UNIVERSIDADE FEDERAL FLUMINENSE

#### BRUNA TOLEDO GUEDES

## Um Algoritmo de Clusterização Espectral para Agrupamento Inteligente em Redes sem Fio Densas

NITERÓI 2022

#### UNIVERSIDADE FEDERAL FLUMINENSE

#### BRUNA TOLEDO GUEDES

### Um Algoritmo de Clusterização Espectral para Agrupamento Inteligente em Redes sem Fio Densas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Mestre em Computação. Área de concentração: Ciência da Computação.

Orientador: Diego Gimenez Passos

Co-orientador: Fernanda Gonçalves de Oliveira Passos

> NITERÓI 2022

Ficha catalográfica automática - SDC/BEE Gerada com informações fornecidas pelo autor

G924a	Guedes, Bruna Toledo Um algoritmo de clusterização espectral para agrupamento inteligente em redes sem fio densas / Bruna Toledo Guedes ; Diego Gimenez Passos, orientador ; Fernanda Gonçalves de Dimeiro Dagaco acorientadora Niterói 0202
	74 f. : il.
	Dissertação (mestrado)-Universidade Federal Fluminense, Niterói, 2022.
	DOI: http://dx.doi.org/10.22409/PGC.2022.m.15830176742
	<ol> <li>Rede de computadores. 2. Padrão IEEE 802.11ah. 3. Clusterização. 4. Produção intelectual. I. Passos, Diego Gimenez, orientador. II. Passos, Fernanda Gonçalves de Oliveira, coorientadora. III. Universidade Federal Fluminense. Instituto de Computação. IV. Título.</li> </ol>
	CDD -

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

#### BRUNA TOLEDO GUEDES

Um Algoritmo de Clusterização Espectral para Agrupamento Inteligente em Redes sem Fio Densas

> Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Mestre em Computação. Área de concentração: Ciência da Computação

Aprovada em fevereiro de 2022.

BANCA EXAMINADORA

Prof. Diego Gimenez Passos - Orientador, UFF

Termon da Parsos

Prof. Fernanda Gonçalves de Oliveira Passos -Co-orientador, UFF

Prof. Igo Monteiro Moraes, UFF

Rafaelli de Cancalho Continho Prof. Rafaelli de Carvalho Coutinho,  $\rm CEFET/RJ$ 

> Niterói 2022

Aos meus pais.

### Agradecimentos

Agradeço aos meus pais, que sempre me incentivaram e ajudaram em todos os momentos da minha vida.

Aos meus orientadores Fernanda Passos e Diego Passos que são fonte de inspiração e admiração para mim. À Fernanda agradeço pelos ensinamentos e dedicação, desde minha graduação, e pelo imenso incentivo para ingressar no mestrado. Ao Diego agradeço por sua orientação, pelos ensinamentos, pela atenção e paciência durante o desenvolvimento da dissertação.

Ao Professor Célio que também é um exemplo e inspiração para mim. Agradeço pelos ensinamentos, por ser sempre prestativo e atencioso.

À Professora Andressa Fernandes e à Prefeitura de Niterói, pela bolsa que concedeu o apoio financeiro, através do Projeto de Pesquisa Cidade Inteligente e Inclusão Financeira: uma moeda digital para o Município de Niterói. Agradeço também ao Professor Tiago Neves pelas contribuições no Projeto.

Aos membros da banca examinadora, pela disponibilidade.

Aos meus amigos pelo apoio e amizade em especial a Thais Marcolino, Bruno de Sá, Yuri Oliveira, Michelle Tizuka e Cleomar Oliveira.

À todos os demais pesquisadores e docentes que compartilharam seu tempo e conhecimento.

### Resumo

A densidade das redes sem fio vem aumentando com a popularização dos dispositivos móveis. As Redes sem Fio Densas (ou DWN, do inglês Dense Wireless Networks) apresentam desafios como a atual escassez espectral e a crescente demanda por capacidade. O mecanismo de *Restricted Access Window* (RAW) foi introduzido pela emenda IEEE 802.11ah para melhorar o desempenho das DWN. O RAW restringe o número de estações que podem acessar o canal, separando-as em grupos. Apesar do mecanismo RAW formar grupos balanceados, distribuindo igualmente as estações, o agrupamento é feito de forma arbitrária. Dessa forma, aumentam as possibilidades de existirem terminais ocultos no mesmo grupo. A presença dos terminais ocultos pode aumentar o número de colisões e diminuir a vazão da DWN. Neste trabalho, foi considerado o uso de duas estratégias de agrupamento na formação dos grupos RAW para diminuir os terminais ocultos, aumentando o desempenho da DWN. A primeira estratégia utiliza o K-Means considerando a posição geográfica das estações e a segunda utiliza a clusterização espectral considerando a informação Received Signal Strength Indicator (RSSI) das estações. O algoritmo de clusterização espectral foi proposto para, na implementação, suprir a ausência da posicão geográfica de cada estação. Também foi comparado o desempenho dos algoritmos de clusterização com o método de agrupamento padrão usado no IEEE 802.11ah e com o algoritmo Hidden Matrix-based Regrouping (HMR). Os resultados das simulações indicam que os algoritmos de clusterização superam significativamente o agrupamento padrão e o HMR em termos de taxa de colisão, vazão e atraso.

**Palavras-chave**: Redes Sem Fio Densas, IEEE 802.11ah, Janela de Acesso Restrito, Clusterização Espectral, Comunicação sem Fio

### Abstract

The density of wireless networks has been increasing with the popularization of mobile devices. Dense Wireless Networks (DWN) present challenges such as the current spectral scarcity and the growing demand for capacity. The Restricted Access Window (RAW) mechanism was introduced by the IEEE 802.11ah amendment to improve DWN performance. RAW restricts the number of stations that can access the channel by separating them into groups. Although the RAW mechanism forms balanced groups, distributing the stations equally, the grouping is done in an arbitrary way. This increases the chances that there are hidden terminals in the same group. The presence of hidden terminals can increase the number of collisions and decrease the throughput of DWN. In this work, the use of two clustering strategies in forming RAW groups was considered to decrease the hidden terminals, increasing the performance of DWN. The first strategy uses K-Means considering the geographical position of the stations and the second uses spectral clustering considering the Received Signal Strength Indicator (RSSI) information of the stations. The purpose of the spectral clustering algorithm is to increase the feasibility of the implementation of the method in the absence of the geographic position of each station. The performance of the clustering algorithms was also compared with the standard clustering method used in IEEE 802.11ah and with the Hidden Matrix-based Regrouping (HMR) algorithm. The simulations results indicate that the clustering algorithms significantly outperform the standard clustering and HMR in terms of collision rate, throughput, and delay.

**Keywords**: Dense Wireless Networks, IEEE 802.11ah, Restricted Access Window, Spectral Clustering, Wireless Communication

# Lista de Figuras

2.1	Exemplo do problema do terminal oculto.	6
2.2	Dispositivos atribuídos nos <i>slots</i> de um mesmo grupo [63]	9
4.1	Exemplos de agrupamentos gerados nas simulações [21]	19
4.2	Grafo de conectividade das estações ponderado pelo RSSI de cada enlace	20
4.3	Fluxo da estratégia baseada na clusterização espectral	22
4.4	Exemplos de grafos simples não ponderados e ponderados e suas respectivas matrizes de adjacência.	22
4.5	Exemplos de matrizes de grau para grafos ponderados e não ponderados.	23
4.6	Cálculo da matriz laplaciana.	23
4.7	Exemplo de autovalores e autovetores.	24
5.1	Fluxo do processo de execução de uma instância de simulação	29
6.1	Taxa de colisão versus número de estações	34
6.2	Média de atraso do pacote <i>versus</i> número de estações. Simulações feitas com 5 grupos	35
6.3	Média de atraso do pacote <i>versus</i> número de estações. Simulações feitas com 10 grupos.	36
6.4	Média de atraso do pacote <i>versus</i> número de estações. Simulações feitas com 15 grupos.	36
6.5	Vazão agregada da rede versus carga oferecida da rede	38
6.6	Taxa de colisão <i>versus</i> número de estações. Comparação entre o desempe- nho do K-Means Espectral com e sem pós-processamento, K-Means Origi- nal, assim como com o agrupamento arbitrário	30
	in the second control of the second s	50

6.7	Vazão agregada da rede $versus$ carga oferecida da rede. Comparação entre	
	o desempenho do K-Means Espectral com e sem pós-processamento, assim	
	como com o agrupamento arbitrário.	40
6.8	Taxa de colisão versus número de estações. Comparação entre o desempe-	
	nho do HMR e K-Means Espectral	43
6.9	Vazão agregada da rede <i>versus</i> carga oferecida da rede. Comparação entre o	
	desempenho do HMR e K-Means Espectral nas simulações com 100 estações.	44
6.10	Vazão agregada da rede <i>versus</i> carga oferecida da rede. Comparação entre o	
	desempenho do HMR e K-Means Espectral nas simulações com 400 estações.	44
6.11	Média de atraso do pacote <i>versus</i> número de estações. Comparação entre	
	o desempenho do HMR e K-Means Espectral.	45
6.12	Taxa de colisão versus número de estações. Comparação entre o desem-	
	penho do K-Means Espectral e do agrupamento arbitrário para simulações	
	com limiar de -94 dBm	47
6.13	Vazão agregada da rede versus carga oferecida da rede. Comparação entre	
	o desempenho do K-Means Espectral e do agrupamento arbitrário para	
	simulações com limiar de -94 dBm.	48
6.14	Taxa de colisão versus número de estações. Comparação entre o desem-	
	penho do K-Means Espectral e do agrupamento arbitrário nas simulações	
	com limiar de -94 d B m e dimensão do cenário de 1500 m $\times$ 1500 m	48
6.15	Vazão agregada da rede versus carga oferecida da rede. Comparação en-	
	tre o desempenho do K-Means Espectral e do agrupamento arbitrário nas	
	simulações com limiar de -94 d B m e dimensão do cenário de 1500 m $\times$	
	1500 m	49
6.16	Número de pares de estações <i>versus</i> dimensão do cenário. Relação entre a	
	quantidade de terminais ocultos com conectividade com o ponto de acesso	
	e a diferença da vazao entre K-Means Espectral e o arbitrário para o limiar	50
	-94 ubm	50
6.17	Número de pares de estações <i>versus</i> dimensão do cenário. Relação entre a	
	quantidade de terminais ocultos com conectividade com o ponto de acesso	
	-70 dBm.	52
	· · · · · · · · · · · · · · · · · · ·	_

# Lista de Tabelas

3.1	Artigos que propõem esquemas de agrupamento RAW	16
5.1	Principais parâmetros do simulador	28
5.2	Principais parâmetros do K-Means Espectral	30
6.1	Principais parâmetros da simulação	33
6.2	Tabela da quantidade de terminais ocultos de simulações com 100 estaçõese 15 grupos.	41
6.3	Tabela da quantidade de terminais ocultos de simulações com 400 estações	
	e 15 grupos	42

# Lista de Abreviaturas e Siglas

AID	:	Association Identifier;
AP	:	Access Point;
CCAThres	:	Clear Channel Assessment Threshold;
CR	:	$Cognitive \ Radio;$
CSMA/CA	:	Carrier Sense Multiple Access With Collision Avoidance;
DCF	:	Distributed Coordination Function;
DWN	:	Dense Wireless Networks;
EDCA	:	Enhanced Distributed Channel Access;
GPS	:	Global Positioning System;
HMR	:	Hidden Matrix-based Regrouping;
IoT	:	Internet of Things;
NIC	:	Network Interface Card;
PU	:	Primary User;
RAW	:	Restrict Access Window;
RPS	:	RAW Parameter Set;
RSSI	:	Received Signal Strength Indicator;
SDN	:	Software-Defined Networking;
TIM	:	Trafic Indication Map;

# Sumário

1	Intr	rodução 1					
	1.1	Organização do Texto	3				
<b>2</b>	Fun	damentação teórica	5				
	2.1	Visão geral dos conceitos	5				
	2.2	Redes sem fio densas 6					
	2.3	Mecanismo Restricted Access Window	8				
	2.4	Clusterização	10				
3	Tra	balhos Relacionados	13				
	3.1	Mecanismo Restricted Access Window	13				
	4 Novas estratégias de agrupamento RAW baseadas no K-Means						
4	Nov	vas estratégias de agrupamento RAW baseadas no K-Means	17				
4	<b>Nov</b> 4.1	vas estratégias de agrupamento RAW baseadas no K-Means Algoritmo K-Means	<b>17</b> 17				
4	<b>Nov</b> 4.1 4.2	vas estratégias de agrupamento RAW baseadas no K-Means Algoritmo K-Means	<ol> <li>17</li> <li>17</li> <li>19</li> </ol>				
4	Nov 4.1 4.2	vas estratégias de agrupamento RAW baseadas no K-Means         Algoritmo K-Means         Clusterização espectral e motivação         4.2.1       Suposições e Notação	<ol> <li>17</li> <li>17</li> <li>19</li> <li>21</li> </ol>				
4	Nov 4.1 4.2	vas estratégias de agrupamento RAW baseadas no K-Means         Algoritmo K-Means         Clusterização espectral e motivação         4.2.1       Suposições e Notação         4.2.2       Visão geral do método	<ol> <li>17</li> <li>17</li> <li>19</li> <li>21</li> <li>21</li> </ol>				
4	Nov 4.1 4.2	vas estratégias de agrupamento RAW baseadas no K-Means         Algoritmo K-Means	<ol> <li>17</li> <li>19</li> <li>21</li> <li>21</li> <li>22</li> </ol>				
4	Nov 4.1 4.2	xas estratégias de agrupamento RAW baseadas no K-Means         Algoritmo K-Means	<ol> <li>17</li> <li>17</li> <li>19</li> <li>21</li> <li>21</li> <li>22</li> <li>24</li> </ol>				
4	Nov 4.1 4.2	vas estratégias de agrupamento RAW baseadas no K-Means         Algoritmo K-Means	<ol> <li>17</li> <li>19</li> <li>21</li> <li>21</li> <li>22</li> <li>24</li> <li>27</li> </ol>				
4	Nov 4.1 4.2 Imp 5.1	Algoritmo K-Means	<ol> <li>17</li> <li>19</li> <li>21</li> <li>21</li> <li>22</li> <li>24</li> <li>27</li> <li>27</li> </ol>				

	5.3	Algoritmo de K-Means Espectral 30		
6 Avaliação de Desempenho				32
	6.1	Descri	ção das simulações	32
	6.2	Result	ados de simulações	33
		6.2.1	Comparação dos algoritmos de clusterização com o agrupamento arbitrário	33
		6.2.2	Efeito da função de pós-processamento do K-Means Espectral	38
		6.2.3	Comparação do algoritmo K-Means Espectral com o algoritmo HMR	41
		6.2.4	Impacto do limiar de detecção de portadora	46
		6.2.5	Influência do aumento do cenário	49
7	Con	clusão		53
	7.1	Trabal	hos Futuros	55
Re	Referências 56			

### Capítulo 1

### Introdução

Nos últimos anos aconteceu uma expansão do uso de dispositivos móveis com comunicação sem fio. As redes sem fio se tornaram uma alternativa amplamente utilizada, substituindo, em muitos cenários, as redes convencionais com fio [14]. Aliado a isso, a busca por conectividade cresce cada vez mais, contribuindo para a disseminação das *Dense Wireless Networks* (DWN). Exemplos de ambientes propensos às DWNs incluem estádios, shoppings e locais de conferências. Nesses locais, há uma grande concentração de pessoas compartilhando a mesma rede sem fio [66] — ou, pelo menos, o mesmo meio sem fio. As DWNs contribuem para o aumento do atraso no acesso ao canal e da taxa de colisão de pacotes [5].

