

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
DAS - DEPARTAMENTO DE AUTOMAÇÃO E  
SISTEMAS**

Daniel Siedersberger

**ABORDAGENS PARA FORMAÇÃO DE TOPOLOGIA  
EM AGRUPAMENTOS EM ÁRVORE PARA RSSF DE  
LARGA ESCALA**

Florianópolis

2019



Daniel Siedersberger

**ABORDAGENS PARA FORMAÇÃO DE TOPOLOGIA  
EM AGRUPAMENTOS EM ÁRVORE PARA RSSF DE  
LARGA ESCALA**

Dissertação submetida ao Programa  
de Pós-Graduação em Engenharia de  
Automação e Sistemas para a obtenção  
do Grau de Mestre.

Orientador: Prof. Carlos Barros Mon-  
tez, Dr

Coorientador: Prof. Erico Meneses Leão,  
Dr

Florianópolis

2019

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Siedersberger, Daniel

Abordagens para Formação de Topologia em  
Agrupamentos em Árvore para RSSF de Larga Escala /  
Daniel Siedersberger ; orientador, Carlos Barros  
Montez, coorientador, Erico Meneses Leão, 2019.  
118 p.

Dissertação (mestrado) - Universidade Federal de  
Santa Catarina, Centro Tecnológico, Programa de Pós  
Graduação em Engenharia de Automação e Sistemas,  
Florianópolis, 2019.

Inclui referências.

1. Engenharia de Automação e Sistemas. 2. Rede de  
sensores sem fio. 3. Agrupamentos em árvore. 4.  
Formação de topologia. 5. IEEE 802.15.4. I. Montez,  
Carlos Barros. II. Leão, Erico Meneses. III.  
Universidade Federal de Santa Catarina. Programa de  
Pós-Graduação em Engenharia de Automação e Sistemas.  
IV. Título.

Daniel Siedersberger

**ABORDAGENS PARA FORMAÇÃO DE TOPOLOGIA  
EM AGRUPAMENTOS EM ÁRVORE PARA RSSF DE  
LARGA ESCALA**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Automação e Sistemas.

Florianópolis, 02 de abril 2019.

---

Werner Kraus Junior  
Coordenador do PGEAS  
Universidade Federal de Santa Catarina

**Banca Examinadora:**

---

Prof. Carlos Barros Montez, Dr  
Universidade Federal de Santa Catarina  
Presidente/Orientador

---

Prof. Tiago Semprebom, Dr  
Instituto Federal de Santa Catarina

---

Prof. Frank Siqueira, Dr  
Universidade Federal de Santa Catarina

---

Prof. Marcelo Ricardo Stemmer, Dr  
Universidade Federal de Santa Catarina



Este trabalho é dedicado aos meus queridos familiares e amigos.





## **AGRADECIMENTOS**

Sou extremamente grato pela oportunidade de realizar este trabalho, o qual me permitiu ter muitas experiências e aprendizados. Agradeço aos professores Carlos Barros Montez e Erico Meneses Leão pela orientação realizada e todo suporte oferecido. Agradeço também à minha família por ter me dado todo apoio necessário, e minha companheira Paola por ter estado sempre ao meu lado. Por fim, agradeço a CAPES pelo apoio financeiro concedido.



*Education is not the learning of facts, it's  
rather the training of the mind to think.*

(Albert Einstein)



## RESUMO

Topologias em agrupamentos em árvore (do Inglês, *cluster tree*) são comumente empregadas em aplicações de Redes de Sensores sem Fio em larga escala, onde os nodos são implantados de forma aleatória. Contudo, o uso dessa topologia apresenta uma série de desafios em relação à conectividade e ao escalonamento de mensagens, sendo necessários métodos eficazes para formação e operação da mesma. Neste contexto, o presente trabalho implementa e avalia o funcionamento de dois métodos de formação de topologia. Primeiramente, um esquema baseado nos protocolos *Spanning Tree* e *Shortest Path First* é implementado e avaliado sobre diversas condições. Com base nos resultados observados, este trabalho propõe a arquitetura ACArch, que é composta pela combinação de métodos de descobrimento de vizinhança, formação da rede e escalonamento de mensagens. Essa proposta visa formar uma rede flexível, com melhor desempenho e capaz de se adaptar a diversos cenários. Resultados das simulações realizadas mostram uma significativa melhora no desempenho da rede quando comparada a outros protocolos.

**Palavras-chave:** RSSF. Agrupamentos em árvore. Formação de topologia.



## ABSTRACT

Cluster tree topologies are commonly adopted in large-scale Wireless Sensor Networks applications, where nodes are randomly deployed. However, the use of this topology presents several challenges regarding connectivity and message schedulability. In this context, the present work implements and evaluates the operation of two cluster tree topology formation methods. Firstly, a scheme based on the Spanning Tree and Shortest Path First protocols is implemented and evaluated on several conditions. Then, based on the observed results, this work proposes the ACArch architecture, that is composed of the combination of methods of neighborhood discovery, network formation and message scheduling. This proposal aims to build a flexible network, with better performance and able to adapt for different scenarios. Simulation results showed significant improvements in the performance of the network when compared to others protocols.

**Keywords:** WSN. Cluster tree. Topology formation.





## LISTA DE FIGURAS

Figura 1	Rede de Sensores Sem Fio. ....	23
Figura 2	Exemplo da topologia em agrupamentos em árvores. ....	32
Figura 3	Estrutura do <i>Superframe</i> . ....	34
Figura 4	Algoritmo <i>Slotted CSMA-CA</i> . ....	37
Figura 5	Arquitetura do padrão ZigBee. ....	38
Figura 6	Exemplo de relações em uma rede em agrupamentos em árvore. ....	39
Figura 7	Esquema de implementação do modelo CT-SIM. ....	42
Figura 8	Esquema de escalonamento <i>Bottom-Up</i> . ....	44
Figura 9	Experimento para definição da distância limite das relações entre nodos. ....	70
Figura 10	(a) Grafo $G = (V, E)$ . (b) Árvore Geradora Mínima ponderada por distância (MST). (c) Árvore de Menor Caminho ponderada por distância (SPD). (d) Árvore de Menor Caminho ponderada por número de saltos (SPH). ....	72
Figura 11	Avaliação das árvores construídas pelos métodos MST, SPD e SPH em relação à (a) Número de agrupamentos, (b) Profundidade Média e (c) Distância Média. ....	73
Figura 12	Análise de escalonamento para os métodos MST, SPD e SPH. ....	75
Figura 13	Distribuição de energia consumida e atraso fim-a-fim para o método MST. ....	77
Figura 14	Distribuição de Energia consumida e Atraso fim-a-fim para o método SPD. ....	78
Figura 15	Distribuição de Energia consumida e Atraso fim-a-fim para o método SPH. ....	78
Figura 16	Média entre os nodos de um cenário específico do (a) Consumo de Energia e do (b) Atraso fim-a-fim dos métodos MST, SPD e SPH. ....	79
Figura 17	Média entre 10 cenários diferentes do (a) Consumo de Energia e do (b) Atraso fim-a-fim dos métodos MST, SPD e SPH. ....	80
Figura 18	Relação entre energia consumida e pacotes transmitidos para: (a) Cenário específico, (b) Diversos cenários. ....	80
Figura 19	Adaptative Clustering Architecture. ....	85

Figura 20 Exemplo da etapa de descobrimento. ....	86
Figura 21 Experimento para definição do número de <i>slots</i> utilizado na etapa de Descobrimento de Vizinhança. ....	92
Figura 22 Experimento para definição do parâmetro <i>RSSI_limit</i> . .	94
Figura 23 Análise da variação do parâmetro <i>maxChildren</i> para (a) Taxa de sucesso, (b) Atraso fim-a-fim e (c) Consumo de energia...	94
Figura 24 Representação da formação para os métodos (a) ACArch, (b) SPH e (c) LEACH. ....	96
Figura 25 Número médio de agrupamentos formados para os cenários de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	97
Figura 26 Número médio de saltos entre todos nodos para os cenários de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	97
Figura 27 Número médio de nodos órfãos para os cenários de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	98
Figura 28 Análise de escalabilidade dos métodos ACArch e SPH. .	99
Figura 29 Energia consumida pelos nodos para os métodos (a) ACArch, (b) SPH e (c) LEACH. ....	102
Figura 30 Análise do consumo médio de energia com variação de nodos para um cenário de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	102
Figura 31 Análise do consumo global médio de energia pelo tempo decorrido de um cenário específico (a) para toda simulação e (b) para o período de configuração. ....	103
Figura 32 Atraso fim-a-fim médio das mensagens de cada nodo para os métodos (a) ACArch, (b) SPH e (c) LEACH. ....	104
Figura 33 Análise do atraso fim-a-fim com variação de nodos para um cenário de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	105
Figura 34 Taxa de sucesso média de cada nodo para os métodos (a) ACArch, (b) SPH e (c) LEACH. ....	106
Figura 35 Análise da taxa de sucesso com variação de nodos para um cenário de (a) 50x50 $m^2$ e (b) 100x100 $m^2$ . ....	106

## LISTA DE TABELAS

Tabela 1	Parâmetros para implantação dos nodos. ....	42
Tabela 2	Parâmetros para formação dos agrupamentos. ....	43
Tabela 3	Parâmetros para configuração do superframe. ....	46
Tabela 4	Parâmetros para configuração do tráfego de monitora- mento. ....	47
Tabela 5	Comparação dos protocolos de formação de topologia ..	60
Tabela 6	Parâmetros de formação de topologia em árvore. ....	70
Tabela 7	Resultados médios das características de formação. ....	73
Tabela 8	Parâmetros de escalonamento de topologia em árvore. .	74
Tabela 9	Parâmetros utilizados na comunicação da rede. ....	76
Tabela 10	Parâmetros de descobrimento de vizinhança empregados pela ACArch. ....	92
Tabela 11	Parâmetros utilizados na formação da rede. ....	93
Tabela 12	Parâmetros de comunicação do protocolo LEACH. ....	100
Tabela 13	Parâmetros utilizados na comunicação da rede. ....	101



## LISTA DE ABREVIATURAS E SIGLAS

ACArch	Adaptative Clustering Architecture
ACK	Acknowledgement
BE	Backoff Exponent
BFS	Breadth-First Search
BI	Beacon Interval
BO	Beacon Order
BS	Base Station
CAP	Contention Access Period
CCA	Clear Channel Assessment
CFP	Contention-Free Period
CH	Cluster Head
CSMA-CA	Carrier Sense Multiple Access with Collision Avoidance
CT-SIM	Cluster Tree Simulation Model
CW	Contention Window
FFD	Full Function Device
GPS	Global Positioning System
GTS	Garanteed Time Slot
IEEE	Institute of Electircal and Electronics Engineers
ISM	Industrial, Scientific and Medical
LEACH	Low-Energy Adaptive Clustering Hierarchy
LQI	Link Quality Indicator
LR-WPAN	Low-Rate Wireless Personal Area Network
MAC	Medium Access Control
MST	Minimum Spanning Tree
NB	Number of Backoffs
OSI	Open Systems Interconnection
PAN	Personal Area Network
PHY	Physical Layer
PRR	Probabilistic Round Robin
RFD	Reduced Function Device
RSSF	Rede de Sensores Sem Fio
RSSI	Received Signal Strength Indicator

SD	Superframe Duration
SDA	Superframe Duration Allocation
SO	Superframe Order
SPD	Shortest Path by Distance
SPH	Shortest Path by Hop
SPT	Shortest Path First
TDMA	Time Division Multiple Access

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	23
1.1 CONTEXTUALIZAÇÃO	23
1.1.1 Aplicações	24
1.1.2 Descrição do Problema	25
1.2 MOTIVAÇÃO E OBJETIVOS	28
1.2.1 Objetivos Específicos	29
1.3 ORGANIZAÇÃO DO TEXTO	29
<b>2 FUNDAMENTAÇÃO TEÓRICA</b>	31
2.1 PADRÃO IEEE 802.15.4	31
2.1.1 Topologias de Rede	31
2.1.2 Camada Física (PHY) do IEEE 802.15.4	33
2.1.3 Subcamada de Controle de Acesso ao Meio (MAC)	34
2.1.4 CSMA-CA	35
2.2 ZIGBEE	36
2.2.1 Camada de Rede ZigBee	37
2.2.1.1 Topologia em Agrupamentos em Árvore	38
2.2.1.2 Roteamento	39
2.2.2 Camada de Aplicação ZigBee	40
2.3 MODELO DE SIMULAÇÃO PARA REDES EM AGRUPAMENTOS EM ÁRVORE	40
2.3.1 Formação da Rede	41
2.3.1.1 Implantação dos Nodos	42
2.3.1.2 Formação dos Agrupamentos	43
2.3.1.3 Esquema de Endereçamento	43
2.3.1.4 Escalonamento dos Períodos Ativos	44
2.3.1.5 Configuração do Superframe	44
2.3.2 Comunicação	47
2.3.3 Coleta de Estatísticas	48
2.4 CONCLUSÕES DO CAPÍTULO	48
<b>3 TRABALHOS RELACIONADOS</b>	51
3.1 FORMAÇÃO DE TOPOLOGIA BASEADA EM AGRUPAMENTOS	51
3.2 FORMAÇÃO DE TOPOLOGIA BASEADA EM ÁRVORES	55
3.3 ESQUEMAS DE COMUNICAÇÃO BASEADOS NO PADRÃO IEEE 802.15.4	57
3.4 COMPARAÇÃO ENTRE TRABALHOS RELACIONADOS	58
3.5 CONCLUSÕES DO CAPÍTULO	61

<b>4 ESQUEMAS DE FORMAÇÃO DE TOPOLOGIAS EM ÁRVORES</b>	63
4.1 MODELO DO SISTEMA	63
4.1.1 Principais Pressupostos	64
4.2 DESCRIÇÃO DO MÉTODO	66
4.2.1 Árvore Geradora Mínima	66
4.2.2 Árvore de Menor Caminho	67
4.2.3 Esquema de comunicação	68
4.3 SIMULAÇÃO E RESULTADOS	69
4.3.1 Características de Formação	69
4.3.2 Análise de Escalabilidade	74
4.3.3 Análise de Desempenho	76
4.4 CONSIDERAÇÕES FINAIS	81
<b>5 ACARCH: ADAPTATIVE CLUSTERING ARCHITECTURE</b>	83
5.1 DESCRIÇÃO DO MODELO	83
5.1.1 Principais Pressupostos	83
5.2 DESCRIÇÃO DA ARQUITETURA	84
5.2.1 Descobrimto de Vizinhaça	85
5.2.2 Associação de Filhos	88
5.2.3 Seleção de Cluster-Heads	89
5.3 SIMULAÇÕES E RESULTADOS	90
5.3.1 Análise da formação da rede	91
5.3.1.1 Definição dos parâmetros de descobrimento	91
5.3.1.2 Definição dos parâmetros de formação	92
5.3.1.3 Análise dos resultados	95
5.3.2 Análise de Escalabilidade	98
5.3.3 Análise de Desempenho	100
5.4 CONSIDERAÇÕES FINAIS	107
<b>6 CONCLUSÕES E TRABALHOS FUTUROS</b>	109
6.1 VISÃO GERAL DO TRABALHO	109
6.2 PRINCIPAIS CONTRIBUIÇÕES	110
6.3 TRABALHOS FUTUROS	110
<b>REFERÊNCIAS</b>	113



# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

O grande avanço tecnológico em sistemas micro eletromecânicos, na comunicação sem fio e na eletrônica digital, vem contribuindo para uma ampla difusão no uso de sistemas de sensores distribuídos (DARGIE; POELLABAUER, 2010). Com o crescente aumento na capacidade de sensoriamento, processamento, armazenamento e comunicação, dispositivos sensores têm se tornado presentes em uma gama cada vez maior de aplicações de monitoramento (AKYILDIZ; VURAN, 2010).

Quando grupos de sensores com uma infraestrutura de comunicação sem fio são dispostos em um ambiente físico a fim de monitorar cooperativamente as condições do mesmo (Figura 1), forma-se uma Rede de Sensores Sem Fio (RSSF). Além da capacidade de se comunicarem uns com os outros, os nodos enviam os dados para uma estação base (BS, do Inglês *Base Station*), a qual processa todos os dados remotamente. A BS pode transmitir os dados coletados via Internet ou diretamente a um usuário final para que sejam analisados e armazenados (DARGIE; POELLABAUER, 2010; AKYILDIZ; VURAN, 2010).

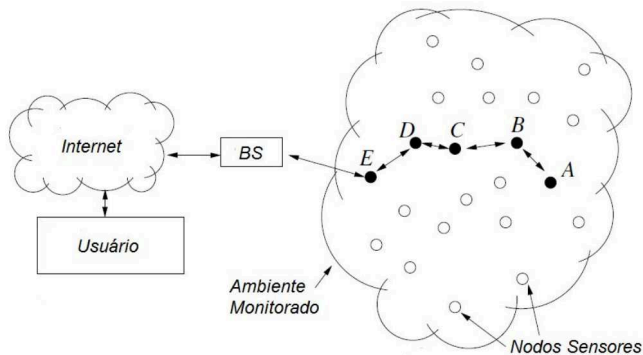


Figura 1: Rede de Sensores Sem Fio.

Os nodos que compõem uma RSSF são dispositivos pequenos e de baixo custo, capazes de coletar informações físicas do ambiente, processá-las localmente e transmiti-las aos seus vizinhos. Para isso, eles são constituídos, basicamente, pelos seguintes componentes:

- **Unidade de Sensoriamento:** geralmente inclui diversos tipos de transdutores que coletam informações físicas do ambiente, tais como, temperatura, pressão, nível, umidade, entre outras.
- **Unidade de Processamento:** essa unidade é a controladora principal do nodo, a qual gerencia todos outros componentes, executa operações de sensoriamento, executa algoritmos e colabora com os nodos da rede via comunicação sem fio.
- **Memória Interna:** utilizada para o armazenamento de dados locais, como por exemplo, identificação de vizinhos e de rotas de transmissão.
- **Módulo de Comunicação:** constituído pelo transceptor de rádio, que através da implementação dos procedimentos de modulação, é responsável pela comunicação entre os nodos.
- **Fonte de Energia:** provê energia para o funcionamento dos nodos sensores, normalmente constituída de baterias, as quais possuem capacidade limitada.

### 1.1.1 Aplicações

As RSSF podem ser formadas por diversos tipos de sensores, capazes de monitorar uma variedade enorme de condições do ambiente. Com isso, é possível implementar uma larga gama de aplicações, dentre as quais, se destacam as seguintes:

- **Aplicações industriais:** Devido ao baixo custo, fácil instalação, fácil manutenção, entre outras vantagens, as RSSF são uma alternativa promissora no contexto de aplicações industriais (AKYILDIZ; VURAN, 2010). Por outro lado, como muitas dessas aplicações exigem confiabilidade nos dados e garantia de tempo real, o uso dessas redes passa a ser um desafio nesse contexto, pois são prejudicadas por ruídos, obstáculos e interferências (ZHAO, 2011). Ainda assim, já existe uma grande variedade de aplicações de sucesso como, por exemplo, a descrita em (SHU et al., 2016) que utiliza sensores de gás para fazer o monitoramento de vazamentos com a intenção de delimitar zonas de perigo e zonas seguras dentro de uma indústria.
- **Aplicações militares:** As características de rápida instalação, auto-organização e tolerância a falhas fazem das RSSF uma infra-

estrutura de sensoriamento promissora adequada para aplicações militares. Alguns exemplos de sistemas que podem ser implementados com essas redes incluem os de comando, controle, comunicação, computação, inteligência, vigilância, reconhecimento e rastreamento (AKYILDIZ; VURAN, 2010).

