



**ATRIBUIÇÃO DINÂMICA DE IDENTIFICADORES AID
EM REDES IEEE 802.11AH BASEADA EM AGRUPAMENTO
VIA ALGORITMO K-MEANS**

EDUARDO DA COSTA OLIVEIRA

**MONOGRAFIA DE GRADUAÇÃO EM ENGENHARIA DE REDES DE
COMUNICAÇÃO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

Faculdade de Tecnologia

Universidade de Brasília

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**ATRIBUIÇÃO DINÂMICA DE IDENTIFICADORES AID
EM REDES IEEE 802.11AH BASEADA EM AGRUPAMENTO
VIA ALGORITMO K-MEANS**

EDUARDO DA COSTA OLIVEIRA

Orientador: PROF. DR. MARCELO MENEZES DE CARVALHO, ENE/UNB

**MONOGRAFIA DE GRADUAÇÃO EM ENGENHARIA DE REDES DE
COMUNICAÇÃO**

BRASÍLIA-DF, 23 DE SETEMBRO DE 2022.

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**ATRIBUIÇÃO DINÂMICA DE IDENTIFICADORES AID
EM REDES IEEE 802.11AH BASEADA EM AGRUPAMENTO
VIA ALGORITMO K-MEANS**

EDUARDO DA COSTA OLIVEIRA

MONOGRAFIA DE GRADUAÇÃO ACADÊMICO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM ENGENHARIA DE REDES DE COMUNICAÇÃO.

APROVADA POR:

Prof. Dr. Marcelo Menezes de Carvalho, ENE/UnB
Orientador

Prof. Dr. Paulo Roberto de Lira Gondim, ENE/UnB
Examinador interno

Enga. Ma. Stephanie Miranda Soares, ENE/UnB
Examinador interno

BRASÍLIA, 23 DE SETEMBRO DE 2022.

FICHA CATALOGRÁFICA

EDUARDO DA COSTA OLIVEIRA

Atribuição Dinâmica de Identificadores AID em Redes IEEE 802.11ah Baseada em Agrupamento via Algoritmo K-Means

2022xv, 60p., 201x297 mm

(ENE/FT/UnB, Bacharel, Engenharia de Redes de Comunicação, 2022)

Monografia de Graduação - Universidade de Brasília

Faculdade de Tecnologia - Departamento de Engenharia Elétrica

REFERÊNCIA BIBLIOGRÁFICA

EDUARDO DA COSTA OLIVEIRA (2022) Atribuição Dinâmica de Identificadores AID em Redes IEEE 802.11ah Baseada em Agrupamento via Algoritmo K-Means. Monografia de Graduação em Engenharia de Redes de Comunicação, Publicação xxx/AAAA, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 60p.

CESSÃO DE DIREITOS

AUTOR: Eduardo da Costa Oliveira

TÍTULO: Atribuição Dinâmica de Identificadores AID em Redes IEEE 802.11ah Baseada em Agrupamento via Algoritmo K-Means.

GRAU: Bacharel ANO: 2022

É concedida à Universidade de Brasília permissão para reproduzir cópias desta monografia de Graduação e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor se reserva a outros direitos de publicação e nenhuma parte desta monografia de Graduação pode ser reproduzida sem a autorização por escrito do autor.

Eduardo da Costa Oliveira

Agradecimentos

À minha família, por todo o suporte dado a mim que contribuiu, indiretamente, para a produção deste trabalho. Aos professores que tive durante minha formação, pelos ensinamentos. Ao meu orientador, por ter me guiado no processo de desenvolvimento deste trabalho. Às amigas que fiz na universidade, por me tranquilizarem nos momentos difíceis. À Universidade de Brasília, pela estrutura e pelo ensino público, gratuito e de qualidade.

Resumo

O padrão de comunicação de redes sem fio IEEE 802.11ah, proposto para ser implementado em redes de internet das coisas (IoT), opera em faixas sub-1GHz e suporta redes com até 8191 estações associadas a um único ponto de acesso (AP). Para a sub-camada de controle de acesso ao meio (MAC), o padrão propõe um conjunto de novos mecanismos para suportar os requisitos de vazão, justiça, latência e eficiência energética de redes altamente densas e com limitações de energia. Exemplos de mecanismos propostos são a Janela de Acesso Restrito (RAW, do inglês *Restricted Access Window*) e a atribuição dinâmica de identificadores de associação (AID, do inglês *Association Identifier*).

O mecanismo de RAW divide as estações em grupos, e cada grupo tem um tempo reservado para acessar o canal. O tempo de acesso de cada grupo pode ser subdividido em compartimentos de tempo (“slots”) de mesma duração, que são restritos a sub-grupos de estações. Com isto, a probabilidade de colisões diminui, o que melhora o desempenho da rede. A operação do RAW é definida por um conjunto de parâmetros (como instante de início de RAW e o número de *slots* em cada RAW) que é enviado nos quadros de *beacon* transmitidos periodicamente via *broadcast*. O conjunto de parâmetros RAW (RPS, do inglês *RAW Parameter Set*) também inclui dois valores de AID para cada grupo: AID inicial e AID final. Uma estação só está em um determinado grupo se o valor do seu AID estiver dentro do intervalo definido por esses dois limitantes. A norma IEEE 802.11ah não indica qual abordagem deve ser seguida para a escolha dos parâmetros RAW. Sendo assim, fica a critério de fabricantes e desenvolvedores aplicar estratégias com base nos seus próprios requisitos de desempenho.

Para dar suporte a abordagens dinâmicas de agrupamento, este trabalho propõe uma implementação da funcionalidade de atribuição dinâmica de AID no simulador ns-3, que permite que o AID de uma estação seja alterado durante a operação da rede. Com isto, deve ser possível reagrupar as estações de acordo com seu comportamento no decorrer do tempo de operação, possibilitando a aplicação de estratégias de agrupamento em cenários de redes dinâmicas e heterogêneas. Em particular, este trabalho propõe uma estratégia de agrupamento de estações baseada no algoritmo K-Means, a partir de parâmetros de tráfego e condição de canal das estações. O objetivo é agrupar as estações com características similares para competirem pelo canal entre si. A estratégia proposta foi implementada no ns-3, e os resultados das simulações mostram ganhos significativos de justiça de vazão em comparação com resultados de esquemas de agrupamento aleatório.

Abstract

The IEEE 802.11ah standard for wireless networks, proposed to be implemented in Internet of Things (IoT) networks, operates in sub-1GHz bands and supports networks with up to 8191 stations associated with a single access point (AP). For the medium access control (MAC) sub-layer, the standard proposes a set of new mechanisms to support the throughput, fairness, latency and energy efficiency requirements of highly dense and energy-constrained networks. Examples of proposed mechanisms are the Restricted Access Window (RAW) and the dynamic assignment of association identifiers (AID).

The RAW mechanism divides stations into groups, and each group has a reserved time to access the channel. The access time of each group can be subdivided into time compartments (slots) of the same duration, which are restricted to sub-groups of stations. As a result, the probability of collisions decreases, which improves network performance. The RAW operation is defined by a set of parameters (such as RAW start time and the number of slots in each RAW) that are sent in beacon frames transmitted periodically via broadcast. The RAW parameter set (RPS) also includes two association identifier (AID) values for each group: start AID and end AID. A station is only in a certain group if its AID value is within the range defined by these two bounds. The IEEE 802.11ah standard does not indicate which approach should be followed for choosing RAW parameters. Therefore, it is up to manufacturers and developers to apply strategies based on their own performance requirements.

To support dynamic grouping approaches, this work proposes an implementation of the dynamic AID assignment functionality in the ns-3 simulator, which allows a station's AID to be changed during network operation. With this, it should be possible to regroup stations according to their behavior during the operating time, enabling the application of grouping strategies in scenarios of dynamic and heterogeneous networks. In particular, this work proposes a station grouping strategy based on the K-Means algorithm that considers traffic parameters and station channel condition. The goal is to group stations with similar characteristics to compete for the channel with each other. The proposed strategy was implemented in the ns-3 simulator, and the simulation results show significant gains in throughput fairness compared to results from random grouping schemes.

SUMÁRIO

AGRADECIMENTOS	I
RESUMO	II
ABSTRACT	III
1 INTRODUÇÃO	1
1.1 CONTEXTUALIZAÇÃO	1
1.2 DEFINIÇÃO DO PROBLEMA	2
1.3 OBJETIVOS.....	3
1.4 CONTRIBUIÇÕES DO PROJETO	3
1.5 ORGANIZAÇÃO DO MANUSCRITO	4
2 REFERENCIAL TEÓRICO	5
2.1 IEEE 802.11AH PARA IOT	5
2.1.1 CASOS DE USO	6
2.2 CAMADA FÍSICA IEEE 802.11AH	7
2.3 CAMADA DE CONTROLE DE ACESSO AO MEIO (MAC) DO IEEE 802.11AH	8
2.3.1 ASSOCIAÇÃO E AUTENTICAÇÃO RÁPIDAS	8
2.3.2 SEGMENTAÇÃO TIM	10
2.3.3 <i>Target Wake Time</i>	10
2.3.4 JANELA DE ACESSO RESTRITO	11
2.3.5 SETORIZAÇÃO DE GRUPOS.....	14
2.3.6 AID HIERÁRQUICO.....	15
2.3.7 AID DINÂMICO	15
2.3.8 COR DE BSS.....	17
2.3.9 CABEÇALHO MAC CURTO.....	17
2.4 ALGORITMO K-MEANS	18
3 PROPOSTAS DE OTIMIZAÇÃO DO RAW	21
3.1 EFICIÊNCIA ENERGÉTICA.....	21
3.2 LATÊNCIA E QUALIDADE DE SERVIÇO.....	22
3.3 PROBABILIDADE DE COLISÃO E TERMINAIS ESCONDIDOS	22

3.4	VAZÃO.....	23
3.5	JUSTIÇA	24
4	IMPLEMENTAÇÃO DAS PROPOSTAS	26
4.1	SIMULADOR NS-3	26
4.2	ASSOCIAÇÃO DE ESTAÇÕES	28
4.3	IMPLEMENTAÇÃO DO AID DINÂMICO	29
4.4	COLETA DE PARÂMETROS DAS ESTAÇÕES	33
4.5	O ALGORITMO K-MEANS	34
5	AVALIAÇÃO DE DESEMPENHO	37
5.1	CONFIGURAÇÃO DAS SIMULAÇÕES	38
5.2	IMPACTO DA INICIALIZAÇÃO DO K-MEANS	40
5.3	VALIDAÇÃO DO MECANISMO DE ATRIBUIÇÃO DINÂMICA DE AID.....	41
5.4	ESTUDO DE CASO 1	43
5.5	ESTUDO DE CASO 2	45
5.6	ESTUDO DE CASO 3	46
5.7	IMPACTO DO NÚMERO DE ESTAÇÕES.....	47
	CONCLUSÃO.....	50
	REFERÊNCIAS BIBLIOGRÁFICAS.....	52
A	IMPLEMENTAÇÃO DOS ALGORITMOS DE INICIALIZAÇÃO DO K-MEANS	56
A.1	INICIALIZAÇÃO ALEATÓRIA.....	56
A.2	INICIALIZAÇÃO PELA HEURÍSTICA DO PONTO DE MAIOR DISTÂNCIA..	57
A.3	INICIALIZAÇÃO PELA HEURÍSTICA DE ORDENAMENTO	59

LISTA DE FIGURAS

2.1	Processo de inicialização em redes IEEE 802.11ah.	9
2.2	Exemplo de operação da segmentação TIM.....	11
2.3	Exemplo de operação do RAW.	12
2.4	Demonstração de atribuição de RAW slot.	13
2.5	Processo de <i>backoff</i> para controle de acesso ao canal com RAW.....	13
2.6	Transmissão sem cross slot boundary.....	14
2.8	Exemplo de setorização de Grupos.....	15
2.9	Exemplo de configuração do AID hierárquico.	16
2.10	Exemplo de operação do AID dinâmico.	17
2.11	Cabeçalho MAC tradicional.	18
2.12	Cabeçalho MAC curto.....	18
4.1	Estrutura modular da implementação de dispositivos Wi-Fi no ns-3.....	28
4.2	Elemento <i>AID Switch Response</i>	30
4.3	Fluxo da operação executada pela função <i>SendOneBeacon()</i>	32
5.1	Topologia da rede dividida em zonas geográficas.	38
5.2	Resultado do agrupamento do algoritmo K-Means com inicialização aleatória.	41
5.3	Resultado do agrupamento do algoritmo K-Means com a inicialização da heurística da maior distância (<i>MinMax</i>).....	42
5.4	Resultado do agrupamento do algoritmo K-Means com a inicialização da heurística de ordenamento (<i>Sorting</i>).	43
5.5	Exemplo de atribuição dinâmica de AID para uma estação com mobilidade.	44
5.6	Gráfico de Justiça normalizada para diferentes números de de estações.	48

LISTA DE TABELAS

2.1	Comparação das tecnologias de comunicação sem fio para IoT [Tian et al. 2021].	6
2.2	Esquemas de Modulação e Codificação do IEEE 802.11ah para NSS = 1 e GI = 8 μ s.	8
5.1	Parâmetros gerais da camada física para as simulações realizadas.	39
5.2	Parâmetros da camada MAC usados nas simulações.	39
5.3	Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 1.	45
5.4	Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 2.	46
5.5	Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 3.	47

LISTA DE CÓDIGOS FONTE

4.1	Função de envio de Resposta de troca de AID.	30
4.2	Função de recepção da Resposta de troca de AID pela estação.	31
4.3	Rotina que calcula a clusterização do K-means.	35
A.1	Rotina que implementa a inicialização aleatória dos grupos do K-Means.	56
A.2	Rotina que implementa a inicialização dos grupos do K-Means pela heurística do ponto mais distante.	57
A.3	Rotina que implementa a inicialização dos grupos do K-Means pela heurística de ordenamento.	59

LISTA DE TERMOS E SIGLAS

k	Número de grupos
N	Número de pontos de entrada do algoritmo K-Means
N_b	Número de blocos da hierarquia de AID
N_p	Número de páginas da hierarquia de AID
N_{RAW}	Número de <i>slots</i> em um grupo RAW
ACS	<i>Slots</i> de controle de associação
AID	Identificador de Associação
AP	Ponto de Acesso
BLE	<i>Bluetooth</i> de baixa energia
BSS	Conjuntos Básicos de Serviço
CAC	Controle Centralizado de Autenticação
CSB	<i>Cross Slot Boundary</i>
CSMA/CA	Acesso múltiplo com detecção de portadora com prevenção de colisão
DAC	Controle de Autenticação Distribuído
DCF	Função de Coordenação Distribuída
DP	Desvio Padrão
DTIM	Mapa de indicação de tráfego de entrega
EDCA	Acesso distribuído ao canal aprimorado
GI	Intervalo de Guarda
IoT	Internet das Coisas
ISM	Industrial, científica e médica

LPWAN	Rede de longa distância de baixa potência
MAC	Controle de acesso ao meio
MCS	Esquema de modulação e codificação
MIMO	Múltiplas entradas e múltiplas saídas
NSS	Número de fluxos espaciais
OFDM	Multiplexação por divisão de frequências ortogonais
PLCP	Procedimento de convergência da camada física
QoS	Qualidade de Serviço
RAW	Janela de Acesso Restrita
RPS	Conjunto de Parâmetros RAW
S1G	Abaixo de 1GHz
TDMA	Acesso Múltiplo por Divisão de Tempo
TIM	Mapa de Indicação de Tráfego
TWT	<i>Target Wake Time</i>
TXOP	Oportunidade de Transmissão
WPAN	Rede de área pessoal sem fio

Capítulo 1

Introdução

1.1 Contextualização

O paradigma de Internet das Coisas (IoT, do inglês *Internet of Things*) representa uma revolução no modo em que seres humanos interagem com a tecnologia e como os dispositivos conectados à rede comunicam-se entre si. IoT envolve o desenvolvimento de ambientes inteligentes, tais como casas inteligentes, sistema de saúde inteligente, sistema de automação na agricultura e cidades inteligentes. Para estes cenários, são implementados diversos dispositivos movidos à bateria (como sensores, atuadores e controladores) interconectados que fazem coleta e transferência de dados. Com isto, surgem requisitos de comunicação que não podem ser totalmente atendidos por tecnologias tradicionais, havendo, assim, a necessidade de se buscar tecnologias de comunicação que sejam energeticamente eficientes, capazes de transmitir a diferentes taxas e com raio de cobertura de sinal maior.

As tecnologias de comunicação sem fio voltadas para IoT são classificadas entre duas categorias: redes de área pessoal sem fio (WPAN, do inglês *Wireless Personal Area Networks*) e rede de longa distância de baixa potência (LPWAN, do inglês *Low Power Wide Area Network*). As tecnologias WPAN (como *Bluetooth* de baixa energia e *Zigbee*) conseguem atingir valores de vazão relativamente altos (por volta de milhares de kb/s) a curtas distâncias de transmissão (em torno de dezenas de metros). Em contrapartida, as tecnologias LPWAN (como LoRa, SigFox, NB-IoT, eMTC e Wi-SUN) conseguem transmitir a longas distâncias (isto é, até vários quilômetros), porém, atingindo valores baixos de vazão (de alguns kb/s) [Tian et al. 2021]. Existe, portanto, um compromisso entre vazão e distância de transmissão quando comparam-se as diferentes tecnologias de comunicação.

Em princípio, pode-se pensar no padrão IEEE 802.11 legado como a carta na manga para cobrir lacuna deixada em tecnologias de comunicação sem fio para IoT, devido ao seu grande sucesso em redes residenciais e comerciais, além de ser encontrado nos mais diversos dispositivos eletrônicos e aplicado nos mais variados cenários. Apesar disso, o padrão IEEE 802.11 não obteve tanto sucesso em Internet das Coisas, pois as otimizações para esta tec-

nologia foram feitas para redes com número pequeno de estações, transmitindo a altas taxas, e não para redes altamente densas operando a taxas mais baixas, como redes de sensores IoT [Park 2015]. Tendo em vista o futuro recente de IoT e a lacuna existente nas tecnologias de comunicação, o grupo de trabalho do IEEE 802.11 propôs uma emenda ao padrão de Wi-Fi, nomeada IEEE 802.11ah [IEEE 2017], e divulgada no mercado como *Wi-Fi HaLow*.

O IEEE 802.11ah possui um conjunto vasto de novas funcionalidades, permitindo que até 8191 estações associem-se a um único ponto de acesso operando em frequências abaixo de 1 GHz na faixa ISM (industrial, científica e médica), com um alcance de transmissão de até 1 km e taxas de transmissão de dados entre 150 kb/s e 78 Mb/s com um único feixe espacial, sendo classificado como uma tecnologia LPWAN. Alguns exemplos dos mecanismos introduzidos na camada MAC pelo IEEE 802.11ah são Janela de Acesso Restrito (RAW, do inglês *Restricted Access Window*), Segmentação de Mapa de Indicação de Tráfego (TIM, do inglês *Traffic Indication Map*) e *Target Wake Time* (TWT). O objetivo do mecanismo de RAW é diminuir a ocorrência de colisões e, com isso, melhorar a performance da rede. Já os mecanismos de segmentação TIM e TWT visam melhor eficiência energética e diminuição de sobrecarga.

1.2 Definição do problema

O mecanismo de Janela de Acesso Restrito é proposto para diminuir o número de colisões em redes densas. Trata-se de uma combinação do TDMA com o CSMA/CA que divide as estações associadas em grupos, de modo que cada grupo tenha um tempo específico para competir pelo canal usando DCF ou EDCA. Com isto, a probabilidade de colisões é altamente reduzida, resultando em melhoria no desempenho da rede em termos de vazão, latência e justiça. No entanto, a norma do IEEE 802.11ah não menciona qualquer estratégia de agrupamento de estações a ser seguida [IEEE 2017], de modo que parâmetros como o número de grupos, a duração de cada grupo e a distribuição de estações entre os grupos são definidos por fabricantes e desenvolvedores. Na literatura, diversas abordagens de agrupamento já foram propostas com o objetivo de otimizar alguma medida de desempenho de interesse, como vazão, latência, eficiência energética e justiça [Tian et al. 2021]. De modo geral, o agrupamento de estações depende de muitos parâmetros, o que exige uma abordagem de agrupamento inteligente, capaz de lidar com redes heterogêneas e dinâmicas. As estratégias propostas na literatura são apresentadas baseadas em modelos analíticos ou por meio de simulações computacionais que representam redes reais. Uma ferramenta muito presente em trabalhos que realizam simulações é o ns-3, mais precisamente, a extensão do simulador que suporta o padrão IEEE 802.11ah implementada por Le Tian et.al [Tian et al. 2018].

