Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Técnicas para Melhorar a Eficiência do Sensoriamento Colaborativo em Redes 5G para Áreas Remotas

Gabriel de Carvalho Ferreira

Dissertação apresentada como requisito parcial para
conclusão do Mestrado em Informática

Orientadora
Prof.a Dr.a Priscila América Solis Barreto

Brasília
2020

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Técnicas para Melhorar a Eficiência do Sensoriamento Colaborativo em Redes 5G para Áreas Remotas

Gabriel de Carvalho Ferreira

Dissertação apresentada como requisito parcial para
conclusão do Mestrado em Informática

Prof.a Dr.a Priscila América Solis Barreto (Orientadora)
CIC/UnB

Prof. Dr. Marcelo Menezes de Carvalho
ENE/UnB

Prof. Dr. Geraldo Pereira Rocha Filho
CIC/UNB

Prof.a Dr.a Genaína Nunes Rodrigues
Coordenador do Programa de Pós-graduação em Informática

Brasília, 13 de março de 2020

Dedicatória

Dedico este trabalho a Deus, minha mãe e família, minha orientadora, amigos, professores e todos que me acompanharam nesta saga.

Agradecimentos

Gostaria de agradecer a todos que me ajudaram nesta empreitada.

Agradeço em especial minha orientadora. Profa. Priscila Solis, que me guiou, me encorajou e me deu a oportunidade de participar do projeto 5G-RANGE.

Também agradeço aos parceiros do projeto 5G-RANGE pela oportunidade de trabalhar com eles, em especial pela ajuda dos doutores Heikki Karvonnen e Johanna Vartiainen, da Universidade de Oulu, que proveram as curvas de probabilidade de detecção, fundamentais para o trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

Resumo

A revolução dos *smartphones* em 2007 iniciou a um processo de crescimento exponencial da demanda por serviços de telefonia móvel. O aumento da demanda sem contrapartida da oferta, dependente do espectro disponível provoca uma queda na qualidade dos serviços prestados. As técnicas que usam Rádios cognitivos e acesso dinâmico ao espectro são consideradas fundamentais para otimizar a utilização do espectro e aumentar a quantidade de banda disponível para as redes 5G, ao permitir acesso oportunístico ao espectro licenciado ocioso.

Diversos estudos apontam a subutilização de bandas, especialmente longe das grandes cidades, em que há menor demanda e menor incentivo econômico para a instalação de infraestrutura por parte das operadoras. Esse comportamento é incentivado devido ao processo de licenciamento de bandas em blocos e alocação estática do espectro, em que uma operadora licencia uma banda e junto a ela fica encarregada por dar cobertura a uma área atrelada à licença, enquanto pequenas operadoras locais ficam completamente de fora dos leilões e são impedidas de competir neste mercado.

O acesso dinâmico ao espectro depende de informações que garantam a identificação de transmissões no canal candidato, afim de se reduzir interferência ao detentor da licença do canal.

Algumas das técnicas mais comuns para se detectar a ocupação do canal via sensoriamento do espectro são *carrier-sense* e detecção de energia, dependendo da largura do canal. O sensoriamento colaborativo melhora a capacidade de detecção de uso do canal quando comparado com o sensoriamento individual, visto que diversifica geograficamente a informação disponível para análise. A qualidade do sensoriamento colaborativo depende não só dos sensoriamentos individuais recebidos, mais também da técnica que consolida ou executa a fusão desses resultados. Existem diversos algoritmos de fusão, cada um com vantagens e desvantagens. Algumas das técnicas de fusão clássicas são baseadas em votação k-em-n, em que k sensoriamentos indicando ocupação do canal resultam em uma fusão indicando ocupação do canal. A fusão 1-em-N, OU lógico, resulta em um número alto de falsos positivos, detectando ocupação do canal mesmo quando está desocupado, enquanto minimiza falsos negativos e a não detecção do canal de fato ocupado.

Por fim, é parte do ciclo de sensoriamento colaborativo filtrar sensoriamentos de usuários maliciosos que desejam perturbar não só o resultado do sensoriamento colaborativo como o funcionamento da rede. No caso de uma fusão simples como OU lógico, um único nó malicioso é capaz de inviabilizar por completo o uso oportunístico do espectro ao transmitir resultados falsos indicando que o canal está ocupado quando de fato está livre.

Diante essa problemática, neste trabalho são propostas duas técnicas para melhorar os resultados do sensoriamento colaborativo, a saber : (1) uma técnica baseada em cadeias de Markov que aplicada aos resultados de sensoriamentos individuais, reduz falsos positivos e falsos negativos, além de reduzir o envio de mensagens de controle ; (2) uma técnica baseada na média harmônica para filtragem de resultados de sensoriamentos individuais recebidos, descartando sensoriamentos de nós mais distantes das fontes de interferência, protegendo de ataques Bizantinos. Ambas as técnicas são avaliadas em cenários de 5G na área rural, em que encontra-se a maior porção de bandas do espectro subutilizadas, candidatas ao acesso oportunístico.

A fim de permitir a avaliação das técnicas propostas, foram realizadas diversas alterações para o modelo de pilha de rede LTE implementado no simulador de redes a nível de sistemas ns-3. As alterações incluem os procedimentos de sensoriamento do espectro individual feito pelos dispositivos de usuários (UEs), a transmissão dos resultados para o ponto de acesso (eNodeB), a fusão dos resultados recebidos e utilização do resultado de fusão no escalonamento de recursos para os dispositivos. Os sensoriamentos individuais são obtidos a partir de curvas de probabilidade de detecção e probabilidade de falsos positivos feitos através de medições em experimentos ou através de simulações a nível de camada física-enlace. As curvas são carregadas durante a configuração inicial da simulação, sendo interpoladas conforme necessário. As curvas podem ser tanto baseadas em distância euclidiana quanto em relação sinal ruído e interferência (SINR). O sensoriamento individual consiste em utilizar a probabilidade de detecção relacionada a um dado valor de SNR ou de distância euclidiana é utilizada para gerar uma amostra aleatória a partir de um gerador com distribuição de Bernoulli. O procedimento se repete a cada 1 milissegundo no ciclo padrão de indicação do subquadro LTE.

A técnica baseada em cadeias de Markov se baseia em um Teorema Central do Limite, em que a média de um certo número de amostras uniformemente distribuídas tende a se aproximar ou ao valor real da distribuição de probabilidade fonte ou ao valor central da distribuição. Em outras palavras, ao amostrar uniformemente uma distribuição desconhecida com número suficiente de amostras, encontra-se uma boa aproximação para o valor real que é procurado. Este princípio é aplicado para o sensoriamento individual do espectro, em que o valor do último sensoriamento é comparado com o resultado atual, e

quando idêntico aumenta o grau de certeza de que este resultado é de fato o real. Quando os resultados diferem, o grau de certeza é perdido. Quando um dado limiar de certeza é ultrapassado, o resultado do sensoriamento que é de fato transmitido para o eNodeB é substituído pelo valor do último sensoriamento. A modelagem deste processo estocástico binomial é feita baseado no lançamento de N moedas, em que apenas o caso em que N resultados iguais consecutivos levam à troca do valor transmitido, sendo facilmente modelado como uma cadeia de Markov de $N - 1$ estados.

Já a técnica baseada em média harmônica se baseia no fato de que as estações próximas das fontes de interferência são mais confiáveis que estações distantes, baseando-se nas curvas de probabilidade de detecção. Com isto, é necessário eliminar os resultados de sensoriamentos informados por usuários maliciosos com alguma informação adicional que sirva de prova que seu sensoriamento reportado é falso. Uma das maneiras de se mitigar informações falsas é utilizando a média harmônica dos CQIs reportados, permitindo identificar UEs mais afetados pela fonte de interferência e descartar todos os resultados por UEs pouco afetadas, mais afastadas da fonte. Para poder se confiar no CQI reportado pelos UEs, é necessário medir a quantidade de retransmissões feitas para cada uma delas. Uma taxa de retransmissões próxima de 10% indica um CQI adequado, enquanto taxas próximas de 0% indicam CQI reportado abaixo do real e taxas acima de 10% indicam CQI reportado acima do real. O limiar de retransmissões é definido nos padrões 3GPP.

A avaliação das propostas foi feita em duas partes: primeira parte com a validação do modelo proposto para o sensoriamento colaborativo no modelo do padrão LTE do simulador, e a segunda parte avaliando o desempenho das técnicas propostas. Durante a validação, foi confirmado o comportamento esperado do sensoriamento colaborativo (sensoriamentos individuais, transmissão dos resultados e fusões) em termos de taxas de falsos positivos e taxas de falsos negativos quando comparado com os modelos matemáticos. Na avaliação do desempenho das técnicas propostas foram avaliadas acurácia, taxas de falsos positivos e taxas de falsos negativos. Em ambos os casos, foram utilizados cenários inspirados em zonas rurais, com: baixo número de nós (10, 20, 50, 100); uma célula com 50 quilômetros de raio; canal de 20 MHz na banda 5 com portadora em 870 MHz; eNodeB transmitindo à 53 dBm; UEs transmitindo à 23 dBm; eNodeB e UEs com antenas com 9 dBi de ganho; detentor da licença do canal (PU) transmitindo à 40 dBm ou 35 dBm; um PU por subcanal de 5 MHz; algoritmos de fusão simples. O cenário de validação foi pouco realístico, com UEs dispersas ao longo de um certo raio fixo de distância do PU, garantindo uma mesma probabilidade de detecção para todos os UEs. Os cenários de avaliação das técnicas foram separados em dois conjuntos, um menos realístico com dispersão aleatória pela célula, outro mais realístico com dispersão aleatória dos PUs pela célula e dispersão aleatória de grupos de UEs pela célula, formando *clusters* de UEs.

Os resultados mostram que as técnicas propostas aumentam a acurácia em relação à técnica clássica de fusão de resultados do sensoriamento colaborativo (fusão OU lógico, ou 1-em-N), reduzindo falsos positivos em até 790 vezes, de 63.23% para 0.08% no cenário com dispersão aleatória dos UEs e sem atacantes. Neste mesmo cenário houve um aumento de 0% para 0.47% do número de falsos negativos, sem impactar severamente o detentor da licença do canal. Nos cenários com atacantes, todas as fusões simples apresentam resultados ruins, com ou sem a técnica das cadeias de Markov, até 100% de falsos positivos, inviabilizando o acesso oportunístico. Já a técnica da média harmônica apresenta bom grau de proteção contra atacantes, em especial nos cenários com mais dispositivos. Sem a técnica baseada em Markov e no cenário com 100 UEs, dos quais 10 atacantes, conseguiu reduzir falsos positivos da fusão OU de 100% para 60%, sem aumentar significativamente o número de falsos negativos. Quando as duas técnicas são combinadas, o número de falsos positivos cai para 5% enquanto falsos negativos sobem para 18%. Nos cenários com menos UEs e com *clusters*, falsos negativos são consistentemente mais altos, porém superiores às fusões 2-em-N, 3-em-N e E utilizando a técnica de Markov no cenário sem atacantes. Em todos os cenários, a técnica baseada em cadeias de Markov também reduziu a taxa média de notificação dos quadros em 2 ordens de grandeza, economizando banda do canal de controle licenciado. Esses resultados permitem concluir que ambas as técnicas são efetivas para o cenário rural para a qual foram propostas.

Também se depreende que o número de estados da cadeia de Markov e da técnica da média harmônica podem ser alterados para se trocar alcance da detecção por certeza da detecção e nível de proteção contra atacantes por falsos negativos, respectivamente.

Como trabalhos futuros, cabem a adaptação da técnica para: incluir cenários urbanos, mais densos, utilizando técnicas de amostragem; utilização de técnicas de localização (e.g. Time-of-Arrival, Angle-of-Arrival) para segmentação da célula em setores; melhorar a técnica da média harmônica para reduzir falsos negativos mantendo o mesmo nível de proteção contra atacantes.

Palavras-chave: redes 5G, acesso oportunístico, acesso dinâmico ao espectro, sensoriamento colaborativo, rádio cognitivo, rural

Abstract

The smartphone revolution of 2007 started an exponential demand growth of mobile connectivity. The ever increasing demand requires an increase in supply, which is depends in the amount of available spectrum. The amount of available spectrum however is limited, curving supply growth and reducing the quality of services perceived by the users. Cognitive radio and dynamic spectrum access are essential to increase the spectrum utilization and amount of available bandwidth in 5G networks, by opportunistically accessing unused licensed spectrum.

The dynamic spectrum access depends on channel information that guarantees the detection of transmissions in the candidate channel, as a means of reducing interference to the channel licensee.

The collaborative spectrum sensing increases the channel usage detection capacity when compared to individual spectrum sensing, as there is more geographically diverse information for analysis and decision-making. The quality of the collaborative sensing depends not only on the individual sensing that feeds information into it, but also on the technique that fuses those results into the final sensing result.

Two techniques to improve the collaborative spectrum sensing results are proposed in this dissertation: (1) a technique based in Markov chains to smooth consecutive individual spectrum sensing results, reducing both false positives and false negatives, while enabling the reduction of sensing reports by skipping sensing reports with the same results; (2) a technique based in the harmonic mean of the channel quality indicator, used to filter the received individual spectrum sensing, discarding nodes far from the source of interference, mitigating against Byzantine attacks. Both techniques are evaluated in rural 5G scenarios, which are the best place to use opportunistic access due to the amount of unutilized and underused spectrum bands.

In order to evaluate the proposed techniques, a set of modifications to the LTE network stack model of the ns-3 system-level simulator is proposed. The modifications include a complete collaborative sensing cycle, including: the individual spectrum sensing procedure, performed by user equipment's (UEs); the transmission of control messages to the access point (eNodeB), the fusion of the received results and utilization of the free

spectrum for the UEs. The individual spectrum sensing is performed by interpolating probability of detection curves and false positive probability, which are produced either by experimental measurements or by link-layer simulations.

The evaluation of the proposals was made in two parts: first part focusing on the validating the collaborative spectrum sensing cycle implementation and integration to the LTE model; second part focusing on the performance of the proposed techniques. The collaborative spectrum sensing cycle (individual sensing, sensing report and fusion) was validated and closely follows the mathematical model. The evaluation of the techniques included accuracy of the fused result, false positive and false negative rates.

The results show the techniques are effective in increasing the accuracy of the collaborative sensing when compared to the standalone classic fusion techniques (OR fusion, or 1-out-of-n). There were reductions in false positives rates of up to 790 times, from 63.23% to 0.08% in the scenario with randomized dispersion of UEs across the cell and without attackers. In the same scenario, the false negatives increased from 0% to 0.47%, which does not severely impact the licensee with interference. All classic fusions behave very poorly in scenarios with attackers, with and without the Markov chain technique. False positive rates soar to as high as 100%, making the opportunistic access impossible. The harmonic mean-based technique reduces the false positives, providing good protection against attackers especially in scenarios with more UEs. The harmonic mean alone reduced false positives for the OR fusion from 100% to 60% without significantly impacting false negatives in the scenario with 100 UEs and 10 attackers. When both techniques are used together, the rate of false positives fall to 5% while false negatives increase to 18%.

Scenarios with less UEs and distributed in clusters tend to have higher false negative rates when both techniques are used, but false positives are consistently lower than other classical fusions (e.g. 2-out-of-N, 3-out-of-N and AND). The Markov chain technique effectively reduced the sensing report rate by 2 orders of magnitude, saving up scarce control bandwidth. These results allow us to conclude that the both techniques are effective for the rural scenario they were proposed.

Keywords: 5G networks, dynamic spectrum access, opportunistic access, collaborative spectrum sensing, cognitive radio, rural

Contents

List of Figures

List of Tables

Acronyms

3GPP	3rd Generation Partnership Project.
AoA	Angle-of-Arrival.
AP	Access Point.
AWGN	Additive White Gaussian Noise.
CCC	Common Control Channel.
CDMA	Code Division Multiple Access.
CLT	Central Limit Theorem.
CQI	Channel quality Indicator.
CR	Cognitive Radio.
CS	Carrier-Sensing.
CSMA	Carrier-Sense Multiple Access.
CSMA/CA	Carrier Sense Multiple Access-Collision Avoidance.
CSS	Collaborative Spectrum Sensing.
CTS	Clear-To-Send.
DCI	Downlink Control Information.
DSA	Dynamic Spectrum Access.
ED	Energy Detection.
ELM	Extreme Learning Machine.
eMBB	Enhanced Mobile Broadband / Extreme Mobile Broadband.
eNB	Evolved Node B.
EPC	Evolved Packet Core.
ETSI	European Telecommunications Standards Institute.

FC	Fusion Center.
FCC	Federal Communications Commission.
FDD	Frequency Division Duplex.
FDMA	Frequency-Division Multiple Access.
FM	Frequency Modulation.
FNO	Fixed Network Operator.
FSM	Finite State Machine.
FSPL	Free-Space Path Loss.
FT	Fusion Technique.
GAA	General Authorized Access.
GBR	Guaranteed Bit Rate.
gNB	5G Node B.
GPRS	General Packet Radio Service.
GSM	Global System for Mobile Communications.
GSMA	GSM Association.
GTP	GPRS Tunneling Protocol.
HARQ	Hybrid Automated Repeat reQuest.
HSS	Home Subscriber Server.
IEEE	Institute of Electrical and Electronics Engineers.
KNN	K-Nearest Neighbors.
LBT	Listen-Before-Talk.
LSA	License Shared Access.
LTE	Long Term Evolution.
MAC	Medium Access Control.
MarkCh	Markov Chain.
MCS	Modulation and Coding Scheme.
MIMO	Multiple-Input and Multiple-Output.
MME	Mobility Management Entity.
MMTC	Massive Machine Type Communications.

mmWave	Millimeter Waves.
MNO	Mobile Network Operator.
MonCar	Monte Carlo.
Mu MIMO	Multi-user MIMO.
NN	Neural Network.
NR	New Radio.
ns-3	Network Simulator 3.
OFDM	Orthogonal Frequency-Division Multiplexing.
OFDMA	Orthogonal Frequency-Division Multiple Access.
P-GW	PDN Gateway.
PCRF	Policy Control and Charging Rule Function.
PD	Preamble Detection.
PHY	Physical layer.
PU	Primary User.
QAM	Quadrature Amplitude Modulation.
QCI	QoS Class Identifier.
QoS	Quality of Service.
RB	Resource Block.
RBG	Resource Block Group.
RNTI	Radio Network Temporary Identifier.
RTS	Request-To-Send.
S-GW	Serving Gateway.
SC-FDMA	Single Carrier FDMA.
SDCSS	Semi-Distributed Cooperative Spectrum Sensing.
SDF	Service Data Flow.
SINR	Signal to Interference plus Noise Ratio.
SMAs	Spectrum Management Agencies.
SNR	Signal to Noise Ratio.
SS	Spectrum Sensing.

SSDF	Spectrum Sensing Data Falsification.
STA	Station.
SU	Scheduling Unit.
SVM	Support-Vector Machine.
TDD	Time Division Duplex.
TDMA	Time Division Multiple Access.
TFT	Traffic Flow Template.
ToF	Time-of-Flight.
TVWS	TV Whitespace.
UCI	Uplink Control Information.
UE	User Equipment.
URLLC	Ultra-Reliable and Low Latency Communications.
WCDMA	Wide-band CDMA.
WIBA	Window-Based Energy Detector.

Chapter 1

Introduction

The mobile phones, created by Motorola in 1973, revolutionized the world offering phone calls over a Frequency Modulation (FM) analog signal between the cellphone and the cell. Considering the potential impact of this technology, governmental agencies and companies created standards to guarantee interoperability, reduce costs and simplify the operation of mobile networks [1]. The European Union created a mobile network work-group, Global System for Mobile Communications (GSM), that is part of the European Telecommunications Standards Institute (ETSI) [1].

The latest widely deployed mobile standard is the Long Term Evolution (LTE), one of the 4G standards [1]. The LTE network is based on IP packets, uses a newer and more efficient Medium Access Control (MAC) technique, that can aggregate multiple channels to increase user throughput and can use multiple antennas, known as Multiple-Input and Multiple-Output (MIMO), to increase the number of streams transmitted/received by the users. The MAC techniques are known as Orthogonal Frequency-Division Multiple Access (OFDMA), used in the downlink, Single Carrier FDMA (SC-FDMA), user in the uplink. The combination of the new features in the LTE networks allowed unprecedented levels of efficiency and availability of radio resources, improving the data rates in the networks and making possible the use of multimedia applications [1].

While the mobile networks advanced, so did the rest of the technology industry. The plummeting costs of data storage, transmission and processing [16, 17], along with a surge in component reliability and management complexity, paved the way for on-demand infrastructure as a service, which became known as the cloud.

The highly flexible architecture of the cloud along with new radio technology, inspired mobile carriers, governments and standard agencies to work on the new architecture for mobile networks, known as 5G [18, 16]. On the radio side, the 5G increases the amount of available spectrum by including new lower frequency and higher frequency bands [19].

However, the additional high frequency bandwidth is not useful for underserved remote areas such as rural zones. Remote areas have low population density and are not economically viable for mobile operators with the default 4G/5G infrastructure. High transmission power and antenna gain, along with low frequency bands are required for long-range transmissions and super large cells.

The long-range transmissions are quite common in services like broadcast radio and TV, emergency and military telecommunication systems, radio astronomy, among others. These service providers pay steep prices for the spectrum band licenses in Spectrum Management Agencies (SMAs) auctions. The licenses indicate the target use of the band, the region in which the license is valid and are usually associated with coverage goals. The coverage goals and prices prevent smaller network operators from providing local mobile services, even though part of the spectrum remains underused, as large operators can fulfill their coverage goals by serving larger urban centers [20, 21, 22].

In order to increase the amount of available bandwidth in remote areas, increase competition in the market and provide better service for those communities, it is necessary to either change how current licenses work or provide options to access underused spectrum opportunistically [23, 24, 25, 26]. Some SMAs have proposed new licensing models, General Authorized Access (GAA), which gives priority of access to the channel licensee, or Primary User (PU), instead of exclusivity, allowing non-licensees, or Scheduling Units (SUs), to use the channel when it is free. The GAA requires that the SUs stop transmitting as soon as the PU indicates the intent to use the channel [27], which can be done via radio signaling or a centralized database managed by the SMAs.

The opportunistic use of underused spectrum is usually done via Dynamic Spectrum Access (DSA) techniques, where SUs listens to candidate channels, select the less used ones and opportunistically access any of them [23]. The SUs is required to detect the PU transmissions as soon as possible and back off from the channel, avoiding interference [23, 24, 25, 28, 29, 26].

The detection of the channel state by the PUs is done using a spectrum sensing technique implemented on the radio Physical layer (PHY), along with processing logic in the MAC layer [30]. However, relying on the spectrum sensing techniques is challenging, as they are affected by a multitude of factors including the radio itself, noise caused by the ambient temperature and unknown sources, multipath due to topography of the area, fading due to blocking, channel bandwidth and frequency plus transmission power from the source [30]. Spectrum sensing techniques end up trading off range of detection, probability of detecting a transmission when one is happening, probability of detection, and probability of detecting a transmission when none is happening, probability of false positives.

The spectrum sensing can also be performed by multiple nodes, improving the detection range and reliability of the spectrum sensing, both essential for opportunistic spectrum access and use in super large cells [30]. This technique is known as Collaborative Spectrum Sensing (CSS), which relies on the sensing nodes sharing their results with the other nodes, which then merge the results using a Fusion Technique (FT) [30]. The FT trades the same properties of the individual spectrum sensing and is also responsible for filtering bad reports [8, 31]. The CSS can be performed in a distributed fashion or a centralized one, being the centralized the most appropriate for mobile networks due to their protocol design [30].

Multiple FTs have been proposed over the years, trying to improve the accuracy of the CSS. The classic FTs are based on voting, also known as k-out-of-n [32]. As machine learning techniques such as Support-Vector Machines (SVMs), Neural Networks (NNs) and K-Nearest Neighbors (KNN) became more popular, more FTs using them were proposed as in KNN in [33, 9]. More recently, FTs using and Extreme Learning Machine (ELM) were proposed in [34, 35], increasing detection accuracy even further, even with a relatively small number of sensing nodes. The machine learning techniques have increased accuracy at the cost of less flexibility (e.g. fixed width and trained for a specific scenario) and more processing power, increasing the energy usage and the latency between the sensing and avoidance, in case it is necessary.

Other FTs focus on mitigating other issues, like filtering of bad reports from malicious users [36, 37, 38, 39]. The forged reports can be used to mislead the FT in order to disrupt the opportunistic access necessary to maintain the SU network, or cause interference to the PU.

Considering the above mentioned problematic, this work presents two techniques to improve the accuracy of the individual spectrum sensing technique of a classic FT in 5G rural networks. The first technique enables DSA using the TV Whitespace (TVWS) while protecting channel licensees with low rate of missed detections (false negatives), incurring in a small latency between the start of the PU transmission and the interference avoidance, while also reducing spectrum sensing reporting overhead. The second technique protects against false positive reports from malicious users, by tracking their reported channel quality, which is unreliable due to self reporting nature, but made reliable anchoring its results with error rates of the medium access control layer using the technique proposed in [40].

1.1 Goal

This work objective is to enable better spectrum utilization by dynamically accessing unused spectrum and provide protection against attackers in 5G rural networks.

1.2 Specific Goals

The specific goals of this work are:

- develop a technique to improve the confidence of spectrum sensing reports, reducing false alarms and false negatives, while reducing the amount of control transmissions.
- develop a technique capable of reducing the impact of malicious user attacks.
- implement and verify the techniques in a simulation tool

The contributions of this work are the proposed techniques which aim to improve spectrum usage altogether with security issues in 5G networks. Also, another contribution of this work are the contributions made to an open source simulation tool. Several changes were implemented in the LTE stack to enable the DSA for the data channel by using the LTE cross-carrier scheduling, that can easily be ported to the new 3rd Generation Partnership Project (3GPP) standard, 5G-NR. These developments are available for further improvement and development in the NS-3 code [41]

1.3 Work structure

This work is structured in 5 chapters. Chapter 2 presents the literature review for the mobile network's evolution, 4G and 5G, medium access control, spectrum sensing and collaborative spectrum sensing technologies, and how these are used in dynamic spectrum access. In Chapter 3, the proposed techniques for collaborative spectrum sensing techniques are presented and discussed. Chapter 4 describes the methodology for performance evaluation and the simulation scenarios, plus the simulation results and their analysis. The work conclusions are presented in Chapter 5.

Chapter 2

Literature Review

This chapter presents a review of the theoretical concepts that are fundamental for the development of this work. Also, the chapter presents a compilation of the latest related research works of interest in the area.

2.1 Mobile networks

The growth of telecommunications after the second world war, the establishment of the internet in the 1980s and its popularization in the 1990s created a huge demand for mobile communications [1]. The first generation of mobile networks (1G) erupted in the end of the 1980s with two different standards, one in the USA and other in the EU [1]. Both standards used analog radio transmissions to transmit voice [1]. Devices were bulky, batteries were heavy and non very portable, the call quality was poor and mainly used as second or last option for wired communications.

Soon after, the EU created an agency responsible for telecommunication standard development and certification, known as European Telecommunications Standards Institute (ETSI) [1]. The agency formed a group with academic and industry partners, which proposed and approved the second generation of mobile standards (2G), the Global System for Mobile Communications (GSM) [1]. The group became known as the GSM Association (GSMA). The GSM standard enabled a multitude of new services, including text and voice messages and an internet-like service, along with interoperability between mobile carriers [1].

The 2G dropped analog transmissions in favor of digital ones. As the demand for the internet access grew in the early 2000s, two groups were formed to study and propose the next generation [1, 42]. The first group, the 3rd Generation Partnership Project (3GPP), ended up with the most support from the industry and the Wide-band CDMA (WCDMA) prevailed. The new standard introduced packet-switching for data and kept

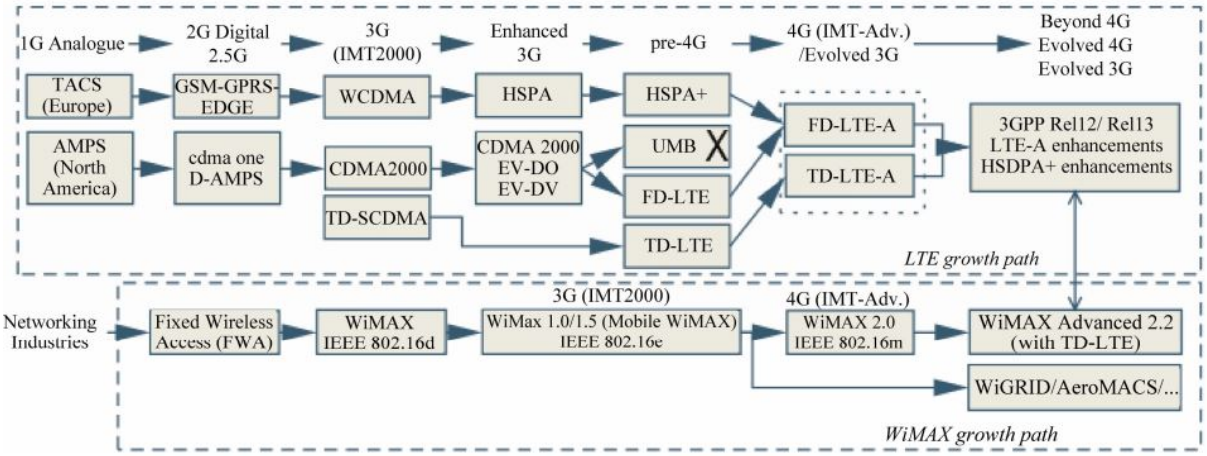


Figure 2.1: Mobile network standards evolution[1]

circuit-switching for voice. The Code Division Multiple Access (CDMA) also increased the spectrum usage efficiency when compared to the Time Division Multiple Access (TDMA) used previously [1, 42].

With the advent of smartphones in late 2000s, bringing the Web and internet-services to the user hands, there was a huge surge in demand for more bandwidth [42]. The 3GPP continued working on improving the standards and keeping it backwards compatible to previous standards, allowing mobile operators to slowly upgrade their infrastructure [42].

In 2008, the 3GPP released the 4G standards, known as Long Term Evolution (LTE) [43]. The LTE introduces a multitude of improvements, including complete deprecation of circuit switching, new MAC protocols that allows for much greater granularity of spectrum allocation, increasing even more utilization [44, 45]. Some improvements were the increase in bits per symbol with a bigger Quadrature Amplitude Modulation (QAM) constellation, allowing for greater efficiency; multiple simultaneous transmissions streams to a single, with Multiple-Input and Multiple-Output (MIMO), or multiple users, with Multi-user MIMO (Mu MIMO), for simultaneous data streams to a single or multiple users; increased carrier aggregation numbers, vastly increasing the maximum bandwidth capacity for the users, giving the ability to better serve traffic bursts [43, 46].

The Figure 2.1 shows the evolution path of each standard for every generation up to the 4G. The LTE standard was later updated to LTE-Advanced and LTE-Advanced Pro, which further increased the maximum number of aggregated carriers, along with QAM constellation size and improvements to the resource scheduling and synchronization protocols to reduce the latency between request for more band and the response [47].

New applications and services brought as well new users' profiles and traffic patterns to the Internet. For example, on-demand video stream services like with Netflix, YouTube

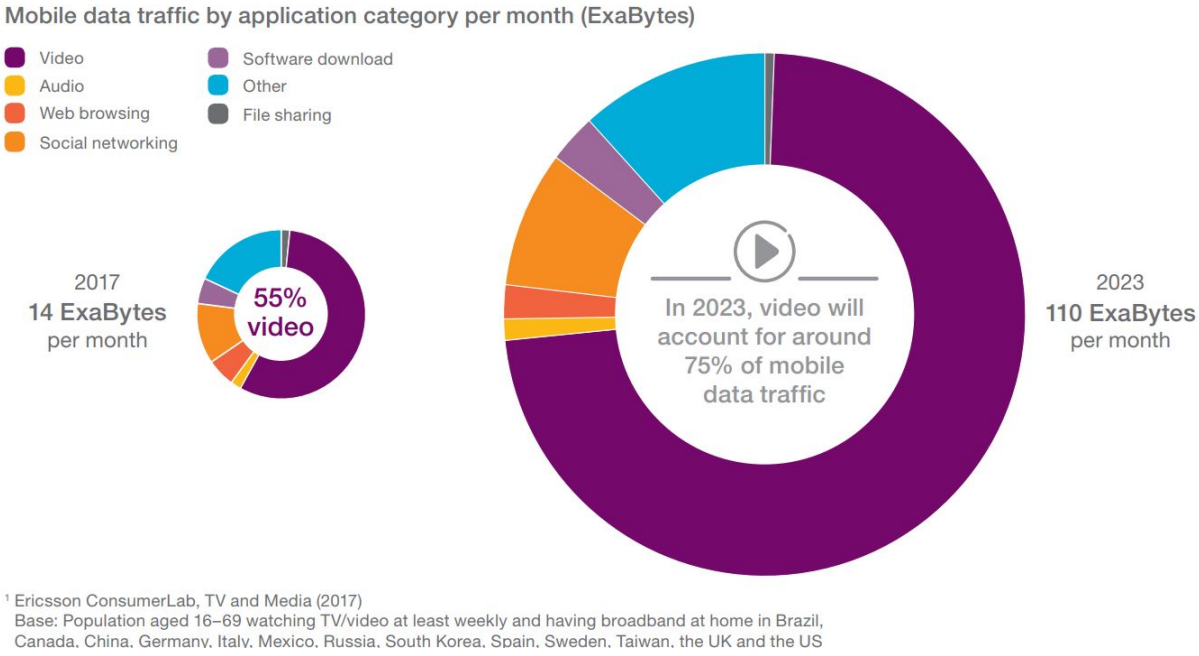


Figure 2.2: Project mobile data growth per application group [2]

and others changed the main traffic growth factor for network operators, projected to increase up to 10 times in the next 6 years [2], as shown in Figure 2.2.

