR $r3 = 0.15$	MobCR $r1 = 0.05$	MobCR $r2 = 0.1$	MobCR $r3 = 0.15$
1^{st}	Nakagami	Non-CSS	0.8281	0.8612	0.8827	0.6218	0.6838	0.7428
2^{nd}		And	0.7207	0.7537	0.7702	0.4660	0.5340	0.5810
3^{rd}		Or	0.9950	0.9950	0.9983	0.9682	0.9773	0.9773
4^{th}		KMeans	0.9967	0.9966	0.9966	0.9788	0.9803	0.9818
1^{st}	Rician	Non-CSS	0.8397	0.8777	0.9025	0.7318	0.7712	0.7909
2^{nd}		And	0.7719	0.8033	0.8182	0.5362	0.5756	0.6373
3^{rd}		Or	0.9967	0.9999	1.0000	0.97575	0.9803	0.9818
4^{th}		KMeans	1.0000	1.0000	1.0000	0.9863	0.9879	0.9909
1^{st}	Rayleigh	Non-CSS	0.5917	0.6247	0.6528	0.5348	0.5620	0.5954
2^{nd}		And	0.3272	0.4165	0.4500	0.2954	0.3848	0.4439
3^{rd}		Or	0.9719	0.9802	0.9851	0.9303	0.9450	0.9681
4^{th}		KMeans	0.9736	0.9851	0.9868	0.9560	0.9670	0.9700

Height-noise environment, Env2/1

1^{st}	Nakagami	Non-CSS	0.4850	0.5966	0.6650	0.4030	0.4810	0.5160
2^{nd}		And	0.4285	0.5040	0.5487	0.2239	0.2904	0.3237
3^{rd}		Or	0.8615	0.9005	0.9250	0.7700	0.8623	0.8950
4^{th}		KMeans	0.8876	0.9388	0.9586	0.8850	0.9107	0.9319
1^{st}	Rician	Non-CSS	0.5370	0.6430	0.7055	0.4136	0.4879	0.5530
2^{nd}		And	0.4770	0.5480	0.6100	0.2015	0.2712	0.3303
3^{rd}		Or	0.8980	0.9370	0.9640	0.7651	0.8742	0.9015
4^{th}		KMeans	0.9322	0.9669	0.9802	0.9045	0.9469	0.9560
1^{st}	Rayleigh	Non-CSS	0.3867	0.4297	0.4743	0.3820	0.4409	0.4636
2^{nd}		And	0.1965	0.2380	0.3060	0.1515	0.1742	0.2409
3^{rd}		Or	0.7960	0.8644	0.9000	0.7166	0.7924	0.8196
4^{th}		KMeans	0.6810	0.8331	0.8942	0.8181	0.8697	0.9010

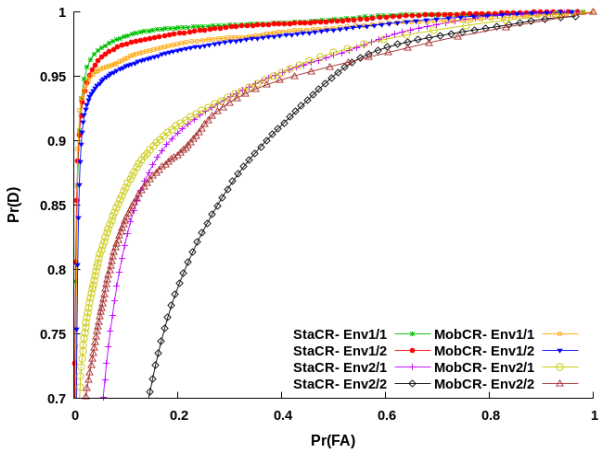


FIGURE 7. StaCR vs. MobCR of KMeans-base CSS for different noisy environments, [Rayleigh, N = 5SUs, M = 1500s].

for future research. In fact, the standards that govern these parameters constitute another aspect worth considering as well.

C. LIMITATIONS RELATED TO THE SECURITY ASPECT

Whereas different levels of SUs' cooperation are considered, we assume that all SUs that participate in the sensing process provide honest information about what has been observed from their point of view. However, the security aspect may be considered. Of course, if we assume scenarios when some adversary nodes deliberately manipulate the sensing data in order to falsy determine the channel state.

VIII. CONCLUSION

This work focused on the performance of KMeans-based techniques in mobile CRN. The effect of the type of fading channel was also investigated. A small-scale CRN was adopted and simulated using the well-known ns3 simulation platform. The general $\kappa - \mu$ fading channel was considered. The $\kappa - \mu$ fading signal was sampled using the well-known rejection sampling method. As ns3 is a discrete-event network simulator, there was a need to develop a noise model using python language. Performance measurements have been performed for the Non-CSS as well as for different types of CSS techniques.

Our experimental results reveal that the KMeans-based and the Or-based techniques with the stationary scenarios

provide the best comparable performance. The And-based and Non-CSS techniques provide the worst performance in stationary scenarios. In mobile CR, the And-based and the Non-CSS techniques provide highly degraded performance. Also, the performance of the KMeans-based and the or-based techniques is better as compared to the And-based and the Non-CSS techniques, but at same time not better than the stationary case. Further, the results show that at least 3+ collaborative SUs and at about 1500 samples are needed in order to improve the performance of the KMeans-based and the Or-based techniques. Finally, we found the performance of the KMeans-based technique is stable in high-noise environment as compared to the And-based and the Or-based techniques.

