MHM: A Novel Collaborative Spectrum Sensing Method based on Markov-chains and Harmonic Mean for 5G Networks

Gabriel Ferreira*, Priscila Solis Barreto †, Geraldo P. Rocha Filho†,
Marcos F. Caetano†, Heikki Karvonen§, Johanna Vartiainen§

Department of Computer Science, University of Brasilia, Brasilia, Brazil
Email: *gabrielcarvfer@aluno.unb.br, †{pris,geraldof,mfcaetano}@unb.br
Centre for Wireless Communications, Oulu University, Oulu, Finland
Email: §{heikki.karvonen,johanna.vartiainen}@oulu.fi

***Abstract*—Cognitive radios and spectrum sensing are considered fundamental for spectrum optimization in 5G networks. Collaborative spectrum sensing improves detection by collecting data from different nodes and increasing the amount of information available for accurate channel state detection. However, malicious nodes can report wrong information, disturbing the collaborative sensing results and network operation. This paper presents two techniques: (1) a Markov chain-based technique that improves spectrum sensing accuracy while reducing the reporting control traffic; (2) a harmonic mean-based technique that discards less relevant sensing reports, mitigating Byzantine attacks. The two techniques were evaluated in a simulation scenarios based on rural areas. The results show that the proposed techniques increase the accuracy of a classic hard-combining fusion technique, reducing false positives and reporting overhead while improving network resilience to malicious nodes.**

***Index Terms*—collaborative spectrum sensing, byzantine attacks, MAC Layer, dynamic spectrum access, simulation, rural areas.**

I. Introduction

The massive use of mobile devices along with new internet services has generated an increasing demand for the already scarce spectrum of frequencies available. It is necessary to better use the available spectrum in order to provide appropriate levels of service. While 5G in urban areas is expected to grow substantially when compared to 4G, rural areas have been left behind due to low economic incentives.

Mobile network costs are driven by several factors such as technology, spectrum availability and licensing, topography and population distribution. Most of the spectrum is already allocated for different applications [1], free bands are rare and expensive, preventing entrance of new players in less lucrative areas, even though studies show that much of licensed spectrum is underused in non-urban areas [1].

Cognitive Radio (CR) enables dynamic spectrum access, which mitigates spectrum scarcity and increases the spectrum efficiency and utilization [2]. There are three types of wireless channel users: Primary User (PU), Secondary User (SU) and malicious user/attacker. PUs are exclusive users of their licensed bands, while SUs typically use unlicensed bands and want to use licensed bands in the absence of the PU. Malicious users try to deny the SUs from accessing the spectrum by cheating them. One of the CR techniques used to improve the spectrum utilization is the Dynamic Spectrum Access (DSA), where a non-licensee (SU), opportunistically uses a licensed band. The SUs is required to avoid interfering with the licensee by transmitting only when the PU is not detected, and back off from the channel as soon as PU transmissions are detected. The correct execution of the DSA increases the amount of bandwidth for the SUs applications [3]. Several Spectrum Management Agencies (SMAs) are evaluating alternatives to static spectrum allocation, including temporary licenses, like the license-shared-access (LSA) [4].

In CRs, one common approach is to integrate Collaborative Spectrum Sensing (CSS) techniques with DSA [1]. CSS is based on a detection function that runs on individual sensing nodes that reports their results to a central node Fusion Center (FC). The FC performs a fusion procedure to consolidate the reports and makes a single global decision about the channel state [5]. The fusion scheme is a key component of CSS and several research works approach this topic [6]. Another challenging issue in CSS is security, as it relies on information provided by third parties and has no way to validate it, what is known as the Byzantine generals problem. Malicious nodes may send false sensing data to the FC, increasing the probability of wrong results that lead to poor network performance or even unavailability. Some research works provide techniques to increase resilience to Byzantine attacks [7], but none of them focus on improving the CSS performance and reducing the reporting overhead along with Byzantine attack mitigation.