A atribuição das estações aos grupos RAW é feita por meio do identificador de associação (AID, do inglês *Association Identifier*), o qual é atribuído a cada estação no processo de associação, e as estações em um mesmo grupo RAW devem ter os valores de AID se-

quenciais. No IEEE 802.11 legado, o valor do AID é fixo, não sendo alterado enquanto a estação estiver associada. Já o padrão IEEE 802.11ah permite a atribuição dinâmica deste identificador, sendo possível alterá-lo enquanto a estação estiver associada, por iniciativa da própria estação ou por iniciativa do ponto de acesso.

Muitas das propostas revisadas para este trabalho não consideram a atribuição dinâmica de AID combinada às estratégias de agrupamento propostas. Logo, as abordagens ficam limitadas a agrupar apenas estações com valores de AID em sequência, tornando as soluções pouco dinâmicas para serem aplicadas em cenários reais.

1.3 Objetivos

Tem-se como objetivo para este trabalho a aplicação da operação de atribuição dinâmica de AID combinada com estratégias de agrupamento, de modo que seja possível modificar a distribuição de estações por grupos RAW durante a operação da rede de acordo com algum critério de interesse. Como o mecanismo de AID dinâmico não está implementado na extensão do ns-3 com suporte ao padrão IEEE 802.11ah proposta por Le Tian, tem-se como objetivo implementar esta funcionalidade ao simulador como contribuição para a ferramenta.

Também é objetivo deste trabalho aplicar uma estratégia de agrupamento RAW de estações baseada no algoritmo K-Means, com base em parâmetros de tráfego e condição de canal medidos no AP. A estratégia deve ser implementada no simulador ns-3 para a validação do mecanismo implementado e a avaliação de medidas de desempenho de vazão média e justiça.

1.4 Contribuições do Projeto

As principais contribuições deste trabalho são:

- Implementação inédita do mecanismo de atribuição dinâmica de AID em redes IEEE 802.11ah no simulador ns-3, como modificação da implementação proposta por [Tian et al. 2018].
- Proposta e implementação de uma estratégia de agrupamento baseada no algoritmo K-Means com base em parâmetros de tamanho de pacote, potência de sinal recebido e taxa nominal de transmissão.
- Avaliação do desempenho da estratégia de agrupamento proposta a partir de medidas de vazão média e justiça de Jain. Os resultados são comparados aos obtidos por estratégias baseadas em agrupamento aleatório. As medidas de desempenho obtidas mostram que a estratégia de agrupamento baseada em K-Means apresenta ganhos de justiça de vazão quando comparada a métodos de agrupamento aleatório.

- Submissão dos resultados obtidos neste trabalho como artigo convidado para a conferência *The 18th International Conference on Mobility, Sensing and Networking (MSN)* com o título “*K-Means Based Grouping of Stations with Dynamic AID Assignment in IEEE 802.11ah Networks*” [Oliveira et al. 2022].

1.5 Organização do Manuscrito

Este trabalho está organizado como segue. No capítulo 2, são apresentados os principais conceitos da fundamentação teórica necessária para o desenvolvimento e a compreensão deste trabalho. No capítulo 3, é feita a revisão das estratégias de agrupamento propostas na literatura para otimizar medidas de desempenho de interesse. No capítulo 4, são descritos os detalhes de implementação da funcionalidade de atribuição dinâmica de AID, bem como a implementação da coleta de parâmetros da rede e da estratégia de agrupamento baseada em K-Means. No capítulo 5, é feita a verificação das funcionalidades apresentadas e a avaliação de desempenho com base nas medidas de vazão média e justiça.

Capítulo 2

Referencial Teórico

2.1 IEEE 802.11ah para IoT

Em internet das coisas, os ambientes são compostos por objetos interconectados com acesso à internet que realizam atividades voltadas para o bem estar humano e ao desenvolvimento econômico. O paradigma de IoT pode ser aplicado em diversos setores, como na área da saúde, na agricultura, em infraestrutura urbana, na manufatura e na segurança. Os dispositivos inteligentes (como sensores, atuadores e controlares) representam a unidade principal em um ambiente IoT e são responsáveis por operações de monitoramento, sensoriamento, controle e atividades de atuação, além da troca de dados com aplicações e servidores inteligentes [Rose et al. 2015].

As tecnologias de comunicação para IoT devem ser capazes de operar eficientemente em redes com alto número de dispositivos com diferentes requisitos de alcance, características de tráfego de dados heterogêneas e com limitações de distância [Al-Sarawi et al. 2017]. Essas tecnologias podem ser classificadas entre duas categorias: rede de área pessoal sem fio (WPAN) e rede de longa distância de baixa potência (LPWAN).

Tecnologias WPAN alcançam distâncias de até 100 metros, transmitindo dados a taxas de até 2 Mb/s. Os exemplos mais populares de tecnologias WPAN são *Bluetooth* de baixa energia (BLE) e *ZigBee*. O BLE [Gomez et al. 2012] opera em 2,4 GHz na banda ISM, podendo transmitir a uma taxa de transmissão de dados de 2 Mb/s. Para a transmissão de dados, uma conexão é criada entre duas entidades nomeadas mestre e escravo, seguindo uma topologia em estrela, de modo que um mestre pode se conectar a diferentes escravos. O Zigbee é baseado no padrão IEEE 802.15.4, desenvolvido para prover comunicação de baixo custo para redes com requisitos de baixa taxa de dados, com baterias duradouras e dispositivos de rede seguros [Al-Sarawi et al. 2017].

Tecnologias LPWAN têm grande alcance de cobertura, podendo alcançar até 10 quilômetros de distância, dependendo do ambiente de implementação [Raza et al. 2017]. Isto é possível através da operação em faixas Sub-1GHz e o uso de esquemas especiais de modu-

lação. Outra característica de tecnologias LPWAN é o baixo consumo energético, de modo que as baterias dos dispositivos podem durar até 10 anos. Alguns exemplos de tecnologias LPWAN são: LoRa, SigFox, NB-IoT, eMTC, Wi-SUN e IEEE 802.11ah. A Tabela 2.1 apresenta uma breve comparação entre as tecnologias de comunicação para IoT.

Tabela 2.1: Comparação das tecnologias de comunicação sem fio para IoT [Tian et al. 2021].

Tecnologia	Frequência	Largura de banda	Alcance	Taxa de dados
BLE	2,4 GHz (ISM)	2 MHz	30 m	1 - 2 Mb/s
Zigbee	2,4 / Sub GHz (ISM)	5 MHz	100 m	10 - 250 kb/s
LoRa	Sub GHz (ISM)	125/500 kHz	20 km	300 - 500 kb/s
SigFox	Sub GHz (ISM)	100 kHz	50 km	100 b/s
NB-IoT (5G)	Sub GHz (licenciadas)	180 kHz	15 km	250 b/s
eMTC (5G)	Sub GHz (licenciadas)	1,4 MHz	-	1 Mb/s
Wi-SUN	Sub GHz (ISM)	200 kHz - 1,2 MHz	1 km	50 kb/s - 2,4 Mb/s
IEEE 802.11ah	Sub GHz (ISM)	1 - 16 MHz	1 km	150 kb/s - 78 Mb/s

A norma IEEE 802.11ah é uma emenda ao padrão IEEE 802.11 (Wi-Fi) que opera nas faixas sub 1GHz, desenvolvido pelo grupo de trabalho IEEE 802.11ah (TGah) para redes IoT com exigências de longo alcance e alta taxa de entrega de dados, tendo em vista as limitações de WLANs IEEE 802.11 para Internet das Coisas [Khorov et al. 2015]. O padrão é especificado pela norma em [IEEE 2017], onde são descritas as características físicas e as modificações na camada MAC do IEEE 802.11 legado [IEEE 2016].

2.1.1 Casos de Uso

Devido às suas características da camada física e funcionalidades da camada MAC, o IEEE 802.11ah atende a vários cenários de uso em sistemas IoT. De acordo com [Sun et al. 2013], as principais aplicações desta tecnologia podem ser classificadas em três categorias: Redes de sensores, Redes *backhaul* e Wi-Fi de alcance estendido.

2.1.1.1 Redes de Sensores

Com sua capacidade de alcançar longas distâncias a altas taxas de transmissão e com baixo consumo de energia, o IEEE 802.11ah é um destaque para a implantação de redes de sensores e monitoramento. Alguns exemplos de aplicações nas quais o padrão pode ser aplicado são: medidores inteligentes, monitoramento de ambientes e automação residencial

e industrial, cuidado com saúde. Neste cenário, os sensores associados a um ponto de acesso IEEE 802.11ah podem ser espalhados a longas distâncias para transmitir ou receber dados.

2.1.1.2 Redes *backhaul*

O padrão IEEE 802.15.4g é consistentemente adotado na indústria por ter baixo custo de manutenção. No entanto, o alcance de transmissão e a taxa de dados são muito inferiores se comparado a outras tecnologias. Para tirar proveito das vantagens do 802.15.4g e do IEEE 802.11ah ao mesmo tempo, é possível fazer uma conexão *backhaul* entre os dispositivos dessa arquitetura e servidores remotos por intermédio do IEEE 802.11ah. Assim, os dados coletados por dispositivos IEEE 802.15.4g são agregados em um roteador 15.4g-11ah que encaminha o fluxo agregado a um ponto de acesso IEEE 802.11ah para transmitir os dados para a internet.

2.1.1.3 Wi-Fi de Alcance Estendido

O IEEE 802.11ah pode ser uma opção para estender o alcance de uma rede Wi-Fi e reduzir *offloading* de tráfego em redes móveis graças ao longo alcance de transmissão, podendo ser aplicado em ambientes ao ar livre. Esta alternativa pode ser muito benéfica, principalmente em países com ampla disponibilidade de canal Sub-1GHz [Sun et al. 2013].

2.2 Camada Física IEEE 802.11ah

As especificações da camada física do IEEE 802.11ah são baseadas nas características do padrão IEEE 802.11ac, adaptadas para as frequências sub-1GHz. As larguras de banda suportadas são 1, 2, 4, 8 e 16 MHz, com as bandas de 1 e 2 MHz sendo obrigatórias. O padrão utiliza frequências não licenciadas na faixa sub 1GHz (como 863 - 868 MHz na Europa, 902 - 928 MHz na América do Norte e 755 - 787 MHz na China). As larguras de banda mais estreitas permitem o aumento da cobertura com menor consumo de energia em comparação a tecnologias Wi-Fi tradicionais [Tian et al. 2021].

Assim como o 802.11ac, o 802.11ah usa multiplexação por divisão de frequências ortogonais (OFDM), transmissão com múltiplas entradas e múltiplas saídas (MIMO, do inglês *Multiple Input Multiple Output*) e MIMO multiusuário no canal de descida (DL MU-MIMO, do inglês *Downlink Multiple User MIMO*) [Khorov et al. 2015, Sun et al. 2013]. O número de fluxos espaciais (NSS) pode ser um valor entre 1 e 4 e o intervalo de guarda pode ser $8 \mu s$ ou $4 \mu s$. Além disso, o padrão herda os 10 esquemas de modulação e codificação (MCS) com algumas adaptações de confiabilidade e taxa de transmissão do 802.11ac. Os esquemas de modulação e codificação para NSS = 1 e GI = $1 \mu s$ são listados na Tabela 2.2 para diferentes larguras de banda. Para a banda de 20 MHz, o esquema MCS9 não está disponível, assim

como para a banda de 20 MHz do 802.11ac. Para aumentar o alcance de canal em 1 MHz, o padrão define o MCS10, uma modificação do MCS0 que aplica repetição.

Tabela 2.2: Esquemas de Modulação e Codificação do IEEE 802.11ah para NSS = 1 e GI = 8 μ s.

Índice MCS	Modulação	Codificação	Taxa (kb/s)				
			1 MHz	2 MHz	4 MHz	8 MHz	16 MHz
0	BPSK	1/2	300	650	1350	2925	5850
1	QPKS	1/2	600	1300	2700	5850	11700
2	QPSK	3/4	900	1950	4050	8775	17550
3	16-QAM	1/2	1200	2600	5400	11700	23400
4	16-QAM	3/4	1800	3900	8100	17550	35100
5	64-QAM	2/3	2400	5200	10800	23400	46800
6	64-QAM	3/4	2700	5850	12150	26325	52650
7	64-QAM	5/6	3000	6500	13500	29250	58500
8	64-QAM	3/4	3600	7800	16200	35100	70200
9	64-QAM	5/6	4000	/	18000	39000	78000
10	BPSK	1/2 com repetição 2x	150	/	/	/	/

O IEEE 802.11ah suporta três diferentes unidades de dados PLCP : *SIG_IM*, usado para a largura de banda de 1 MHz; *SIG_SHORT* para as demais bandas em transmissões multiusuários e *SIG_LONG* para transmissões multiusuários e único usuário (SU) formado por feixes. Mais detalhes das características da camada física do IEEE 802.11ah podem ser consultados na norma [IEEE 2017].

2.3 Camada de Controle de Acesso ao Meio (MAC) do IEEE 802.11ah

A maior parte das inovações introduzidas no padrão IEEE 802.11ah está na camada MAC. Este padrão inclui um conjunto de novas funcionalidades com o objetivo de melhorar a eficiência energética, permitir o crescimento em escala da rede, e reduzir a sobrecarga da rede [Khorov et al. 2015]. Tais funcionalidades são aplicadas para otimização do tempo de autenticação e associação, conforme descrito na Seção 2.3.1, eficiência energética (Seções 2.3.2 e 2.3.3), controle de acesso ao meio (Seções 2.3.4 e 2.3.5), organização do AID (Seção 2.3.6), AID Dinâmico (Seção 2.3.7) e coexistência de BSSs (Seção 2.3.8).

2.3.1 Associação e Autenticação Rápidas

Ao ser inicializada, uma estação em uma rede sem fio inicia o processo de configuração inicial, no qual ela monitora o canal para identificar a presença de pontos de acesso e inicia o procedimento de autenticação, seguido do processo de associação. Após receber o quadro de *beacon* transmitido pelo ponto de acesso, a estação passa a saber da existência da rede e inicia o processo de autenticação. Para isto, a estação envia um requisição de associação (*Auth*

Req). Ao receber a requisição, o ponto de acesso pode aceitar ou não autenticar a estação, e este resultado é enviado na resposta de associação (*Auth Resp*). Caso a autenticação seja aceita, a estação então passa para o processo de associação, enviando, assim, uma requisição de associação (*Assoc Req*) ao AP, anunciando as funcionalidades por ela suportadas (chamadas de capacidades). Ao receber a requisição, o ponto de acesso responde com uma resposta de associação (*Assoc Resp*). Caso a associação tenha sido aceita, a resposta de associação deve conter o identificador de associação atribuído à estação. Todo este processo é ilustrado na Figura 2.1.

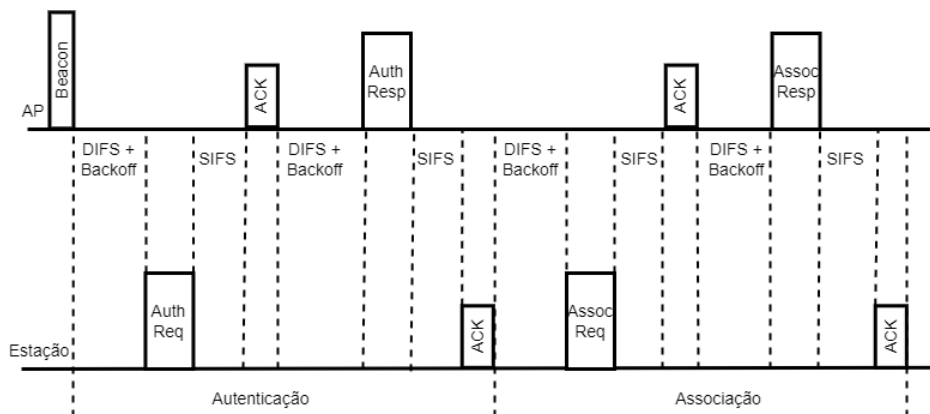


Figura 2.1: Processo de inicialização em redes IEEE 802.11ah. A estação inicia transmitindo uma requisição de autenticação (*Auth Req*) após notar a presença do AP ao receber o quadro de *beacon*. Após receber a requisição, o AP responde com uma resposta de autenticação (*Auth Resp*). Após o fim do processo de autenticação, a estação envia uma requisição de associação (*Assoc Req*). Ao receber a requisição, o AP envia uma resposta de associação (*Assoc Resp*).

Em redes com poucas estações, a função de coordenação distribuída (DCF, do inglês *Distributed Coordination Function*), baseada no CSMA/CA, é suficiente para o controle de acesso ao canal em situações que várias estações tentam associar-se ao mesmo tempo. No entanto, em redes de IoT, onde espera-se um grande número de dispositivos sem fio, o cenário em que todas as estações tentam associar-se ao mesmo tempo está muito mais propenso a colisões, fazendo com que o processo de associação e autenticação seja muito mais demorado. Como resolução deste problema, o IEEE 802.11ah apresenta dois novos mecanismos de controle de associação rápida: Controle Centralizado de Autenticação (CAC) e Controle de Autenticação Distribuído (DAC) [Tian et al. 2021].

No CAC, o ponto de acesso controla, de forma dinâmica, quais estações são permitidas a enviar requisições de associação em determinados períodos. Isto é feito através do envio de valores de limiar por meio dos *beacons* transmitidos em modo *broadcast* periodicamente. Assim que uma estação é inicializada, ela gera um valor aleatório entre 0 e 1022, e tenta enviar uma requisição de associação somente se este valor for menor que o limiar recebido do ponto de acesso. Caso contrário, a estação espera pelo próximo *beacon* com um outro valor de limiar e repete o mesmo procedimento de decisão.

No DAC, o ponto de acesso divide o intervalo de *beacon* em sub-intervalos chamados

slots de controle de associação (ACS). Cada estação escolhe um intervalo de *beacon* e um ACS para enviar a requisição de associação. Caso haja falha na transmissão do quadro, a estação tenta retransmitir no próximo p -ésimo intervalo de *beacon* e q -ésimo ACS, em que p e q são gerados com base no mecanismo de *backoff* exponencial binário.

2.3.2 Segmentação TIM

Nos padrões 802.11 tradicionais, o ponto de acesso transmite elementos TIM em cada *beacon* (chamado TIM *beacon*), anunciando para quais estações há pacotes na fila esperando para serem transmitidos. Ao receber o TIM, a estação verifica se está no conjunto de estações anunciadas, e caso não esteja ela volta para o estado de dormência (“*doze*”). Caso contrário, a estação envia um quadro PS-Poll requisitando os dados em espera na fila. Para as estações que recebem um baixo fluxo de dados, acordar periodicamente para receber *beacons* TIM não é energeticamente eficiente.

Para tornar o TIM ainda mais eficiente, o IEEE 802.11ah propõe o mecanismo de segmentação TIM, que divide as estações em grupos e cada um desses grupos é descrito pelo próprio TIM *beacon*. O *beacon* com o mapa de indicação de tráfego de entrega (DTIM) é usado para sinalização de tráfego em nível de grupos, enquanto o TIM sinaliza o tráfego em nível de estações de um determinado grupo. O DTIM é enviado periodicamente e as estações de todos os grupos devem acordar para recebê-lo e verificar se há pacotes pendentes para os seus respectivos grupos TIM. Caso haja, a estação se programa para acordar posteriormente para receber o TIM do seu grupo e verificar se há dados na fila do AP destinados a ela. Caso não haja pacotes para a estação, ela volta a dormir até o próximo DTIM *beacon*.

Um exemplo de operação deste mecanismo, presente na norma [IEEE 2017], é mostrado na Figura 2.2. No primeiro instante, o ponto de acesso transmite o DTIM e o TIM do primeiro grupo. O mapa de bits do DTIM aponta que há pacotes na fila apenas para o grupo TIM 7. As estações que não estão no TIM 7 verificam que não há pacotes para os seus respectivos grupos e voltam para o estado de dormência até o próximo DTIM. Já as estações 1538, 1539 e demais estações do grupo 7 acordam para receber o TIM 7 apontando que há pacotes na fila do AP apenas para a estação 1538. As estações do grupo que não estão sinalizadas no TIM voltam a dormir, enquanto a estação 1538 acorda para receber os dados em espera.

2.3.3 Target Wake Time

O *Target Wake Time* (TWT) é um mecanismo de economia de energia voltado para estações que trocam pacotes esporadicamente. Com este mecanismo, uma estação pode negociar com o ponto de acesso um intervalo de atividade chamado período de serviço (SP) TWT. Fora deste período, a estação permanece no estado de economia por muito tempo, não recebendo nem mesmo os *beacons* enviados pelo AP.