The quick evolution from 2G to 4G brought many issues to mobile operators, which got indebted not only to redeploy different generations of the network, but also to buy expensive spectrum licenses that also required coverage expansion [48, 49]. To solve the bandwidth and latency demand, while keeping costs down, the 3GPP and ETSI proposed a revamped architecture for the 5G New Radio (NR) standard. The new architecture is based on the cloud concept, which replaces most of the hardware in favor of virtualized network functions running on the edge and core networks, connected to different cells and the internet [50, 51]. The standard defines support to different access technologies, including the so called non-3GPP networks [50, 52]. The bandwidth and latency demands are solved by including even more spectrum to the available pool, most of it from the millimeter wave spectrum range [19] and dividing the infrastructure into slices, which is a core concept of the standard that aims to optimize network usage for different applications [50, 52, 51].

In the following sections, essential concepts of LTE/NR components used in this work are explained and further detailed.

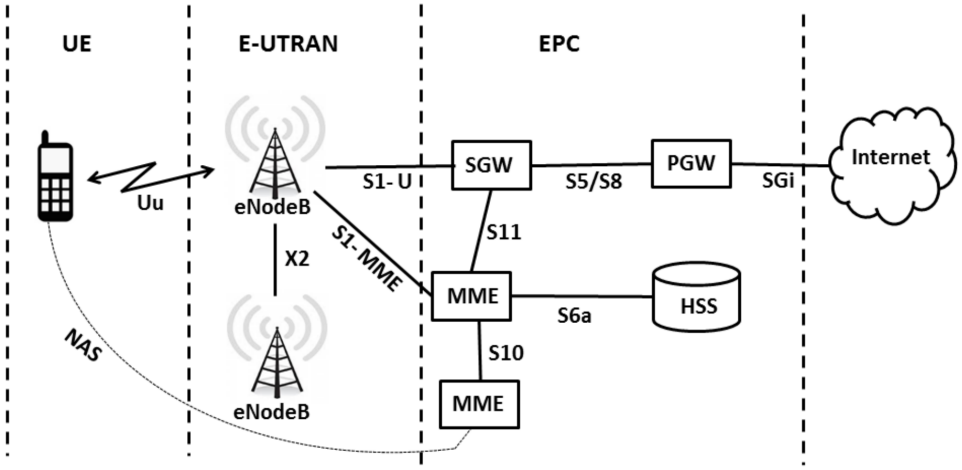


Figure 2.3: LTE Architecture [3]

2.1.1 4G LTE and 5G NR

The 4G LTE and 5G NR are the latest mobile network standards, which share a myriad of similarities. The main differences between them comes from the virtualized nature of the NR infrastructure, focused on cost reduction, maintainability and upgradeability, while LTE relies on dedicated hardware for most of its components.

2.1.1.1 Architecture and Stack Overview

Both LTE and NR architectures can be divided into the radio access entities and the backhaul entities, also called Evolved Packet Core (EPC) [44, 53]. The radio access entities include the User Equipment (UE) and the base station. The base station is known as Evolved Node B (eNB) in the LTE and as 5G Node B (gNB) in NR [44, 53].

The LTE architecture is shown in Figure 2.3, showing the UE-eNB connection, along with: the Serving Gateway (S-GW), that tunnels internet packets transferred between the UE and the PDN Gateway (P-GW); the P-GW, that acts as an intermediary between the external server and the UE; the Mobility Management Entity (MME), that manages UE roaming, authentication, connection management and configures eNBs; Policy Control and Charging Rule Function (PCRF), that intercepts user traffic and bills accordingly; Home Subscriber Server (HSS), which manages roaming and authentication of users connected to the network [44].

Differently from the approach of Institute of Electrical and Electronics Engineers (IEEE) 802.11 [54] based networks, that rely on different frames to identify control, management and data transmissions in each channel, both LTE and NR logically separates control and data channels in regular intervals [47]. Aggregated carriers can be used exclusively for data if using cross-carrier scheduling, where control data is transmitted in

each channel while the data transmissions can be scheduled to another or both [47]. This difference in channel usage depends on the centralized MAC of mobile standards, versus the distributed Medium Access Control (MAC) in wireless networks that share unlicensed channels. Those differences will be explored later in Section 2.2.

2.1.1.2 Frame structure

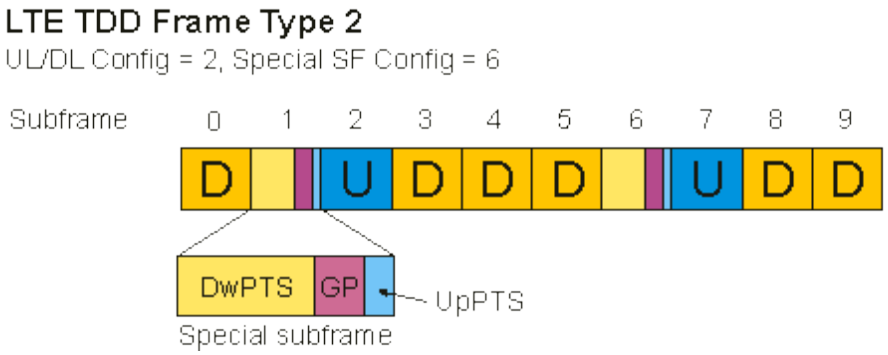
The LTE frame comes in two distinct versions: one version for Time Division Duplex (TDD), where downlink and uplink are alternated in time; and another version for Frequency Division Duplex (FDD), where both downlink and uplink happen simultaneously [4, 47]. The Figures 2.4(a) and 2.4(b) show both the TDD and the FDD versions of an LTE frame. The frames are split into subframes, which serves as the synchronization base for the LTE protocol [4]. The subframes are further split into Resource Blocks (RBs), that are grouped into Resource Block Groups (RBGs) for scheduling purposes [4]. Figure 2.4(c) shows the internals of the FDD frame, including the matrix of 84 Orthogonal Frequency-Division Multiplexing (OFDM) symbols that make up each RB [4].

The number and index of the RBs that will be allocated for each UE is defined by the eNB and transmitted in a control message called Downlink Control Information (DCI) and Uplink Control Information (UCI) for download and upload [55]. When assigning RBs to the users, the eNB starts by excluding RBs reserved for essential control transmissions (allocation map, synchronization signal, localization pilot, and others) and for retransmissions with Hybrid Automated Repeat reQuest (HARQ) [55]. After that, the resource scheduler in the MAC layer of the eNB assigns the RBs based on a given scheduling policy and a QoS Class Identifier (QCI) [55, 56].

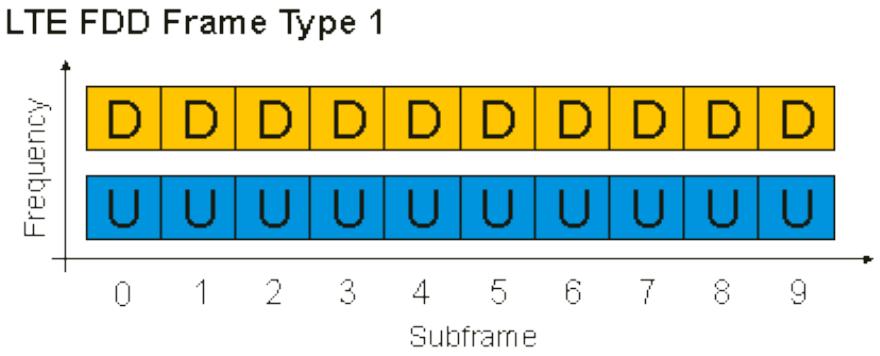
2.1.2 Quality of Service

Quality of Service (QoS) is essential for telecommunication services that are used by several critical applications (weather and public safety alert systems, emergency calls and telesurgery) and other important applications (finance services, autonomous vehicles, etc.) [57]. In LTE, QoS requirements of a GPRS tunnel between the UE and the P-GW, known as a Bearer in 3GPP standards, is given by a QCI number [44, 47]. The General Packet Radio Service (GPRS) tunnel uses the GPRS Tunneling Protocol (GTP) to create a virtual connection between two endpoints. Figure 2.5 shows how the Bearer is built and the QoS is enforced throughout the network, first on the P-GW with Traffic Flow Templates (TFTs) that associate services with QCI levels [44, 5, 47].

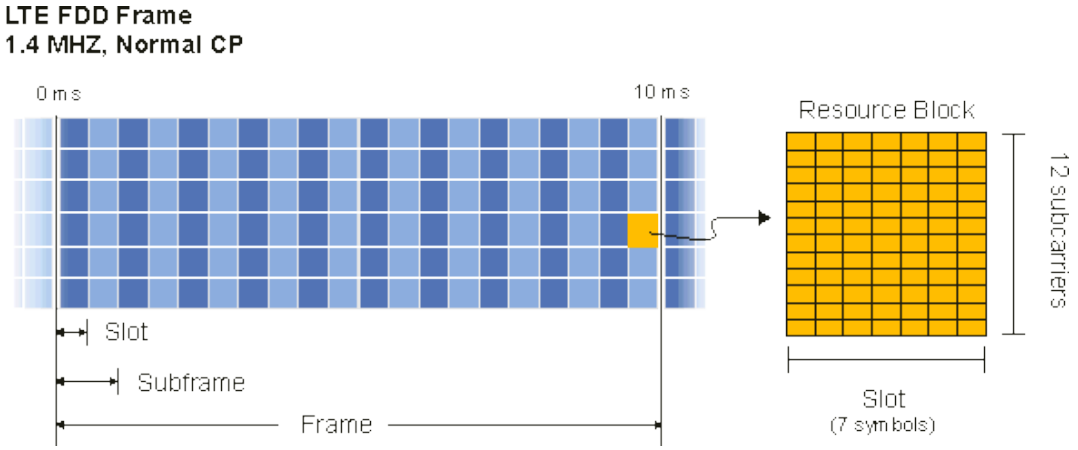
The QCIs are separated into two large groups, one Guaranteed Bit Rate (GBR) and the other non-GBR [44, 6]. The QCIs also specifies the maximum acceptable latency,



(a) LTE TDD frame subdivided into downlink (D), uplink (U) and special reconfiguration subframes



(b) LTE FDD frames for downlink and uplink with their respective subframes



(c) LTE Frame subdivisions

Figure 2.4: LTE frame structure and internals [4]

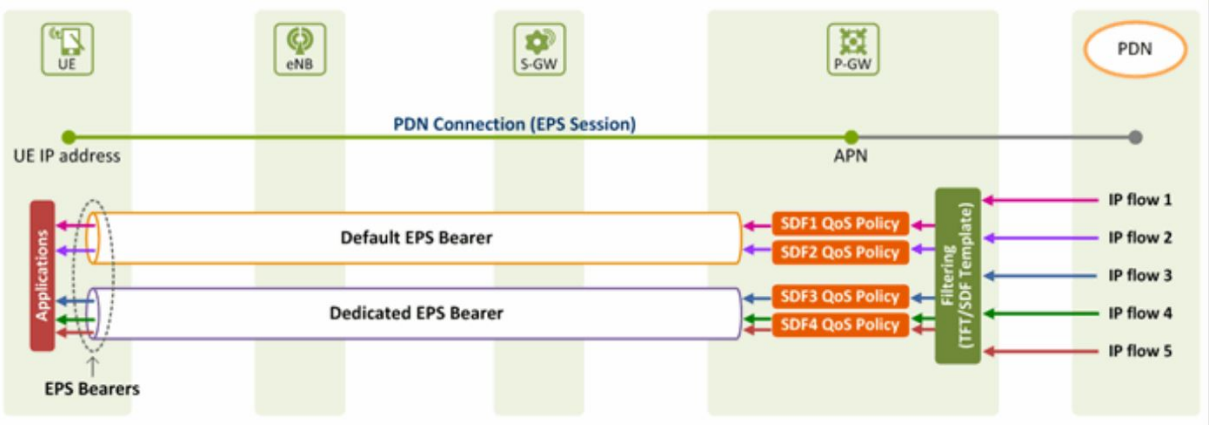


Figure 2.5: LTE Bearer and Service Data Flow (SDF) level QoS architecture [5]

maximum rate of failures and the priority of the bearer in case of congestions [44, 6]. In the 5G NR standard, the CQIs are also separated in three broad groups associated with the user demands: the combination of low latency and high failure rate is known as Ultra-Reliable and Low Latency Communications (URLLC), while the combination of higher latency, throughput and failure rates is known as Enhanced Mobile Broadband / Extreme Mobile Broadband (eMBB), while Massive Machine Type Communications (MMTC) Quality of Service (QoS) levels depend on each application [57, 53]. The different QCI values, related QoS metrics and services are shown in Figure 2.6.

2.2 Medium Access Control

The Medium Access Control (MAC) layer and protocols is responsible for controlling the access of different nodes to a given channel. MAC protocols are divided into 3 major groups: the contention-less protocols, which include polling and token passing techniques; the contention or random access protocols, which include the Aloha and CSMA techniques; the channelization protocols, which include Frequency-Division Multiple Access (FDMA), TDMA, CDMA and other techniques found in most mobile network standards [58]. Contention protocols are distributed, with every node controlling its own access to the channel, commonly used in shared and unlicensed channels [58]. Channelization protocols are centralized, with one or few of the nodes controlling the access of other nodes to the channel. Channelization is commonly used in licensed channels, where the licensee has exclusive access to the channel. In the following subsections we explore the distributed and centralized MAC protocols [58].

QCI	Resource Type	Priority Level	Packet Delay Budget (NOTE 13)	Packet Error Loss Rate (NOTE 2)	Example Services
1 (NOTE 3)	GBR	2	100 ms (NOTE 1, NOTE 11)	10^{-2}	Conversational Voice
2 (NOTE 3)		4	150 ms (NOTE 1, NOTE 11)	10^{-3}	Conversational Video (Live Streaming)
3 (NOTE 3, NOTE 14)		3	50 ms (NOTE 1, NOTE 11)	10^{-3}	Real Time Gaming, V2X messages Electricity distribution - medium voltage (e.g. TS 22.261 [51] clause 7.2.2) Process automation - monitoring (e.g. TS 22.261 [51] clause 7.2.2)
4 (NOTE 3)		5	300 ms (NOTE 1, NOTE 11)	10^{-6}	Non-Conversational Video (Buffered Streaming)
65 (NOTE 3, NOTE 9, NOTE 12)		0.7	75 ms (NOTE 7, NOTE 8)	10^{-2}	Mission Critical user plane Push To Talk voice (e.g., MCPTT)
66 (NOTE 3, NOTE 12)		2	100 ms (NOTE 1, NOTE 10)	10^{-2}	Non-Mission-Critical user plane Push To Talk voice
67 (NOTE 3, NOTE 12)		1.5	100 ms (NOTE 1, NOTE 10)	10^{-3}	Mission Critical Video user plane
75 (NOTE 14)		2.5	50 ms (NOTE 1)	10^{-2}	V2X messages
5 (NOTE 3)	Non-GBR	1	100 ms (NOTE 1, NOTE 10)	10^{-6}	IMS Signalling
6 (NOTE 4)		6	300 ms (NOTE 1, NOTE 10)	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7 (NOTE 3)		7	100 ms (NOTE 1, NOTE 10)	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8 (NOTE 5)		8	300 ms (NOTE 1)	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
9 (NOTE 6)		9			
69 (NOTE 3, NOTE 9, NOTE 12)		0.5	60 ms (NOTE 7, NOTE 8)	10^{-6}	Mission Critical delay sensitive signalling (e.g., MC-PTT signalling, MC Video signalling)
70 (NOTE 4, NOTE 12)		5.5	200 ms (NOTE 7, NOTE 10)	10^{-6}	Mission Critical Data (e.g. example services are the same as QCI 6/8/9)
79 (NOTE 14)		6.5	50 ms (NOTE 1, NOTE 10)	10^{-2}	V2X messages
80 (NOTE 3)		6.8	10 ms (NOTE 10, NOTE 15)	10^{-6}	Low latency eMBB applications (TCP/UDP-based); Augmented Reality

Figure 2.6: QoS Class Identifiers (QCIs) and associated characteristics [6]

2.2.1 Centralized MAC protocols

The centralized MAC protocols rely on a node that centralizes the control of the channel in each region, assigning transmission slots to other nodes based on a scheduling policy. The transmission slots can be either separated in multiple ways in respect to frequency, time and bandwidth [58].

Mobile networks rely mostly in licensed spectrum, which guarantees exclusive access to the channel. Also, the centralized MAC offers a multitude of advantages when compared to the distributed MAC, such as higher channel utilization due to the lack of guard periods, collisions and back-off times, all of them used in distributed MAC protocols.

The assignment of radio resources to the nodes is commonly assigned with some form of resource matrix. In the 4G LTE protocol, the resource allocation is transmitted in a control message known as DCI for the downlink and UCI for the uplink. In the following subsection we detail the resource allocation message in the 4G LTE standard.

2.2.1.1 Downlink and Uplink Control Information (DCI and UCI) of 4G LTE

The DCI and DCI are control messages broadcasted by the eNBs to the UEs every subframe (1ms) [44]. These messages contain a list of Radio Network Temporary Identifiers (RNTIs), which identifies a given associated UE [44]. For each entry of UE-RNTI values, there is either a bitmap containing allocated RBGs to that UE or a tuple containing an offset to the first allocated RBG plus a few subsequent RBGs [55].

There are multiple DCI versions, each of them with different formats representing the same basic information [55], which includes: the format type, resource block assignment, Modulation and Coding Scheme (MCS) index, requests for asynchronous channel quality testing and other. One example of the DCI formats is shown in Figure 2.7, listing the number of bits used for each field of the DCI Format 8 from Release 8, that introduced the LTE [7, 55].

2.2.2 Distributed MAC protocols

The distributed MAC protocols rely on the cooperation of the nodes in each channel, where each node tries to avoid interfering with others and monopolizing the channel. Due to its nature, this protocol is used in shared and unlicensed channels, where there is no licensee/primary user (PU) to protect from interference.

The main distributed MAC protocols are based on the Listen-Before-Talk (LBT) principle allowing for coexistence of different technologies on the same channel [59, 60, 61]. Each node starts by listening to the channel for a given amount of time. If no transmission is detected in each period, the node proceeds to transmission. If a transmission or a

Format 0 (Release 8) - C-RNTI, SPS C-RNTI	
Field Name	Length
Flag for format0/format1A differentiation	1
Hopping flag	1
N_ULhop	1 (1.4 Mhz) 1 (3 Mhz) 1 (5 Mhz) 2 (10 Mhz) 2 (15 Mhz) 2 (20 Mhz)
Resource block assignment	5 (1.4 Mhz) 7 (3 Mhz) 7 (5 Mhz) 11 (10 Mhz) 12 (15 Mhz) 13 (20 Mhz)
MCS and RV	5
NDI (New Data Indicator)	1
TPC for PUSCH	2
Cyclic shift for DM RS	3
UL index (TDD only)	2
Downlink Assignment Index (DAI)	2
CSI request (1 or 2 bit)	1 or 2

Figure 2.7: DCI Format 0 Release 8 [7]

collision is detected, the transmitting node stops all transmissions and waits for a random back-off time. Then the cycle restarts. Congested channels are bound to cause collisions as back-off times end up aligning themselves.

The most used protocol in the category is the Carrier-Sense Multiple Access (CSMA) [62, 63] and its derivatives, which are specializations of the LBT. The CSMA listens to the central frequency of the channel, which is dedicated to signaling patterns, instead of the entire channel bandwidth.

2.2.3 Spectrum Sensing

The Spectrum Sensing (SS) is a set of techniques used to assess whether a channel is occupied or not. Some of the techniques can also detect which modulations and protocols are being used in case the channel is occupied. The sensing can be performed either individually or in collaboration with other nodes, which are detailed in the following subsections.

2.2.3.1 Individual Spectrum Sensing

The individual spectrum sensing is performed by the physical (PHY) and medium access control (MAC) layers of the network interface of a node. The sensing is based on the sampling of the perceived spectrum of a given channel and some form of post processing. The techniques are usually divided into two sets, based on the sensed bandwidth: nar-

rowband sensing is usually associated with Carrier-Sensing (CS), while wideband sensing is usually associated with Energy Detection (ED).

2.2.3.1.1 Carrier-Sensing Carrier-Sense (CS) is a set of one of the most basic techniques of SS, based on comparing the power levels in the vicinity of the central frequency (carrier) of a given channel against the noise-floor[64]. The CS is indicated for narrowband signals [64]. In the IEEE 802.11 and 802.15 standards, the CS variant called Preamble Detection (PD) is used [64]. The PD searches for not only a power peak, but also a transmission pattern to identify the preamble segment of frames [64]. Ubiquitous, CS is used by both wired (e.g. Ethernet) and wireless protocols (e.g. Wi-Fi).

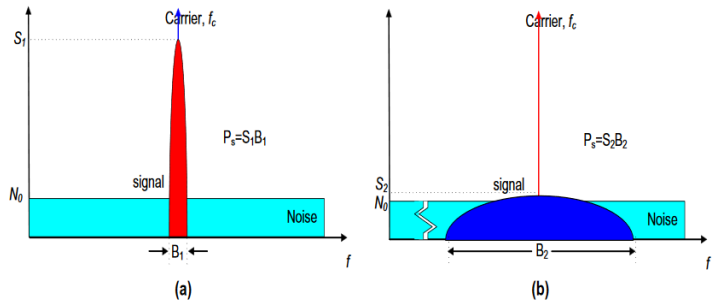
2.2.3.1.2 Energy Detection The Energy Detection (ED) is another set of techniques of SS, based on the detection of a wideband transmission by comparing spectral density to noise floor [64]. As the transmission power is distributed over a wider band, the power peaks are much closer to the noise floor, reducing the probability of detection [64]. Figure 2.8(a) shows the spectral density of both narrowband and wideband transmissions. Figure 2.8(b) compares the performance between the PD Carrier-Sensing (CS) technique applied to the narrowband scenario versus the ED applied to the wideband scenario.

2.2.3.1.3 Window-Based Energy Detector (WIBA) The Window-Based Energy Detector (WIBA), an Energy Detection-based technique proposed in [65, 66], uses a window-based technique to sample the sensed spectrum over time. The wideband spectrum is split into overlapping segments [65], as shown in Figure 2.9(a). The total energy $E_l = \sum_i Y_i(l)$ of each segment is computed [65]. The power threshold $T_h = \gamma \frac{1}{L} \sum_{l=1}^{L} E_l$ is computed using the segments power spectral density [65]. The computed threshold T_h is compared against each segment total energy E_l, and if $E_l > T_h$, the segment is said to be occupied by a transmission [65].

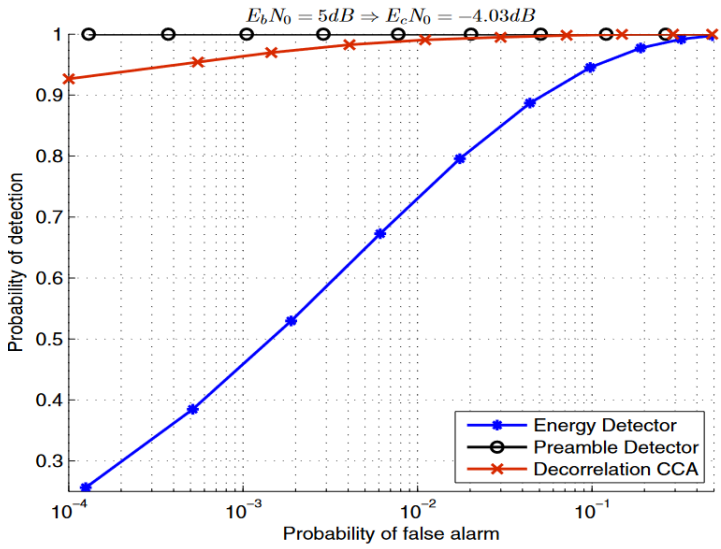
Figure 2.9(b) shows the probability of detection curve of the WIBA method based on the SNR levels of a given scenario. The WIBA performance is proven to outperform other simple energy detection techniques, while being able to detect transmissions even when the power spectral density is lower than the noise-floor (SNR < 0) [66].

After performing the SS procedures of one of the previous techniques, the node that performed the sensing proceeds to the decision making process. The decision making determines if the SS results indicate whether a transmission was detected or not. This decision is usually done based on detection thresholds, but other techniques such as machine learning can also be used using the SS result as one of the inputs.

The individual SS provides satisfactory information to prevent most collisions in uncrowded channels for short range transmissions in unlicensed channels. However, the

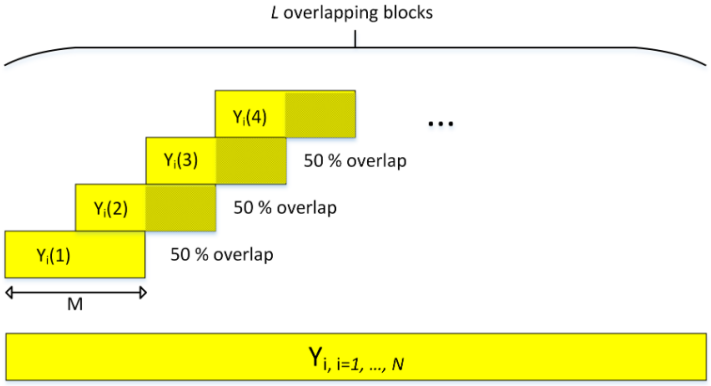


(a) Spectral densities in narrowband and wideband transmissions [64]

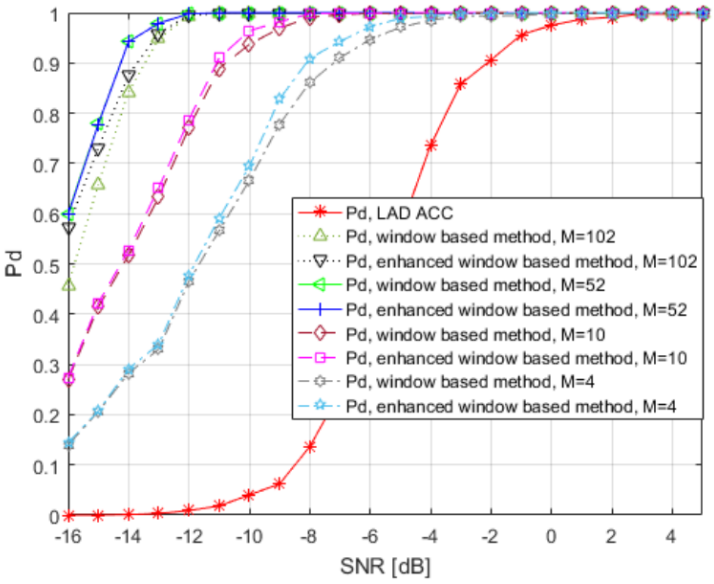


(b) Probability of detection comparison between Preamble Detection versus Energy Detection [64]

Figure 2.8: Spectral densities of narrow and wideband transmissions and probability of detection for Carrier-Sense and Energy Detection techniques



(a) Overlapped segmentation of the channel bandwidth [65]



(b) Signal-to-Noise-Ratio (SNR) versus Probability of detection [65]

Figure 2.9: Segmentation of the band into overlapping blocks for the WIBA and SNR-based probability of detection curves

limited information of the channel state is not enough to ensure the protection of licensees (Primary Users (PUs)) during opportunistic use of their channels, especially in long-distance scenarios (e.g. rural and arctic wide-range access networks).

2.2.3.2 Collaborative Spectrum Sensing

Collaborative Spectrum Sensing (CSS) schemes are used to share the individual SS results with different nodes, increasing the amount of information available for the decision making process [67]. The CSS can be described as a series of steps that are executed in cycles, that can be triggered periodically or on-demand. A generic CSS cycle has the following steps:

1. the sensing nodes execute the individual SS technique
2. individual SS result sharing
3. reception of shared individual SS results
4. fusion of individual SS results using a Fusion Technique (FT)
5. decision making of whether the channel is occupied or not

The CSS can be either distributed or centralized, based on how the sensing information is shared and how the decision making is made [67].

2.2.3.2.1 Distributed Collaborative Sensing In the Distributed CSS, each node is responsible for the collection of shared sensing reports, Step 3, fusion with their own sensing results, Step 4, making the decision about the channel state, Step 5 and which action to take upon the decision result [8, 31].

Multiple distributed CSS schemes have been proposed, most of them focusing on the already distributed-by-design Wi-Fi networks [8, 31]. The SS result sharing, Step 2, is made via broadcast to nearby nodes, shown in Figure 2.10. Nodes can also choose whether to forward the results from other nodes or not.

The Distributed CSS shares pros and cons with distributed networks: they are more resilient, but less efficient. In the context of collaborative sensing, a distributed scheme: requires the transmission of more messages, nodes have smaller range, there is a bigger delay between the detection of a transmission by one node and the reception of that report by a node far from the sensor that originally detected it, distributed decision making can result in conflicts that result in interference between either sensing nodes or the channel licensee, in case of a licensed channel. On the resiliency side, nodes in distributed schemes can identify and prioritize trustworthy sensing reports and ignore nodes that are known to report wrong sensing results.

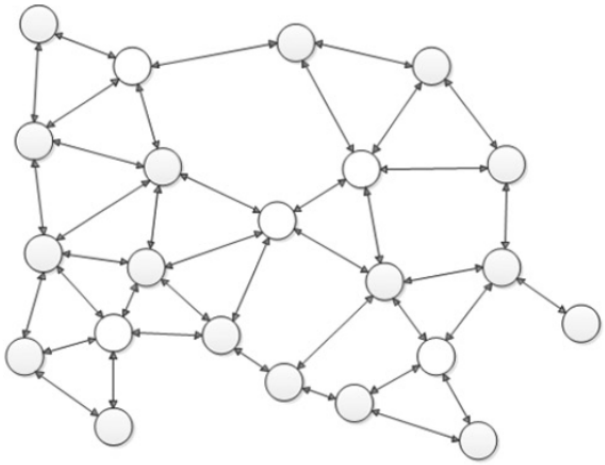


Figure 2.10: Distributed collaborative sensing [8]

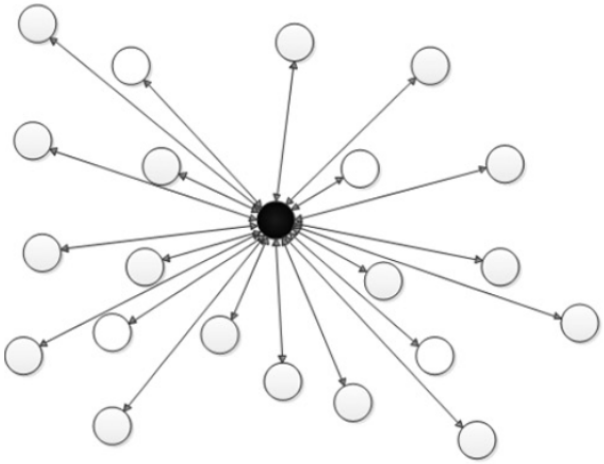


Figure 2.11: Centralized collaborative sensing [8]

2.2.3.2.2 Centralized Collaborative Sensing In the Centralized CSS, the sensing nodes only transmit their sensing results to the central node during Step 2. The sensing results are collected by the central node during Step 3, followed by the fusion, Step 4, and the decision making, Step 5. After the decision making, the fused results are either shared directly via broadcast or indirectly via the resource allocation matrix.

Figure 2.11 depicts the transmission of the individual SS results to the central node. The processing may include false reports filtering, weighting of sample relevance [8, 31]. When the individual sensing results are processed, they proceed to the same fusion technique used in distributed CSS, and the result is shared across the nodes [8, 31].

One example of how a Centralized CSS cycle operates is shown in Figure 2.12. Figure 2.12(a) show the individual spectrum sensing executed by the sensing nodes (UEs), Step 1. The sensing is followed by the report of their sensing results to the central node (eNB), Step 2, as a bit per channel (Hard-combining). In Figure 2.12(b), the central node receives

the reports, Step 3. The collected sensing reports are then fused, Step 4, and the central node decides if the channel is occupied or not, Step 5, as shown in Figure 2.12(c).

Centralized CSS is indicated for use with centralized MAC protocols, where there is already an implicit source of synchronization, where control messages are exchanged between the access point (Access Point (AP) for Wi-Fi, eNB for LTE) and stations (Stations (STAs) for Wi-Fi, UEs for LTE). A few additional fields to those messages can carry which station should perform and report the individual sensing, while the results of the sensing can also be appended to already existing control messages, reducing the number of transmissions and collisions when compared to a distributed MAC implementation.