In this paper we propose a novel approach based on two techniques using Markov-chains to improve CSS fusion accuracy, while optimizing reporting overhead and increasing resilience against attackers, using a hard-combining fusion tech-

Annex to ISBN 978-3-903176-28-7©2020 IFIP

nique. Both techniques are evaluated in simulation scenarios of rural areas using CSS to opportunistically access underused bands. The results show that the proposed techniques have better accuracy than classic hard-combining fusion techniques.

This paper is organized as follows: Section II presents related works. Section III presents and discusses the proposed techniques. Section IV describes the simulation scenarios and analyses the simulation results. Finally, Section V presents the conclusions and future works.

II. Related Works

Markov-chain based solutions have been used in CR methods for the CSMA/CA protocol [8], where the medium access control is distributed. The Spectrum Sensing (SS) results are used to control the protocol back-off timer using a Markov-chain to guarantee a certain level of certainty that the channel is free before transmitting. Other CSMA/CA solutions [2], propose that each node executes its own SS and broadcast it to nearby nodes. Each node, then fuses the available reports and performs the CSMA/CA protocol on the detected free channel.

Regarding CSS, correctly identifying which nodes contribute more for the sensing is fundamental to improve the efficiency and the accuracy in PU detection [9]. Validating the User Equipments (UEs) reports with against retransmission metrics is one of the way to identify collisions with the PU. However, several factors influence the detection probability like shadowing, fading effects, and noise uncertainty. Choosing which UEs need to report may require additional information, like localization and energy constraints. Other ways to reduce overlapped sensing reports include selecting users with low correlation levels between them or to select devices with higher gain antennas, that can collect other local user reports with a side-channel and forward them to the FC [10]. In [11], the authors study grouping methods (random, reference-based, statistic-based and distance-based) to select adequate UEs to report, where only the random ones do not require localization information. In [12], the UEs are selected based on the received transmission power and distance to the Evolved Node-B (eNB).

Regarding malicious users/attackers, the resilience to Byzantine attacks is studied in [7], in which mitigation is done using voting, consensus techniques and machine learning. However, these works focus exclusively on the resilience instead of CSS performance. In [13], the authors propose a mitigation technique for false reporting of the Channel Quality Indicator (CQI) in the 3GPP LTE standard. The mitigation technique uses the number of retransmissions to determine if the reported CQI is inconsistent with the channel behavior.

The previous works show that Markov-chain based techniques are not unheard in CR and SS research, but are not commonly used, especially in CSS. To the best of our knowledge, in the related works listed to date, none addresses the resilience to malicious users/attackers considering optimization of control channel transmissions in CSS reporting, while maintaining low false positives and negatives for the collaborative spectrum sensing in remote area scenarios.

III. Proposal

This section presents two proposed techniques to increase CSS performance and resilience to attackers. The first technique filters noise of the hard-combining individual SS using a Markov-chain, reducing unnecessary reporting and saving Common Control Channel (CCC) bandwidth. The second technique filters less relevant reports from the CSS fusion, using a Markov-chain and the harmonic mean, to improve the resilience against attacks.

The techniques are based on the following premises:

- **P.1** the individual SS depends unique and exclusively on the UE to detect the state of the channel, which are represented as the probability of detection p_d and probability of false positives p_{fp} functions.
- **P.2** UEs, eNB and the PUs are completely or almost static.
- **P.3** $p_d >> 1 - p_d$ if the PU is active.
- **P.4** $p_{fp} << 1 - p_{fp}$ if the PU is not active.
- **P.5** the sensing period $t_{sensing}$ is much smaller than the PU transmission period $t_{putransmission}$.