A busca perene por conectividade certamente desempenha um papel nesta popularização das redes sem fio. No entanto, não se deve ignorar os méritos de padrões como o IEEE 802.11, assim como a terceira, quarta e quinta gerações de tecnologias celulares: baixo custo e bom desempenho são fatores decisivos na revolução da comunicação sem fio nas últimas décadas.

Atualmente, vários dispositivos móveis possuem Wi-Fi e 3G/4G/5G integrados. Portanto, qualquer concentração de pessoas provavelmente também é uma concentração de interfaces sem fio potenciais. Além disso, essas interfaces provavelmente estarão ativas, pois os dispositivos móveis geralmente executam várias aplicações de segundo plano que geram/consomem dados.

No entanto, a aplicação massiva da comunicação sem fio não se restringe aos dispositivos pessoais. A *Internet of Things* (IoT), por exemplo, tem como objetivo conectar todos os tipos de objetos do cotidiano [33]. Considerando as características de certos objetos, esse ambicioso requisito só pode ser atendido com o emprego de tecnologias sem fio. Esses fatores explicam a tendência recente de redes sem fio densas, que são caracterizadas por um grande número de estações em uma área geográfica limitada [26]. Dada a escassez de espectro eletromagnético utilizável — ou seja, a quantidade de faixas de frequência disponíveis para comunicação prática — e a natureza inerentemente compartilhada do meio sem fio, o aumento da densidade de tais redes representa um grande problema de escalabilidade para administradores de rede e pesquisadores.

Em tecnologias baseadas em contenção, conforme o número de estações concorrentes aumenta, as colisões se tornam predominantes [8]. Mais estações também tendem a aumentar o número de terminais ocultos, o que agrava ainda mais o problema. Terminais ocultos podem transmitir simultaneamente, porque essas estações estão distantes uma da outra e, portanto, não podem detectar que outras transmissões já estão em andamento, resultando em colisões. Este problema diminui drasticamente a vazão da rede [29].

Vários padrões de comunicação sem fio recentes tentam fornecer melhor suporte para cenários DWN. O IEEE 802.11ah é um deles [57]. Este padrão suporta um grande número de dispositivos, fornecendo mecanismos mais avançados para controlar o acesso ao meio e economizar energia. Particularmente, para lidar com a alta probabilidade de colisões em redes muito densas, o IEEE 802.11ah apresenta o mecanismo *Restricted Access Window* (RAW) [45]. O RAW restringe o número de estações que podem acessar o canal. Este mecanismo divide as estações em grupos e atribui intervalos de tempo a cada grupo. Apenas as estações de um grupo terão permissão para acessar o canal durante o intervalo de tempo de seu grupo [57]. O RAW atribui estações a grupos em rodízio de acordo com seu *Association IDentification* (AID) [45] — os AIDs são atribuídos de acordo com o momento em que cada estaçõo se associa à rede.

Embora esse agrupamento atinja o objetivo de reduzir os níveis de contenção, ele não considera os terminais ocultos. Neste trabalho, será mostrado como o algoritmo K-Means pode ser usado para agrupar as estações mais próximas — geograficamente — diminuindo a probabilidade de terminais ocultos serem colocados no mesmo grupo e melhorando o desempenho de uma DWN. Embora os resultados das simulações demonstrem que o método é promissor, na prática geralmente é difícil obter e manter a localização geográfica das estações da rede, tornando sua implantação muitas vezes inviável. Em um cenário real onde as pessoas estão usando *smartphones* associados a uma rede, por exemplo, à medida que se movem, a rede não tem informações diretas sobre a posição das estações em nenhum momento. Na ausência de um GPS (*Global Positioning System*) ou sistema de localização semelhante, até o próprio dispositivo pode não saber com precisão sua própria posição. Portanto, para uma implementação prática, o ideal seria usar informações que estão diretamente disponíveis para as — ou mensuráveis pelas — estações da rede. Devido à dificuldade de se obter a posição exata de cada estação, optou-se por utilizar as informações já disponíveis para as estações da rede. Uma possibilidade é usar o *Received Signal Strength Indicator* (RSSI), um valor que normalmente pode ser obtido em quadros recebidos na maioria das NICs (*Network Interface Card*) sem fio típicas e tende a se correlacionar com a distância entre as estações.

Motivado por isso, este trabalho propõe também um algoritmo de agrupamento para gerar os grupos RAW baseados no RSSI entre cada par de estações, ao invés de suas posições geográficas. Diferentemente das posições geográficas, o RSSI é um dado simples de ser obtido à medida que as estações recebem fisicamente as transmissões umas das outras no meio sem fio compartilhado. Com base nos valores RSSI medidos, foi empregada uma técnica de clusterização espectral, seguida de um pós-processamento que equaliza o tamanho dos grupos. Por meio de simulações em cenários densos, foi mostrado que o agrupamento proposto resulta em aumento da vazão da rede, devido a uma menor probabilidade de colisão causada pela diminuição dos terminais ocultos dentro dos grupos em comparação tanto com a estratégia de agrupamento usada no IEEE 802.11ah quanto com o algoritmo *Hidden Matrix-based Regrouping* (HMR) [64].

Em suma, este trabalho contribui para o estado da arte apresentando uma nova estratégia para agrupamento RAW com base na clusterização espectral. Outra contribuição é a implementação de um algoritmo de pós-processamento, após a clusterização, para balancear os grupos. As simulações também revelam vários padrões interessantes sobre a relação entre agrupamentos RAW e o desempenho de rede.

#### 1.1 Organização do Texto

O texto deste trabalho está dividido da seguinte forma. O Capítulo 2 apresenta uma visão geral dos conceitos básicos utilizados no trabalho, aborda os desafios encontrados em redes sem fio densas, fundamentos teóricos do mecanismo RAW e os algoritmos de clusterização. O Capítulo 3 faz uma revisão sobre trabalhos da literatura que analisam ou propõem formas de agrupamento RAW com o objetivo de aumentar o desempenho da rede em termos de vazão e eficiência energética. O Capítulo 4 descreve as estratégias de agrupamento RAW propostas baseadas no K-Means. O Capítulo 5 detalha as implementações do simulador de rede e da estratégia de agrupamento baseada na clusterização

espectral. O Capítulo 6 mostra as simulações que foram realizadas com o objetivo de analisar e comparar o desempenho dos algoritmos de clusterização propostos em relação ao método de agrupamento padrão usado no IEEE 802.11ah e ao algoritmo HMR [64]. Por fim, o trabalho é concluído e são abordadas melhorias que podem ser desenvolvidas futuramente.

### Capítulo 2

### Fundamentação teórica

Este capítulo apresenta uma visão geral dos conceitos básicos utilizados no trabalho, aborda os desafios encontrados em redes sem fio densas e fundamentos teóricos do mecanismo RAW, que é um técnica de acesso ao meio baseada em agrupamento [62]. Trata também dos algoritmos de clusterização que fazem parte das estratégias de agrupamento propostas pelo trabalho que serão descritos no Capítulo 4.

#### 2.1 Visão geral dos conceitos

Uma rede sem fio possibilita a conexão das estações e a transmissão de informações sem a utilização de cabos. Em várias tecnologias, a comunicação é feita por ondas eletromagnéticas propagadas por meios não-guiados [31]. As estações sem fio são os sistemas finais que executam aplicações, como *notebooks*, *smartphones* ou computadores de mesa [31]. A troca de informações ou quadros entre uma estação e outra é intermediada pelo ponto de acesso, sendo o responsável por receber um quadro de uma estação emissora e enviar para a estação receptora. As estações se conectam ao ponto de acesso por meio de enlaces de comunicação sem fio [31].

Nas redes densas, o grande número de estações competindo para acessar o canal e enviar os quadros influencia no aumento das transmissões feitas simultaneamente. Em tecnologias como o Wi-Fi, antes de acessar o canal, a estações realizam um procedimento chamado de detecção de portadora para verificar se alguma transmissão já está sendo realizada. Caso as estações realizem a detecção de portadora ao mesmo tempo, o canal será identificado como livre e as transmissões serão feitas simultaneamente. As estações que estão ocultas também podem transmitir ao mesmo tempo por estarem muito distantes e dessa forma não perceberem as transmissões que estão sendo feitas. A Figura 2.1 mostra um cenário com estação oculta. Nesse exemplo, as estações A e C estão no alcance de B, mas A e C não estão no alcance uma da outra, sendo consideradas estações ocultas. Caso a estação C envie um quadro para a estação B, a estação A, ao fazer a detecção de portadora, não consegue identificar essa transmissão por estar muito longe da estação C. Dessa forma, a estação A envia um quadro para a estação B ao mesmo tempo, acarretando uma colisão, sendo necessária a retransmissão desses quadros, aumentando o atraso e diminuindo a vazão da rede.



Figura 2.1: Exemplo do problema do terminal oculto.

#### 2.2 Redes sem fio densas

O termo *Rede Sem Fio Densa* tem sido usado com significados diversos ao longo dos anos. Cerca de vinte anos atrás, durante o auge da pesquisa sobre protocolos de roteamento para redes *ad hoc*, a densidade da rede comumente referia-se ao grau de conectividade típico de estações em uma topologia de rede sem fio de múltiplos saltos [20]. Alguns autores, entretanto, empregam o termo para denotar grandes redes infraestruturadas com muitos pontos de acesso (AP) [55]. Em particular, uma rede densa neste contexto se referiria à oferta tipicamente abundante de APs para cada estação cliente: em qualquer ponto da área de cobertura, uma estação normalmente tem vários APs disponíveis para associar-se. Em um sentido mais geral, a densidade pode se referir simplesmente à concentração de rádios operando nas mesmas frequências dentro do alcance um do outro. Segundo essa definição, esses rádios podem não estar conectados à mesma rede ou usar a mesma tecnologia. De todo modo, dada a natureza compartilhada do meio sem fio, a alta densidade pode levar a uma série de problemas diferentes.

Em tal cenário, talvez a questão mais fundamental seja a coexistência [51]: em outras

palavras, como diferentes redes podem operar em uma mesma região compartilhando uma determinada banda de frequência sem prejudicar o desempenho umas das outras. Isso é particularmente importante para tecnologias que operam em bandas não licenciadas, como IEEE 802.11 e IEEE 802.15.4 — implementadas comercialmente como Wi-Fi e Zigbee —, por exemplo. A coexistência entre diferentes tecnologias de comunicação sem fio é geralmente auxiliada pelo emprego de técnicas de espalhamento espectral [32]. Ao espalhar um sinal original por uma largura de banda consideravelmente maior, ele não apenas se torna mais imune a interferências, mas também tem menos probabilidade de causar interferências prejudiciais a terceiros.

Outra ferramenta para melhorar a coexistência é a canalização. Ao dividir a banda de frequência na qual uma determinada tecnologia opera em vários canais ortogonais, as redes próximas podem operar em faixas de frequência não sobrepostas, evitando interferências [19]. No entanto, a densidade das redes sem fio atuais atingiu níveis tão altos que a atribuição de canais é atualmente um problema difícil. As redes Wi-Fi de média a grande escala geralmente empregam controladores centralizados que periodicamente coletam dados sobre o uso do canal e aplicam algoritmos para realizar uma seleção automática e, espera-se, boa — de canal que reduz a interferência [38]. Um algoritmo genético pode ser empregado para melhorar as atribuições de canais ao longo de uma série de iterações consecutivas de etapas evolutivas, conforme apresentado em [60].

A dificuldade de selecionar canais é apenas um sintoma da escassez geral do espectro. O grande número de dispositivos sem fio, a busca aparentemente interminável por taxas de transmissão mais altas, as diversas características de propagação de diferentes bandas de frequência e o modelo tradicional de alocação de banda fixa são alguns dos fatores que levaram às condições espectrais de superlotação enfrentados ultimamente [24]. Nesse cenário, os Rádios Cognitivos (ou CR, do inglês *Cognitive Radio*) têm sido vistos como uma possível solução. Os CRs são transceptores capazes de operar em uma faixa relativamente ampla de bandas. O termo *cognitivo* está relacionado às capacidades de detecção avançadas dos CRs, permitindo-lhes detectar frequências ociosas que podem ser usadas oportunisticamente para aumentar a capacidade de comunicação. Em particular, os CRs podem tirar vantagem de partes licenciadas ociosas do espectro, usando-as para comunicação até que um Usuário Primário (ou PU, do inglês Primary User) — um dispositivo que possui a licença para operar naquela frequência — seja detectado. Essa detecção é um dos desafios dos CRs, porque deve acontecer rapidamente para evitar a interrupção da comunicação dos PUs. Como um exemplo recente da literatura, o trabalho em 1 usa uma estrutura baseada em lógica *fuzzy* para detectar frequências com alocação dinâmica.

Enquanto uma solução efetiva para a escassez espectral não é alcançada, o controle de acesso ao meio torna-se um tópico importante devido à tendência de concentração de estações operando no mesmo canal dentro do alcance de rádio umas das outras. Enquanto as tecnologias celulares geralmente utilizam protocolos baseados em divisão no tempo, frequência ou código, essas abordagens são muito inflexíveis para certos cenários, especialmente aqueles com tráfego intermitente e/ou imprevisível. Como consequência, os protocolos baseados em contenção são a escolha para várias tecnologias populares [6].

Apesar de potencialmente melhorar a utilização de recursos, a suscetibilidade a colisões dos protocolos baseados em contenção pode prejudicar gravemente o desempenho. Em particular, as colisões se tornam predominantes conforme o número de estações concorrentes aumenta [7] — claramente uma combinação ruim para a tendência atual de redes densas. A possibilidade adicional de estações ocultas — ou seja, estações que não podem detectar as transmissões umas das outras — agrava ainda mais os problemas. Além das colisões, as estações próximas ainda precisam compartilhar a largura de banda normalmente limitada dos canais sem fio. Mesmo que as estações sigam um escalonamento perfeito para acessar o meio, os recursos limitados se tornam um obstáculo.

#### 2.3 Mecanismo Restricted Access Window

Com o objetivo de aumentar a eficiência da rede, reduzindo colisões, atrasos e aumentando a vazão, as estações podem ser reunidas em grupos. Em vez de todas terem permissão de transmitir ao mesmo tempo, cada grupo de estações poderia possuir um intervalo de tempo para realizar as transmissões [56]. Como explicado no capítulo anterior, o mecanismo RAW do padrão IEEE 802.11ah adota essa ideia [27].

No IEEE 802.11ah, o ponto de acesso reserva um período de tempo, durante o intervalo de *beacon*, chamado de RAW [45]. Os grupos de estações são atribuídos de maneira uniforme pelo ponto de acesso. Depois disso, cada período RAW é dividido em um ou mais *slots* de tempo [57].

O ponto de acesso transmite periodicamente o quadro *beacon* contendo o elemento RPS (*RAW Parameter Set*) que especifica as configurações do RAW, como o número de grupos RAW, o número de estações atribuídas a cada um deles, a duração e o número de *slots* em cada grupo RAW. As estações obtêm essas informações e acessam o canal em seu período de tempo RAW correspondente [58]. Dentro do período RAW, o ponto de acesso pode atribuir as estações nos *slots* ou as estações podem selecionar aleatoriamente um *slot* [2]. As estações só poderão tentar acessar o canal dentro do seu *slot*, como mostrado na Figura 2.2 que foi feita com base no artigo [63].



Figura 2.2: Dispositivos atribuídos nos slots de um mesmo grupo [63].