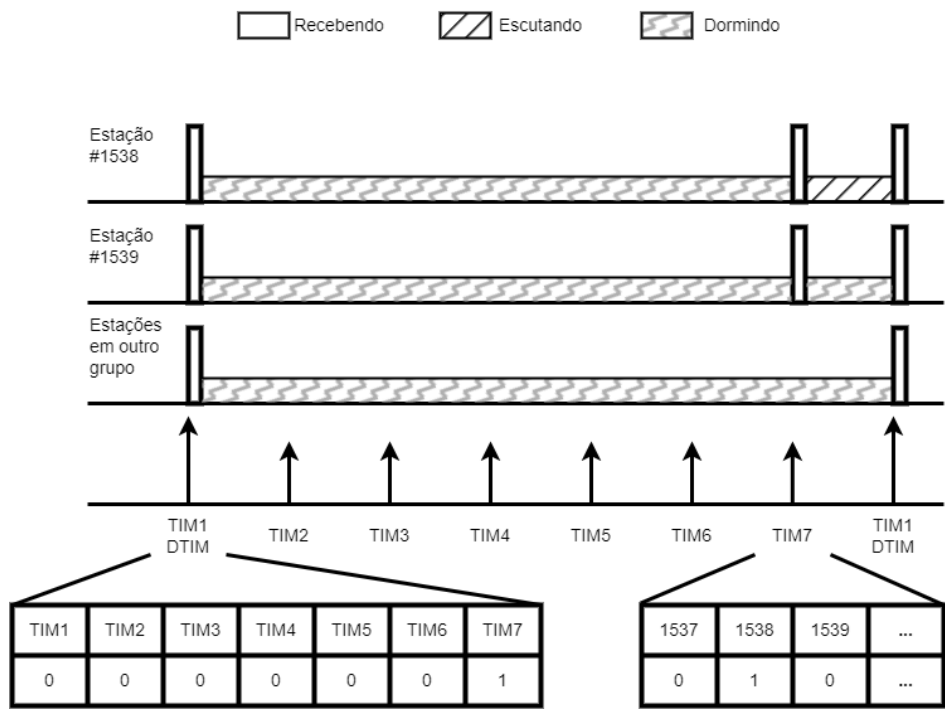


Figura 2.2: Exemplo de operação da segmentação TIM.

O processo de negociação pelo TWT SP poder ser iniciado tanto pelo ponto de acesso quanto pela estação através da troca de mensagens contendo parâmetros do TWT. Alguns desses parâmetros são *Target Wake Time*, indicando quando o intervalo TWT inicia; *Minimum Wake Duration*, indicando o valor mínimo de TWT SP; *Wake Interval*, indicando o tempo médio entre TWT SPs sucessivos e *Flow ID*, indicando um dentre os 8 possíveis TWTs que podem ser configurados pela estação em um mesmo fluxo [Khorov et al. 2015].

2.3.4 Janela de Acesso Restrito

Para minimizar o número de colisões, e com isto melhorar o desempenho em redes com grande número de estações, o IEEE 802.11ah propõe o método de controle de acesso ao meio chamado Janela de Acesso Restrito (RAW). Este método restringe o número de estações que podem competir pelo canal simultaneamente, dividindo-as em grupos e reservando janelas de tempo para cada grupo [IEEE 2017]. Assim, as estações de cada grupo somente têm acesso ao canal por intervalos específicos usando a função de coordenação distribuída (DCF) ou a função de acesso distribuído ao canal aprimorado (EDCA) .

O RAW divide o tempo de acesso ao meio em intervalos, com alguns deles sendo atribuídos a grupos RAW, enquanto o restante é compartilhado entre todas as estações. A Figura 2.3 retrata um exemplo de operação do RAW. A cada intervalo de *beacon*, o ponto de acesso anexa um conjunto de parâmetros RAW (RPS, do inglês *RAW Parameter Set*) ao quadro de *beacon* transmitido a todas as estações, contendo as informações relacionadas ao RAW, tais como quais são as estações de cada grupo e seus instantes de tempo de início de acesso. O RPS informa o conjunto de estações em cada grupo a partir do valor inicial e o valor final de

AID daquele grupo, e portanto as estações de um mesmo grupo devem ter valores de AID em sequência. Além disso, cada grupo RAW consiste de um ou mais slots no quais as estações são distribuídas de modo *round-robin* [Tian et al. 2021]. Outras informações como número de slots (N_{RAW}), formato de slot e contagem de duração de slot (C) fazem parte do RPS, e a duração (em micro-segundos) dos RAW *slots* é calculada por

$$D = 500 \mu s + C \times 120 \mu s. \quad (2.1)$$

O parâmetro C pode ser um número binário composto por $y = 11$ ou $y = 8$ bits, dependendo se o valor do formato do *slot* é 1 ou 0, respectivamente. O parâmetro N_{RAW} é um número de $14 - y$ bits. Para $y = 1$, o valor máximo de C é 2047, sendo a duração máxima de slot igual a 246,14 ms, e o número máximo de *slots* igual a 8. Para $y = 8$, o valor máximo de C é igual a 255 (correspondendo à duração de 31,1 ms), e o número máximo de *slots* é igual a 64.

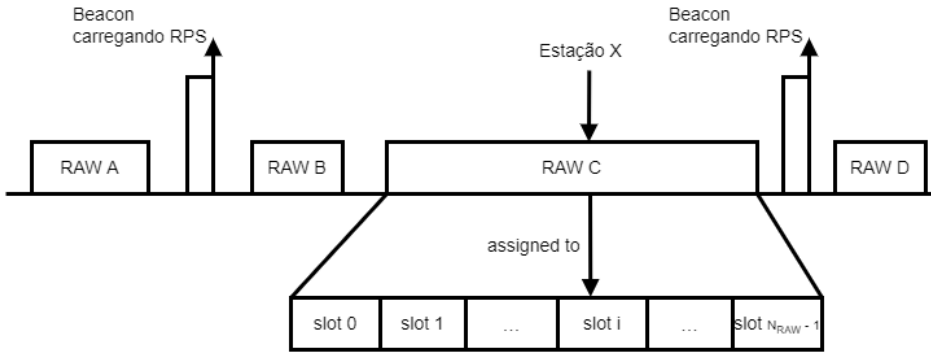


Figura 2.3: Exemplo de operação do RAW.

Ao receber o RPS, a estação define o índice i_{slot} do seu slot por meio da função de mapeamento

$$i_{slot} = (x + N_{offset}) \bmod N_{RAW}, \quad (2.2)$$

em que x pode ser o índice da posição do AID ou o próprio valor do AID da estação. O valor de x será o índice de posição caso o RAW seja restrito às estações cujos os bits de AID no elemento TIM são iguais a 1, conforme ilustrado na Figura 2.4. Caso contrário, x será igual ao valor do AID da estação. A variável N_{offset} representa o *offset* usado na função de mapeamento para fins de aumentar a justiça entre as estações [IEEE 2017]. A norma define que a estação deve usar os 2 octetos menos significantes do campo FCS do quadro de *SIG Beacon* como valor de N_{offset} .

Dentro de um RAW *slot*, o controle de acesso ao meio é feito por DCF ou EDCA. No entanto, diferente dos outros padrões IEEE 802.11, as estações 802.11ah usam duas funções de *backoff* distintas, conforme ilustrado na Figura 2.5. Nota-se que as estações STA 1 e STA 2 estão no mesmo grupo RAW, porém a STA 1 está no *slot 0*, enquanto a STA 2 está no *slot 1*. A primeira função de *backoff* é usada fora do RAW, durante o período de acesso livre

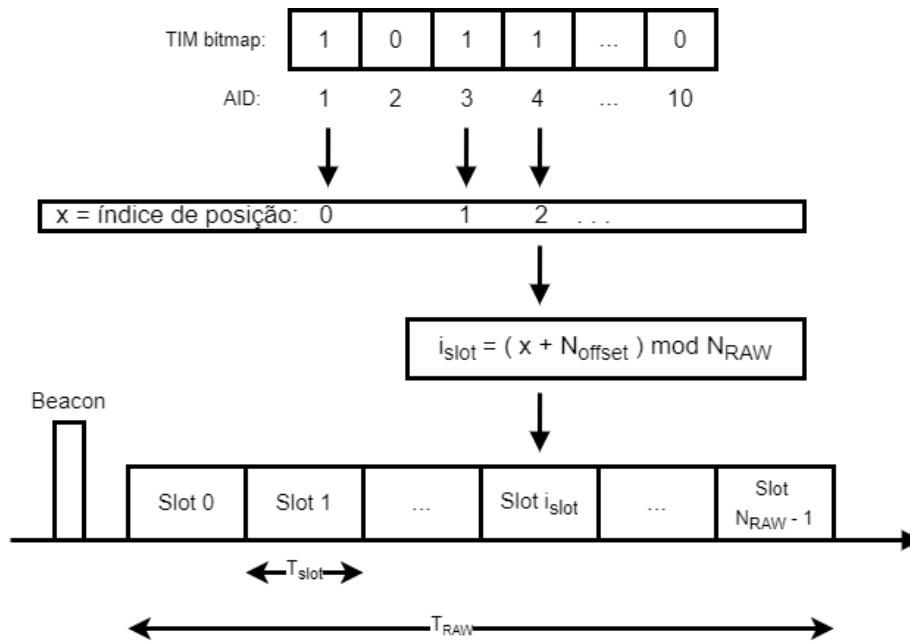


Figura 2.4: Demonstração de atribuição de RAW slot para o caso em que o RAW é restrito às estações com bit de AID igual a 1 no elemento TIM.

no qual todas as estações podem competir pelo canal. Após o recebimento do *beacon*, com o início da janela RAW, a estação 1 começa a competir pelo canal, assim como as demais estações do slot 0. As estações do *slot 0* salvam seus respectivos parâmetros (como contador de backoff, limite de tentativa, janela de contenção) da primeira função de *backoff*, e passam a competir pelo canal considerando uma segunda função de *backoff* com novos valores de estado. O mesmo acontece para a estação 2 no início do *slot 1*. Ao fim do período de RAW, todas as estações retomam o estado da primeira função de *backoff* e voltam a utilizá-la para o controle de acesso ao canal.

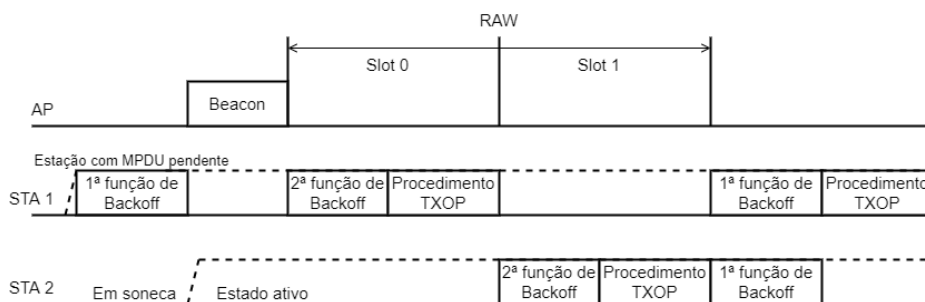


Figura 2.5: Processo de *backoff* para controle de acesso ao canal com RAW. As estações STA 1 e STA 2 estão no mesmo grupo RAW, porém em *slots* diferentes. A primeira função de *backoff* é utilizada fora do período do RAW, tendo seu estado interrompido e salvo para ser retomado após o fim do RAW. A segunda função de *backoff* é usada dentro do período RAW.

Outro campo importante do RPS transmitido às estações é o *Cross Slot Boundary (CSB)*. A Figura 2.6 ilustra o acesso para transmissões dentro de um RAW *slot* com o CSB desabilitado. Neste cenário, a estação não pode iniciar a transmissão de um quadro após um determi-

nado tempo de segurança, que deve ser no mínimo o valor da Oportunidade de Transmissão (TXOP) da estação [Tian et al. 2021]. A Figura 2.7 ilustra o cenário com o CSB habilitado. Neste caso não há um tempo de segurança, e a estação é permitida a iniciar a transmissão até o final do seu RAW *slot*, mesmo que esta transmissão cruze o limite entre seu RAW *slot* e o RAW *slot* seguinte.

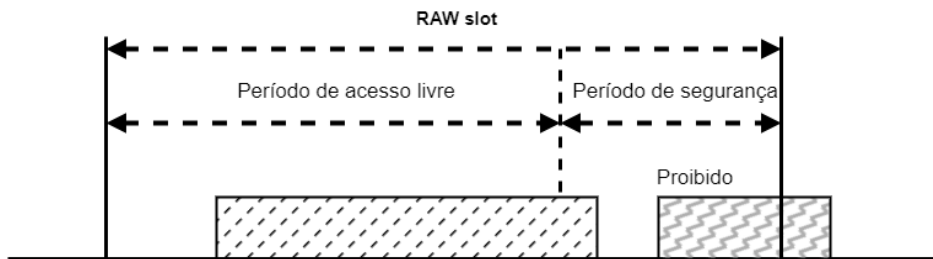


Figura 2.6: Transmissão sem cross slot boundary. A estação só pode iniciar a transmissão de um quadro se o tempo restante do RAW *slot* for maior que o tempo necessário para a transmissão do quadro e a recepção de uma resposta imediata.

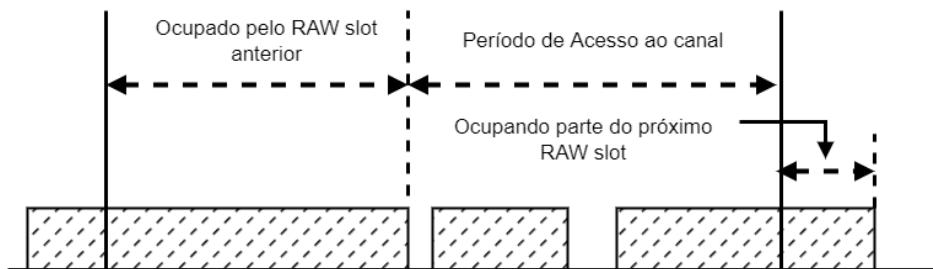


Figura 2.7: Transmissão com cross slot boundary. A estação pode iniciar a transmissão de um quadro em qualquer instante dentro RAW *slot*, e é permitido que a estação cruze o limite do RAW *slot* para completar a transmissão em andamento.

2.3.5 Setorização de grupos

Devido ao longo alcance e ao grande número de estações associadas, o problema do nó escondido é muito crítico em redes IEEE 802.11ah. Para mitigar este problema, o padrão propõe a setorização de estações com base em suas posições para controle de acesso ao canal. Este método baseia-se na combinação de multiplexação por divisão do espaço e do tempo [Khorov et al. 2015].

O ponto de acesso divide o espaço em setores, e cada estação associada é incluída em um setor no processo de associação por via do envio do identificador do setor ao qual a estação pertence. Com isto, o ponto de acesso divide o tempo de canal em intervalos de setores, e cada intervalo é iniciado após a transmissão do *beacon* que informa quais setores estão permitidos a acessar o canal dentro deste intervalo, conforme ilustrado na Figura 2.8. As estações que não pertencem aos setores anunciados são proibidas de acessar o canal. E as estações que recebem o identificador de grupo igual a 0 não pertencem a nenhum setor, e podem competir pelo canal a qualquer intervalo de beacon [IEEE 2017]. O ponto de acesso

também pode definir intervalos de setores omnidirecionais, durante os quais qualquer estação associada tem permissão para transmitir.

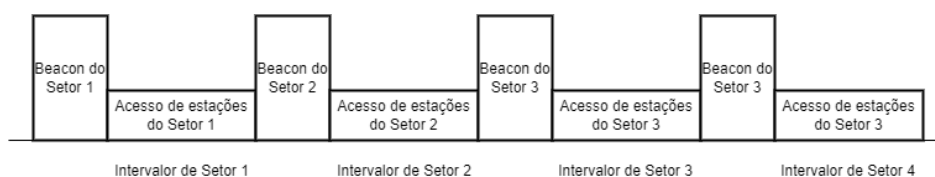


Figura 2.8: Exemplo de setorização de Grupos. Cada *beacon* transmitido às estações informa quais setores têm permissão de competir pelo canal no intervalo de *beacon* seguinte.

Diferente do RAW, em que não há sobreposição do tempo de canal de estações de diferentes grupos, dois ou mais setores podem acessar o canal no mesmo Intervalo de Setor.

2.3.6 AID Hierárquico

O identificador de associação (AID) é um valor único de 14 bits que é atribuído a cada uma das estações no processo de associação. Em padrões de Wi-Fi tradicionais, o AID só pode assumir valores entre 1 e 2007, pois 0 e 2008-16383 são valores reservados. O AID 0, por exemplo, é reservado para sinalizar tráfego de grupos.

Para suportar redes com grande número de dispositivos, o AID do padrão IEEE 802.11 passa a ser um número de 13 bits que pode assumir até 8191 valores. Para simplificar as operações com este número de estações associadas, o padrão propõe o mecanismo de organização hierárquica de AID. Neste mecanismo, os AIDs de 13 bits são organizados em uma estrutura com os seguintes 4 níveis: páginas de 2 bits, blocos de 5 bits, sub-blocos de 3 bits e estações de 3 bits. Assim, as estações são divididas em N_p páginas com N_b blocos cada, com cada bloco contendo 8 sub-blocos de 8 estações cada. Os valores de N_p e N_b são configuráveis, e um exemplo de configuração com $N_p = 4$ e $N_b = 32$ é mostrado na Figura 2.9. Seguir esta hierarquia no referenciamento de estações agrupadas de acordo com suas características reduz a sobrecarga na rede [Tian et al. 2021].

2.3.7 AID Dinâmico

No IEEE 802.11 legado, o AID é um valor fixo atribuído unicamente a cada estação enquanto ela estiver associada. O padrão IEEE 802.11ah prevê a funcionalidade de atribuição dinâmica do AID às estações. Ou seja, é possível alterar os valores de AID das estações enquanto elas estão associadas, de acordo com a especificação na norma [IEEE 2017].

Durante a associação, a estação e o ponto de acesso trocam informações de capacidades, o que inclui a informações de suporte ao AID dinâmico, indicado pelo campo *Dynamic AID* igual a 1. Estações que suportam AID dinâmico são definidas na norma como estações D-AID. Na operação desta funcionalidade, estações D-AID trocam mensagens de requisição

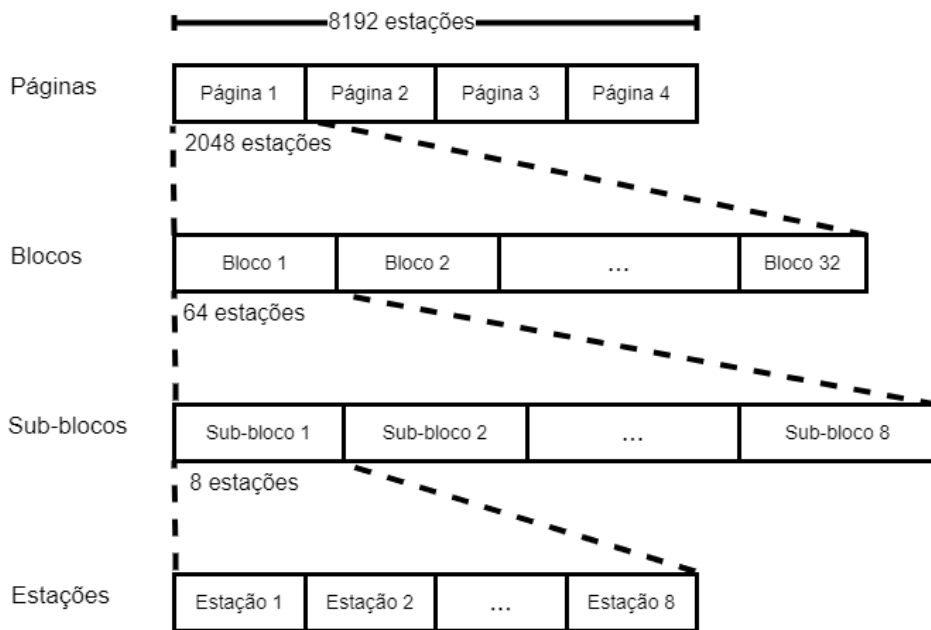


Figura 2.9: Exemplo de configuração do AID hierárquico. Os 8192 valores possíveis de AID são divididos em 4 páginas, com cada página composta por 32 blocos. Cada bloco é dividido em 6 sub-blocos contendo 8 valores de AID cada.

de troca de AID (*AID Switch Request*) e resposta de troca de AID (*AID Switch Response*).