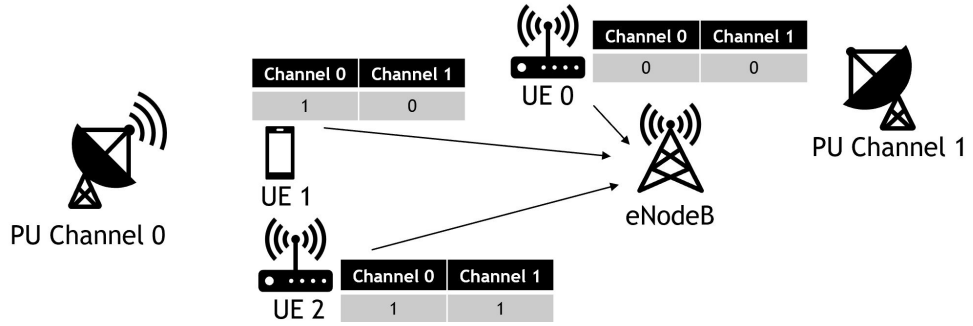
A centralized approach also comes with its flaws: single point of failure; latency between the sensing and the reception of the CSS results transmitted by the AP; lost of access opportunity due to the latency between sensing and result.

2.2.3.2.3 Spectrum Sensing Reporting The transmission of individual sensing results may require different amounts of data depending on the report type. Hard combining fusion techniques use a single bit to represent a channel state (occupied or not), while Soft-combining fusion techniques use n bits to represent different probabilities of occupation of the channel [68, 69, 70].

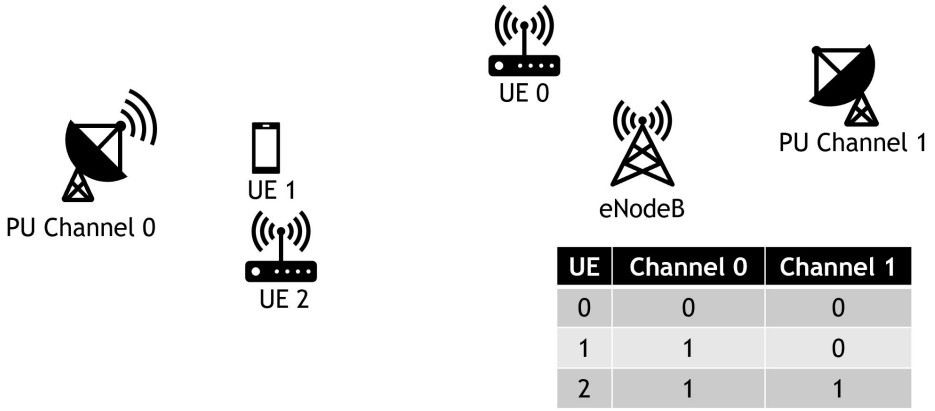
2.2.3.3 Fusion Techniques

The FTs are fundamental for collaborative sensing schemes, as they are used to merge the individual sensing results from multiple nodes into a single final sensing result [67], as shown in Figure 2.12(c). The final sensing result is computed by each node or a few nodes in Decentralized collaborative sensing schemes, while a single central node computes the final result in Centralized collaborative schemes. FTs can be based on multiple principles: statistical, with counting/voting and temporal analysis; pattern-matching and clustering through machine learning [67].

As the FT merges individual sensing results, it directly affects the outcome of decisions taken based on the fused result. The different FTs trade false positives (e.g. detection a transmission when no node is transmitting) and false negatives (e.g. not detecting a transmission when a node is transmitting) in different ways and should be selected based on the use case [67]. In the CSS for opportunistic access of licensed spectrum, mitigating false negatives is more important than mitigating false positives, as one causes interference to the channel licensee while the other wastes unused spectrum [67]. Other metrics also impacted by the FTs are: the processing power, important for battery-powered mobile devices; the delay between the sensing and the broadcasting of the fusion result, also for channel licensee protection; reporting overhead, which depends on the number of reported

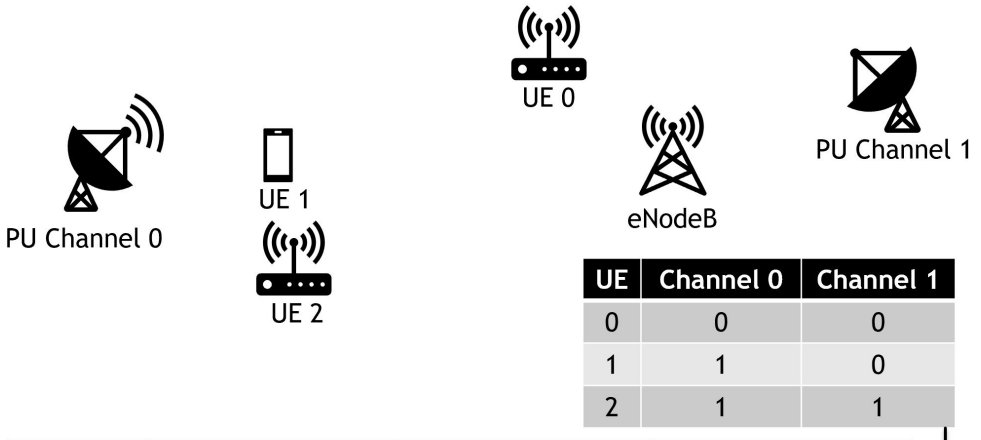


(a) Individual spectrum sensing and reporting of sensing nodes, Steps 1 and 2 of the CSS cycle



UE	Channel 0	Channel 1
0	0	0
1	1	0
2	1	1

(b) Report collection by the central node, Step 3 of the CSS cycle



UE	Channel 0	Channel 1
0	0	0
1	1	0
2	1	1

Fusion Technique	OR		AND		2-out-of-3	
	Channel 0	Channel 1	Channel 0	Channel 1	Channel 0	Channel 1
Fusion result	1	1	0	0	1	0
Ground truth	1	0	1	0	1	0
Side effect	-	Unused spectrum	Collision with the licensee	-	-	-

(c) Fusion of individual sensing results by the central node and channel state decision, Steps 4 and 5 of the CSS cycle

Figure 2.12: Example of partial Collaborative Spectrum Sensing Cycle

channels, number of reporting nodes and the number of bits to represent the state of the channels, which depends on both the chosen fusion and the SS technique.

Multiple FTs have been proposed over the years and a few of them are described next.

2.2.3.3.1 k-out-of-n rule The k-out-of-n rule for fusion is based on the probability of an event happening k times out of n tests [67, 32]. That probability is given by a binomial distribution $\binom{n}{k} p_{di}^k (1 - p_{di})^{n-k}$. The probability of an event happening at least k times out of n tests is given by them sum of binomial distributions with $l = k..n$ events with same outcome out of n samples, as in Eq. 2.1.

$$P_d(d) = \sum_{l=k}^{n} \binom{n}{l} p_{di}(d)^l (1 - p_{di}(d))^{n-l} \tag{2.1}$$

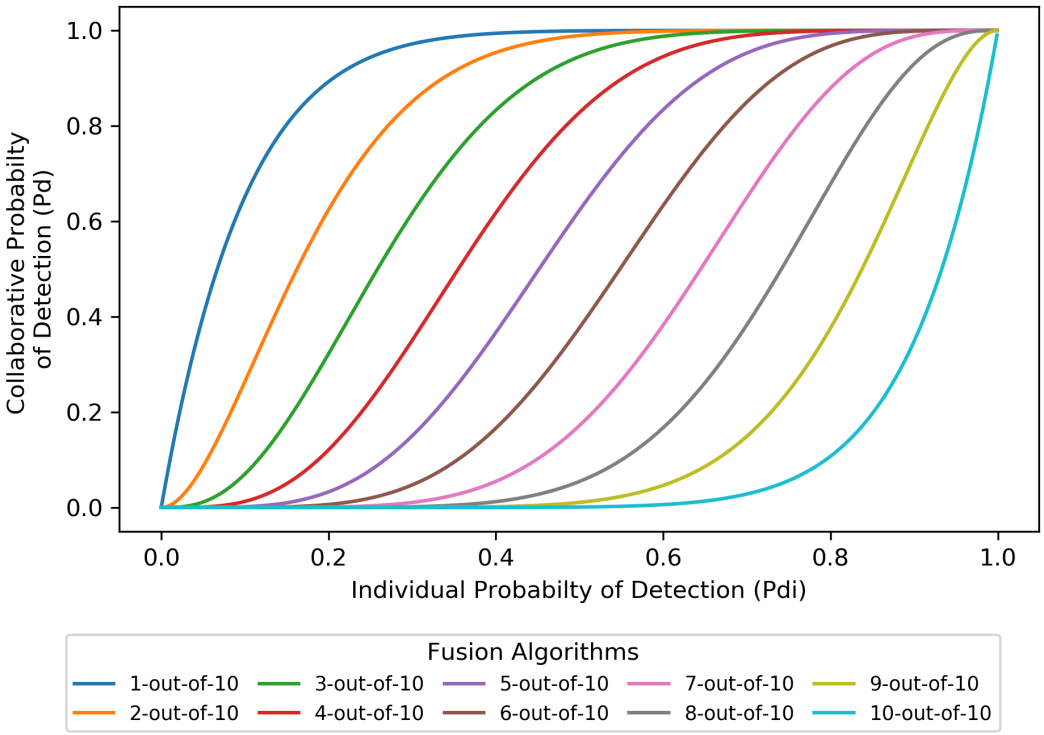


Figure 2.13: Collaborative probability of detection using the k-out-of-n fusion rule for different values of k

As an example of the behavior of the k-out-of-n rule fusion, the Figure 2.13 presents the fused probability $P_d(d)$ of at least k events of detecting an occurring transmission. The number of sensing nodes n is set to $n = 10$, the threshold of nodes reporting a detection k ranges from $k = 1..n$ and the probability of detection of each spectrum sensing event for the sensing nodes $p_{di}(d)$ ranges of $p = 0..1$. The 1-out-of-n rule is also known as the logic OR rule, while n-out-of-n rule is known as logic AND rule.

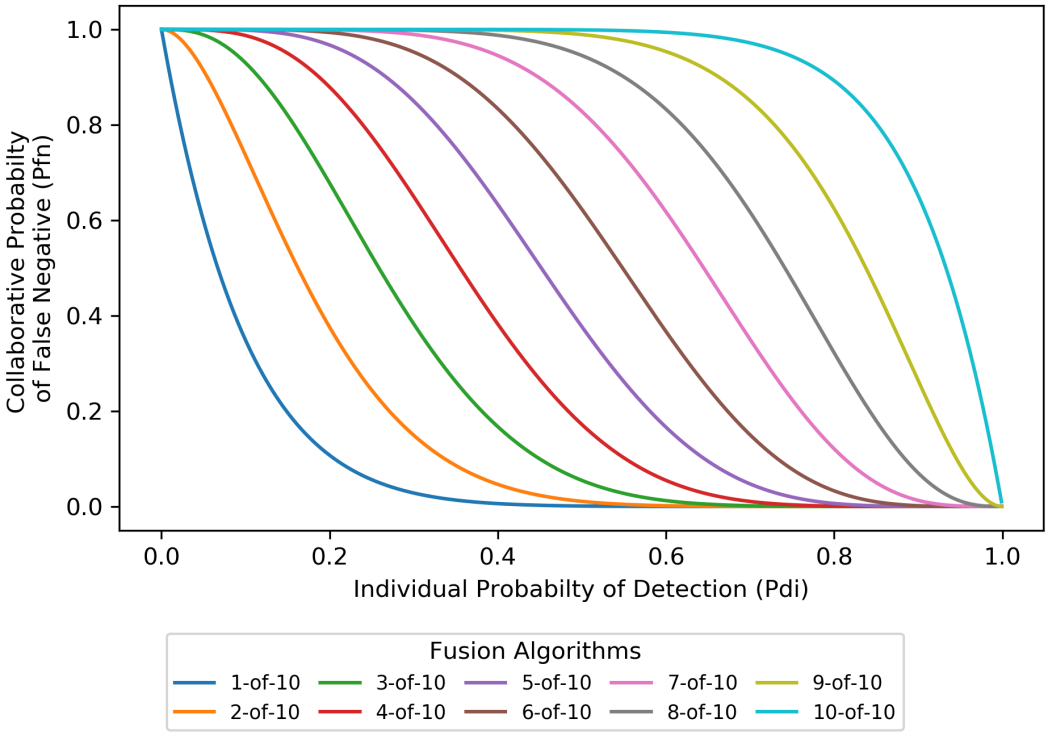


Figure 2.14: Collaborative probability of false negative using the k-out-of-n fusion rule for different values of k

It can be implied from Figure 2.13 that an increasing individual probability of detection p_{di} increases the collaborative probability of detection P_d. At the same time, a small threshold k leads to higher collaborative probability of detection P_d and lower probability of false negatives P_{fn} (not detecting a transmission), as shown in Figure 2.14.

The trade-off seen in the number of minimum events k however affects the collaborative probability of detection P_d when there is no transmission to be detected. The individual probability of detection of the spectrum sensing technique p_{di} in the equation is replaced by the probability of false positives p_{fp} (when a transmission is detected when its not happening) of the same spectrum sensing technique. When this happens, the decrease in the threshold k results in an increase of the probability of false alarms p_{fp}, as shown in Figure 2.15. The increase in false alarms reduces the chances to opportunistically access the licensed channel.

The simplicity of the implementation and modeling make it one of the most used FT schemes, even though it is not the most accurate.

2.2.3.3.2 Majority rule The majority rule says that the fusion result P_d is defined by most reports, as shown in Eq. 2.2 [32]. In this equation, the probability of detection $p_{di}(d)$ is based on the distance d_k to each of the N sensing nodes (UEs). It is a special

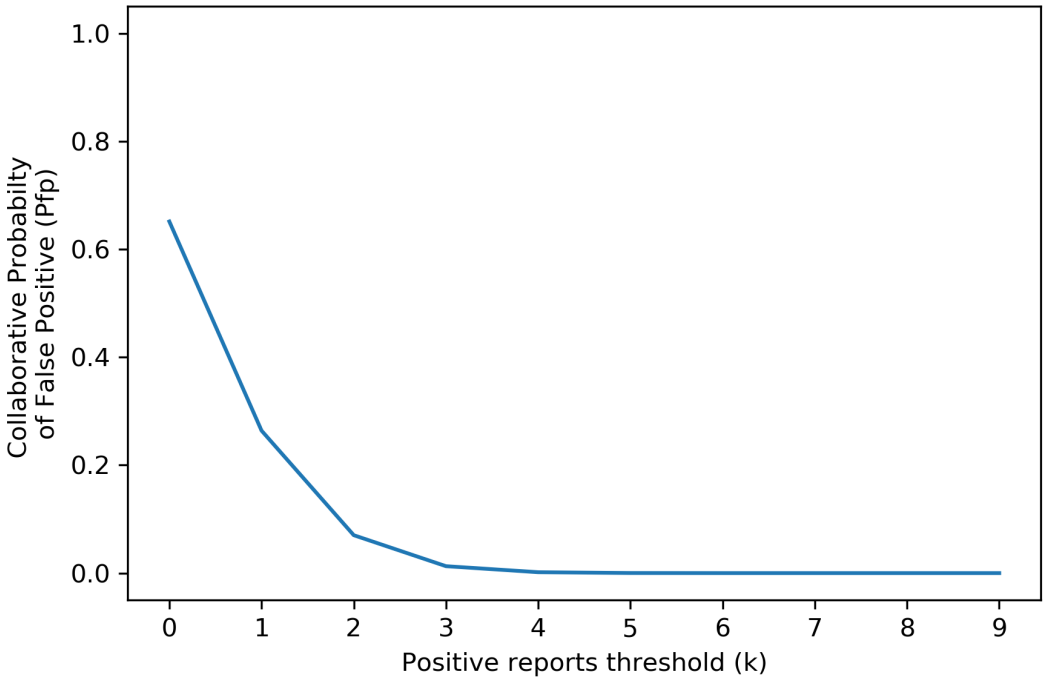


Figure 2.15: Collaborative probability of false positives using the k-out-of-n fusion rule for different values of k, assuming a fixed individual probability of false positive of 0.1

case of the k-out-of-n rule.

$$P_d = \begin{cases} 1, for \sum_{k=1}^{N} p_{di}(d_k) > \frac{N}{2} \\ 0, otherwise \end{cases} \tag{2.2}$$

2.2.3.3.3 Machine learning Multiple fusion techniques have been proposed over the years using different machine learning techniques such as Neural Networks (NNs), Support-Vector Machines (SVMs) or K-Nearest Neighbors (KNN) [67].

In [33, 9], a shallow and narrow feed-forward neural network is trained with both individual sensing report (local decision in Figure 2.16) and the reported Signal to Noise Ratio (SNR) of every reporting node as the input and the real channel state as the collaborative fusion result target (output) in a supervised learning approach. Figure 2.16 shows the topology of the NN proposed in [9].

The technique is simple, but scales very poorly for a big range of sensing nodes, as they would require the usage of multiple NNs with different lengths, or a different topology. It can be effective if the network behaves as expected in the simulations that were used to train the NN, but may not be as effective in the real world as sudden temperature fluctuations, multi-path effects due to blocking, and topography may impact the measurements by the UEs and reflect on the fusion result.

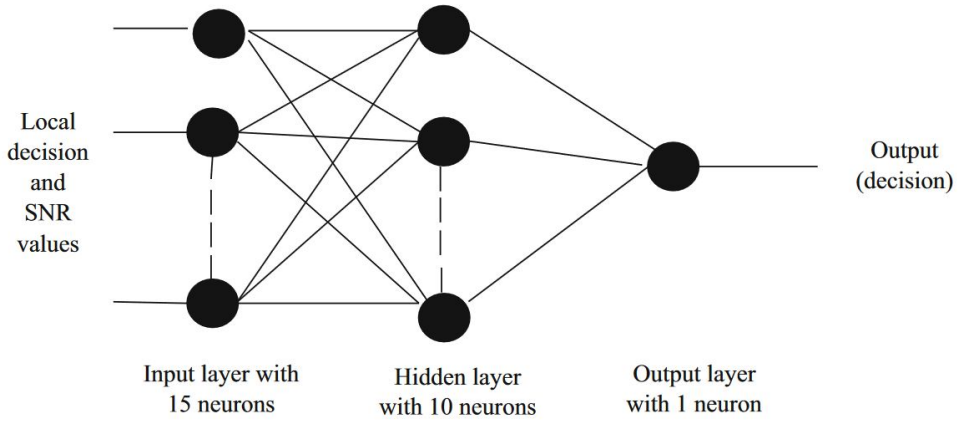


Figure 2.16: Artificial neural network topology proposed in [9] for the fusion of individual sensing results in a Centralized CSS scheme

In [10], two fusion techniques are proposed, one using SVM and the other using KNN. The SVM technique is used to improve the ED spectrum sensing technique results by analyzing them over time, while the KNN classifier is used for distributed voting between nearby nodes. The technique has higher accuracy than compared techniques in the scenarios simulated by the authors, as shown in Figure 2.17. The simulated scenarios are inspired by urban networks, with high density of both PUs and Scheduling Units (SUs) distributed within a small cell, with high mobility. The computational costs are ignored.

More recently, other machine learning techniques started being used in new FTs proposals, such as in [34] and Extreme Learning Machines (ELMs) in [35]. As with other techniques, they are used for clustering the sensing results and deciding based on the clustering. The ELM used in [35] follows a similar strategy proposed as the one proposed in [34] but using a different NN topology and achieving similar results.

The CNN proposed in [34], shown in Figure 2.18(a), aggregates sensing results reported by the sensing nodes per channel subcarrier, forming a binary matrix that is fed into the CNN. The CNN then processes the convolutions, producing the results that feed the $FC1$ layer. This layer feeds the next and a softmax function pass, that normalizes the results r into weights $wr_i \in [0,1] / \sum_{i=1}^{N} r_i = 1$. The results of the softmax layer are then used as weights to define if the channel is idle with Eq 2.3. The technique has better detection accuracy than other tested techniques in the scenario used by the authors, as shown in Figure 2.18(b).

$$\frac{e^{wr_0 \times r}}{\sum_{i=0}^{1} wr_i \times r} > \frac{e^{wr_1 \times r}}{\sum_{i=0}^{1} wr_i \times r} \tag{2.3}$$

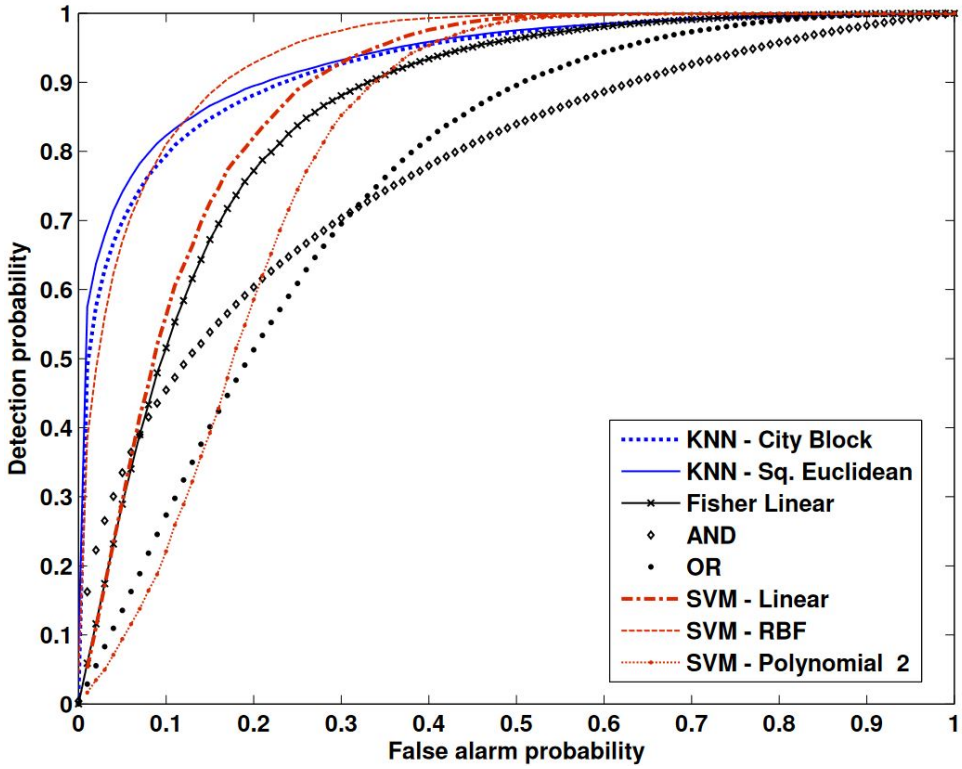
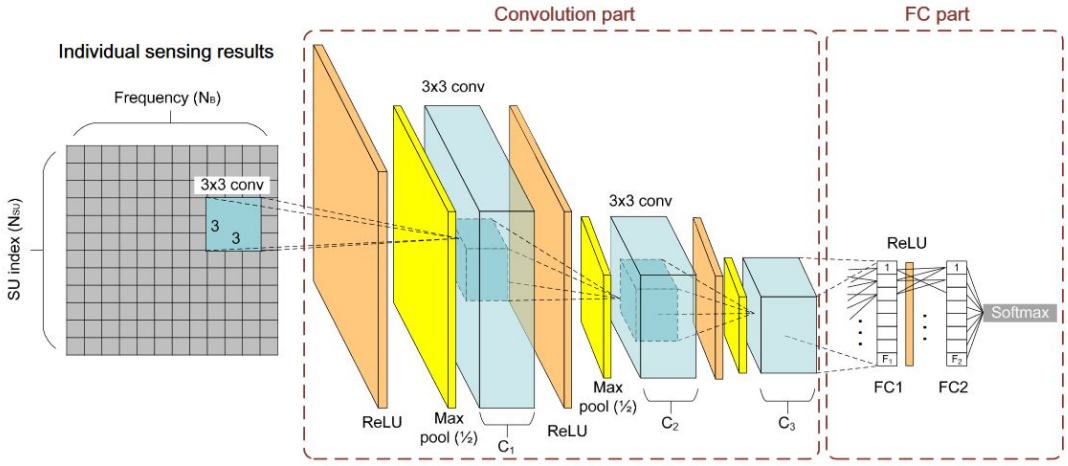
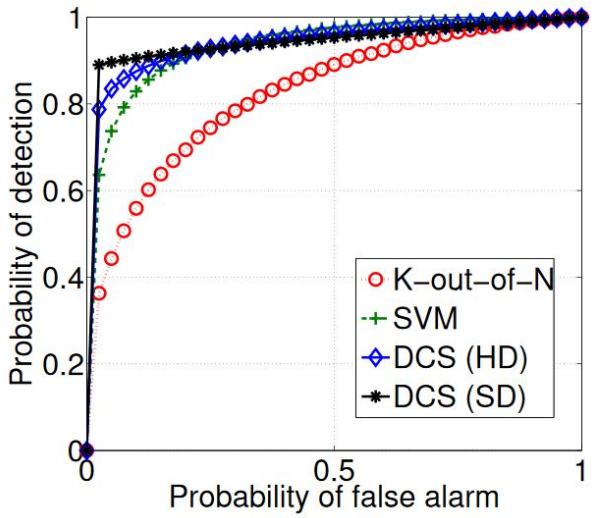


Figure 2.17: Comparison between different fusion techniques with the ones proposed in [10]. Scenarios assume PUs are divided into senders and receivers uniformly spaced with 100m between each other, detection radius of sensing nodes is 300 m, 49 sensing nodes are organized in a 7x7 grid, 1000 sensing samples are used for the classifier training.



(a) Convolutional Neural Network (CNN) model for deep collaborative sensing [34]



(b) Performance comparison of CNN and other FTs [34]. Assuming a scenario with 32 sensing nodes and a single PU in an area of 200 m x 200 m, with 200 samples per node and a noise floor of $-164dBm/Hz$

Figure 2.18: Convolutional Neural Network topology for collaborative sensing fusion and performance comparison

2.2.3.3.4 Additional fusion techniques There are multiple FTs implementations based on the techniques previously mentioned which improve on specific aspects, such as: the accuracy of different fusion techniques in [68, 69, 70]; reducing power consumption of the reporting devices [71, 8]; and reducing the traffic overhead of the reporting in [72, 73, 74].

2.2.3.4 Mitigation of Byzantine attacks

Byzantine generals problem was described in [75] as a problem of consistency due to a faulty component of a system. A reliable system requires that these inconsistencies should be mitigated to prevent erratic behavior [75]. The conflicting information provides a path to identify, isolate and recover the faulty component [75, 76]. However, identifying which information comes from the faulty component is the fundamental issue of the Byzantine generals problem. If every component of the system is trustworthy, then the consensus between components and identification is reasonably simple. If the components are not trustworthy, then it may be impossible to reach the consensus.

The collaborative spectrum sensing schemes, distributed or centralized, can be targeted by byzantine attackers, which can transmit forged sensing reports to disrupt the opportunistic use of the spectrum or cause harmful interference [76, 39]. Byzantine attacks in CSS is also known as Spectrum Sensing Data Falsification (SSDF). A multitude of different Byzantine attacks and defense mechanisms against them have been studied for both distributed and centralized schemes, with probabilistic (realistic) and non-probabilistic (unrealistic) falsification of results [77], but it is impossible to completely protect against all of them. Most of the defense mechanisms are based on assigning trust to the nodes and filtering nodes below a certain threshold.

According to [36, 38], the Byzantine attacks on the CSS context can be classified as: Always True, Always False, Dynamic. The Always True attack happens when a malicious node, or attacker, always report that a PU is transmitting, while the Always False attack happens when a malicious node always report that no PU is transmitting[36, 38]. Dynamic attacks are more sophisticated and harder to identify, where the malicious node change the forging pattern of reports to disguise itself from the central node [38].

2.2.4 Spectrum Allocation

Spectrum Management Agencies (SMAs) are responsible for managing how the spectrum is used and by whom it is used. Spectrum bands are commonly auctioned to a single licensee known as the PU, resulting in the static allocation of the spectrum. The PU has the exclusivity of use of that band in each area by a given time period, usually tens of

years. An example of such spectrum allocation is shown in Figure 2.19, produced by the Brazilian SMA known as Anatel (National Telecommunications Agency). The large green chunks in Figure 2.19 represent bands reserved for analog broadcasters (TV, AM and FM radio).

As the number of applications requiring access to spectrum grows and spectrum remains underused in the countryside, the academia, small Fixed Network Operators (FNOs) (FNO) left out of the mobile business due to the lack of available spectrum, and some SMAs started proposing either dynamic spectrum allocation.

The dynamic spectrum allocation would rely on an on-demand temporary licenses with a short expiration date, to increase the utilization rate of spectrum and prevent Mobile Network Operators (MNOs) from buying spectrum to keep competition away[27]. This temporary license is based on the License Shared Access (LSA). The other proposed option, mixes Dynamic Spectrum Access (DSA) and LSA elements [27], replacing exclusivity licenses with a shared one based on tiered access that guarantees the priority to the PUs, while allowing higher and lower priority to SUs. This option is known as General Authorized Access (GAA) [27].

2.2.5 Opportunistic Access

The lack of free air spectrum prevents MNOs from increasing the total bandwidth available for their users. The concern was partially solved by increasing compatibility of the 5G NR stack with additional spectrum in the FM, the 6GHz and the Millimeter Waves (mmWave) bands [19]. Most of the additional spectrum is in the mmWave bands, which are adequate for dense deployment due to their shorter range [19]. At the same time, most of the spectrum licensees underutilize their spectrum by transmitting only or almost exclusively in bigger cities, where unlicensed spectrum is crowded. In remote areas, most of the spectrum, licensed and unlicensed, remains unused [20, 21, 22].

Proposals have been made to opportunistically access the underused spectrum, making sure not to disturb the PU of the channel when it is transmitting [23]. The proposals usually call for geographical databases which are queried for channel use in each region by prospecting users. Other proposals call not only for a query system, but a full temporary license authorization system, giving exclusivity over a channel for a given amount of time. Others call for supplementary techniques, such as using SS to validate and update the information provided by the database.

In the following subsections, we detail the dynamic spectrum access technique and its enablers.

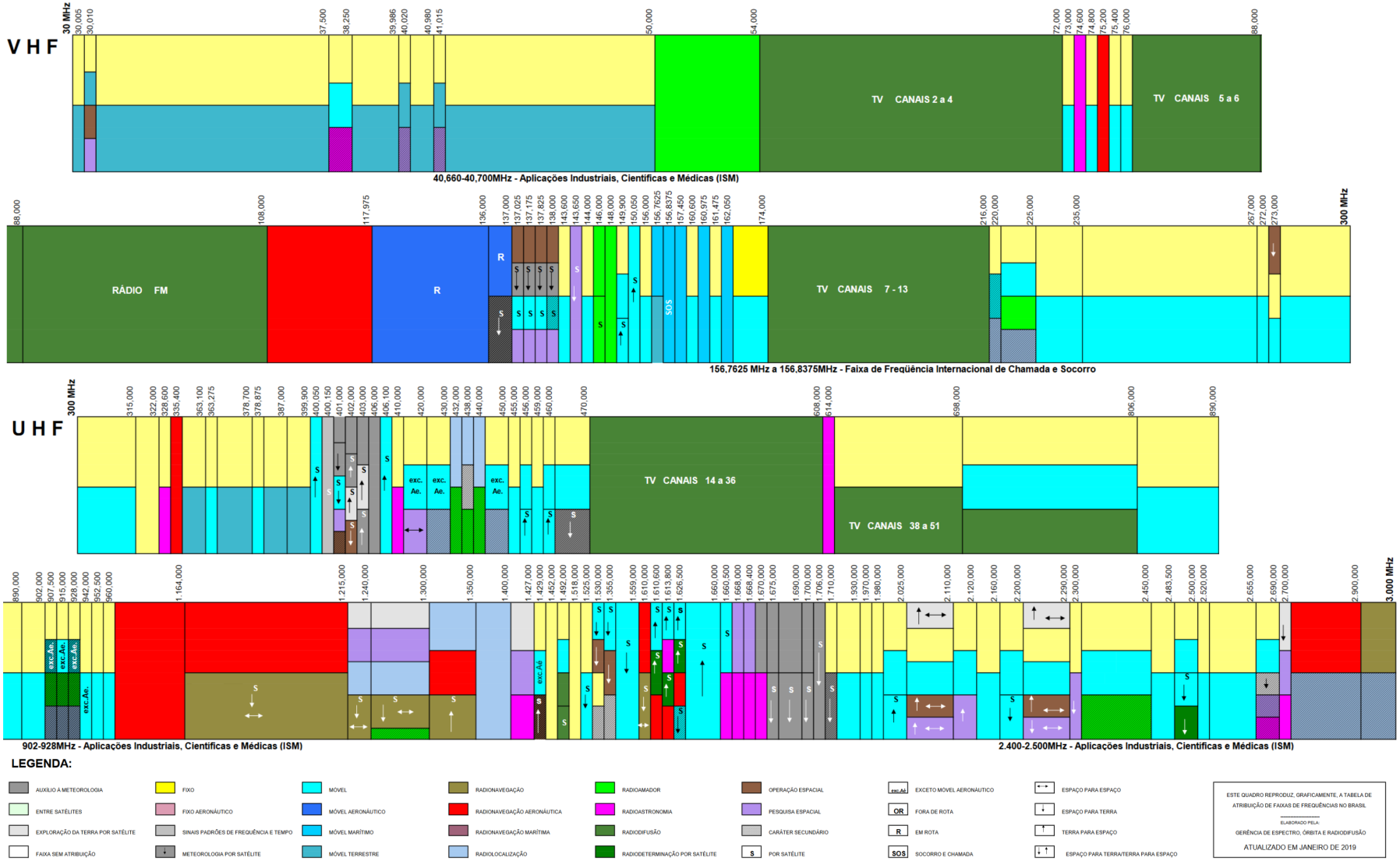


Figure 2.19: Spectrum allocation of VHF and UHF bands in Brazil [11]

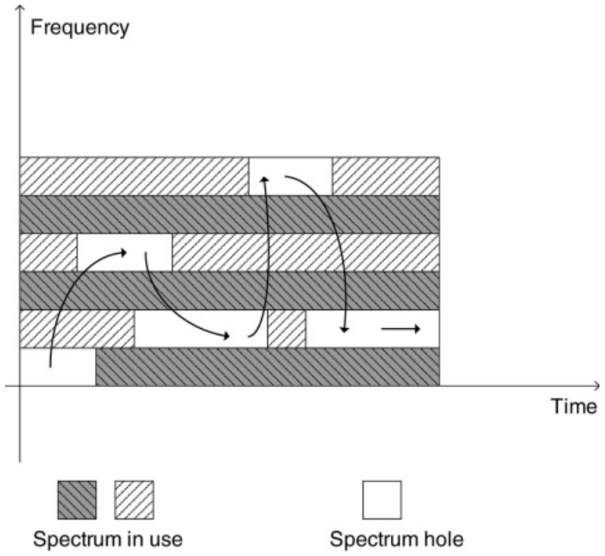


Figure 2.20: Dynamic Spectrum Access of unused spectrum (spectrum holes [12]

2.2.5.1 Dynamic Spectrum Access (DSA)

The DSA is a form of spectrum access based on the opportunistic access to the licensed channels by one or more unlicensed users [23, 24, 25, 26]. The SUs preemptively backs off from transmitting when senses activity from the channel licensee, known as the PU. The SU back off prevents interference to the PU, while increasing spectrum utilization [23, 78, 24]. To prevent disruption to PUs protocols that are not prepared for coexistence, the delay between the start of the PU transmission and the SUs back off should be minimized as much as possible. The correct execution of the DSA technique, should increase the amount of available bandwidth for the SUs [28, 29].