Escolhido um *slot* dentro do período RAW do seu grupo, o acesso ao meio é feito pelo mecanismo Enhanced Distributed Channel Access (EDCA) [4, 28]. No início do slot RAW, o dispositivo gera um valor inicial para o contador de *backoff*, um valor aleatório entre o intervalo de  $[0, W_0 - 1]$  da janela de contenção [4, 28]. Em particular, o dispositivo realiza a detecção de portadora ao tentar transmitir pela primeira vez para verificar se o meio está livre [62]. A Figura. 2.2 mostra que o dispositivo 1 escolheu o *slot* 1. Quando ele fizer a detecção de portadora, perceberá que o canal está livre e poderá fazer a transmissão, já que não existe nenhum outro dispositivo nesse *slot*. Os dispositivos 2, 3 e 4 escolheram o slot 3: o dispositivo 2 está transmitindo e o 3 e 4 também desejam fazer uma transmissão. Quando mais de um dispositivo em um mesmo *slot* deseja transmitir, eles devem entrar em regime de *backoff* para evitar colisões [28]. Os dispositivos 3 e 4 devem ouvir o canal, fazendo a detecção de portadora até a transmissão que está sendo feita terminar e o canal ficar livre. Quando o meio estiver livre, os tempos de backoff são decrementados. O dispositivo que terminar de decrementar seu tempo de *backoff* primeiro poderá realizar a transmissão e os demais devem parar de decrementar seus tempos até que a transmissão termine [43].

Apesar de o mecanismo RAW distribuir igualmente as estações nos grupos, formando grupos balanceados, as estações são atribuídas em rodízio de acordo com seu AID, sem considerar as suas posições geográficas. Dessa forma, dependendo das característica da rede, aumentam as possibilidades de existirem terminais ocultos no mesmo grupo. A presença dos terminais ocultos pode aumentar o número de colisões e diminuir a vazão da rede. Evitar a inclusão de terminais ocultos na formação dos grupos RAW é justamente o objetivo dos métodos de agrupamento propostos nesse trabalho.

#### 2.4 Clusterização

A clusterização é uma técnica de aprendizado de máquina não supervisionada que tem por objetivo agrupar um conjunto de dados semelhantes em *clusters* [23]. Para fazer o agrupamento é necessário usar uma medida para determinar se os dados são semelhantes ou diferentes. Essa medida de similaridade é baseada, normalmente, em medidas de distância como, por exemplo, a distância Euclidiana e a distância de Manhattan [23]. Os *clusters* são formados de modo que os pares de dados em um mesmo *cluster* possuem um valor mínimo de distância e os pares de dados em *clusters* diferentes possuem distância máxima [23]. De acordo com [23, 44], os algoritmos de agrupamento podem ser classificados como: métodos hierárquicos, métodos baseados em densidade e métodos de particionamento [23].

Os métodos hierárquicos criam uma hierarquia de *clusters*, fazendo a junção deles ou dividindo-os em *clusters* menores [16]. Nesses métodos, existem duas abordagens: a abordagem aglomerativa e divisiva. Na abordagem aglomerativa, os *clusters* são gerados de cima para baixo [17]. O ponto de partida são os dados individuais sendo posteriormente mesclados para criar uma estrutura semelhante a uma árvore [17]. Inicialmente, cada dado é atribuído a um *cluster* próprio, totalizando *n clusters*. Em cada etapa, a distância entre cada par de *clusters* é calculada e em seguida os pares mais próximos são unidos em um único *cluster* [37]. Na abordagem divisiva, o agrupamento é feito de baixo para cima [17]. Inicialmente, os dados são agrupados em um único *cluster* e, em cada etapa, os dados são divididos em dois *clusters* menores [16]. O processo se repete até quando uma condição de término seja alcançada como, por exemplo, um número k de *clusters* seja formado [39]. Nas duas abordagens, o resultado é exibido graficamente como uma árvore, chamada de dendrograma, que registra a sequência das divisões dos dados nos *clusters* [25].

Os métodos baseados em densidade visam formar os *clusters* na medida em que o número de dados na sua vizinhança tenha uma proximidade determinada. O agrupamento é feito nas regiões que possuem um maior número de dados e devem ser separadas das regiões que possuem menor concentração de dados, ou seja que possuem baixa densidade [49]. A ideia principal desses métodos é que, para cada dado de um *cluster*, deve existir um número mínimo de outros dados em um determinado raio, ou seja, o número de dados pertencentes a esse agrupamento deve exceder a algum limite [12]. O DBSCAN é um dos algoritmos de clusterização baseado em densidade mais utilizados [35]. Esse algoritmo exige como entrada o número mínimo de pontos (MinPts) e o epsilon que é o raio do círculo criado em torno de cada ponto de dado, para medir a densidade [34]. Os pontos de dados são classificados como central, borda e ruído. O ponto é classificado como central quando possui em sua vizinhança pelo menos o número mínimo de pontos. Os pontos de borda são aqueles que não possuem o número mínimo de pontos, porém possuem pelo menos um ponto central em sua vizinhança. Os pontos de ruído são aqueles que não possuem o número mínimo de pontos e também não possuem um ponto central [48].

A abordagem de particionamento é considerada a mais popular [23]. Nessa abordagem, os agrupamentos são formados para otimizar um critério de particionamento, como uma função de dissimilaridade, ou seja, uma função que recebe dois dados e retorna a distância entre eles, de modo que os dados dentro de um *cluster* sejam "semelhantes", enquanto os dados em *clusters* separados sejam "diferentes" [44]. Os métodos de particionamento são vantajosos em aplicações que necessitam de um número fixo de *clusters* [44]. O bem conhecido algoritmo K-Means é exemplo de algoritmo que utiliza o método de particionamento [23].

O algoritmo K-Means tem por objetivo agrupar os pontos de dados em k partições, onde cada partição representa um *cluster* [23]. Os dados de entrada de um algoritmo clássico do K-Means são o número k de *clusters* e o conjunto de dados. O algoritmo inicializa escolhendo aleatoriamente os dados que serão o centro dos *clusters*, chamados de centroides. Em seguida, é calculada a distância de cada dado para cada centroide. Para isso, pode ser usada a medida da distância euclidiana, por exemplo. Os pontos serão associados ao centroide mais próximo a eles. Em seguida, é computada uma média de todos os pontos em cada centroide para definir a posição dos novos centroides. O processo se repete com o cálculo da distância entre os dados e os novos centroides [52]. O processo termina quando não é mais possível fazer atualizações nas mudanças dos dados para outros *clusters*.

A clusterização espectral é um algoritmo de particionamento de grafo [9]. O problema de agrupamento de grafos já foi abordado em vários trabalhos na literatura [61, 10, 42]. Na clusterização espectral, a similaridade dos dados é representada na forma de um grafo, denominado grafo de similaridade, constituído por um conjunto de vértices que representam os dados e um conjunto de arestas que os conectam [59]. O problema de agrupamento pode ser reformulado usando uma abstração de grafo na qual se deseja encontrar uma partição do grafo de forma que os pesos entre os vértices do mesmo grupo sejam altos e os pesos dos vértices em grupos diferentes sejam baixos [59]. As técnicas espectrais são aplicadas em vários contextos, como por exemplo em segmentação de imagem, classificação de páginas web e redução de dimensionalidade do espaço [50]. Os algoritmos K-Means e clusterização espectral foram escolhidos como estratégias de agrupamento RAW e serão detalhados no Capítulo 4.

### Capítulo 3

### Trabalhos Relacionados

Este capítulo apresenta uma revisão sobre trabalhos da literatura que abordam temas relacionados com esta dissertação. A Seção 3.1 apresenta os trabalhos que analisam ou propõem formas de agrupamento RAW com o objetivo de reduzir os terminais ocultos e aumentar o desempenho da rede em termos de vazão e eficiência energética.

#### 3.1 Mecanismo Restricted Access Window

Em DWNs, entender e melhorar o desempenho do protocolo CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*) é um tópico importante e bem estudado. Bianchi *et al.* [8], por exemplo, apresentam um modelo de desempenho do CSMA/CA e mostram que as colisões aumentam rapidamente com o crescimento do número de estações concorrentes — mesmo na ausência de terminais ocultos — reduzindo assim o desempenho da rede em termos de vazão.

O mecanismo RAW foi introduzido pela emenda IEEE 802.11ah para reduzir a ocorrência de colisões [56]. O RAW tenta diminuir a contenção e colisões do canal limitando o número de estações que podem tentar acessar o canal. Este mecanismo faz isso distribuindo estações em grupos e atribuindo *slots* exclusivos para cada grupo. O RAW possui uma série de parâmetros cuja escolha não está totalmente definida no padrão, mas pode afetar bastante o desempenho da rede [45]. Dois exemplos desses parâmetros são a definição da duração do *slot* e do número de grupos — que determina o número médio de estações competindo em cada grupo. Wang *et al.* [63], por exemplo, investigam o problema de otimizar a eficiência energética do RAW para comunicações de *uplink* aplicando a teoria da probabilidade. Um algoritmo é proposto para identificar a duração ideal do *slot* e o número ideal de estações concorrentes em cada grupo do ponto de vista energético. Kureev *et al.* [30] mostram que o consumo de energia das estações pode ser controlado conjuntamente usando a segmentação do TIM (*Traffic Indication Map*) e RAW. Os autores utilizam um modelo matemático para demonstrar que, configurando os parâmetros desses mecanismos, é possível diminuir o consumo de energia. Nawaz *et al.* [41] propõem um modelo no qual a duração do *slot* RAW é escolhida de acordo com o tamanho do grupo, a fim de aumentar a vazão da rede.

Por outro lado, alguns autores estudaram como diferentes estratégias de agrupamento podem influenciar o desempenho do RAW. Por exemplo, Zheng et al. [65] analisam o desempenho do Group-Synchronized Distributed Coordination Function (GS-DCF) em DWNs, uma função de coordenação distribuída do padrão IEEE 802.11ah. Ao usar o RAW, o GS-DCF pode melhorar a vazão da rede. No artigo [65], foi comparado o rendimento do GS-DCF usando esquemas de agrupamento centralizado e descentralizado. No agrupamento centralizado, o ponto de acesso distribui de forma uniforme as estações nos grupos e também aloca uniformemente os slots RAW aos grupos. No entanto, em uma rede, o número de estações pode variar com tempo. Considerando um agrupamento centralizado, o ponto de acesso deverá reagrupar as estações com frequência, podendo ocasionar um alto *overhead* de controle. Em um agrupamento descentralizado, esse problema pode ser minimizado pois permite que as estações escolham aleatoriamente o grupo ao qual devem pertencer. Os autores mostram que o RAW aumenta a vazão em redes densas e diminui a contenção do canal. Além disso, os autores mostram que a vazão do esquema de agrupamento centralizado foi semelhante a do agrupamento descentralizado. Porém, nas simulações, não foram considerados terminais ocultos, o que afetaria o desempenho da rede.

Bhandari *et al.* [5] propõem um mecanismo de agrupamento de estações em uma rede IoT densa com o objetivo de diminuir o número de estações concorrentes e ocultas, melhorando assim a eficiência de uso do canal. Nessa proposta, um *cloud-center* fornece ao ponto de acesso as informações da localização das estações, facilitando o processo de clusterização. A área de cobertura do ponto de acesso é dividida em setores. Dentro de cada setor, as estações são uniformemente divididas em grupos. Por meio de simulações, os autores mostram que essa estratégia de agrupamento melhora a vazão e o atraso da rede em relação à *Distributed Coordination Function* (DCF) do IEEE 802.11. Embora o ponto de acesso faça a divisão das estações em grupos de acordo com o número de estações e suas localizações, o artigo não propõe um mecanismo para evitar terminais ocultos dentro dos grupos. Além disso, o *cloud-center* deve de alguma forma conhecer as posições geográficas das estações, o que pode ser inviável na prática, pois as posições das estações podem estar em constante mudança e, em geral, as próprias estações podem não ter meios de saber sua própria localização.

Em [64], os autores propõem um algoritmo para diminuir o número de terminais ocultos em uma rede. O algoritmo, chamado de *Hidden Matrix-based Regrouping*, ou simplesmente HMR, move as estações entre grupos para diminuir o número de terminais ocultos a partir do agrupamento gerado pelo RAW do padrão IEEE 802.11ah. Primeiro, um grupo é selecionado e uma lista de terminais ocultos neste grupo é criada. A estação que possui mais terminais ocultos é selecionada para ser movida para outro grupo, desde que não seja um terminal oculto de nenhuma estação desse grupo — caso contrário, outro grupo de destino será escolhido. Se todos os grupos foram considerados alvos potenciais e em todos eles a estação selecionada possui terminais ocultos, a estação permanece no grupo atual e é removida da lista. Este processo é repetido para todas as estações da lista de terminais ocultos do grupo atual e, a seguir, o mesmo procedimento é aplicado a todos os grupos restantes. O HMR é um pré-processamento heurístico que já parte de um agrupamento potencialmente ruim, podendo eventualmente não encontrar boas soluções, conforme ilustrado no Capítulo 6.

A Tabela 3.1 lista os artigos citados neste capítulo, que propõem ou analisam esquemas de agrupamento, indicando se consideram duas características importantes das estratégias de agrupamento propostas neste trabalho: se a solução depende das posições geográficas das estações e se considera os terminais ocultos como sendo o problema a ser minimizado. Alguns artigos possuem os mesmos objetivos deste trabalho, como a redução dos terminais ocultos e o aumento do desempenho da rede em termos de vazão. Porém, dentre as propostas atuais encontradas na literatura, nenhuma atende simultaneamente aos critérios desejados. Dentre as pesquisas que foram feitas, não foram encontrados artigos que aplicaram algoritmos de clusterização baseados em IA para agrupamento RAW com o objetivo de reduzir os terminais ocultos em redes densas. Portanto, um novo método de clusterização que trata do problema dos terminais ocultos se torna relevante. A estratégia de agrupamento RAW proposta baseada na clusterização espectral é apresentada no Capítulo 4. A estratégia considera uma informação que está disponível para as estações da rede para fazer a agrupamento, possibilitando sua implementação na prática e se mostrou eficaz em melhorar o desempenho da rede, diminuindo as taxas de colisões e aumentando a vazão.

Trabalho	Depende do	Considera	Qual objetivo a ser otimizado
	conhecimento	terminais	
	das posições	ocultos	
Bhandari <i>et al.</i>	Sim	Não	Reduzir a contenção do canal.
Yoon, Seo e Bahk	Não	Sim	Reduzir o número de termi-
			nais ocultos
Zheng <i>et al.</i>	Não	Não	Analisar o desempenho do
			GS-DCF em uma DWN.
Wang <i>et al.</i>	Não	Não	Otimizar a eficiência energé-
			tica em comunicações uplink,
			identificando os números ide-
			ais de estações e de slots em
			cada RAW.
Kureev et al.	Não	Não	Aumentar o rendimento das
			estações através de um mo-
			delo matemático que permite
			encontrar parâmetros ideais
			para o TIM e o RAW.
Nawaz <i>et al.</i>	Não	Não	Aumentar a vazão da rede
			propondo um novo modelo
			RAW, no qual o quadro é
			dividido em dois subquadros
			e a duração dos slots em
			cada subquadro é escolhido
			de acordo com o tamanho do
			grupo.

Tabela 3.1: Artigos que propõem esquemas de agrupamento RAW.

### Capítulo 4

## Novas estratégias de agrupamento RAW baseadas no K-Means

Este capítulo apresenta os detalhes das duas estratégias de agrupamento RAW propostas neste trabalho, baseadas no K-Means, os motivos das escolhas destas estratégias e a metodologia de avaliação das estratégias.

#### 4.1 Algoritmo K-Means

A primeira estratégia de agrupamento RAW proposta por este trabalho foi utilizar o algoritmo de clusterização K-Means para agrupar as estações evitando colocar terminais ocultos em um mesmo grupo. A escolha do K-Means foi feita por ser o algoritmo de clusterização mais utilizado na literatura e devido à possibilidade de agrupar as estações de acordo com uma característica similar [52]. No caso deste trabalho, o agrupamento é feito utilizando as posições geográficas das estações. Assim, é possível reunir as estações mais próximas em um mesmo grupo e as mais distantes em grupos diferentes. Como as estações mais próximas serão agrupadas, a probabilidade de existir terminais ocultos em um mesmo grupo diminui e consequentemente aumenta o desempenho da DWN.