From the premises **P.1**, the probability of detection is given by the Equation 1. The variable d refers to the distance between the UE and the PU while pu_{active} indicates whether the PU is active at a given time t. Considering the premise **P.2**, $d(t) = d(0) \forall t \in [0, \infty)$. The probability of k positive samples out of n-sized population is given by the Binomial distribution, shown in Equation 2.

$$p_{sense}(t, d, pu_{active}) = \begin{cases} p_d(d(t)), & \text{if } pu_{active}(t) = 1. \\ p_{fp}, & \text{otherwise.} \end{cases} \quad (1)$$

$$\binom{n}{k} p_{sense}^k (1 - p_{sense})^{n-k} \quad (2)$$

For premises **P.3** and **P.4**, the probability $\lim_{k \to n, n \to 0} \binom{n}{k} p_{sense}^k (1 - p_{sense})^{n-k} \simeq p_{sense}$ for a small n-sized population (e.g. $n \in [2, 3, 4]$). The probability $\lim_{k \to n, n \to \infty} \binom{n}{k} p_{sense}^k (1 - p_{sense})^{n-k} \simeq 0$. The side-effect of this behavior can be seen in the probability curve, which models the real spectrum sensing technique, $P_d(d)$ in Figure 4, where the source probability of detection (blue squares) is further attenuated as the number consecutive samples with same results $k = n$ grows.

We use the Markov chain described in Section to aggregate the UEs individual sensing results. This is valid if the premise **P.5** is correct, which means that multiple individual sensing procedures are executed for each PU transmission. The Markov chain transition states with probability P for each consecutive value, while different values lead to state transitions with probability $1-P$. After at least N consecutive results, the Markov chain reaches the final state S.

As each consecutive state K increases, we accumulate the certainty $C_{accum} = 1 - P^K$ of the previous K consecutive events is lost and the counting begins from the start. With a given threshold L for the accumulated probability, the number of required states S can be calculated. For example, if we model using a coin toss with $P = 0.5$ and we define a certainty threshold of $L = 0.9$, the number of number of states of the Markov chain is $S = 5$, with 4 state transitions and an accumulated certainty $C_{accum} > L$. The number of state transitions is then used to determine the attenuation effect of the number of states in the source probability P. In our case, the effects of the different state transitions ($K = N$) to the source probability of detection curve is shown in Figure 4, as calculated with Equation 2 and $P = P_{sense}$.

We consider that the sensing results have short-range statistical dependence, then, for the first technique our approach is to use a simple Markov chain with S states in the SS, which represents the counting of same consecutive results up to a given S state. The adaptation of the SS standard procedure is illustrated in Figure 1, with two additional checks used to implement the Markov chain. For each cycle, the current sensing result (R_{sense}) is compared to the previous one (R_{prev}). In case both are equal, certainty is accumulated (P_{accum}). With hard-combining, the sensing results are binary, and we model as a coin toss, which accumulates certainty in $\frac{1-P_{accum}}{2}$ steps. When the certainty P_{accum} is bigger than a given threshold (e.g. 90%), the sensing result R_{sense} is assumed to be correct and the variable that holds the last reported value sent to the FC (R_{markov}) is updated with the current sensing result. This value is then transmitted to the FC. If the result is different from the previous, the accumulated probability P_{accum} of the previous K consecutive events is lost, and the counting begins from the starting point.

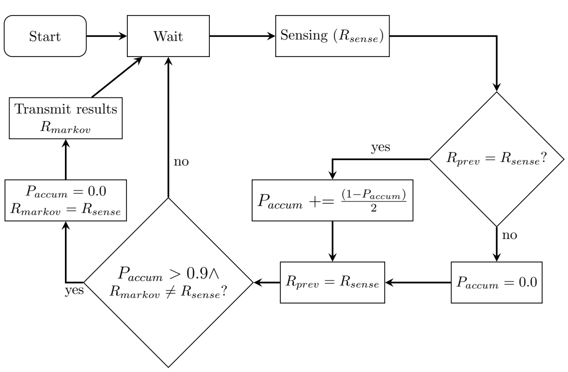


Fig. 1: Individual Spectrum Sensing with the Markov-chain

The second technique uses Markov chain based mechanism like in the first technique, which is used to discard reports from UEs that are less relevant to fusion, as shown in Figure 2. The filtering policy is based on the Harmonic mean of the CQI reported by the UEs, using Hybrid-ARQ metrics as a trust anchor. The first check in Figure 2 halves, following the same coin toss model, the relevance $relev$ of UEs that are either far from the PU or approximating to the eNB, since they are considered potential attackers. The second check increases the relevance $relev$ of UEs that reported no PU presence and have a stable CQI.