- **Aplicações de monitoramento ambiental:** As RSSF podem ser empregadas tanto em aplicações de monitoramento ambiental em ambientes fechados, tais como os encontrados em construções e escritórios; como em ambientes abertos, tais como ambientes terrestres ou marinhos (OLIVEIRA; RODRIGUES, 2011; XU et al., 2014). Em ambos os casos, o uso de uma RSSF constituída com uma grande quantidade de nodos espalhados na área monitorada pode ser vantajoso por prover uma boa cobertura e precisão espacial.
- **Aplicações em saúde:** Uma grande quantidade de aplicações na área da saúde tem sido proposta utilizando RSSF, incluindo a monitoração de pacientes com epilepsia, cardíacos, em reabilitação e idosos. Diversas pesquisas na área vem se esforçando para construir sistemas que disponibilizem informações sobre doenças e prevenções, integrem as infra-estruturas de saúde, monitorem a saúde dos pacientes, alertem os responsáveis quando uma intervenção é necessária e reduzam a inconveniência e custo de visitas ao paciente (DARGIE; POELLABAUER, 2010).

### 1.1.2 Descrição do Problema

Sistemas eficientes para o monitoramento de grandezas em áreas com grandes dimensões têm sido propostos graças ao desenvolvimento de nova tecnologias e ao suporte algorítmico relacionados ao conceito das RSSF. Essas redes apresentam diversas características vantajosas quando comparadas com as soluções convencionais, tais como: monitoramento descentralizado próximo da grandeza de interesse, implantação rápida e baixo custo.

Normalmente, em aplicações de RSSF de larga escala, a topologia em agrupamentos em árvore é utilizada, pois, devido à sua estrutura hierárquica em grupos, ela permite maior cobertura e escalabilidade. Isso porque, neste tipo de aplicação, existe a necessidade de se enviar, geralmente através de vários saltos na rede, os dados monitorados pelos sensores na direção de uma estação base, onde estes irão receber seus

devidos tratamentos.

Neste contexto, surge uma série de desafios a serem superados, tais como:

- **Consumo energético:** Tipicamente, as aplicações em larga escala utilizam baterias como fonte de energia, e sua substituição ou recarga é impraticável. Portanto, os principais desafios associados ao projeto de RSSF estão relacionados com a capacidade energética limitada dos nodos e o tempo de vida da rede. No projeto de uma RSSF, a questão da eficiência energética precisa levar em consideração desde o tipo de *hardware* utilizado na camada física, os protocolos da subcamada MAC, os protocolos de roteamento na camada de rede, até o próprio algoritmo usado na aplicação que executa em cada nodo.
- **Topologia e autogerenciamento:** A instalação dos nodos em um ambiente de forma eficiente e que possibilite a monitoração dos fenômenos de interesse é um grande desafio. Por exemplo, sensores utilizados para monitoração de ambientes em larga escala são, geralmente, implantados de forma aleatória. Com isso, os nodos precisam se adaptar automaticamente utilizando as informações do ambiente e de seus vizinhos para iniciar sua comunicação e monitoramento. Além disso, o fato dos nodos estarem dispostos em um ambiente dinâmico impõe mudanças sobre a rede. Como exemplo, a conectividade entre os nodos pode ser alterada devido a obstáculos móveis, interferência ou falhas. Ainda, dependendo da aplicação, a rede pode sofrer mudanças periódicas quando, por exemplo, há um escalonamento entre os nodos, fazendo com que alguns deles sejam desligados para economizar energia (DARGIE; POELLABAUER, 2010; AKYILDIZ; VURAN, 2010).
- **Escalabilidade:** No sentido de aumentar a confiabilidade da comunicação e também permitir uma cobertura adequada da área monitorada, geralmente há uma redundância nos nodos espalhados na rede (FELSKE et al., 2013). Apesar de prover melhor tolerância a falhas e possuir vantagens em relação à cobertura e à conectividade, a implantação com uma maior densidade de nodos torna a escalabilidade da rede um grande desafio, devido ao alto número de mensagens geradas (AKYILDIZ; VURAN, 2010; LI et al., 2013).
- **Agrupamento de nodos:** A adoção de agrupamentos em uma RSSF traz diversas vantagens ao desempenho da rede, porém

apresenta uma série de desafios. Na maioria dos trabalhos, a seleção dos coordenadores é feita com objetivo de balancear o consumo energético da rede, porém estes não buscam manter a área de cobertura persistente (SORO; HEINZELMAN, 2009). Além disso, caso não haja uma rotação entre os nodos para cumprirem o papel de coordenadores da rede, os nodos coordenadores irão trabalhar mais e esgotar suas energias precocemente. A maioria dos protocolos de agrupamento considera o número de nodos fixo, resultando em um desbalanceamento na distribuição da carga de trabalho nos coordenadores e, conseqüentemente, na má distribuição do consumo de energia. Por fim, conforme (LEAO et al., 2017a), a dessincronização entre agrupamentos pode acarretar sérios problemas de colisão como, por exemplo, quando dois ou mais coordenadores tentam transmitir suas mensagens ao mesmo tempo. Portanto, os maiores desafios relacionados ao agrupamento de nodos incluem a seleção e rotação dos coordenadores, a seleção do número de nodos apropriado em cada agrupamento, a sincronização entre eles, e a manutenção da área de cobertura do ambiente monitorado.

Além desses desafios, por essas redes serem constituídas por dispositivos de baixo custo, a comunicação entre nodos é usualmente feita através de um único canal. Contudo, além de colisão entre *frames*, problemas relacionados ao consumo desnecessário de energia, tais como escuta inútil (*overhearing*), escuta ociosa (*idle listening*) e *overemitting*, precisam ser contornados por protocolos de controle de acesso ao meio (MAC). Resumidamente, escuta inútil se refere ao problema de um nodo que recebe *frames* destinados a outros, escuta ociosa ocorre quando um nodo permanece algum tempo escutando um canal ocioso, e *overemitting* quando um nodo envia *frames* para outro que não está pronto para recebê-los.

Nas RSSF, protocolos MAC usualmente lidam com essas questões através de um escalonamento do acesso ao canal envolvendo uma combinação de TDMA (*Time Division Multiple Access*) e de um *duty-cycle*. O TDMA divide o tempo de acesso ao canal usado pelos dispositivos em compartimentos (*slots*) de tempo, cada qual usado exclusivamente por um nodo; já o *duty-cycle* divide o tempo de comunicação dos nodos em períodos de atividade e de inatividade. Enquanto o TDMA evita colisões entre *frames*, o *duty-cycle*, de forma complementar, economiza energia evitando que um nodo fique ativo desnecessariamente.

Essas técnicas, por outro lado, possuem custos associados que poderiam ser evitados em aplicações onde poucos nodos transmitem si-

multaneamente. Nesse caso, protocolos baseados em contenção, como o CSMA-CA (*Carrier Sense Multiple Access with Collision Avoidance*), seriam suficientes. Dessa forma, muitos protocolos MAC existentes adotam abordagens mistas envolvendo CSMA-CA, TDMA e *duty-cycle*. Um exemplo é o padrão IEEE 802.15.4 (IEEE, 2011) que define especificações da camada física (PHY) e da subcamada MAC para redes de baixa potência, voltadas para aplicações com baixa demanda de taxa de dados e alcance limitado de comunicação (IEEE, 2011).

Por fim, dado esse conjunto de problemas relacionados à RSSF em larga escala, surge a seguinte questão:

- Como formar uma rede escalonável e flexível, capaz de manter a sua cobertura sem degradar seu desempenho em relação ao consumo de energia, aos atrasos fim-a-fim das mensagens e à taxa de sucesso na entrega dos dados?

## 1.2 MOTIVAÇÃO E OBJETIVOS

O monitoramento de ambientes de larga escala vêm sendo empregado de maneira crescente em vários setores industriais, urbanos, rurais, entre outros. Com isso, as RSSF têm se tornado uma infraestrutura popular utilizada para suportar tais aplicações.

Geralmente, em aplicações de monitoramento, diversos nodos são dispostos aleatoriamente para coletar informações sobre uma determinada área. Essas informações são enviadas periodicamente para BS, responsável pelo processamento e tratamento das mesmas.

Conforme descrito na Seção 1.1.2, devido à imprevisibilidade da localização dos nodos sensores no ambiente e no número de mensagens que são transmitidas pela rede, torna-se necessário a implementação de métodos capazes de superar tais problemas.

Neste contexto, o presente trabalho visa realizar uma análise sobre os principais fatores que integram a formação da topologia em agrupamentos em árvore e, com base nisso, propor uma arquitetura de gerenciamento desse tipo de topologia em redes compatíveis com o padrão IEEE 802.15.4, capaz de maximizar a taxa de sucesso da comunicação, reduzir os atrasos das mensagens, garantir a conectividade da rede e reduzir a energia consumida.

### 1.2.1 Objetivos Específicos

Para atingir o objetivo principal deste trabalho, os seguintes objetivos específicos foram definidos:

- 1) Analisar e compreender as principais especificações dos padrões IEEE 802.15.4 e ZigBee;
- 2) Analisar técnicas de formação existentes aplicadas à topologia em agrupamentos em árvore;
- 3) Implementar e analisar métodos de formação da topologia em agrupamentos em árvore integrados com um protocolo de comunicação baseado no padrão IEEE 802.15.4, utilizando técnicas centralizadas, tais como: árvore geradora mínima (*minimum spanning tree*) e árvore de menor caminho (*shortest path tree*);
- 4) Propor e implementar uma arquitetura que integre métodos de formação de topologia descentralizados com um protocolo de comunicação baseado no padrão IEEE 802.15.4;
- 5) Avaliar, por meio de simulações, a arquitetura proposta.

## 1.3 ORGANIZAÇÃO DO TEXTO

Este documento está organizado em seis capítulos. No segundo capítulo são descritos os principais fundamentos teóricos relacionados a este trabalho, tais como, os padrões IEEE 802.15.4/ZigBee e um modelo de comunicação e escalonamento de mensagens baseado nos mesmos. Logo após, no terceiro capítulo, o detalhamento e a comparação dos trabalhos relacionados são realizados. No quarto capítulo é feita uma avaliação sobre as principais características de algumas abordagens centralizadas, que são comumente utilizadas para formação de topologia em agrupamentos em árvore. Com base nessa avaliação, o quinto capítulo apresenta uma proposta de arquitetura descentralizada para formação de topologia. Por fim, o sexto capítulo apresenta as conclusões do trabalho, assim como sugestões para trabalho futuros.





## 2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais conceitos que formam a base teórica deste trabalho. Primeiramente, as principais características dos padrões IEEE 802.15.4 e ZigBee são descritas. Logo após, são apresentados o modelo de simulação de redes CT-SIM (LEAO et al., 2017b) e o esquema de escalonamento de *beacons* Load-SDA (LEAO et al., 2017a), os quais são utilizados como base na implementação e operação dos métodos propostos no presente estudo.

### 2.1 PADRÃO IEEE 802.15.4

O padrão IEEE 802.15.4 define a camada física (PHY) e a subcamada de acesso ao meio (MAC) para redes sem fio de área pessoal e baixa taxa de dados (LR-WPAN – *Low-Rate Wireless Personal Area Network*). Essas redes são projetadas para serem simples e possuir baixo custo de comunicação. Seus principais objetivos são permitir fácil instalação, transferência de dados confiável e tempo de vida da bateria prolongado (IEEE, 2011).

Segundo (IEEE, 2011), dois tipos de dispositivos podem participar de uma rede IEEE 802.15.4, são eles: dispositivos de função completa (FFD – Full Function Device) e dispositivos de função reduzida (RFD – Reduced Function Device). Enquanto os FFD podem exercer funções que exigem maior capacidade de seus recursos, como roteamento de mensagens, processamento de dados e coordenação de outros dispositivos, os RFD tem sua capacidade reduzida, possibilitando apenas atividades simples de coleta e envio de dados. Embora consuma menos energia, cada RFD possui a desvantagem de necessitar sempre estar associado a um FFD.

#### 2.1.1 Topologias de Rede

Dependendo dos requisitos da aplicação, o padrão IEEE 802.15.4 pode operar em topologia estrela ou par a par (P2P). Na topologia estrela, todos os dispositivos estabelecem uma comunicação direta com um coordenador PAN, o qual é responsável por iniciar, terminar e rotear todas as comunicações da rede. A topologia estrela é inadequada em aplicações de larga escala pois sua cobertura está limitada pelo alcance

de transmissão dos dispositivos. Por outro lado, na topologia par a par, cada dispositivo é capaz de se comunicar com qualquer outro dentro do seu alcance de transmissão, e dessa forma permite que formações mais complexas sejam implementadas. A topologia em agrupamentos em árvore é um caso especial da par a par, onde os nodos são agrupados e coordenados de forma centralizada por coordenadores de agrupamento, conhecidos como *cluster-heads* (CHs), conforme ilustrado na Figura 2. Os CHs são responsáveis pela formação, comunicação e sincronização de seu respectivo agrupamento, e são interconectados através de relações de pai e filho, formando uma estrutura de rede hierárquica em árvore.

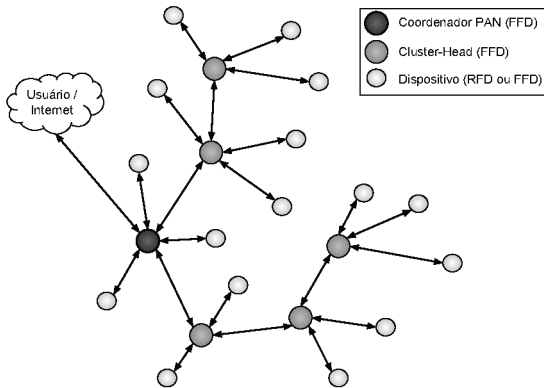


Figura 2: Exemplo da topologia em agrupamentos em árvores.

O primeiro agrupamento de uma rede em agrupamentos em árvore é formado pelo coordenador PAN, o qual é responsável pelo gerenciamento das atividades da rede e normalmente assume o papel de nodo coletor (estação base).

Na Figura 2, o coordenador está representado com uma ligação à Internet. Esse tipo de configuração, apesar de não ser obrigatória, é muito comum nos atuais sistemas de Internet das Coisas (*IoT - Internet of Things*).

Os CHs são responsáveis por coletar os dados e prover serviços de sincronização aos dispositivos associados em seu agrupamento. Os métodos de agrupamento e de seleção dos CHs não são especificados no padrão IEEE 802.15.4, e podem obedecer a uma série de critérios que variam conforme os requisitos da aplicação.

### 2.1.2 Camada Física (PHY) do IEEE 802.15.4

Nas redes IEEE 802.15.4, a camada PHY é responsável por uma série de funções, entre elas, as principais são:

- Ativação ou desativação do transceptor de rádio, que possibilita a transmissão e recepção de dados, e permite a economia de energia através do acionamento do modo *sleep*;
- Detecção de energia do canal, que permite ao dispositivo obter o pico de energia em cada canal para, no caso do coordenador PAN, selecionar o canal de operação da rede;
- Indicação de qualidade da conexão (*LQI – Link Quality Indicator*) e indicação de força do sinal recebido (*RSSI - Received Signal Strength Indicator*), que permite aos protocolos MAC verificar a qualidade da comunicação entre dois dispositivos, possibilitando o estabelecimento de conexões com qualidade apropriada para comunicação;
- Avaliação de canal livre (*CCA – Clear Channel Assessment*), que permite aos protocolos MAC verificar se o canal se encontra ocupado ou não, sendo utilizado para evitar colisões de transmissão;

Existem três bandas de frequências diferentes em que o padrão IEEE 802.15.4 pode operar, são elas:

- 868–868,6 MHz (banda de 868 MHz)
- 902–928 MHz (banda de 915 MHz)
- 2400–2483,5 MHz (banda de 2,4 GHz)

A banda de 868 MHz é utilizada na Europa para um certo número de aplicações, incluindo RSSF. As outras duas bandas são parte das bandas de frequência industriais, científicas e médica (*ISM – Industrial, Scientific and Medical*). Enquanto a banda de 915 MHz é utilizada principalmente na América do Norte, a banda de 2,4 GHz é utilizada em todo mundo. Além disso, a taxa de dados nominal máxima em que o padrão IEEE 802.15.4 pode operar é de 250 Kbps.

### 2.1.3 Subcamada de Controle de Acesso ao Meio (MAC)

As principais características da subcamada MAC são: gerenciamento de *beacon*, acesso ao canal, gerenciamento de compartimentos de tempo garantidos (GTS – *Guaranteed Time Slot*), validação de quadros e associação e dissociação de dispositivos.

Além disso, o padrão permite o uso opcional de uma estrutura chamada de *superframe*, a qual é ilustrada na Figura 3. Quando essa estrutura é utilizada, a rede opera em modo *beacon* habilitado, caso contrário, a rede opera em modo *beacon* desabilitado. O formato da estrutura do *superframe* é definido pelo coordenador.

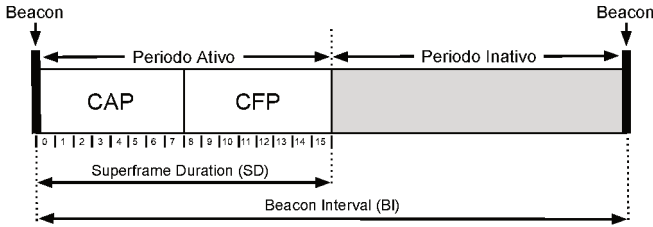


Figura 3: Estrutura do *Superframe*.

O *superframe* é limitado por quadros de *beacon*, os quais são transmitidos periodicamente pelo coordenador da PAN para sincronizar os nodos, identificar a rede e descrever a estrutura do *superframe*. O intervalo de tempo entre dois *beacons* é chamado *Beacon Interval* (BI), sendo constituído por uma porção ativa e outra inativa (opcional). A porção ativa é utilizada para os nodos transmitirem seus dados, e a inativa é utilizada para preservar os recursos energéticos. O período ativo corresponde ao *Superframe Duration* (SD) e pode ser dividido em duas partes, um período de acesso com contenção (CAP – *Contention Access Period*) e um período livre de contenção (CFP – *Contention Free Period*), sendo este último opcional.