As estações IEEE 802.11ah podem ser agrupadas de acordo com suas características de serviço, perfil de tráfego ou modo de energia. Ao perceber uma mudança de tráfego, nas características de serviço ou na carga restante da bateria, uma estação pode transmitir uma requisição de troca do seu AID para um novo valor, de modo que ela passe a acessar o canal em um novo modo aplicado a um grupo diferente de estações [IEEE 2017]. Em uma requisição de troca de AID, a estação D-AID informa sua característica de serviço, caso tenha havido mudança, além do novo intervalo de escuta ao canal do grupo ao qual a estação espera ser incluída (sendo também um campo opcional). O ponto de acesso então responde com uma resposta de troca de AID contendo o novo AID para a estação. Um exemplo da operação de AID dinâmico é mostrado na Figura 2.10. A estação inicia com um valor de AID atribuído na associação entre 200 e 299, pertencendo ao grupo TIM 2, com intervalo de despertar de 2 intervalos de *beacon*. Em um determinado instante, a estação percebe uma mudança de tráfego e então envia uma mensagem de requisição de troca de AID ao AP para que seja transferido para o grupo TIM 3. O ponto de acesso responde com uma mensagem de resposta de troca de AID contendo um novo valor de AID entre 300 e 399. Após esta operação, a estação passa a escutar o canal no mesmo modo que as estações do TIM 3, com intervalo de despertar igual a 3 intervalos de *beacon*.

Tabém é possível que o ponto de acesso mude o AID de uma estação D-AID para trocá-la de grupo sem que a estação precise enviar uma mensagem de requisição de troca. Para isto, o AP deve suportar a funcionalidade de AID dinâmico *não solicitado*, que também deve ser informada nas capacidades das estações a partir do campo *Unsolicited Dynamic AID* igual a 1. Neste modo, o ponto de acesso envia um mensagem de resposta de troca de

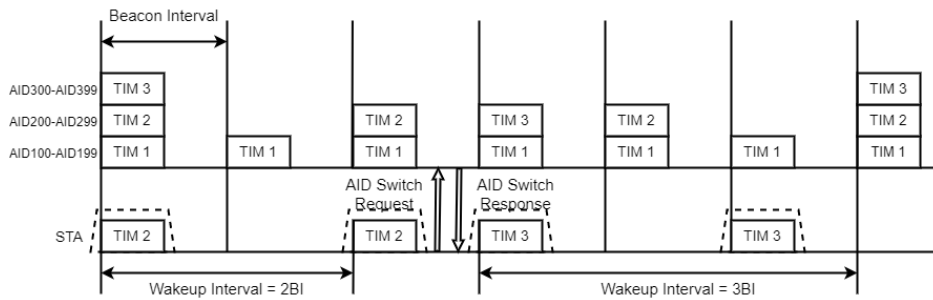


Figura 2.10: Exemplo de operação do AID dinâmico.

AID não solicitada a uma determinada estação com um novo valor de AID. Ao receber esta mensagem, a estação muda o seu AID para o valor informado na mensagem, desde que ela também tenha suporte para AID dinâmico não solicitado.

2.3.8 Cor de BSS

Como redes IEEE 802.11ah podem cobrir um raio de até 1km, os cenários em que há alta densidade de dispositivos, a operação de varios Conjuntos Básicos de Serviço (BSS) na mesma frequência, e na mesma área, pode resultar em interferências causadas pela coexistência de diferentes BSSs. Para mitigar este problema, o IEEE 802.11ah propõe o recurso de Cor de BSS. Este método atribui a cada BSS uma cor única (em termos de bits designados no campo LSIG no cabeçalho da camada física) [Tian et al. 2021, Baños-Gonzalez et al. 2016]. Durante a recepção de um pacote, caso a estação perceba que o pacote tenha uma cor diferente da cor da sua BSS, ela interrompe a recepção do pacote para reduzir consumo de energia e evitar interferência.

2.3.9 Cabeçalho MAC curto

No IEEE 802.11 legado, o tamanho do cabeçalho MAC é de 30 bytes, conforme mostrado na Figura 2.11, além de 4 bytes de Sequência de Checagem de Quadro (FCS) [IEEE 2016]. Para quadros de dados com 100 bytes, o cabeçalho representa mais de 30% do tamanho do quadro, e essa porcentagem é ainda maior para quadros menores, o que pode representar uma sobrecarga na rede. Para reduzir esta sobrecarga, o IEEE 802.11ah propõe um novo formato de cabeçalho MAC para quadros de dados, de gerenciamento e de controle, com tamanho entre 10 e 24 bytes, como mostrado na Figura 2.12. Neste novo formato, os campos de Duração/ID, Controle QoS e Controle HT são excluídos, e o endereço de 6 bytes é substituído por um identificador curto (SID) de 2 bytes. Para *downlink*, o endereço 1 (que é o endereço de destino no MAC legado) é substituído pelo AID de destino. E para *uplink*, o endereço 2 (que no cabeçalho MAC legado é o endereço da fonte) é substituído pelo AID da fonte.

O IEEE 802.11ah pode usar os dois formatos de cabeçalhos, e para distinguir o tipo do ca-

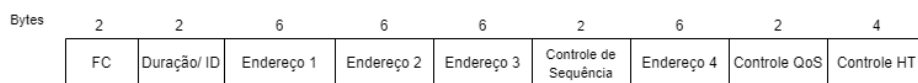


Figura 2.11: Cabeçalho MAC tradicional.

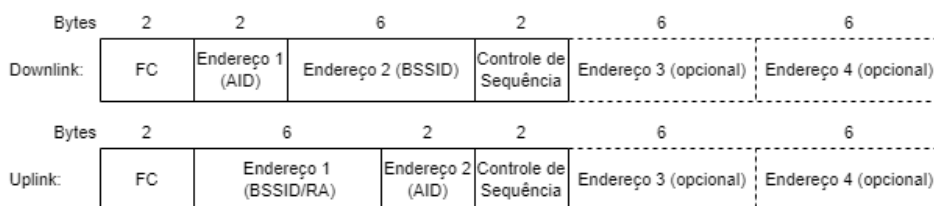


Figura 2.12: Cabeçalho MAC curto.

beçalho, os bits de versão no campo de controle de quadro (FC) foram modificados de 00 para 01, tratando-se da primeira alteração na versão do protocolo de Wi-Fi [Khorov et al. 2015].

Assim como o cabeçalho MAC legado, o cabeçalho curto também pode conter um quarto endereço. Porém, diferente dos outros padrões IEEE 802.11, no 802.11ah apenas dois endereços são obrigatórios, e não três. Os três endereços são necessários apenas quando o quadro é enviado entre estações do mesmo BSS via o ponto de acesso. Quatro endereços (receptor, transmissor, fonte e destino) só são necessários para redes *multihop*.

2.4 Algoritmo K-means

O K-Means é um algoritmo de aprendizado de máquina não supervisionado que faz o agrupamento de N pontos de dados em k grupos [Alpaydin 2014, Sieranoja 2019]. Apesar de existirem algoritmos de agrupamento conhecidos que, em muitas situações, resultam em melhores resultados de agrupamento, o K-Means ainda é um algoritmo popular por ser simples de implementar e por ter sido objeto de diversos estudos, sendo, portanto, bem conhecido.

O algoritmo é executado de modo iterativo, não sendo possível resolver o problema de otimização analiticamente. Inicialmente, k pontos são escolhidos aleatoriamente como centroides para os k grupos iniciais. Em seguida, é executada a fase de atribuição, onde cada ponto é incluído no grupo com o centroide mais próximo, com base na distância euclidiana dada por $d_{c_i} = \|\mathbf{x}^t - \mathbf{c}_i\|$, em que \mathbf{x}^t é o vetor que representa um ponto de dado e \mathbf{c}_i é o centroide do i -ésimo grupo. Após a atribuição de todos os pontos, a etapa de atualização é executada, em que o centroide de cada grupo é recalculado como a média de todos os pontos atribuídos a este grupo na etapa de atribuição. As etapas de atribuição e atualização são executadas a cada iteração, a um número máximo de iterações pré-definido.

Uma limitação do algoritmo K-Means é que o resultado final do agrupamento tem grande dependência dos centroides escolhidos na etapa de inicialização, já que trata-se de um procedimento de busca local e, portanto, os centroides não se movimentam pelos grupos se esti-

verem muito distantes [Alpaydin 2014]. Dessa forma, uma má escolha inicial dos centroides faz com que o agrupamento fique preso em mínimos locais inferiores. Em [Sieranoja 2019], é feito o estudo de como a inicialização dos centroides e o número de repetições impactam o resultado final do algoritmo. No trabalho, são comparados diferentes métodos de inicialização, incluindo a atribuição aleatória citada anteriormente. Os resultados mostram que técnicas como a Heurística do Ponto Mais Distante (*Maxmin*) e a Heurística de Ordenação (*Sorting*) conseguem apresentar grande melhoria no agrupamento. A qualidade do agrupamento é avaliada conforme duas medidas de desempenho: o Índice de Centroide, que corresponde ao número de centroides que foram corretamente calculados, e a taxa de sucesso, que mede quantos pontos foram corretamente atribuídos aos seus respectivos grupos.

A Heurística do Ponto Mais Distante [Gonzalez 1985] inicia escolhendo um ponto arbitrário como primeiro centroide e, em sequência, adiciona os demais centroides um a um. A cada etapa, o próximo centroide é o ponto mais distante (max) do centroide mais próximo existente (min). Este algoritmo tem complexidade $O(k^2N)$, que pode ser reduzida para $O(kN)$, dependendo do modo de implementação. Na implementação com complexidade $O(kN)$, para cada ponto é mantido um ponteiro para o seu centroide mais próximo. A cada novo centroide escolhido, é calculada a distância entre este centroide e cada um dos pontos de entrada. Para cada ponto, se a distância for menor que a distância do centroide anterior mais próximo, o ponteiro do ponto é atualizado para o novo centroide.

Na Heurística de Ordenação, os pontos são ordenados de acordo com um determinado critério. Esta operação tem complexidade $O(N \log N)$, sendo menor que $O(kN)$ assumindo que $\log N \leq k$ [Sieranoja 2019]. Os k centroides são escolhidos utilizando-se uma das seguintes heurísticas:

- Primeiros k pontos.
- Primeiros k pontos não permitindo que os novos pontos estejam mais próximos que ϵ dos pontos já escolhidos.
- Cada (N/k) -ésimo ponto.

Além disso, diferentes critérios de ordenação podem ser considerados:

- Distância do ponto central.
- Densidade.
- Centralidade.
- Atributo com maior variância.

Os dois algoritmos apresentaram bons resultados em [Sieranoja 2019] e com baixa complexidade de tempo de execução e de implementação. Além desses, Sieranoja, S. apresenta outros algoritmos com bons resultados, porém com maior complexidade.

Resumo

O padrão de comunicação IEEE 802.11ah opera nas faixas sub-1GHz e foi desenvolvido para ser aplicado em redes de internet das coisas. O padrão suporta até 8191 estações associadas a um único AP, que consegue transmitir a longas distâncias. O 802.11ah herda as características físicas do 802.11ac, porém para as frequências sub-1GHz, e portanto suporta 10 esquemas de modulação e codificação (MCS). A norma IEEE 802.11ah propõe um conjunto de novas funcionalidades para a camada MAC com os objetivos de melhor eficiência energética, melhoria da vazão, diminuição da latência e diminuição do número de colisões. Este capítulo apresentou os principais conceitos teóricos relacionados ao IEEE 802.11ah que serviram como base para o desenvolvimento deste trabalho. Adicionalmente, foi apresentada a definição do algoritmo de agrupamento K-Means, que é proposto como uma estratégia de agrupamento das estações em grupos RAW.

Capítulo 3

Propostas de Otimização do RAW

Como a norma do IEEE 802.11 ah não especifica nenhuma abordagem obrigatória para o agrupamento de estações, existem diversos trabalhos na literatura que propõem estratégias de otimização dos parâmetros RAW baseadas em uma ou mais medidas de desempenho de interesse. Em [Tian et al. 2021], as abordagens propostas são classificadas de acordo com os objetivos dos algoritmos em otimizar as seguintes medidas de desempenho: eficiência energética, latência, probabilidade de colisão e de transmissão bem-sucedida, colisão por nós escondidos, qualidade de serviço (QoS), vazão e justiça. Esta classificação não é restrita, pois uma mesma abordagem pode ter dois ou mais objetivos para otimização. Nas seções seguintes deste capítulo, serão abordados alguns trabalhos propostos para cada um dos objetivos listados.

3.1 Eficiência Energética

A eficiência energética para fluxo *uplink* é avaliada para um único RAW em [Wang et al. 2015]. São considerados dois cenários: 1) atribuição de uma única estação por slot e 2) atribuição de várias estações ao mesmo slot. Para os dois casos, é calculada a probabilidade de um dispositivo transmitir um único pacote com sucesso para estimar o consumo energético. Com isto, o trabalho propõe um modelo que estima a eficiência energética baseado no número de dispositivos por RAW e o tempo de RAW slots. O modelo não considera o ordenamento de estações para permutação e agrupamento de estações a partir das medidas de vazão e consumo energético.

Em [Kai et al. 2019], é assumida a distribuição uniforme de estações em RAW slots. Um modelo da cadeia de Markov é aplicado para a análise de vazão e consumo de energia, e para argumentar que a eficiência energética mínima de um grupo é maximizada quando as características de tráfego das estações de um mesmo grupo são as mesmas. O Algoritmo Ótimo de Agrupamento por Tráfego (OTGA) é proposto com base no método de Ramificar e Limitar (BBM) para determinar o melhor esquema de agrupamento por tráfego. Além disso,

é proposto o algoritmo heurístico de mapeamento de sensor de tráfego (HTMA) que agrupa as estações de acordo com sua característica de tráfego.

3.2 Latência e Qualidade de Serviço

Em [Khorov et al. 2020], é feito o estudo de otimização do RAW para cenários de ambientes com transmissão de alertas de eventos de emergência, como vazamento químico, aumento do nível de radiação, incêndios e inundação em fábricas ou usinas nucleares. Para estes ambientes, o atraso e a confiabilidade de entrega são de extrema importância. Portanto, o trabalho propõe um modelo matemático para estes cenários que faz a seleção dos parâmetros RAW de modo a diminuir o compartilhamento de tempo de canal.

Em [Mahesh and Harigovindan 2019], as estações são agrupadas de acordo com seus requisitos de vazão para transmitir dados durante o período RAW. Esses requisitos são estimados pelas próprias estações antes do processo de associação e são então enviados ao AP no campo opcional do cabeçalho MAC curto. Com as informações de todas as estações, o ponto de acesso determina os parâmetros RAW (tempo de RAW, número de slots, etc). Em seguida, são definidos valores de prioridade para cada grupo criado. O número de slots reservados para cada grupo depende do valor de prioridade estimado.

3.3 Probabilidade de Colisão e Terminais Escondidos

Com o objetivo de eliminar o problema de colisões em redes IEEE 802.11ah, [Nabuuma and Alsusa 2017] propõe um método de configuração dos temporizadores de *backoff* baseada nos valores de AID das estações nos grupos. No trabalho, é assumido que o mecanismo de Setorização em grupos está habilitado, resolvendo, assim, o problema de nós escondidos. No modelo, as estações devem detectar a presença das demais estações através do recebimento de sinais a uma potência mínima de detecção. Essas informações são passadas ao AP que, com isto, cria um grafo baseado na potência recebida. O agrupamento RAW é feito a partir do algoritmo *Welsh Powell*. Cada estação, ao ser incluída em um grupo RAW, deve utilizar valores de *backoff* baseados nos índices de seus respectivos AIDs no mapa de bits. Este trabalho não considera a atribuição dinâmica de AID para o recálculo do modelo em cenários com mobilidade.

Para mitigar o problema de colisões causadas por nós escondidos, [Damayanti et al. 2016] propõe uma solução formada por dois componentes: um esquema de mitigação de cadeia de colisão e um algoritmo de agrupamento consciente com reconhecimento de dispositivos escondidos. O esquema de mitigação de colisão detecta colisões ao monitorar blocos de energia no canal medindo a duração de tempo das colisões. Com isto, é possível detectar a presença de nós escondidos ao comparar a duração da

colisão com um limiar definido, e esta informação é transmitida em *broadcast* a todos os dispositivos para mitigar a cadeia de colisão detectada. O agrupamento é feito a partir de tabelas pré-definidas de sensibilidade da portadora. A configuração dos grupos pode ser fixa por toda a operação ou pode haver um reagrupamento das estações. Porém, o trabalho não menciona o uso de atribuição dinâmica de AID para o segundo caso.

Em [Hu 2019], o agrupamento de nós é inicialmente formulado como um Problema de Programação Inteira 0/1 (0/1 ILP) para minimizar pares de estações ocultas em redes IEEE 802.11 ah. Diferentemente de outras abordagens já propostas que consideram apenas estações ocultas em um mesmo RAW *slot*, esta abordagem é baseada nas características de tráfego das estações. O objetivo é minimizar o número de pares de estações ocultas em um mesmo grupo RAW.

3.4 Vazão

Com o objetivo de melhorar o resultado de vazão em redes IEEE 802.11ah, [Nawaz et al. 2017] propõe um modelo analítico para determinar a duração dos RAW slots com base no tamanho (em número de dispositivos) de sub-grupos em um mesmo RAW. Cada intervalo de RAW é dividido em dois sub-intervalos formados por um conjunto de RAW slots. Neste caso, é possível que dois slots em um mesmo RAW possuam durações diferentes, o que não está de acordo com as especificações da norma [IEEE 2017].

Em [Chang et al. 2019], os dispositivos são agrupados em um número fixo de grupos de acordo com suas demandas de tráfego, definidas pelo taxa de amostragem (considerando que os sensores transmitem dados periodicamente ao AP), o tamanho de pacotes e a taxa de transmissão. O agrupamento é feito pelo algoritmo de agrupamento com reconhecimento de tráfego ganancioso, que busca maximizar a utilização de canal do grupo com menor expectativa de utilização do canal. Comparado à estratégia de agrupamento aleatória uniforme, o algoritmo apresenta melhores resultados de utilização do canal. No entanto, a atribuição dinâmica não é mencionada para a realocação de estações em ambientes com dinamismo de tráfego e de condições de canal.

Em [Tian et al. 2017a], o algoritmo de otimização RAW adaptável ao tráfego (TAROA) é proposto para ajustar os parâmetros RAW em tempo real com base nas condições de tráfego da rede. O algoritmo foi desenvolvido para redes com tráfego heterogêneo e dinâmico, em que as estações podem alterar seu intervalos de transmissão a qualquer momento. Os dois principais procedimentos executados pelo algoritmo são: (1) a estimativa do intervalo de transmissão de cada nó da rede e (2) a atribuição das estações aos grupos RAW de acordo com os valores estimados. Além disso, o número ótimo de estações por RAW slot é estimado por meio da execução de simulações no ns-3. Uma extensão do TAROA (nomeada E-TAROA) é proposta em [Tian et al. 2017b], em que o algoritmo estimador de tráfego é melhorado ao explorar o campo “*More Data*” do cabeçalho MAC curto. Nos dois traba-

lhos, a atribuição de estações aos grupos RAW é limitada às estações com valores de AID sequenciais, não havendo uso do mecanismo de atribuição dinâmica de AID.

Em [Mahesh and Harigovindan 2019], o agrupamento de estações é baseado nas taxas de transmissão nominais (definidas pelos esquemas de codificação e modulação MCS). A análise de vazão e eficiência energética é feita em um cenário onde a área de cobertura da rede é divididas em zonas geográficas com diferentes características de tráfego. As estações com mesma taxa de transmissão são colocadas no mesmo grupo, resultando em melhor performance de vazão agregada e economia de energia.

3.5 Justiça

Uma solução de agrupamento baseada na taxa de dados é proposta em [Mahesh et al. 2020], que explora o uso de uma rede neural de mapas auto-organizados (SOM) para o agrupamento de estações. Os valores das taxas de transmissão são obtidos por meio do cabeçalho PLCP, antes do procedimento de associação, e o algoritmo de agrupamento só é executado após todas as estações estarem associadas. Esta solução não se aplica a cenários com tráfego dinâmico, pois uma vez que os grupos são definidos, eles não são modificados pelo resto da operação. O reagrupamento das estações seria possível com a atribuição dinâmica de AIDs, podendo ser aplicada a cenários com MCS adaptativo.

Em [Sangeetha and Babu 2020], a informação de taxa contida no cabeçalho PLCP também é usada como parâmetro de agrupamento de estações. À medida que as estações se associam ao AP, elas são agrupadas de acordo com sua taxa de dados nominal, e o valor de AID atribuído à estação é derivado do índice do grupo ao qual ela foi inserida. Neste cenário não há reagrupamento das estações, logo esta solução não suporta redes com tráfego dinâmico.