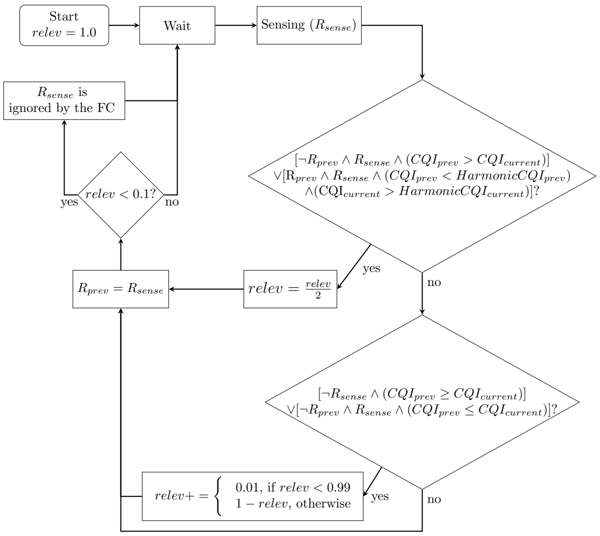


Fig. 2: Harmonic mean-based spectrum sensing report filtering scheme

IV. Simulation Results

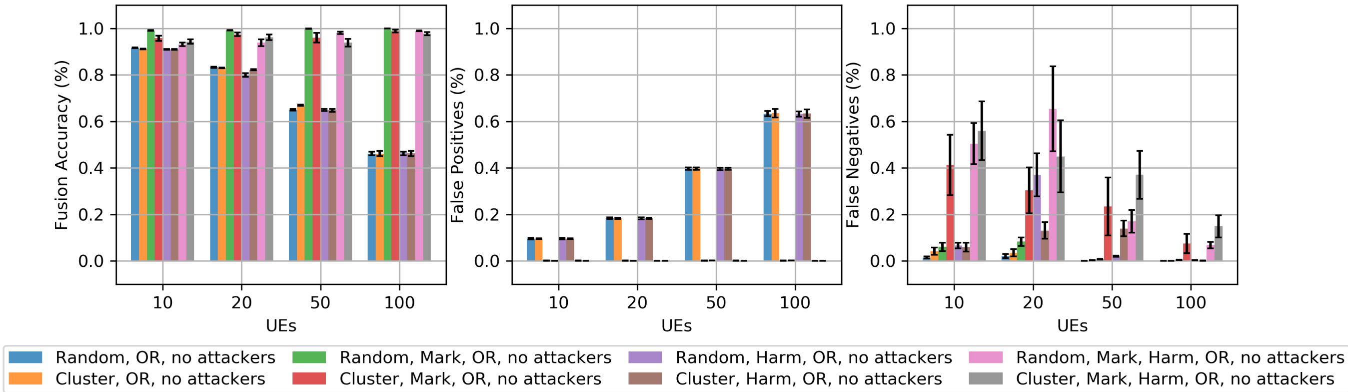
To evaluate the proposal described in the Section III, simulations of remote areas with a super-cells with 50km radius cell serving them created based on previous studies using the NS-3 simulator [14]. The simulation parameters are listed in the Table I. The SS probability of detection probability used are based on the link-layer results of the WIBA [14], also from the 5G-RANGE scenarios, and is shown in Figure 4.