Os parâmetros que definem o tamanho do BI e do SD são, respectivamente, o *Beacon Order* (BO) e o *Superframe Order* (SO) e estão relacionados da seguinte maneira:

$$BI = aBaseSuperframeDuration \times 2^{BO} \quad (2.1)$$

$$\text{para } 0 \leq BO \leq 14$$

e

$$SD = aBaseSuperframeDuration \times 2^{SO} \quad (2.2)$$

para  $0 \leq SO \leq BO \leq 14$

De acordo com (IEEE, 2011), o parâmetro *aBaseSuperframeDuration* possui o valor de 960 símbolos, sendo que cada símbolo é representado por um número específico de bits relacionado com tipo de modulação utilizada pelo transceptor de rádio. Como exemplo, quando é utilizada a banda 2,4 GHz, um símbolo equivale ao envio de 4 bits, o que corresponde ao tempo de  $16\mu s$ . Nesse caso, o valor de *aBaseSuperframeDuration* é de  $15360\mu s$ , e valores válidos de BI variam no intervalo desde  $15,3ms$  até o máximo de 4,2 minutos, aproximadamente.

Quando a estrutura do *superframe* é utilizada, o CAP começa imediatamente depois do primeiro beacon, seguido do CFP. Durante o CAP os nodos transmitem seus dados utilizando o mecanismo CSMA-CA. Logo após, no CFP, um número de *slots* de tempo garantidos (GTS – Guaranteed Time Slot) é dividido através do esquema TDMA.

As topologias em agrupamentos em árvores em conformidade com padrão IEEE 802.15.4 utilizam o modo *beacon-enabled* para comunicação e sincronização de seus agrupamentos. Cada agrupamento possui seu período ativo no qual deve acontecer a comunicação. Apesar de permitir a operação, este padrão não define quaisquer métodos para formação e comunicação em redes que empregam essa topologia. Contudo, alguns desses métodos são apresentados nas próximas seções.

### 2.1.4 CSMA-CA

Para prover acesso ao canal de comunicação durante o CAP, o padrão IEEE 802.15.4 emprega o mecanismo CSMA-CA. Existem duas variações deste mecanismo: *slotted* e *unslotted*. Quando a estrutura do *superframe* é utilizada, o mecanismo *slotted* CSMA-CA é empregado. Neste mecanismo, sempre que um dispositivo quer transmitir, ele escuta o meio realizando um CCA. Se o canal estiver ocupado, o dispositivo irá esperar um período de tempo aleatório para tentar de novo. Este período de espera é chamado de recuo ou *backoff*, sendo um múltiplo inteiro da unidade *aUnitBackoffPeriod* definida pelo padrão IEEE, que equivale a 20 símbolos.

O algoritmo empregado pelo *slotted* CSMA-CA, ilustrado na Figura 4, funciona da seguinte maneira:

1. Antes de qualquer tentativa de transmissão, três parâmetros são

inicializados, Número de *Backoffs* (NB), Janela de Contenção (CW – *Contention Window*) e Expoente de *Backoff* (BE – *Backoff Exponent*). Conforme o padrão IEEE 802.15.4, esses parâmetros são inicializados com os valores zero, dois e *macMinBE*, respectivamente.

2. Após a inicialização dos parâmetros, o nodo recua por um período de tempo aleatório entre o intervalo  $[0, 2^{BE} - 1]$ .
3. Assim que o período expira, o nodo executa duas CCAs. O número de CCAs necessárias para garantir que o canal esteja livre é controlado pela variável CW.
4. Caso o canal permanecer livre durante os dois CCAs a transmissão pode ocorrer. Caso contrário, o valor de BE é incrementado em um, e o nodo realiza um *backoff* de tempo aleatório entre um intervalo ainda maior. O valor de BE pode ser incrementado até atingir o valor máximo definido pela variável *macMaxBE*. Se o valor de BE atingir seu máximo, ele não pode ser alterado ao menos que a transmissão do pacote tenha sucesso ou falhe. Neste caso, o valor de BE é reniciado para o valor de *macMinBE*.
5. A cada tentativa de transmissão, o valor de NB também é incrementado por um. A falha na transmissão ocorre quando esse valor atinge um número máximo de *backoffs* definido pela variável *macMaxCSMABackoffs*.
6. Se a transmissão ocorre, um pacote de confirmação (ACK) é enviado pelo receptor do pacote. Caso o pacote de ACK não chegue, o nodo tenta retransmitir seu pacote de dados. Essa retransmissão pode ocorrer por um número máximo de vezes, definida pela variável *macMacFrameRetries*.

Dependendo da aplicação, a utilização de mensagens de confirmação não é necessária. Sendo assim, a etapa 6 do funcionamento do mecanismo CSMA-CA é opcional.

## 2.2 ZIGBEE

ZigBee é um padrão que define um conjunto de protocolos de comunicação para redes sem fio de baixa taxa de dados e curtas distâncias (SPECIFICATION, 2008). Suas principais aplicações são aquelas onde a energia é limitada pelo uso de baterias e, portanto, requerem

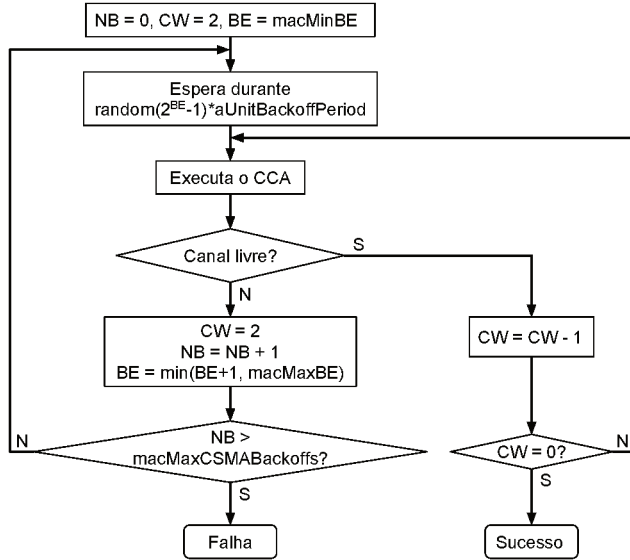


Figura 4: Algoritmo *Slotted* CSMA-CA.

baixo consumo energético, baixa taxa de dados e baixo custo para sua operação.

O padrão ZigBee emprega o IEEE 802.15.4 em sua Camada Física (PHY) e de Controle de Acesso ao Meio (MAC). Portanto, um dispositivo compatível com o padrão ZigBee também é compatível com o IEEE 802.15.4 (FARAHANI, 2011). Além disso, o padrão define as camadas de Rede e de Aplicação, incluindo serviços de segurança, conforme ilustrado na Figura 5.

### 2.2.1 Camada de Rede ZigBee

A camada de Rede ZigBee fornece dois tipos de serviços: serviço de dados e serviço de gerenciamento. O serviço de dados é responsável pela execução de alguns métodos que as camadas adjacentes utilizam para o encaminhamento e recebimento de pacotes de dados, como por exemplo, controle de retransmissões, controle do número de sequência, seleção do mecanismo de comunicação utilizado (*broadcast*, *multicast*, ou *unicast*), entre outros. O serviço de gerenciamento é responsável, principalmente, pelas seguintes tarefas: formação da rede,

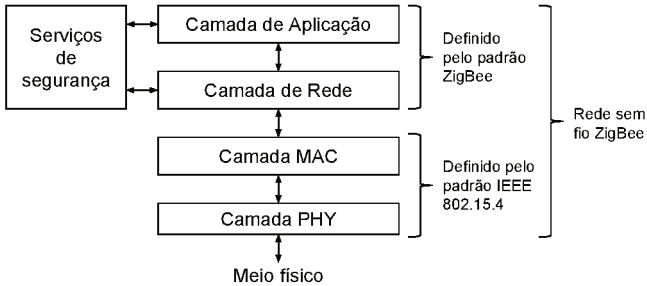


Figura 5: Arquitetura do padrão ZigBee.

descobrimiento de redes e vizinhos dentro da área de cobertura, associação e dissociação dos nodos, definição e manutenção do roteamento dos pacotes e configuração do superframe.

#### 2.2.1.1 Topologia em Agrupamentos em Árvore

Como uma extensão ao IEEE 802.15.4, a camada de Rede ZigBee define mecanismos para lidar com redes de topologia baseada em árvore. Conforme apresentado na Figura 6, uma rede em árvore começa a partir do coordenador PAN, que representa a raiz da rede. O coordenador PAN e os CHs podem assumir o papel de dispositivo pai, aceitando ou rejeitando pedidos de associação de outros dispositivos na rede. Aqueles nodos que são aceitos e se conectam ao seu pai, são chamados de dispositivos filhos. Todas as mensagens geradas pelos filhos só podem ser roteadas até a raiz (coordenador PAN) através dos pais.

A profundidade de rede é definida como o número de saltos mínimo requerido para que um pacote atinja a raiz da rede através das relações entre pai e filho (FARAHANI, 2011). O coordenador PAN possui profundidade de rede igual a 0, os nodos associados diretamente ao coordenador PAN possuem profundidade 1 e assim por diante.

Um dos mecanismos oferecidos pelo padrão ZigBee para redes em árvore é o de alocação de endereços, também conhecido como endereçamento distribuído padrão (SPECIFICATION, 2008). Nesse mecanismo, o coordenador PAN provê para cada pai, um subconjunto de endereços de rede para serem endereçados aos seus respectivos filhos.

A alocação de endereço é realizada com base nos seguintes parâmetros:



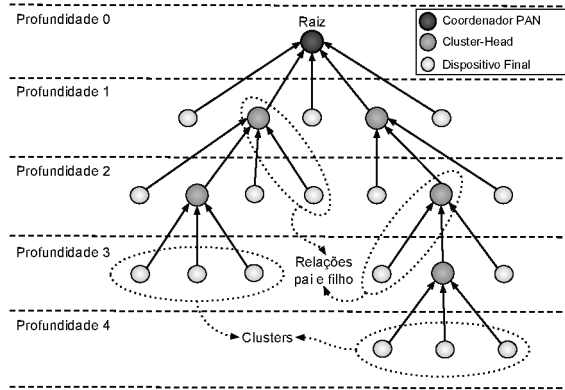


Figura 6: Exemplo de relações em uma rede em agrupamentos em árvore.

- $Lm$ : Máxima profundidade da rede.
- $Cm$ : Número máximo de filhos permitidos para cada pai.
- $Rm$ : Número máximo de CHs filhos permitidos para cada pai.
- $d$ : Profundidade do nodo na rede.

Os parâmetros  $Lm$ ,  $Cm$  e  $Rm$  podem ser definidos pelo usuário antes da inicialização da rede. O padrão ZigBee não define nenhum valor específico para esses parâmetros, pois são dependentes da aplicação.

#### 2.2.1.2 Roteamento

Roteamento é o processo de seleção do caminho sobre o qual as mensagens irão ser transmitidas para seu destino final (FARAHANI, 2011). O coordenador PAN e os CHs são os responsáveis por executar as tarefas de seleção e manutenção desses caminhos na rede.

Cada caminho na rede é constituído por uma série de relações entre os nodos. Cada relação é composta por dois nodos e está associada a um custo. Esse custo é normalmente associado a uma probabilidade de sucesso na entrega das mensagens, porém, podem haver outras maneiras de defini-lo.

O padrão ZigBee permite que diferentes métodos sejam utilizados para estimar a probabilidade de sucesso na entrega de mensagens,

porém, define que a estimativa inicial seja baseada na média da LQI, a qual indica a energia do sinal ou a relação sinal-ruído entre dois nodos. Sendo assim, quanto maior a probabilidade de sucesso, menor o custo da relação.

Quando diferentes caminhos são comparados para seleção, utiliza-se a soma dos custos de cada relação que formam o mesmo. O caminho com menor custo tem maiores chances de entregar as mensagens com sucesso e portanto é escolhido. Uma tabela de roteamento é mantida em cada coordenador para determinar as rotas que as mensagens devem tomar para um destino particular.

O mecanismo de roteamento para redes em árvore é chamado roteamento hierárquico. Nele, os nodos possuem relação somente com seu pai e/ou filho e as mensagens devem obedecer os caminhos definidos por essas relações.

### **2.2.2 Camada de Aplicação ZigBee**

A camada de aplicação ZigBee consiste em três partes: Subcamada de Suporte de Aplicação, Objetos de Dispositivos ZigBee, e Estrutura de Aplicação. Cada parte provê um conjunto de funcionalidades específicas.

A Subcamada de Suporte de Aplicação (SSA) é responsável por manter as tabelas utilizadas, pelo encaminhamento de mensagens entre dispositivos relacionados, pelo gerenciamento de endereços, pela conversão e mapeamento dos endereços IEEE e por suportar o transporte confiável de dados. Os Objetos de Dispositivos ZigBee utilizam a camada de rede e a SSA para implementar os dispositivos da rede, definindo seu papel, descobrir os outros dispositivos da rede, iniciar ou responder requisições e executar tarefas relacionadas a segurança. Por último, a Estrutura de Aplicação ZigBee é o ambiente em que cada objeto de aplicação é hospedado.

## **2.3 MODELO DE SIMULAÇÃO PARA REDES EM AGRUPAMENTOS EM ÁRVORE**

O conjunto formado pelos padrões IEEE 802.15.4/ZigBee é um dos mais utilizados na implementação de RSSF. Entretanto, devido a sua alta complexidade, o projeto desse tipo de rede é uma tarefa desafiadora, sendo custoso realizar testes e avaliações em cenários reais. Neste

contexto, para permitir uma avaliação experimental satisfatória, modelos de simulação são usualmente utilizados. Com base nisso, o presente estudo emprega o modelo de simulação CT-SIM (LEAO et al., 2017b). Esse modelo emprega o conjunto de padrões IEEE 802.15.4/ZigBee, fornecendo mecanismos importantes para avaliação de redes de larga escala em agrupamentos em árvore.

Baseado em um estudo sobre diversas ferramentas de simulação de RSSF, os autores do modelo CT-SIM selecionaram o simulador Castalia para implementação do mesmo, o qual é uma ferramenta de código aberto (*open-source*) e baseada em eventos discretos (BOULIS et al., 2011).

O Castalia foi desenvolvido no NICTA (*National ICT Australia Ltd*) para simulação de RSSF, Redes de Área do Corpo (BANs – *Body Area Networks*) e redes de dispositivos embarcados de baixa potência em geral. Esse simulador é baseado na plataforma OMNeT++, sendo amplamente utilizado por pesquisadores e desenvolvedores para testar seus protocolos.

Alguns motivos para sua popularidade é porque o Castalia implementa um modelo de canal sem fio avançado baseado em dados medidos empiricamente. Também provê modelos de rádios reais, comportamento realístico dos nodos e um modelo de consumo energético. Ademais, o Castalia é adaptativo e extensível através da linguagem de programação C++.

Basicamente, o modelo de simulação CT-SIM é constituído por três etapas, nomeadas como: formação da rede, comunicação, e coleta de estatísticas. Cada etapa é configurada conforme uma série de parâmetros, os quais podem ser definidos pelo usuário. A Figura 7 ilustra o esquema de implementação do modelo CT-SIM.

### 2.3.1 Formação da Rede

A etapa de formação da rede é responsável pela definição da topologia em agrupamentos em árvore e dos principais mecanismos utilizados na operação da rede. Essa etapa é executada sequencialmente através dos seguintes passos: implantação dos nodos, formação dos agrupamentos, esquema de endereçamento, escalonamento dos períodos ativos e configuração do superframe.

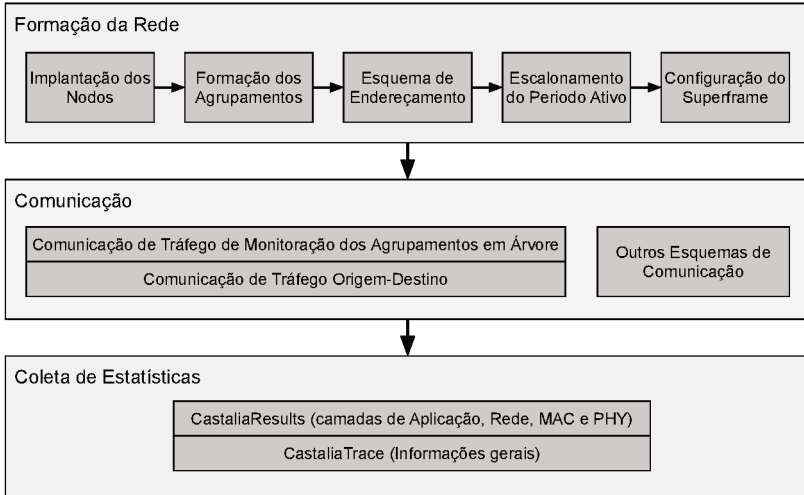


Figura 7: Esquema de implementação do modelo CT-SIM

### 2.3.1.1 Implantação dos Nodos

Na implantação dos nodos, são definidos o tamanho do cenário monitorado, o número de nodos sensores e as suas respectivas localizações no ambiente. Os parâmetros utilizados para implantação estão definidos na Tabela 1.

Tabela 1: Parâmetros para implantação dos nodos.

Parâmetros	Descrição
<i>numNodes</i>	Número de nodos sensores.
<i>deployment</i>	Tipo de implantação.
<i>xCoor</i>	Coordenada x para um nodo específico.
<i>yCoor</i>	Coordenada y para um nodo específico.
<i>field_x</i>	Tamanho do eixo x do cenário.
<i>field_y</i>	Tamanho do eixo y do cenário.

O Castalia fornece diversos tipos de implantação de nodos no cenário. Neste trabalho é empregada a implantação aleatória uniformemente distribuída.

### 2.3.1.2 Formação dos Agrupamentos

A formação dos agrupamentos é iniciada pelo coordenador PAN através da difusão de uma mensagem de formação, a qual contém o identificador (ID) do agrupamento e outras informações básicas. Os nodos que recebem a mensagem de formação enviam uma mensagem de requisição de associação ao coordenador utilizando o mecanismo CSMA-CA. As requisições são aceitas pelo coordenador somente se forem enviadas dentro de um limite de tempo predefinido, e se os nodos cumprirem os critérios exigidos.

Os autores do modelo (LEAO et al., 2017b), assim como o padrão ZigBee, utilizam um número máximo de filhos por agrupamento e a LQI como critério de associação. Caso o nodo seja aceito, uma mensagem de confirmação é enviada a ele.

Após a associação dos nodos, o coordenador PAN executa uma função de seleção de CHs. O principal objetivo dessa função é selecionar os CHs com base em critérios específicos. Basicamente, os autores utilizam o valor RSSI e o número máximo de CHs que podem ser selecionados por agrupamento como critérios de seleção.

Depois de serem selecionados, cada CH continua a formação da rede, executando as mesmas etapas para formar seus próprios agrupamentos. Após certo tempo, o processo de formação é finalizado com a árvore construída. A Tabela 2 descreve os parâmetros utilizados.

Tabela 2: Parâmetros para formação dos agrupamentos.

Parâmetros	Descrição
<i>formationTime</i>	Tempo de duração para requisições de associação.
<i>maxChildren</i>	Número máximo de filhos por agrupamento.
<i>maxCCH</i>	Número máximo de CHs por agrupamento.
<i>maxDepth</i>	Profundidade máxima da rede.