Most od DSA proposals are focused on bands that remain unused after the transition from analog to digital television [79, 80]. The unused bands are called TV whitespaces (TVWS) [81], which researchers have advocated for unlicensed use after the TV transition finishes while MNOs pressure for repurposing them as additional 5G spectrum [82, 83, 84]. Those bands provide wide-coverage and requires less power, ideal of mobile connections [80].

2.2.5.2 Cognitive Radio (CR)

The CR is an enabler technology for DSA [85, 22], being generally defined as a radio that can sense its environment and adapting to it [86]. The Federal Communications Commission (FCC) [78] defines CR as "a radio that can change its transmitter parameters based on interaction with the environment in which it operates. This interaction may

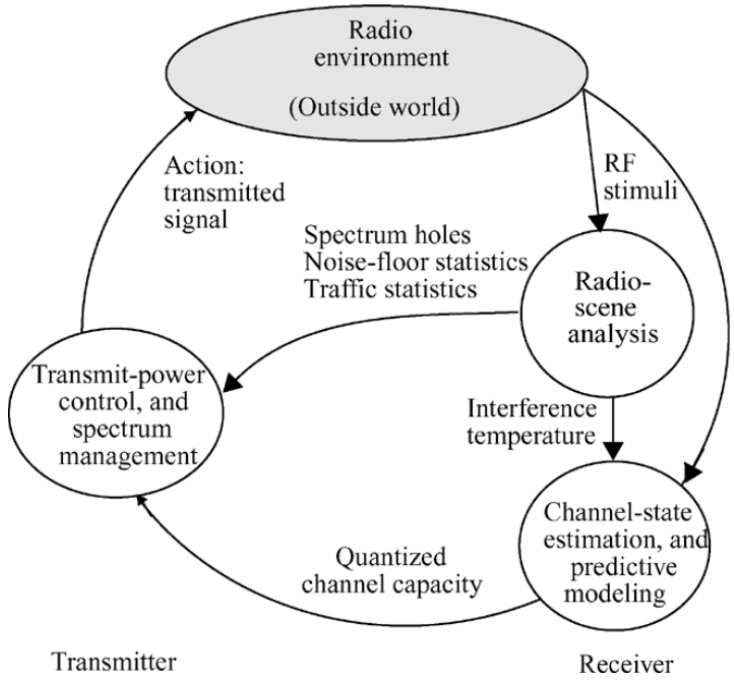


Figure 2.21: Cognitive Radio cycle [13]

involve active negotiation or communications with other spectrum users and/or passive sensing and decision making within the radio".

A CR usually operates in a three step cycle: sensing, decision and judging. During the sensing step, the radio collects channel information, either individually or by cooperative actions[86], shown in Figure 2.21 as the Radio-scene analysis. During the decision step, the radio checks the current parameters and last sensing results and tries to act to increase one of its performance metrics, increasing transmission power, using a different MIMO configuration, reducing errors by choosing a more conservative coding and modulation scheme[86]. The decision step is shown in Figure 2.21 as the transmit-power control and spectrum management state. During the judging step, the previous action taken is compared to the measured outcome and the previous configuration is reverted if the outcomes are worse [86]. This step is represented by the channel-state estimation and predictive modeling in Figure 2.21.

The Dynamic Spectrum Sharing (DSS) enabled by CR and DSA leads to higher spectrum efficiency and utilization [31, 72, 87].

2.2.5.3 Geographical Databases

The geographical databases are an enabler technology proposed for the DSA, reducing costs on specific spectrum sensing hardware required by Cognitive Radio (CR). It contains a set of base stations per region, along with their associated coordinates and transmission

parameters (e.g. antenna type, gain, height, direction and inclination; transmission power; modulation and coding schemes; channels) [88].

and mobile operators already maintain similar databases for licensing, maintenance and infrastructure sharing purposes, making it a viable choice. Some studies propose the possibility of offering temporary licenses for users that want to use unused licensed channels in a region [89]. The users would be able to request a spectrum band using sending a request to the geographical database, as shown in Fig. 2.22(a), and the database could either respond with either a denial for the request or the spectrum response containing the transmission parameters, as shown in Fig. 2.22(b).

2.3 Statistical tools

Some of the most generic tools to model/extract the distribution of random processes are the Markov Chains and Monte-Carlo simulations. We detail them in the following subsections.

2.3.1 Central Limit Theorem (CLT)

There are multiple theorems that state the existence of a Central Limit Theorem (CLT) [90]. Those theorems may vary depending on the context they are used, although the Laplacian definition has been widely adopted as a general case and improved over the years [90]. Poisson improved on the CLT relating it to the Law of Great Numbers, previously known as Bernoulli's Theorem, which states that, for individually and identically distributed samples, the mean value of samples $\bar{X} = \frac{1}{n}\sum_{i=0}^{n} x_i$ converges to the either the real value X or the mean value of the source probability distribution Φ as the number of samples n increase [90].

In modern times, after the contributions of Chebyschev, Markov and Lindberg, the CLT became completely generalized [91]. They have shown that even in cases that the samples are not individually and identically distributed, if the samples correlation $c(k) = \langle x_{n+k} + x_n \rangle$ between samples x_i has a zero mean and can be summable $\sum_{k=0}^{\infty} c(k) < \infty$, then the CLT limit in Equation 2.4 holds [91]. The CLT is given by Equation 2.5, where S_N is the mean value of the samples, N is the number of samples, and the effective standard deviation $\sigma_{eff}^2 = \sigma^2 + 2\sum_{k=0}^{\infty} c(k)$ [91].

$$\lim_{N\to\infty} P(a \leq \frac{S_N}{\sqrt{N\sigma_{eff}^2}} \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx \tag{2.4}$$

$$\langle S_n \rangle \simeq 2\sigma_{eff}^2 N \tag{2.5}$$

Parameter	Content	Data Structure
database query		
Device descriptor	 Manufacturer serial number ruleset ID FCC ID Device type Radio Access Technology (RAT)	 string list string string string
geolocation	Point, Region, Center Latitude, Longitude	
Antenna characteristics	antenna height,antenna type antenna direction, radiation pattern antenna gain, antenna polarization	list
Device owner		vcard
Device capabilities	 authority Max location change Max polling sec Rule set IDs	 string float int list

(a) Geographic database request for spectrum [89]

Parameter	Content	Data Structure
database query		
Time stamp	 start time stop time ID	 string list
Spectrum schedule	 Bandwidth frequency range Bandwidth	
Spectrum report	antenna height antenna type antenna direction radiation pattern antenna gain antenna polarization	list
rule set ID		vcard
Device capabilities	 authority Max location change Max polling sec Rule set IDs	 string float int list
Location	geolocation and selected from the spectrum request	
maxLocationChange	renew the spectrum request	

(b) Geographic database response for a spectrum request [89]

Figure 2.22: Geographic database request and response for temporary spectrum band licenses

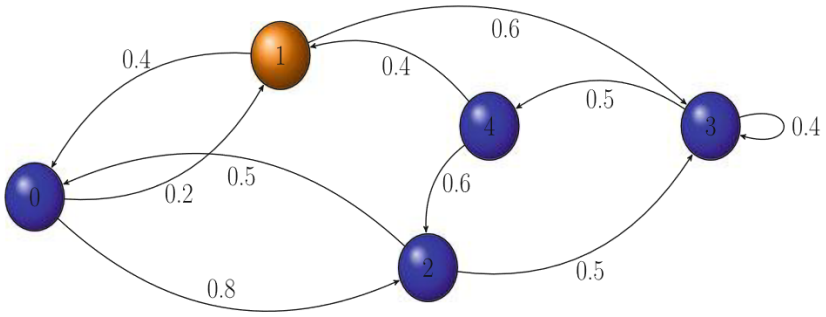


Figure 2.23: Markov Chain modeled as a Finite State Machine [14]

2.3.2 Markov Chain (MCh)

A Markov Chain (MarkCh) is a tool to model a stochastic process. A process is said to be a Markov process or have a Markov property if the probability of distribution of the following state S_{n+1} is dependent only on the previous state S_n at the time n and not on previous states S_k for $k = n - 1, ..., 0$ [14, 92].

The MarkCh itself can be modeled as a combinational (memoryless) Finite State Machine (FSM) and a set of transitions bounded to the stochastic process probability distribution [14]. The probability distribution of the process can be obtained by observing transitions from the events. The sum of the state transitions exiting from each state adds up to 1. Figure 2.23 depicts a MarkCh as a FSM, with 5 states and 10 state transitions.

2.3.3 Monte-Carlo (MC)

The Monte Carlo (MonCar) methods are a set of stochastic modeling techniques that uses random events to simulate complex probabilistic events [93]. Random events are used to extract a good representation of an underlying complex distribution of a given process [93].

2.3.3.1 Mathematical background

According to [94], the MonCar methods have two aims: generate samples from a given distribution $P(x)$, and estimate expectations of functions under the distribution P, as in $\Phi = \langle\phi(x)\rangle \equiv \int d^N x P(x)\phi(x)$.

The sampling goal consists of evaluating a P^* function which we assume is distant from the target P distribution by a multiplicative constant $Z = \int d^N x P^*$, as in $P(x) = P^*(X)/Z$ [94]. However, Z is unknown and hard to compute in both the continuous and discrete versions [94].

The second goal of estimating the expectation of a function $\phi(x)$ can be made using the discrete version of the sampling function above [94]. However, the distribution of $\phi(x)$ or $P^*(x)$ may have a distribution that is concentrated in a few spots of a large state space, and sampling enough to cover them requires many samples [94].

To reduce the number of required samples, different methods are used. Both importance and rejection methods make use of a proposal density $Q(x)$ to generate random samples x^r and study $\hat{\Phi} \equiv \frac{\sum_r w_r \phi(x^r)}{\sum_r w_r}$ with $w_r \equiv \frac{P^*(x^r)}{Q^*(x^r)}$ [94]. The importance sampling requires that the samples are in the typical set of P, otherwise the simulation will require a lot of samples, unless Q is a good approximation of P [94]. For rejection sampling, Q should also be a good approximation of P, otherwise a large sample size will be required, and most samples will get rejected[94].

2.3.3.2 Simulations

Systems simulations can be classified as static, if the system does not evolve over time (time-invariant), or dynamic, if the system evolve over time (time-variant) [95, 96].

The MonCar equivalents for the static and dynamic simulation models are the MonCar MarkCh and MonCar sequential models, respectively. Sequential MonCars are used for scenarios where future results depend on previous states, while Markov Chains are used for scenarios where future results only depend on a stationary probability and the current state [93].

The main use cases of MonCar simulations are: the estimation of values through stochastic methods (e.g. calculating the Pi value by calculating the fraction of random 2-D coordinates samples inside of a circle circumscribed by a square with same width divided by the number of total samples), machine-learning (e.g. gradient descent), optimization problems, simulation of a given system, and visualization of complex systems through sampling [93].

2.4 Related Works

2.4.1 Markov chains in individual and collaborative spectrum sensing

Markov chain based solutions have been used in CR methods for the Carrier Sense Multiple Access-Collision Avoidance (CSMA/CA) protocol, as proposed in [15], which uses a non-collaborative approach. The SS results and the exponential back off are used to determine the probability of transmission and probability of collision in a channel [15]. The joint stochastic model is made using a bi-dimensional MarkCh composed of two components

A, derived from the CSMA/CA stochastic process for the exponential back off, and Pi, derived from the proposed stochastic process of the spectrum sensing technique [15]. The authors prove that both stochastic processes are independent and the probability of a transmission happening can be expressed as $\tau = \sum_{i=0}^{m} b_{i,0} \sum_{s=0}^{C-1} s_c$, where $b_{i,0}$ is the stationary probability of the CSMA/CA back off, s_c is the stationary probability of the SS indicating whether a channel c is idle or not and m is the number of the back off stages [15]. They also derive the probability of collision with at least one of the PU nodes transmitting as $p_c = 1 - (1 - \tau)^{n-1}$, where $n - 1$ is the number of other PUs sharing the same channel. The proposed MarkCh FSM is depicted in 2.24. The proposal does not work with CSS, which reduces the amount of information the nodes could have access to, increasing the chance of collisions and interference.

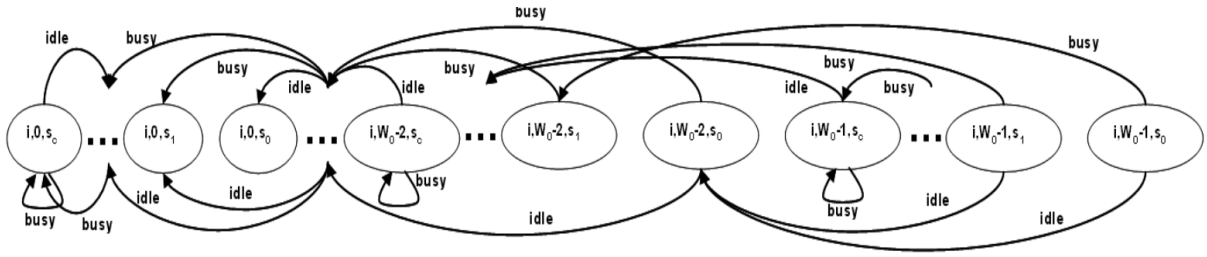


Figure 2.24: Markov chain to model the joint stochastic process of the CSMA/CA and spectrum sensing [15]

A different proposal is made in [97], using a CSMA/CA based reporting scheme on top of a Common Control Channel (CCC) licensed channel. In the proposed solution, a sensing node performs the sensing on a candidate channel, then switches to the CCC channel where it performs the PD [97]. If the CCC is occupied, the sensing nodes switches to the next candidate channel and repeats the cycle [97]. If the CCC is free, the sensing node broadcasts the sensing report [97]. Authors did not explore fusion techniques, nor specified the saved bandwidth of their technique [97].

Yet another technique for CSMA/CA is presented in [72], where a Semi-Distributed Cooperative Spectrum Sensing (SDCSS) technique is proposed along with a p-persistent CSMA/CA, which according to the authors outperforms the CSMA PD [72]. Their proposed SDCSS technique consists in assigning a set of channels to be sensed by each sensing node at every sensing cycle [72]. After sensing all the assigned channels, each sensing node transmits its sensing results for other nodes, which collect the reported data [72]. After all nodes finish transmitting their reports, a new sensing cycle starts, and every node by itself decides each channel is idle or busy depending on the received reports [72]. After all nodes identifies empty channels, if the node decides to access a channel, it reserves a time slot with a 4-way handshake with Request-To-Send (RTS)/Clear-To-Send (CTS) [72].

Authors do not mention bandwidth savings or overall probability of detection, focusing instead on estimated throughput.

2.4.2 Localization and statistical sampling techniques to improve fusion

Regarding CSS, correctly identifying the nodes that have better performance in sensing has a fundamental role for improving efficiency and accuracy in PU detection. Then, the approach is to find an efficient relationship between the amount of data transmitted on the CCC and the detection performance. One way is to define that the UEs with the highest detection probability have the best detection performance [98]. 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 [99] or to select devices with higher gain antennas, that can collect other local user reports with a side-channel and forward them to the Fusion Center (FC) [8].

In [100], the authors study grouping methods (random, reference-based, statistic-based and distance-based) to select adequate UEs to report in order to save bandwidth, energy and improve the detection probability. From the different grouping methods studied, only the random clustering did not require localization information [100]. Not requiring localization information can save power and increase the number of scenarios in which the technique can be used (e.g. single tower cells without specialized hardware required in Angle-of-Arrival (AoA) and Time-of-Flight (ToF) techniques) [100]. The authors show that the random clustering outperforms the other grouping methods [100].

Other grouping technique for CSS is presented in [101], where the sensing nodes are grouped into clusters and the node from each cluster that is the closest from the FC becomes the cluster head. The cluster head receives energy sensing reports from the nodes within its cluster, fuses them with an energy fusion technique, and forwards the result to the FC [101]. The FC uses an OR fusion to consolidate the cluster reports [101]. The technique can save a lot of bandwidth on the CCC if a side-channel is used for intracluster transmissions [101]. Authors claim their technique probability of detection rates are an order of magnitude higher than traditional CSS for the same scenario [101]. The authors don not mention the reliability and protection against attackers, which can be mitigated with more advanced fusion techniques but not with the traditional OR fusion.

2.4.3 Spectrum sensing trust anchor

To circumvent the lack of reliability in the SS reporting, we can use additional information to validate the report, like tracking the CQI levels of a UE. However, the CQI levels are self-reported, which means a malicious attacker can also report wrong values on purpose. A technique to mitigate false reporting of Channel quality Indicator (CQI) in LTE networks is proposed in [40]. The technique is because the standard implies the existence of a correlation between retransmissions and the CQI levels [40]. If the number of retransmissions is higher than 10%, either the quality of channel deteriorated in the meantime or the reported CQI is higher than the optimal CQI, which indicate a possible fraud [40]. If the number of retransmissions is low but the CQI remains low, there is also indication the CQI is fraudulent.

2.4.4 Mitigation of Byzantine attacks

One of most crucial points on CSS is that relying on third-party sensing results ends up opening opportunities for malicious users that may report wrong information to disrupt the network, especially in an opportunistic access scenario.

A framework for detection of false sensing reports is proposed in [36], along with a clustering technique to assess the sensing node trustworthiness. The clustering technique is used to calculate the probability of each sensing node to be legitimate and discard bad reports before fusion [36]. The authors assume that if the sensing period is smaller than the transmission period to be detected, the probability of legitimate users having the same consecutive result is higher than malicious user, depending on the rate they flip their results [36]. This premise is also used in this dissertation.

Another clustering technique, Fuzzy-C means, is proposed in [37], also used to generate weights for the sensing node reports, which are summed and compared against a threshold to determine the final sensing result[37]. The basic premise of the work is that the sensing period is far smaller than the primary user activity period, from which we can assume the number of consecutive results for the sensing nodes should be stable throughout the entire period, while attackers should fluctuate during stable periods [37]. The technique shows good results, with an error rate an order of magnitude smaller than the traditional reputation technique they compare against [37]. The authors, however, overlook shadowing effects, which may severely affect the fluctuation of the sensing results from the legitimate nodes.

Still on protection against Byzantine attackers, [38] proposes a sequential test that eliminates false reports in both mobile and static scenarios. The proposed technique starts by collecting sensing reports from the sensing nodes and assignment of a trust value based

on the correctness of the reports over a period [38]. To circumvent cases where most of the sensing nodes are malicious, the FC verifies if any scheduled transmission over an opportunistically accessed resource was correctly received by the recipient [38]. If the data was correctly received, the fusion decision from the previous period is confirmed as correct [38]. The author effectively mitigates the Byzantine attack problem, but the premise assumes that the PU transmissions will interfere on the data reception, which may not be the case for all sensing nodes.

Another technique to identify Byzantine attackers is proposed in [39]. The technique uses an initial training period to collect normal operation of the network and build a defense reference, measuring the reputation of each sensor node [39]. The reputation of each node is computed by comparing their report to the majority rule based fusion [39]. The evaluation period follows the training period, where no node tagged as Byzantine can participate in the fusion [39].

2.5 Chapter review

This chapter reviewed concepts regarding the evolution of mobile networks, with a short introduction to mobile network architecture, the two main components associated with the radio access (UE and eNodeB), the radio frame and radio resources subdivisions.

The chapter also reviewed distributed and centralized MAC techniques, including the resource matrix for the LTE standard, known as DCI/UCI for downlink/uplink channels. Spectrum sensing techniques are then presented along with the concepts of collaborative spectrum sensing, which aggregates sensing results of multiple nodes into a single result based on different fusion techniques. The collaboration may be either distributed, with nodes broadcasting their sensing results to nearby nodes, or centralized, with nodes transmitting their results to a single FC.

After that, statistical tools are presented, including: the central limit theorem, which states the average of uniformly spaced samples of an unknown random distribution tend to either the real value or the real average of the distribution; Markov chains, which can be used to model stochastic processes; and Monte Carlo techniques, usually used to model a complex system with a large number of parameters or with high computational complexity.

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. Even though clustering techniques are used to save up bandwidth and improve the reliability of the system, authors ignore the possibility of fraudulent users, which would tank their results. To the best of our knowledge, in the related works listed to date, none addresses DSA in a rural

scenario, also focusing in resilience to malicious users/attackers considering optimization of control channel transmissions in CSS reporting, while maintaining low false positives and negatives in channel detection.

In the next chapter, our proposed techniques to improve both the individual and collaborative spectrum sensing will be presented.

Chapter 3

Markov Chains for Improving Collaborative Sensing

This chapter is divided into five sections: Section 3.1 lists and justifies the assumptions the proposed techniques are based on. Sections 3.2.2 and 3.3.1 describe the proposed techniques to increase the Collaborative Spectrum Sensing (CSS) accuracy and the resilience against attackers. Section 3.4 describes how the collaborative spectrum sensing cycle was implemented on top of the Long Term Evolution (LTE) stack using the Network Simulator 3 (ns-3) simulator for evaluation of the proposed techniques.

The first technique filters noise of the hard-combining individual Spectrum Sensing (SS) using a Markov chain, reducing unnecessary reporting transmissions and saving Common Control Channel (CCC) bandwidth. The second technique filters less relevant reports from the CSS fusion, improving resilience.

3.1 Premises and Reasoning

Both of the proposed techniques are based on the following premises:

P.1 the individual SS depends unique and exclusively on the User Equipment (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, Evolved Node B (eNB) and the Primary Users (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 $t_{putransmission}$.

Premise **P.1** assumes that the individual SS technique results will reflect the vision of a given channel by each sensing UE, which is to be expected from the individual SS. Those techniques results can be compiled into probability of detection curves $p_d(d)$ or $p_d(SINR)$, based on either distance d between the UE and the PU or the perceived Signal to Interference plus Noise Ratio (SINR). The technique results are produced by the SS technique either with measurements and/or link-layer simulations.

Premise **P.2** derives from the fact that most of the population underserved by internet connectivity lives in remote and sparsely populated areas, as in a rural scenario. Using a super-cell with 50km radius proposed in the 5G-RANGE project, we assume the UEs act as relay nodes and are static, due to high gain antenna alignment requirements. We also assume the PUs are static broadcasting stations. The eNB is also assumed to be static, as proposed in the 5G-RANGE project. A moving eNB, as in the Google's project Loon would require additional validation and adaptation of the proposed techniques, especially if moving at high speed.

Premises **P.3** and **P.4** assumes that an UE nearby a PU has high probability of correctly detecting its transmission pattern, which derives from the fact that the probability of correctly detecting transmissions increases the sensor and source get closer.

Premise **P.5** assumes the sensing period is smaller than the PU transmission period, which results in multiple sensing cycles during a single transmission period. As the ratio of sensing cycles per transmission period increases, the amount of sensed information increases, and temporal behavior can be analyzed.

3.2 Individual Spectrum-Sensing Smoothing technique

3.2.1 Fundamentals

From the premises **P.1**, the probability of detection is given by the Equation 3.1. 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 3.2.

$$p_{sense}(t, d, pu_{active}) = \begin{cases} p_d(d(t)), & \text{if } pu_{active}(t) = 1. \\ p_{fp}, & \text{otherwise.} \end{cases} \tag{3.1}$$

$$\binom{n}{k} p_{sense}^k (1 - p_{sense})^{n-k} \tag{3.2}$$

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 3.1, where the source probability of detection (blue squares) is further attenuated as the number consecutive samples with same results $k = n$ grows.

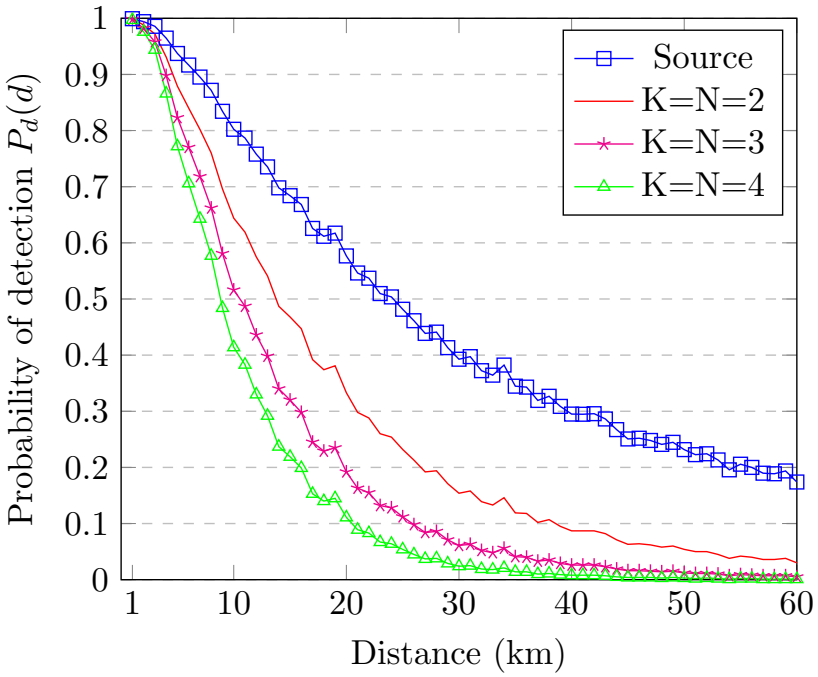


Figure 3.1: Probability of detection curve based on distance

In our proposal, to aggregate the UEs individual sensing results, we use the Markov Chain (MarkCh) presented in Figure 3.3. This is valid if, and only if, the premise **P.5** is truth, which means that multiple individual sensing procedures are executed for each PU transmission. The MarkCh from Figure 3.3 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 MarkCh 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

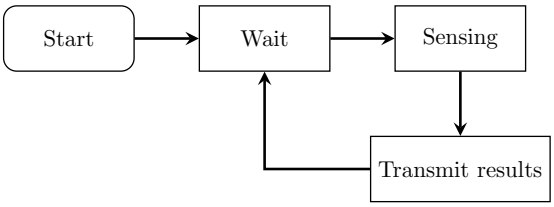


Figure 3.2: Steps of the Individual Spectrum Sensing

certainty threshold of $L = 0.9$, the number of number of states of the MarkCh 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 3.1, as calculated with Equation 3.2 and $P = P_{sense}$.

3.2.2 Proposal of Markov Chain-based technique

The standard SS procedure is described in Figure 3.2, which shows a waiting cycle, followed by the sensing method and transmission of results to the Fusion Center (FC). We consider that the sensing results have short-range statistical dependence, then, for the first technique our approach is to use a simple MarkCh with S states in the SS as shown in Figure 3.3, 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 3.4, with two additional checks used to implement the MarkCh. 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.

At the same time, the number of reports increases along with the number of sensing nodes (UEs). The non-transmission of a same consecutive sensing result reduces the

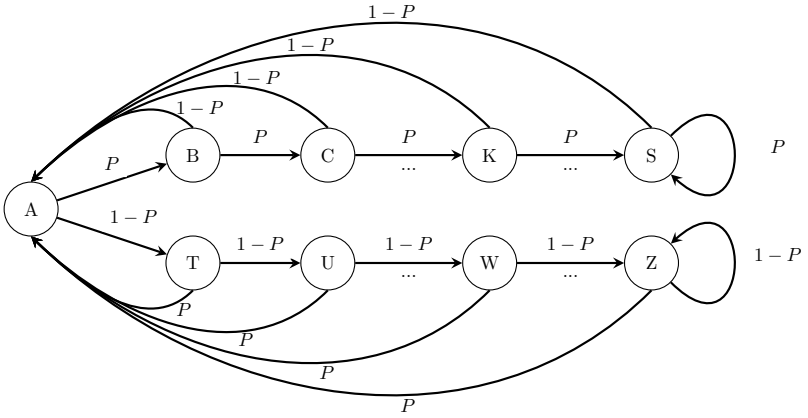


Figure 3.3: Markov chain with $S = Z$ depth levels

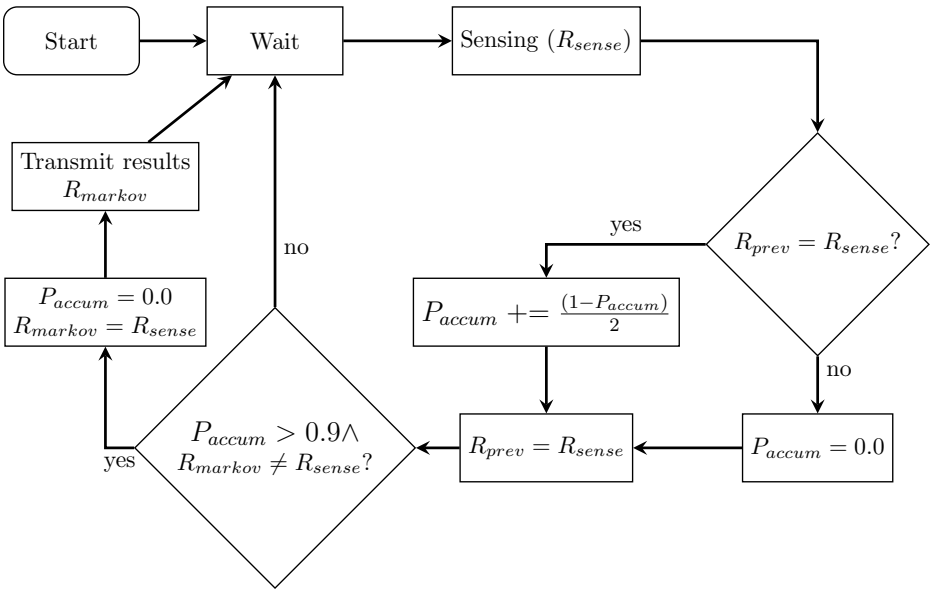


Figure 3.4: Steps of the Individual Spectrum Sensing with the Markov chain

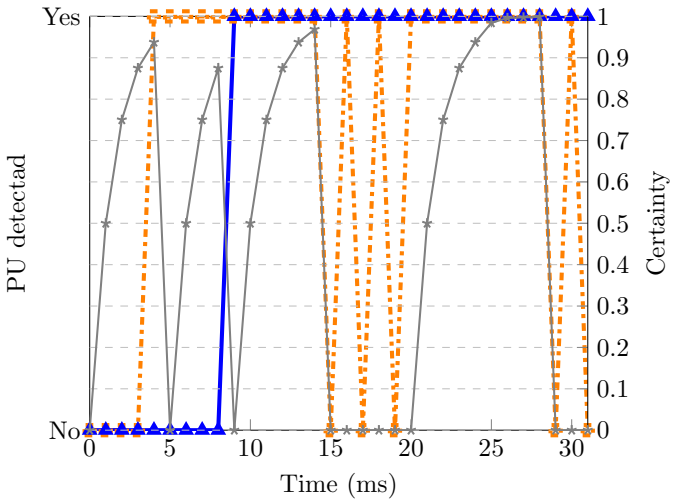


Figure 3.5: Comparison between the untreated sensing and the sensing with the 5-state Markov chain

amount of reporting transmissions, where the FC assumes that the channel state remained the same for a user that did not transmit in the current cycle. As a side-effect of our proposal, the number of transmissions is further reduced as shown in Figure 3.5. The sensing results of a given UE are represented in the dot-dashed orange line. By not transmitting equal consecutive results, the number of reports fall from 31 to 10. With our proposal, the amount of reports fall from 31 to 1, as the sensing results never reach a minimum amount of certainty to switch state for a second time. The convergence can be sped up by reducing the number of states in the MarkCh Finite State Machine (FSM), also reducing the certainty of the result and increasing false positives if using the OR fusion.