	Scenario / Simulation parameters	Markov and Harmonic technique evaluation
General	Simulation time	10 s (10^4 subframes)
	Propagation model	5G-RANGE
	Band	5 ($\sim$ 850 MHz)
	Number of channels	4
	Channel bandwidth	3x5.2 MHz + 1x4.4MHz
	PUs per channel	1
	Noise floor	-174dBm/Hz
PU	Tx power	40 dBm
	Tx period	[1-5] s
	Tx duty cycle	[0.1-0.4]
eNB	Tx power	53 dBm
	Antenna gain	9 dBi
	Fusion techniques	[OR, AND, [2,3]-out-of-n UEs]
UE	Number of UEs	[10, 20, 50, 100]
	Number of attackers	[0, 1, 2, 5, 10]
	Tx power	23 dBm
	Antenna gain	9 dBi

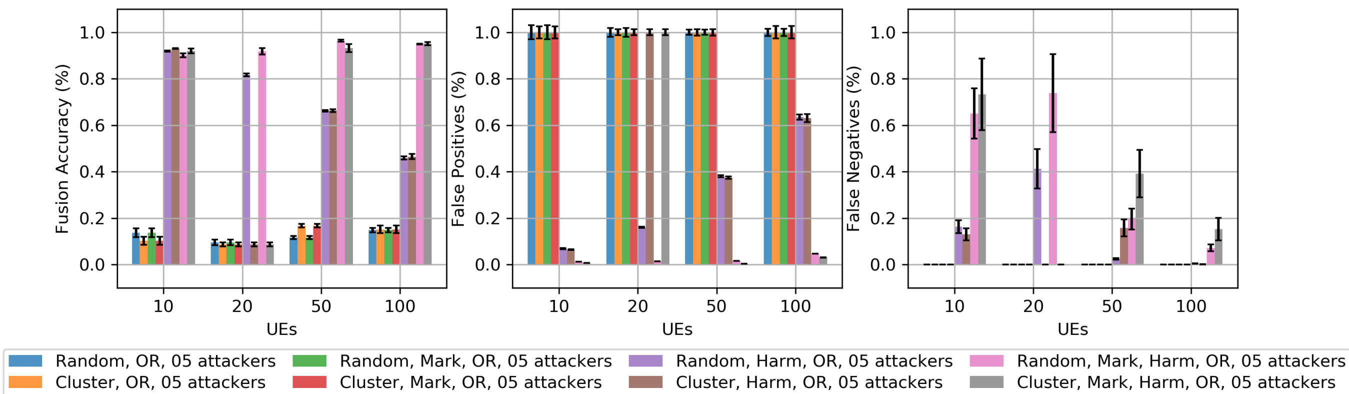
TABLE I: Simulation parameters for the different scenarios

The PUs are distributed randomly throughout the cell and the UEs are distributed either the same or into randomly placed clusters of 5 km radius. The randomly distributed scenario may be unrealistic for such a large cell, but provides the best-case scenario for most fusion techniques, as it provides more diverse data. The clustered scenario is more realistic and represents micro-regions, such as small villages.

Take SF as the number of simulated subframes, SF_{active} and $SF_{inactive}$ as the number of subframes the PU transmitted



(a) OR fusion with no attackers



(b) OR fusion with 5 attackers

Fig. 3: OR fusion results with different numbers of attackers, on clustered and randomized UEs distributions throughout the cell (Cluster and Random), with and without the Markov technique (Mark), with and without the Harmonic technique (Harm).

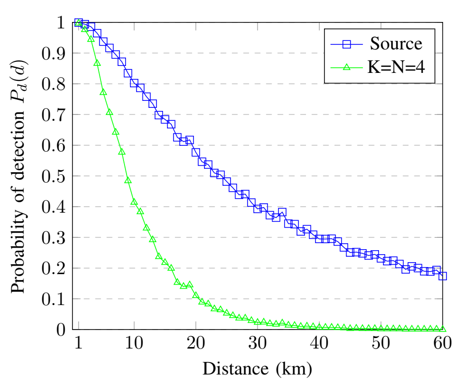


Fig. 4: Probability of detection curve based on distance (blue squares)

or not, FP and FN as the number of fusions indicating the PU presence when it was not transmitting, or the inverse. Three performance metrics are used to evaluate the techniques: false positive ratio ($P_{fp} = \frac{FP}{SF_{inactive}}$), false negative ratio ($P_{fn} = \frac{FN}{SF_{active}}$) and fusion accuracy ($Acc = \frac{SF-FN-FP}{SF}$).
Figure 3 shows the compiled results for the same scenarios, where Figures 3a, 3brepresent simulations with 0 and 5 attackers, varying UEs within 10, 20, 50 and 100. The results include the standalone fusion, along with its combinations with the proposed techniques. The figures illustrate how the Markov-chain proposal reduces false positives while keeping false negatives low (if there is no attacker) and how the harmonic-mean based technique mitigates attacks (behaving as the standard fusion in the scenario without attackers).