### 2.3.1.3 Esquema de Endereçamento

O esquema de endereçamento utilizado no modelo CT-SIM é baseado no esquema definido pelo padrão ZigBee apresentado na Seção 2.2.1, onde cada agrupamento possui um bloco de endereços e cada nodo sensor um endereço único. A diferença entre esses esquemas de endereçamento é que, no CT-SIM, somente após os agrupamentos serem formados, os CHs enviam uma requisição de endereço para o coordena-

dor PAN baseada no número de filhos associados. Depois de receber as requisições dos CHs da rede, o coordenador PAN especifica os blocos de endereços para cada agrupamento e os envia através de uma mensagem de resposta de endereçamento.

#### 2.3.1.4 Escalonamento dos Períodos Ativos

Após o processo de endereçamento, é realizado o escalonamento dos períodos ativos. Esse escalonamento emprega um esquema de divisão do tempo para evitar colisões de comunicação entre os agrupamentos. O coordenador PAN é responsável por executar essa função, definindo, para cada agrupamento, seus intervalos de transmissão dentro do seu BI. Dessa maneira, apenas um agrupamento permanece ativo por vez.

Além disso, o modelo prevê duas abordagens de escalonamento: *top-down* e *bottom-up*. Na primeira, o fluxo de mensagens é direcionado de cima pra baixo, ou seja, do coordenador PAN para os nodos. Na segunda, o fluxo é direcionado dos nodos para o coordenador PAN. A Figura 8 ilustra o esquema de escalonamento *bottom-up*, no qual as transmissões começam a partir dos agrupamentos mais profundos da rede.

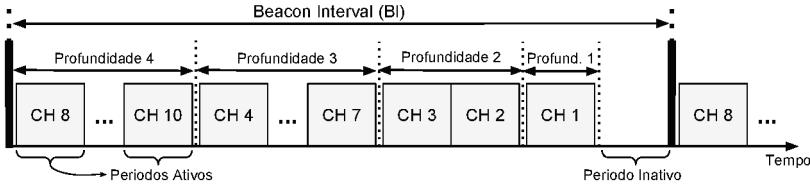


Figura 8: Esquema de escalonamento *Bottom-Up*.

Cada período ativo é representado pela estrutura do superframe. Sua duração pode ser definida individualmente, para cada agrupamento, através dos parâmetros BO e SO.

#### 2.3.1.5 Configuração do Superframe

Em um estudo realizado recentemente (LEÃO et al., 2016), os autores mostram que, para redes de topologia em agrupamentos em árvore, os padrões IEEE 802.15.4/ZigBee não definem mecanismos satisfató-

rios para alocação da duração do superframe (SD). Nesse sentido, os autores do modelo CT-SIM (LEAO et al., 2017b) definem três esquemas de alocação:

1. Esquema de alocação padrão: todos os agrupamentos possuem a mesma duração de superframe (SD) e os mesmos intervalos de *beacons*, sendo definidos pelo parâmetro BO (*beaconOrder*). Esse esquema segue as recomendações dos padrões IEEE 802.15.4/Zig-Bee.
2. Esquema de alocação *Proportional-Load*: cada agrupamento possui um SD proporcional à carga imposta sobre eles, considerando todos seus descendentes (filhos, filhos dos filhos e assim por diante). Para utilização desse esquema, a periodicidade de transmissão, de cada nodo, deve ser conhecida.
3. Esquema de alocação *Proportional-Node*: cada agrupamento possui um SD proporcional ao número de nodos descendentes. Esse esquema é utilizado em casos quando a periodicidade de transmissão de mensagens não é conhecida.

Considerando que, neste trabalho, a periodicidade de transmissão é conhecida, o esquema *Proportional-Load* é utilizado. Para empregar este esquema, um conjunto de restrições devem ser respeitadas para que a rede opere adequadamente (LEAO et al., 2017a).

Conforme as equações 2.1 e 2.2, que definem os parâmetros BI e SD respectivamente, o BI precisa ser suficientemente grande para garantir que todos as durações dos superframes ( $SD_j$ ), de cada CH  $j$ , possam ser escalonadas. Com base nisso, tem-se a seguinte equação:

$$BI \geq \sum_{j=1}^{N_{CH}} SD_j \quad (2.3)$$

Outra restrição está relacionada à periodicidade das mensagens. Considerando que cada nodo gera uma mensagem de dados com período  $P_i$ , e que as mensagens só possam ser transmitidas no período ativo do respectivo agrupamento, tem-se uma restrição direta do valor de BI sobre a periodicidade das mensagens. Mais especificamente, caso o valor de BI seja maior que o período de geração de mensagens de um determinado nodo, as mensagens geradas irão se acumular no *buffer*, pois entre um período ativo e outro, podem haver duas ou mais mensagens geradas. Portanto, considerando o conjunto  $\mathbb{P} = \{P_1, P_2, \dots, P_M\}$

como o período das mensagens de todos os  $M$  nodos, temos a seguinte restrição:

$$BI \leq \min\{\mathbb{P}\} - \delta, \quad (2.4)$$

onde  $\min\{\mathbb{P}\}$  é o menor período entre todos os nodos e  $\delta$  é uma constante que representa o tempo mínimo necessário para transmitir uma mensagem através do mecanismo CSMA-CA e receber uma confirmação. Com isso, as equações 2.3 e 2.4 podem ser resumidas da seguinte forma:

$$\sum_{j=1}^{N_{CH}} SD_j \leq BI \leq \min\{\mathbb{P}\} - \delta \quad (2.5)$$

Além de possuir essas restrições, a definição do valor de BI está relacionada diretamente aos atrasos das mensagens e ao consumo energético dos nodos. Como exemplo, considerando que as restrições impostas pela Equação 2.5 sejam satisfeitas, temos que, se o valor de BI definido é baixo, ou seja, mais próximo do valor da soma dos SDs (Equação 2.3) do que do mínimo período de geração de mensagens (Equação 2.4), teriam-se maiores gastos de energia e menores atrasos. Isso porque, a frequência dos períodos ativos seria maior e os nodos passariam mais tempo com seus rádios ativos. Por outro lado, as mensagens passariam menos tempo armazenadas esperando para serem enviadas.

Contudo, dependendo da aplicação, a definição do valor de BI deve ser feita com cuidado, pois é fundamental no funcionamento da rede. Os principais parâmetros utilizados para configuração do superframe são apresentados na Tabela 3.

Tabela 3: Parâmetros para configuração do superframe.

Parâmetros	Descrição
<i>superframeOrder</i>	valor de SO para definição de SD (Equação 2.2).
<i>beaconOrder</i>	valor de BO para definição de BI (Equação 2.1).

Importante ressaltar que as restrições apresentadas neste trabalho são para o escalonamento do tipo *bottom-up*, uma vez que o objetivo do mesmo está relacionado com monitoramento periódico do ambiente. Para ver com mais detalhes os esquemas apresentados nessa seção, consultar os trabalhos referentes (LEÃO et al., 2016; LEO et al., 2017a, 2017b).



### 2.3.2 Comunicação

Conforme as definições do padrão IEEE 802.15.4, o modelo CT-SIM implementa dois tipos de comunicação: direta e indireta. Na comunicação direta, na topologia em agrupamentos e árvore, os nodos transmitem periodicamente os dados coletados para seu pai. Na comunicação indireta, os nodos podem requisitar alguma informação através de uma mensagem de requisição específica, ou também, o coordenador pode indicar mensagens pendentes através do endereço de um determinado nodo em seu *beacon*, as quais serão requisitadas por eles ao receber esse *beacon*.

A partir desses tipos de comunicação, os autores do modelo CT-SIM implementaram dois tipos de tráfego: tráfego de monitoramento e tráfego de origem para destino (*source-to-destination*).

O tráfego de monitoramento é caracterizado pela transmissão periódica de dados gerados pelos nodos e encaminhados para o coordenador PAN através dos mecanismos de comunicação direta. Esse tipo de comunicação costuma ser denominada *convergecast*, por ser oposta à comunicação *broadcast*, na qual o coordenador envia mensagens para todos os nodos da rede. A taxa de geração desses dados, para cada nodo, é definida pelo usuário na configuração da simulação.

O tráfego de origem para destino é caracterizado por um tráfego também periódico de mensagens, porém, o destino dessas mensagens pode ser qualquer nodo da rede. Por esse motivo, esse tráfego emprega os mecanismos de comunicação direta e indireta.

Uma vez que o objetivo do presente trabalho é avaliar técnicas de formação de rede, por motivos de simplicidade, apenas o tráfego de monitoramento é utilizado. A Tabela 4 apresenta os principais parâmetros necessários para sua configuração.

Tabela 4: Parâmetros para configuração do tráfego de monitoramento.

Parâmetros	Descrição
<i>macMinBE</i>	mínimo expoente de <i>backoff</i> do CSMA-CA.
<i>macMaxBE</i>	máximo expoente de <i>backoff</i> do CSMA-CA.
<i>macMaxCSMABackoffs</i>	número máximo de <i>backoffs</i> do CSMA-CA.
<i>macMaxFrameRetries</i>	número máximo de tentativas de envio.
<i>packet_rate</i>	taxa de geração de pacotes (pacotes/segundo).

### 2.3.3 Coleta de Estatísticas

As informações geradas pelas simulações são de grande importância para análise do comportamento da rede. Essa análise possibilita que protocolos complexos sejam projetados de maneira eficiente e com baixo custo.

O modelo CT-SIM utiliza os recursos disponibilizados pelo simulador Castalia para coleta e análise das informações geradas nas simulações. Um desses recursos, chamado de *CastaliaResults*, é um *script* responsável pela leitura do arquivo de resultados e pela apresentação estruturada dos mesmos.

Além disso, o Castalia fornece mecanismos para que o projetista implemente funções que colem dados relevantes sobre a simulação. Esses dados são exibidos posteriormente em um arquivo de resultados gerado. Entre as principais informações coletadas, estão: atrasos fim-a-fim, taxa de perda de pacotes, taxa de sucesso, número de pacotes gerados, energia consumida, número de agrupamentos formados, entre outros.

Outra forma de analisar o comportamento da rede é através do arquivo de texto *CastaliaTrace*. Esse arquivo contém registros de todos os eventos que ocorrem durante a simulação. O usuário é livre para selecionar quais registros deseja salvar no arquivo. Entre esses eventos, estão: recebimento e encaminhamento de mensagens, ativação e desativação do transceptor de rádio, associação e dissociação de nodos, entre outros.

Para mais informações sobre os mecanismos de coleta de dados disponibilizados pelo simulador Castalia, pode-se consultar o manual (BOULIS et al., 2011).

## 2.4 CONCLUSÕES DO CAPÍTULO

Este capítulo apresentou o padrão IEEE 802.15.4, o qual especifica a camada Física e a subcamada de Controle de Acesso ao Meio para redes LR-WPAN. Também, foram descritas suas principais funcionalidades, assim como a estrutura do superframe e o mecanismo CSMA-CA utilizados para comunicação. Também foi apresentado o padrão Zig-Bee, que complementa o IEEE 802.15.4, definindo as camadas de Rede e de Aplicação. Além disso, foram expostos os principais conceitos da topologia em agrupamentos em árvore. Por fim, o modelo CT-SIM foi descrito, onde foram apresentados os conceitos básicos do simula-

dor Castalia, assim como os mecanismos empregados para formação da rede, comunicação e análise dos resultados.

Os fundamentos teóricos expostos neste capítulo constituem a base para a implementação e entendimento deste trabalho. Nos próximos capítulos serão apresentados diversos métodos que empregam esses conceitos.



### 3 TRABALHOS RELACIONADOS

Neste capítulo são apresentados os trabalhos relacionados à formação de topologia para monitoração periódica de ambientes. Conforme previamente mencionado, a formação de topologia é uma etapa fundamental para uma operação adequada da rede, porém, também é um grande desafio.

Entre os trabalhos relacionados à formação de topologia, podemos destacar aqueles que tratam sobre protocolos hierárquicos, que utilizam esquemas baseados em agrupamentos ou que focam apenas na formação de árvores. Nesse sentido, ambas abordagens utilizam critérios específicos de formação para melhorar o desempenho da rede.

Quando existe a formação de agrupamentos, pode-se dividi-la em duas etapas: seleção de *cluster-heads* (CH) e associação dos filhos. Ambas etapas podem assumir diferentes critérios para formação, como, por exemplo, qualidade do sinal (LQI e RSSI), energia residual, localização, vizinhança, entre outros.

No caso da abordagem não baseada em agrupamentos, é comum a utilização do conceito de custos nas relações de comunicação entre os nodos, podendo ser definidos com base em diversos critérios.

Além da formação da rede, esquemas de comunicação eficientes também são fundamentais para um bom funcionamento da rede. Neste contexto, muitos trabalhos utilizam o padrão IEEE 802.15.4 como base. Entretanto, existem trabalhos que propõem alterações no padrão, com objetivo de melhorar o seu desempenho.

#### 3.1 FORMAÇÃO DE TOPOLOGIA BASEADA EM AGRUPAMENTOS

O uso de topologias baseadas em agrupamentos é uma maneira eficiente de prover cobertura da rede, permitir maior escalabilidade e reduzir o consumo energético, entre outras vantagens. Devido a isso, diversos protocolos foram propostos, entre eles, temos:

- **Younis e Fahmy (2004):** O trabalho proposto implementa um protocolo híbrido e distribuído de formação de agrupamentos chamado HEED (*Hybrid Energy-Efficient Distributed*), cujo principal objetivo é formar uma rede de forma a minimizar o consumo de energia da rede. Para isso, o protocolo faz uso de dois parâmetros. Primeiramente, utiliza-se como base, a energia residual

dos nodos para a seleção dos *cluster-heads* (CHs). Logo após, considera-se um cálculo de custo de comunicação interna do agrupamento, que pode ser uma função relacionada à proximidade ou ao número de vizinhos.

Basicamente, o primeiro e o segundo parâmetro são utilizados respectivamente para uma seleção probabilística de um conjunto inicial de CHs e para que os nodos escolham apenas um CH para se associar, já que podem estar dentro do alcance de mais de um candidato a CH. O alcance de cada nodo varia de acordo com o nível de potência de rádio utilizado.

No caso de múltiplos candidatos a CH, será selecionado aquele que possui o menor custo de comunicação *intrachuster*. Esse custo é uma função das propriedades do agrupamento, como o tamanho, e se o mecanismo de variação de potência é permitido ou não dentro dos mesmos.

- **Heinzelman et al. (2000):** Nesse trabalho é proposto o protocolo LEACH (*Low-Energy Adaptive Clustering Hierarchy*), um dos protocolos mais conhecidos para formação de topologia baseada em agrupamentos. Esse protocolo é distribuído e visa minimizar o uso de energia global da rede através da distribuição de carga entre todos os nodos da rede em diferentes pontos no tempo.

Para isso, o protocolo implementa os seguintes mecanismos: coordenação e controle local de configuração e operação dos agrupamentos; rotação aleatória dos CHs baseada em probabilidade; e agregação ou compressão de dados locais.

A operação do LEACH é dividida em *rounds*, inicializados em uma fase de ajuste, quando os agrupamentos são formados, seguidos de uma fase estável, quando ocorre a comunicação e os dados são transferidos para a estação base. Quando os agrupamentos estão sendo formados, cada nodo decide se será ou não um CH naquele *round* específico. Essa decisão é baseada em um valor de probabilidade pré-definido e no número de vezes que o nodo já foi CH, buscando não sobrecarregar nenhum nodo.

Cada nodo que se elege CH no *round* corrente difunde uma mensagem de anúncio para os seus vizinhos. Os nodos que recebem essas mensagens armazenam as informações de força de sinal recebido e a identificação do CH para decidir a qual agrupamento irão se juntar. Após essa decisão, os nodos informam o CH esco-

lhido através de um pedido de associação. Todos os CHs possuem comunicação direta com a estação base e, portanto, os nodos associados estão a dois saltos da mesma.

Baseado no número de nodos no seu agrupamento, o CH cria um escalonamento TDMA e informa a cada nodo quando eles poderão transmitir suas mensagens. Em seguida, uma vez que os agrupamentos e o escalonamento TDMA estão fixados, a transmissão de dados pode começar. Essas transmissões utilizam a menor quantidade de energia possível, e os nodos que não são CH podem desligar seus rádios até a chegada de seu tempo de transmissão alocado pelo TDMA. Durante as transmissões, os CHs devem permanecer com seus rádios ligados para receber todos os dados de seu agrupamento. Após o recebimento, eles executam processamento local, se necessário, e enviam seus dados para estação base. Para não ocorrerem colisões entre agrupamentos na fase de comunicação, os autores consideram a utilização de um esquema de comunicação baseado em Acesso Múltiplo por Divisão de Código (CDMA – *Code Division Multiple Access*), no qual os pacotes são codificados para cada agrupamento.

Apesar das vantagens do protocolo LEACH, ele não oferece nenhuma garantia de qualidade na escolha dos CHs, uma vez que essa escolha é feita aleatoriamente. Com isso, os mesmos autores propuseram anos mais tarde (2002) o protocolo LEACH centralizado (LEACH-C). Nesse protocolo, um algoritmo central localizado na estação base (BS - *Base Station*) escolhe os CHs e forma os agrupamentos de forma melhor distribuída pela rede.

Durante a formação, os nodos informam sua localização e seu nível de energia à BS. Com base nisso, calcula-se a média de energia, descartando a possibilidade dos nodos abaixo da média serem CHs. Além disso, o protocolo conta com o algoritmo de otimização *simulated annealing* para resolver o problema NP-difícil de encontrar o número de agrupamentos ótimo com base nas informações recebidas dos nodos. Após a formação, o protocolo LEACH-C opera da mesma maneira que o protocolo LEACH.

Experimentos mostraram que, devido ao LEACH-C dispor do conhecimento global para formação da rede, ele apresenta maiores taxas de sucesso e menor consumo energético que o LEACH.

- **Ding et al. (2016):** O trabalho apresenta a proposta do algoritmo de formação RTBDG (*Real-Time Big Data Gathering*).

A proposta possui o objetivo de reduzir o consumo energético da rede, visando prolongar seu tempo de vida. Para isso, o algoritmo constrói agrupamentos e define o roteamento da rede com base na energia residual de cada nodo e na força do sinal entre eles (RSSI).

Assumindo que a estação base tem um alcance de transmissão capaz de enviar uma mensagem para qualquer nodo da rede, ela inicializa a formação através de uma mensagem contendo informações de configuração e os TDMA's de cada nodo. Quando um nodo recebe a mensagem, ele obtém o valor RSSI utilizado para estimar a qualidade do sinal entre ele e a estação base. Caso essa estimativa obedeça certa condição, o nodo é capacitado para se comunicar diretamente a BS. Além disso, todos nodos calculam uma estimativa de se tornar CH com base na sua energia residual e valor RSSI com a estação base.

Logo após, os nodos difundem essas estimativas para seus vizinhos. Aqueles que tiverem a maior estimativa da sua vizinhança se tornam CHs. Por fim, os nodos que restaram se juntam ao CH que possui a maior estimativa recebida.

O algoritmo RTBDG foi desenvolvido para analisar os riscos presentes em ambientes industriais. Assim, os nodos coletam os dados de forma regular, armazenando e verificando seus valores. Caso os dados coletados estejam dentro de valores considerados normais são apenas armazenados. Caso os valores não sejam considerados normais, ou os nodos cheguem ao limite da capacidade de armazenamento, os dados são transmitidos para a estação base.