Resumo

Uma das funcionalidades propostas pelo padrão IEEE 802.11ah para a camada MAC é a janela de acesso restrito (RAW). Este esquema divide as estações associadas em grupos com períodos reservados para competir pelo canal dentro de um intervalo de *beacon*. O RAW opera a partir da definição e transmissão do conjunto de parâmetros RAW (RPS), que inclui atributos como o número de slots em que o período de RAW é dividido, o instante de início de cada RAW, as estações em cada RAW, etc. A norma não especifica como esses parâmetros devem ser definidos, e nem como as estações devem ser distribuídas entre os grupos. Neste capítulo foram apresentadas algumas propostas de estratégias de agrupamento com o objetivo de otimizar uma ou mais medidas de desempenho em redes IEEE 802.11ah. As abordagens podem ser baseadas em consumo energético, tempo de *backoff*, detecção

de colisões, qualidade de serviço, características de tráfego e taxa de transmissão. Pela revisão bibliográfica realizada, foi identificado que nenhum trabalho até agora fez uso do procedimento de atribuição dinâmica de AID e não apresentou propostas para agrupamento dinâmico de estações baseado neste mecanismo.

Capítulo 4

Implementação das Propostas

Este capítulo apresenta a implementação do recurso previsto na norma IEEE 802.11ah para alteração dinâmica de AID de estações via envio de mensagens do tipo *AID Switch Response*. A implementação da técnica é feita no simulador ns-3 como contribuição inédita ao módulo específico existente para o IEEE 802.11ah. Adicionalmente, foi implementado e incorporado o algoritmo K-Means para funcionar como estratégia de agrupamento de estações tendo como entrada um conjunto de parâmetros (diretos e indiretos) obtidos das estações.

A implementação proposta envolve a alteração de classes (do C++) que englobam as operações realizadas na camada MAC do modelo de redes Wi-Fi do ns-3. As alterações foram feitas em cima da implementação proposta por Le Tian et al. [Tian et al. 2016], que é baseada na versão 3.25 do simulador, porém com suporte ao modelo de redes IEEE 802.11ah. O algoritmo K-Means foi incorporado ao simulador como um módulo que é diretamente acionado na implementação da camada MAC seguindo a lógica de operação do agrupamento de estações. Levando em conta as limitações deste algoritmo K-Means, três técnicas de inicialização dos centroides foram implementadas e comparadas para que seja possível tirar o melhor proveito do algoritmo de agrupamento.

4.1 Simulador NS-3

O ns-3 é um simulador de rede a eventos discretos de código aberto e gratuito, desenvolvido para fins educacionais e de pesquisa [Riley and Henderson 2010, nsnam 2021]. O simulador é escrito em C++ e possui um conjunto vasto de modelos que simulam elementos de redes de computadores, como nós, dispositivos de rede, canais de comunicação, protocolos de comunicação, cabeçalhos e pacotes. Esses modelos são projetados com o objetivo de serem realistas, de modo que os modelos implementados possam ser emulados para serem utilizados em cenários reais. Além disso, o simulador fornece ferramentas que auxiliam desde a definição de cenários de simulações até o rastreamento de valores de parâmetros, medidas de desempenho e variáveis.

O módulo Wi-Fi do ns-3 é um dos mais populares do simulador em termos de números de pesquisas em que é utilizado [nnsam 2021]. Este módulo possui uma implementação acurada da camada MAC conforme as especificações da norma IEEE 802.11. A implementação da camada MAC em um nível mais inferior inclui modelos de funções de acesso ao meio (DCF e EDCA), de proteção de quadros (como RTS/CTS) e de reconhecimento (ACK/BlockACK). Já a implementação da camada MAC em nível superior inclui a geração de *beacon*, a operação de sondagem, o processo de associação entre estação e ponto de acesso e algoritmos de controle de taxa. A implementação da norma IEEE 802.11ah inclui as especificações da camada física dos padrões 802.11a/b/e/g/n/ac/ax/be, que são modelam as características específicas de cada emenda, bem como as operações e funções gerais da camada física (como modelos de perda e atraso de propagação).

No ns-3, o nó é uma abstração para qualquer ponto final em uma rede. Um mesmo nó pode possuir um ou mais dispositivos Wi-Fi, e diferentes dispositivos podem operar em diferentes modos e sob operação definida por diferentes padrões. O esquema da Figura 4.1 ilustra a arquitetura de um dispositivo Wi-Fi, representado pela classe *WifiNetDevice*. Esta classe engloba o conjunto de operações da camada MAC e da camada física de dispositivos Wi-Fi. A camada MAC é implementada em duas sub-camadas, *Mac High* e *Mac Low*. A camada *Mac High* refere-se aos modos de operação e funções do nó na rede (ponto de acesso, estação ou modo ad-hoc), além de cuidar dos procedimentos de associação, troca de beacon, etc. Na versão modificada por Le Tian et al. [Tian et al. 2016, Tian et al. 2018], é nesta camada que estão implementadas as operações do mecanismo de RAW, Segmentação TIM e configuração do TWT do IEEE 802.11ah.

A camada *MAC Low* lida com enfileiramento, fragmentação e retransmissão de pacotes. A classe *FrameExchangeManager* implementa sequências de trocas de quadros introduzidos pelo padrão IEEE 802.11 em uso. Além disso, cuida da agregação de quadros, retransmissões, proteção e reconhecimento. A classe *ChannelAccessManager* implementa as funções de DCF e EDCA. As classes *Txop* e *QosTxop* lidam com as filas de pacotes. A classe *Txop* pode ser usada pela camada MAC superior que não têm QoS habilitado, e para transmissão de quadros cujo o padrão determina que sejam enviados por DCF, como quadros de gerenciamento. Já a classe *WosTxop* é usada pela camada MAC superior com QoS habilitado.

As principais classes que implementam a camada física são *WifiPhy* e *WifiChannel*. A classe *WifiPhy* cuida do recebimento e envio de quadros, consulta os modelos de interferência e erro para verificar a probabilidade de recebimento de um quadro baseado na razão sinal-interferência-mais-ruído (*Signal-to-interference-plus-noise ratio* - SINR). A classe *WifiChannel* cuida de alterar o sinal para todas as entidades físicas consultando o modelo de perda de propagação e o modelo de atraso de propagação. Mais detalhes da implementação dos modelos citados podem ser encontrados na documentação oficial do simulador [nnsam 2021].

Para a proposta deste trabalho, foram alteradas e adicionadas novas classes à implementação do IEEE 802.11ah feita por Le Tian et. al [Tian et al. 2016] no ns-3, versão 3.25. Os

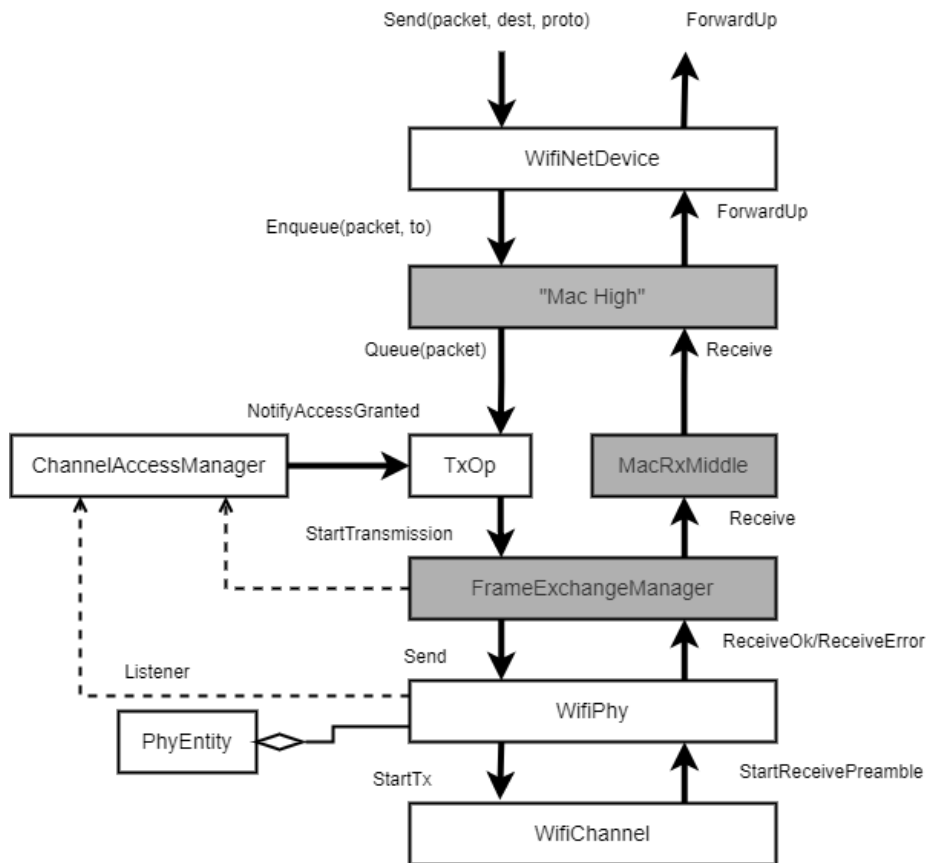


Figura 4.1: Estrutura modular da implementação de dispositivos Wi-Fi no ns-3. Cada bloco corresponde a uma classe que inclui um conjunto de operações das camadas física e MAC do IEEE 802.11. Em cinza, estão os blocos das classes que foram alteradas no desenvolvimento deste trabalho.

módulos em destaque na Figura 4.1 foram alterados para dar suporte à operação de AID Dinâmico, a execução do algoritmo de K-Means e agrupamento resultante, e a medição de parâmetros de tráfego e de condição do canal.

4.2 Associação de Estações

Na versão do simulador desenvolvida por Le Tian et al., o processo de associação ocorre conforme o procedimento de associação rápida descrito na Seção 2.3.1. No entanto, a implementação não inclui a etapa anterior de autenticação rápida especificada pela norma. Ou seja, as estações partem para o envio da mensagem de requisição de associação assim que inicializam. Ao receber a requisição de associação, o ponto de acesso avalia se a associação é possível e, em seguida, cria um quadro de resposta de associação para enviar à estação. Caso a associação seja possível, um valor de AID é atribuído a esta estação e incluído na mensagem de resposta. Por conveniência, os desenvolvedores optaram por utilizar os 13 bits menos significativos do endereço MAC da estação como seu número de AID. Todo este procedimento é executado por funções da classe *ApWifiMac*, que representa as operações na

camada MAC superior de um ponto de acesso.

Neste trabalho, o esquema de escolha do AID baseado no endereço MAC foi substituído pela escolha de um valor de AID contido em um bloco aleatório dentre N_b blocos, seguindo a estrutura de AID apresentada na Seção 2.3.6. O valor de N_b foi definido como o número de grupos RAW pré-configurados para as simulações. Dessa forma, os parâmetros de AID inicial e AID final de cada grupo foram configurados, respectivamente, como o primeiro e o último AID de um bloco de AID correspondente. Ou seja, para o RAW 1, o AID inicial é definido como 1 (pois o AID 0 é reservado) e o AID final é definido como 63. Isto implica que cada grupo RAW é limitado a 64 estações, com exceção do primeiro grupo que suporta no máximo 63. Essas alterações foram feitas na função *SendAssocResp()* da classe *ApWifiMac*. O mapeamento de grupos RAW em blocos de AID foi implementado como código fixo desta função. Sendo assim, para simular situações em que o número de estações por RAW seja maior que 64, é necessário fazer modificações no código da implementação proposta.

4.3 Implementação do AID Dinâmico

A implementação do padrão IEEE 802.11ah no ns-3 feita por Le Tian et. al [Tian et al. 2016, Tian et al. 2018] não inclui a operação do mecanismo de atribuição dinâmica de AID. Sendo assim, neste trabalho foi feita a implementação da operação deste mecanismo de acordo com as especificações previstas na norma.

Foi visto na Seção 2.3.7 que a operação de troca de AID pode ser iniciada pela estação (por meio do envio da requisição de AID) ou pelo próprio ponto de acesso (por meio do envio da mensagem de resposta de AID). Como o objetivo deste trabalho é o agrupamento de estações de modo dinâmico por parte do ponto de acesso, apenas a operação de AID Dinâmico *não solicitado* foi implementada. Neste cenário, apenas mensagens de *AID Switch Response* são trocadas.

A norma do IEEE 802.11ah classifica os quadros de *AID Switch Request* e *AID Switch Response* como quadros do tipo *Action*, e sub-tipo *Unprotected SIG Action*. Esta sub-tipo de quadros foi criado para suportar trocas de mensagens das funcionalidades SIG, como AID dinâmico, configuração de TWT e Setorização de Grupos. Os quadros de *AID Switch Request* e *AID Switch Response* foram implementados neste trabalho de acordo com a categorização prevista na norma, e estão representadas no código do ns-3 pelas classes *AIDSwitchRequestHeader* e *AIDSwitchResponseHeader*, respectivamente. É importante notar que apesar do quadro de requisição de troca de AID ter sido implementado no simulador, a operação de troca de AID iniciada pela estação não está implementada. Ou seja, as mensagens de *AID Switch Request* não são enviadas pelas estações e não existem rotinas que interpretam essas mensagens no ponto de acesso.

A Figura 4.2 mostra os campos do elemento *AID Switch Response* que foram incluídos como atributos da classe de cabeçalho. O campo ID do elemento identifica o tipo de

elemento da mensagem. Para os quadros de *AID Response*, este campo tem valor igual a 211. O campo Tamanho é o comprimento (em bytes) do elemento, cujo valor é igual a 5. O campo AID/AID de Grupo carrega o valor do AID a ser atribuído à estação de destino. O campo Contagem de troca de AID define a contagem regressiva, em número de intervalos de *beacon*, para a estação, após receber a mensagem de troca, modificar o seu AID para o novo valor atribuído. O campo Intervalo de resposta AID define o intervalo de escuta, em número de intervalos de *beacon*, no qual a estação TIM deve acordar para receber quadros de *beacon SIG* [IEEE 2017].

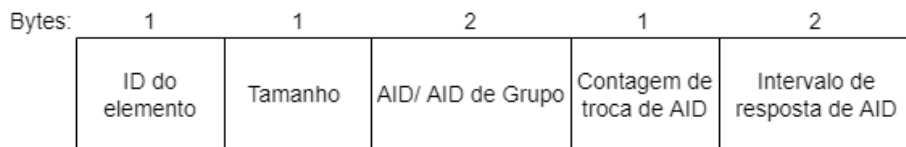


Figura 4.2: Elemento *AID Switch Response*.

O envio das mensagens de *AID Switch Response* é feito pela rotina *SendAIDSwitchResponse()*, implementada como método da classe *ApWifiMac*, conforme mostrado na listagem 4.1. A função recebe o endereço MAC da estação de destino e o valor de AID a ser atribuído a essa estação. Na linha 5, um objeto do cabeçalho MAC é criado e o tipo do quadro é definido como tipo *Action* na linha 7. Nas linhas 8-10 são definidos os endereços de origem e de destino. Nas linhas 12-17 é criado o cabeçalho do quadro de *AID Switch Response* e os valores dos campos são atribuídos. O campo *AID Switch Count* é definido como 0, para que a estação altere imediatamente o seu AID para o valor recebido, e o intervalo de resposta é definido como 1, porque a segmentação TIM está desabilitada em nossos estudos. Logo, todas as estações acordam a cada período de *beacon* para receber os quadros de *SIG Beacon*. Nas linhas 20-23, o cabeçalho do tipo *Action* é formatado, e nas linhas 25-27 é feito o encapsulamento de um novo pacote com os cabeçalhos criados. Na linha 29, o quadro é adicionado à fila para ser transmitido pelo canal por meio mecanismo de acesso ao canal distribuído (DCA). Na linha 30 é feita a atualização do AID na classe de gerenciamento RAW das estações associadas. É nesta classe que estão implementadas as operações relacionadas ao agrupamento RAW.

Listagem 4.1: Função de envio de Resposta de troca de AID.

```

1 void
2 ApWifiMac::SendAIDSwitchResponse (Mac48Address to, uint16_t aid){
3     NS_LOG_FUNCTION ( this );
4
5     WifiMacHeader hdr;
6     if ( m_slgSupported ){
7         hdr.SetAction ();
8         hdr.SetAddr1 ( to );
9         hdr.SetAddr2 ( GetAddress ( ) );
10        hdr.SetAddr3 ( GetAddress ( ) );

```

```

11
12     AIDSwitchResponseHeader respHdr;
13
14     respHdr.SetElementID(211);
15     respHdr.SetLength(5);
16     respHdr.SetAID(aid);
17     respHdr.SetAIDSwitchCount(0);
18     respHdr.SetAIDResponseInterval(1);
19
20     WifiActionHeader actionHdr;
21     WifiActionHeader::ActionValue action;
22     action.unprotectedSIG = WifiActionHeader::AID_SWITCH_RESPONSE;
23     actionHdr.SetAction(WifiActionHeader::SIG, action);
24
25     Ptr<Packet> packet = Create<Packet> ();
26     packet->AddHeader (respHdr);
27     packet->AddHeader (actionHdr);
28
29     m_dca->Queue(packet, hdr);
30     m_SlgRawCtr.SetSensorAid(to, aid);
31 }
32 }

```

A norma não especifica em quais momentos os quadros de requisição e resposta de AID devem ser enviados, logo esta decisão é tomada conforme as necessidades de operação do ponto de acesso. Sendo assim, optou-se pelo envio de mensagens de resposta de AID a cada 10 intervalos de *beacon*, imediatamente após a execução do algoritmo de agrupamento RAW, conforme mostrado na Figura 4.3. Como pode ser visto, a variável *count* é inicializada no começo da simulação com valor igual a 0, e é incrementada a cada *beacon* enviado. As mensagens de troca de AID são enviadas apenas às estações que foram movidas de grupo RAW e, portanto, precisam ter seus números de AID alterados para valores dentro do intervalo de AID do novo grupo. A chamada à função ocorre dentro da função *SendOneBeacon()*, pois todas as estações estão acordadas neste período para receberem quadros de *SIG Beacon*, e, portanto, podem receber as mensagens de troca de AID.

Foi implementado o tratamento de quadros de *AID Switch Response* pelas estações na classe *StaWifiMac*. Um trecho da função *Receive()* é mostrado na listagem 4.2. O pacote recebido é avaliado pelo tipo (linhas 7 e 11), e para os quadros do tipo *Action*, é verificado se a estação está associada ao BSS correspondente (linha 13). Nas linhas 15-17 é extraído o cabeçalho do tipo *Action*. Se o quadro for do tipo *AID Switch Response*, os campos de interesse são extraídos (linhas 24-28) e a mudança de AID é agendada de acordo com o valor do campo *AID Switch Count* (linha 30). O campo de intervalo de resposta não é utilizado porque a segmentação TIM não é utilizada, e, portanto, a estação já está agendada para acordar no próximo intervalo de *beacon*.

Listagem 4.2: Função de recepção da Resposta de troca de AID pela estação.