O Algoritmo 1 detalha a versão específica do K-Means utilizada no trabalho. Os dados de entrada são o número k de *clusters* e o conjunto de estações com suas respectivas posições geográficas. O algoritmo começa definindo as posições dos centroides iniciais dos *clusters* de forma aleatória, sendo atribuída a mesma posição geográfica de estações do conjunto de estações escolhidas aleatoriamente. Neste trabalho, considerou-se um máximo até 20 iterações: se os *clusters* não estabilizarem, o processo é interrompido na vigésima tentativa. Após a definição dos centroides iniciais, as iterações iniciam. Em seguida, é

```
Algoritmo 1: Algoritmo K-Means.
```

```
Entrada: K (Número de clusters), C (Conjunto de estações com as posições
 geográficas).
Saída: Estações agrupadas em k cluster, de acordo com a distância euclidiana
 for 0 \le i \le K do
   // Define uma estação aleatória como posição inicial do cluster
   cluster[i].centroide \leftarrow rand(C.posicao)
end
// Repita até os clusters estabilizarem ou até alcançar as 20
   iterações
for 0 < iteration < 20 do
   foreach E \in C do
      // calcula a distância entre estação E e o centroide do
         primeiro cluster
      menorDistancia = getDistance(E, cluster[0].centroide)
      indice \leftarrow 0
      // Ignorar o primeiro cluster
      for 1 \leq i < K do
         // calcula a distância entre estação {\it E} e o centroide i
         distancia = getDistance(E, cluster[i].centroide)
         // se a distância da estação E para o centroide i for
             menor do que a distância para o primeiro cluster então
             muda a estação para o centroide i
         if distancia < menorDistancia then
            indice \leftarrow i
             menorDistancia = distancia
         end
      end
      cluster[indice].append(E)
   end
   mudouPosicao = 0
   // para cada cluster
   for 0 \le i < K do
      // Cálculo da média da posição de todas as estações em cada
          centroide
      novaPosicao = calculateCentroid()
      if novaPosicao ! = cluster/i/.centroide then
         cluster[i] = novaPosiçao
         mudouPosicao = 1
      end
   end
   if mudouPosicao == 0 then
   | break
   end
end
```

calculada a distância das estações para cada centroide. A estação será atribuída para o *cluster* mais próximo. Em seguida é calculada a média das posições de todas as estações em cada *cluster* para reposicionar os centroides e, novamente, cada estação é atribuída ao centroide mais próximo. Depois é verificado se houve mudança na posição dos centroides. Se os centroides pararam de se mover, então a repetição é interrompida, antes do limite de 20 iterações. Apesar de os resultados das simulações demonstrarem que o método é promissor (vide Capítulo 6), na prática, é difícil de obter e manter a localização geográfica de estações da rede, inviabilizando sua implantação.

A Figura 4.1 ilustra a comparação dos resultados dos agrupamentos baseados no RAW sem utilizar o K-Means (arbitrário) e utilizando o K-Means para uma das instâncias de simulação com 1000 estações. O agrupamento gerado pelo mecanismo RAW está sendo chamado de arbitrário por não considerar as posições geográficas das estações.



Figura 4.1: Exemplos de agrupamentos gerados nas simulações [21].

Uma vez que o RAW usa uma estratégia *round-robin* simples sobre os AIDs — que não estão correlacionados com as posições das estações — há uma distribuição uniformemente arbitrária de estações atribuídas a grupos na área simulada. Por outro lado, o K-Means cria 5 sub-regiões muito distintas, cada uma atribuída a um grupo diferente.

#### 4.2 Clusterização espectral e motivação

O K-Means utiliza uma medida para calcular a distância entre as estações [23]. Neste trabalho, foi utilizada a distância euclidiana. A melhor forma de se obter a distância entre uma estação e o ponto de acesso é através da sua posição geográfica. Assim como no trabalho de Bhandari *et al.* [5], obter as informações das posições geográficas das estações pode não ser possível em um cenário real, tornando a implementação do agrupamento baseado em K-Means inviável na prática. Em um cenário real em que pessoas estão utilizando *smartphones* associados a uma rede, por exemplo, quando elas se locomovem, a rede não recebe em nenhum momento a informação de suas posições. Mesmo em cenários com estações estáticas, as posições geográficas não são trocadas. Para uma implementação prática, o ideal seria usar uma informação que está disponível para as estações da rede.

Nesse contexto, a clusterização espectral é capaz de criar o agrupamento de estações desejado a partir de informações como o *Received Signal Strength Indication* (RSSI). Devido à já mencionada dificuldade em obter a posição geográfica exata de cada estação, optou-se por usar no trabalho a clusterização espectral como a segunda estratégia de agrupamento RAW, considerando informações já disponíveis para as estações da rede. Assim, em vez de usar a posição geográfica, essa estratégia usa o RSSI, uma informação que pode ser obtida trivialmente para quadros recebidos na maioria das *Network Interface Card* (NICs) sem fio típicas. A partir do RSSI, é possível construir um grafo interrelacionando as estações, de forma que duas estações sejam mais fortemente conectadas no grafo à medida que o RSSI entre elas seja mais alto. A Figura 4.2 mostra um exemplo de um grafo no qual os vértices representam as estações da rede, as arestas representam os enlaces que conectam as estações e os pesos das arestas são os RSSIs (em dBm).



Figura 4.2: Grafo de conectividade das estações ponderado pelo RSSI de cada enlace.

O RSSI é fácil de se obter porque essa informação é medida sempre que uma interface recebe um quadro de outra interface. Entretanto, o RSSI não pode ser usado da mesma forma que foram utilizadas as localizações geográficas das estações. Enquanto a localização geográfica da estação é uma característica da própria estação, o RSSI é uma característica da relação entre duas estações. Como o RSSI está, geralmente, relacionado à distância — tende a ser maior para estações que estão mais próximas e mais baixo para estações mais distantes —, o comportamento esperado é que a saída da clusterização espectral calculada sobre a informação RSSI seja semelhante à do algoritmo K-Means sobre as posições geográficas.

#### 4.2.1 Suposições e Notação

A entrada do algoritmo é um grafo ponderado, conforme exemplificado na Figura 4.2. Este grafo foi denotado por G = (V, E) onde os vértices (elementos de V) deste grafo são estações da rede e as arestas (elementos de E) são os enlaces sem fio. Note que, embora o foco desse trabalho sejam as redes infraestruturadas, nas quais todas as comunicações são intermediadas pelo AP, o grafo de conectividade utilizado para a clusterização considera arestas também entre duas estações. Embora essas arestas não sejam usadas para comunicação na rede, elas indicam as relações de conectividade e interferência entre as estações, importantes para representar os possíveis terminais ocultos existentes.

Em nossa proposta, os pesos de aresta usados para medir a relação entre as estações são os respectivos RSSI de cada enlace viável. Mais formalmente, assume-se que existe uma função  $w : E \mapsto \mathbb{R}$  que atribui um peso (RSSI) a cada aresta. Na prática, foi presumido que, quando uma estação decodifica fisicamente um quadro de outra estação, ela mede o RSSI. De tempos em tempos, cada estação relatará essas informações de RSSI ao ponto de acesso. As informações serão compiladas em um grafo pelo ponto de acesso ou por um servidor controlador da rede — e o algoritmo de agrupamento será executado.

#### 4.2.2 Visão geral do método

Um fluxograma da estratégia baseada na clusterização espectral é mostrado na Figura 4.3. A entrada do algoritmo é um grafo constituído pelas estações da rede (vértices) e os enlaces (arestas) que conectam essas estações. Em cada aresta há um peso associado representando o RSSI do respectivo enlace. Na primeira etapa, é executado o algoritmo de clusterização espectral, que foi denominado de K-Means Espectral. O K-Means Espectral é composto pela decomposição espectral e pela execução do algoritmo K-Means. Na decomposição espectral, o grafo é representado por uma matriz, a partir da qual são calculados autovalores e autovetores. Um subconjunto dos autovetores são selecionados e utilizados na execução do K-Means para fazer o agrupamento. Em seguida, é executada a segunda etapa dessa estratégia, a função de pós-processamento, para fazer o balanceamento das estações. Como resultado, são obtidos os grupos formados de acordo com a informação do RSSI. Essas etapas são detalhadas na próxima subseção.



Figura 4.3: Fluxo da estratégia baseada na clusterização espectral.

#### 4.2.2.1 K-Means Espectral

No algoritmo K-Means Espectral, primeiramente é feita a decomposição espectral que começa com a construção da matriz de adjacência A, representando o grafo: as linhas e colunas correspondem aos vértices, enquanto o valor de cada entrada representa a existência — ou a falta — da aresta correspondente [11]. Essa etapa é exemplificada na Figura 4.4: a Figura 4.4(a) apresenta um grafo não ponderado e a Figura 4.4(b), um grafo ponderado com suas respectivas matrizes de adjacência.



Figura 4.4: Exemplos de grafos simples não ponderados e ponderados e suas respectivas matrizes de adjacência.

Em seguida, é calculada a matriz de graus D, que é uma matriz diagonal cujas entradas denotam o número de arestas que cada vértice possui — no caso de um grafo não ponderado — ou a soma dos pesos das arestas que esse nó possui — no caso de um grafo
ponderado [54]. A Figura 4.5 mostra as matrizes de grau dos grafos não ponderados e ponderados da Figura 4.4.



Figura 4.5: Exemplos de matrizes de grau para grafos ponderados e não ponderados.

Um conceito central do agrupamento espectral é a matriz laplaciana, que captura relações de similaridade do grafo original [13]. Matematicamente, a matriz laplaciana L é definida como:

$$L = D - A, \tag{4.1}$$

onde A denota a matriz de adjacência do grafo original e D denota a matriz de grau associada. A Figura 4.6 exemplifica o cálculo de uma matriz laplaciana associada ao grafo ponderado da Figura 4.4(b).

Grafo	М	atriz	de gi	rau	Ν	1atri	z de	adjao	cência	a	Mat	riz la	placi	ana	
3 <sup>2</sup> 4 <sup>5</sup> 2 7 1	7 0 0 0	0 5 0 0	0 0 9 0	0 0 0 7		0 0 7 0	0 0 0 5	7 0 0 2	0 5 2 0		7 0 -7 0	0 5 0 -5	-7 0 9 -2	0 -5 - 2 7	

Figura 4.6: Cálculo da matriz laplaciana.

Uma vez que a matriz laplaciana é calculada, o próximo passo é determinar seus autovetores e autovalores, pois eles capturam várias propriedades do grafo original [13]. Uma das propriedades encontradas é em relação a conectividade do grafo que é definida pelo segundo menor autovalor da matriz laplaciana: se esse autovalor for positivo, o grafo estará conectado [36, 47]. Os autovetores calculados a partir da matriz laplaciana podem ser organizados em uma nova matriz, onde as linhas representam pontos em um espaço vetorial — que, por sua vez, representam as estações do grafo original — e as colunas formadas pelos autovetores representam as coordenadas das estações [22].

Supondo que o grafo original esteja conectado, exatamente um dos autovalores será zero [15]. O segundo menor autovalor — que é positivo para um grafo conectado [18] — é conhecido como valor de Fiedler. O valor de Fiedler se aproxima do corte mínimo necessário do grafo para separá-lo em dois componentes. O autovetor associado a este segundo autovalor pode ser usado para particionar o grafo original em dois grupos. De forma mais geral, os autovetores associados aos k-1 menores autovalores não nulos podem ser usados para particionar o grafo em k clusters [40, 53].

Após ser feita a decomposição espectral, o algoritmo K-Means é executado para o agrupamento das estações. O conjunto de dados de entrada para o K-Means é um subconjunto da matriz de autovetores da matriz laplaciana. Para obter esse subconjunto, apenas os k - 1 autovetores associados aos k - 1 menores autovalores não nulos são selecionados. A Figura 4.7 mostra um exemplo dos autovalores com seus respectivos autovetores calculados a partir da matriz laplaciana, cada autovalor possui 6 autovetores. Em um exemplo considerando 3 *clusters*, com exceção do autovalor nulo, os dois autovalores marcados em azul são os dois menores autovalores, e os seus respectivos autovetores, também marcados em azul, devem ser organizados em uma nova matriz para serem utilizados como entrada no algoritmo K-Means.

Autovetores					
-0,4082483	-0,19067	-0,30212403	-0,14014626	0,38850398	0,73152105
-0,4082483	0,09380835	0,70072259	-0,51041662	0,23356282	0,13581039
-0,4082483	-0,10966983	0,198989	0,12463272	-0,82639351	0,28853022
-0,4082483	-0,33466188	0,18110579	0,72430906	0,30663967	-0,26435571
-0,4082483	-0,31437074	-0,53448682	-0,38203205	-0,12965797	-0,53485374
-0,4082483	0,8555641	-0,24420653	0,18365314	0,02734501	-0.08503141

Autovetores

Autovalores

0	1307,297415	1188,140832	990,6603398	786,788537	830,08669102

Figura 4.7: Exemplo de autovalores e autovetores.

#### 4.2.2.2 Função de pós-processamento

No final do processamento do K-Means, os grupos nem sempre serão balanceados — i.e., com o mesmo número de estações em cada grupo. Porém, para o mecanismo RAW, é importante que os grupos sejam balanceados, pois agrupamentos altamente desbalanceados levam a grandes grupos com alta competição pelo uso do meio e, portanto, mais colisões [21]. Por este motivo, nossa proposta inclui uma função de pós-processamento onde o balanceamento é realizado movendo certas estações de *clusters* maiores — referidos como *doadores* — para os menores — referidos como *receptores*. Observe, no entanto, que essa tarefa não é trivial, pois não se deve mover um nó para um determinado *cluster* se outras estações neste *cluster* estiverem ocultas a ele — pois isso também levaria a mais colisões. Assim, em algumas situações não será possível ter um equilíbrio perfeito devido à possibilidade de criação de terminais ocultos.

Para evitar a inclusão de terminais ocultos nos grupos, no pós-processamento, um nó só é transferido se não estiver oculto a nenhum outro nó do *cluster* receptor. Caso contrário, outro nó neste grupo doador é selecionado e a mesma verificação é executada novamente. Esse procedimento é repetido até que um nó do *cluster* doador seja eleito para migrar para o *cluster* receptor ou não haja mais estações para analisar no *cluster* doador. As migrações tentam equilibrar os *clusters* para que atinjam o número ideal de estações ou a melhor distribuição possível, evitando a criação de terminais ocultos. No entanto, a ordem de realização das trocas pode fazer diferença nessa distribuição, e nem sempre será possível obter o melhor equilíbrio.

O Algoritmo 2 detalha este procedimento de pós-processamento. O algoritmo começa com o cálculo do número ideal de estações por *cluster*, que é obtido dividindo o número total de estações pelo número de *clusters*. Em seguida, as tentativas de transferência de estações começam, dos *clusters* doadores (que possuem mais estações do que a quantidade ideal) para os *clusters* receptores (que possuem menos estações do que a quantidade ideal). Para cada *cluster* receptor, todos os *clusters* doadores são atravessados. Para que um nó seja considerado eleito, ele não pode ser um terminal oculto para as estações do *cluster* receptor. Portanto, todas as arestas entre o nó doador e as estações do *cluster* receptor serão verificadas. Se houver pelo menos uma aresta com peso abaixo do *CCAThres* (*Clear Channel Assessment Threshold* usado pelo IEEE 802.11), caracterizando um terminal oculto, então o nó doador não será elegível, e o algoritmo irá mover para o próximo nó no agrupamento de doadores. Caso contrário, o nó doador não é um terminal oculto e, portanto, será transferido para o *cluster* receptor. As tentativas de balanceamento continuam até que todos os *clusters* receptores e doadores sejam percorridos.