- **Bholowalia e Kumar (2014):** Os autores propuseram o protocolo *EBK-means*, o qual utiliza o algoritmo de agrupamento K-means para formação dos agrupamentos de nodos. O algoritmo é aplicado sobre as coordenadas X e Y dos nodos dentro do cenário. O método *Elbow* é empregado para definir adequadamente o número  $k$  de agrupamentos na rede.

A proposta é utilizar o K-means para formação da rede e o protocolo LEACH para o roteamento, onde o primeiro é usado para seleção dos CHs e, logo em seguida, o segundo assume o envio das mensagens para o coordenador PAN. Assim que o valor de  $k$  é definido, os CHs são determinados com base na proximidade das suas localizações e das médias do K-means. Os nodos restantes escolhem os CHs que possuem um menor custo de comunicação



(RSSI). Após a primeira seleção, o protocolo LEACH assume, fazendo com que os CHs mudem dinamicamente a fim de distribuir a energia.

Nesse trabalho, os autores optam por executar os algoritmos na BS, de forma centralizada. Isso porque, além da BS ser conectada diretamente a uma fonte de energia ilimitada, ela possui capacidade computacional para executar esses algoritmos em menos tempo.

- **Muruganathan et al. (2005):** Os autores propõem o método centralizado BCDCP (*Base Station Controlled Dynamic Clustering Protocol*), no qual a BS é responsável pela execução de todas as tarefas de formação dos agrupamentos e de roteamento. Os autores justificam o emprego de uma abordagem centralizada como maneira de economizar energia e diminuir custos, pois com todas as atividades sendo realizadas pela BS, os nodos podem ser simples e de baixo custo.

Durante a formação, a BS recebe informações sobre o estado de energia de todos os nodos. Com base nisso, a média de energia dos nodos é calculada e um conjunto  $S$  é formado com aqueles acima dela. Logo após, com base na localização dos nodos, um algoritmo de separação é executado sobre o conjunto  $S$  para definição dos agrupamentos e seus respectivos CHs. Essa definição tem objetivo de balancear o consumo de energia através da formação de agrupamentos de mesmo tamanho e uniformemente distribuídos pelo cenário.

Depois de formar os agrupamentos, a BS define o roteamento através de uma árvore geradora mínima, visando minimizar o consumo de energia da rede.

### 3.2 FORMAÇÃO DE TOPOLOGIA BASEADA EM ÁRVORES

Uma maneira eficiente de definir o roteamento entre os nodos ou até mesmo entre agrupamentos é através de topologias em árvore. Para isso, é comum a utilização dos conceitos de grafos, árvores geradoras mínimas (MST – *Minimum Spanning Tree*) e árvores de menor caminho (SPT – *Shortest Path Tree*). Nesse contexto, muitos protocolos foram propostos, entre eles, estão:

- **Gagarin et al. (2010):** O trabalho propõe um método distribuído de roteamento baseado na utilização de MSTs sobre um

grafo ponderado não-direcional. O objetivo principal do trabalho é encontrar uma topologia que minimize a energia consumida entre todos os nodos da rede.

A fim de satisfazer essas condições, considera-se um subconjunto de MSTs criadas a partir de um grafo  $G$ , onde cada aresta de  $G$ , que representa uma relação entre dois nodos, é ponderada pela energia de transmissão consumida pelo nodo de origem. Com isso, o método visa encontrar uma MST que minimize a energia máxima consumida pela rede. Além disso, o método também tenta reduzir a profundidade da rede.

Para atingir seus objetivos, o algoritmo implementado pelos autores utiliza uma combinação dos métodos *Depth-first*(DF) e *Breadth-first*(BF). Primeiramente, divide-se a matriz de adjacência que representa o grafo  $G$  em  $k$  agrupamentos, nos quais aplica-se o algoritmo BF para formação das árvores geradoras. Logo em seguida, os agrupamentos mais próximos são unidos, formando caminhos até a BS.

- **Huang et al. (2006):** Nesse trabalho, os autores propõem um protocolo chamado *Dynamic Minimal Spanning Tree Routing Protocol* (DMSTRP). A principal ideia deste método é utilizar o conceito de MSTs para comunicação dentro dos agrupamentos.

Primeiramente, o protocolo DMSTRP constrói agrupamentos de forma similar ao BCDCP (MURUGANATHAN et al., 2005), no qual, através das informações sobre o estado de energia e localização de todos os nodos, a BS define os agrupamentos e seus respectivos CHs de forma energeticamente balanceada.

A principal diferença desse algoritmo está na definição da comunicação que ocorre dentro dos agrupamentos (*intracluster*). Enquanto que a maioria dos protocolos utiliza uma abordagem de um salto entre o CH e seus filhos, o DMSTRP constrói uma MST para conectar todos os nodos com base na distância entre eles.

Os autores comparam seu protocolo com o BCDCP e o LEACH em relação ao tempo de vida da rede e ao atraso que os CHs sofrem para coletar as mensagens de seus filhos.

- **Bechkit et al. (2012):** Os autores propõem um método baseado na abordagem SPT, na qual todos os nodos transmitem seus dados para um nodo coletor (*sink*). Desse maneira, o principal objetivo do trabalho é construir uma árvore com raiz no *sink*, no

qual os custos dos caminhos de qualquer nodo até a raiz sejam mínimos.

Para isso, os autores utilizam uma abordagem distribuída baseada no algoritmo de *Bellman-Ford* (BF). Além disso, um novo esquema de definição dos custos considerando a profundidade dos nodos é proposto. Esse esquema visa atribuir pesos diferentes as relações mais próximas do *sink*, pois elas são mais sobrecarregadas com as tarefas de roteamento.

Os autores não assumem um critério específico para atribuição dos custos. Entretanto, o método apresentado permite o uso de apenas um critério para formação, podendo variar de acordo com os objetivos da rede.

- **Lachowski et al. (2015):** O método EBF (*Efficient Bellman-Ford*) possui objetivo de formar agrupamentos em árvore de maneira eficiente baseado no algoritmo *Bellman-Ford* Distribuído (DBF). O algoritmo DBF funciona de forma iterativa para que os nodos obtenham os custos das relações entre eles até o coordenador PAN e assim gerar a árvore geradora de menor caminho. Entretanto, esse algoritmo não é adequado para uso em RSSF devido ao alto número de mensagens trocadas e ao desconhecimento do término de sua execução.

Neste contexto, os autores propõem algumas alterações para redução no número de transmissões de mensagens. Porém, conforme esse número é reduzido, a qualidade da árvore gerada cai. Além de não definir os custos que devem ser utilizados nas relações, o trabalho não avalia o funcionamento da rede em relação ao desempenho operacional, apenas ao de formação, tais como: tempo total de formação, número de mensagens transmitidas durante a formação e profundidade média da árvore gerada.

### 3.3 ESQUEMAS DE COMUNICAÇÃO BASEADOS NO PADRÃO IEEE 802.15.4

Conforme descrito anteriormente, além da formação de topologia, são necessários esquemas de comunicação eficientes que possibilitem o funcionamento adequado da rede. Para isso, muitos trabalhos utilizam o padrão IEEE 802.15.4 como base.

Embora o padrão disponibilize mecanismos que contribuam para o desempenho da rede, existem alguns pontos fracos que podem ser

melhorados (KHANAFER et al., 2014). Nesse sentido, os seguintes trabalhos são apresentados:

- **Alberola e Pesch (2012):** Esse trabalho propõe o algoritmo DCLA (*Duty Cycle Learning Algorithm*), o qual define uma configuração para que os nodos alcancem um desempenho ótimo em diferentes condições. O algoritmo é executado nos nodos coordenadores que, através de estatísticas coletadas durante a execução, estimam os tráfegos de entrada e definem os *duty cycles* de forma automática. Sua principal vantagem é a capacidade de auto configuração sem intervenção manual.
- **Gilani et al. (2013):** Os autores introduzem um protocolo híbrido adaptado para melhorar a vazão e o consumo energético do IEEE 802.15.4. O protocolo emprega o conceito de TDMA no CAP do superframe, dividindo o CAP em CSMA-CA e TDMA com base nas informações das filas dos nodos e no nível de colisões. A principal vantagem de utilizar TDMA no CAP é a redução no número de nodos que participam da contenção, diminuindo as colisões e economizando bateria. Além de infringir o padrão IEEE 802.15.4 com respeito à especificação do *superframe*, a principal desvantagem é que os nodos precisam esperar o tempo definido pelo TDMA, implicando em atrasos maiores.
- **Anastasi et al. (2011):** O trabalho proposto tem como objetivo encontrar a melhor configuração para o padrão IEEE 802.15.4. Os autores argumentam que o protocolo MAC do padrão sofre com problemas de confiabilidade, pois seu desempenho é fraco em relação à taxa de sucesso na entrega das mensagens. Com isso, diversas simulações empregando configurações diferentes para os parâmetros do mecanismo CSMA-CA são realizadas. O trabalho conclui que o desempenho do padrão pode ser melhorado com o ajuste adequado dos parâmetros. A principal vantagem deste método é a simplicidade, pois não são necessárias alterações no padrão IEEE 802.15.4.

### 3.4 COMPARAÇÃO ENTRE TRABALHOS RELACIONADOS

Conforme apresentado, existem diversos trabalhos que abordam protocolos de formação de topologia e de comunicação. Os protocolos de formação visam a construção de topologias baseadas em árvores, agrupamentos ou em ambos. Dessa maneira, cada trabalho possui uma

série de características específicas, que podem divergir dependendo dos objetivos. Neste contexto, a Tabela 5 apresenta uma comparação entre os principais trabalhos que empregam métodos para formação de topologia. Para isso, cada trabalho é classificado quanto às seguintes categorias:

- **Modelo de formação:** a formação pode acontecer de forma centralizada ou distribuída, dependendo se a definição da topologia é realizada somente pela estação base ou se é realizada distribuídamente pelos nodos da rede, respectivamente.
- **Abordagem base:** a topologia pode ser formada buscando a formação de agrupamentos ou árvores, podendo também empregar ambos os conceitos.
- **Critério(s) de formação:** dependendo dos objetivos, um ou mais critérios podem ser utilizados para formação.
- **Métrica(s) de avaliação:** para avaliação dos métodos de formação, diferentes métricas podem ser utilizadas.

Conforme apresentado na Tabela 5, os trabalhos que empregam um modelo centralizado para formação da rede, normalmente, consideram que a BS possui conhecimento global da mesma, como por exemplo, a localização de todos nodos. Porém, dependendo da aplicação, formar a rede de maneira centralizada, assim como obter o conhecimento global da mesma, pode ser inviável, principalmente em aplicações de larga escala. Por isso, muitos trabalhos empregam um modelo de formação distribuído.

Em relação a abordagem base, enquanto alguns trabalhos se baseiam apenas na formação de árvores e outros na formação de agrupamentos, existem aqueles que empregam as duas abordagens. Para formação de árvores, geralmente, são utilizados métodos bem conhecidos, como por exemplo, MST e SPT. A formação de agrupamentos é frequentemente dividida em duas etapas, associação de filhos e seleção de CHs. Além disso, uma combinação das duas abordagens também pode ser realizada.

Como se pode observar na Tabela 5, diversos critérios de formação e métricas de avaliação podem ser utilizados. Na maioria dos trabalhos, emprega-se critérios e métricas relacionados a energia, porém, dependendo da aplicação, outras também são consideradas.

Tabela 5: Comparação dos protocolos de formação de topologia

Protocolo	Modelo de formação		Abordagem base		Critério(s) de formação	Métrica(s) de avaliação
	Centralizado	Distribuído	Árvores	Agrupamentos		
HEED		✓		✓	-Energia residual -Vizinhança	-Consumo de energia -Número de agrupamentos -Tempo de formação -Tempo de vida
LEACH		✓		✓	-Probabilidade -Valor RSSI	-Consumo de energia -Tempo de vida
LEACH-C	✓			✓	-Energia Residual -Localização -Energia Residual -Localização	-Consumo de energia -Taxa de sucesso -Tempo de vida
RTBDG		✓		✓	-Energia residual -Valor RSSI -Energia residual -Valor RSSI	-Consumo de energia -Tempo de vida
EBK-means	✓			✓	-Localização -Valor RSSI	-Tempo de formação -Tempo de vida
BSDCP	✓		✓	✓	-Energia residual -Localização	-Consumo de energia -Tempo de vida
Gagarin et al. (2010)		✓	✓		-Localização -Profundidade	-Consumo de energia -Peso total da árvore
DMSTRP	✓		✓	✓	-Energia residual -Localização	-Tempo de vida -Atraso fim-a-fim
EBF			✓	✓	-NA	-Número de mensagens transmitidas para formação -Tempo de formação -Profundidade média da árvore gerada
<b>Estudo</b>	✓		✓		-Localização -Profundidade	-Atraso fim-afim -Consumo de energia -Número de agrupamentos
<b>Proposta</b>		✓	✓	✓	-Valor RSSI -Vizinhança	-Profundidade média -Taxa de sucesso -Tempo de vida

Também são apresentados na Tabela 5, um estudo e uma proposta, os quais fazem parte do presente trabalho. Resumidamente, o estudo assume um modelo de formação simples e centralizado, no qual a BS (neste caso coordenador PAN) possui o conhecimento global sobre a localização dos nodos. Por outro lado, a proposta emprega um modelo de formação distribuído, onde nenhum tipo de conhecimento global é assumido. Enquanto diferentes critérios de formação são empregados, ambos são avaliados em relação às mesmas métricas.

### 3.5 CONCLUSÕES DO CAPÍTULO

A partir do estudo realizado, foi possível notar que a maioria dos trabalhos encontrados visam reduzir o consumo energético da rede. Embora o consumo de energia seja um dos principais desafios para RSSF, existem diversos outros fatores que são importantes, como por exemplo, taxa de sucesso, atraso fim-a-fim, cobertura, entre outros. Nesse sentido, poucos trabalhos abrangem esses fatores em seus objetivos.

Alguns protocolos possuem objetivo de construir a topologia de forma eficiente, rápida e com menor número de mensagens possível, por exemplo. Entretanto, apesar de avaliar seus métodos em relação a sua rapidez ou número de mensagens necessárias para formação, esses trabalhos não avaliam o desempenho do monitoramento da rede.

Em relação aos esquemas de comunicação baseados no padrão IEEE 802.15.4, os trabalhos de Alberola e Pesch (2012) e Gilani et al. (2013) apresentam esquemas automáticos, que não precisam de intervenção humana para configuração dos parâmetros utilizados pelo padrão. De forma mais simples, o trabalho de Anastasi et al. (2011) propõe que sejam realizadas simulações, com diferentes configurações, antes da inicialização da rede, para definição da configuração mais adequada. Embora este último não seja tão eficiente, ele oferece uma forma mais simples de melhorar o desempenho do padrão IEEE 802.15.4.

Nesse sentido, com objetivo de obter uma rede que opere de forma eficiente em relação a diversas métricas, um estudo e uma proposta são realizados e apresentados nos próximos capítulos. Enquanto o estudo visa a compreensão dos principais fatores de formação e operação da rede, a proposta se baseia na compreensão obtida para definir uma arquitetura de formação eficiente.





## 4 ESQUEMAS DE FORMAÇÃO DE TOPOLOGIAS EM ÁRVORES

Uma Rede de Sensores Sem Fio geralmente é composta por um grande número de nodos sensores, densamente distribuídos em um ambiente, com objetivo de monitorar um ou mais fenômenos de interesse (AKYILDIZ et al., 2002). A interação entre os nodos, assim como seu funcionamento e disposição no cenário constituem um sistema complexo. Dessa forma, o emprego de abstrações computacionais e conceitos matemáticos, como por exemplo, árvores e grafos, são fundamentais para uma boa implementação e compreensão desse sistema.

Nesse contexto, o presente capítulo apresenta um esquema de formação de topologia baseado no conceito de árvores geradoras mínimas e árvores de menor caminho. Conforme apresentado na Seção 3.1, é comum o uso desses tipos de métodos para formação de topologia, pois apresentam uma maneira eficiente de economizar energia e prover escalabilidade à rede (SARAVANAN; MADHESWARAN, 2016; MURUGANATHAN et al., 2005; LACHOWSKI et al., 2015).

Além de propor esquemas de formação, este capítulo também possui objetivo de apresentar um estudo sobre os principais fatores que, de alguma forma, contribuem ou prejudicam o desempenho de uma RSSF. Para tal estudo, diferentes algoritmos e critérios de formação são utilizados, diversas simulações são conduzidas e, por fim, resultados são analisados.

O capítulo começa com a definição do modelo do sistema, onde são expostos os principais conceitos e fundamentos utilizados na proposta. Em seguida, os métodos empregados para formação da topologia são apresentados. Após a descrição dos métodos, são apresentadas as simulações e os resultados obtidos. Por fim, são feitas algumas considerações finais.

### 4.1 MODELO DO SISTEMA

Nesse modelo assume-se que o coordenador PAN possui conhecimento global sobre a localização de todos os nodos da rede, além de informações sobre os enlaces que os interligam. Com base nisso, a partir dessas informações, a topologia é definida de forma centralizada.

As árvores são obtidas a partir de um grafo ponderado  $G = (V, E)$ , onde  $V$  é um conjunto finito de vértices que representam os

nodos e  $E$  é um conjunto de arestas de  $G$ , que representa os enlaces. Cada aresta representa uma relação binária ( $E \subset V \times V$ ) entre os nodos (CORMEN et al., 2009). Ainda, assume-se que cada aresta  $(u, v) \in E$  possui associado um peso  $w(u, v)$  que especifica o custo da relação entre  $u$  e  $v$ .

Essa ponderação dos custos das arestas de  $G$  pode ser realizada através de diversos critérios distintos, ou uma combinação deles. Esses critérios podem ser relacionados à distância entre os nodos, ao número de saltos até a raiz (coordenador PAN), ao número de vizinhos, à energia consumida, entre outros.

A tentativa de encontrar uma árvore que minimize os custos das relações é considerado um problema NP-difícil (SARAVANAN; MADHESWARAN, 2016). Quando é utilizada uma combinação de mais de um critério para definição dos custos das relações em  $E$ , é comum o emprego de técnicas de otimização (SARAVANAN; MADHESWARAN, 2016), as quais estão fora do escopo deste trabalho. Dessa forma, por razões de simplicidade, os esquemas apresentados nesse capítulo consideram um único critério para ponderação do grafo  $G$ .

Além disso, o simulador Castalia oferece diversas opções de configuração para definição do modelo de comunicação. Entre essas opções, para simplificar a análise dos métodos, este trabalho considera os modelos de disco unitário (*unit disc model*) e de interferência aditiva. O primeiro modelo considera que todas as transmissões realizadas dentro de uma determinada distância ocorram perfeitamente, exceto em caso de interferência. O segundo modelo adotado assume que, se houver duas transmissões simultâneas dentro de um mesmo alcance, existem duas possibilidades: (1) ocorre colisão entre os pacotes e nenhum deles é recebido ou (2) o pacote enviado pelo transmissor de sinal mais forte é recebido, caso seja forte o suficiente.