REFERENCES

[1] A. M. Wyglinski, M. Nekovee, and T. Hou, *Cognitive Radio Communications and Networks: Principles and Practice*. New York, NY, USA: Academic, 2009.
[2] Y. Arjoune and N. Kaabouch, "A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions," *Sensors*, vol. 19, no. 1, p. 126, Jan. 2019.
[3] Y. Arjoune and N. Kaabouch, "Principles and challenges of cooperative spectrum sensing in cognitive radio networks," in *Handbook of Cognitive Radio*. W. Zhang, Ed. USA: Springer, 2017.
[4] S. Khamayseh and A. Halawani, "Cooperative spectrum sensing in cognitive radio networks: A survey on machine learning-based methods," *J. Telecommun. Inf. Technol.*, vol. 3, pp. 36–46, Oct. 2020.
[5] V. Kumar, D. C. Kandpal, M. Jain, R. Gangopadhyay, and S. Debnath, "K-mean clustering based cooperative spectrum sensing in generalized $\kappa - \mu$ fading channels," in *Proc. 2nd Nat. Conf. Commun. (NCC)*, Mar. 2016, pp. 1–5.
[6] A. Paul and S. P. Maity, "Kernel fuzzy C-means clustering on energy detection based cooperative spectrum sensing," *Digit. Commun. Netw.*, vol. 2, no. 4, pp. 196–205, Nov. 2016.
[7] S. Zhang, Y. Wang, P. Wan, Y. Zhang, X. Li, and J. Li, "A cooperative spectrum sensing method based on information geometry and fuzzy C-means clustering algorithm," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 17, Dec. 2019.
[8] C. Sun, Y. Wang, P. Wan, and Y. Du, "A cooperative spectrum sensing algorithm based on principal component analysis and K-medoids clustering," in *Proc. 33rd Youth Academic Annu. Conf. Chin. Assoc. Autom. (YAC)*, May 2018, pp. 835–839.
[9] K.-J. Lei, Y.-H. Tan, X. Yang, and H.-R. Wang, "A K-means clustering based blind multiband spectrum sensing algorithm for cognitive radio," *J. Central South Univ.*, vol. 25, no. 10, pp. 2451–2461, Oct. 2018.
[10] G. C. Sobabe, Y. Song, X. Bai, and B. Guo, "A cooperative spectrum sensing algorithm based on unsupervised learning," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, 2017, pp. 1–6.
[11] Y. Wang, Y. Zhang, P. Wan, S. Zhang, and J. Yang, "A spectrum sensing method based on empirical mode decomposition and K-means clustering algorithm," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–10, Jul. 2018.
[12] 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.*, Apr. 2016, pp. 1–6.
[13] E. Ghazizadeh, D. Abbasi-moghadam, and H. Nezamabadi-pour, "An enhanced two-phase SVM algorithm for cooperative spectrum sensing in cognitive radio networks," *Int. J. Commun. Syst.*, vol. 32, no. 2, p. e3856, Jan. 2019.
[14] O. P. Awe and S. Lambotharan, "Cooperative spectrum sensing in cognitive radio networks using multi-class support vector machine algorithms," in *Proc. 9th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2015, pp. 1–7.
[15] S. U. Jan, V.-H. Vu, and I. Koo, "Throughput maximization using an SVM for multi-class hypothesis-based spectrum sensing in cognitive radio," *Appl. Sci.*, vol. 8, no. 3, p. 421, 2018.
[16] O. P. Awe, A. Deligiannis, and S. Lambotharan, "Spatio-temporal spectrum sensing in cognitive radio networks using beamformer-aided SVM algorithms," *IEEE Access*, vol. 6, pp. 25377–25388, 2018.
[17] 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.
[18] M. R. Vyas, D. K. Patel, and M. Lopez-Benitez, "Artificial neural network based hybrid spectrum sensing scheme for cognitive radio," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7.
[19] J. Oksanen, J. Lundén, and V. Koivunen, "Reinforcement learning based sensing policy optimization for energy efficient cognitive radio networks," *Neurocomputing*, vol. 80, pp. 102–110, Mar. 2012.
[20] M. A. Aref, S. Machuzak, S. K. Jayaweera, and S. Lane, "Replicated Q-learning based sub-band selection for wideband spectrum sensing in cognitive radios," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Jul. 2016, pp. 1–6.
[21] 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.
[22] M. D. Yacoub, "The $\kappa - \mu$ distribution and the $\eta - \mu$ distribution," *IEEE Antennas Propag. Mag.*, vol. 49, no. 1, pp. 68–81, Jun. 2007.

MURAD A. ABUSUBAIH received the B.Sc. degree (Hons.) in computer systems engineering from Palestine Polytechnic University, Palestine, in 1999, the M.Sc. degree in communications engineering from the Jordan University of Science and Technology, in 2001, and the Ph.D. degree in telecommunications engineering from the Technical University of Berlin, Germany, in 2009. During 2004 to 2009, he worked as a Research Assistant at the Telecommunication Networks Group, Technical University of Berlin. He has been working on policies and protocols for wireless networks. He was involved in different projects: cognitive radio, localization in wireless networks, and MESH networking. He has published several research articles in journals and conferences. Currently, he is interested in the development of mechanisms and policies for wireless networks management using machine learning and artificial intelligence tools. He has been involved in different international projects, especially with EU.

SUNDOUS KHAMAYSEH received the B.Sc. degree in information technology and communications from Al-Quds Open University, Hebron, Palestine, in 2014. She is a graduate student of the master's program in informatics with Palestine Polytechnic University, Hebron. Her research interests include employing innovative and intelligent solutions based on machine learning and its methodologies to develop wireless communication technologies and enhance their capabilities.