Os detalhes da implementação do algoritmo K-Means Espectral e da função de pósprocessamento são descritos no Capítulo 5.

Algoritmo 2: Função de pós-processamento.				
Entrada: nos (número de estações da rede), clusters (número de clusters),				
<i>clustersFormados</i> (saída do K-Means).				
Saída: uma versão atualizada e mais equilibrada dos <i>clusters</i> .				
$quantidadeIdeal \leftarrow nos \div clusters$				
for each receptor $\in$ clustersFormados do				
$eleito \leftarrow true$				
while receptor. NumDeEstações $<$ quantidadeIdeal do				
foreach $doador \in clustersFormados$ do				
$\mathbf{h} = \mathbf{h} + \mathbf{h} + \mathbf{h}$				
<b>foreach</b> $dNo \in doador$ <b>do</b>				
<b>foreach</b> $rNo \in receptor$ <b>do</b>				
$\prod_{i=1}^{n} (aNO, FNO) \in E ana \ w(aNO, FNO) < CCAIntes then$				
break				
end				
end				
if eleito then				
move $dNo$ to receptor				
break				
$ $ $ $ end				
$ $ $ $ end				
end				
end				
if não eleito then				
break; // impossível mover qualquer nó				
end				
end				
end				

# Capítulo 5

## Implementação Prática

A primeira estratégia de agrupamento RAW baseada na clusterização K-Means foi implementada usando a informação das posições geográficas das estações, de acordo com o Algoritmo 1 apresentado no Capítulo 4. Para realizar os testes, foi implementado um simulador de redes sem fio para verificar se o desempenho obtido com a estratégia definida é melhor do que o agrupamento RAW. Os programas criados nesta primeira estratégia foram reutilizados para a segunda, com os devidos ajustes.

Este capítulo apresenta os detalhes da implementação da estratégia baseada na clusterização espectral, de acordo com as duas etapas explicadas no Capítulo 4: K-Means Espectral e o pós-processamento. Além disso, são apresentados os detalhes da implementação do simulador e dos procedimentos de simulação. Todos os programas foram implementados em Python, versão 2.7.

### 5.1 Simulador

Para a realização dos testes das estratégias de agrupamento, foi utilizado um programa simulador de redes sem fio. O simulador foi desenvolvido em Python especificamente para esse trabalho<sup>1</sup>. O programa foi desenvolvido para simular uma rede utilizando o mecanismo RAW de forma parametrizável. Pode ser simulado sem usar nenhuma forma de agrupamento, com o agrupamento arbitrário (RAW) ou grupos pré-definidos. Desta forma, pode-se usar os grupos criados pelas estratégias de agrupamento propostas neste trabalho para substituir os grupos arbitrários do RAW e observar o comportamento da rede com esta formação alternativa.

 $<sup>{}^1\</sup>rm Disponível\ em:\ https://github.com/BrunaToledoGuedes/ProgramasDissertacao.$ 

Além de definir a forma ou ausência de agrupamento, existem outros parâmetros básicos de entrada utilizados nas simulações que são: o número de estações da simulação, o número de grupos, a altura e largura do cenário e o valor da semente. Desta forma é possível configurar o simulador para vários ambientes distintos de uma rede sem fio. Estes parâmetros são mostrados na Tabela 5.1.

Parâmetros	Descrição
-n	Número de estações na simulação
-g	Número de grupos na simulação
-H/-W	Altura e largura do cenário em metros
-S	Entrada do gerador de número aleatório
-pE	Gera arquivo que contém os links com os RSSI
-pP	Gera arquivo que contém as coordenadas das estações
-G	Leitura do arquivo gerado pela estratégia de agrupamento

Tabela 5.1: Principais parâmetros do simulador.

As primeiras simulações foram feitas para testar a primeira estratégia de agrupamento, na qual utiliza-se o K-Means considerando a posição geográfica das estações. Essa informação é gerada pelo simulador, que exporta um arquivo texto, lido pelo programa implementado com o algoritmo K-Means.

Para a segunda estratégia de agrupamento, o programa que implementa a clusterização espectral precisa receber informações do RSSI. No simulador, essas informações entre enlaces que conectam as estações são obtidas através de um modelo Two-Ray Ground [46]. Um arquivo é gerado contendo as relações das informações RSSIs entre as estações. Esse arquivo representa o grafo e é usado pelo algoritmo de clusterização para a construção da matriz de adjacência.

Com o simulador, foi possível realizar os testes observando o comportamento específico de uma rede sem fio utilizando o mecanismo RAW, com os grupos sendo criados de forma arbitrária, considerando as suas posições geográficas ou considerando os RSSIs. A única variável diferente entre as simulações de uma determinada semente está no tipo de agrupamento, indicando que os resultados obtidos nas simulações refletem somente a forma de agrupar as estações.

### 5.2 Visão Geral do Procedimento de Simulação

Utilizando o simulador descrito na seção anterior, foi possível criar uma metodologia de avaliação. Através do simulador foi possível analisar o desempenho da rede. O fluxograma do processo de execução de uma instância de simulação é mostrado na Figura 5.1. Neste fluxo, são apresentadas as atividades realizadas para medir o desempenho de uma rede utilizando o mecanismo RAW. As simulações são feitas com agrupamento arbitrário e estratégias baseadas no K-Means. Primeiro, deve ser executado o simulador com agrupamento arbitrário. Além de gerar o log da simulação, a execução gera também um arquivo contendo as posições geográficas das estações ou o RSSI dos enlaces que conectam essas estações, dependendo da estratégia a ser comparada. Na segunda atividade é executado o algoritmo com a respectiva estratégia de agrupamento baseada no K-Means e é gerado um arquivo com os grupos e as respectivas estações. Em seguida, uma nova simulação da rede sem fio é executada, desta vez considerando os agrupamentos criados na atividade anterior. Por último, um programa que faz a leitura dos logs da simulação é executado, para cada log gerado, e produz os resultados, com as taxas de colisão, atraso e vazão.

Para fazer a simulação com agrupamento arbitrário, é necessário passar os seguintes parâmetros do ambiente: o número de estações, o número de grupos, o valor da semente, o tamanho do cenário em metros e o nome do arquivo de saída que conterá a posição geográfica ou as informações RSSI. Para simular a rede com agrupamento baseado no K-Means, a única diferença é que, no lugar do arquivo de saída, deverá ser informado o nome do arquivo de entrada, com seu respectivo parâmetro. Este arquivo contém os grupos e estações criados pelo K-Means.



Figura 5.1: Fluxo do processo de execução de uma instância de simulação.

### 5.3 Algoritmo de K-Means Espectral

O algoritmo K-Means Espectral é utilizado para obter o agrupamento das estações da rede de acordo com o RSSIs. A Tabela 5.2 mostra os principais parâmetros que devem ser passados ao executar esse algoritmo. Os principais parâmetros de entrada são o número de estações, o número de grupos, a dimensão do cenário e o arquivo de saída a ser gerado pelo simulador de rede que contém os relacionamentos entre as estações, com as informações RSSI dos enlaces que conectam as estações.

Parâmetros	Descrição
-n	Número de estações na simulação
-g	Número de grupos na simulação
-H/-W	Altura e largura do cenário em metros
-p	Leitura do arquivo que contém os enlaces com os RSSI
-c	Leitura do arquivo que contém as coordenadas das estações

Tabela 5.2: Principais parâmetros do K-Means Espectral.

No algoritmo K-Means Espectral, é feita a decomposição espectral para o tratamento das informações RSSI. Primeiro, é calculada a representação do grafo de rede na forma dos autovetores e autovalores da matriz laplaciana, conforme descrito na Subseção 4.2.2.1. A biblioteca Numpy, que é amplamente usada para trabalhar com vetores multidimensionais, facilitou a implementação desta etapa. A biblioteca Numpy foi utilizada para construir as matrizes de adjacência, grau, laplaciana, e para calcular os autovalores e autovetores a partir da matriz laplaciana.

Após esse passo, um subconjunto dos autovetores é selecionado para particionar o grafo em K clusters. Para particionar o grafo em k clusters são selecionados os autovetores associados aos k - 1 menores autovalores.

A matriz formada pelos autovetores associados aos k - 1 menores autovalores é um dos parâmetros de entrada para a execução do algoritmo K-Means na segunda etapa. As linhas da matriz representam as estações do grafo original e suas colunas representam as coordenadas das estações no espaço vetorial [22]. Foi utilizada uma implementação de código aberto do K-Means<sup>2</sup>. A escolha por esta implementação se deu por conta da sua clareza, muito bem documentada, facilitando a realização dos ajustes necessários para alcançar os objetivos da estratégia baseada na clusterização espectral. Após o K-Means, será executada a função de pós-processamento, conforme descrito na Subseção 4.2.2.2, para equilibrar o número de estações nos *clusters*, finalizando a execução do algoritmo

 $<sup>^2</sup> Disponível \ em: \ \texttt{https://gist.github.com/PSJoshi/c3da7bb6b5643eabe8ae}.$ 

#### K-Means Espectral.

# Capítulo 6

## Avaliação de Desempenho

Este capítulo descreve as simulações que foram realizadas com o objetivo de analisar e comparar o desempenho dos algoritmos baseados no K-Means em relação ao método de agrupamento padrão usado no IEEE 802.11ah. Também foram feitas simulações sem utilizar nenhum tipo de agrupamento. O desempenho do K-Means Espectral e K-Means utilizando as posições geográficas, que foi chamado de K-Means Original, também foi comparado com o algoritmo da literatura HMR [64]. As principais métricas de comparação utilizadas foram as taxas de colisão, atraso e vazão. Vários cenários com diferentes níveis de densidade de estações foram avaliados, resultando em diferentes proporções de ocorrência de terminais ocultos. Cenários com diferentes níveis de densidade foram simulados com o objetivo de analisar o desempenho das estratégias de agrupamento em relação as métricas avaliadas à medida que a quantidade de estações aumenta.

### 6.1 Descrição das simulações

A rede foi simulada usando um programa Python que foi desenvolvido com base no framework SimPy, mencionado na Subseção 5.1. A Tabela 6.1 resume os principais parâmetros de simulação. Considerou-se uma arquitetura de rede infraestruturada com o ponto de acesso no centro do cenário. As simulações foram feitas em cenários com tamanho de  $280m \times 280$  m.

As estações são posicionadas aleatoriamente no cenário. Assumiu-se o uso de canais de 2 MHz e todas as interfaces usam taxa de transmissão de 650 kb/s. Cada ponto dos gráficos corresponde à média de 10 repetições da mesma simulação com sementes aleatórias diferentes. Isso nos permitiu calcular intervalos de confiança de 95%, que são mostrados em todos os gráficos como barras de erro — embora, em alguns casos, as barras de erro não sejam visíveis devido ao seu pequeno tamanho. O número total de estações varia para cada simulação, entre 90 até 1000 estações. Nas simulações, foi considerada uma rede estática: as estações não mudam de posições e os valores de RSSI entre as estações permanecem os mesmos durante cada simulação.

Parâmetros	Valores
Arquitetura da rede	Número de estações na simulação
Dimensão do cenários	$280m \times 280m$
Posição do AP	Centro do cenário
Números de slots RAW	5, 10 e 15
Taxa de dados da interface	$650 \mathrm{~kb/s}$
Tamanho do pacote	520 B
Tempo da simulação	20 s
Repetições	10
Intervalo de confiança	95%

Tabela 6.1: Principais parâmetros da simulação

#### 6.2 Resultados de simulações

Nesta seção são mostradas as simulações realizadas com o objetivo de analisar o desempenho dos algoritmos K-Means Espectral e K-Means Original em relação ao método de agrupamento padrão usado no IEEE 802.11ah e ao algoritmo HMR. Também foram feitas simulações sem utilizar nenhum tipo de agrupamento.

#### 6.2.1 Comparação dos algoritmos de clusterização com o agrupamento arbitrário

As simulações mostradas a seguir foram feitas para comparar o desempenho dos algoritmos K-Means Espectral, K-Means Original com o agrupamento arbitrário. As métricas de comparação utilizadas foram as taxas de colisão, atraso e vazão. A Figura 6.1 mostra as taxas de colisão em função do número de estações associadas ao ponto de acesso para grupos com tamanhos  $k \in \{5, 10, 15\}$ . Os agrupamentos baseados no K-Means Original e K-Means Espectral tiveram um desempenho melhor do que o agrupamento arbitrário, apresentando menores taxas de colisão. Essa diferença é acentuada em cenários com menos estações. Para um exemplo extremo, com 100 estações e 15 grupos, o agrupamento arbitrário resultou em uma taxa de colisão de 67%, enquanto o K-Means Original resultou em 7% e o K-Means Espectral resultou em 9%. À medida que o número de estações



Figura 6.1: Taxa de colisão versus número de estações.

aumenta, o número de grupos disponíveis torna-se insuficiente para lidar de forma satisfatória com o nível de competição, independentemente da estratégia de agrupamento. Como consequência, a diferença de desempenho é reduzida.

Essa grande diferença de desempenho entre o agrupamento arbitrário e as estratégias baseadas em K-Means é esperada. Com o agrupamento arbitrário, duas estações podem estar no mesmo grupo, mas muito distantes uma da outra. Quando uma dessas estações começa a transmitir, o sinal pode ser recebido pela outra estação abaixo do limiar de detecção de portadora. Assim, a segunda estação julga que o meio está livre e pode começar a transmitir, causando uma colisão.

Também é importante destacar que o K-Means Espectral teve um desempenho muito semelhante ao K-Means Original para a maioria dos cenários avaliados. Diferenças de desempenho significativas entre as duas abordagens só podem ser percebidas quando apenas cinco grupos estavam disponíveis, o que torna mais difícil evitar que terminais ocultos sejam colocados nos mesmos grupos. Isso é digno de nota porque o K-Means Original tem à sua disposição as posições geográficas precisas das estações, uma informação que normalmente não está disponível.

Os benefícios dos agrupamentos baseados no K-Means também podem ser vistos nas Figuras 6.2, 6.3 e 6.4. Elas mostram os atrasos médios do pacote com cada estratégia de



Figura 6.2: Média de atraso do pacote versus número de estações. Simulações feitas com 5 grupos.

agrupamento para grupos com tamanhos  $k \in \{5, 10, 15\}$ , respectivamente, bem como sem RAW.

Este atraso do pacote é medido a partir do momento em que o pacote é gerado na estação de origem até o momento em que é recebido com sucesso pelo ponto de acesso. Os agrupamentos baseados no K-Means Original e espectral resultaram em um melhor desempenho com menos atraso em relação ao agrupamento arbitrário, exceto para o gráfico da Figura 6.2, em que são usados 5 grupos. Neste caso os três tipos de agrupamentos não precisam esperar muito para poder transmitir, porque são poucos grupos, diminuindo o atraso para os três tipos de agrupamento em relação aos gráficos das Figuras 6.3 e 6.4.

Quando a quantidade de grupos aumenta para 10 e 15, nos gráficos das Figuras 6.3 e 6.4, os três tipos de agrupamento precisam esperar mais tempo para transmitir, aumentando assim o atraso. Para esses dois gráficos, a diferença do atraso aumentou entre o K-Means Original e K-Means Espectral em relação ao arbitrário, porque, quando foram usados 10 e 15 grupos, tanto para o K-Means Original quanto o espectral foram apresentadas taxas de colisão mais baixas em relação ao agrupamento arbitrário. Isso contribui para um menor atraso, pois, quando ocorrem colisões, todo o tempo de transmissão do pacote é perdido. Além disso, as colisões levam a tempos de *backoff* no CSMA/CA,



Figura 6.3: Média de atraso do pacote versus número de estações. Simulações feitas com 10 grupos.