#### 4.1.1 Principais Pressupostos

Este trabalho assume um sistema formado por um conjunto de nodos, organizados em uma árvore e dispostos conforme uma distribuição aleatória uniforme sobre uma área retangular. Além disso, foram assumidos os seguintes pressupostos para o modelo:

- (i) A área monitorada possui dimensões de tal forma que os dados transmitidos pelos nodos, geralmente, não alcançam diretamente o coordenador PAN com apenas um salto.
- (ii) Os nodos são estáticos, possuem as mesmas capacidades de pro-

cessamento e comunicação, utilizam a mesma potência de transmissão e são cientes de suas localizações.

- (iii) A aplicação tem um padrão de comunicação periódico, isto é, os nodos coletam dados periodicamente do ambiente, transmitem aos seus pais que, por sua vez, retransmitem com destino ao coordenador PAN.
- (iv) Os canais de comunicação são simétricos e têm sua qualidade proporcional à distância entre os nodos.
- (v) Existe um número suficiente de nodos no momento da implantação da rede, de forma que nenhum nodo fique desconectado da árvore quando é aplicado um método de formação adequado.
- (vi) Os dados não sofrem nenhum tipo de agregação, codificação ou fusão durante as suas transmissões.

As propriedades (i), (ii) e (iii) são usuais em diversas aplicações de RSSF, onde dispositivos de baixo custo são implantados de forma aleatória, o nível de potência dos rádios é predefinido igualmente, a localização é obtida através de GPS ou estimativas, têm seus dados coletados periodicamente, e os transmitem ao coordenador PAN através de múltiplos saltos. Além disso, o fato dos nodos serem homogêneos faz com que o método de formação de topologia seja mais flexível, possibilitando que qualquer nodo possa assumir o papel de pai.

A propriedade (iv) é assumida para simplificar a formação da topologia, pois uma vez que o coordenador conhece as localizações dos nodos, ele é capaz de definir as relações e o roteamento adequado para eles.

A propriedade (v) garante a cobertura da rede e uma conectividade adequada entre os nodos. Por fim, o item (vi) se refere a questões que, por simplicidade, não são assumidas no modelo, mas que, para tornar a aplicação de monitoramento mais eficiente, poderiam ser adicionadas posteriormente em trabalhos futuros.

É importante mencionar que a reconfiguração da topologia tem como um dos objetivos aumentar o tempo de vida da rede, uma vez que os nodos escolhidos como retransmissores situados perto do coordenador consomem mais energia do que os outros. Portanto, se não houver rotação desses nodos, as baterias dos mesmos acabarão mais cedo. Apesar disso, o presente estudo visa a redução do consumo energético global da rede, e por esse motivo, assume-se que a presença de um mecanismo de rotação não interfere na avaliação do mesmo.

## 4.2 DESCRIÇÃO DO MÉTODO

Nesta seção, são descritos os principais fundamentos e algoritmos utilizados para formação da topologia. Primeiramente, para definir as relações entre os nodos, o coordenador PAN, através da informação global da rede, constrói um grafo  $G = (V, E)$ . Para isso, leva-se em consideração as seguintes restrições:

1. Somente os nodos que possuem uma distância menor que um limiar predefinido ( $d_{limit}$ ) podem se relacionar.
2. Para ponderação do grafo  $G$  deve ser considerada a distância entre os nodos ou o número de saltos até o coordenador PAN.

A primeira restrição é assumida para que as relações entre os nodos tenham uma qualidade mínima de sinal, possibilitando uma comunicação adequada. A segunda refere-se à ponderação de  $G$ , ou seja, dependendo da aplicação, podem ser utilizados diferentes critérios, porém, por questões de simplicidade, apenas um critério deve ser escolhido por vez.

Depois da ponderação de  $G$ , o coordenador PAN constrói uma árvore, com base nos pesos atribuídos às relações. Nessa fase, um dos seguintes métodos pode ser utilizado: Árvore Geradora Mínima e Árvore de Menor Caminho.

### 4.2.1 Árvore Geradora Mínima

Uma Árvore Geradora Mínima, também conhecida por *Minimum Spanning Tree* (MST), consiste em um subconjunto acíclico de um grafo  $G$  qualquer que conecte todos os seus vértices e cuja soma total dos pesos é minimizada (CORMEN et al., 2009). Resumidamente, o principal objetivo de uma MST é encontrar, em um grafo ponderado, um subgrafo que conecte todos os vértices com menor custo possível e sem a presença de ciclos (PATEL; PATEL, 2015).

Existem diversos métodos para construção de uma MST, entre eles, os mais conhecidos são os algoritmos de Prim, Kruskal e Boruvka. Devido à árvore possuir raiz no coordenador PAN, o método mais adequado para construção da MST é o algoritmo de Prim (CORMEN et al., 2009), sendo apresentado no Algoritmo 1.

O Algoritmo 1 é inicializado com a entrada dos seguintes parâmetros: o grafo  $G$ , a matriz de pesos das arestas  $W$  e a raiz da árvore  $r$ .

---

**Algoritmo 1:** Algoritmo de Prim.

---

**Entrada:**  $G, W, r$   
**Saída:** árvore geradora mínima  
1 **para** cada  $u \in V(G)$  **faça**  
2      $u.peso = \infty$ ;  
3      $u.pai = \text{nulo}$ ;  
4 **fim**  
5  $r.peso = 0$ ;  
6  $Q = V(G)$ ;  
7 **enquanto**  $Q \neq \emptyset$  **faça**  
8      $u = \text{Extraí Peso Mínimo em } Q$ ;  
9     **para** cada  $v$  adjacente( $u$ ) **faça**  
10         **se**  $v \in Q$  **então**  
11              $v.peso = W(u,v)$ ;  
12              $v.pai = u$ ;  
13         **fim**  
14     **fim**  
15 **fim**

---

Logo após (linhas 1-5), são definidos os atributos de *peso* e *pai* de cada vértice  $V$  de  $G$ , com os valores infinito e nulo respectivamente, exceto para o nodo raiz (coordenador PAN), o qual é definido o valor zero para o atributo *peso* e nulo para o *pai*. Na sequência (linha 6), o grafo  $Q$  recebe os vértices de  $G$  com todas essas atribuições. Em seguida (linhas 7 e 8), enquanto existirem vértices em  $Q$ , são extraídos aqueles com menor peso. Após a extração de  $u$  (linhas 10-12), os pesos e os pais dos vértices adjacentes a  $u$  e presentes no grafo  $Q$  são atualizados com os atributos corretos.

Considerando a primeira iteração do algoritmo, o único vértice com peso diferente de infinito é a raiz, a qual é o primeiro vértice a ser extraído. Com isso, todos os vértices que são adjacentes, ou seja, que possuem relação com o coordenador PAN, têm seus pesos e pais atualizados. Na próxima iteração, o vértice de menor peso relacionado à raiz é selecionado, extraído de  $Q$  e tem seus vértices adjacentes atualizados. O algoritmo executa até que todos os vértices sejam extraídos de  $Q$ . Assim, no final da execução, tem-se uma MST constituída pelos vértices com seus respectivos pais e pesos atualizados.

#### 4.2.2 Árvore de Menor Caminho

Enquanto que o método MST consiste em encontrar uma árvore que minimize a soma total dos custos, a árvore de menor caminho (SPT

– *Shortest Path Tree*) visa encontrar o menor custo de cada caminho entre os vértices e a raiz da árvore. Isto implica que uma SPT não é necessariamente uma MST. O custo total de um caminho é o somatório de cada relação presente nele.

Um dos algoritmos mais conhecidos para formação de uma SPT é o algoritmo de Dijkstra, o qual utiliza o conhecimento global da rede para calcular os caminhos de menor custo entre um nó e todos outros nós da rede (KUROSE; ROSS, 2013).

O algoritmo de Dijkstra é similar ao de Prim, apresentado no Algoritmo 1. Porém, ao invés de considerar apenas o peso de uma relação individual, o algoritmo de Dijkstra considera a soma de todos os pesos presentes no caminho. O Algoritmo 2 apresenta o método proposto por Dijkstra.

---

**Algoritmo 2:** Algoritmo de Dijkstra.

---

```

Entrada:  $G, W, r$ 
Saída: árvore de menor caminho
1 para cada  $u \in V(G)$  faça
2    $u.peso = \infty;$ 
3    $u.pai = \text{nulo};$ 
4 fim
5  $r.peso = 0;$ 
6  $Q = V(G);$ 
7 enquanto  $Q \neq \emptyset$  faça
8    $u = \text{Extrai Peso Mínimo em } Q;$ 
9   para cada  $v \text{ adjacente}(u)$  faça
10    se  $v \in Q$  então
11       $v.peso = u.peso + W(u,v);$ 
12       $v.pai = u;$ 
13    fim
14  fim
15 fim

```

---

A diferença entre os algoritmos 1 e 2 está na linha 11 de ambos, onde, ao invés de atualizar o peso dos vértices adjacentes a  $u$  apenas com o custo da relação  $(u, v)$ , atualiza-se com a soma de todo custo do caminho.

### 4.2.3 Esquema de comunicação

Conforme apresentado na Seção 2.3, essa proposta utiliza o esquema de comunicação proposto no modelo CT-SIM, empregando também, a técnica de alocação da duração do superframe proporcional à

carga imposta pelos nodos descendentes (Load-SDA).

Apesar da proposta deste capítulo não ser baseada em agrupamentos, podemos considerar as relações hierárquicas de maneira similar ao modelo CT-SIM. Dessa forma, os mesmos esquemas de endereçamento, escalonamento, configuração do superframe e comunicação podem ser considerados.

## 4.3 SIMULAÇÃO E RESULTADOS

Através de uma série de simulações, esta seção apresenta uma comparação entre os dois métodos de formação de árvore descritos anteriormente. Em cada um dos métodos, dois critérios podem ser utilizados para ponderação do grafo  $G$ , são eles: distância entre os nodos e número de saltos até o coordenador PAN.

A comparação entre os métodos é dividida em duas partes. Primeiramente, são apresentados os principais parâmetros em relação à disposição dos nodos e ao cenário, juntamente com as diferentes formações de topologia geradas por cada método. Em seguida, o funcionamento da rede é avaliado em relação à sua escalabilidade, a qual emprega o modelo CT-SIM (LEAO et al., 2017b) para o escalonamento de *beacons*. Por fim, são avaliadas as métricas de taxa de sucesso na transmissão dos pacotes até o coordenador PAN, atraso fim-a-fim na entrega dos pacotes e consumo energético.

Com objetivo de verificar o desempenho e a robustez dos métodos de formação para diversas situações, são realizadas um total de 10 simulações com as disposições dos nodos diferentes. Dessa forma, as inconsistências dos resultados são reduzidas.

### 4.3.1 Características de Formação

A disposição aleatória dos nodos é muito comum em aplicações de RSSF, pois dependendo do número de sensores e do ambiente a ser monitorado, uma instalação em pontos geográficos exatos pode ser impraticável. Nesse sentido, a fim de aproximar as simulações da realidade, uma distribuição aleatória uniforme é utilizada para definir as localizações dos nodos no ambiente. A Tabela 6 descreve os principais parâmetros utilizados nessa fase da simulação.

De acordo com a Tabela 6, o parâmetro  $d\_limit$  representa a máxima distância que os nodos podem se relacionar. O valor desse

Tabela 6: Parâmetros de formação de topologia em árvore.

Parâmetro	Descrição	Valor
<i>numNodes</i>	Número de nodos sensores, incluindo o coordenador PAN.	101
<i>deployment</i>	Tipo de implantação disponibilizada pelo Castalia.	<i>uniform</i>
<i>field_x</i>	Tamanho do eixo x do cenário (metros).	100
<i>field_y</i>	Tamanho do eixo y do cenário (metros).	100
<i>d_limit</i>	Distância máxima permitida para relação entre nodos (metros)	20

parâmetro foi definido com base no experimento representado pela Figura 9. Nesse experimento, foi medida a taxa de sucesso de transmissão entre um nodo e seu coordenador, para diversas distâncias diferentes. Com isso, optou-se por escolher o valor 20, pois permite transmissões equilibradas entre distância e qualidade de sinal.

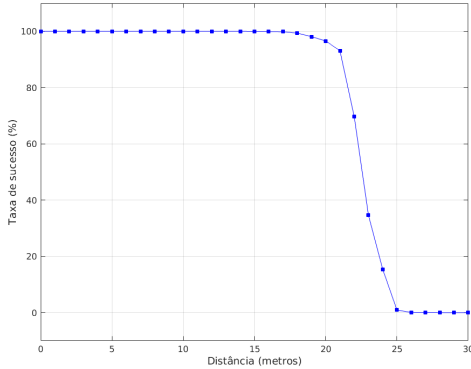


Figura 9: Experimento para definição da distância limite das relações entre nodos.

Conforme descrito no início desta seção, diferentes critérios podem ser utilizados para formação da topologia. Nesse sentido, dois critérios típicos são considerados, distância entre os nodos e número de saltos até o coordenador PAN. Enquanto o primeiro critério visa conectar os nodos considerando a sua proximidade, a qual é proporcional à qualidade do sinal, o segundo critério tende a reduzir o número de saltos necessários para as mensagens chegarem no coordenador PAN,



consequentemente diminuindo a quantidade de retransmissões.

Quando aplicamos o critério de distância para construção da árvore geradora mínima, o algoritmo de Prim possui o objetivo de diminuir as distâncias entre todas relações individuais, formando uma árvore com custo total minimizado. Por outro lado, quando aplicamos o mesmo critério para construção de uma árvore de menor caminho, o algoritmo de Dijkstra visa minimizar o custo total de cada caminho e não das distâncias individuais ou da árvore como um todo.

Diferentemente do critério de distância, o critério de número de saltos até o coordenador PAN resulta na mesma árvore para ambos algoritmos. Isso porque, a construção da árvore geradora mínima para redução do número de saltos pode empregar a abordagem de busca em largura (BFS – *Breadth-First Search*), uma vez que existem muitos pesos iguais atribuídos às relações. Tal abordagem é utilizada com a estratégia de priorizar a expansão da largura da árvore, formando caminhos com menores número de saltos, sendo a base para os algoritmos utilizados em árvores de menor caminho, como o algoritmo de Dijkstra (CORMEN et al., 2009; LACHOWSKI et al., 2015). Dessa forma, o critério de número de saltos é utilizado apenas com o algoritmo de Dijkstra. Porém, devido a esse critério atribuir muitos pesos repetidos às relações, diversas árvores diferentes podem ser formadas. Sendo assim, com o objetivo de melhorar a formação, a distância entre os nodos é empregada como critério de desempate. Ou seja, caso existam dois ou mais nodos com o mesmo número de saltos até a raiz, será escolhido aquele com menor distância.

A partir de um grafo  $G$ , três árvores distintas são construídas, as quais são apresentadas na Figura 10. Nos gráficos da figura, o coordenador PAN está centralizado no ambiente monitorado, no qual as arestas representam as relações entre os nodos.

Conforme se pode observar, a Figura 10(a) mostra o grafo  $G$  utilizado para construção das árvores em um cenário com 100 nodos, onde as arestas representam todas as relações com distâncias menores que  $d\_limit$ . Nas figuras 10(b), 10(c) e 10(d), são apresentadas as diferentes árvores formadas para cada um dos três métodos utilizados.

Para avaliar as formações resultantes de cada método, são analisadas as seguintes características de formação: profundidade média da rede, número de agrupamentos e distância média das relações. Além disso, foram realizadas 10 simulações com diferentes disposições em cenários com 100 nodos mais o coordenador PAN. Os resultados médios de todas simulações estão presentes na Tabela 7.

Devido ao resultado médio das simulações nem sempre repre-

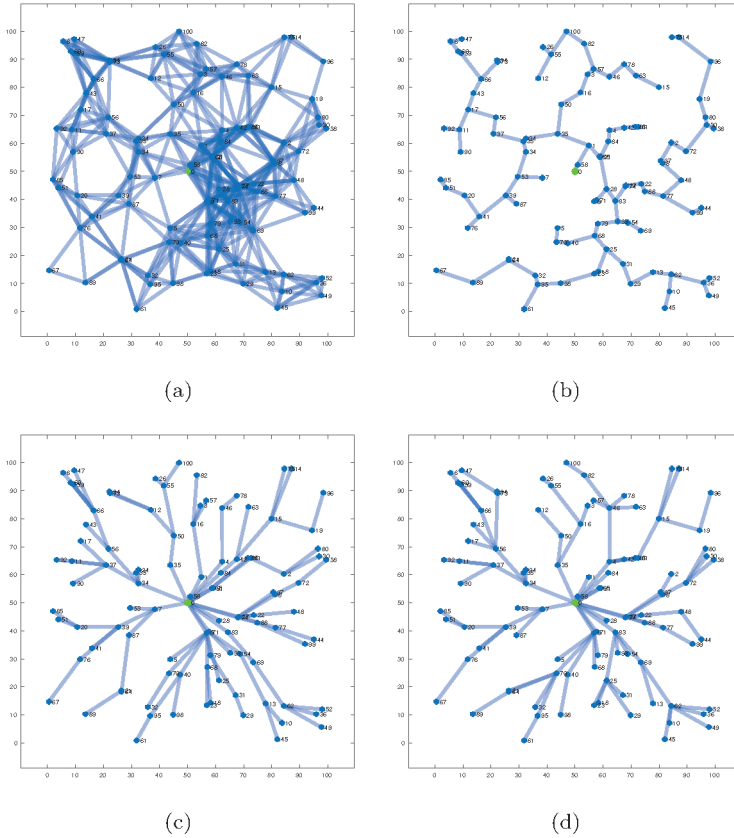


Figura 10: (a) Grafo  $G = (V, E)$ . (b) Árvore Geradora Mínima ponderada por distância (MST). (c) Árvore de Menor Caminho ponderada por distância (SPD). (d) Árvore de Menor Caminho ponderada por número de saltos (SPH).

sentar consistentemente os resultados obtidos em cada simulação, a Figura 11 apresenta em forma de gráfico de barras uma comparação visual entre os métodos de Árvore Geradora Mínima (MST *Minimum Spanning Tree*), Árvore de Menor Caminho ponderada por Distância (SPD *Shortest Path by Distance*) e Árvore de Menor Caminho ponderada por Saltos (SPH *Shortest Path by Hop*). Cada gráfico representa uma métrica de avaliação e as barras representam os resultados médios de cada método com seu respectivo desvio padrão anexado. O desvio

Tabela 7: Resultados médios das características de formação.

Método de Formação	Profundidade Média	Número Médio de Pais	Distância média das relações
Árvore Geradora Mínima ponderada por distância (MST)	14,71 saltos	77,8 pais	6,8 metros
Árvore de Menor Caminho ponderada por distância (SPD)	2,95 saltos	43,7 pais	13,5 metros
Árvore de Menor Caminho ponderada por número de saltos (SPH)	2,88 saltos	34,8 pais	12,7 metros

padrão representa a variação obtida entre as 10 simulações executadas sobre cada um dos métodos.