The performance for the OR fusion is shown in Figure 3a. For the case with 100 UEs, our proposal was able to reduce the false positives to a minimum (from 0.6323 ± 0.0110 to 0.0008 ± 0.0002 in the random scenario, a 790x reduction), at the cost of increasing false negatives (from 0.0000 ± 0.0000 to 0.0047 ± 0.0005, in the random scenario, and to 0.0744 ± 0.0419, in the clustered scenario). The large difference in the false negative values of the random and clustered scenarios shows how the number of UEs and their distribution within the cell may influence the results when using the Markov-chain based technique.

Figure 3a, shows that the accuracy and false positives are very similar for both the randomized and clustered cases. However, false negatives are wildly different. Figure 3b shows the scenario with 5 attackers. Both the OR fusion with or without the Markov-chain based technique behaves poorly, with nearly 100% of false positives. False negatives with the

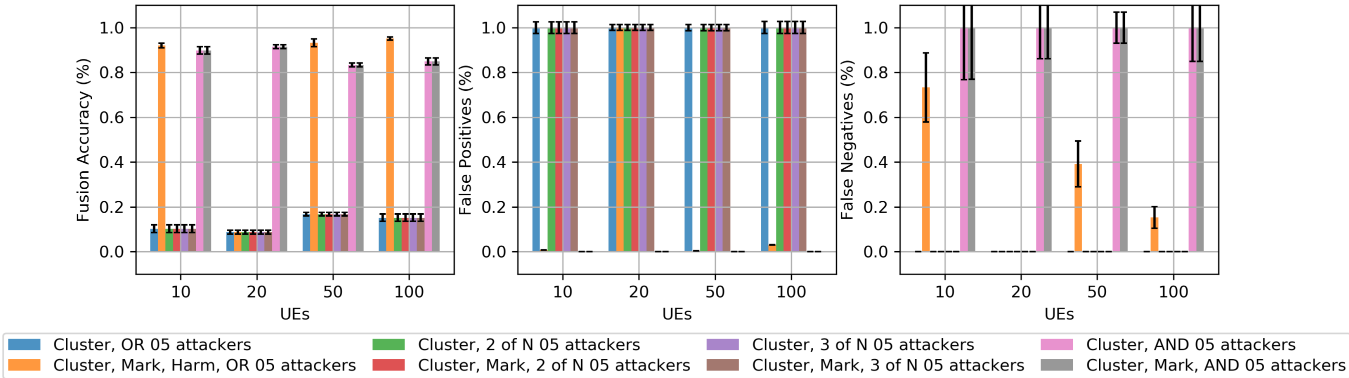


Fig. 5: Performance of different fusion techniques, including the combination of our proposed techniques in orange.

harmonic mean technique follow the same behavior as in the previous scenario.

To estimate the number of UEs to guarantee a certain level of false negatives, the data was fitted to an exponential curve using the least-squares method for both clustered scenario and random scenario. The resulting curves are $0.622063e^{-0.0127897x}$, with $R^2 \approx 0.937$, and $0.761142e^{-0.0226437x}$, with $R^2 \approx 0.995$. Using a 10% false negative threshold would require 143 and 90 UEs, respectively.

Figure 5 shows the comparison of simulation results using different fusion techniques. The effect of the harmonic-mean based filtering is very clear on Figure 5, where only the OR fusion mitigated the attackers action, except for the scenario with 20 UEs grouped into clusters.