Figura 6.4: Média de atraso do pacote versus número de estações. Simulações feitas com 15 grupos.

causando ainda mais atrasos.

Por outro lado, quando não foi utilizado nenhum agrupamento (Sem RAW), o atraso foi menor em relação a todas as variantes RAW, mesmo com níveis de colisão mais elevados. Nesse caso, a ausência das restrições rígidas impostas pelo RAW no momento em que cada estação pode tentar transmitir dá uma grande vantagem quando não é usado nenhum tipo de agrupamento nesta métrica. As estações não precisam esperar pela vez de seu grupo para fazer as transmissões, causando uma redução significativa no atraso dos poucos quadros que foram realmente entregues. Tal como acontece com os resultados mostrados na Figura 6.1, foi visto novamente que o uso de informações geográficas precisas pelo K-Means Original não se traduz em melhor desempenho em comparação com o K-Means Espectral. Em vez disso, o K-Means Espectral supera o K-Means Original conforme o número de estações aumenta. Isso pode ser explicado pelo procedimento de pós-processamento utilizado com o K-Means Espectral, que resulta na redução de terminais ocultos dentro dos *clusters*, ao migrar as estações de um grupo para outro.

Como outro experimento, as simulações foram repetidas, mas agora variando a carga da rede, manipulando a taxa de geração de pacotes em cada estação. A Figura 6.5 mostra uma comparação da vazão agregada obtida em função da carga oferecida da rede. Nessas simulações, foram utilizados 100 estações e 15 grupos.

Para cargas de rede extremamente baixas, todas as estratégias se comportam de maneira semelhante. Mesmo sem RAW, a rede é perfeitamente capaz de suportar a carga oferecida. No entanto, à medida que a carga aumenta, quando não é usado nenhum agrupamento (Sem RAW) não é possível sustentar o desempenho e as abordagens baseadas em RAW mostram suas vantagens. Nesse caso, os agrupamentos baseados em K-Means Original e Espectral resultam em melhor capacidade da rede devido à redução de colisões entre terminais ocultos (Figura 6.1) e menor atraso (Figura 6.4), atingindo uma vazão superior em comparação ao agrupamento arbitrário. Essa diferença da vazão entre as abordagens arbitrária e baseada em K-Means torna-se significativa para todas as cargas oferecidas acima de 1 Mb/s e é maximizada para uma carga de 100 Mb/s, quando o K-Means Espectral atinge uma vazão 35,8% mais alta. Comparando as duas abordagens baseadas em K-Means, novamente foi obtido um desempenho muito semelhante para todas as cargas.



Figura 6.5: Vazão agregada da rede versus carga oferecida da rede.

#### 6.2.2 Efeito da função de pós-processamento do K-Means Espectral

Além do algoritmo K-Means Espectral proposto não precisar das posições geográficas das estações, outra vantagem desse algoritmo é a utilização da função de pós-processamento para balancear os grupos. Para quantificar os efeitos desse pós-processamento nos resultados, os experimentos anteriores foram repetidos, mas agora executando simulações com o agrupamento espectral sem pós-processamento e comparando com os resultados obtidos com o pós-processamento.

A Figura 6.6 mostra o efeito nas taxas de colisão da função de pós-processamento. A figura também mostra os resultados obtidos pelo K-Means Original e agrupamento arbitrário para referência. As simulações foram realizadas com 15 grupos. Como esperado, o K-Means Espectral com pós-processamento alcançou uma taxa de colisão mais baixa do que o K-Means Espectral sem o pós-processamento. Isso ocorre porque os grupos ficam mais equilibrados após o pós-processamento. No cenário de 200 estações e 15 grupos, por exemplo, o número ideal de estações em cada *cluster* é de aproximadamente 13. Sem o pós-processamento, o tamanho dos grupos gerados pelo K-Means Espectral teve um desvio padrão de 12,9. Após o pós-processamento, esse desvio padrão foi reduzido para 2,1.



Figura 6.6: Taxa de colisão *versus* número de estações. Comparação entre o desempenho do K-Means Espectral com e sem pós-processamento, K-Means Original, assim como com o agrupamento arbitrário.

Tal como acontece com todas as outras estratégias de agrupamento, a diferença entre usar ou não o pós-processamento torna-se menor à medida que aumenta o número de estações, uma vez que o número fixo de grupos disponíveis torna-se insuficiente para reduzir efetivamente o nível de competição e evitar a presença de terminais ocultos em cada grupo. O K-Means Espectral sem pós-processamento obteve taxas de colisões maiores em relação ao K-Means Original para cenário com até 700 estações por gerar grupos menos equilibrados que o K-Means Original. O K-Means Espectral com pós-processamento obteve taxas de colisões semelhantes em relação ao K-Means Original, mesmo não usando as posições precisas das estações, por gerar grupos mais equilibrados, diminuindo a competição entre as estações no momento das transmissões dos quadros e consequentemente reduzindo o número de colisões.

Na Figura 6.6, deve-se notar também que, mesmo sem o pós-processamento, o K-Means Espectral é capaz de reduzir bastante as colisões com relação ao agrupamento arbitrário. Aqui, novamente, isso pode ser explicado pela suscetibilidade do agrupamento arbitrário a colocar estações ocultas dentro de um mesmo grupo, enquanto o agrupamento espectral tende a separá-los. Apesar de resultar em grupos perfeitamente equilibrados, a presença de terminais ocultos intragrupo aumenta severamente as colisões com o agrupamento arbitrário.

A Figura 6.7 mostra o efeito do pós-processamento em relação à vazão agregada da rede. As simulações foram realizadas com 100 estações e 15 grupos. Os resultados da vazão são consistentes principalmente com os dados de colisão mostrados na Figura 6.6.



Figura 6.7: Vazão agregada da rede *versus* carga oferecida da rede. Comparação entre o desempenho do K-Means Espectral com e sem pós-processamento, assim como com o agrupamento arbitrário.

Com o pós-processamento, o agrupamento espectral atinge maior rendimento, pois consegue balancear as estações entre os grupos, evitando grupos com (proporcionalmente) muitas estações e, consequentemente, diminuindo a taxa de colisão. Portanto, mais pacotes serão entregues, resultando em uma vazão mais alta em comparação com K-Means sem pós-processamento. No entanto, à medida que a carga aumenta, a rede fica saturada e a vazão para de aumentar. Nesse ponto, a diferença entre usar ou não o pós-processamento torna-se menor – embora existente. Também de forma semelhante ao comportamento representado na Figura 6.6, mesmo sem pós-processamento, o agrupamento espectral é capaz de quase sempre superar o agrupamento arbitrário, como mostra a Figura 6.7. Conforme a carga da rede aumenta, os ganhos de ambas as versões do K-Means Espectral tornam-se substanciais devido à menor probabilidade de terminais ocultos no mesmo *cluster*. Isso compensa consideravelmente qualquer vantagem possível que o agrupamento arbitrário possa ter devido ao seu equilíbrio inerente.

### 6.2.3 Comparação do algoritmo K-Means Espectral com o algoritmo HMR

No Capítulo 3 foi descrito o algoritmo HMR [64], uma solução da literatura para diminuir os terminais ocultos com o objetivo de melhorar o desempenho da rede. Neste algoritmo, a informação do quadro PS-POLL é usada para verificar se as estações são terminais ocultos entre si. As estações enviam esse quadro PS-POLL para o ponto de acesso no momento em que elas estão preparadas para receber os dados [3]. Para comparar o desempenho dos algoritmos K-Means Original e espectral com o algoritmo HMR, em relação às métricas de colisão, atraso e vazão, foi implementada uma versão do algoritmo HMR levando em considerando a mesma forma de identificação dos terminais ocultos da nossa proposta, na qual foi usado como critério o limiar de detecção de portadora. As Tabelas 6.2 e 6.3 mostram as quantidades de terminais ocultos após a execução do agrupamento arbitrário e dos algoritmos HMR e K-Means Espectral (depois do pós-processamento) para 100 estações, considerando 15 grupos e para 400 estações, considerando 15 grupos, respectivamente.

	100 estações					
Cluster	Agrupamento arbitrário (RAW)	Algoritmo HMR	Clusterização espectral			
0	3	4	0			
1	6	4	0			
2	5	1	0			
3	2	0	0			
4	4	2	0			
5	4	2	0			
6	4	0	0			
7	4	0	0			
8	8	2	0			
9	7	1	0			
10	3	0	0			
11	4	0	0			
12	6	0	0			
13	3	0	0			
14	3	0	0			
Total	66	16	0			

Tabela 6.2: Tabela da quantidade de terminais ocultos de simulações com 100 estações e 15 grupos.

	400  estações					
Cluster	Agrupamento arbitrário (RAW)	Algoritmo HMR	Clusterização espectral			
0	93	166	0			
1	73	85	0			
2	64	37	0			
3	70	20	0			
4	72	14	0			
5	124	52	0			
6	93	18	0			
7	62	26	0			
8	83	12	0			
9	60	0	0			
10	95	0	0			
11	90	4	0			
12	39	0	0			
13	67	0	0			
14	74	5	0			
Total	1159	439	0			

Tabela 6.3: Tabela da quantidade de terminais ocultos de simulações com 400 estações e 15 grupos.

Como não leva em consideração nenhum critério para evitar que estações ocultas fiquem no mesmo grupo, o agrupamento arbitrário foi o que resultou no maior número de terminais ocultos. Ao executar o algoritmo HMR, houve uma grande redução dos terminais ocultos em relação ao agrupamento arbitrário. A diminuição dos terminais ocultos, no entanto, foi ainda maior com a execução do algoritmo K-Means Espectral, conseguindo remover todos os terminais ocultos dos cenários com 100 e 400 estações. Isso acontece porque este último organiza as estações considerando sua relação de conectividade de rádio.

O resultado do K-Means Espectral é mais eficiente do que do algoritmo HMR, também utilizado para eliminar os terminais ocultos. Isso acontece porque raramente o algoritmo HMR conseguirá fazer um remanejamento eficiente, por conta da distribuição arbitrária, que é o ponto de partida deste algoritmo. O algoritmo HMR faz a movimentação das estações entre os grupos para diminuir o número de terminais ocultos a partir do agrupamento feito pelo mecanismo RAW. Por ser um agrupamento arbitrário, os grupos formados podem possuir muitos terminais ocultos. Isso dificulta a eficiência do algoritmo em tentar reduzir esses terminais ocultos. O algoritmo K-Means Espectral utiliza o algoritmo tradicional do K-Means para fazer a distribuição das estações pelos grupos levando em considerando a informação RSSI, por isso consegue diminuir uma quantidade maior de terminais ocultos em relação ao algoritmo HMR.

A Figura 6.8 mostra as taxas de colisão em função do número de estações associadas ao ponto de acesso para grupos com tamanhos  $k \in \{5, 10, 15\}$ . O algoritmo K-Means Espectral obteve taxas de colisão mais baixas em relação ao algoritmo HMR para 100, 400 e 600 estações. Isso é uma consequência direta da sua menor quantidade de terminais ocultos em relação ao algoritmo HMR. Para 400 estações e 15 grupos, por exemplo, o algoritmo HMR resultou em 439 pares de terminais ocultos, enquanto o K-Means Espectral conseguiu eliminar todas as ocorrências. Em decorrência da diminuição dos terminais ocultos, há uma queda substancial na taxa de colisão em relação ao algoritmo HMR.



Figura 6.8: Taxa de colisão *versus* número de estações. Comparação entre o desempenho do HMR e K-Means Espectral.

As Figuras 6.9 e 6.10 mostram a vazão agregada em função da carga oferecida da rede. As simulações do gráfico da Figura 6.9 foram feitas com 100 estações e 15 grupos, enquanto as da Figura 6.10 foram feitas com 400 estações e 15 grupos. Para a maioria dos casos nos dois gráficos, o algoritmo K-Means Espectral atingiu uma vazão maior em relação ao algoritmo HMR devido à diminuição dos terminais ocultos – já que mais pacotes são entregues com sucesso, aumentando a vazão. Porém, quando foram usadas 400 estações, a diferença da vazão foi substancialmente maior entre o algoritmo HMR e o K-Means Espectral. Conforme o número de estações aumenta, a eficiência do agrupamento feito pelo HMR piora, comparado ao algoritmo K-Means Espectral. À medida que a



Figura 6.9: Vazão agregada da rede *versus* carga oferecida da rede. Comparação entre o desempenho do HMR e K-Means Espectral nas simulações com 100 estações.



Figura 6.10: Vazão agregada da redeversus carga oferecida da rede. Comparação entre o desempenho do HMR e K-Means Espectral nas simulações com 400 estações.

quantidade de estações aumenta, a quantidade de terminais ocultos nos grupos também aumenta quando o algoritmo HMR foi executado. O K-Means Espectral conseguiu formar grupos com uma quantidade menor de terminais ocultos em relação ao HMR devido ao agrupamento ser feito utilizando o RSSI.

A Figura 6.11 mostra o atraso médio dos pacotes com cada estratégia de agrupamento para 100, 400, 600 estações e 15 grupos bem como sem RAW.



Figura 6.11: Média de atraso do pacote *versus* número de estações. Comparação entre o desempenho do HMR e K-Means Espectral.

O algoritmo K-Means Espectral obteve atraso menor em relação ao algoritmo HMR para todas as quantidades de estações. No gráfico da Figura 6.8, a maior diferença da taxa de colisão entre o K-Means Espectral e o algoritmo HMR foi para as simulações em que foram usadas 400 estações. Em relação à métrica de atraso, as simulações usando 400 estações também resultaram em uma diferença maior do atraso para esses dois algoritmos. Como as taxas de colisão do K-Means Espectral foram menores em relação ao HMR, menos pacotes serão retransmitidos devido às perdas causadas pelas colisões e, dessa forma, o atraso será menor também.

#### 6.2.4 Impacto do limiar de detecção de portadora

Considerando os modelos adotados pelo simulador utilizado, o limiar de sensibilidade do rádio (*i.e.*, a potência mais baixa para que um sinal recebido pelo rádio seja notado) é de -94 dBm. Ainda considerando os parâmetros de propagação adotados no simulador, para cenários com dimensões de 280 m  $\times$  280 m, as estações que estão a uma distância de 280 metros recebem um sinal com potência de -75 dBm. A distância máxima (diagonal do cenário) em que uma estação pode estar da outra é de aproximadamente 396 metros. Para esta distância, o sinal recebido será de -81 dBm. Assim, para esse cenário, independentemente da distância entre duas estações, o sinal recebido será sempre acima do limiar de sensibilidade de -94 dBm. Isso significa que, na prática, não existem terminais ocultos — em termos físicos. Entretanto, o limiar de detecção de portadora do IEEE 802.11ah é de -70 dBm, o que cria terminais ocultos "lógicos" — *i.e.*, nós cuja transmissão é fisicamente percebida, mas ignorada —, aumentando a probabilidade de colisões (com o objetivo de, possivelmente, obter ganhos em termos de reuso espacial).

Para entender como o limiar de detecção de portadora afeta o desempenho da rede e se relaciona com as estratégias de agrupamento RAW, diferentemente das simulações anteriores feitas com o limiar de detecção de portadora de valor -70 dBm definido pelo padrão IEEE 802.11ah, nessa seção são apresentados resultados de simulações feitas com o limiar de valor -94 dBm.

As Figuras 6.12 e 6.13 mostram os resultados da taxa de colisão e vazão. No gráfico da Figura 6.12, são usados grupos com tamanhos  $k \in \{5, 10, 15\}$ , enquanto no gráfico da Figura 6.13 são usados 100 estações e 15 grupos. Como as estações conseguem perceber as transmissões que estão sendo realizadas e dessa forma também conseguem evitar as colisões, o desempenho do K-Means Espectral e do agrupamento arbitrário foram parecidos em relação às taxas de colisão e a vazão agregada da rede nos gráficos das Figuras 6.12 e 6.13. No gráfico da Figura 6.13, o agrupamento arbitrário consegue atingir uma vazão um pouco maior que o K-Means Espectral. Como neste caso não existem terminais ocultos, o melhor agrupamento possível é um agrupamento que seja totalmente balanceado, ou seja, que tenha o mesmo número de estações em cada grupo. Nesse objetivo particular, a natureza round-robin do agrupamento arbitrário sempre consegue o melhor desempenho.