3.3 Collaborative Spectrum-Sensing Filtering technique

3.3.1 Proposal of Harmonic Mean-based filtering technique

The second technique uses MarkCh based mechanism like the one described in 3.2.1, which is used to discard reports from UEs that are less relevant to fusion, as shown in Figure 3.6. The filtering policy is based on the Channel quality Indicator (CQI) reported by the UEs, as described in [40]. The filtering works as follows: a high CQI indicate a high Signal to

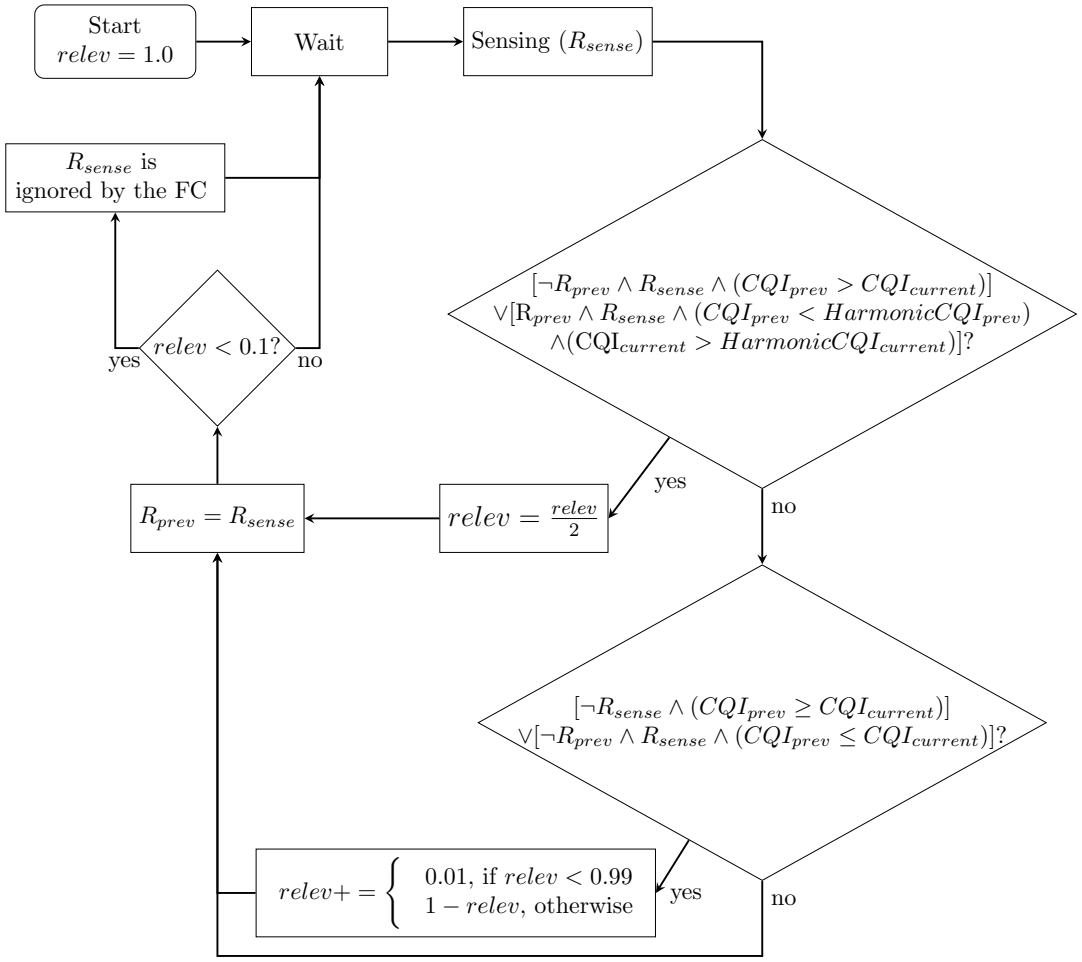


Figure 3.6: Harmonic mean-based spectrum sensing report filtering scheme

Noise Ratio (SNR), while a low CQI refers to a low SNR. As the PU starts transmitting, it interferes with the UEs and CQI drops, especially for those UEs closer to the PU. The harmonic mean of the CQIs is used to find the UEs closer to the PU if the PU is active, and UEs further from the eNB if the PU is inactive, as the mean tends to get smaller, as illustrated in Figure 3.7. The first check in Figure 3.6 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, which is the expected behavior to take advantage of the Dynamic Spectrum Access (DSA). It also checks for UEs that just started sensing the PU and had a brief spike in their CQI, which is rare, but may happen when the UE is shielded from the PU interference.

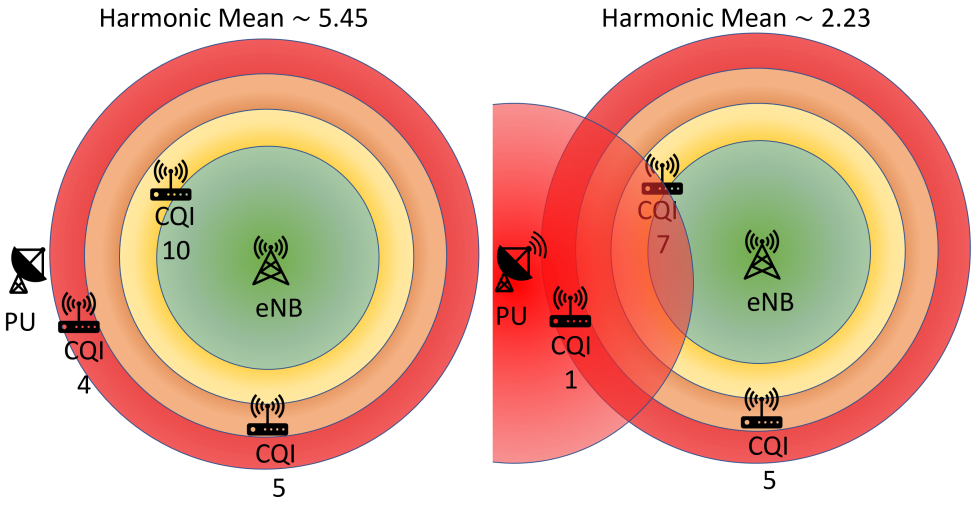


Figure 3.7: The harmonic-mean based filtering favors nodes distant from the eNB when the PU is inactive. Otherwise, it favors nodes closest to the PU.

3.4 Collaborative sensing cycle implementation on the LTE stack

The evaluation of the proposed techniques required the implementation of the collaborative sensing cycle on a wireless network stack. The stack of the LTE was chosen due to similarities between it and the new 5G-NR stack. The simulator used was the ns-3 [41].

The proposed techniques rely on a physical layer spectrum sensing technique, which is implemented in form of a function that interpolate probability of detection curves. The probability of detection curves can be obtained from either real measurements or link-layer simulations, which was our case. The probability curves indicate the probability of an UE detecting a PU transmission and can be based either on the Euclidian distance between UEs and PUs, or SNR levels measured by the UEs. A random number generator following a Bernoulli distribution with the interpolated probability represents if the channel is sensed as idle or occupied. After the sensing procedure, the sensing results are transmitted to the eNB using the CCC. The eNB stores the results and fuses the results together during the subframe indication. Multiple fusion techniques were implemented, in special the classic k-out-of-n variants.

To implement the collaborative sensing in the ns-3 simulator, several changes are required. Most changes are concentrated in the LteSpectrumPhy model, that simulates the interaction between the physical layer models (LteUePhy and LteEnbPhy) and the channel model (MultiModelSpectrumChannel). Since the LTE model implemented in ns-3 does not support cross-carrier scheduling, one is logically split into multiple subchannels by dividing groups of Resource Block Groups (RBGs). After that, the average SINR for each

subchannel is calculated and the sensing procedure is executed. The SensingProcedure can use either the previously calculated SINR for each channel, or the distance to the PUs, calculated within the channel model. The SensingProcedure proceeds to the interpolation of the probability of detection for each subchannel, then uses the probabilities to generate random samples from a Bernoulli distribution. The samples are used to prepare a bitmap for each channel containing sensing results, plus additional information only used for analysis of the results, including: flags for false positive, false negative, fake report of an attacker and finally if the sensing result is the same of the previous procedure.

The LteSpectrumPhys StartRx requires changes to identify subchannels with peaks of power from non-LTE transmissions, that is stored along with the distance between transmitter and receiver and the PU distance registry. The StartRx function also requires changes to forward the sensing results to the MAC layer, that will enqueue a new control message containing the sensing report data, that will then be transmitted to the eNB.

The collaborative spectrum sensing requires additional changes to both UE and eNB physical and Medium Access Control (MAC) layers to properly forward the spectrum sensing control messages to the eNBs MAC layer, where the individual sensing reports are stored for later use. During the subframe indication, the eNB prepares information required by the resource scheduler, one of them a map containing the empty RBGs. Before calling the scheduler, the fusion technique processes the most up-to-date sensing reports and creates a mask which indicates that the RBGs is occupied by the PU transmission. The mask is passed down to the resource scheduler and is used to prevent allocation to the RBGs, reducing the probability of collisions with the PU.

The 3.8 shows how the individual spectrum sensing on the LteSpectrumPhy works. In the initial StartRx node, there is the forwarding of the latest sensing results to the MAC layer. Starting from the StartRx and moving to the right, there is the common data path for control messages, that also triggers the sensing, guaranteeing it is synced with the eNB. Starting from the StartRx and moving down, there is the common data path for non-LTE messages, where the subchannel in which the PU is transmitting is be identified and stored along with its range to the current UE using a ResetPuPresence, used to both store and remove distance registries.

The Figure 3.9 shows how the collaborative spectrum sensing on the eNB works. The control message sent by the UEs is received by the eNBs LteSpectrumPhy and them forwarded to the LteEnbPhy and LteEnbMac, where the sensing report is stored. Also, in the eNBs MAC, the DoSubframeIndication prepares information for the resource scheduler and then calls the scheduler. Just before calling the scheduler, the MergeSensingReports function prepares the fusion of the individual sensing reports using the latest available data. This function also collects false positives and negatives data for later use. The

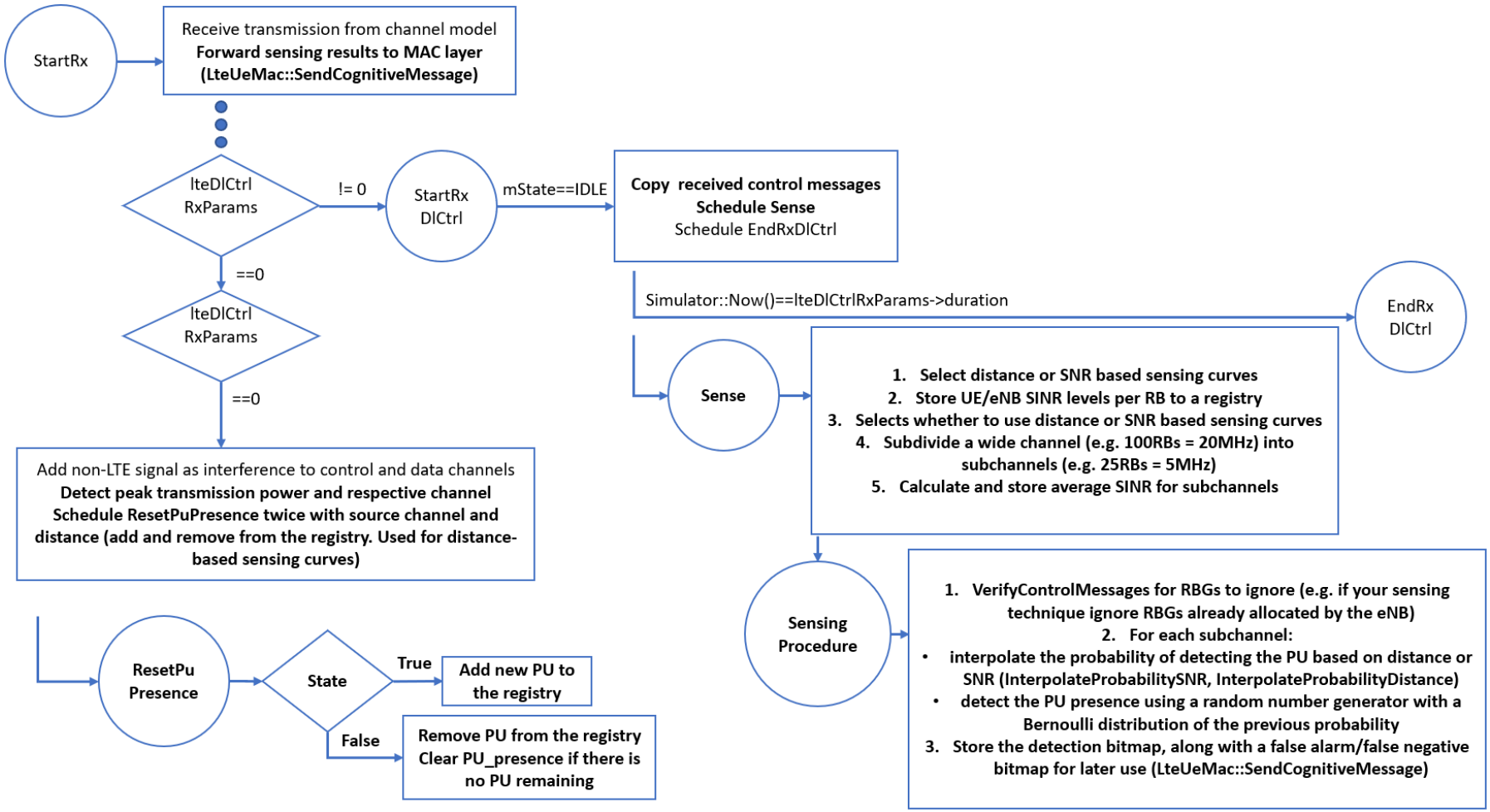


Figure 3.8: Changes for the LteSpectrumPhy model required by the proposed individual spectrum sensing cycle

fusion result is a bitmap, used by the scheduler to tag RBGs that should not be scheduled for the UEs.

The complete CSS cycle is shown in 3.10, with changes highlighted in bold text, starting with: StartRX, from SpectrumChannel (5) to the UE SpectrumPhy (6), which triggers the spectrum sensing function and saves the result to a local variable; SendingSensingReport, from the UE SpectrumPhy (6) to the UE MAC (8), which triggers the creation of the control message containing the sensing report with the results of the previous spectrum sensing; SendLteControlMessage, from UE MAC (8) to UE MAC (8), which copies the message to the eNB MAC (1) buffer; ReceiveLteCtrlMsgList, from eNB SpectrumPhy (4) to eNB Phy (3), which receives the spectrum sensing report and forwards to the eNB MAC (1); DoSubframeIndication, from eNB MAC (1), which triggers the MergeSensingReports, that fuses the reports and collect performance metrics, then triggers DoSchedDlTriggerReq of the eNB Scheduler (2), that use the fused results to mark RBGs as occupied or not before allocating resources to the UEs.

The above described changes to the ns-3 simulator were published in [102] and the source code is publicly available along with the usage instructions in `https://gabrielcarvfer.github.io/NS3/COLAB/`.

3.5 Chapter review

In this chapter we presented a Markov chain-based technique to improve the reliability of individual spectrum sensing, at the cost of reduced detection range, and the Harmonic mean-based technique that discards sensing results from UEs that are most likely distant from the source of interference, at the cost of reduced probability of detection. We also propose changes to the LTE model in the ns-3 simulator to implement a simple CSS cycle that enables the simulations of the proposed techniques in more realistic conditions than most works reviewed in Chapter 2.

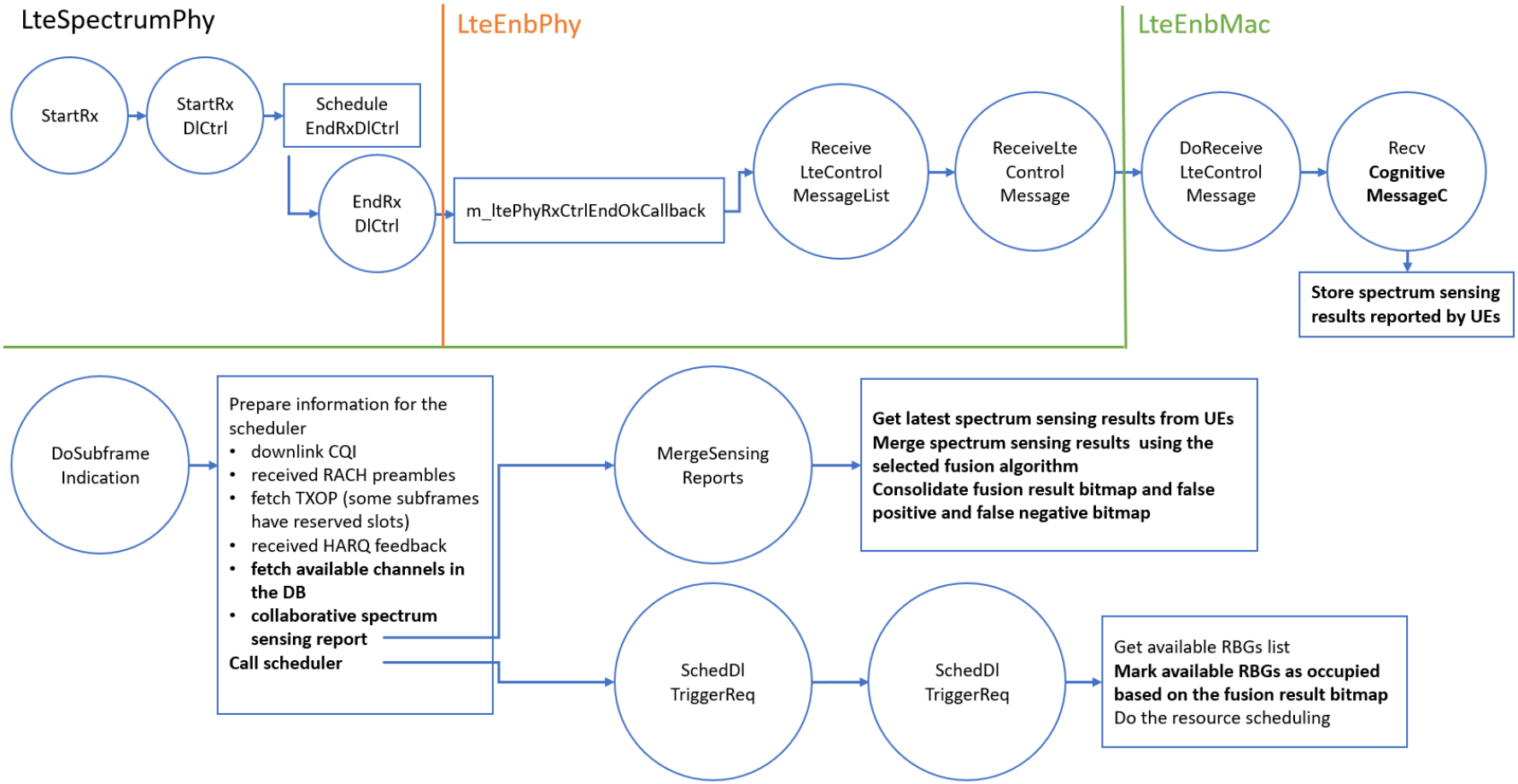


Figure 3.9: Changes for the eNodeB model required by the proposed collaborative sensing cycle.

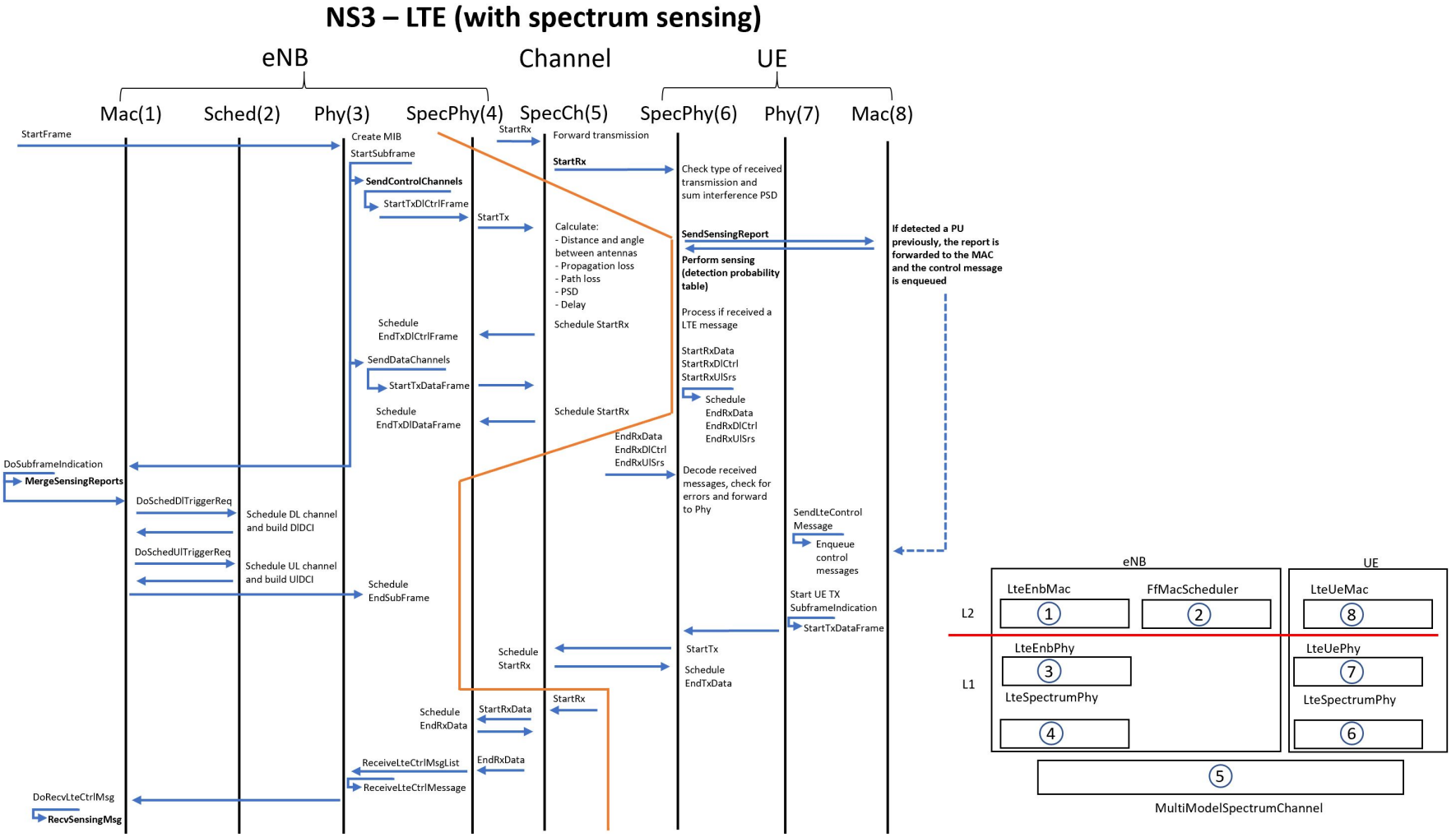


Figure 3.10: CSS cycle impact on the LTE stack event-flow in ns-3

Chapter 4

Evaluation and Simulations Results

This chapter presents the methodology and simulations that were developed during the research and implementation of this work. First, we describe the methodology in Section 4.1, then, in the following sections, for each of the simulations we describe the scenario and its results.

4.1 Methodology

4.1.1 Tools and methods

The simulations were ran using a modified ns-3 simulator, nicknamed COLAB [1]. The first release was based on the ns-3.29 and was used for the CSS validation. The second release, based on the ns-3.30.1, was used for the rest of simulations.

For the validation scenarios, we ran a simulation for each scenario and each fusion technique, using a fixed seed, then compared the results against the mathematical models. All parameters were kept the same, especially the distance of the nodes to one of the PUs, used to interpolate the individual probability of detection from the input curves.

For evaluation, we follow best practices from [103]. A Monte Carlo simulation approach is used, performing a simulation for each of the 10 scenarios generated with the same fusion technique (4), Markov and Harmonic techniques (4), number of UEs (4) and distribution (2), with different number of attackers (3), resulting in a total of 3840 simulations. Each simulation ran for 10 seconds of simulated time post the transient (initial exchange of messages to setup the LTE network), resulting in 10000 individual sensing cycles executed per UE plus 10000 collaborative sensing cycles of the eNB, per simulation. The other simulation parameters (transmission power, antenna gains, PU transmission duty cycle, transmission interval and packet size, user application, position of the eNB,

[1] `https://gabrielcarvfer.github.io/NS3/COLAB/`

channel bandwidth, carrier frequency, simulation time and motion speed) are fixed. Errors from derived metrics are properly propagated and the margins of error are calculated using a 95% confidence interval using the t-Student distribution.

Most of the other fixed parameters (transmission power, antenna gains, position of the eNB, channel bandwidth, carrier frequency and motion speeds) are based on a long range rural network, as proposed in the 5G-RANGE project [99]. The PU transmission interval and duty cycle were randomly generated. The link-layer individual spectrum sensing technique is modeled after the Window-Based Energy Detector (WIBA) [65], which is the proposed Spectrum Sensing (SS) technique for the 5G-RANGE project [99]. The spectrum sensing technique is modeled based on the probability of detection curve in Fig. 3.1. As this dissertation focuses strictly on CSS performance, we use an unrepresentative user application workload and ignore the resulting general network performance metrics like throughput and latency.

The proposed techniques and implemented tools can be used in the future with more realistic scenarios and traffic workloads if the premises are still valid.

4.1.2 Performance Metrics

In order to assess how much the proposed techniques impacted on the collaborative sensing results, we use three different metrics:

False positives which represent the probability of detecting a Primary User (PU) transmission when it is not transmitting. It can be calculated as the ratio of subframes where the PU is incorrectly detected and the number of subframes where the PU was not present, as shown in Eq. 4.1. This metric is used to compare how much spectrum will remain subutilized with each different fusion technique.

False negatives which represent the probability of not detecting a PU transmission when it is transmitting. It can be calculated as the ratio of subframes where the PU was active but was not detected and the number of subframes in which the PU was active, Eq. 4.2. This metric is used to compare how much interference to the PU can be expected from each fusion technique.

Fusion accuracy which represents the probability of correctly detecting the PU transmissions. It can be calculated as the ratio of subframes where the fusion result matches the real PU activity/inactivity, Eq. 4.3. This metric shows a combination of both False Positives and False Negatives metrics, but it can be skewed in favor of one of the two metrics depending on the PU transmission pattern and fusion technique (e.g. OR fusion

	Simulation parameters \ Scenario	Spectrum sensing cycle validation	Markov technique evaluation	Harmonic technique evaluation
General	Simulation time	10 s (10^4 subframes)		
	Propagation model	FSPL	5G-RANGE	
	Band	5 ($\sim$ 850 MHz)		
	Number of channels	4		
	Channel bandwidth	3x5.2 MHz + 1x4.4MHz		
	PUs per channel	1		
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	[10, 20, 50, 100]	
	Number of attackers	0	[0, 1, 2, 5, 10]	
	Tx power	23 dBm		
	Antenna gain	9 dBi		

Table 4.1: Simulation parameters for the different scenarios

technique produces false positives while AND fusion technique produces false negatives. In a scenario with more PU activity, the false positives of the OR fusion will become more apparent, while false negatives of the AND fusion become more apparent in scenarios with less PU activity).

$$P_{fp} = \frac{\#falsePositives}{\#subframesWithInactivePU} \tag{4.1}$$

$$P_{fn} = \frac{\#falseNegatives}{\#subframesWithActivePU} \tag{4.2}$$

$$acc = \frac{\#totalSubframes - \#falseNegatives - \#falsePositives}{\#totalSubframes} \tag{4.3}$$

4.2 Collaborative Spectrum Sensing cycle validation

4.2.1 Validation scenario

The first step to perform simulations in different scenarios is to validate the cognitive cycle implementation. For this, a simulation was done in a simple scenario that we denote as Validation Scenario, shown in Fig. 4.1, with the following parameters:

a) 20 MHz Additive White Gaussian Noise (AWGN) channel, broken into 4x5 MHz channels, noise floor of -174 dBm/Hz, Free-Space Path Loss (FSPL) path loss model. In real scenarios, the TV Whitespace (TVWS) has 24 MHz bandwidth divided into 3 x 8

MHz (Europe) or 4 x 6 MHz (Brazil) data channels. In the simulation, to abstract the physical layer and keep compatibility with the LENA module, the 20 MHz bandwidth emulates the 24 MHZ TVWS channel and translates 100 LTE Resource Blocks (RBs), that were grouped into 50 RBGs of size 2 (also required changes to the DCI Type 0, used in the NS-3/LTE model). Those RBGs were then grouped into 4 subchannels of 3 x 13 RBGs (5.2 MHz) and 1 x 11 RBGs (4.4 MHz).

b) the channel is in band 5 (869 MHz). gNB transmits with 53 dBm and is equipped with 9 dBi gain antennas, resulting in a 50 km radius cell. UEs transmit with 23 dBm and is equipped with 9 dBi gain antennas. The PUs transmits with 40 dBm. The UEs are placed at a fixed 30 km radius from the PU0, while PUs 1-3 are placed at a fixed 40 km from the PU0. The PU0 is 35 km away from the gNB. This scenario characterizes a rural area with a 50 km cell, one gNB and 4 PUs channels.

c) the spectrum sensing curves used in the simulation are based on the WIBA link-level simulation[65], collecting data from 10 UEs. The PU detection probability curves used are shown in Fig. 3.1. To evaluate the different possibilities, at the fusion center in the gNB, 4 different fusion techniques were implemented (OR, 2-out-of-10, 3-out-of-10, 4-out-of-10).

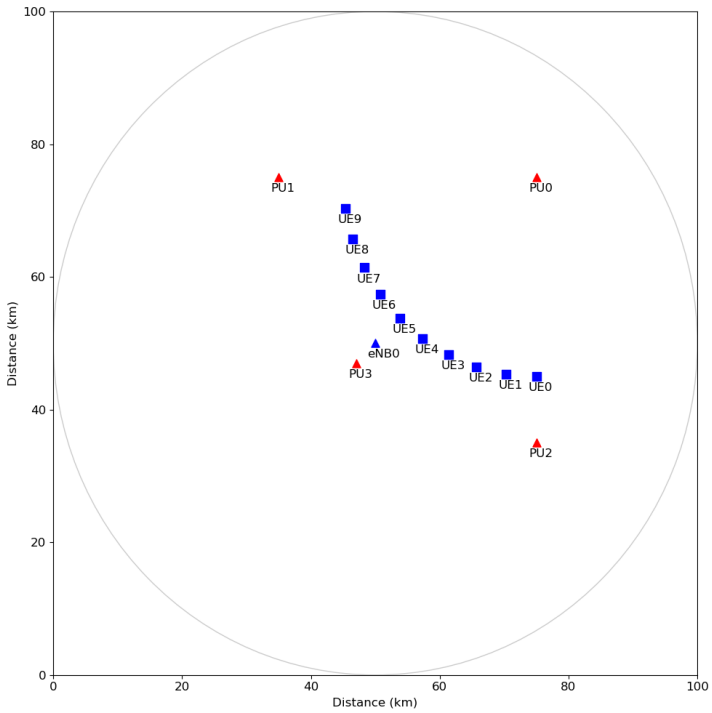


Figure 4.1: Topology of the Validation Scenario for the CSS

4.2.2 Results and analysis

The results for different fusion techniques of the simulated scenario are shown in Table 4.2. False positives are the ratio of subframes where the PU was not present but was detected. False negatives show the ratio of subframes where a PU was present but was not detected and the number of subframes that the PU was present. False positives result in wasted radio resources while false negatives result in interference to the PU. The false positives for all channels are below 10% for all the types of fusion rules while false negatives increase with higher detection thresholds.

The simulation results of the cognitive cycle are shown in Figure 4.2, based on the topology shown in Figure 4.1, that contains 10 UEs, 1 eNB and 4 PUs, one for each channel. The first lines of Figure 4.2 show the perceived PSD of each of the PUs in their respective channel. The second line shows the individual SNR levels of each UEs for each one of the four channels. The third line show the results of the RBGs scheduling, where blank spaces are unused spectrum, blue spaces are scheduled spectrum and red spaces are avoided spectrum. The third line directly reflects the collaborative sensing fusion behavior, being capable of increasing or decrease the false alarms and false negatives. As false negatives implicate in collisions with the PU, the main target with the collaborative sensing fusion should be decreasing false negatives while keeping false positives low.

As the UEs are equidistant from the PU, their probability of detection is identical. We can calculate the theoretical results using Equation 2.1 and compared to the simulated ones. The results of this comparison are shown in Table 4.3, which shows simulated results are on par with the mathematical model, indicating that the model is properly implemented.