V. Conclusions and Future Work

This paper presented two simple techniques based on Markov chains to improve collaborative spectrum sensing in networks for rural areas. The results show that the proposed techniques reduce the overhead of detection reports, enhancing control channel efficiency and at the same time, reducing false positives and false negatives, which enables accurate opportunistic use of the licensed spectrum and protects PUs transmissions. Further, the proposal improves network security by providing resilience for byzantine attacks in the spectrum sensing procedures. As future work we plan to integrate into the simulated scenarios techniques for resource scheduling using machine learning.

Acknowledgements

This work was supported by the 5G-RANGE (www.5g-range.eu), approved in the 4th EU-BR call in TICs, financed by the CTIC/RNP/MCTIC of Brazil and by the European Commission Research and Innovation Action H2020-EUB-2017 – EU-BRAZIL Joint Call Grant ICT-777137.

References

[1] F. Hu, B. Chen, and K. Zhu, "Full spectrum sharing in cognitive radio networks toward 5g: A survey," *IEEE Access*, vol. 6, pp. 15 754–15 776, 2018.

[2] L. T. Tan and L. B. Le, "Joint cooperative spectrum sensing and MAC protocol design for multi-channel cognitive radio networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, jun 2014.

[3] M. Szydelko and M. Dryjanski, "3gpp spectrum access evolution towards 5g," *EAI Endorsed Transactions on Cognitive Communications*, vol. 3, no. 10, p. 152184, feb 2017.

[4] S. Basnet, B. A. Jayawickrama, Y. He, E. Dutkiewicz, and M. D. Mueck, "Opportunistic access to PAL channel for multi-RAT GAA transmission in spectrum access system," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, jun 2017.

[5] K.-L. A. Yau, J. Qadir, C. Wu, M. A. Imran, and M. H. Ling, "Cognition-Inspired 5G Cellular Networks: A Review and the Road Ahead," *IEEE Access*, vol. 6, pp. 35 072 – 35 090, 2018.

[6] Y. Fu, F. Yang, and Z. He, "A quantization-based multibit data fusion scheme for cooperative spectrum sensing in cognitive radio networks," *Sensors*, vol. 18, no. 2, p. 473, feb 2018.

[7] L. Zhang, G. Nie, G. Ding, Q. Wu, Z. Zhang, and Z. Han, "Byzantine attacker identification in collaborative spectrum sensing: A robust defense framework," *IEEE Transactions on Mobile Computing*, vol. 18, no. 9, pp. 1992–2004, sep 2019.

[8] F. Foukalas, G. T. Karetsos, and P. Chatzimisios, "Cross-layer design of csma/ca with spectrum sensing for cognitive radio networks," in *ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems*, Aug 2013, pp. 1–5.

[9] Z. Khan, J. J. Lehtomäki, M. Mustonen, and M. Matinmikko, "Sensing order dispersion for autonomous cognitive radios," in *Proceedings of the 6th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications*. IEEE, 2011.

[10] K. Cichon, A. Kliks, and H. Bogucka, "Energy-efficient cooperative spectrum sensing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1861–1886, 2016.

[11] A. C. Malady and C. R. C. M. da Silva, "Clustering methods for distributed spectrum sensing in cognitive radio systems," in *MILCOM 2008 - 2008 IEEE Military Communications Conference*. IEEE, nov 2008.

[12] S. Wang, H. Liu, and K. Liu, "An improved clustering cooperative spectrum sensing algorithm based on modified double-threshold energy detection and its optimization in cognitive wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2015, pp. 1–7, 2015.

[13] K. Pelechrinis, P. Krishanmurthy, and C. Gkantsidis, "Trustworthy operations in cellular networks: The case of PF scheduler," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 292–300, feb 2014.

[14] G. Ferreira, P. S. Barreto, M. Caetano, E. Alchieri, J. Vartiainen, H. Karvonen, M. Matinmikko-Blue, and J. Seki, "A tool for developing collaborative sensing and cognitive mac layer solutions for 5g in rural areas," in *6th International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, 2019.