As Figuras 6.14 e 6.15 mostram os resultados das simulações feitas com o cenário de dimensão de 1500 m  $\times$  1500 m usando o mesmo limiar de detecção de portadora com valor de -94 dBm.



Figura 6.12: Taxa de colisão *versus* número de estações. Comparação entre o desempenho do K-Means Espectral e do agrupamento arbitrário para simulações com limiar de -94 dBm.

As métricas de comparação foram as mesmas utilizadas nos gráficos das Figuras 6.12 e 6.13 assim como a quantidades de grupos, e a única diferença foi a dimensão do cenário. Neste cenário de 1500 m × 1500 m, as estações podem receber o sinal abaixo do limiar -94 dBm. Isso significa que, na prática, existem terminais ocultos. Em particular, levando em conta os modelos e parâmetros das simulações, isso acontece para estações que estão a uma distância igual ou maior a 831 metros. Esses sinais são ignorados, e as estações não percebem que uma transmissão já está sendo realizada e transmitem também, causando colisões. Nesse caso, o K-Means Espectral realiza um agrupamento mais inteligente considerando o RSSI, evitando terminais ocultos em uma mesmo grupo. Dessa forma, as taxas de colisão do K-Means Espectral foram mais baixas em relação ao agrupamento arbitrário, como mostrado na Figura 6.14. Consequentemente, o K-Means Espectral também conseguiu um desempenho melhor em relação à vazão, como mostrado na Figura 6.15.

Duas principais conclusões podem ser tiradas desses resultados. Em primeiro lugar, chama a atenção o fato de que o descasamento entre o limiar de sensibilidade do rádio e o limiar de detecção de portadora prejudica bastante o desempenho do agrupamento arbitrário do RAW em cenários mais compactos. Por exemplo, no cenário de 280 m  $\times$  280 m com 100 nós e 15 grupos, quando o limiar de detecção de portadora é de -70 dBm (vide Figura 6.1), o agrupamento arbitrário resulta em taxas de colisão acima



Figura 6.13: Vazão agregada da redeversus carga oferecida da rede. Comparação entre o desempenho do K-Means Espectral e do agrupamento arbitrário para simulações com limiar de -94 dBm.



Figura 6.14: Taxa de colisão versus número de estações. Comparação entre o desempenho do K-Means Espectral e do agrupamento arbitrário nas simulações com limiar de -94 dBm e dimensão do cenário de 1500 m  $\times$  1500 m.



Figura 6.15: Vazão agregada da rede versus carga oferecida da rede. Comparação entre o desempenho do K-Means Espectral e do agrupamento arbitrário nas simulações com limiar de -94 dBm e dimensão do cenário de 1500 m  $\times$  1500 m.

de 60%. Por outro lado, alterando o limiar para -94 dBm e mantendo-se os demais parâmetros fixos, a taxa de colisão cai para cerca de 5% (vide Figura 6.12). Enquanto em redes pouco densas esse descasamento frequentemente auxilia no desempenho, reduzindo o impacto do problema do terminal exposto — i.e., permitindo certas transmissões que seriam erroneamente inibidas pela detecção de portadora do CSMA/CA —, em DWN, esse descasamento parece ser desaconselhável.

Por outro lado, esses resultados também evidenciam que apenas a alteração do limiar de detecção de portadora não é suficiente para garantir um bom desempenho do agrupamento arbitrário em todos os casos. A partir do momento em que a área de cobertura do AP cresce além do limite de sensibilidade do rádio e terminais ocultos — físicos surgem, um agrupamento apropriado se torna importante.

#### 6.2.5 Influência do aumento do cenário

As simulações mostradas a seguir foram feitas para analisar o desempenho do K-Means Espectral — em relação ao agrupamento arbitrário —, à medida que o tamanho do cenário aumenta. Para esses experimentos, foi usado o limiar de detecção de portadora de -94 dBm de forma a casar com o limiar de sensibilidade dos rádios. Os resultados que se seguem fazem referência ao número de terminais ocultos envolvendo estações que estão ambas no alcance do ponto de acesso. Essa contabilização foi feita dessa forma porque, para os cenários maiores, em alguns casos, o posicionamento aleatório das estações fez com que essas ficassem longe demais do ponto de acesso.

O gráfico da Figura 6.16 mostra a relação entre a quantidade de pares de terminais ocultos com conectividade com o ponto de acesso e a diferença da vazão entre o K-Means Espectral e o agrupamento arbitrário à medida que se aumenta a dimensão do cenário. As simulações foram realizadas com 100 estações e 15 grupos, em cenários quadrados variando de 200 m  $\times$  200 m a 2500 m  $\times$  2500 m. A vazão foi calculada para uma carga oferecida da rede de 10<sup>4</sup> kb/s. O gráfico mostra que, à medida que a quantidade de terminais ocultos que possuem conectividade com o ponto de acesso aumenta, a diferença da vazão entre o K-Means Espectral e o arbitrário aumenta também, possuindo dessa forma uma correlação forte positiva igual a 0,70.



Figura 6.16: Número de pares de estações *versus* dimensão do cenário. Relação entre a quantidade de terminais ocultos com conectividade com o ponto de acesso e a diferença da vazão entre K-Means Espectral e o arbitrário para o limiar -94 dBm.

As simulações com as dimensões entre 200 m  $\times$  200 m a 1200 m  $\times$  1200 m foram as que resultaram em uma diferença menor da vazão entre o K-Means Espectral e o arbitrário — inclusive, valores negativos, indicando superioridade do desempenho do agrupamento arbitrário. Esses cenários não possuem uma grande quantidade de terminais ocultos. Por este motivo, a distribuição mais equilibrada possível dos grupos será mais eficiente do que o agrupamento feito pelo K-Means Espectral para evitar os terminais ocultos nos grupos. O agrupamento arbitrário sempre distribui da forma mais balanceada possível. O K-Means Espectral nem sempre consegue o melhor balanceamento, mesmo com poucos terminais ocultos. As simulações feitas para os cenários entre 1250 m  $\times$  1250 m a 1600 m  $\times$  1600 m foram as que resultaram na maior diferença da vazão entre o K-Means Espectral e o arbitrário. Nestes cenários, existem grandes quantidades de pares de terminais ocultos com conectividade com o ponto de acesso. Logo, o K-Means Espectral consegue fazer uma melhor distribuição desses terminais ocultos nos grupos do que o arbitrário. Dessa forma, são feitas menos transmissões colidentes e a vazão aumenta.

Quando a dimensão do cenário começa a ficar muito grande, entre 1800 m  $\times$  1800 m a 2500 m  $\times$  2500 m, muitas estações passam a estar muito longe uma das outras, mas também do ponto de acesso, deixando de ter conectividade. Por conta dessa falta de conectividade com o ponto de acesso, esses pares, a rigor, deixam de ser terminais ocultos. Dessa forma, a distribuição eficiente dos terminais ocultos realizada pelo K-Means Espectral não faz diferença, porque os pacotes não serão entregues ao ponto de acesso. O agrupamento bem balanceado da abordagem arbitrária, por outro lado, faz diferença nesses cenários maiores, aumentando a vazão em relação ao K-Means Espectral, resultando em uma diferença negativa de vazão. Nestes cenários muito grandes, os pacotes enviados pelas estações que não possuem conectividade com o ponto de acesso são perdidos, seja por colisões ou por estarem longe demais do ponto de acesso. A distribuição equilibrada dos grupos feita pelo agrupamento arbitrário evita a formação de grupos com uma grande quantidade de estações, diminuindo a concorrência do acesso ao meio. Dessa forma o número de colisões diminui e a vazão aumenta.

No gráfico da Figura 6.17, as simulações foram feitas com os mesmos parâmetros do gráfico da Figura 6.16, exceto pelo fato de que o valor do limiar de detecção de portadora foi alterado para -70 dBm. O objetivo aqui é verificar como o desempenho relativo do K-Means Espectral e do agrupamento arbitrário se comporta com o aumento do tamanho do cenário, considerando o limiar definido pelo padrão.

Neste gráfico, a quantidade de terminais ocultos que possuem conectividade com o ponto de acesso e a diferença da vazão também possuem correlação forte positiva – aqui, com valor de 0,80. As simulações com cenários com dimensões de 300 m  $\times$  300 m a 1200 m  $\times$  1200 m resultaram em uma maior diferença na vazão entre o K-Means Espectral e o arbitrário. Nestes cenários, existe a possibilidade das estações receberem sinais com RSSI abaixo de -70 dBm. Então, nestes cenários, existem terminais ocultos — ainda que lógicos — e, além disso, eles possuem uma maior quantidade de terminais ocultos com conectividade com o ponto de acesso em relação aos cenários de dimensões maiores, principalmente entre 1400 m  $\times$  1400 m a 2500 m  $\times$  2500 m metros.



Figura 6.17: Número de pares de estações *versus* dimensão do cenário. Relação entre a quantidade de terminais ocultos com conectividade com o ponto de acesso e a diferença da vazão entre K-Means Espectral e o arbitrário para o limiar -70 dBm.

Nos cenários em que existe uma maior quantidade de terminais ocultos com conectividade com o ponto de acesso, o K-Means Espectral consegue agrupar as estações evitando colocar terminais ocultos em um mesmo grupo, considerando a informação RSSI, diferentemente do agrupamento arbitrário. Por esse motivo, o K-Means Espectral obteve uma vazão maior em relação ao agrupamento arbitrário, para os cenários de 300 m  $\times$  300 m a 1200 m  $\times$  1200 m. A diferença máxima da vazão chegou a 160 kb/s e 148 kb/s, para os cenários de 300 m  $\times$  300 m e 400 m  $\times$  400 m, respectivamente.

À medida que a dimensão dos cenários cresce, principalmente de 1400 m  $\times$  1400 m a 2500 m  $\times$  2500 m metros, as estações acabam ficando muito longe uma das outras, os terminais ocultos passam a não ter conectividade com o ponto de acesso e a vantagem do K-Means Espectral diminui. A distribuição das estações feita pelo K-Means Espectral, evitando a existência dos terminais ocultos em um mesmo grupo, não será mais eficiente porque os terminais ocultos estão muito longe do ponto de acesso e os pacotes enviados por eles serão perdidos por falta de conectividade, diminuindo o seu desempenho em termo de vazão em relação ao agrupamento arbitrário. Nota-se, nessa faixa de cenários, que a diferença de vazão chega a ficar negativa, indicando superioridade do agrupamento arbitrário. Nesse caso, isso ocorre pelo mesmo motivo observado no gráfico da Figura 6.16, um melhor balanceamento resultante do agrupamento arbitrário.

# Capítulo 7

# Conclusão

Neste trabalho, foi considerado o uso de estratégias de agrupamento RAW baseadas em algoritmos de clusterização. Na primeira estratégia de agrupamento RAW, o algortimo de clusterização K-Means foi utilizado para agrupar as estações mais próximas, considerando as suas posições geográficas, com o objetivo de evitar colocar terminais ocultos em um mesmo grupo. A segunda estratégia de agrupamento foi baseada na clusterização espectral, utilizando o RSSI entre as estações em vez de suas localizações geográficas. Dessa forma, a segunda se torna mais viável de implementar na prática, pois as informações RSSI podem ser facilmente obtidas na maioria das redes, ao contrário das posições geográficas das estações, exigidas pela abordagem inicial baseada no K-Means. Após o agrupamento, a abordagem baseada na clusterização espectral inclui uma função de pós-processamento que tenta equalizar o tamanho dos grupos, onde o balanceamento é realizado movendo estações de *clusters* maiores para os menores, tomando cuidado para não introduzir novos pares de terminais ocultos aos grupos.

Para avaliar as estratégias de agrupamento propostas, foram conduzidas simulações extensivas em cenários de redes densas. Nessas simulações, as estratégias de agrupamento foram comparadas com o algoritmo da literatura HMR e com o mecanismo RAW usado no IEEE 802.11ah. As simulações mostraram que a aplicação dos algoritmos de agrupamento propostos podem melhorar o desempenho da rede quando comparados com o HMR e RAW, em termos de todas as métricas avaliadas, que foram as taxas de colisões, vazão e atraso. Além disso, foi mostrado que a estratégia de agrupamento baseada na clusterização espectral atua em um nível semelhante — às vezes até superior — à estratégia baseada no K-Means [52], apesar de não exigir a localização geográfica precisa das estações. Além disso, os resultados sugerem que, embora a estratégia baseada na clusterização espectral possa fornecer ganhos substanciais em relação ao agrupamento arbitrário, o desempenho é ainda melhorado pelo pós-processamento. Em particular, o pós-processamento é necessário para que a estratégia baseada na clusterização espectral atinja o mesmo nível de desempenho da estratégia baseada no K-Means. Nas simulações em que foram analisadas a redução dos terminais ocultos, ao aplicar a estratégia baseada na clusterização espectral foi possível reduzir um número bem maior de terminais ocultos em comparação ao HMR. Também foram feitas simulações para analisar como o limiar de detecção de portadora afeta o desempenho da rede e se relaciona com as estratégias de agrupamento RAW. Essas simulações mostraram que em um cenário pequeno de 280 m  $\times$  280 m quando o limiar é alterado de -70 para -94, os terminais ocultos passam a não existirem e o desempenho da estratégia baseada na clusterização espectral em relação ao RAW se mostram parecidos em relação às taxas de colisão e de vazão agregada da rede. Quando o tamanho do cenário aumenta para 1500 m  $\times$  1500 m os terminais ocultos passam a existirem e o desempenho da estratégia baseada na clusterização espectral se mostrou superior ao RAW em relação às taxas de colisão e de vazão agregada da rede. Além disso, foi analisada a influência do tamanho do cenário. Quando é usado o limiar no valor de -94, a taxa de vazão nas simulações feitas com a estratégia baseada na clusterização espectral foi superior em relação ao RAW, somente nos cenários com tamanhos entre 1250 m  $\times$  1250 m a 1600 m  $\times$ 1600 m. Quando é usado o limiar de -70, à medida que o tamanho do cenário aumenta, a taxa de vazão nas simulações feitas com a estratégia baseada na clusterização espectral diminui em relação ao RAW, mas permanece superior ao RAW até cenários com tamanho de 1300 m  $\times$  1300 m. À medida que o tamanho do cenário aumenta muitas estações passam a estar longe do ponto de acesso, deixando de ter conectividade. Por isso, o ganho obtido com o uso da estratégia baseada na clusterização espectral é perdido porque os pacotes não serão entregues ao ponto de acesso. O melhor balanceamento dos grupos feito pelo RAW resultou em melhor desempenho em cenários muito grandes.

Presume-se que as estações podem obter informações RSSI ao decodificar o quadro de outra estação. De tempos em tempos, essa informação deve ser reportada ao ponto de acesso que é responsável por construir o grafo de conectividade das estações e executar o algoritmo de agrupamento. Dependendo do tamanho do grafo, o custo de processamento da estratégia baseada na clusterização espectral pode ser relativamente alto. Nesse caso, possivelmente uma solução baseada em *Software-Defined Networking* (SDN) pode ser usada, onde os dados de conectividade coletados pelo ponto de acesso podem ser retransmitidos para um controlador de rede com mais recursos para processamento. O tempo de processamento da estratégia baseada na clusterização espectral para um cenário com 100 estações foi de 0.40 segundos e para um cenário com 400 estações foi de 1 minutos e 57 segundos.