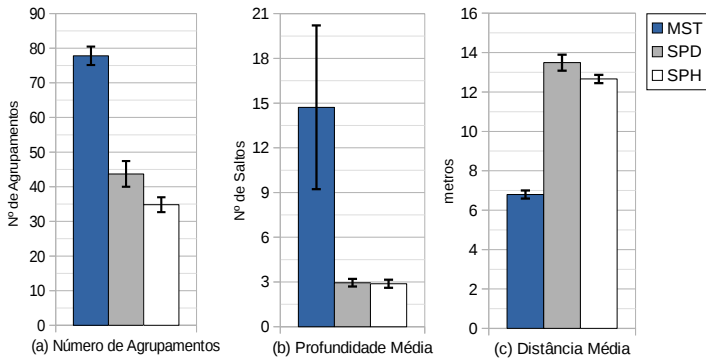


Figura 11: Avaliação das árvores construídas pelos métodos MST, SPD e SPH em relação à (a) Número de agrupamentos, (b) Profundidade Média e (c) Distância Média.

Como é possível notar na Figura 11, o método MST apresenta um número de agrupamentos e uma profundidade média significativamente maiores que os apresentados pelos outros dois métodos. Essa alta quantidade de agrupamentos, assim como a profundidade, são consequência da tentativa de encurtar as distâncias entre as relações, resultando na menor distância média entre os métodos. Além disso, o desvio padrão em relação à profundidade do método MST é bastante alto, correspon-

dendo à sensibilidade do mesmo a diferentes disposições dos nodos.

Ainda na Figura 11, nota-se que os métodos SPD e SPH apresentam resultados similares. Porém, é interessante observar que mesmo que o primeiro utilize a ponderação por distância, o método SPH possui menor distância média. Isso porque, enquanto que o método SPD considera os custos dos caminhos inteiros e não individualmente, o SPH utiliza a distância individual das relações como critério de desempate para construção dos caminhos.

### 4.3.2 Análise de Escalabilidade

Conforme descrito na Seção 1.1.2, uma vez que o aumento do número de nodos permite maior conectividade e redundância de mensagens, a escalabilidade é um fator fundamental para cobertura e confiabilidade da rede.

A Tabela 8 apresenta os parâmetros considerados para o escalonamento da rede. A taxa de geração de pacotes igual a 0,05 pacotes por segundo equivale a um período de geração de pacote igual a 20 segundos (1/0,05). Conforme a Equação 2.5, o valor de BI não pode ser maior que o período de geração de pacotes (20 segundos), por isso, para satisfazer essa condição, o valor de BO é definido para que o valor de BI seja equivalente a 16 segundos. O rádio empregado na simulação é baseado em um modelo real, especificado no arquivo *CC2420.txt*, o qual define a taxa máxima de transmissão (*phyDataRate*) e o número de bits equivalentes a um símbolo (*phyBitsPerSymbol*).

Tabela 8: Parâmetros de escalonamento de topologia em árvore.

Parâmetro	Descrição	Valor
<i>packet_rate</i>	Taxa de geração de pacotes (pacotes/segundo).	0,05
<i>beaconOrder</i>	Parâmetro BO que define o valor de BI	10
<i>RadioParametersFile</i>	Arquivo de especificações do rádio.	<i>CC2420.txt</i>
<i>phyDataRate</i>	Taxa máxima de transmissão (kbps)	250
<i>phyBitsPerSymbol</i>	Quantidade de bits equivalentes a um símbolo	4

Nesse contexto, baseado na Equação 2.5 do modelo CT-SIM (Seção 2.3), é realizada uma análise de escalabilidade de cada método. Para isso, são comparados os somatórios dos períodos ativos (SDs) es-

calonados para cada formação. Esses somatórios equivalem ao tempo necessário para que todas as mensagens possam ser transmitidas para o coordenador PAN. Mais especificamente, para respeitar as restrições impostas pelo modelo, quanto maior o somatório de SDs, maior deve ser o valor de BI e, consequentemente, menor deve ser a taxa de geração de pacotes, a qual é o inverso do período ( $P = 1/\text{taxa de geração}$ ). A Figura 12 apresenta, para cada método, a média dos somatórios dos SDs com o desvio padrão representando a variação para 10 simulações com diferentes disposições.

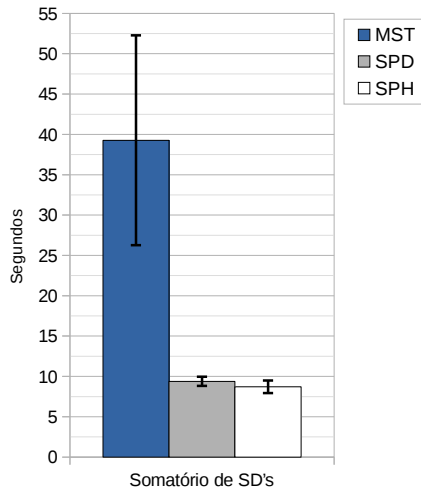


Figura 12: Análise de escalonamento para os métodos MST, SPD e SPH.

Conforme apresentado na Figura 12, o método MST não satisfaz as condições do modelo, uma vez que, de acordo com a Equação 2.5, o somatório dos SDs é maior que o valor de BI (16 segundos). Mais especificamente, o método MST não é escalonável para essa taxa de geração de pacotes, pois o valor de BI não pode ser maior que o período de geração de pacotes definido (20 segundos). Portanto, para tornar esse método escalonável, o período de geração dos pacotes e o valor de BI deveriam ser redefinidos. Os outros métodos satisfazem as condições do modelo, apresentando somatórios similares. Além disso, conforme pode ser visto nos desvios padrões anexados aos resultados, os métodos SPD e SPH possuem pouca sensibilidade à disposição dos nodos no cenário, significando maior robustez.

### 4.3.3 Análise de Desempenho

Conforme apresentado na Seção 3.1, muitos trabalhos se preocupam em reduzir o consumo de energia da rede. Apesar dessa métrica de desempenho ser uma das mais importantes, ela não é a única. Dependendo da aplicação, outras métricas de desempenho também deveriam ser atendidas, como por exemplo, taxa de sucesso na entrega das mensagens e atraso fim-a-fim entre a geração e entrega das mensagens ao coordenador PAN.

Enquanto o atraso fim-a-fim das mensagens representa o tempo que demora para os dados medidos pelos nodos chegarem ao coordenador PAN e possam ser analisados, a taxa de sucesso representa a relação entre os pacotes gerados pelos nodos e os recebidos pelo coordenador PAN. Esta relação possibilita uma avaliação indireta da conectividade e da cobertura da rede, pois caso alguns nodos gerem seus pacotes de dados e não consigam transmiti-los por falta de conectividade, tanto a taxa de sucesso como a cobertura da mesma são reduzidas.

Nesse contexto, para avaliação do desempenho de operação da rede, uma série de simulações foram realizadas. A Tabela 9 mostra os principais parâmetros utilizados para essas simulações.

Tabela 9: Parâmetros utilizados na comunicação da rede.

Parâmetro	Descrição	Valores		
		MST	SPD	SPH
<i>sim-time-limit</i>	Tempo de duração da simulação (segundos).	7200	7200	7200
<i>packet_rate</i>	Taxa de geração de pacotes (pacotes/segundo).	0,016	0,05	0,05
<i>beaconOrder</i>	Parâmetro que define a duração de BI através da Equação 2.1.	11	10	10
<i>macMinBE</i>	Expoente de <i>backoff</i> mínimo, que determina a duração dos <i>backoffs</i> do mecanismo CSMA-CA.	5	7	7
<i>macMaxBE</i>	Expoente de <i>backoff</i> máximo, que determina a duração dos <i>backoffs</i> do mecanismo CSMA-CA.	6	8	8

Entre os parâmetros mostrados, existem algumas diferenças entre os valores definidos para o método MST e os outros métodos. De acordo com a análise de escalabilidade realizada anteriormente, não é possível definir a mesma taxa de transmissão para todos os métodos. Com isso, enquanto que os métodos SPD e SDS possuem um período

de geração de pacotes de 20 segundos, o MST possui um período de 62,5 segundos para satisfazer as restrições do modelo CT-SIM. Pelo mesmo motivo, os BOs também são diferentes. Além disso, como foi apresentado na Seção 3.3, devido ao esquema de comunicação utilizado pelos nodos empregar o mecanismo CSMA-CA, alguns parâmetros devem ser adequados conforme a topologia utilizada. Sendo assim, os valores dos parâmetros *macMinBE* e *macMaxBE* foram definidos com base em diversos experimentos.

Com base nas configurações definidas, as figuras 13, 14 e 15 apresentam, para um cenário específico de 100 nodos, gráficos da distribuição dos valores de energia consumida e atraso fim-a-fim, para os métodos MST, SPD e SPH, respectivamente. A taxa de sucesso não é mostrada, pois todos os métodos apresentaram taxas muito próximas de 100%, ou seja, quase todas as mensagens geradas foram entregues ao coordenador PAN.

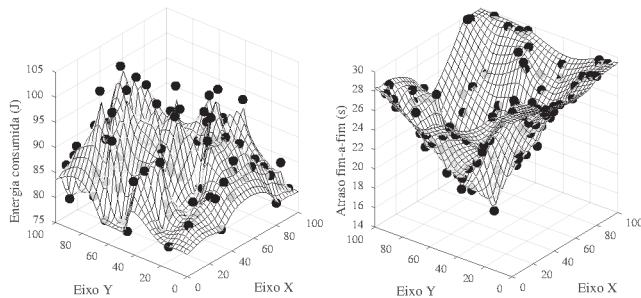


Figura 13: Distribuição de energia consumida e atraso fim-a-fim para o método MST.

Como se pode observar na Figura 13, enquanto os menores consumos de energia estão presentes nos nodos localizados nas extremidades do ambiente, os maiores estão nos nodos mais próximos do centro. Esse comportamento é justificado pelo coordenador PAN estar localizado no centro do cenário, onde o fluxo de mensagens é maior. Por outro lado, a distribuição do atraso fim-a-fim entre os nodos apresenta um comportamento inverso ao do consumo energético, pois as mensagens enviadas pelos nodos mais próximos ao coordenador precisam de menos saltos para serem entregues.

Conforme ilustrado nas figuras 14 e 15, os métodos SPD e SPH apresentam comportamentos similares ao do MST. Apesar disso, existem algumas diferenças que podem ser notadas. A principal diferença

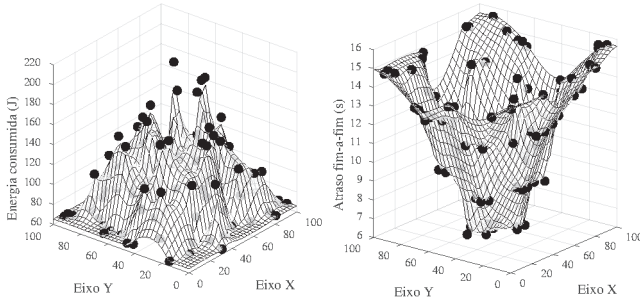


Figura 14: Distribuição de Energia consumida e Atraso fim-a-fim para o método SPD.

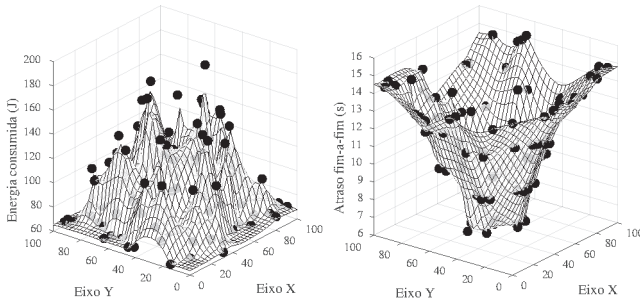


Figura 15: Distribuição de Energia consumida e Atraso fim-a-fim para o método SPH.

está relacionada aos valores encontrados. A Figura 16 apresenta, em forma de gráfico de barras, as médias e os respectivos desvios padrões entre todos nodos da rede para os valores de consumo de energia e de atraso fim-a-fim. Enquanto a média representa um valor global da rede para uma determinada métrica, o desvio padrão ilustra a distribuição dessa métrica entre os nodos.

Como é possível observar nas figuras 13 e 16(a), o método MST possui menores valores de variação e consumo de energia entre os nodos da rede em relação aos outros métodos. Esses resultados são explicados pelo fato de sua taxa de transmissão de pacotes ser consideravelmente menor do que a taxa dos outros métodos. Enquanto que cada nodo em uma rede formada pelo método MST entrega, em média, 107 pacotes ao coordenador PAN, os nodos das redes formadas pelos métodos SPD e SPH entregam, em média, 337 pacotes.



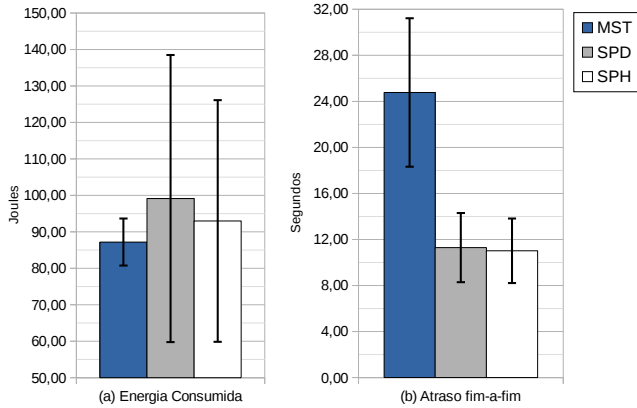


Figura 16: Média entre os nodos de um cenário específico do (a) Consumo de Energia e do (b) Atraso fim-a-fim dos métodos MST, SPD e SPH.

A Figura 16(b) mostra o atraso que as mensagens sofrem a partir de sua geração, até sua entrega ao coordenador PAN. Pode-se notar que os nodos do método MST possuem um atraso médio de aproximadamente 25 segundos, mais que o dobro da média dos outros dois métodos. Esse maior atraso está relacionado com a profundidade da árvore gerada e com o BI definido, cujos valores são maiores.

Para avaliar como os critérios de desempenho variam em relação à disposição dos nodos, 10 simulações com cenários diferentes foram realizadas para cada método.

A Figura 17 apresenta os valores médios similares aos anteriores, significando que o desempenho não sofre muita influência com a alteração do cenário. Porém, conforme mostram os desvios padrões anexados em cada resultado, existe uma variação associada à disposição dos nodos no cenário, principalmente no método SPD em relação à energia consumida e no método MST em relação ao atraso fim-a-fim.

Por fim, devido às diferenças das taxas de transmissão entre os métodos, a Figura 18 apresenta uma relação entre energia consumida e o número de pacotes transmitidos. Essa relação possibilita uma melhor comparação do consumo de energia entre os métodos, uma vez que o consumo de energia mais significativo é relacionado à comunicação entre os nodos.

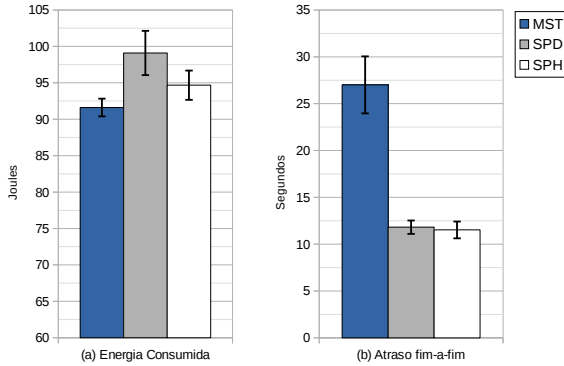


Figura 17: Média entre 10 cenários diferentes do (a) Consumo de Energia e do (b) Atraso fim-a-fim dos métodos MST, SPD e SPH.

Enquanto a Figura 18(a) apresenta a média do consumo de energia entre os nodos para um cenário específico, a Figura 18(b) apresenta a média entre diversos cenários. Com isso, tem-se que, respectivamente, os desvios padrões representam a variação do consumo entre os nodos e entre as diferentes disposições.

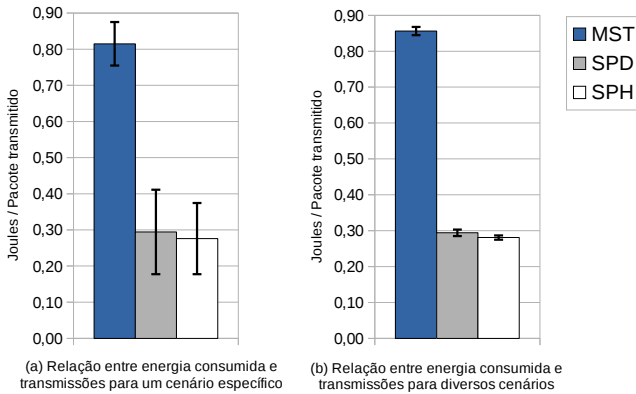


Figura 18: Relação entre energia consumida e pacotes transmitidos para: (a) Cenário específico, (b) Diversos cenários.

Como se pode notar na Figura 18, o método MST apresenta um consumo de energia significativamente maior que os outros métodos. Isso porque, devido a sua maior profundidade média, os pacotes de-

vem ser retransmitidos mais vezes para alcançar o coordenador PAN, conseqüentemente aumentando o consumo de energia de toda rede.

#### 4.4 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados dois métodos frequentemente utilizados na construção de topologias em árvore, são eles: MST (*Minimum Spanning Tree*) e SPT (*Shortest Path Tree*). Ainda, foi discutido o uso de diferentes custos baseados em distância e profundidade. A combinação desses conceitos resultaram em três métodos distintos, denominados MST ponderado por distância, SPD (*Shortest Path by Distance*), e SPH (*Shortest Path by Hop*). Esses métodos foram descritos e avaliados, considerando suas características de formação, escalabilidade e desempenho na fase de comunicação da rede.

A avaliação mostrou que o método MST compromete a escalabilidade e o desempenho da rede, pois constrói caminhos demasiadamente profundos em relação aos outros métodos. Por outro lado, os métodos SPD e SPH apresentaram melhores resultados, com comportamentos similares. Porém, devido ao objetivo de reduzir o número de saltos, considerando um critério de desempate por distância, o método SPH apresentou melhores resultados em relação à escalabilidade e desempenho do que os outros. No que se refere à taxa de sucesso nas transmissões, todos os métodos tiveram resultados de aproximadamente 100%. Isso se deve à utilização da comunicação escalonada conforme o modelo CT-SIM.

Por fim, como o objetivo deste capítulo foi o de avaliar o comportamento da rede para diferentes topologias, optou-se por assumir uma abordagem centralizada, na qual, a partir de uma visão global da rede, o coordenador PAN pôde defini-la completamente, de diversas formas. Dessa maneira, foi possível verificar os principais aspectos da topologia que interferem no desempenho da rede. Entre esses aspectos, destacam-se a profundidade da rede e o número de agrupamentos, os quais interferem tanto na escalabilidade quanto no desempenho da rede.