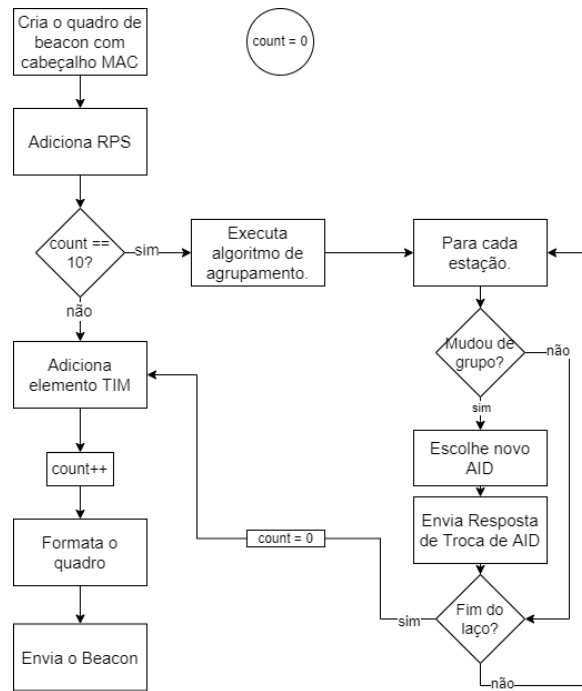


Figura 4.3: Fluxo da operação executada pela função *SendOneBeacon()*.

```

1
2 void
3 StaWifiMac::Receive (Ptr<Packet> packet, const WifiMacHeader *hdr, double
4   signalDbm, double noiseDbm, uint32_t rate)
5 {
6   ...
7   if (hdr->IsData ())
8   {
9     ...
10  }
11  else if (hdr->IsAction () && m_dynamicAidSupported)
12  {
13    if (IsAssociated () && hdr->GetAddr2 () == GetBssid ())
14    {
15      Ptr <Packet> packetCopy = packet->Copy ();
16      WifiActionHeader actionHdr;
17      packetCopy->RemoveHeader (actionHdr);
18
19      if ( actionHdr.GetAction () . unprotectedS1G == WifiActionHeader::
20        AID_SWITCH_RESPONSE)
21      {
22        AIDSwitchResponseHeader respHdr;
23        packetCopy->RemoveHeader (respHdr);
24
25        uint8_t elementID = respHdr.GetElementID ();
26        uint16_t aid = respHdr.GetAID ();
27        uint8_t aidSwitchCount = respHdr.GetAIDSwitchCount ();

```

```

27     uint8_t length = respHdr.GetLength();
28     uint8_t responseInterval = respHdr.GetAIDResponseInterval();
29
30     Simulator::Schedule(aidSwitchCount * beaconInterval, &StaWifiMac
31         ::SetAID, this, aid);
32     return;
33 }
34 }
35
36 [...]
37 }

```

4.4 Coleta de Parâmetros das Estações

Pensando em possíveis abordagens de agrupamento baseadas em métricas da rede, foi feita a modificação das funções de recebimento de pacotes pelas estações para que houvesse o repasse de métricas de interesse para a implementação da camada MAC superior do ponto de acesso. No ns-3, as medidas de potência do sinal recebido (em dBm) e taxa nominal do cabeçalho PLCP (em bits/s) capturadas na camada física são repassadas à camada *MAC Low*. Essas variáveis não são encaminhadas à camada *MAC High*, que é onde está implementada a operação de RAW. Logo, foi necessário ajustar as funções das classes *MACRxMiddle* e *FrameExchangeManager* (destacadas na Figura 4.1) para que houvesse o repasse dessas variáveis para o nível *Mac High*. Com isto, para qualquer quadro recebido, são capturados os valores de potência do sinal recebido e a taxa nominal presente no cabeçalho PLCP. Além disso, outro parâmetro de interesse é o tamanho do pacote recebido, que é diretamente calculado no recebimento do quadro. Para um pacote recebido no AP no instante t , vindo da estação i , o valor do atributo $j = \{1, 2, 3\}$, que é salvo para o cálculo do agrupamento, é dado por

$$v_{j,t}^i = 0,2 v_{j,t-T_R}^i + 0,8 v_{j,c}^i \quad (4.1)$$

em que v_{j,T_R}^i é o último valor salvo do parâmetro j da estação i no instante T_R , e $v_{j,c}^i$ é a medida obtida pelo pacote que acabou de ser recebido. É dado um peso maior ao novo valor medido, para que mudanças de comportamento das estações sejam rapidamente percebidos e refletidos no agrupamento.

4.5 O Algoritmo K-Means

O algoritmo K-means¹ foi implementado como parte do módulo Wi-Fi do ns-3. O código é composto por três classes principais:

- *Point*: representa um ponto com d dimensões. Os principais atributos desta classe são: o endereço MAC da estação correspondente (m_mac), os valores das coordenadas do ponto (m_values) e o identificador do grupo ao qual o ponto pertence ($m_clusterId$).
- *Cluster*: representa um conjunto de pontos. Os atributos dessa classe são: o identificador do conjunto ($m_clusterId$), o vetor do centroide ($m_centroid$) e o vetor de pontos do grupo (m_points).
- *KMeans*: compreende as operações do algoritmo de agrupamento. Os atributos dessa classe são: o número de grupos (m_k), o número máximo de iterações (m_iter), o número de dimensões dos pontos ($m_dimensions$), o número total de pontos (m_total_points), o conjunto de k grupos ($m_clusters$) e o modo de inicialização ($m_initialization$).

O algoritmo é executado pela função apresentada na listagem 4.3 a cada 10 intervalos de *beacon* logo antes da transmissão do *SIG Beacon* a todas as estações. As coordenadas dos pontos recebidos correspondem aos valores médios dos três parâmetros coletados no recebimento de pacotes (intensidade de sinal, taxa nominal de transmissão e tamanho do pacote). Nas linhas 5-10, o valor de k é definido como $\min(m_k, m_total_points)$, pois o número de grupos deve ser no mínimo igual ao número de pontos. Na linha 12 o valor de $m_dimensions$ é atualizado. Na linha 14 é feita a normalização dos pontos de entrada, dada por

$$norm(x_j^i) = \frac{x_j^i - \min(\mathbf{x}_j)}{\max(\mathbf{x}_j) - \min(\mathbf{x}_j)}, \quad (4.2)$$

em que x_j^i é atributo $j = \{1, 2, 3\}$ do ponto i , e $\mathbf{x}_j = \{x_j^0, x_j^1, \dots, x_j^N\}$ é o vetor com os dos parâmetros j de todos os N pontos. Nas linhas 16-21, um dos métodos de inicialização é executado (aleatório, *MinMax* ou *Sorting*), criando k grupos com os centroides iniciais. As implementações detalhadas dos três métodos de inicialização do algoritmo K-Means são apresentadas no Apêndice A.

Nas linhas 24-54 são executadas as iterações do algoritmo. A etapa de atribuição é executada na linha 27, onde cada ponto tem seu $m_clusterId$ atualizado com o identificador do grupo com centroide mais próximo. Além disso, o conjunto de pontos (m_points) de cada grupo é atualizado. Na linha 28 é executada a etapa de atualização, onde os centroides são

¹A implementação em C++ foi baseada no código em <https://github.com/marcoscastro/kmeans>.

atualizados como a média de todos os pontos dos seus respectivos grupos. O laço é interrompido caso nenhum ponto tenha mudado de grupo ou se o número máximo de iterações for atingido (linhas 30-33).

Listagem 4.3: Rotina que calcula a clusterização do K-means.

```
1
2 void
3 KMeans::run(std::vector<Point> &all_points)
4 {
5     m_total_points = all_points.size();
6     uint16_t K;
7     if (m_k > m_total_points)
8         K = m_total_points;
9     else
10        K = m_k;
11
12    m_dimensions = all_points[0].getDimensions();
13
14    Normalize(all_points);
15
16    if (m_initialization == "sort")
17        SortingAttrInitialization(all_points, K);
18    else if (m_initialization == "minmax")
19        MinMaxInitialization(all_points, K);
20    else if (m_initialization == "random")
21        RandomInitialization(all_points, K);
22
23    int iter = 1;
24    while (true)
25    {
26        bool done = true;
27        done = Clusterize(all_points);
28        updateCentroids();
29
30        if (done || iter >= m_iters)
31        {
32            break;
33        }
34        iter++;
35    }
36 }
```

Resumo

A estratégia de agrupamento proposta neste trabalho foi implementada e testada por meio do simulador de redes ns-3. A implementação partiu da extensão do simulador proposta

por [Tian et al. 2016] que inclui o modelo de camada física do IEEE 802.11ah, bem como algumas funcionalidades da camada MAC (como segmentação TIM, operação do RAW e associação rápida). Dado que o agrupamento de estações no RAW depende dos valores de AID das estações, este trabalho propõe uma implementação inédita da operação de atribuição dinâmica de AID prevista em norma, de modo que uma estratégia de agrupamento dinâmico possa ser aplicada. Além disso, é proposta a aplicação do algoritmo de agrupamento K-Means como estratégia para distribuir as estações entre os grupos RAW de acordo com parâmetros medidos a partir do recebimento de pacotes de dados. Os parâmetros de interesse são potência do sinal recebido, taxa nominal de transmissão (definida pelo modo MCS) e tamanho do pacote.

Capítulo 5

Avaliação de Desempenho

Considerando um cenário em que as estações são atribuídas aleatoriamente aos grupos RAW, em um mesmo grupo pode haver estações com características de tráfego e de canal muito distintas. Nesta situação, estações transmitindo a altas taxas, com alta carga e em boas condições de canal, competem pelo meio com estações transmitindo a taxas mais baixas, com menor carga e em piores condições de canal. As estações transmitindo a menores taxas são mais lentas, e, portanto, tomam mais tempo de canal para transmitir uma mesma quantidade de dados que uma estação com maior taxa. Além disso, as más condições de canal geram perdas e retransmissões, o que resulta em desperdício do tempo de acesso ao canal. Por um lado, as estações mais rápidas podem dominar completamente o canal, prejudicando, assim, o desempenho das estações mais lentas. Por outro lado, os valores de vazão média dos diferentes grupos não devem reproduzir o resultado de acordo com as taxas nominais de transmissão das estações.

A aplicação do algoritmo K-Means ao problema de agrupamento é proposta para verificar o desempenho das estações IEEE 802.11ah quando são agrupadas de acordo com suas características de tráfego e suas condições de canal. Neste cenário, as estações competem pelo canal com outras estações em condições similares. Espera-se, com isto, que seja possível tirar proveito das estações configuradas para transmitir a maiores taxas e em melhores condições, e ao mesmo tempo dar um acesso ao meio justo às estações em piores condições. Para avaliar o desempenho da rede, e comparar o método com agrupamento por K-means à abordagem de agrupamento aleatório citada anteriormente, duas medidas de desempenho são consideradas: 1) a vazão média de cada grupo de estações e 2) a justiça de Jain normalizada pela taxa nominal. A vazão deve nos indicar a performance para os diferentes grupos considerando-se as diferentes configurações abordadas, enquanto a justiça normalizada deve nos indicar o quão próximos os valores de vazão de cada estação estão das taxas nominais configuradas. Ou seja, o quão próximo cada estação chegou do seu objetivo máximo. Para cada medida de desempenho são tomadas a média e o desvio padrão (DP) calculados a partir da execução de uma mesma simulação para diferentes valores de semente de gerador de números aleatórios. Com isto é possível avaliar o nível de estabilidade das diferentes abor-

dagens em estudo.

A avaliação das estratégias de agrupamento foi feita a partir de resultados de simulações executadas no ns-3. Uma configuração geral foi adotada para todas as simulações de estudos de caso. Foram executadas simulações com o agrupamento pelo algoritmo K-means com três diferentes métodos de inicialização para verificar qual dos métodos melhor serve para a estratégia de agrupamento proposta. Uma simulação com mobilidade foi executada para verificar-se o funcionamento do algoritmo de agrupamento por K-means em conjunto com a atribuição dinâmica de AID. Para análise foram coletadas medidas de vazão e justiça normalizada para três diferentes cenários considerando-se três estratégias de agrupamento: agrupamento aleatório estático (fixo), agrupamento aleatório dinâmico e agrupamento baseado em K-Means. Por fim, foram executadas simulações com diferentes números de estações na rede para verificar o impacto do número de estações na justiça observada para as três diferentes abordagens.

5.1 Configuração das Simulações

Foram executadas simulações para diferentes cenários, porém a configuração geral da topologia foi mantida fixa, com as estações distribuídas uniformemente entre quatro zonas geográficas concêntricas $R_i, i \in \{1, 2, 3, 4\}$, com o ponto de acesso no centro, como ilustrado na Figura 5.1. Na Zona 1 estão espalhadas as estações mais próximas do ponto de acesso dentro de um raio de 50 metros. As estações das demais zonas (2, 3 e 4) são distribuídas em regiões em formato de anel com largura de 50 metros. O objetivo desta configuração é avaliar o impacto da distância (medida pela intensidade do sinal recebido) no agrupamento RAW com o K-means. No entanto, esta setorização de estações não corresponde à Setorização de grupos descrita da Seção 2.3.5, pois o agrupamento de estações por zonas em formato de anel não mitiga o problema do nó oculto, e o controle de acesso ao meio neste caso é feito pelo agrupamento RAW.

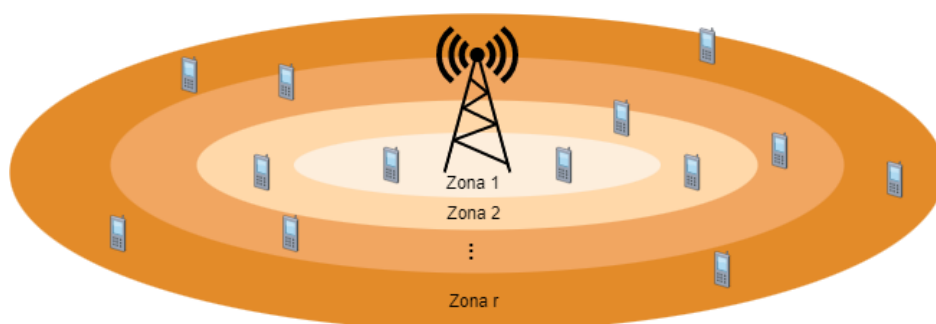


Figura 5.1: Topologia da rede dividida em zonas geográficas.

As estações foram posicionadas em pontos aleatórios dentro de suas respectivas zonas, e, por padrão, foi utilizado o modelo de mobilidade de posição constante, de modo que não houvesse movimentação das estações pela topologia. Nas simulações é assumida uma rede

outdoor com modelo de propagação para ambiente macro de redes IEEE 802.11ah proposto por Hamzi, A. em [Hazmi et al. 2012] dado por

$$L(d) = 8 + 37.6 \log_{10}(d) + 21 \log_{10} \left(\frac{f}{900 \text{ MHz}} \right), \quad (5.1)$$

em que d é a distância de transmissão-recepção e f é a frequência da onda portadora. O modelo de erro assumido é o *YansErrorModel* [Pei and Henderson 2009] implementado no ns-3 para diferentes modulações e assumido para simulações com IEEE 802.11ah. Neste modelo, os pacotes são divididos em fragmentos com eventos de início e fim de recepção. Para cada fragmento, é calculada a relação sinal-interferência-mais-ruído entre a potência do pacote recebido e a soma do ruído e da interferência de todos os pacotes recebidos. Com isto, é calculada a taxa de sucesso baseada na taxa de erro de bit (BER). A Tabela 5.1 lista os principais parâmetros da camada física configurados nas simulações realizadas.

Tabela 5.1: Parâmetros gerais da camada física para as simulações realizadas.

Parâmetro	Valor (Estação/AP)
Frequência	900 MHz
Largura de banda do canal	2 MHz
Potência de Transmissão	0 dBm / 0 dBm
Ganho de transmissão	0 dB / 3 dB
Figura de ruído	6,8 dB
Modelo de perda de propagação	Outdoor, macro
Modelo de taxa de erro	YansErrorModel
Modos MCS	MCS0 (650 kbps), MCS1 (1300 kbps), MCS2 (1950 kbps), MCS3 (2600 kbps)

Para a camada MAC, os parâmetros foram configurados conforme os valores da Tabela 5.2. O intervalo de *beacon* de 199,84 ms foi dividido em quatro grupos RAW, cada grupo consistindo de 2 RAW slots. A duração do RAW slot foi configurada de modo que o tempo total de RAW ocupe todo o intervalo de *beacon*. Além disso, o CSB foi desabilitado, de modo que as estações não fossem permitidas de acessar o canal fora dos seus intervalos de slot. Por fim, a troca de quadros de RTS/CTS para controle de acesso foi desabilitada.

Tabela 5.2: Parâmetros da camada MAC usados nas simulações.

Parâmetro	Valor
Intervalo de beacon (ms)	199,84
Número de grupos RAW	4
Número de RAW slots por grupo	2
Tempo de RAW slot (ms)	24,98
Tempo total de RAW (ms)	199,84
Cross slot boundary	não habilitado
RTS/CTS	não habilitado

Todas as estações foram configuradas com aplicações UDP transmitindo pacotes a cada

2 ms. O tamanho dos pacotes foi deixado como uma variável de entrada, pois trata-se de um parâmetro de interesse para o agrupamento com o algoritmo K-means.

As principais medidas de desempenho de interesse para as estratégias abordadas são a vazão média por grupo RAW e a justiça de Jain normalizada. A vazão média para uma determinada estação é calculada como número total de bits entregues dentro do tempo total de canal disponível para esta estação. Ou seja, é a soma de todos os bits de informação entregues com sucesso ao AP dividida pelo tempo total ocupado pelo slot RAW correspondente por toda a simulação. No caso da vazão média por grupo RAW, o cálculo é dado pelo número total de bits entregues dentro do tempo de RAW daquele grupo dividido pelo tempo total do RAW. Dado que s_i é a vazão média obtida pela estação i e r_i é a taxa nominal configurada nesta estação, define-se a razão $u_i = s_i/r_i$ tal que a justiça normalizada seja dada por

$$J = \frac{(\sum_{i=1}^m u_i)^2}{m \sum_{i=1}^m u_i^2}, \quad (5.2)$$

em que m é o número total de estações.

5.2 Impacto da Inicialização do K-Means

Na Seção 2.4 foi visto que o resultado do agrupamento do algoritmo K-means é fortemente dependente do método de inicialização dos centroides. A escolha de pontos aleatórios aumenta as chances do algoritmo ficar preso em um mínimo local com resultado de agrupamento pouco preciso. Tendo isso em vista, os três seguintes métodos de inicialização foram implementados no simulador: aleatório, heurística da maior distância (*MinMax*) e heurística do ordenamento.

Para verificar o desempenho dos três métodos, foram executadas simulações da topologia em quatro zonas geográficas com características de tráfego distintas. A cada zona foram adicionadas 20 estações, totalizando 80 estações na rede. O objetivo dessas simulações é visualizar o resultado do agrupamento para $k = 4$, considerando que idealmente os nós com as mesmas características de tráfego e em condições de canal similares devem ser agrupados nos mesmos conjuntos. As estações foram configuradas com os seguintes modos MCS: MCS3 para as estações de R_1 , MCS2 para as estações de R_2 , MCS1 para as estações de R_3 e MCS0 para as estações de R_4 . O tamanho de pacotes foi definido como 512 bytes para todas as estações.

A Figura 5.2 mostra o resultado do agrupamento do K-means com a inicialização aleatória. Como apenas quatro tipos de MCS foram configurados, a taxa nominal assume apenas quatro valores possíveis. Nota-se que a configuração dos grupos não segue a segregação das estações de acordo com suas zonas. O agrupamento incorreto acontece com maior frequência para as estações das zonas mais próximas, pois há maior variação de potência da recepção, tendo em vista que o modelo de propagação assumido é logarítmico.

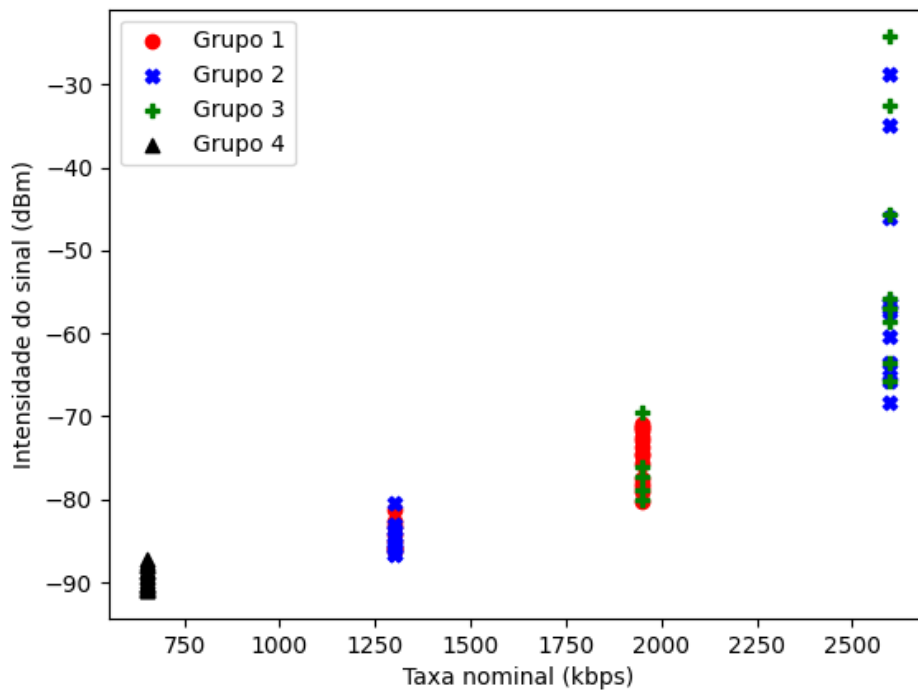


Figura 5.2: Resultado do agrupamento do algoritmo K-Means com inicialização aleatória. Cada ponto do gráfico representa uma estação cujo os valores de intensidade de sinal e taxa nominal foram mensurados.

A Figura 5.3 mostra o resultado do agrupamento dos mesmos pontos ao utilizar-se a inicialização da heurística da maior distância. Todas as estações da zona 4 foram colocadas no mesmo grupo (Grupo 4). No entanto, os nós das demais zonas ficaram espalhados por diferentes grupos, caracterizando um agrupamento ineficiente.