#### 7.1 Trabalhos Futuros

Como trabalho futuro seria interessante estudar melhor o comportamento do tamanho dos grupos usados nos cenários. Foram feitas simulações com 5, 10 e 15 grupos. Nos resultados das simulações mostrados no Capítulo 6, foi constatado que nos cenários com muitas estações — como 800 ou 1000 —, o tamanho dos grupos se tornou insuficiente para lidar de forma satisfatória com o nível de competição, independentemente da estratégia de agrupamento. Provavelmente, para esses cenários, seria preciso aumentar a quantidade de grupos para ser obtido um bom desempenho da rede em termos das taxas de colisão, atraso e vazão.

A função de pós-processamento pode ser melhorada para obter um balanceamento ainda mais uniforme e reduzir ainda mais o número de terminais ocultos nos grupos. O pós-processamento apresentado neste artigo é um algoritmo guloso e poderia ser melhorado usando técnicas de otimização mais sofisticadas, como GRASP, simulated annealing e ILS. Além disso, seria interessante implementar uma solução completa, incluindo as técnicas de obtenção dos dados RSSI nas estações e um protocolo para o envio periódico desses dados ao ponto de acesso. Isso permitiria a avaliação dos *overheads* de comunicação da solução.

As propriedades da clusterização espectral também podem ser exploradas posteriormente. Em particular, o conceito de *eigengap* [61] pode ser usado para selecionar automaticamente um número adequado de grupos.

## Referências

- AGARWAL, A.; DUBEY, S.; KHAN, M. A.; GANGOPADHYAY, R.; DEBNATH, S. Learning based primary user activity prediction in cognitive radio networks for efficient dynamic spectrum access. In 2016 International Conference on Signal Processing and Communications (SPCOM) (2016), IEEE, pp. 1–5.
- [2] AHMED, N.; HUSSAIN, M. I. A distributed channel access mechanism for IEEE 802.11 ah. In *IEEE world forum on Internet of Things (WF-IoT). IEEE* (2016), pp. 1–6.
- [3] ARNOSTI, S. Estratégias para agrupamento de nós e alocação dinâmica de recursos em redes IEEE 802.11ah. Master's thesis, Universidade Estadual de Campinas, Campinas, SP, 2020.
- [4] BANKOV, D.; KHOROV, E.; LYAKHOV, A.; FAMAEY, J. Resource allocation for machine-type communication of energy-harvesting devices in Wi-Fi Halow networks. *Sensors 20*, 9 (2020), 2449.
- [5] BHANDARI, S.; SHARMA, S. K.; WANG, X. Device grouping for fast and efficient channel access in IEEE 802.11ah based IoT networks. In 2018 IEEE International Conference on Communications Workshops (ICC Workshops) (New York, NY, USA, 2018), IEEE, pp. 1–6.
- BIANCHI, G. IEEE 802.11-saturation throughput analysis. *IEEE communications letters 2*, 12 (1998), 318–320.
- [7] BIANCHI, G. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on selected areas in communications 18*, 3 (2000), 535–547.
- [8] BIANCHI, G.; FRATTA, L.; OLIVERI, M. Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs. In Proceedings of PIMRC'96-7th International Symposium on Personal, Indoor, and Mobile Communications (1996), vol. 2, IEEE, pp. 392–396.
- [9] BOEDIHARDJO, M.; DENG, S.; STROHMER, T. A performance guarantee for spectral clustering. arXiv preprint arXiv:2007.05627 (2020).
- [10] BRANDES, U.; GAERTLER, M.; WAGNER, D. Experiments on graph clustering algorithms. In *European Symposium on Algorithms* (2003), Springer, pp. 568–579.
- [11] CAO, S.; LU, W.; XU, Q. Grarep: Learning graph representations with global structural information. In Proceedings of the 24th ACM international on conference on information and knowledge management (2015), pp. 891–900.

- [12] CASTRO, M. A. T. A. Agrupamento-"clustering". Projeto de Licenciatura. Instituto Superior de Engenharia do Porto (2003), 6–7.
- [13] CHIN, C.-F.; SHIH, A. C.-C.; FAN, K.-C. A novel spectral clustering method based on pairwise distance matrix. J. Inf. Sci. Eng. 26, 2 (2010), 649–658.
- [14] CISCO, U. Cisco annual internet report (2018–2023) white paper. Cisco: San Jose, CA, USA (2020).
- [15] COMMISSION, E.; CENTRE, J. R.; COUTSOMITROS, C.; GUTIÉRREZ, E.; BONO, F. Data analysis of non-standard time series : the role of graph Laplacians and covariance matrices in data and processes of complex systems. Publications Office, 2016.
- [16] DING, C.; HE, X. Cluster merging and splitting in hierarchical clustering algorithms. In 2002 IEEE International Conference on Data Mining, 2002. Proceedings. (2002), IEEE, pp. 139–146.
- [17] FARRELLY, C. M.; SCHWARTZ, S. J.; AMODEO, A. L.; FEASTER, D. J.; STEIN-LEY, D. L.; MECA, A.; PICARIELLO, S. The analysis of bridging constructs with hierarchical clustering methods: An application to identity. *Journal of Research in Personality 70* (2017), 93–106.
- [18] FIEDLER, M. Algebraic connectivity of graphs. Czechoslovak mathematical journal 23, 2 (1973), 298–305.
- [19] GAST, M. 802.11 wireless networks: the definitive guide. "O'Reilly Media, Inc.", 2005.
- [20] GE, Y.; KUNZ, T.; LAMONT, L. Quality of service routing in ad-hoc networks using OLSR. In 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the (2003), IEEE, pp. 9–pp.
- [21] GUEDES, B. T.; PASSOS, F. G.; PASSOS, D. The role of artificial intelligence in current and future dense wireless networks. In 2019 1st Blockchain, Robotics and AI for Networking Security Conference (BRAINS) (2019), IEEE, pp. 1–8.
- [22] HEFEEDA, M.; GAO, F.; ABD-ALMAGEED, W. Distributed approximate spectral clustering for large-scale datasets. In *Proceedings of the 21st international symposium* on High-Performance Parallel and Distributed Computing (2012), pp. 223–234.
- [23] IRANI, J.; PISE, N.; PHATAK, M. Clustering techniques and the similarity measures used in clustering: A survey. *International journal of computer applications* 134, 7 (2016), 9–14.
- [24] JI, Z.; LIU, K. R. Cognitive radios for dynamic spectrum access-dynamic spectrum sharing: A game theoretical overview. *IEEE Communications Magazine* 45, 5 (2007), 88–94.
- [25] KAMESHWARAN, K.; MALARVIZHI, K. Survey on clustering techniques in data mining. International Journal of Computer Science and Information Technologies 5, 2 (2014), 2272–2276.

- [26] KATTI, S.; RAHUL, H.; HU, W.; KATABI, D.; MÉDARD, M.; CROWCROFT, J. XoRs in the air: Practical wireless network coding. In *Proceedings of the 2006* conference on Applications, technologies, architectures, and protocols for computer communications (2006), pp. 243–254.
- [27] KHOROV, E.; KROTOV, A.; LYAKHOV, A. Modelling machine type communication in IEEE 802.11 ah networks. In 2015 IEEE international conference on communication workshop (ICCW) (2015), IEEE, pp. 1149–1154.
- [28] KHOROV, E.; LYAKHOV, A.; YUSUPOV, R. Two-slot based model of the IEEE 802.11 ah restricted access window with enabled transmissions crossing slot boundaries. In 2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM) (2018), IEEE, pp. 1–9.
- [29] KHURANA, S.; KAHOL, A.; JAYASUMANA, A. P. Effect of hidden terminals on the performance of IEEE 802.11 MAC protocol. In *Proceedings 23rd Annual Conference* on Local Computer Networks. LCN'98 (Cat. No. 98TB100260) (New York, NY, USA, 1998), IEEE, pp. 12–20.
- [30] KUREEV, A.; BANKOV, D.; KHOROV, E.; LYAKHOV, A. Improving efficiency of heterogeneous Wi-Fi networks with joint usage of tim segmentation and restricted access window. In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) (2017), IEEE, pp. 1–5.
- [31] KUROSE, J. F.; ROSS. Redes de Computadores e a Internet: uma abordagem topdown. Pearson, 2013.
- [32] KUROSE, J. F.; ROSS, K. W. Redes de Computadores e a Internet. Addison-Wesley, 2005.
- [33] LI, S.; DA XU, L.; ZHAO, S. The Internet of Things: a survey. Information Systems Frontiers 17, 2 (2015), 243–259.
- [34] LIM, J.; KOOK, J.; KIM, J. DBSCAN-D: A density-based clustering method of directionality. International Journal of Applied Engineering Research 12, 13 (2017), 3927–3932.
- [35] LIU, J.; HUANG, J. Z.; LUO, J.; XIONG, L. Privacy preserving distributed dbscan clustering. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops* (2012), pp. 177– 185.
- [36] LUCIŃSKA, M.; WIERZCHOŃ, S. T. Clustering based on eigenvectors of the adjacency matrix. International Journal of Applied Mathematics and Computer Science 28, 4 (2018).
- [37] MADHULATHA, T. S. An overview on clustering methods. arXiv preprint ar-Xiv:1205.1117 (2012).
- [38] MAGALHÃES, L. C. S.; BALBI, H. D.; CORRÊA, C.; VALLE, R. D. T. D.; STAN-TON, M. Scifi-a software-based controller for efficient wireless networks. *Proceedings* and report of the 6th UbuntuNet Alliance annual conference (2013).
- [39] MELO, V. V. D. Clustering de artigos científicos em uma ferramenta inteligente de apoio à pesquisa. Tese de Doutorado, Universidade de São Paulo, 2014.
- [40] MERCADO, P.; TUDISCO, F.; HEIN, M. Spectral clustering of signed graphs via matrix power means. In *International Conference on Machine Learning* (2019), PMLR, pp. 4526–4536.
- [41] NAWAZ, N.; HAFEEZ, M.; ZAIDI, S. A. R.; MCLERNON, D. C.; GHOGHO, M. Throughput enhancement of restricted access window for uniform grouping scheme in IEEE 802.11 ah. In 2017 IEEE international conference on communications (ICC) (2017), IEEE, pp. 1–7.
- [42] NOACK, A. Energy models for graph clustering. J. Graph Algorithms Appl. 11, 2 (2007), 453–480.
- [43] PASSOS, D.; BALBI, H.; CARRANO, R. Tecnologias de redes. In Escola Superior de Redes (2016).
- [44] POPAT, S. K.; EMMANUEL, M. Review and comparative study of clustering techniques. International journal of computer science and information technologies 5, 1 (2014), 805–812.
- [45] RAEESI, O.; PIRSKANEN, J.; HAZMI, A.; LEVANEN, T.; VALKAMA, M. Performance evaluation of IEEE 802.11 ah and its restricted access window mechanism. In 2014 IEEE international conference on communications workshops (ICC) (2014), IEEE, pp. 460–466.
- [46] RAPPAPORT, T. S. Wireless communications-principles and practice, (the book end). *Microwave Journal* 45, 12 (2002), 128–129.
- [47] RITTENHOUSE, M. L. Properties and Recent Applications in Spectral Graph Theory. Tese de Doutorado, Virginia Commonwealth University, 2008.
- [48] RODRIGUES, R. D. Detecção de outliers baseada em caminhada determinística do turista. Tese de Doutorado, Universidade de São Paulo, 2018.
- [49] RODRIGUES, W. C.; FADEL, A. C.; SEMAAN, G. S.; DE MOURA BRITO, J. A. Um novo método baseado em grade e densidade com tratamento de ruídos para a identificação do número ideal de grupos. In *Simpósio Brasileiro de Pesquisa Operacional* (2016).
- [50] SAERENS, M.; FOUSS, F.; YEN, L.; DUPONT, P. The principal components analysis of a graph, and its relationships to spectral clustering. In *European conference on* machine learning (2004), Springer, pp. 371–383.
- [51] SHERMAN, M.; MODY, A. N.; MARTINEZ, R.; RODRIGUEZ, C.; REDDY, R. IEEE standards supporting cognitive radio and networks, dynamic spectrum access, and coexistence. *IEEE Communications Magazine* 46, 7 (2008), 72–79.
- [52] SHUKLA, S.; NAGANNA, S. A review on k-means data clustering approach. International Journal of Information and Computation Technology 4, 17 (2014), 1847–1860.

- [53] SMITH, L. M.; LERMAN, K.; GARCIA-CARDONA, C.; PERCUS, A. G.; GHOSH, R. Spectral clustering with epidemic diffusion. *Physical Review E 88*, 4 (2013), 042813.
- [54] SPIELMAN, D. Spectral graph theory. Combinatorial scientific computing 18 (2012).
- [55] TENG, J.; XU, C.; JIA, W.; XUAN, D. D-scan: Enabling fast and smooth handoffs in ap-dense 802.11 wireless networks. In *IEEE INFOCOM 2009* (2009), IEEE, pp. 2616–2620.
- [56] TIAN, L.; DERONNE, S.; LATRÉ, S.; FAMAEY, J. Implementation and validation of an IEEE 802.11 ah module for ns-3. In *Proceedings of the Workshop on ns-3* (2016), pp. 49–56.
- [57] TIAN, L.; FAMAEY, J.; LATRÉ, S. Evaluation of the IEEE 802.11 ah restricted access window mechanism for dense IoT networks. In 2016 IEEE 17th international symposium on a world of wireless, mobile and multimedia networks (WoWMoM) (2016), IEEE, pp. 1–9.
- [58] TIAN, L.; LOPEZ-AGUILERA, E.; GARCIA-VILLEGAS, E.; MEHARI, M. T.; DE PO-ORTER, E.; LATRÉ, S.; FAMAEY, J. Optimization-oriented RAW modeling of IEEE 802.11 ah heterogeneous networks. *IEEE Internet of Things Journal 6*, 6 (2019), 10597–10609.
- [59] TRIVEDI, S.; PARDOS, Z. A.; SÁRKÖZY, G. N.; HEFFERNAN, N. T. Spectral clustering in educational data mining. In *EDM* (2011).
- [60] VAEZPOUR, E.; DEHGHAN, M. A multi-objective optimization approach for joint channel assignment and multicast routing in multi-radio multi-channel wireless mesh networks. Wireless personal communications 77, 2 (2014), 1055–1076.
- [61] VON LUXBURG, U. A tutorial on spectral clustering. Statistics and computing 17, 4 (2007), 395–416.
- [62] WANG, Y.; CHAI, K. K.; CHEN, Y.; SCHORMANS, J. Energy efficient window control scheme for IEEE 802.11 ah (Wi-Fi Halow) based networks. J. Electr. Eng 5 (2017), 242–252.
- [63] WANG, Y.; LI, Y.; CHAI, K. K.; CHEN, Y.; SCHORMANS, J. Energy-aware adaptive restricted access window for IEEE 802.11 ah based smart grid networks. In 2015 IEEE international conference on smart grid communications (SmartGridComm) (2015), IEEE, pp. 581–586.
- [64] YOON, S.; J.SEO; BAHK, S. Regrouping algorithm to alleviate the hidden node problem in 802.11 ah networks. *Computer Networks 105* (2016), 22–32.
- [65] ZHENG, L.; NI, M.; CAI, L.; PAN, J.; GHOSH, C.; DOPPLER, K. Performance analysis of group-synchronized DCF for dense IEEE 802.11 networks. *IEEE Tran*sactions on Wireless Communications 13, 11 (2014), 6180–6192.
- [66] ZHONG, Z.; KULKARNI, P.; CAO, F.; FAN, Z.; ARMOUR, S. Issues and challenges in dense WiFi networks. In 2015 International Wireless Communications and Mobile Computing Conference (IWCMC) (2015), IEEE, pp. 947–951.