	Channel 0		Channel 1		Channel 2		Channel 3	
Fusion	False Positives (%)	False Negatives (%)	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)
OR	9.2	0.4	9.9	0.3	9.3	0.0	9.7	0.0
2-out-of-10	0.4	5.0	0.4	2.2	0.4	0.3	0.5	0.1
3-out-of-10	0.1	17.9	0.0	3.9	0.0	3.9	0.0	0.3
4-out-of-10	0.0	38.1	0.0	27.3	0.0	14.4	0.0	2.8

Table 4.2: Collaborative sensing results for the validation scenario.

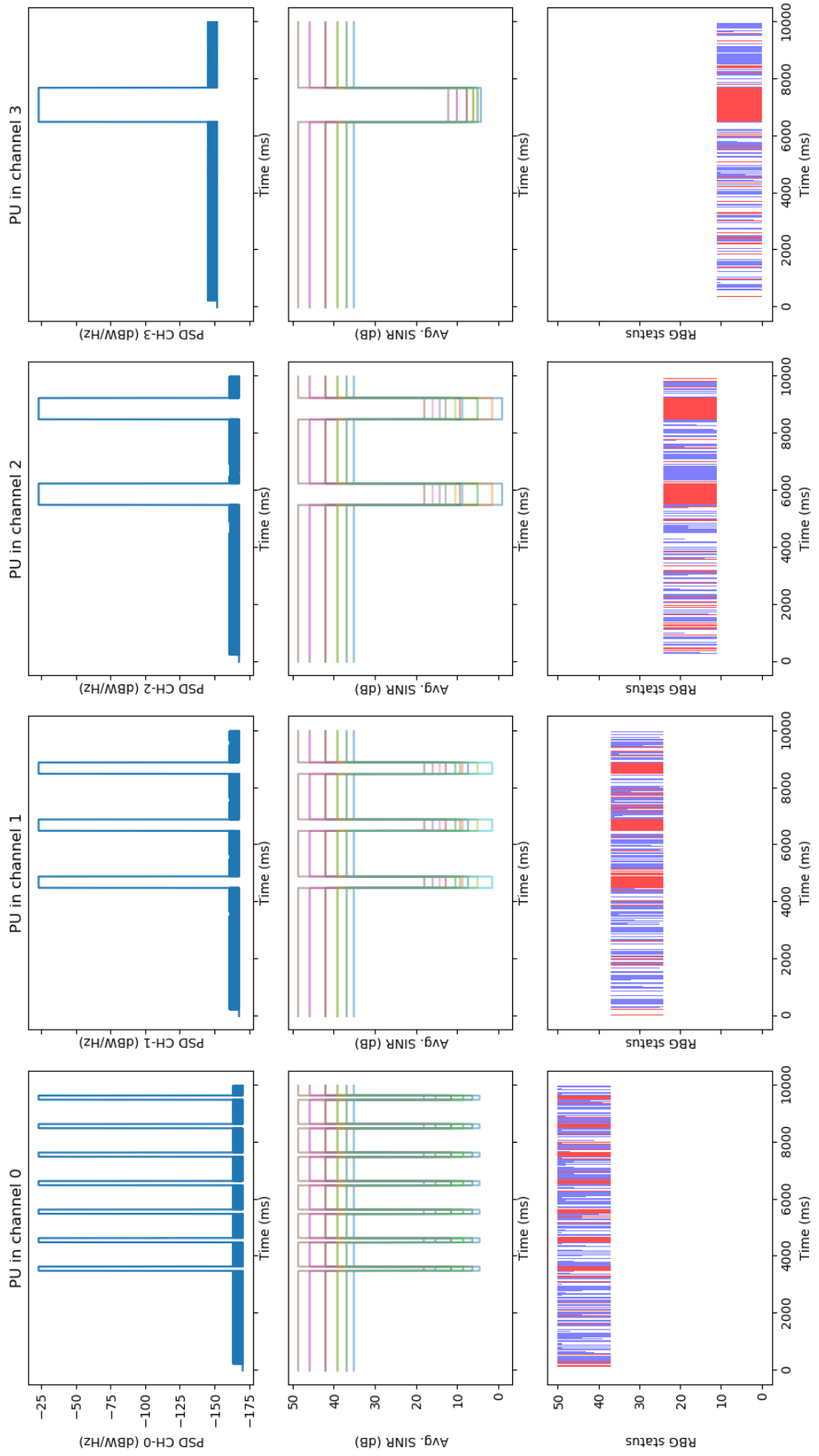


Figure 4.2: Results of the collaborative sensing using OR fusion and radio resource scheduling for the simulated scenario.

Fusion	Channel 0			
	Simulated		Theoretical	
	False Positives (%)	False Negatives (%)	False Positives (%)	False Negatives (%)
OR	9.2	0.4	9.6	0.6
2-out-of-10	0.4	5.0	0.4	4.6
3-out-of-10	0.1	17.9	0.0	16.7
4-out-of-10	0.0	38.1	0.0	38.2

Table 4.3: Comparison between simulated and theoretical results for channel 0.

4.3 Proposed techniques evaluation

4.3.1 Evaluation scenarios

To evaluate the proposed techniques, we developed in NS-3 a scenario that emulates a 5G rural area with a 50km radius cell per Evolved Node B (eNB) and a set of parameters, that were defined based on previous studies[104], as follows:

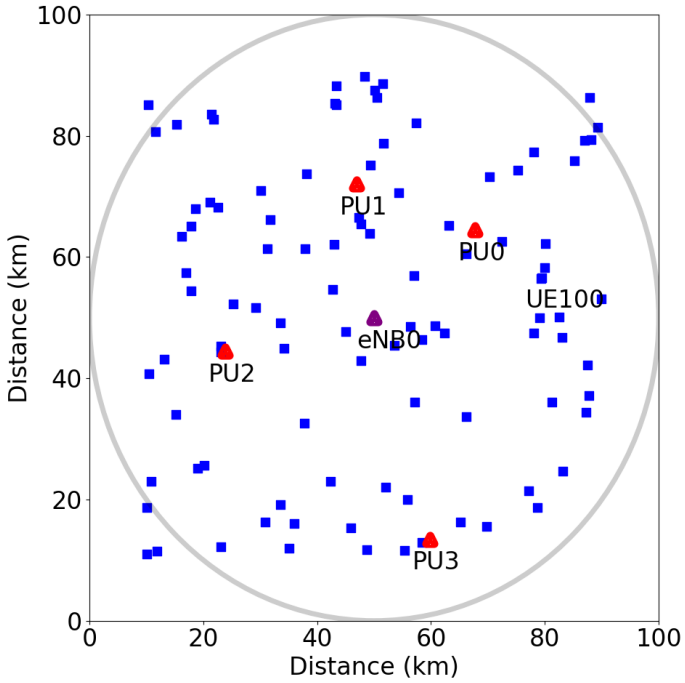
a) a 20 MHz channel with a 5G propagation model for remote areas [99], broken into 4x5 MHz TVWS channels, and a noise floor of -174 dBm/Hz.

b) the channel uses band 5 (850 MHz). The eNB transmits with 53 dBm and has 9 dBi gain antennas. UEs transmit with 23 dBm and use 9 dBi gain antennas. The PUs transmits with 35 dBm.

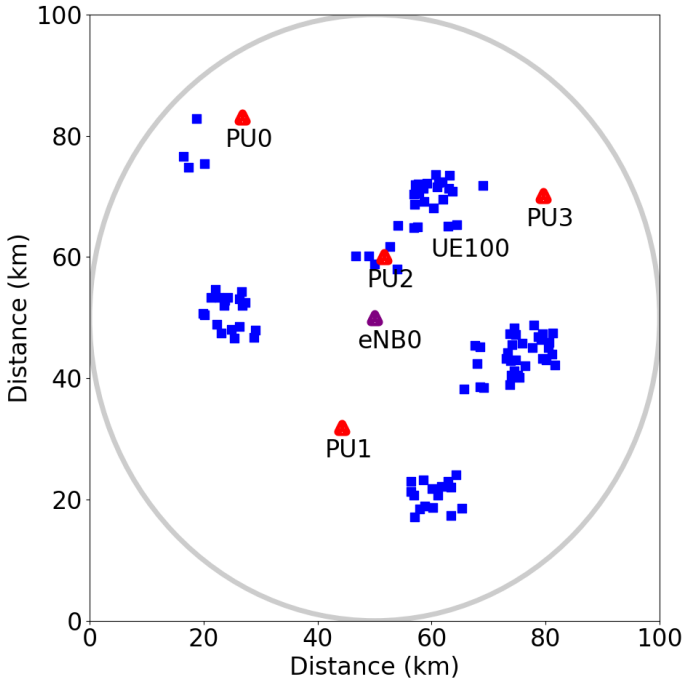
c) the PUs are distributed randomly through the cell and the UEs are distributed either the same or into randomly placed clusters of 5 km radius. The randomly distributed scenario may be unrealistic for such a large cell but provides an interesting approach for evaluating the proposal. The clustered scenario is more realistic and represents a micro-region as small towns/villages. Examples of both randomized and clustered scenarios are illustrated, respectively, in Figures 4.3(a) and 4.3(b). In both scenarios, all the simulations were made with 4 PUs while varying the number of User Equipments (UEs) within 10, 20, 50 and 100. d) we use SS detection probability curves, 3.1, based on the link-layer results using the WIBA [65] sensing technique. The PU detection probability curve is described in Figure 3.5 and was already discussed in Section III.

e) The fusion center uses 4 different techniques: OR, 2-out-of-N, 3-out-of-N and AND.

f) in every simulation, a given number of UEs act as attackers to evaluate the impact considering the proposed techniques using Markov chains and the harmonic mean.



(a) UEs randomly places into the 50km radius cell



(b) UEs placed randomly into randomly placed clusters with 5km radius

Figure 4.3: Scenarios with different UEs placements

4.3.2 Results and analysis

4.3.2.1 Randomized Scenarios

Figure 4.4 shows the compiled results of the simulations using different fusion techniques combinations with an increasing number of UEs per simulation. Each subfigure of Figure 4.4 shows the simulation results for the same scenarios with an increasing number of attackers transmitting fraudulent sensing reports.

The fusion techniques used were:

i. OR/1-out-of-N

ii. OR with the Markov technique, that increases the reliability of the individual spectrum sensing

iii. OR with the Harmonic technique, that filter nodes distant from the PU protecting against attackers

iv. OR with both techniques

The performance for the standalone OR fusion (i.) is shown in blue, OR with the Markov technique (ii.) is shown in orange, OR with Harmonic technique (iii.) is shown in green and the OR with both techniques (iv.) is shown in red.

The baseline scenario without attackers is shown in Figure 4.4(a), where the fusion techniques with the Markov technique (ii. and iv.) performed better than the others without it. The fusion accuracy (how many collaborative fusion results matched the channel state) of the standalone OR and the Harmonic technique (i. and iii.) falls as the number of false positives steadily grows as the number of reporting UEs increases.

The Markov technique (ii.) beats all the other techniques in the scenario without attackers. In the case with 100 UEs, our technique 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.

The Harmonic technique (iii.) behaves similarly to the standalone OR (i.), except in terms of false negatives. In the worst simulated scenario, with 20 UEs, the technique had 0.37 ± 0.092 false negatives, 17x the standalone OR (i.).

The combination of both techniques (iv. performs badly in scenarios with no attackers, with even larger false negatives numbers. As in the Harmonic technique, the worst simulated scenario had 20 UEs, with 31x the standalone OR (i.).

However, false negatives of both Harmonic techniques (iii. and iv.) decrease as the number of UEs increase, providing more spatial information for the fusion. The $0.068 \pm$

0.014 false negatives in the scenario with 100 UEs is reasonably low for opportunistic access.

Figures 4.4(b) and 4.4(c) show the simulation results for the same scenario, but with 2 and 5 attackers within the cell. Both the standalone OR and the Markov technique (i. and ii.) perform badly in this scenario, fusing the fake results, that increases false positives to as high as 100%. Conversely, both techniques using the Harmonic mean (iii. and iv. offer protection against the attacks for large number of UEs, behaving similarly to the scenario without attackers, shown in Figure 4.4(a).

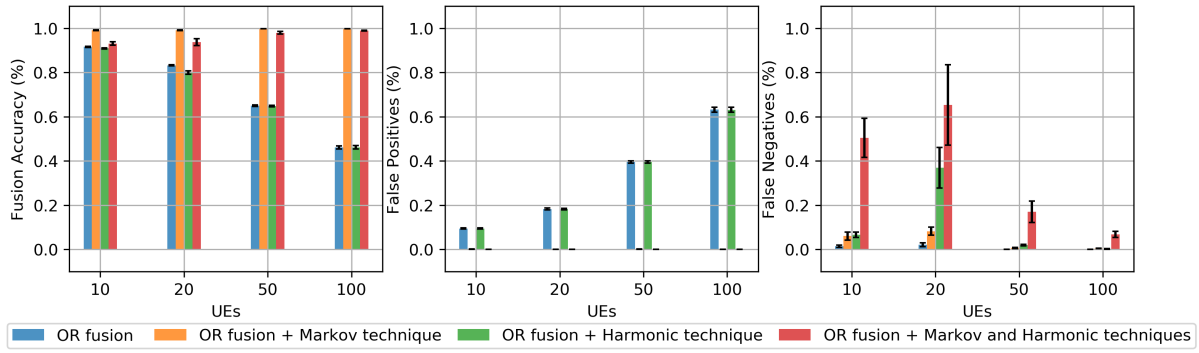
To estimate the number of UEs to guarantee a certain level of false negatives, the data was fitted to a decaying exponential curve using the least-squares method. The resulting curve for the randomized scenario is given by $0.761142e^{-0.0226437x}$, with $R^2 \approx 0.995$. Assuming an interference threshold of up to 0.1 of false negatives, the estimated number of required UEs sensing is of ~ 90 UEs.

The values plotted in the Figure 4.4 are tabled in Tables 4.6, 4.5 and 4.4.

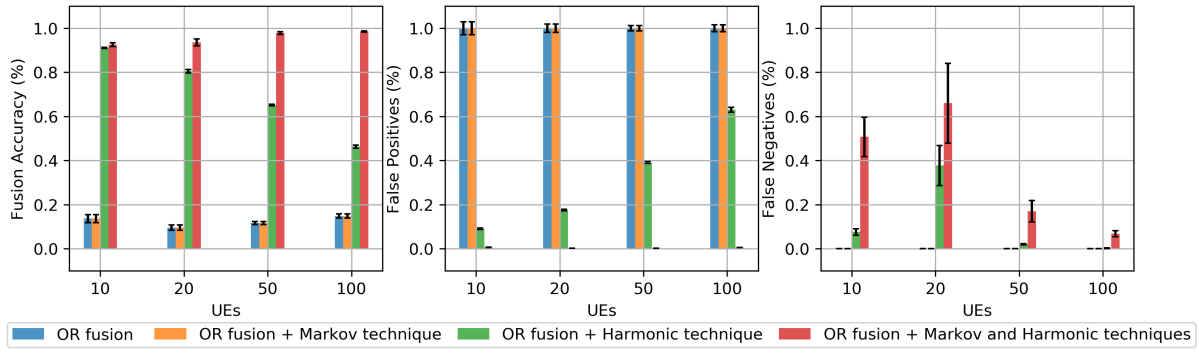
10 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.015 ± 0.005	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.066 ± 0.012	0.076 ± 0.015	0.164 ± 0.028	0.799 ± 0.156
Markov	0.060 ± 0.018	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.504 ± 0.088	0.507 ± 0.089	0.650 ± 0.108	0.896 ± 0.173
20 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.021 ± 0.009	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.370 ± 0.092	0.378 ± 0.091	0.413 ± 0.085	0.498 ± 0.098
Markov	0.083 ± 0.018	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.653 ± 0.182	0.659 ± 0.181	0.738 ± 0.168	0.796 ± 0.179
50 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.020 ± 0.003	0.021 ± 0.003	0.025 ± 0.003	0.038 ± 0.005
Markov	0.007 ± 0.002	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.170 ± 0.048	0.170 ± 0.048	0.195 ± 0.045	0.206 ± 0.049
100 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.002 ± 0.001	0.002 ± 0.001	0.004 ± 0.001	0.004 ± 0.001
Markov	0.005 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov+Harmonic	0.068 ± 0.014	0.068 ± 0.014	0.071 ± 0.014	0.077 ± 0.015

Table 4.4: False negative values for the base scenario and combinations of the proposed techniques in the randomized scenarios

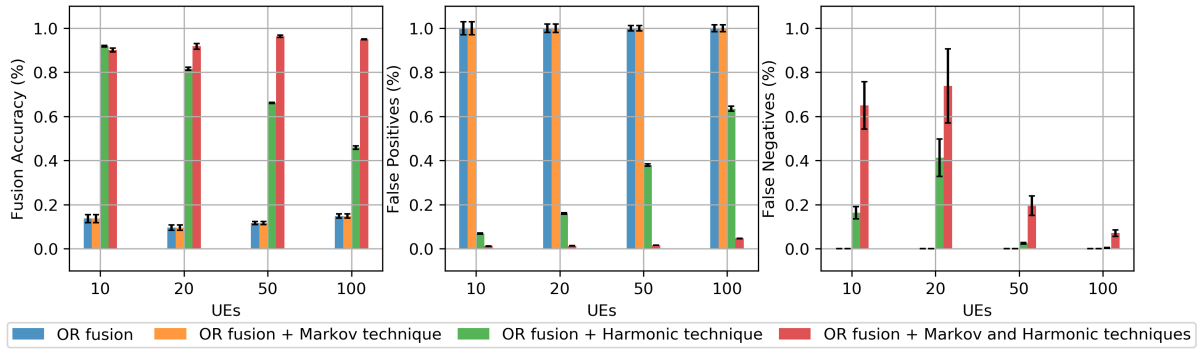
Figure 4.5 depicts the simulation results of the randomized scenarios using multiple fusion techniques: standalone OR, 2-out-of-N, 3-out-of-N and AND. The results are for scenarios with 0, 2, 5 and 10 attackers per simulation, shown respectively in 4.5(a), 4.5(b), 4.5(c) and 4.5(d).



(a) Scenario with no attackers



(b) Scenario with 2 attackers



(c) Scenario with 5 attackers

Figure 4.4: OR fusion results with different numbers of attackers in the randomized scenario

The standalone OR had the best false negatives performance (lower is better) in all scenarios shown in Figure 4.5, as expected. The AND had the best false positives performance (lower is better) in all scenarios, also as expected. However, the Markov-based techniques (ii. and iv.) are, respectively, the best overall performers in terms of accuracy (higher is better) in the scenario without attackers and in scenarios with any tested number of attackers, delivering both protection the the PU with the low false negatives and taking advantage of underused spectrum with the low false positives.

10 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.095 ± 0.003	1.000 ± 0.030	1.000 ± 0.030	1.000 ± 0.030
Harmonic	0.095 ± 0.003	0.091 ± 0.003	0.068 ± 0.002	0.075 ± 0.002
Markov	0.001 ± 0.000	1.000 ± 0.030	1.000 ± 0.030	1.000 ± 0.030
Markov Harmonic	0.001 ± 0.000	0.006 ± 0.000	0.012 ± 0.000	0.066 ± 0.001
20 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.183 ± 0.004	1.000 ± 0.019	1.000 ± 0.019	1.000 ± 0.019
Harmonic	0.183 ± 0.004	0.176 ± 0.004	0.160 ± 0.003	0.117 ± 0.002
Markov	0.001 ± 0.000	1.000 ± 0.019	1.000 ± 0.019	1.000 ± 0.019
Markov Harmonic	0.000 ± 0.000	0.002 ± 0.000	0.013 ± 0.000	0.013 ± 0.000
50 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.396 ± 0.005	1.000 ± 0.012	1.000 ± 0.012	1.000 ± 0.012
Harmonic	0.395 ± 0.005	0.391 ± 0.005	0.380 ± 0.005	0.350 ± 0.005
Markov	0.001 ± 0.000	1.000 ± 0.012	1.000 ± 0.012	1.000 ± 0.012
Markov Harmonic	0.001 ± 0.000	0.002 ± 0.000	0.016 ± 0.000	0.017 ± 0.000
100 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.632 ± 0.011	1.000 ± 0.016	1.000 ± 0.016	1.000 ± 0.016
Harmonic	0.632 ± 0.011	0.631 ± 0.011	0.635 ± 0.011	0.616 ± 0.011
Markov	0.001 ± 0.000	1.000 ± 0.016	1.000 ± 0.016	1.000 ± 0.016
Markov Harmonic	0.001 ± 0.000	0.006 ± 0.000	0.047 ± 0.001	0.047 ± 0.001

Table 4.5: False positive values for the base scenario and combinations of the proposed techniques in the randomized scenarios

4.3.2.2 Clustered scenarios

Figure 4.6 shows the compiled results of the simulations using different fusion techniques combinations with an increasing number of UEs per simulation, as in the randomized scenarios. Each subfigure of Figure 4.6 shows the simulation results for the same scenarios with an increasing number of attackers transmitting fraudulent sensing reports. The fusion techniques used were the same as in the randomized scenarios. The expected difference in these scenarios are the different behavior from the fusion techniques using the Harmonic technique (iii. and iv.), which are dependent on the distribution of the UEs over the cell.

The baseline scenario without attackers is shown in Figure 4.6(a), where the fusion techniques with the Markov technique (ii. and iv.) performed better in terms of accuracy than the others, at the cost of increased false negatives. The fusion accuracy of the standalone OR and the Harmonic technique (i. and iii.) falls as the number of false positives steadily grows as the number of reporting UEs increases, the same with the randomized scenarios.

The Markov technique (ii.) beats all the other techniques in the scenario without

10 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.916 ± 0.002	0.137 ± 0.018	0.137 ± 0.018	0.137 ± 0.018
Harmonic	0.909 ± 0.002	0.911 ± 0.003	0.919 ± 0.003	0.826 ± 0.015
Markov	0.991 ± 0.002	0.137 ± 0.018	0.137 ± 0.018	0.137 ± 0.018
Markov Harmonic	0.931 ± 0.008	0.925 ± 0.008	0.901 ± 0.009	0.820 ± 0.017
20 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.832 ± 0.003	0.095 ± 0.012	0.095 ± 0.012	0.095 ± 0.012
Harmonic	0.800 ± 0.008	0.805 ± 0.008	0.816 ± 0.006	0.847 ± 0.007
Markov	0.991 ± 0.001	0.095 ± 0.012	0.095 ± 0.012	0.095 ± 0.012
Markov Harmonic	0.938 ± 0.015	0.935 ± 0.015	0.918 ± 0.013	0.912 ± 0.014
50 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.650 ± 0.004	0.116 ± 0.007	0.116 ± 0.007	0.116 ± 0.007
Harmonic	0.648 ± 0.004	0.652 ± 0.003	0.661 ± 0.003	0.686 ± 0.003
Markov	0.998 ± 0.000	0.116 ± 0.007	0.116 ± 0.007	0.116 ± 0.007
Markov Harmonic	0.980 ± 0.005	0.978 ± 0.005	0.964 ± 0.005	0.961 ± 0.005
100 UEs, OR Random	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.461 ± 0.007	0.149 ± 0.009	0.149 ± 0.009	0.149 ± 0.009
Harmonic	0.462 ± 0.007	0.463 ± 0.007	0.459 ± 0.007	0.475 ± 0.007
Markov	0.999 ± 0.000	0.149 ± 0.009	0.149 ± 0.009	0.149 ± 0.009
Markov Harmonic	0.989 ± 0.002	0.985 ± 0.002	0.949 ± 0.002	0.949 ± 0.002

Table 4.6: Accuracy values for the base scenario and combinations of the proposed techniques in the randomized scenarios

attackers. In the case with 100 UEs, our technique was able to reduce the false positives to a minimum (from 0.635 ± 0.018 to 0.001 ± 0.000, a 653x reduction), at the cost of increasing false negatives (from 0.0000 ± 0.0000 to 0.148 ± 0.047.

The Harmonic technique (iii.) behaves similarly to the standalone OR (i.), except in terms of false negatives. In the worst simulated scenario, with 50 UEs, the technique had 0.139 ± 0.034 false negatives, 46x the standalone OR (i.).

The combination of both techniques (iv. performs badly in scenarios with no attackers when compared to the standalone Markov technique (ii., even larger false negatives numbers. Differently from what happened in the randomized scenario, the worst simulated scenario in relative to others had 50 UEs, with 123x the standalone OR (i.).

However, false negatives of both Harmonic techniques (iii. and iv.) decrease as the number of UEs increase, providing more spatial information for the fusion. The 0.148 ± 0.047 false negatives in the scenario with 100 UEs is, even though higher than in the randomized scenarios, reasonably low for opportunistic access. To estimate the number of UEs to guarantee a certain level of false negatives, the data was fitted to a decaying exponential curve using the least-squares method. The resulting curve for the clustered

scenario is given by $0.622063e^{-0.0127897x}$, with $R^2 \approx 0.937$. Assuming an interference threshold of up to 0.1 of false negatives, the estimated number of required UEs sensing is of ~ 143 UEs.

Figures 4.6(b) and 4.6(c) show the simulation results for the same scenario, but with 2 and 5 attackers within the cell. Both the standalone OR and the Markov technique (i. and ii.) perform badly in this scenario, fusing the fake results, that increases false positives to as high as 100%. Conversely, both techniques using the Harmonic mean (iii. and iv. offer some protection against the attacks for large number of UEs, behaving similarly to the scenario without attackers, shown in Figure 4.6(a).

The values plotted in the Figure 4.6 are tabled in Tables 4.9, 4.8 and 4.7.

10 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.042 ± 0.016	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.059 ± 0.019	0.069 ± 0.019	0.131 ± 0.026	0.000 ± 0.000
Markov	0.413 ± 0.130	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.559 ± 0.127	0.581 ± 0.129	0.732 ± 0.155	0.000 ± 0.000
20 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.035 ± 0.015	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.130 ± 0.035	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov	0.303 ± 0.098	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.449 ± 0.154	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
50 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.003 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.139 ± 0.034	0.144 ± 0.034	0.157 ± 0.036	0.174 ± 0.036
Markov	0.234 ± 0.125	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov Harmonic	0.370 ± 0.103	0.372 ± 0.104	0.392 ± 0.102	0.401 ± 0.101
100 UEs, OR	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Harmonic	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.001	0.001 ± 0.001
Markov	0.074 ± 0.042	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Markov+Harmonic	0.148 ± 0.047	0.148 ± 0.047	0.152 ± 0.049	0.162 ± 0.052

Table 4.7: False negative values for the base scenario and combinations of the proposed techniques in the clustered scenarios

Figure 4.7 depicts the simulation results of the clustered scenarios using multiple fusion techniques: standalone OR, 2-out-of-N, 3-out-of-N and AND. The results are for scenarios with 0, 2, 5 and 10 attackers per simulation, shown respectively in 4.7(a), 4.7(b), 4.7(c) and 4.7(d).

The standalone OR had the best false negatives performance (lower is better) in all scenarios shown in Figure 4.7, as expected due to the nature of the technique. The AND had the best false positives performance (lower is better) in all scenarios, also as expected. However, the Markov-based techniques (ii. and iv.) are, respectively, the best overall performers in terms of accuracy (higher is better) in the scenario without attackers

10 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.095 ± 0.002	1.000 ± 0.026	1.000 ± 0.026	1.000 ± 0.026
Harmonic	0.095 ± 0.002	0.088 ± 0.002	0.064 ± 0.001	1.000 ± 0.026
Markov	0.000 ± 0.000	1.000 ± 0.026	1.000 ± 0.026	1.000 ± 0.026
Markov Harmonic	0.000 ± 0.000	0.003 ± 0.000	0.006 ± 0.000	1.000 ± 0.026
20 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.183 ± 0.002	1.000 ± 0.013	1.000 ± 0.013	1.000 ± 0.013
Harmonic	0.183 ± 0.002	1.000 ± 0.013	1.000 ± 0.013	1.000 ± 0.013
Markov	0.000 ± 0.000	1.000 ± 0.013	1.000 ± 0.013	1.000 ± 0.013
Markov Harmonic	0.000 ± 0.000	1.000 ± 0.013	1.000 ± 0.013	1.000 ± 0.013
50 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.396 ± 0.005	1.000 ± 0.014	1.000 ± 0.014	1.000 ± 0.014
Harmonic	0.396 ± 0.005	0.391 ± 0.005	0.374 ± 0.005	0.346 ± 0.005
Markov	0.001 ± 0.000	1.000 ± 0.014	1.000 ± 0.014	1.000 ± 0.014
Markov Harmonic	0.000 ± 0.000	0.002 ± 0.000	0.004 ± 0.000	0.010 ± 0.000
100 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.635 ± 0.018	1.000 ± 0.027	1.000 ± 0.027	1.000 ± 0.027
Harmonic	0.634 ± 0.018	0.642 ± 0.018	0.631 ± 0.018	0.611 ± 0.017
Markov	0.001 ± 0.000	1.000 ± 0.027	1.000 ± 0.027	1.000 ± 0.027
Markov Harmonic	0.001 ± 0.000	0.031 ± 0.001	0.031 ± 0.001	0.031 ± 0.001

Table 4.8: False positive values for the base scenario and combinations of the proposed techniques in the clustered scenarios

and in scenarios with any tested number of attackers, delivering both protection the the PU with the low false negatives and taking advantage of underused spectrum with the low false positives. However, the number of required sensing UEs is $\sim 60\%$ higher than in the randomized scenario to achieve the same false negative results, based on the fitted curves.

Figure 4.7 shows the comparison of simulation results with additional fusion techniques. The effect of the Harmonic-mean based and the Markov chain can be seen very clearly on Figures 4.7(c), where only the combination of the techniques and the OR mitigated the attacker's action, except for the scenario with 20 UEs grouped into clusters where at least one of the attackers was closer to the PU than legitimate users and the filtering did not work as intended.

4.3.2.3 Collaborative spectrum sensing report overhead

Finally, for all the scenarios, the Markov-based technique effectively reduced the amount of reports by 2 orders of magnitude, as shown in Figure 4.8, which presents the mean percentage of reported frames per UE for all the simulations. The amount of reporting

10 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.910 ± 0.002	0.102 ± 0.017	0.102 ± 0.017	0.102 ± 0.017
Harmonic	0.908 ± 0.002	0.914 ± 0.002	0.929 ± 0.002	0.102 ± 0.017
Markov	0.958 ± 0.011	0.102 ± 0.017	0.102 ± 0.017	0.102 ± 0.017
Markov Harmonic	0.943 ± 0.009	0.938 ± 0.009	0.919 ± 0.010	0.102 ± 0.017
20 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.830 ± 0.002	0.086 ± 0.008	0.086 ± 0.008	0.086 ± 0.008
Harmonic	0.822 ± 0.003	0.086 ± 0.008	0.086 ± 0.008	0.086 ± 0.008
Markov	0.974 ± 0.008	0.086 ± 0.008	0.086 ± 0.008	0.086 ± 0.008
Markov Harmonic	0.961 ± 0.013	0.086 ± 0.008	0.086 ± 0.008	0.086 ± 0.008
50 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.669 ± 0.003	0.167 ± 0.008	0.167 ± 0.008	0.167 ± 0.008
Harmonic	0.647 ± 0.006	0.650 ± 0.006	0.662 ± 0.007	0.683 ± 0.006
Markov	0.960 ± 0.021	0.167 ± 0.008	0.167 ± 0.008	0.167 ± 0.008
Markov Harmonic	0.938 ± 0.017	0.936 ± 0.017	0.932 ± 0.017	0.925 ± 0.017
100 UEs, OR Cluster	no Attackers	02 Attackers	05 Attackers	10 Attackers
Standard	0.461 ± 0.011	0.152 ± 0.016	0.152 ± 0.016	0.152 ± 0.016
Harmonic	0.462 ± 0.011	0.455 ± 0.011	0.465 ± 0.011	0.481 ± 0.011
Markov	0.988 ± 0.006	0.152 ± 0.016	0.152 ± 0.016	0.152 ± 0.016
Markov Harmonic	0.977 ± 0.007	0.952 ± 0.007	0.951 ± 0.007	0.950 ± 0.007

Table 4.9: Accuracy values for the base scenario and combinations of the proposed techniques in the clustered scenarios

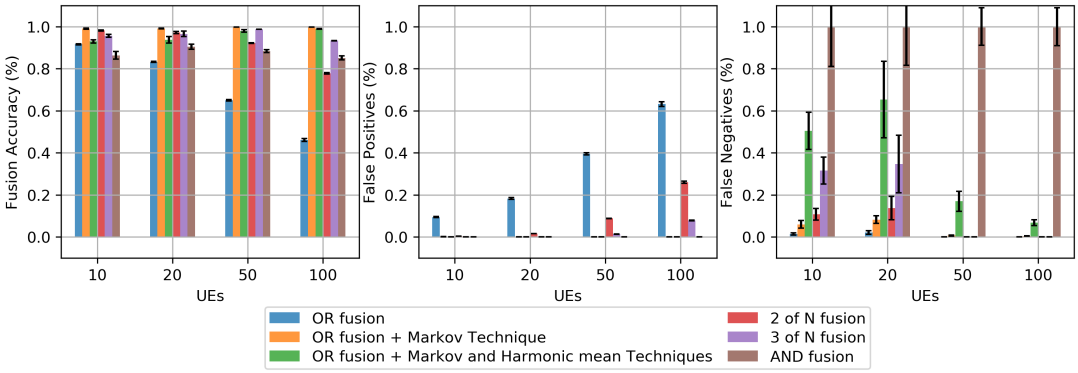
can be further reducing using sampling techniques or using a hybrid fusion approach with both decentralized fusions using a side-channel intraclusters and the centralized fusion using the control-channel interclusters. This reduction shows the effectiveness of the technique in reducing scarce control bandwidth, enabling more nodes to be connected to the same cell and reducing licensing costs.