A Figura 5.4 mostra o resultado do agrupamento por K-Means com a inicialização da heurística de ordenamento. Nota-se que as divisões dos grupos está bem definida de acordo com as diferenças de taxa e de potência. Comparado aos demais métodos, o algoritmo de inicialização *Sorting* aparenta ser a melhor opção a ser usada no cenário apresentado. Sendo assim, o método foi adotado como padrão para todas as simulações apresentadas nas seções seguintes.

5.3 Validação do Mecanismo de Atribuição Dinâmica de AID

Para certificar-se da implementação do algoritmo de agrupamento por K-means e a operação de atribuição dinâmica de AID, considerou-se para simulação um cenário em que há movimentação de estação em algum momento da operação da rede. Neste cenário, todas as estações foram configuradas com o modo MCS 3, com as aplicações transmitindo pacotes

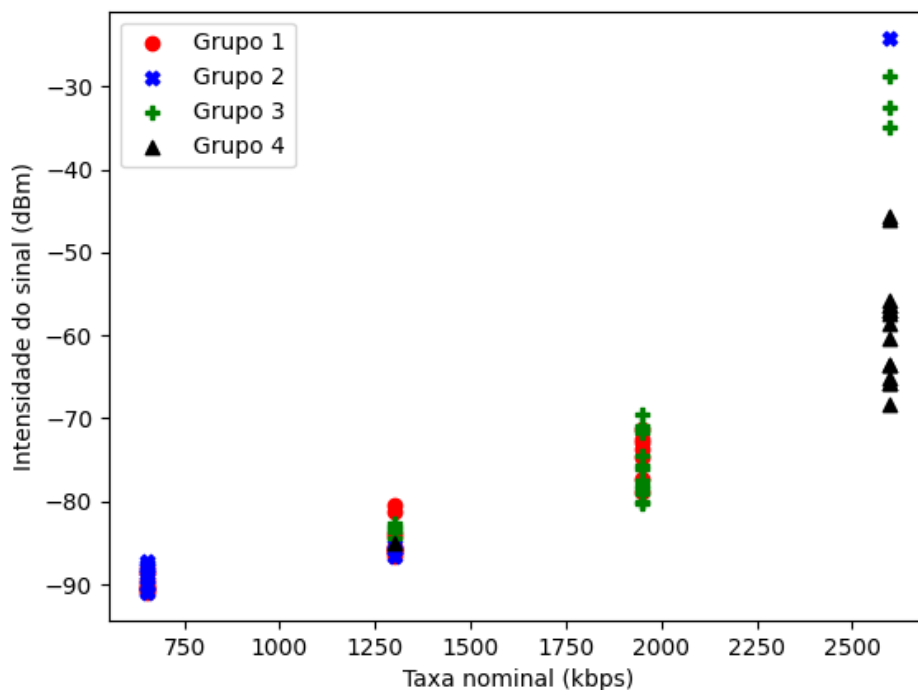


Figura 5.3: Resultado do agrupamento do algoritmo K-Means com a inicialização da heurística da maior distância (*MinMax*).

de 512 bytes. Com isto, é possível avaliar o impacto da mudança do parâmetro de potência de sinal medida e verificar como a implementação proposta reage a esta mudança.

Inicialmente, a estação foi posicionada a 10 metros de distância do ponto de acesso, estando, portanto, dentro da Zona 1. Ao associar-se, a estação recebe um AID aleatório, o que faz com que seu grupo RAW inicial também seja aleatório. Após 10 intervalos de beacon, o algoritmo de agrupamento por K-means é executado, e a estação é atribuída a um grupo de acordo com os parâmetros coletados no recebimento de pacotes de dados. Caso não haja mudança no tráfego (no caso, no comprimento do pacote de dados), na taxa de bit (ou seja, modo MCS), ou na posição da estação, a tendência é que a ela permaneça no mesmo grupo RAW, e, conseqüentemente, permaneça com o mesmo valor de AID. Porém, a estação foi configurada para mover-se instantaneamente para uma nova posição a 190 metros de distância do ponto de acesso aos 50 segundos de simulação. Nesse instante, a estação passa a fazer parte da Zona 4, e a intensidade do sinal recebido medida no AP deve diminuir.

A Figura 5.5 mostra os valores de AID atribuídos à estação com mobilidade no decorrer da simulação com duração de 120 segundos. Inicialmente, o AID 192 é atribuído aleatoriamente à estação, que está incluso no quarto grupo RAW, grupo este que contém as estações com as características da zona 1. Após a execução do algoritmo de agrupamento, a estação continua com o mesmo AID, o que indica que ela permaneceu no mesmo grupo RAW. O que é esperado, pois a estação, de fato, está na zona 1 e, portanto, foi inserida no grupo cujas estações apresentam maior intensidade de sinal, conforme já visto na Figura 5.4. Aos 56

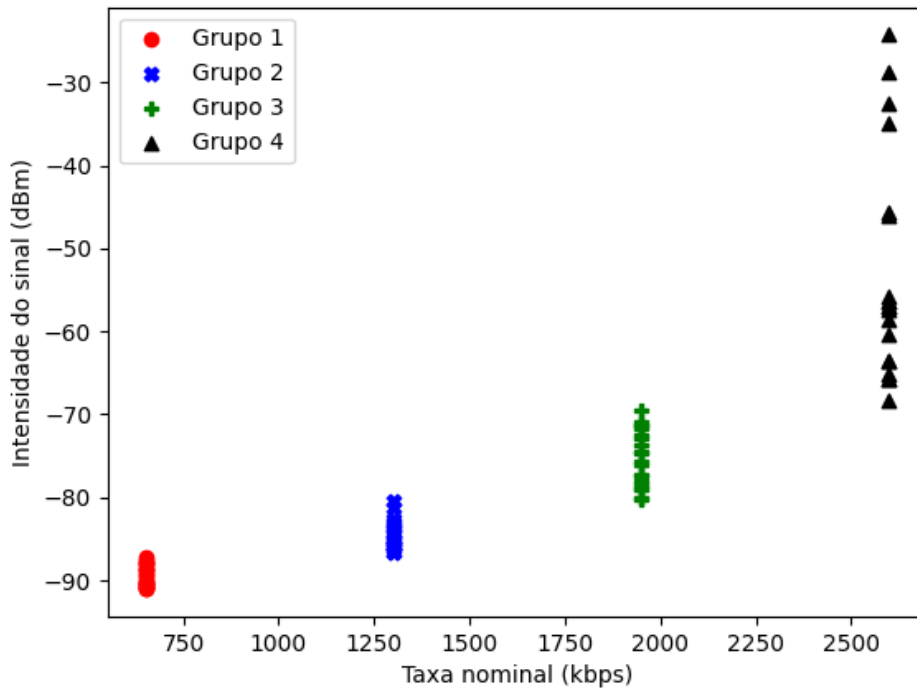


Figura 5.4: Resultado do agrupamento do algoritmo K-Means com a inicialização da heurística de ordenamento (*Sorting*).

segundos de operação, é possível ver que houve mudança do AID da estação, que passa a ter valor 14. Este novo valor de AID está contido no primeiro grupo RAW, que contém as estações com as características da zona 4. Nota-se que houve um atraso de 6 segundos entre a mudança de posição da estação e a alteração do seu valor de AID. Isto se dá pelas seguintes razões: 1) o algoritmo de agrupamento só é executado a cada 10 intervalos de *beacon*. Logo, a mudança da potência de sinal não causa uma alteração imediata dos grupos já criados; 2) o ponto de acesso só consegue perceber esta mudança quando recebe pacotes de dados vindos da estação, porém como ela fica mais distante do AP, é mais provável que ocorram colisões e perdas de pacotes por erro de canal.

Os resultados demonstram que a mudança no valor de potência das estações resulta em alterações no comportamento do algoritmo de agrupamento, o que também envolve a troca dinâmica de AID. O mesmo vale para os atributos relacionados à taxa de dados (MCS) e tamanho de pacote.

5.4 Estudo de Caso 1

Para as próximas simulações, três abordagens de agrupamento foram consideradas: agrupamento aleatório dinâmico, agrupamento aleatório estático e agrupamento por K-means. No agrupamento aleatório dinâmico, a cada 10 intervalos de *beacon* é feita uma atribuição

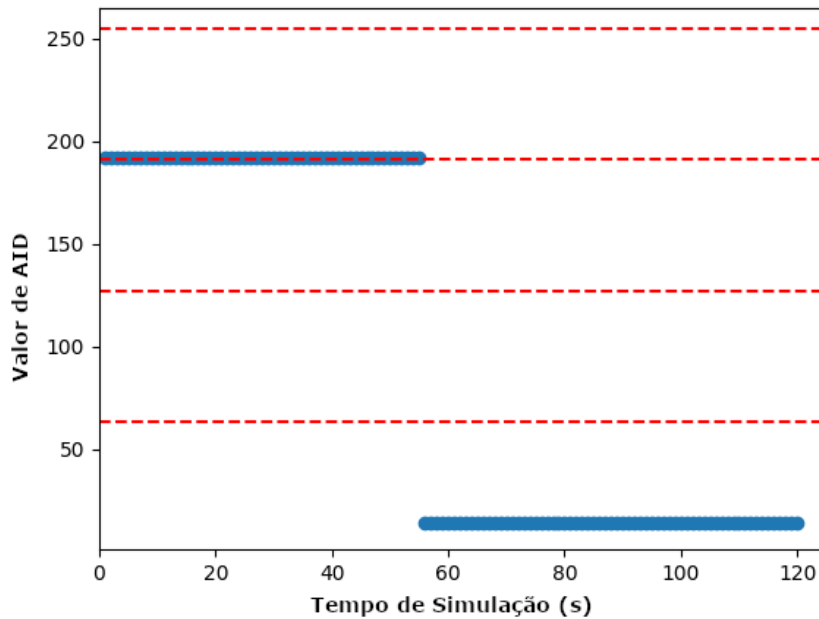


Figura 5.5: Exemplo de atribuição dinâmica de AID para uma estação como resultado de alteração de sua localização na rede. Algoritmo K-Means altera o agrupamento da estação (novo AID) após detectar alteração no valor da potência do sinal recebido desta estação.

aleatória de cada estação a um dos quatro grupos RAW. Caso a estação permaneça no mesmo grupo que já se encontrava, ela continua com o mesmo valor de AID. Caso contrário, um novo AID é escolhido dentro do intervalo de AIDs do novo grupo, e então é transmitido um quadro de resposta de troca AID para a estação com o novo valor de AID. No agrupamento aleatório estático, as estações permanecem com os valores de AID atribuídos no processo de associação por todo o tempo de simulação e, portanto, não sofrem alteração de grupos RAW. A agrupamento por K-means ocorre conforme descrito na Seção 4.5. É importante citar que no agrupamento aleatório dinâmico, assim como no agrupamento por K-Means, há a transmissão em modo *unicast* de quadros de *AID Switch Response* para as estações que trocam de grupo após a execução dos algoritmos. Porém, com o agrupamento baseado em K-Means, o número de transmissões dessas mensagens de gerenciamento é menor, pois os parâmetros medidos tendem a se estabilizar durante a operação, resultando, assim, em menos alterações no agrupamento e, conseqüentemente, menos envio de quadros de *AID Switch Response*. Nos três cenários experimentados, 80 estações foram distribuídas entre as quatro zonas, com 20 estações cada.

No primeiro cenário, as taxas de dados (modos MCS) das estações são definidas pelas zonas às quais elas pertencem. Sendo assim, as estações da zona mais próxima (Zona 1) são configurados com MCS3, as estações da Zona 2 são configuradas com MCS2, as estações da Zona 3 usam MCS1 e as estações da Zona 4 transmitem com MCS0. O tamanho de pacote é mantido fixo igual a 512 bytes para todas as estações.

Os resultados de vazão média para as três abordagens são mostrados na Tabela 5.3. Os

valores de vazão correspondem à média dos resultados de vazão para 10 execuções da simulação com diferentes sementes de números aleatórios. E, para cada valor médio, foi calculado o desvio padrão (DP) das amostras. É possível observar que, nos casos aleatórios, os valores de vazão para os diferentes grupos são bem próximos, o que é esperado, pois não há qualquer critério de agrupamento sendo aplicado. No entanto, é possível notar que o caso dinâmico apresenta resultados de vazão ligeiramente maiores. Isto ocorre pois a redistribuição aleatória a cada 10 intervalos de *beacon* pode resultar em grupos mais homogêneos no decorrer da simulação. Esta uniformidade na vazão atingida pelos diferentes grupos indica que as estações transmitindo a taxas mais baixas competem pelo meio com estações transmitindo a taxas mais rápidas e ocupam maior tempo de canal para transmitir a mesma quantidade de bits, havendo, assim, prejuízo às estações com taxas maiores. Além disso, deve-se considerar que em um mesmo grupo RAW pode haver estações com grande diferença de distância do ponto do acesso. Neste caso, as estações mais próximas do AP têm mais chance de conseguir o acesso ao canal por sofrerem um recuo aleatório (*backoff*) menor com maior probabilidade do que o sofrido por estações mais distantes, em decorrência do maior número de retransmissões que é mais comum em estações sob piores condições de canal. Assim sendo, o acesso ao canal tende a ser menos justo para as atribuições de AID aleatórias.

Tabela 5.3: Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 1.

Grupo	Aleatório (dinâmico)		Aleatório (estático)		K-Means	
	Média	DP	Média	DP	Média	DP
1	565,62	43,36	530,56	77,19	325,28	17,06
2	537,83	33,79	524,34	40,98	473,99	18,06
3	580,89	25,32	521,70	50,94	654,99	7,73
4	635,34	32,29	662,81	47,22	921,94	19,49

Para o caso com K-means, nota-se que o grupo que abriga as estações da Zona 1 apresenta maior vazão média devido à mitigação dos problemas encontrados nos casos aleatórios anteriormente citados. Neste caso, o agrupamento por similaridade faz com que estações transmitindo à mesma taxa, e em distâncias próximas, compitam pelo canal entre si. Isso faz com que a vazão obtida se aproxime mais da taxa nominal das estações. Este resultado é reforçado pelos valores de justiça normalizada obtidos por cada abordagem. Para o caso aleatório dinâmico, a justiça normalizada obtida foi igual a 0,818 (DP = 0,08), enquanto que para o caso dinâmico estático foi igual a 0,736 (DP = 0,07) e para o caso com K-means, obteve-se 0,905 (DP = 0,04).

5.5 Estudo de Caso 2

No segundo cenário, além das taxas de dados, os tamanhos dos pacotes também são definidos pelas zonas as quais elas pertencem. Sendo assim, as estações da zona mais próxima

(Zona 1) são configurados com MCS3 e tamanho da pacote de 1024 bytes. Já as estações da Zona 2 são configuradas com MCS2 e tamanho de pacote de 768 bytes. As estações da Zona 3 usam MCS1 enviando pacotes de 512 bytes e as estações da Zona 4 transmitem pacotes de 256 bytes em MCS0. Este cenário é baseado na decisão de transmitir pacotes maiores quando as condições do canal são melhores, havendo menor probabilidade de perdas. Enquanto que em condições de canal piores, pacotes menores são transmitidos devido à maior probabilidade de perda, o que implica em menor volume de dados em bits retransmitidos. Esta abordagem pode ser aplicada, por exemplo, em redes IEEE 802.11 com fragmentação de quadros, tirando assim melhor proveito deste mecanismo. No entanto, nas simulações realizadas, a fragmentação de quadros não estava habilitada, sendo o tamanho do pacote definido na aplicação e transmitido em um único quadro. Os demais parâmetros de simulação foram mantidos exatamente iguais ao do estudo de caso 1.

Os resultados de vazão para este caso são mostrados na Tabela 5.4. Assim como no estudo de caso 1, os valores foram obtidos para 10 execuções da mesma simulação. Nota-se que, no geral, a vazão média dos grupos são maiores que as obtidas no primeiro cenário. Isto ocorre porque o tamanho médio de pacotes é maior que no caso anterior. Logo, espera-se que o volume de bits entregues seja maior. O comportamento geral dos resultados obtidos para os casos aleatórios é similar aos obtidos no primeiro cenário, com uma distribuição de vazão aproximadamente uniforme entre os grupos. Já os resultados de vazão do K-Means estão mais próximos do comportamento esperado conforme as taxas nominais. Nota-se também que os valores de desvio padrão para o caso aleatório dinâmico são muito menores que os valores obtidos para o caso aleatório estático, demonstrando maior consistência nos resultados de vazão obtidos. Os valores de justiça normalizada obtidos para os casos aleatório dinâmico, aleatório estático e com K-means são 0,843 (DP = 0,02), 0,640 (DP = 0,05) e 0,912 (DP = 0,03), respectivamente. Esses resultados contribuem com o argumento de que o K-means resulta em um agrupamento mais justo e consistente.

Tabela 5.4: Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 2.

Grupo	Aleatório (dinâmico)		Aleatório (estático)		K-Means	
	Média	DP	Média	DP	Média	DP
1	802,57	49,24	736,81	99,43	360,60	22,49
2	747,65	37,11	715,28	86,28	497,55	18,67
3	796,29	27,35	720,19	61,87	740,13	10,75
4	901,16	43,80	886,43	65,13	1232,65	28,79

5.6 Estudo de Caso 3

O terceiro cenário objeto de análise tem o objetivo de avaliar o agrupamento de estações considerando as condições do canal, e definindo que todas as estações transmitam

a uma mesma taxa nominal. Em trabalhos anteriores, a taxa de dados nominal informada no cabeçalho PLCP é o único parâmetro considerado para o agrupamento de estações [Sangeetha and Babu 2020, Badarla and Harigovindan 2021]. Considerar outros parâmetros no agrupamento de estações pode gerar resultados mais consistentes com as características da rede, além de tornar o processo de atribuição mais dinâmico.

Os tamanhos de pacote por zona foram definidos da mesma forma que a configuração do cenário 2. No entanto, a taxa nominal foi deixada constante, com todas as estações configuradas com MCS3. Os resultados de vazão para este cenário executado 10 vezes são mostrados na Tabela 5.5. Nota-se que houve um aumento geral da vazão obtida para todos os casos. Isto ocorre pois todas as estações foram configuradas com MCS3, que transmite a taxas mais altas. Também é possível ver que os resultados de vazão do K-means apresentam maior consistência e homogeneidade, pois além dos valores estarem mais consistentes com as taxas nominais, o desvio padrão é muito menor que para os casos aleatórios, assim como os resultados obtidos nos cenários anteriores. A justiça normalizada para o esquema aleatório dinâmico é igual a 0,744 (DP = 0,02), para o caso aleatório estático é 0,575 (DP = 0,08) e para o K-means igual a 0,844 (DP = 0,02). Nota-se que o K-means continua sendo mais justo que os demais esquemas. No entanto, houve diminuição geral da justiça normalizada para todos os casos. Isto ocorre porque, como todas as estações transmitem a uma mesma taxa, as condições de canal têm maior impacto na entrega da quadros. Sendo assim, a probabilidade de erro de pacote para as estações mais distantes é maior, o que resulta em um número maior de retransmissões.

Tabela 5.5: Vazão média (em kb/s) e desvio padrão (DP) para os esquemas de grupos para o cenário 3.

Grupo	Aleatório (dinâmico)		Aleatório (estático)		K-Means	
	Média	DP	Média	DP	Média	DP
1	982,95	46,60	88,28	109,10	651,35	17,27
2	950,40	23,23	914,21	38,35	816,25	17,78
3	997,55	25,16	915,13	49,44	931,19	15,79
4	1079,42	41,19	1037,83	92,58	1270,42	19,49

5.7 Impacto do Número de estações

Considerando que redes IEEE 802.11ah podem ser altamente densas (suportando até 8192 estações associadas a um único ponto de acesso), foram executadas simulações com diferentes números de estações. Foi assumida a mesma configuração da simulação apresentada na Seção 5.5, variando apenas o número de estações por zona. Foram executadas simulações com 120 e 160 estações uniformemente distribuídas entre as quatro zonas geográficas. Para cada número de estações, a mesma simulação foi executada para 10 valores de sementes para os geradores de número aleatório. A Figura 5.6 mostra os resultados de justiça

normalizada obtida como a média dos valores das obtidos nas 10 execuções. Nota-se que o aumento do número de estações na rede tende a torná-la menos justa para as três abordagens de agrupamento. Como os parâmetros de tempo de RAW, número de grupos e número de RAW slots foram mantidos os mesmos, o aumento do número de estações representa maior competição pelo canal e maior probabilidade de colisões, e este efeito é ainda mais presente em estações com piores condições de canal. Nota-se, ainda, que o impacto negativo do aumento de estações na justiça normalizada é maior para a abordagem aleatória estática. Como nesta abordagem a atribuição permanece fixa durante toda a operação, os grupos definidos no decorrer das simulações são menos homogêneos, e portanto, menos justos em termos de acesso ao canal.

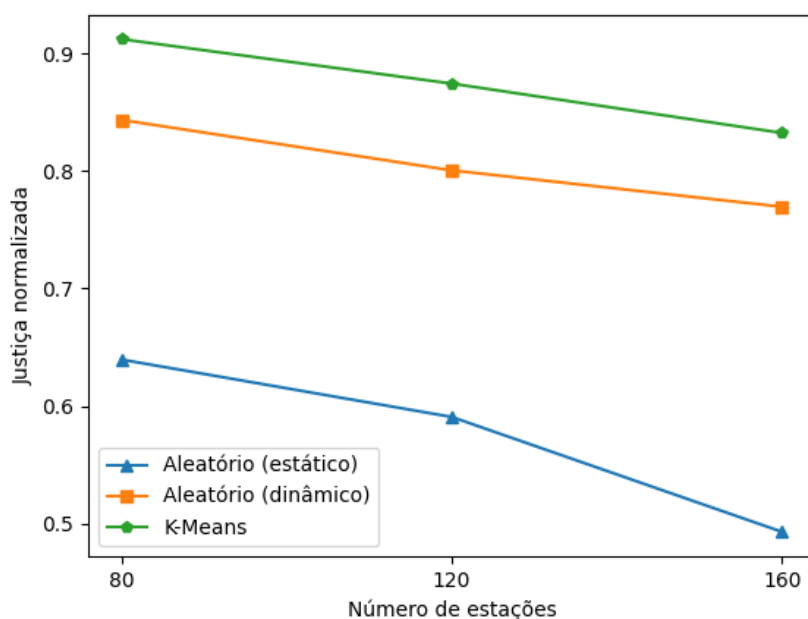


Figura 5.6: Gráfico de Justiça normalizada para diferentes números de de estações.

Resumo

Neste capítulo foram apresentados os resultados obtidos a partir da execução de simulações de diferentes cenários com as implementações propostas. Inicialmente, foi feita a análise comparativa de três métodos de inicialização do algoritmo K-Means aplicados ao cenário de uma rede IEEE 802.11ah heterogênea com estações distribuídas em zonas geográficas bem definidas. O algoritmo da heurística de ordenamento mostrou-se muito mais preciso que o método de inicialização aleatório e a inicialização pela heurística do ponto mais distante. Além disso, um cenário de rede com mobilidade foi simulado para validar o reagrupamento de uma estação que mudou de posição e o funcionamento da troca dinâmica de AID como consequência disso. Por fim, três cenários foram objetos de estudo para verificar o desempenho da estratégia de agrupamento pelo algoritmo K-means quando comparada

a métodos de agrupamento aleatório. As medidas de desempenho consideradas na análise foram vazão média por grupo e justiça normalizada média. Os resultados demonstram que o agrupamento por K-Means agrupa as estações em uma configuração mais justa em termos de vazão, além de apresentar resultados mais homogêneos. Este comportamento se reproduz para situações com maior número de estações associadas.

Conclusão

Neste trabalho foi proposta uma abordagem de agrupamento de estações IEEE 802.11ah baseado no algoritmo K-Means, combinado com a implementação da funcionalidade de atribuição dinâmica de AID. Na abordagem proposta, as estações devem ser agrupadas com base em suas características de tráfego e condições de canal, que são representadas pelas métricas de tamanho médio de pacotes, taxa nominal de transmissão de cada e intensidade média de sinal recebido. A funcionalidade de atribuição dinâmica de AID foi implementada para possibilitar o remanejamento das estações entre os diferentes grupos, tornando assim o agrupamento RAW mais dinâmico. Dessa forma, o método proposto é capaz de perceber mudanças nas características de tráfego ou nas condições de canal de uma estação e, a partir disso, movê-la para um novo grupo que contém estações com características similares.

A abordagem proposta mostrou-se promissora por alcançar bons resultados de justiça de vazão entre grupos RAW, considerando cenários com multi-taxas e características de tráfego e de canal heterogêneas. As simulações realizadas indicam que o uso de diferentes critérios para o agrupamento de estações resulta em grupos mais homogêneos e mais justos em termos da vazão média. Além disso, foi visto que apesar de não ser o algoritmo de agrupamento de melhor performance, o algoritmo K-Means consegue apresentar bons resultados a partir da implementação de métodos de inicialização mais sofisticados, que tendem a evitar que o algoritmo de agrupamento fique preso em mínimos locais e resulte em um agrupamento pouco preciso.

Trabalhos Futuros

Existem muitas outras variáveis que podem ser consideradas pela estratégia de agrupamento RAW. Como exemplo, os parâmetros de RAW (como número de grupos, número de RAW slots, duração de slots) que foram mantidos com valores fixos em nossas simulações, podem ser incluídos como saída de um algoritmo de agrupamento que tem como entrada as mesmas métricas utilizadas pelo algoritmo K-Means. Com isto seria adicionada uma segunda camada de inteligência ao agrupamento, buscando atender as demandas de grupos de estações com características semelhantes de modo mais específico. Além disso, pode-se fazer a modificação da atribuição dinâmica de AID condicionada à operação de segmentação TIM, e com isto ter uma solução mais energeticamente eficiente. Por fim, foi visto

que a nossa implementação está limitada a operar com no máximo 64 estações por grupo RAW, pois as sequências de AID de cada grupo é baseada nos blocos da hierarquia de AID. Pode-se, portanto, expandir a capacidade máximo dos grupos adotando-se outra estratégia de definição dos limites de AID.

Referências Bibliográficas

- [Al-Sarawi et al. 2017] Al-Sarawi, S., Anbar, M., Alieyan, K., and Alzubaidi, M. (2017). Internet of things (IoT) communication protocols: Review. In *2017 8th International Conference on Information Technology (ICIT)*, pages 685–690.
- [Alpaydin 2014] Alpaydin, E. (2014). *Introduction to Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 3 edition.
- [Badarla and Harigovindan 2021] Badarla, S. P. and Harigovindan, V. P. (2021). Restricted access window-based resource allocation scheme for performance enhancement of IEEE 802.11ah multi-rate IoT networks. *IEEE Access*, 9:136507–136519.
- [Baños-Gonzalez et al. 2016] Baños-Gonzalez, V., Afaqui, M. S., Lopez-Aguilera, E., and Garcia-Villegas, E. (2016). IEEE 802.11ah: A technology to face the IoT challenge. *Sensors*, 16(11).
- [Chang et al. 2019] Chang, T.-C., Lin, C.-H., Lin, K. C.-J., and Chen, W.-T. (2019). Traffic-aware sensor grouping for IEEE 802.11ah networks: Regression based analysis and design. *IEEE Transactions on Mobile Computing*, 18(3):674–687.
- [Damayanti et al. 2016] Damayanti, W., Kim, S., and Yun, J.-H. (2016). Collision chain mitigation and hidden device-aware grouping in large-scale IEEE 802.11ah networks. *Computer Networks*, 108:296–306.
- [Gomez et al. 2012] Gomez, C., Oller, J., and Paradells, J. (2012). Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors*, 12(9):11734–11753.
- [Gonzalez 1985] Gonzalez, T. F. (1985). Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306.
- [Hazmi et al. 2012] Hazmi, A., Rinne, J., and Valkama, M. (2012). Feasibility study of IEEE 802.11ah radio technology for IoT and M2M use cases. In *2012 IEEE Globecom Workshops*, pages 1687–1692.
- [Hu 2019] Hu, C.-C. (2019). Approximation algorithms of minimizing hidden pairs in 802.11ah networks. *IEEE Access*, 7:170742–170752.

- [IEEE 2016] IEEE (2016). *IEEE Standard for Information technology—Telecommunications and information exchange between systems Local and metropolitan area networks—Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.*
- [IEEE 2017] IEEE (2017). *IEEE Standard for Information technology—Telecommunications and information exchange between systems - Local and metropolitan area networks—Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Sub 1 GHz License Exempt Operation.*
- [Kai et al. 2019] Kai, C., Zhang, J., Zhang, X., and Huang, W. (2019). Energy-efficient sensor grouping for IEEE 802.11ah networks with max-min fairness guarantees. *IEEE Access*, 7:102284–102294.
- [Khorov et al. 2015] Khorov, E., Lyakhov, A., Krotov, A., and Guschin, A. (2015). A survey on IEEE 802.11ah: An enabling networking technology for smart cities. *Computer Communications*, 58:53–69. Special Issue on Networking and Communications for Smart Cities.
- [Khorov et al. 2020] Khorov, E., Lyakhov, A., Nasedkin, I., Yusupov, R., Famaey, J., and Akyildiz, I. F. (2020). Fast and reliable alert delivery in mission-critical Wi-Fi HaLow sensor networks. *IEEE Access*, 8:14302–14313.
- [Mahesh and Harigovindan 2019] Mahesh, M. and Harigovindan, V. P. (2019). Restricted access window-based novel service differentiation scheme for group-synchronized DCF. *IEEE Communications Letters*, 23(5):900–903.
- [Mahesh et al. 2020] Mahesh, M., Pavan, B. S., and Harigovindan, V. (2020). Data rate-based grouping using machine learning to improve the aggregate throughput of IEEE 802.11ah multi-rate IoT networks. In *2020 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, pages 1–5.
- [Nabuuma and Alsusa 2017] Nabuuma, H. and Alsusa, E. (2017). Enhancing the throughput of 802.11ah sectorized networks using aid-based backoff counters. In *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1921–1926.
- [Nawaz et al. 2017] Nawaz, N., Hafeez, M., Zaidi, S., McLernon, D., and Ghogho, M. (2017). Throughput enhancement of restricted access window for uniform grouping scheme in IEEE 802.11ah. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–7.
- [nsgn 2021] nsgn (2021). ns-3 – a discrete-event network simulator for internet systems. [Online; accessed 6-September-2022].

- [Oliveira et al. 2022] Oliveira, E. C., Soares, S. M., and Carvalho, M. M. (2022). K-means based grouping of stations with dynamic aid assignment in IEEE 802.11ah networks. In *The 18th International Conference on Mobility, Sensing and Networking (MSN), December 14–16, Guangzhou, China, 2022*.
- [Park 2015] Park, M. (2015). IEEE 802.11ah: sub-1-GHz license-exempt operation for the internet of things. *IEEE Communications Magazine*, 53(9):145–151.
- [Pei and Henderson 2009] Pei, G. and Henderson, T. (2009). Validation of ns-3 802.11b PHY model. *ns3*.
- [Raza et al. 2017] Raza, U., Kulkarni, P., and Sooriyabandara, M. (2017). Low power wide area networks: An overview. *IEEE Communications Surveys & Tutorials*, 19(2):855–873.
- [Riley and Henderson 2010] Riley, G. and Henderson, T. (2010). *The ns-3 Network Simulator*, pages 15–34.
- [Rose et al. 2015] Rose, K., Eldridge, S., and Chapin, L. (2015). The internet of things: An overview. *The internet society (ISOC)*, 80:1–50.
- [Sangeetha and Babu 2020] Sangeetha, U. and Babu, A. (2020). Fair and efficient resource allocation in IEEE 802.11ah WLAN with heterogeneous data rates. *Computer Communications*, 151:154–164.
- [Sieranoja 2019] Sieranoja, S. (2019). How much k-means can be improved by using better initialization and repeats? *Pattern Recognition*, 93.
- [Sun et al. 2013] Sun, W., Choi, M., and Choi, S. (2013). IEEE 802.11ah a long range 802.11 WLAN at sub 1 Ghz. *Journal of ICT Standardization*, 1:83–107.
- [Tian et al. 2016] Tian, L., Deronne, S., Latré, S., and Famaey, J. (2016). Implementation and validation of an IEEE 802.11ah module for ns-3.
- [Tian et al. 2017a] Tian, L., Khorov, E., Latré, S., and Famaey, J. (2017a). Real-time station grouping under dynamic traffic for IEEE 802.11ah. *Sensors*, 17(7).
- [Tian et al. 2017b] Tian, L., Santi, S., Latré, S., and Famaey, J. (2017b). Accurate sensor traffic estimation for station grouping in highly dense IEEE 802.11ah networks. In *Proceedings of the First ACM International Workshop on the Engineering of Reliable, Robust, and Secure Embedded Wireless Sensing Systems, FAILSAFE’17*, page 1–9, New York, NY, USA. Association for Computing Machinery.
- [Tian et al. 2021] Tian, L., Santi, S., Seferagić, A., Lan, J., and Famaey, J. (2021). Wi-Fi HaLow for the internet of things: An up-to-date survey on IEEE 802.11ah research. *Journal of Network and Computer Applications*, 182:103036.

- [Tian et al. 2018] Tian, L., Šljivo, A., Santi, S., De Poorter, E., Hoebeke, J., and Famaey, J. (2018). Extension of the IEEE 802.11ah ns-3 simulation module. In *Proceedings of the 2018 Workshop on Ns-3, WNS3 '18*, page 53–60, New York, NY, USA. Association for Computing Machinery.
- [Wang et al. 2015] Wang, Y., Li, Y., Chai, K. K., Chen, Y., and Schormans, J. (2015). Energy-aware adaptive restricted access window for IEEE 802.11ah based smart grid networks. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 581–586.

Apêndice A

Implementação dos Algoritmos de Inicialização do K-Means

Neste apêndice são apresentadas as implementações de três métodos de inicialização do algoritmo K-Means: inicialização aleatória, inicialização pela heurística do ponto mais distante e inicialização pela heurística do ordenamento. Os códigos aqui apresentados fazem parte da classe *K-Means* apresentada na Seção 4.5. Os algoritmos de inicialização recebem um vetor de N pontos e o número de grupos k , e, seguindo seus respectivos critérios, eles escolhem k pontos como centroides iniciais de k grupos.

A.1 Inicialização Aleatória

O algoritmo de inicialização aleatória do K-Means é mostrado na Listagem A.1. Este algoritmo deve escolher k pontos aleatórios como centroides iniciais para os k grupos. Na linha 4 o vetor de grupos obtidos em outra execução do algoritmo é limpo. Na linha 6 o vetor dos pontos aleatórios escolhidos é inicializado sem nenhum ponto. O laço das linhas 8-24 é executado k vezes, e cada iteração corresponde à escolha do centroide para um grupo. Em uma iteração, um laço infinito é executado até que um novo ponto aleatório seja encontrado (linhas 10-23). Na linha 12, o valor de índice de um ponto aleatório é escolhido. Verifica-se nas linhas 14-15 se este ponto já está sendo usado como centroide (ou seja, se o ponto já foi adicionado ao vetor inicializado na linha 6). Caso o ponto ainda não tenha sido usado, seu índice é adicionado ao vetor com os índices dos pontos em uso (linha 17), o identificador do seu grupo é definido como o índice do grupo da iteração (linha 18) e um objeto *Cluster* é criado tendo o ponto escolhido como centroide (linha 20). Uma vez que um ponto foi encontrado, o laço infinito é interrompido e a próxima iteração do laço *for* é executada (linha 21).

Listagem A.1: Rotina que implementa a inicialização aleatória dos grupos do K-Means.

```
1 void
```

```

2 KMeans::RandomInitialization(std::vector<Point> &all_points, uint16_t k)
3 {
4     m_clusters.clear();
5
6     std::vector<uint16_t> used_pointIds;
7
8     for (uint16_t i = 1; i <= k; i++)
9     {
10        while (true)
11        {
12            int index_point = rand() % total_points;
13
14            if (find(used_pointIds.begin(), used_pointIds.end(),
15                used_pointIds) == used_pointIds.end())
16            {
17                used_pointIds.push_back(index_point);
18                all_points[index_point].setCluster(i);
19                Cluster cluster(i, all_points[index_point]);
20                clusters.push_back(cluster);
21                break;
22            }
23        }
24    }
25    Clusterize(all_points);
26 }

```

A.2 Inicialização pela Heurística do Ponto de Maior Distância

O código do algoritmo de inicialização pela heurística do ponto mais distante é mostrado na Listagem A.2. O objetivo deste algoritmo é escolher k pontos iniciais de modo que eles estejam o mais distante possível entre si. Na linha 4 o vetor com os clusters criados em outra execução do algoritmo é limpo. O laço das linhas 5-47 é executado k vezes, em que cada iterações corresponde à escolha do centroide de um novo grupo. Nas linhas 7-19 o primeiro centroide é definido como um ponto aleatório, com processo similar ao que é feito na inicialização aleatória. Nas linhas 15-17 todos os pontos são incluídos no primeiro grupo criado. Nas linhas 24-42 um laço *for* é executado para o número total de pontos. Para cada ponto, calcula-se a distância euclidiana entre ele e o centroide do grupo ao qual ele pertence (linhas 31-36). Se esta distância for a maior distância calculada até o momento, este ponto é salvo como o ponto mais distante e o valor da maior distância é atualizado. Após iterar por todos os pontos, o ponto com maior distância do seu centroide é escolhido como um novo centroide e um novo grupo é criado a partir dele (linhas 43-45). Ao final da execução, o algoritmo tem k grupos criados com os pontos de maior distância como centroides iniciais.

Listagem A.2: Rotina que implementa a inicialização dos grupos do K-Means pela heurística do ponto mais distante.

```

1  void
2  KMeans::MinMaxInitialization(std::vector<Point> &all_points, uint16_t k)
3  {
4      m_clusters.clear();
5      for (uint16_t i = 0; i < k; i++)
6      {
7          if (i == 0)
8          {
9              int index = rand() % m_total_points;
10
11             Cluster cluster(i+1, all_points[index]);
12             m_clusters.push_back(cluster);
13
14
15             for (uint16_t j = 0; j < all_points.size(); j++)
16             {
17                 all_points[j].setCluster(i+1);
18             }
19         }
20         else
21         {
22             uint16_t furthestPoint = 0;
23             double maxDist = 0;
24             for (uint16_t m = 0; m < all_points.size(); m++)
25             {
26                 uint16_t currentClusterId = all_points[m].getCluster();
27
28                 Cluster currentCluster = m_clusters[currentClusterId - 1];
29
30                 double sum = 0;
31                 for (uint16_t n = 0; n < m_dimensions; n++)
32                 {
33                     sum += pow(currentCluster.getCentroidByPos(n) -
34                               all_points[m].getVal(n), 2);
35                 }
36
37                 double dist = sqrt(sum);
38                 if (dist > maxDist)
39                 {
40                     maxDist = dist;
41                     furthestPoint = m;
42                 }
43             }
44             Cluster cluster(i+1, all_points[furthestPoint]);
45             m_clusters.push_back(cluster);
46             Clusterize(all_points);
47         }
48     }
49 }

```



```

47     }
48 }

```

A.3 Inicialização pela Heurística de Ordenamento

A implementação do algoritmo de inicialização pela heurística de ordenamento é mostrada na Listagem A.3. Este algoritmo deve escolher k pontos com base no ordenamento seguindo o critério da distância euclidiana entre os pontos e o ponto zero ($\vec{0}$). O laço das linhas 17-31 calcula a distância do ponto zero para cada um dos pontos de entrada. Na linha 33, os valores de distância são ordenados em ordem crescente. A variável da linha 35 trata-se do intervalo de amostragem dos pontos na ordem crescente. Nas linhas 38-42, um ponto é escolhido a cada N/k pontos.

Listagem A.3: Rotina que implementa a inicialização dos grupos do K-Means pela heurística de ordenamento.

```

1  struct Distance
2  {
3      int index;
4      double value;
5  };
6
7  bool CompareDistance(Distance d1, Distance d2)
8  {
9      return (d1.value < d2.value);
10 }
11
12 void
13 KMeans::SortingAttrInitialization(std::vector<Point> &all_points,
14     uint16_t k)
15 {
16     std::vector<Distance> allValues;
17
18     for (uint16_t i = 0; i < all_points.size(); i++)
19     {
20         double sum = 0.0;
21         for (uint16_t j = 0; j < m_dimensions; j++)
22         {
23             sum += pow(all_points[i].getVal(j), 2);
24         }
25         double d = sqrt(sum);
26
27         Distance dist;
28         dist.index = i;
29         dist.value = d;
30
31         allValues.push_back(dist);

```

```

31     }
32
33     std::sort(allValues.begin(), allValues.end(), CompareDistance);
34
35     uint16_t step = floor(all_points.size() / k);
36
37     m_clusters.clear();
38     for (uint16_t q = 0; q < k; q++)
39     {
40         Cluster cluster(q+1, all_points[allValues[q*step].index]);
41         m_clusters.push_back(cluster);
42     }
43     Clusterize(all_points);
44 }

```