4.4 Chapter review

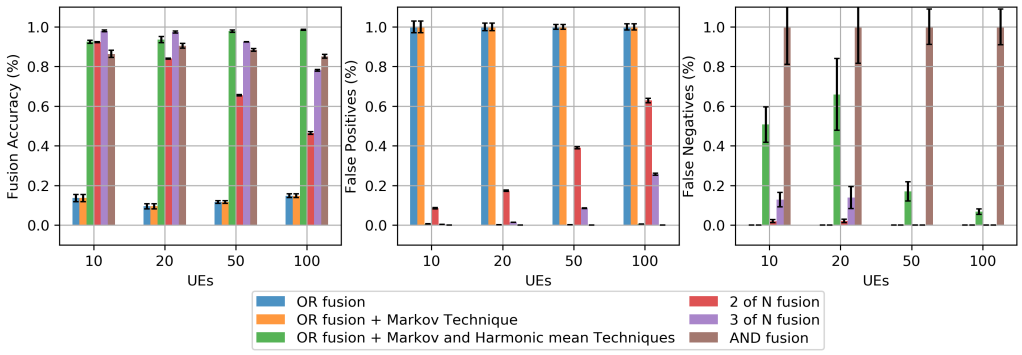
The chapter presented the tools, methods, performance metrics, and the simulations scenarios for validation of the CSS cycle and evaluation of the proposed techniques, along with the results for these scenarios. The implemented CSS cycle results are on par with the expected theoretical results. The proposed techniques show their strength and flaws in different scenarios. The Markov chain-based technique drastically reduces false positives but is not ineffective in the case jammers or attackers send false reports to the fusion center. The harmonic mean-based technique shows equivalent results to the OR fusion in

a scenario without attackers but is very effective in reducing the number of false positives in scenarios with attackers.

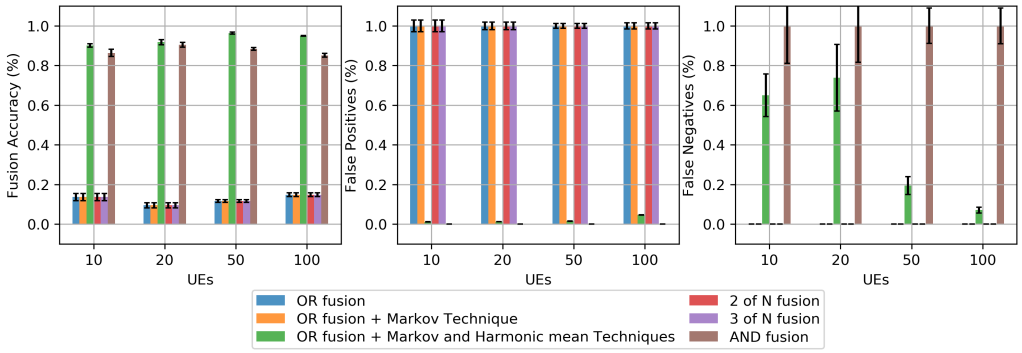
The combination of both techniques behaves similarly to the Markov chain-based one in a scenario without attackers but completely different in the scenario with attackers, where it drastically reduces the false positives at the cost of sharply increasing false negatives. As the number of sensing UEs grow, the number of false negatives drops exponentially.



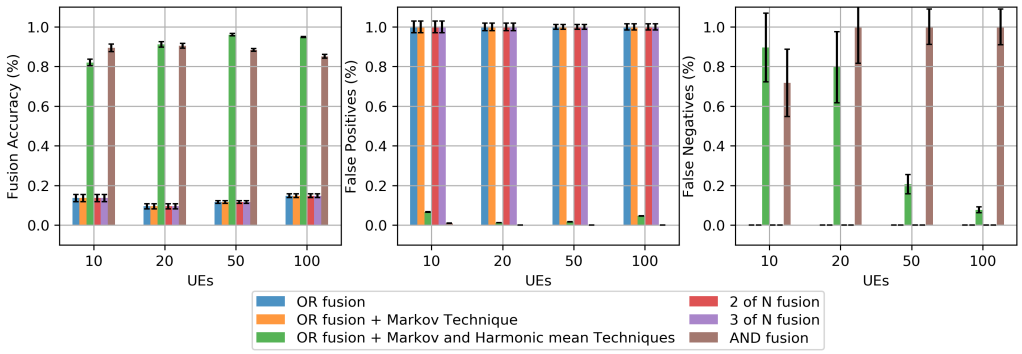
(a) Scenario with no attackers



(b) Scenario with 2 attackers

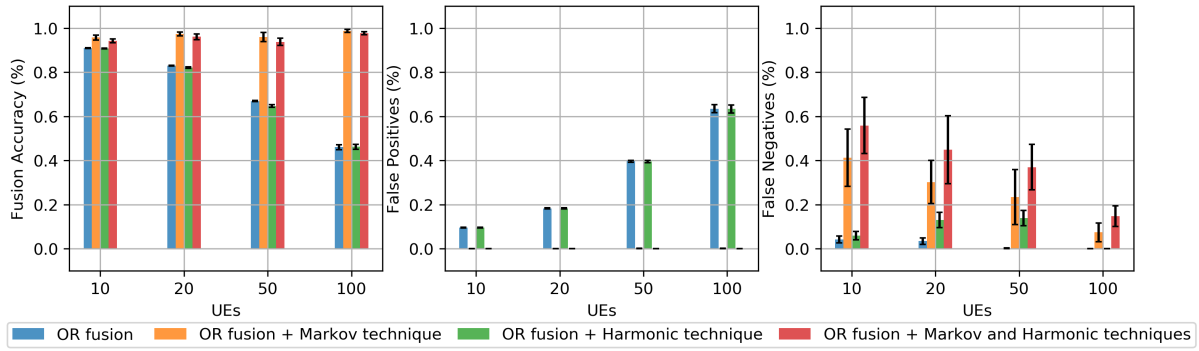


(c) Scenario with 5 attackers

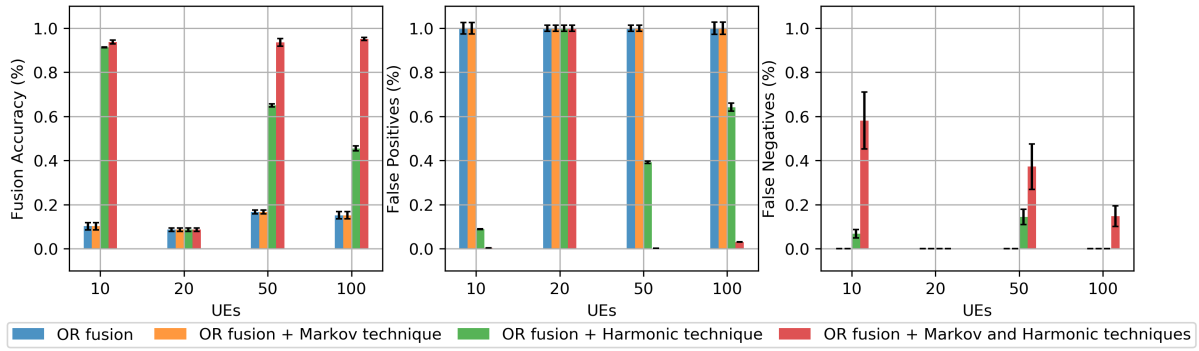


(d) Scenario with 10 attackers

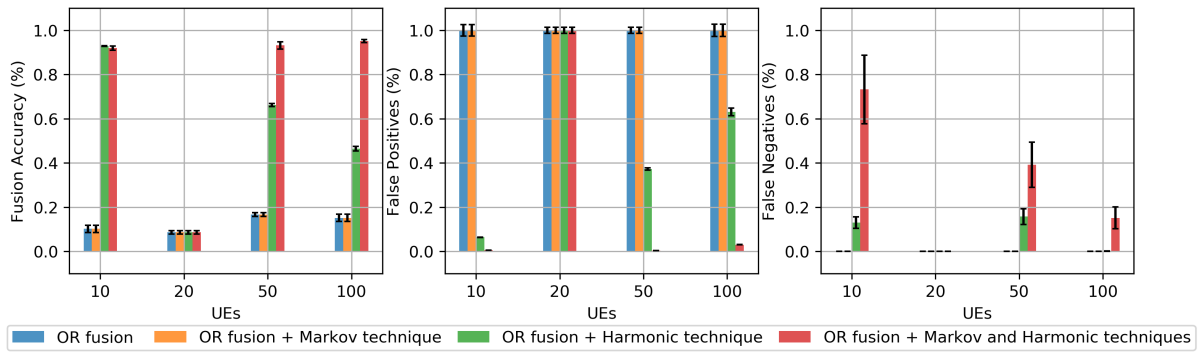
Figure 4.5: Fusion results of different fusion techniques in the randomized scenario.



(a) Scenario with no attackers

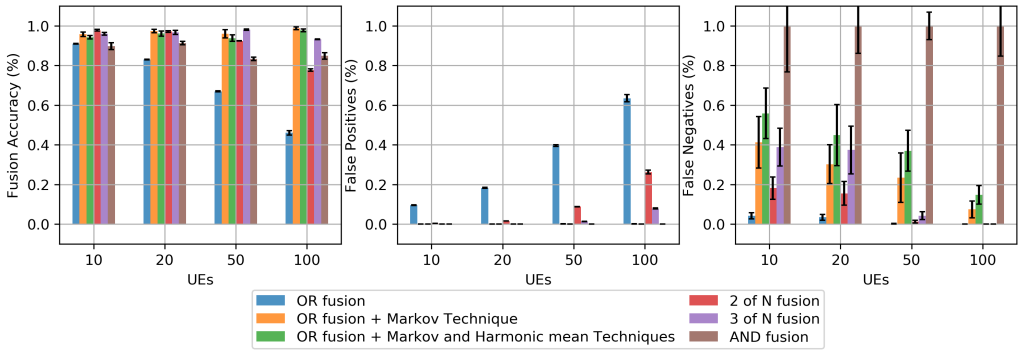


(b) Scenario with 2 attackers

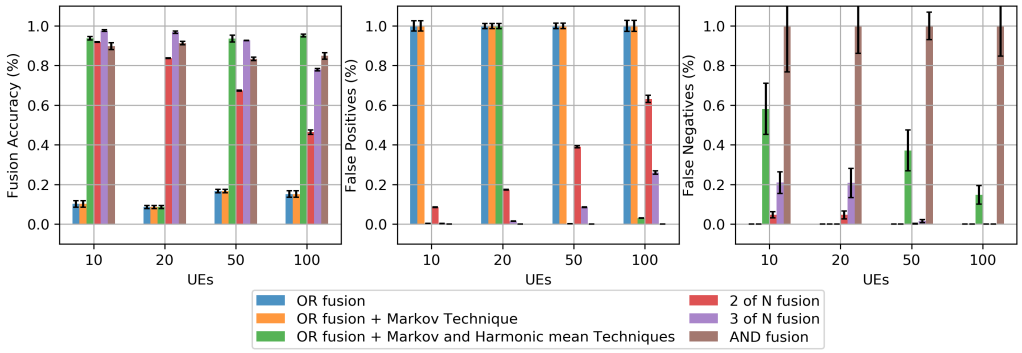


(c) Scenario with 5 attackers

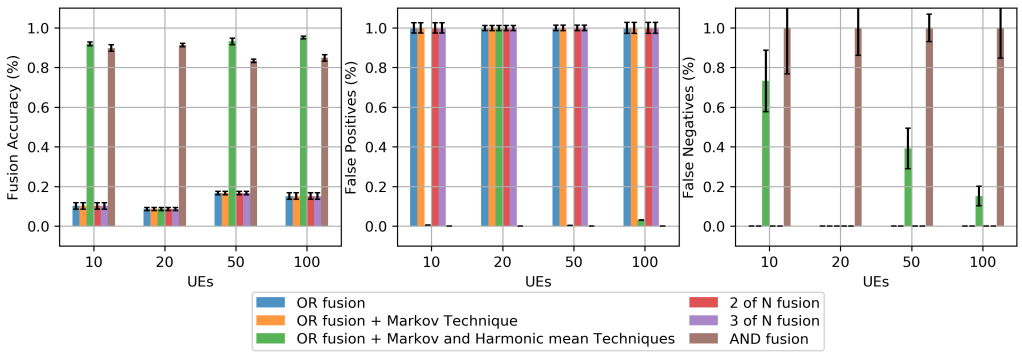
Figure 4.6: OR fusion results with different numbers of attackers in the clustered scenario



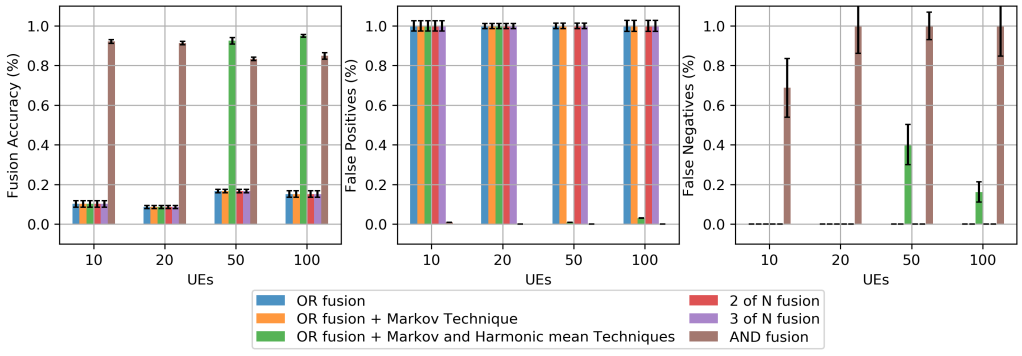
(a) Scenario with no attackers



(b) Scenario with 2 attackers



(c) Scenario with 5 attackers



(d) Scenario with 10 attackers

Figure 4.7: Fusion results of different fusion techniques in the clustered scenario.

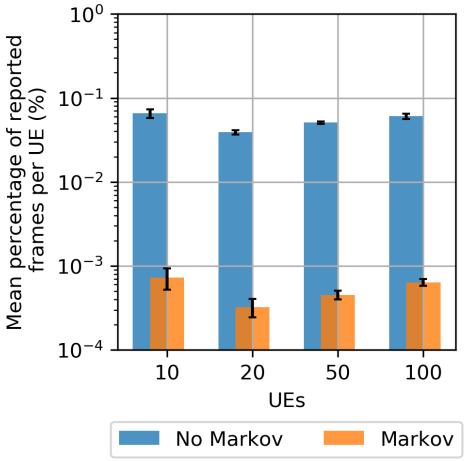


Figure 4.8: Sensing report overhead with and without the proposed Markov chain based mechanism

Chapter 5

Conclusions and Future Work

This dissertation contributions are two techniques to increase the efficiency of collaborative spectrum sensing and dynamic spectrum access in remote areas, in addition to the tooling required to simulate and evaluate future research on the same topic. Even though the techniques might not be as good as more sophisticated techniques, such as machine learning, they require no training, nor additional message exchanges, while keeping the computational costs low.

First, a review of the evolution of mobile networks, their current state and architecture, followed by different medium access control techniques, focused on centralized MAC and OFDMA used in the LTE standard were presented. The resource allocation matrices and scheduling techniques were also reviewed, followed by both individual and collaborative spectrum sensing techniques, which includes the fusion techniques. Statistical tools such as a central limit theorem, Markov chains and Monte Carlo techniques were also are reviewed, for later use at both the proposal itself and for its evaluation with simulations. Related works were reviewed along with their proposals, results and limitations.

Then, two techniques to improve the results of collaborative sensing are proposed: (1) a technique based on Markov chains to smooth the results of individual sensing, reducing numbers of false positives and false negatives, in addition to reducing the sending of control messages with equal sensing results consecutively; (2) a technique based on the harmonic mean for filtering the results of individual sensing, discarding sensing of nodes that are more distant from the sources of interference, protecting from Byzantine attacks. Both techniques are evaluated in rural 5G scenarios, in which there is the largest portion of underused bands, candidates for opportunistic access.

A comparison of the proposed techniques shows that the proposed Markov chain technique reduces the amount of control message exchange containing individual spectrum sensing reports by two orders of magnitude, saving limited bandwidth in the licensed control channel. The miss detections probability and the resulting interference was kept

low in all simulated scenarios, except in the presence of a malicious users, which can severely disturb the opportunistic use of the spectrum. The Harmonic mean-based filtering technique nullified false positive reports, while increasing miss detections, which do cause interference. The amount of miss detections falls sharply as the number of users, in the randomized scenario, or the number of clusters, in the clustered scenario, increases.

Both techniques focus in 5G for rural areas. For that, we implemented and validated a collaborative spectrum sensing cycle on top of the LTE stack of the NS-3 simulator. The implementations include procedures for sensing the individual spectrum performed by user devices (UEs), transmitting the results to the access point (eNodeB), merging the received results and using the merging result in scaling resources for the devices. The individual sensing reports are made from curves of probability of detection and probability of false positives obtained through measurements in experiments or through simulations at the physical link layer. The simulations results match their mathematical model counterparts with minimal fluctuations and the adaptations in the simulator resulted in a new tool, called COLAB that may be useful for future simulations and adaptations in 5G research.

As future work, we intend to address the following topics:

- simulate scenarios with a higher number of users in a denser environment by using sampling techniques for the CSS.
- break up the cell into sectors and use Angle-of-Arrival (AoA) or Time-of-Flight (ToF) to locate the general region of the UEs and use as additional information for the fusion or hard combining.
- reduce miss detections by improving the Harmonic mean-based techniques.

Bibliography

[1] Minges, Michael and Pratikshya Simkhada: *The Evolution To 3G Mobile — Status Report.* https://www.itu.int/itunews/issue/2003/06/thirdgeneration.html, [Online; accessed 19-February-2018]. xv, 1, 5, 6

[2] Ericsson: *Ericsson mobility report.* https://www.ericsson.com/assets/local/mobility-report/documents/2017/ericsson-mobility-report-november-2017.pdf. xv, 7

[3] Kitana, Asem, Issa Traore, and Isaac Woungang: *Impact study of a mobile botnet over lte networks.* J. Internet Serv. Inf. Secur., 6(2):1–22, 2016. xv, 8

[4] Keysight: *Lte physical layer overview.* http://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/lte/content/lte_overview.htm. xv, 9, 10

[5] NETMANIAS: *LTE QoS: SDF and EPS Bearer QoS.* https://www.netmanias.com/en/?m=view&id=techdocs&no=5908&xtag=eps-lte-qos-sdf&xref=lte-qos-sdf-and-eps-bearer-qos, [Online; accessed 2-March-2018]. xv, 9, 11

[6] 3GPP: *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 15).* https://www.3gpp.org/ftp/Specs/archive/23_series/23.203/23203-f50.zip, [Online; accessed 28-February-2020]. xv, 9, 11, 12

[7] ShareTechnote: *DCI.* https://www.sharetechnote.com/html/DCI.html#36_213_8_1, [Online; accessed 29-February-2020]. xv, 13, 14

[8] Cichon, Krzysztof, Adrian Kliks, and Hanna Bogucka: *Energy-efficient cooperative spectrum sensing: A survey.* IEEE Communications Surveys & Tutorials, 18(3):1861–1886, 2016. xv, 3, 18, 19, 28, 38

[9] Jaglan, Reena Rathee, Rashid Mustafa, and Sunil Agrawal: *Scalable and robust ANN based cooperative spectrum sensing for cognitive radio networks.* Wireless Personal Communications, 99(3):1141–1157, jan 2018. xv, 3, 24, 25

[10] Thilina, Karaputugala Madushan, Kae Won Choi, Nazmus Saquib, and Ekram Hossain: *Pattern classification techniques for cooperative spectrum sensing in cognitive radio networks: SVM and w-KNN approaches.* In *2012 IEEE Global Communications Conference (GLOBECOM).* IEEE, dec 2012. xvi, 25, 26

[11] Anatel: *Plano de atribuição, distribuição e destinação de radiofrequências.* https://anatel.gov.br/setorregulado/atribuicao-destinacao-e-distribuicao-de-faixas. xvi, 30

[12] Nikookar, Homayoun: *Wavelets for spectrum sensing in cognitive radio applications.* In *Wavelet Radio*, pages 112–138. Cambridge University Press. xvi, 31

[13] Haykin, S.: *Cognitive radio: brain-empowered wireless communications.* IEEE Journal on Selected Areas in Communications, 23(2):201–220, feb 2005. xvi, 32

[14] Privault, Nicolas: *Discrete-time markov chains.* In *Springer Undergraduate Mathematics Series*, pages 89–113. Springer Singapore, 2018. xvi, 35

[15] Foukalas, F., 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*, pages 1–5, Aug 2013. xvi, 36, 37

[16] Meeker, Mary: *Internet trends 2018.* https://www.kleinerperkins.com/files/INTERNET_TRENDS_REPORT_2018.pdf. 1

[17] McAfee: *Navigating a cloudy sky.* https://www.mcafee.com/enterprise/en-us/solutions/lp/cloud-security-report.html. 1

[18] Abdelwahab, Sherif, Bechir Hamdaoui, Mohsen Guizani, and Taieb Znati: *Network function virtualization in 5g.* IEEE Communications Magazine, 54(4):84–91, apr 2016. 1

[19] Chourbaji, Wassim: *5g spectrum access.* https://www.itu.int/dms_pub/itu-r/oth/0a/0E/R0A0E0000C60001PDFE.pdf, [Online; accessed 28-February-2020]. 1, 7, 29

[20] Valenta, Vaclav, Zbynek Fedra, Roman Marsalek, Genevieve Baudoin, and Martine Villegas: *Towards cognitive radio networks: Spectrum utilization measurements in suburb environment.* In *2009 IEEE Radio and Wireless Symposium.* IEEE, jan 2009. 2, 29

[21] Beckman, Richard, Karthik Channakeshava, Fei Huang, V. S. Anil Vullikanti, Achla Marathe, Madhav V. Marathe, and Guanhong Pei: *Implications of dynamic spectrum access on the efficiency of primary wireless market.* In *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN).* IEEE, apr 2010. 2, 29

[22] Hu, F., B. Chen, and K. Zhu: *Full spectrum sharing in cognitive radio networks toward 5g: A survey.* IEEE Access, 6:15754–15776, 2018, ISSN 2169-3536. 2, 29, 31

[23] Leaves, P.: *Dynamic spectrum allocation in a multi-radio environment: concept and algorithm.* In *Second International Conference on 3G Mobile Communication Technologies (3G 2001).* IEE, 2001. 2, 29, 31

[24] Song, Min, Chunsheng Xin, Yanxiao Zhao, and Xiuzhen Cheng: *Dynamic spectrum access: from cognitive radio to network radio.* IEEE Wireless Communications, 19(1):23–29, feb 2012. 2, 31

[25] Hoffmann, H., P. Ramachandra, I. Z. Kovács, L. Jorguseski, F. Gunnarsson, and T. Kürner: *Potential of dynamic spectrum allocation in LTE macro networks.* Advances in Radio Science, 13:95–102, nov 2015. 2, 31

[26] Kumbhkar, Ratnesh: *Opportunistic access of noncontiguous spectrum.* PhD thesis, Rutgers University, jan 2018. `https://rucore.libraries.rutgers.edu/rutgers-lib/56032/`. 2, 31

[27] Basnet, Shubhekshya, Beeshanga Abewardana Jayawickrama, Ying He, Eryk Dutkiewicz, and Markus Dominik 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. 2, 29

[28] Arafat, Abdullah Omar, Akram Al-Hourani, Nazmus S. Nafi, and Mark A. Gregory: *A survey on dynamic spectrum access for LTE-advanced.* Wireless Personal Communications, 97(3):3921–3941, aug 2017. 2, 31

[29] Szydelko, Michal and Marcin Dryjanski: *3gpp spectrum access evolution towards 5g.* EAI Endorsed Transactions on Cognitive Communications, 3(10):152184, feb 2017. 2, 31

[30] Arjoune, Youness and Naima Kaabouch: *A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions.* Sensors, 19(1):126, jan 2019. 2, 3

[31] Akyildiz, Ian F., Brandon F. Lo, and Ravikumar Balakrishnan: *Cooperative spectrum sensing in cognitive radio networks: A survey.* Physical Communication, 4(1):40–62, mar 2011. 3, 18, 19, 32

[32] Bhatnagar, Charu, Anjali Potnis, Prshant Dwivedy, and Sunil Kumar Meena: *Performance analysis and optimization schemes for cooperative spectrum sensing and information fusion for cognitive radio : A survey.* In *2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech)*. IEEE, apr 2017. 3, 22, 23

[33] Mustafa, Rashid, Reena Rathee Jaglan, and Sunil Agrawal: *Performance evaluation of cooperative spectrum sensing over fading channels based on neural network learning approach.* In *Proceedings of 2nd International Conference on Communication, Computing and Networking*, pages 567–575. Springer Singapore, sep 2018. 3, 24

[34] Lee, Woongsup, Minhoe Kim, and Dong Ho Cho: *Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks.* IEEE Transactions on Vehicular Technology, 68(3):3005–3009, mar 2019. 3, 25, 27

[35] Giri, Manish Kumar and Saikat Majumder: *Extreme learning machine based cooperative spectrum sensing in cognitive radio networks.* In *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, feb 2020. 3, 25

[36] Cai, Yifeng, Yijun Mo, Kaoru Ota, Changqing Luo, Mianxiong Dong, and Laurence Yang: *Optimal data fusion of collaborative spectrum sensing under attack in cognitive radio networks.* IEEE Network, 28(1):17–23, jan 2014. 3, 28, 39

[37] Li, Lei and Chunxiao Chigan: *Fuzzy c-means clustering based secure fusion strategy in collaborative spectrum sensing.* In *2014 IEEE International Conference on Communications (ICC)*. IEEE, jun 2014. 3, 39

[38] Wu, Jun, Tiecheng Song, Yue Yu, Cong Wang, and Jing Hu: *Sequential cooperative spectrum sensing in the presence of dynamic byzantine attack for mobile networks.* PLOS ONE, 13(7):e0199546, jul 2018. 3, 28, 39, 40

[39] Zhang, Linyuan, Guangming Nie, Guoru Ding, Qihui Wu, Zhaoyang Zhang, and Zhu Han: *Byzantine attacker identification in collaborative spectrum sensing: A robust defense framework.* IEEE Transactions on Mobile Computing, 18(9):1992–2004, sep 2019. 3, 28, 40

[40] Pelechrinis, Konstantinos, Prashant Krishanmurthy, and Christos Gkantsidis: *Trustworthy operations in cellular networks: The case of PF scheduler.* IEEE Transactions on Parallel and Distributed Systems, 25(2):292–300, feb 2014. 3, 39, 47

[41] Catalunya, Centre Tecnològic de Telecomunicacions de: *The LENA ns-3 LTE Module Documentation.* `http://networks.cttc.es/wp-content/uploads/sites/2/2014/01/lena-lte-module-doc.pdf`, [Online; accessed 28-February-2018]. 4, 49

[42] Aldmour, Ismat: *LTE and WiMAX: Comparison and future perspective.* Communications and Network, 05(04):360–368, 2013. 5, 6

[43] 3GPP: *LTE.* `https://www.3gpp.org/technologies/keywords-acronyms/98-lte`, [Online; accessed in 28-February-2020]. 6

[44] Palat, Sudeep and Philippe Godin: *The LTE Network Architecture: A comprehensive tutorial.* `http://www.cse.unt.edu/~rdantu/FALL_2013_WIRELESS_NETWORKS/LTE_Alcatel_White_Paper.pdf`, [Online; accessed 21-February-2018]. 6, 8, 9, 11, 13

[45] Firmin, Frédéric: *The Evolved Packet Core.* `https://www.3gpp.org/technologies/keywords-acronyms/100-the-evolved-packet-core`, [Online; accessed 5-May-2019]. 6

[46] 3GPP: *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2.* `http://www.3gpp.org/ftp//Specs/archive/36_series/36.300/36300-f00.zip`, [Online; accessed 1-March-2018]. 6

[47] Erik Dahlman, Stefan Parkvall, Johan Skold: *4G LTE-Advanced Pro and The Road to 5G.* Academic Press, 3rd edition, 2016, ISBN 0128045752,978-0-12-804575-6,9780128046111,0128046112. 6, 8, 9

[48] Bjorkdahl, J, Erik Bohlin, and Sven Lindmark: *Financial assessment of fourth generation mobile technologies.* COMMUNICATIONS AND STRATEGIES., pages 71–96, 2004. 7

[49] Union, International Telecommunication: *The state of broadband: Broadband catalyzing sustainable development.* `https://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-BROADBAND.19-2018-PDF-E.pdf`. 7

[50] GSM Association: *An introduction to network slicing*, nov 2017. `https://www.gsma.com/futurenetworks/wp-content/uploads/2017/11/GSMA-An-Introduction-to-Network-Slicing.pdf`. 7

[51] Ordonez-Lucena, Jose, Pablo Ameigeiras, Diego Lopez, Juan J. Ramos-Munoz, Javier Lorca, and Jesus Folgueira: *Network slicing for 5g with SDN/NFV: Concepts, architectures, and challenges.* IEEE Communications Magazine, 55(5):80–87, may 2017. 7

[52] Katsalis, Kostas, Navid Nikaein, Eryk Schiller, Adlen Ksentini, and Torsten Braun: *Network slices toward 5g communications: Slicing the LTE network.* IEEE Communications Magazine, 55(8):146–154, aug 2017. 7

[53] 5G PPP Architecture Working Group: *View on 5g architecture.* `https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf`. 8, 11

[54] *IEEE standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications.* 8

[55] 3GPP: *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Multiplexing and chennal coding (Release 15).* `https://www.3gpp.org/ftp/Specs/archive/36_series/36.212/36212-f80.zip`, [Online; accessed 28-February-2020]. 9, 13

[56] Nsiri, Bechir, Mallouki Nasreddine, Mahmoud Ammar, Walid Hakimi, and Mhatli Sofien: *Performance comparison of scheduling algorithms for downlink lte system.* 2014. 9

[57] 5G Americas: *5g services and use cases.* `http://www.5gamericas.org/wp-content/uploads/2019/07/5G_Service_and_Use_Cases__FINAL.pdf`, [Online; accessed 28-February-2020]. 9, 11

[58] Rao, S.P.V Subba, S Venkata Chalam, and D Sreenivasa Rao: *A survey on mac protocols for wireless multimedia networks.* International Journal of Computer Science & Engineering Survey, 2(4):57–74, nov 2011. 11, 13

[59] Dwijaksara, Made Harta, Wha Sook Jeon, and Dong Geun Jeong: *A channel access scheme for bluetooth low energy to support delay-sensitive applications.* In *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, sep 2016. 13

[60] Voicu, Andra M., Ljiljana Simic, J. Pierre de Vries, Marina Petrova, and Petri Mahonen: *Analysing wi-fi/LTE coexistence to demonstrate the value of risk-informed interference assessment.* In *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, mar 2017. 13

[61] Bae, Sunghwan and Hongseok Kim: *Towards harmonious coexistence in the unlicensed spectrum: Rational cooperation of operators.* Sensors, 17(10):2432, oct 2017. 13

[62] Kleinrock, Leonard and Fouad Tobagi: *Random access techniques for data transmission over packet-switched radio channels.* In *Proceedings of the May 19-22, 1975, national computer conference and exposition on - AFIPS '75*. ACM Press, 1975. 14

[63] Lam, Simon S: *A carrier sense multiple access protocol for local networks.* Computer Networks (1976), 4(1):21–32, feb 1980. 14

[64] Ramachandran, Iyappan and Sumit Roy: *WLC46-2: On the impact of clear channel assessment on MAC performance.* In *IEEE Globecom 2006*. IEEE, nov 2006. 15, 16

[65] Saarnisaari, Harri and Johanna Vartiainen: *Spectrum window based signal detection at low SNR.* In *2018 International Conference on Military Communications and Information Systems (ICMCIS)*. IEEE, may 2018. 15, 17, 56, 58, 61

[66] Vartiainen, Johanna, Marja Matinmikko-Blue, Heikki Karvonen, Luciano Mendes, Alexandre Matos, and Carlos Silva: *Performance of WIBA energy detector in rural and remote area channel.* In *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, aug 2019. 15

[67] Hossain, E. and K.G.Madushan Thilina: *Cognitive radio networks and spectrum sharing.* In *Academic Press Library in Mobile and Wireless Communications*, pages 467–522. Elsevier, 2016. 18, 20, 22, 24

[68] Teguig, Djamel, B Scheers, and V Le Nir: *Data fusion schemes for cooperative spectrum sensing in cognitive radio networks.* In *2012 Military Communications and Information Systems Conference (MCC)*, pages 1–7. IEEE, 2012. 20, 28

[69] Alvi, Sheeraz Akhtar, Muhammad Shahzad Younis, Muhammad Imran, and Fazal e Amin: *A weighted linear combining scheme for cooperative spectrum sensing.* Procedia Computer Science, 32:149–157, 2014. 20, 28

[70] Fu, Yuanhua, Fan Yang, and Zhiming He: *A quantization-based multibit data fusion scheme for cooperative spectrum sensing in cognitive radio networks.* Sensors, 18(2):473, feb 2018. 20, 28

[71] Maleki, Sina, Sundeep Prabhakar Chepuri, and Geert Leus: *Optimization of hard fusion based spectrum sensing for energy-constrained cognitive radio networks*. Physical Communication, 9:193–198, dec 2013. 28

[72] Tan, Le Thanh and Long Bao Le: *Joint cooperative spectrum sensing and MAC protocol design for multi-channel cognitive radio networks*. EURASIP Journal on Wireless Communications and Networking, 2014(1), jun 2014. 28, 32, 37

[73] Liu, Jian, Lu Yang, Peng Zhang, Zhizhong Zhang, and Yichuan Zheng: *Low-overhead cooperative spectrum sensing technology for cognitive radio networks*. In *International Conference on Cyberspace Technology (CCT 2014)*. Institution of Engineering and Technology, 2014. 28

[74] Meenu Rani, S. B. Dhok and R. B. Deshmukh: *A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications*. IEEE Spectrum, pages 4875 – 4894, 2018. 28

[75] Lamport, Leslie, Robert Shostak, and Marshall Pease: *The byzantine generals problem*. ACM Transactions on Programming Languages and Systems (TOPLAS), 4(3):382–401, jul 1982. 28

[76] Vempaty, Aditya, Lang Tong, and Pramod K. Varshney: *Distributed inference with byzantine data: State-of-the-art review on data falsification attacks*. IEEE Signal Processing Magazine, 30(5):65–75, sep 2013. 28

[77] Zhang, Linyuan, Guoru Ding, Qihui Wu, Yulong Zou, Zhu Han, and Jinlong Wang: *Byzantine attack and defense in cognitive radio networks: A survey*. IEEE Communications Surveys & Tutorials, 17(3):1342–1363, 2015. 28

[78] Federal Communications Commission: *Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies*, dec 2003. `https://apps.fcc.gov/edocs_public/attachmatch/FCC-03-322A1.pdf`. 31

[79] Hong, Xuemin, Cheng Xiang Wang, John Thompson, and Yan Zhang: *Demystifying white spaces*. In *2008 International Conference on Communications, Circuits and Systems*. IEEE, may 2008. 31

[80] Harrison, Kate, Shridhar Mubaraq Mishra, and Anant Sahai: *How much white-space capacity is there?* In *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN)*. IEEE, apr 2010. 31

[81] CWSC: *What is white space?* `https://www.wirelesswhitespace.org/more/what-is-white-space/`. 31

[82] GSMA: *Digital Dividend: Laying the Foundations for Expanded Mobile Service*. `https://www.gsma.com/spectrum/wp-content/uploads/2013/09/dd_qa_rancy_09_13-3.pdf`, [Online; accessed 29-February-2020]. 31

[83] Roberts, Sidney, Paul Garnett, and Ranveer Chandra: *Connecting africa using the TV white spaces: from research to real world deployments*. In *The 21st IEEE International Workshop on Local and Metropolitan Area Networks*. IEEE, apr 2015. 31

[84] Saifullah, Abusayeed, Mahbubur Rahman, Dali Ismail, Chenyang Lu, Ranveer Chandra, and Jie Liu: *SNOW*. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems - SenSys '16*. ACM Press, 2016. 31

[85] Yau, Kok Lim Alvin, Junaid Qadir, Celimuge Wu, Muhammad Ali Imran, and Mee Hong Ling: *Cognition-Inspired 5G Cellular Networks: A Review and the Road Ahead*. IEEE Access, 6:35072 – 35090, 2018. 31

[86] Clancy, Charles, Joe Hecker, Erich Stuntebeck, and Tim O'Shea: *Applications of machine learning to cognitive radio networks*. IEEE Wireless Communications, 14(4):47–52, aug 2007. 31, 32

[87] Wang, Dan, Bin Song, Dong Chen, and Xiaojiang Du: *Intelligent Cognitive Radio in 5G: AI-Based Hierarchical Cognitive Cellular Networks*. IEEE Wireless Communications, 26:54 – 61, 2019. 32

[88] Mfupe, Luzango, Listietsi Montsi, Mjumo Mzyece, and Fisseha Mekuria: *Enabling dynamic spectrum access through location aware spectrum databases*. In *2013 Africon*. IEEE, sep 2013. 33

[89] Park, Jung Min and Jeffrey H. Reed: *Ensuring Operational Privacy of Primary Users in Geolocation Database-Driven Spectrum Sharing*, jun 2013. `https://www.ntia.doc.gov/files/ntia/publications/ssd-summary_report-tr1_1_link_5.pdf`. 33, 34

[90] Fischer, Hans: *A History of the Central Limit Theorem*. Springer New York, 2011, ISBN 9781461427018. 33

[91] Vulpiani, Angelo, Fabio Cecconi, Massimo Cencini, Andrea Puglisi, and Davide Vergni (editors): *Large Deviations in Physics*. Springer Berlin Heidelberg, 2014. 33

[92] Douc, Randal, Eric Moulines, Pierre Priouret, and Philippe Soulier: *Examples of markov chains*. In *Springer Series in Operations Research and Financial Engineering*, pages 27–52. Springer International Publishing, 2018. 35

[93] Barbu, Adrian and Song Chun Zhu: *Monte Carlo Methods*. Springer Singapore, ISBN 9789811329715. 35, 36

[94] Mackay, D. J. C.: *Introduction to monte carlo methods*. In *Learning in Graphical Models*, pages 175–204. Springer Netherlands, 1998. 35, 36

[95] Rubinstein, Reuven Y. and Dirk P. Kroese: *Simulation and the Monte Carlo Method*. Wiley-Interscience, 2007, ISBN 9780470177945. 36

[96] Oppenheim, Alan V. and George C. Verghese: *Signals, Systems and Inference*. Pearson, 2015, ISBN 9780133943283. 36

[97] Li, Xia, Marina Petrova, and Petri Mahonen: *FCSS: CSMA/CA based fast cooperative spectrum sensing over multiband cognitive networks*. In *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)*. IEEE, sep 2012. 37

[98] Khan, Zaheer, Janne J. Lehtomäki, Miia Mustonen, and Marja 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. 38

[99] 5G-Range: *Spectrum sensing to complement databases*. Technical report, oct 2018. `http://5g-range.eu/wp-content/uploads/2018/04/D4.2-Spectrum-Sensing-to-Complement-Databases.pdf`. 38, 56, 61

[100] Malady, Amy C. and Claudio 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. 38

[101] Wang, Shubin, Huiqin Liu, and Kun 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, 2015:1–7, 2015. 38

[102] Ferreira, Gabriel, Priscila Solis, Marcos Fagundes Caetano, Marcus Vinicius Lamar, Eduardo Alchieri, Johanna Vartiainen, and Heikki Karvonen: *COLAB: Módulo LTE de sensoriamento colaborativo e radio cognitivo para o simulador de redes ns-3*. In *Anais Estendidos do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*. Sociedade Brasileira de Computação - SBC, sep 2019. 52

[103] Jain, Raj: *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991, ISBN 9780471503361. 55

[104] 5G-Range: *Deliverable d7.1 - exploitation, communication, dissemination and standardization – part i*. Technical report, Jan 2019. `http:http://5g-range.eu/wp-content/uploads/2018/04/D7.1-Exploitationcommdisseminationstandarization-part-I.pdf`. 61

Annex I

COLAB: Módulo LTE de Sensoriamento Colaborativo e Radio Cognitivo para o Simulador de Redes ns-3

Gabriel C. Ferreira[1], Priscila Solís[1], Marcos F. Caetano[1], Marcos V. Lamar[1], Eduardo A. P. Alchieri[1], Johanna Vartiainen[2], Heikki Karvonen[2]

[1]Departamento de Ciência da Computação – Universidade de Brasília (UnB)
Caixa Postal 4466 CEP 70910-900 – Brasília – DF – Brasil

gabrielcarvfer@aluno.unb.br, {pris,mfcaetano,lamar,alchieri}@unb.br

[2]Centre for Wireless Communications – University of Oulu, Oulu, Finland

{johanna.vartiainen,heikki.karvonen}@oulu.fi

Abstract. *The exponential growth of wireless devices connected to the Internet and the limitation of the available electromagnetic spectrum requires the optimization of resource scheduling and frequency reuse to provide a better quality of service. New applications from the evolution of wireless networks, especially the 5G network, consider the use of cognitive radios and the sensing of the spectrum as fundamental elements in its development. The advance and development of new technologies depends on simulation environments that facilitate and integrate tools to evaluate the various applications and scenarios of 5G networks. This article presents a collaborative sensing module and a cognitive resource scheduler for the LS module of the ns-3 network simulator.*

Resumo. *O crescimento exponencial de dispositivos sem fio conectados à internet e a limitação do espectro electromagnético disponível requer a otimização do escalonamento de recursos e reúso de frequências para prover uma melhor qualidade de serviço. As novas aplicações oriundas da evolução das redes sem fio, principalmente a rede 5G, consideram o uso de rádios cognitivos e o sensoriamento do espectro como elementos fundamentais no seu desenvolvimento. O avanço e desenvolvimento de novas tecnologias depende de ambientes de simulação que facilitem e integrem ferramentas para avaliar as diversas aplicações e cenários das redes 5G. Este artigo apresenta um módulo de sensoriamento colaborativo e um escalonador de recursos cognitivo para o módulo LTE do simulador de rede ns-3.*

1. Introdução

A rápida expansão de dispositivos com conectividade sem fio, aliada a novos serviços na internet, geram forte demanda por espectro para transmissão. Para garantir níveis de qualidade de serviço apropriados aos diferentes tipos de aplicações é necessário o aumento da largura de banda e também o aumento da eficiência no uso do espectro disponível. Espera-se que essas necessidades deverão ser atendidas na próxima geração de redes sem fio, ou seja, a quinta geração (5G) que planeja fornecer uma capacidade aumentada em 1.000 vezes, uma taxa de dados 10-100 vezes maior e 10-100 vezes mais dispositivos conectados, em comparação com as atuais redes sem fio 4G.

Hoje em dia o espectro de radiofrequência está quase totalmente atribuído às diferentes operadoras licenciadas. Para permitir um melhor uso dos recursos de espectro limitado e otimizar a sua utilização são usadas várias tecnologias de multiplexação tais como Acesso Múltiplo por Divisão de Freqüência (FDMA em inglês), Acesso Múltiplo por Divisão de Tempo (TDMA em inglês) e técnicas de Múltiplas Entradas e Múltiplas Saídas (MIMO em inglês). Redes com acesso centralizado (TDMA, CDMA, *Ortoghonal* FDMA) garantem uma maior eficiência espectral ao evitar o desperdício de recursos com colisões, como ocorre no CSMA (*Carrier Sense Multiple Access*) ou nas requisições assíncronas para acesso ao meio (CSMA/*Collision Avoidance*). No controle de acesso ao meio (MAC) centralizado, os pontos de acesso indicam às estações associadas o momento e frequência em que podem transmitir e/ou devem receber uma transmissão, conforme a disponibilidade de recursos de rádio e necessidade das estações. Essas técnicas não se mostram capazes de resolver fundamentalmente o problema de escassez de recursos do espectro causado pela política de alocação fixa [Hu et al. 2018].

Vários estudos mostram [Valenta et al. 2009, Beckman et al. 2010, Hu et al. 2018] que o espectro licenciado é em muitos momentos subutilizado. A tecnologia de Rádio Cognitivo (RC) surgiu como uma proposta para resolver o problema de subutilização do espectro e otimizar a alocação de recursos. O espectro de radiofrequência apresenta um conjunto de faixas não contínuas de espectro não utilizados pelos respectivos usuários licenciados chamados de usuários primários (PUs em inglês). O sensoriamento de espectro é o primeiro passo do compartilhamento de espectro e permite que usuários secundários (SUs em inglês) usem de maneira oportunística as faixas de espectro vazias de acordo com sua necessidade e deixem de acessar o canal tão logo o PU retome sua transmissão. Este compartilhamento é chamado de acesso dinâmico ao espectro (DSA em inglês) e permite o uso de espectro ocioso desde que não interfira com o PU. A correta execução do procedimento permite aumentar a largura de banda disponível para os SUs e é considerada fundamental para o desenvolvimento das redes 5G.

Normalmente, o uso compartilhado do espectro envolve quatro passos: 1) sensoriamento; 2) alocação; 3) acesso e 4) *hand-off*. Trabalhos recentes de pesquisa [Hu et al. 2018] agrupam as técnicas de sensoriamento do espectro em 3 grandes grupos: a) não colaborativos, onde os SUs detectam o PU e evitam o canal, ideal para uma abordagem de protocolos MAC descentralizados; b) base de dados, em que cada SU envia o resultado do sensoriamento para o ponto de acesso ou estação rádio base, que compara resultado com dados históricos para encontrar faixas disponíveis e c) colaborativo, que combina os resultados de detecção de vários SUs para melhorar a confiabilidade da detecção usando algoritmos de fusão.

Amplamente usado no meio acadêmico, o ns-3 [NSNAM 2019] é um simulador de rede de eventos discretos de código aberto. Comparado a outros simuladores de código aberto, o ns-3 oferece alguns recursos de simulação multi-RAT (*Radio Access Technology*) e multi-banda, com Wi-Fi, WiGig, LTE (*Long Term Evolution*) (LTE-A, LAA, LTE-U), entre outros. Existe um interesse de estender as funcionalidades do ns-3 para permitir avaliar as tecnologias que promoverão o desenvolvimento da rede 5G, entre elas, os RC e as técnicas DSA. Existem várias iniciativas nesse sentido, por exemplo o LENA [CTTC 2019], desenvolvido para o LTE no ns-3. Entretanto, no momento em que

a presente ferramenta foi desenvolvida, não existia ainda uma implementação sólida de algoritmos de sensoriamento colaborativo com foco em redes LTE e 5G.

Com essa motivação, este trabalho apresenta o COLAB, um conjunto de adaptações para o módulo LTE do simulador ns-3 que implementa o sensoriamento colaborativo para rádio cognitivo. A implementação e validação deste módulo para o ns-3 utilizando a pilha LTE serve de base para a simulação de uma rede 3GPP [Sultan 2019], facilitando comparações de resultados e portabilidade de novos protocolos para a pilha do 5G-NR. O algoritmo de sensoriamento permite que a torre de telefonia chamada de *evolved NodeB* (eNB) estime o comportamento de um PU e aloque recursos de rádio de maneira a evitar colisões. O protótipo do sensoriamento colaborativo foi desenvolvido em C++ e um conjunto de ferramentas auxiliares para plotagem de resultados foram implementadas em Python.

As contribuições esperadas do COLAB à área de pesquisa em redes 5G são as seguintes: a) suporte ao desenvolvimento de novos algoritmos de escalonamento da camada de enlace, com base nas informações do sensoriamento colaborativo; b) suporte ao desenvolvimento de novas técnicas para caracterização do padrão de acesso ao canal do PU, a fim de permitir a utilização oportunística do canal sem causar prejuízos à transmissão primária; c) verificação do impacto dos mecanismos de sensoriamento da camada física nas camadas superiores e nas métricas das diversas aplicações das redes LTE e 5G; d) facilitar um ambiente para o desenvolvimento de algoritmos de detecção de espectro ideal em conjunto com estratégias de alocação de espectro e mecanismos de acesso ao espectro que atendam aos diferentes requisitos dos quatro cenários de aplicação das redes 5G. Este artigo está organizado como segue: a Seção 2 apresenta a especificação da arquitetura usada no ns-3 para implementar o algoritmo de sensoriamento colaborativo e seu uso com rádio cognitivos. A Seção 3 descreve o roteiro de demonstração da ferramenta. A Seção 4 apresenta as conclusões e trabalhos futuros.

2. Descrição da ferramenta

2.1. Arquitetura e funcionalidades

A implementação da arquitetura para sensoriamento colaborativo adotada no módulo LTE do simulador ns-3 é dispersa pelas múltiplas classes que implementam as diferentes camadas da pilha do protocolo, conforme mostrado na Figura 1. À camada LteSpectrumPhy, que corresponde à parte de transmissão e recepção de sinais, no método de recepção (`StartRX`) foi acrescida uma parte para integração dos dados recebidos e o encaminhamento para uma função de sensoriamento (`Sense`). A função `StartRxDlCtrl` foi modificada para permitir o recebimento de mensagems de controle com resultados do sensoriamento e o seu encaminhamento para a camada PHY do eNB (eNodeB, implementado no LteEnbPhy).

À camada MAC do módulo do UE (*User Equipment*), foi acrescida uma função para montagem da mensagem de controle com resultados do sensoriamento (`SendCognitiveMessageC`). Ainda à camada MAC do eNB, foram acrescidas funções para receber mensagens de sensoriamento das diferentes estações (`ReceiveCognitiveMessage`) e fazer fusão dos resultados (`mergeSensingReports`). Assim também, a função `DoSubframeIndication`

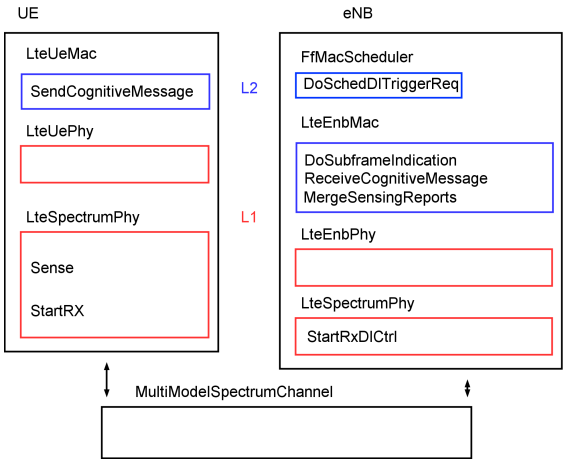


Figura 1. Implementação de pilha de rede LTE com sensoriamento no ns-3.

do LTE foi modificada, para implementar a fusão de resultados antes de executar o escalonamento de recursos. Ainda na camada MAC do eNB, a função `DoSchedDlTriggerReq` dos escalonadores passou a receber um parâmetro adicional, contendo um *bitmap* informando os grupos de blocos de recursos (RBGs em inglês) ocupados, prevenindo seu uso pelo escalonador.

2.2. Funcionamento dos Módulos Desenvolvidos

A Figura 2 mostra um esquema dos passos do algoritmo desenvolvido no plano de controle para a operação do sensoriamento colaborativo em um canal licenciado. A primeira versão dessa implementação no módulo exemplifica o acesso do usuário primário (PU) ao canal em um dado momento, seguida pela notificação da UE para o eNB sobre a presença do PU, a agregação das notificações pelo eNB e o seu efeito no escalonamento de recursos para prevenir transmissões que interfiram com o PU.

Na camada física (LteSpectrumPhy), está disposta a função de sensoriamento do canal (`Sense`), que executa o sensoriamento. Dado que o ns-3 não emula a camada física no caso da rede LTE, o algoritmo de sensoriamento do canal utiliza os parâmetros de SINR e uma curva de probabilidades de detecção de usuário primário, gerados externamente em um simulador de camada de enlace. O protótipo de algoritmo de sensoriamento implementado (`sensingProcedure`), verifica a relação SINR de cada um dos RBGs, ou a relação entre o SINR médio dos RBGs, e estima as chances de um PU estar transmitindo baseado na curva de probabilidade. A Figura 3 mostra um exemplo de curva de probabilidade de detecção que serve de entrada no ns-3, calculada em um ambiente simulado com o uso de vários pontos de amostragem e parâmetros do canal físico de interesse[5G-Range 2018]. A probabilidade de falso positivo é Pfa = 0.01. O módulo da camada física dos eNBs (LteEnbPhy) foi modificado para redirecionar mensagens com informações do sensoriamento cognitivo para a camada MAC do eNB associado.

Na camada MAC dos UEs (LteUeMac), se encontra a função de notificação da detecção de PU via canal de controle (`SendCognitiveMessageC`), ou via canal de dados (`SendCognitiveMessage`). A notificação contém o número do quadro e sub-quadro, mais um *bitmap* indicando os RBGs em que a transmissão do PU foi detectada. A camada MAC dos eNBs (LteEnbMac) foi modificada para receber as men-

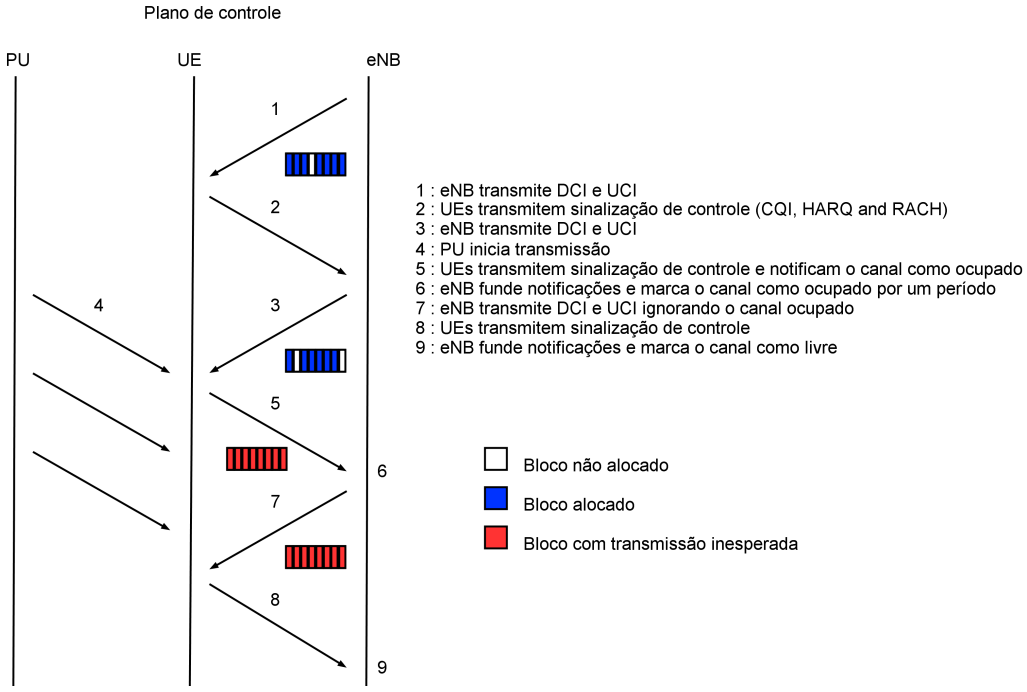


Figura 2. Diagrama do sensoriamento colaborativo

sagens do sensoriamento colaborativo (`RecvCognitiveMessage`, via canal de dados, ou `RecvCognitiveMessageC`, via canal de controle), guardando em seguida os resultados dos sensoriamentos dos UEs em registros que facilitam a posterior consulta. Ainda na camada MAC dos eNBs (LteEnbMac), dentro da função de indicação do sub-quadro (`DoSubframeIndication`) e antes do disparo de cada escalonamento (`SchedDlTriggerReq`), foi implementado um método que agrega os dados do sensoriamento colaborativo. O resultado dessa agregação é um *bitmap* de RBGs a serem ignoradas pelo escalonamento, evitando interferir com o usuário primário.

Um exemplo dos resultados da execução da ferramenta é mostrado na Figura 5, utilizando a topologia da Figura 4, que contém 10 dispositivos (UEs) e um eNodeB. O primeiro gráfico da Figura 5 mostra a potência de transmissão medida do PU no canal. No segundo gráfico, é mostrado o resultado do sensoriamento individual dos UEs ao longo da simulação. No terceiro gráfico, é mostrada a relação SINR média dos RBGs para cada um dos UEs ao longo da simulação, visualmente indicando a presença da transmissão do PU. No quarto gráfico, são plotados os resultados do escalonamento, com espaços em branco indicando RBGs livres, azul indicando RBGs escaladas (transmissões de *downlink*), enquanto RBGs com transmissões do PU são indicadas em vermelho, onde transmissões LTE cessam. Este gráfico reflete diretamente o algoritmo de agregação das informações de sensoriamento colaborativo advindo dos UEs, podendo aumentar ou reduzir o número de falsos positivos e/ou falsos negativos.

2.3. Aplicações e Casos de uso

A presente ferramenta pode ser utilizada como um suporte para o desenvolvimento e avaliação dos seguintes casos de uso:

- **Novos algoritmos de escalonamento para a camada MAC**: algoritmos clássicos

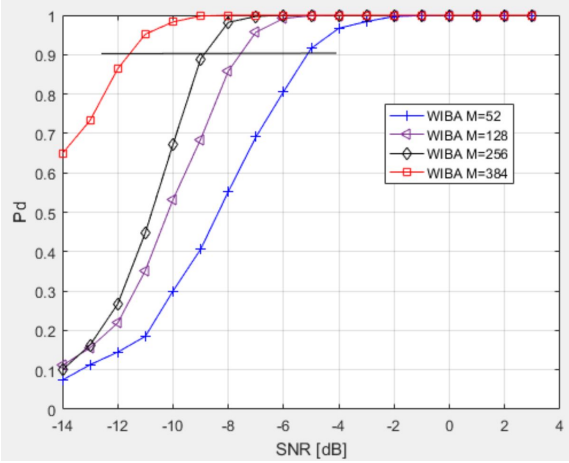


Figura 3. Curvas de probabilidade de detecção do PU vs SNR [5G-Range 2018]

de escalonamento tipicamente assumem que o canal é licenciado e, portanto, o eNB é usuário primário. O sensoriamento colaborativo implementado na ferramenta provê informações suficientes para a modificação e implementação de novos algoritmos que levem em conta a possibilidade de usar oportunisticamente um canal licenciado. A ferramenta assume que o canal de controle fica em banda licenciada e não sofre interferência de terceiros.

- **Caracterização do padrão de acesso do PU ao canal e rádio cognitivo**: o acesso à *whitespaces*, é tipicamente regido por bases de dados controlados por agências controladoras do espectro. As bases de dados porém, não contam com comportamento detalhado do padrão de acesso do PU, o que dificulta o reúso do espectro por terceiros. Neste sentido, a ferramenta oferece a possibilidade da implementação de algoritmos de aprendizado para identificar e caracterizar o padrão de acesso do PU, a fim de prever seus acessos e identificar oportunidades de acesso.
- **Efeitos colaterais no sistema baseados no extrapolação do sensoriamento real**: a ferramenta pode ser utilizada para verificar como diferentes mecanismos de sensoriamento da camada física, seja real ou simulada, afetam a rede como todo. O método `sensingProcedure` carrega uma curva de probabilidade da detecção de um PU, que pode ser substituída facilmente; permite a detecção para cada RBG ou no canal como todo, dependendo da necessidade do algoritmo de sensoriamento utilizado; pode ser feito apenas em RBGs não alocados para transmissão, evitando RBG escaladas para *downlink* (DCIs), *uplink* (UCIs) e retransmissões HARQ (RAR).

3. Roteiro de Demonstração

O código fonte, manual e vídeos estão disponíveis em `https://gabrielcarvfer.github.io/NS3/COLAB/`.

A demonstração da ferramenta se dará em um conjunto de cenários com e sem acesso dinâmico ao espectro, onde uma rede LTE utiliza oportunisticamente um canal licenciado para a transmissão de dados, coexistindo com o PU do canal. O cenário simulado difere do padrão LTE-LAA, onde a rede LTE faz uso de um canal não licenciado utilizando o princípio LBT (*Listen-Before-Talk*). A detecção de transmissões do

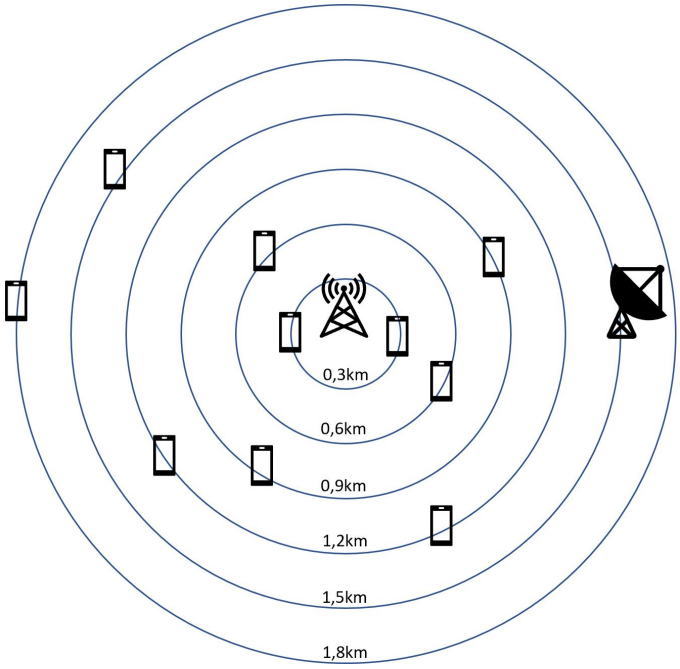


Figura 4. Topologia do cenário de simulação para sensoriamento colaborativo

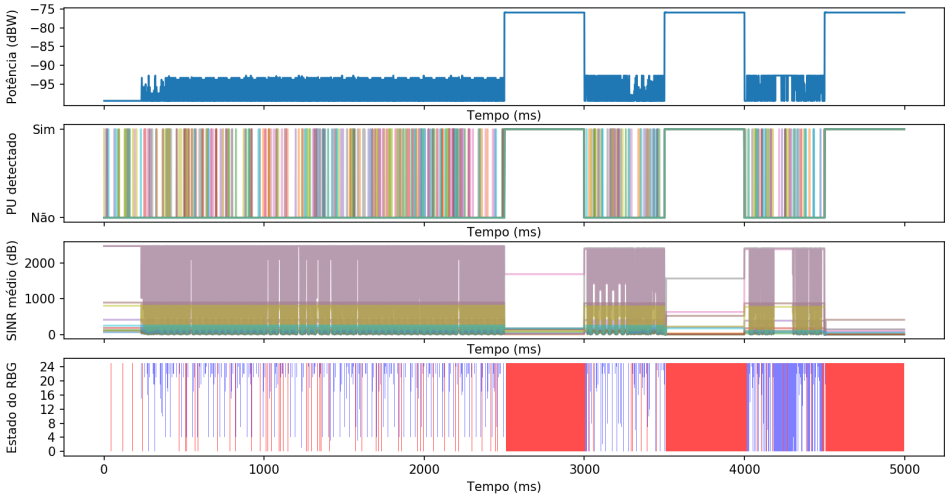


Figura 5. Resultados do sensoriamento colaborativo e reflexo no escalonamento de recursos

PU através do sensoriamento colaborativo e cessão das transmissões no canal licenciado, prevem e/ou reduzem a geração de interferência que afete o PU. A topologia da rede utilizada nos cenários é a mesma utilizada na Seção 2.2, mostrada na Figura 4. A curva de probabilidade utilizada é a mostrada na Figura 3.

Os cenários exemplificados na demonstração são 04, resultantes das combinações de um canal AWGN (*Additive white Gaussian noise*) de 20MHz, com ruído de fundo de -174dBm/Hz, com e sem a presença do usuário primário, com e sem DSA. As métricas comparadas são latência, vazão, interferência percebida à transmissão do usuário primário (medição de SINR) e taxa de utilização do canal.

4. Conclusão e trabalhos futuros

A escassez de recursos de rádio e a demanda por melhor cobertura e qualidade de serviço tornam evidentes a necessidade do desenvolvimento e aprimoramento de tecnologias que permitam o uso oportunístico do espectro. O COLAB, a ferramenta ora apresentada, pode ajudar na pesquisa, desenvolvimento e teste de algoritmos para redes cognitivas e acesso oportunístico integrado à pilha LTE do ns-3. Como trabalhos futuros, serão investigados novos métodos para aprimorar o ciclo cognitivo de sensoriamento, algoritmos de inteligência artificial para predição do comportamento do PU e alocação de RBGs, algoritmos de fusão dos dados de sensoriamento, melhorando os resultados da ferramenta e das redes simuladas.

5. Agradecimentos

Este trabalho foi realizado com suporte do projeto 5G-Range (www.5g-range.eu), aprovado na 4a. Chamada EU-BR em TICs, financiada pelo CTIC/RNP/MCTIC.

Referências

5G-Range (2018). Spectrum sensing to complement databases. Technical report. `http://5g-range.eu/wp-content/uploads/2018/04/D4.2-Spectrum-Sensing-to-Complement-Databases.pdf`.

Beckman, R., Channakeshava, K., Huang, F., Vullikanti, V. S. A., Marathe, A., Marathe, M. V., and Pei, G. (2010). Implications of dynamic spectrum access on the efficiency of primary wireless market. In *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN)*. IEEE.

CTTC (2019). The first release of 5g-lena is available. `http://www.cttc.es/the-first-release-of-5g-lena-is-available/`.

Hu, F., Chen, B., and Zhu, K. (2018). Full spectrum sharing in cognitive radio networks toward 5g: A survey. *IEEE Access*, 6:15754–15776.

NSNAM (2019). ns-3 network simulator. `https://www.nsnam.org/`.

Sultan, A. (2019). 21.915 release 15. Technical report. `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3389`.

Valenta, V., Fedra, Z., Marsalek, R., Baudoin, G., and Villegas, M. (2009). Towards cognitive radio networks: Spectrum utilization measurements in suburb environment. In *2009 IEEE Radio and Wireless Symposium*. IEEE.