## 5 ACARCH: ADAPTATIVE CLUSTERING ARCHITECTURE

Neste capítulo é apresentada a proposta ACArch (*Adaptive Clustering Architecture*), uma arquitetura de gerenciamento de topologia em agrupamentos em árvore para redes em conformidade com o padrão IEEE 802.15.4. A ACArch é constituída por uma série de métodos, que visam a construção de uma rede robusta, capaz de se adaptar a cenários diferentes, sem que haja perdas significativas em seu desempenho. Para atingir esses objetivos, considera-se as conclusões obtidas do estudo apresentado no Capítulo 4.

O capítulo começa com a descrição do modelo adotado para formação e simulação da rede, quando são apresentados os principais fundamentos e pressupostos. Logo em seguida, a arquitetura é descrita, apresentando todos os métodos empregados e seus respectivos funcionamentos. Por fim, as simulações e os resultados são apresentados.

### 5.1 DESCRIÇÃO DO MODELO

Diferentemente dos métodos apresentados no Capítulo 4, a presente proposta busca executar seus procedimentos de maneira descentralizada, sem considerar aspectos de ciência de localização e de visão global da rede. Isso torna a ACArch mais apropriada a cenários reais, nos quais, uma formação de modo centralizado com conhecimento global da rede pode ser inviável.

Além disso, assim como no capítulo anterior, na simulação usando o Castalia, este capítulo considera os modelos de disco unitário e de interferência aditiva para todos os métodos e experimentos apresentados.

#### 5.1.1 Principais Pressupostos

O método proposto neste capítulo assume um sistema formado por um conjunto de nodos de tamanho conhecido, dispostos conforme uma distribuição aleatória uniforme desconhecida, sobre uma área retangular, e organizados em uma árvore de agrupamentos. Além disso, o modelo considera os seguintes pressupostos:

- (i) A área monitorada possui dimensões de tal forma que os dados transmitidos pelos nodos, geralmente, não alcançam diretamente

o coordenador com apenas um salto.

- (ii) Os nodos são estáticos, possuem as mesmas capacidades de processamento e comunicação, utilizam a mesma potência de transmissão e não são cientes de suas localizações.
- (iii) A aplicação tem um padrão de comunicação periódico, isto é, os nodos coletam dados periodicamente do ambiente, transmitem aos seus pais que, por sua vez, retransmitem com destino ao coordenador PAN.
- (iv) Os canais de comunicação são simétricos.
- (v) Os dados não sofrem nenhum tipo de agregação, codificação ou fusão durante as suas transmissões.

As propriedades assumidas para a ACArch são similares às apresentadas na Seção 4.1, porém existem algumas diferenças importantes. As justificativas para os itens (i), (ii) e (iii) são as mesmas apresentadas na Seção 4.1, com exceção da ciência da localização.

A propriedade (iv) é justificada pelo uso do valor de RSSI (*Received Signal Strength Indication*) como forma de mensurar a força do sinal entre dois nodos, o qual pode ser medido por um nodo no momento da recepção de um *frame* enviado por outro nodo. O valor de RSSI é considerado como uma métrica adequada para avaliar a qualidade do canal de comunicação, e pode ser assumido como simétrico nos rádios modernos (LEVIS et al., 2006).

Por último, a justificativa para a propriedade (v) é a mesma da propriedade (vi), apresentada na Seção 4.1.

## 5.2 DESCRIÇÃO DA ARQUITETURA

A ACArch pode ser dividida em três etapas subsequentes: descobrimento de vizinhança, formação da rede e escalonamento de *beacons*. Conforme a Figura 19, as etapas interagem por meio de estruturas de dados. Neste caso, as estruturas são duas tabelas e um arranjo hierárquico que representam as vizinhanças, os nodos já associados e as relações dos agrupamentos em árvore, respectivamente. Além disso, a segunda etapa é constituída de dois métodos, associação de filhos e seleção de CHs, os quais são responsáveis pela formação da estrutura da rede.

Primeiramente, a ACArch é inicializada por um método de descobrimento de vizinhança, onde cada nodo define sua tabela de vizinhança contendo os nodos dentro do seu respectivo alcance de rádio. Logo após, com base nessas tabelas, a etapa de formação da rede é

executada. Essa execução ocorre de forma iterativa, começando pela associação de filhos, seguida pela seleção de CHs, que associam mais filhos e selecionam outros CHs, e assim por diante. Como a disposição dos nodos é desconhecida, não se sabe a quantidade de iterações necessárias para cobrir toda rede. Dessa forma, uma estimativa de tempo deve ser feita. Com isso, assume-se um valor superestimado que permita a formação completa da rede. Por fim, o escalonamento de *beacons* (Seção 2.3) é executado sobre essa estrutura, definindo os endereços, o escalonamento da comunicação e a alocação da duração do *superframe*. Dessa forma, a rede está preparada para iniciar o monitoramento do ambiente.

Todos os métodos que constituem a ACArch são apresentados a seguir, com exceção do escalonamento de *beacons*, o qual já foi descrito na Seção 2.3.

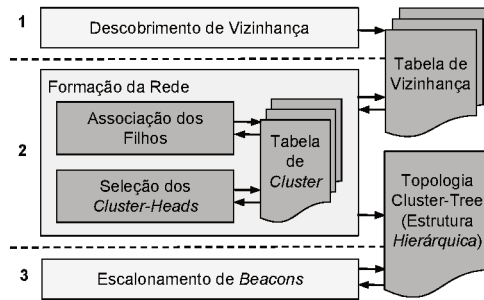


Figura 19: Adaptative Clustering Architecture.

### 5.2.1 Descobrimto de Vizinhança

Dependendo do número de nodos e do tamanho da rede, a aquisição do conhecimento global da mesma por um ou mais nodos pode ser uma tarefa bastante custosa, capaz de sobrecarregar a rede, causando gastos excessivos de energia e tempo. Entretanto, o uso desse conhecimento na etapa de formação da rede contribui significativamente na qualidade da topologia formada. Nesse sentido, como primeira etapa da construção da rede, a ACArch emprega um método de descobrimto de vizinhança. O objetivo do método é que os nodos obtenham, de forma eficiente, o conhecimento local sobre sua vizinhança.

Essa etapa é fundamental na formação da rede, pois permite

uma implementação eficiente dos esquemas de comunicação e gerenciamento de energia. Existem diversas técnicas para descobrimento da vizinhança. Entretanto, conforme apresentado por (SUN et al., 2014), os *protocolos de aniversário* são recomendados para redes estáticas que possuem implantações aleatórias, pois apresentam maior flexibilidade, simplicidade e eficiência para esse tipo de aplicação (MCGLYNN; BORBASH, 2001). Esses protocolos são inspirados no conceito do "paradoxo do aniversário", o qual determina que, dado um grupo de  $x$  pessoas escolhidas aleatoriamente, haverá uma probabilidade de que, pelo menos, duas delas tenham a mesma data de aniversário. Como exemplo, se 23 pessoas forem escolhidas aleatoriamente, a probabilidade de pelo menos duas delas terem a mesma data de aniversário é de 50%. Extrapolando para 57 pessoas, a probabilidade chega a 99%. Para alcançar 100% de probabilidade, são necessárias pelo menos 367 pessoas.

Esse tipo de protocolo pode ser empregado no descobrimento da vizinhança da seguinte forma: (1) os nodos dividem seu tempo em  $n$  slots; (2) para cada slot existe uma probabilidade do nodo transmitir um pacote de descobrimento ( $p_t$ ), escutar o meio ( $p_e$ ) ou dormir ( $p_d$ ) para economizar energia; (3) descobrimentos acontecem apenas quando um único nodo transmite e um ou mais nodos escutam; caso dois ou mais nodos dentro do mesmo raio de transmissão transmitam juntos, ocorre uma colisão e não há descobrimento; (4) após cada descobrimento, são armazenados em uma tabela, chamada tabela de vizinhança ( $T_v$ ), o identificador ( $ID_v$ ) do nodo transmissor e a identificação da força do sinal recebido ( $RSSI_v$ ).

A Figura 20 apresenta um exemplo com quatro nodos dentro do mesmo raio de transmissão. No primeiro slot de tempo, o nodo 4 descobre o nodo 3. No segundo, ocorre uma colisão e não há descobrimento, pois os nodos 1 e 4 tentam transmitir juntos. No terceiro, todos dormem. Por fim, no quarto e no quinto slots, os nodos 1 e 4 descobrem o nodo 3, e o nodo 3 descobre o 2, respectivamente.

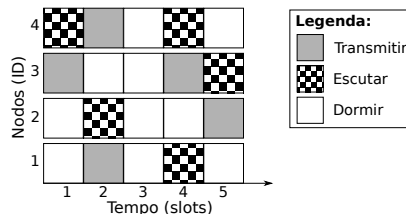


Figura 20: Exemplo da etapa de descobrimento.



Como é possível notar no exemplo, nenhum dos nodos atinge o conhecimento total da vizinhança, pois nem todos os vizinhos dentro do raio de transmissão são descobertos. Essa situação pode melhorar com a variação das probabilidades  $p_t$ ,  $p_e$  e  $p_d$  ou do aumento no número  $n$  de *slots*. Porém, é importante ressaltar que deve haver um balanceamento entre essas variáveis, no sentido de que não ocorram muitas colisões ou um excesso de *slots*, pois assim o consumo de energia e os atrasos nos envios de mensagens aumentariam, respectivamente. Além disso, de acordo com a análise feita por (MCGLYNN; BORBASH, 2001), protocolos que usam descobrimento de vizinhança não têm degradação significativa em seus desempenhos, ainda que não haja descobrimento total da vizinhança.

Para maximizar a eficiência do descobrimento da vizinhança, o presente trabalho emprega o protocolo de aniversário proposto por McGlynn e Borbash (2001), nomeado PRR (*Probabilistic Round Robin*). Esse protocolo visa aumentar as chances de descobrimento através das seguintes atribuições de probabilidade:

$$\begin{aligned} p_t &= \frac{1}{N} \\ p_e &= 1 - \frac{1}{N} \\ p_d &= 0 \end{aligned} \tag{5.1}$$

Conforme mostra o conjunto de equações 5.1, as probabilidades são definidas com base no número de vizinhos  $N$ . Entretanto, como a distribuição de nodos é aleatória, não se sabe, de fato, o número de vizinhos  $N$  para os nodos. Por isso, os autores sugerem superestimar esse número. Além disso, o protocolo PRR não considera a probabilidade do nodo dormir, pois embora economize energia, diminui as chances de descobrimento.

Nessa etapa, apenas os nodos que possuem RSSI maior que um valor predefinido (*RSSI\_limit*) são armazenados na tabela de vizinhança. Isso porque, de acordo com Levis et al. (2006), a partir de um determinado valor de RSSI, as chances de perda de pacotes aumentam significativamente. A descrição de como é feita a definição desse valor é apresentada mais adiante neste capítulo.

### 5.2.2 Associação de Filhos

Assim como os métodos do Capítulo 4, que utilizam uma distância limite para estabelecer relações entre os nodos, essa etapa visa garantir a conectividade da rede através de um limite mínimo de RSSI, o qual é definido através do parâmetro *RSSI\_limit*. Outro fator importante considerado é a preservação do desempenho do método CSMA-CA, que de acordo com Khanafer et al. (2014), sofre degradações conforme o número de nodos disputando acesso ao meio aumenta. Essa preservação é atingida através da definição de um parâmetro limite para o tamanho dos agrupamentos formados, chamado *maxChildren*.

A etapa de associação de filhos começa após a etapa de descobrimento ter sido finalizada. Sua inicialização é realizada pelo coordenador PAN através dos seguintes passos:

1. Uma mensagem de formação é difundida no meio;
2. Ao receber esta mensagem, os nodos encaminham para o coordenador um pedido de associação;
3. Além de ter que respeitar o número máximo de associações permitidas por agrupamento, somente serão aceitos os nodos que estiverem armazenados na tabela de vizinhança do coordenador, ou seja, possuam um *RSSIs* maior que o definido no parâmetro *RSSI\_limit*;
4. Uma mensagem de confirmação de associação é enviada para os nodos aceitos;
5. Após aceitos, os nodos selecionados como CHs continuam a formação da rede enquanto que os demais dormem até o fim da etapa de formação.

Após a formação do primeiro agrupamento, o coordenador PAN executa a etapa de seleção de CHs, informando aos nodos selecionados, através de uma tabela anexada à mensagem de confirmação, quais nodos já foram associados à rede. Essa tabela, nomeada de tabela de *cluster* ( $T_c$ ), é formada pelos *IDs* dos nodos já associados àquele ramo da árvore. Ela é utilizada posteriormente para auxiliar na escolha dos próximos CHs.

Assim como o coordenador PAN, os CHs selecionados enviam mensagens de formação e aguardam pedidos de associação de seus vizinhos. Os pedidos de associação encaminhados ao coordenador *c* contêm

a identificação do nodo ( $ID_i$ ) e a sua respectiva tabela de vizinhança ( $T_{v_i}$ ) formada na etapa anterior, também utilizada para auxiliar a escolha dos próximos CHs.

Ao receber um pedido de associação de um nodo qualquer, o coordenador executa os passos descritos anteriormente. Caso um nodo não seja aceito, nenhuma ação é tomada, deixando o nodo livre para tentar associação a outro agrupamento posteriormente.

### 5.2.3 Seleção de Cluster-Heads

Como foi mostrado no Capítulo 4, as propriedades de profundidade e número de agrupamentos formados na rede influenciam diretamente no número de retransmissões e no escalonamento dos pacotes, respectivamente. Enquanto que o número de transmissões está relacionado principalmente ao escalonamento e ao consumo de energia, a profundidade da rede também relaciona-se com o atraso fim-a-fim sofrido pelas mensagens. Com base nisso, o objetivo dessa etapa é reduzir ao máximo essas propriedades, a fim de manter a escalabilidade e o bom desempenho da rede.

Nesse contexto, o método proposto nessa etapa se baseia no problema de cobertura de conjuntos (*set-covering problem*), o qual é um problema de otimização que modela muitos sistemas de alocação de recursos. Segundo (CORMEN et al., 2009), esse problema é NP-difícil e, para geração de resultados úteis, podem ser utilizadas heurísticas de aproximação. Considerando que todos os nodos possuem uma conectividade mínima, através dessas heurísticas é possível encontrar um número reduzido de agrupamentos capaz de cobrir todos nodos da rede.

O problema de cobertura de conjuntos pode ser resumido da seguinte forma. Dado um conjunto finito  $X$  e uma família  $\Omega$  de subconjuntos de  $X$ , tal que todo elemento de  $X$  pertence a pelo menos um subconjunto em  $\Omega$ , tem-se que  $\Omega$  cobre  $X$ . Se  $\Omega$  é constituído por um número  $n$  de subconjuntos, o objetivo do problema de cobertura de conjuntos é encontrar um  $n$  mínimo capaz de cobrir todo conjunto  $X$ .

Na aplicação desse problema no método de seleção de CHs assume-se que existe um conjunto finito de nodos  $X$  representado por uma família  $\Omega$  de subconjuntos. Cada subconjunto, chamado de  $R_i$ , é constituído das relações de um nodo  $i$  encontradas em  $Tv_i$ . Além disso, a família  $\Omega$  é formada apenas pelos subconjuntos dos nodos associados ao agrupamento  $C$ , ou seja, assumindo  $Tv_i$  como um conjunto, temos que  $R_i = \{ID_j \mid j \in Tv_i\}$  e  $\Omega = \{R_i \mid i \in C\}$ . Neste contexto, a seleção

de CH tem como objetivo selecionar um número mínimo de nodos  $i$  em  $C$ , cujas relações em  $R_i$  cubram todo conjunto  $X$ . Essa seleção é executada pelos coordenadores através do Algoritmo 3.

---

**Algoritmo 3:** Seleção de cluster-heads.

---

**Entrada:**  $X, \Omega$   
**Saída:**  $\beta$

```

1  $U \leftarrow X, \beta \leftarrow \emptyset$ 
2 enquanto  $U \neq \emptyset$  faça
3   |   seleciona o conjunto  $R_i \in \Omega$  que maximize  $R_i \cap U$ 
4   |    $U \leftarrow U - R_i$ 
5   |    $\beta \leftarrow \beta \cup i$ 
6 fim
7 retorna  $\beta$ 
```

---

Através do conjunto  $X$  e da família de subconjuntos  $\Omega$ , o Algoritmo 3 executa a seleção dos nodos que possuem a maior cobertura em  $U$  (linha 3). Porém, uma vez selecionado o nodo com maior cobertura, suas relações devem ser removidas do conjunto  $U$  (linha 4) para não serem contabilizadas novamente. Essa seleção é realizada até que o conjunto  $U$  seja totalmente coberto (linhas 2-6). Por fim, o nodo selecionado é agrupado ao conjunto  $\beta$  (linha 5), o qual é retornado ao final da seleção (linha 7).

### 5.3 SIMULAÇÕES E RESULTADOS

Para avaliação da proposta ACArch, uma série de simulações foi conduzida através do simulador Castalia, utilizando o modelo CT-SIM. Essas simulações visam a comparação da ACArch com os métodos SPH, descrito na Seção 4.2 e LEACH, descrito na Seção 3.1.

Devido ao método SPH utilizar os mesmos mecanismos de comunicação que a ACArch, a comparação entre eles permite analisar as mudanças que os diferentes esquemas de formação causam no desempenho da rede. Por outro lado, o protocolo LEACH permite uma comparação completa, possibilitando uma análise mais detalhada do comportamento da arquitetura proposta.

Assim como o capítulo anterior, as comparações entre os métodos são realizadas em relação às características de formação, à escalabilidade e ao desempenho da rede. Nesse último, serão avaliadas as métricas de taxa de sucesso, atraso fim-a-fim e consumo de energia.

Por fim, para avaliar os métodos em diversas situações, o número de nodos também é variado. Essa variação pode ter vantagens ou desvantagens para rede. Por um lado, o aumento no número de nodos

aumenta também a conectividade da rede. Por outro lado, o número de colisões também aumenta, podendo degradar a escalabilidade e o desempenho da rede.

### 5.3.1 Análise da formação da rede

Nesta etapa, as principais propriedades e parâmetros relacionados à formação da rede são apresentados para cada método. Primeiramente, são descritos os principais parâmetros utilizados para configuração das simulações. Logo após, uma análise sobre a quantidade de agrupamentos, profundidade média e número de nodos órfãos das redes formadas é realizada para diversos cenários, incluindo densidade e disposições de nodos variadas.

#### 5.3.1.1 Definição dos parâmetros de descobrimento

Conforme já mencionado anteriormente, o descobrimento de vizinhança é uma etapa fundamental para formação adequada da rede. A definição dos parâmetros de probabilidade  $p_t$ ,  $p_e$  e  $p_d$  foi realizada conforme o conjunto de equações 5.1. Considerando um número estimado  $N$  de 70 vizinhos, os seguintes valores foram obtidos:

$$\begin{aligned} p_t &= \frac{1}{70} \approx 0,015 \\ p_e &= 1 - \frac{1}{70} \approx 0,985 \\ p_d &= 0 \end{aligned}$$

Para definição do número de *slots*, o experimento representado pela Figura 21 foi realizado. Nele, variou-se o número de *slots* enquanto o número de descobrimentos era medido.

Conforme observado na Figura 21, a partir de aproximadamente 2000 *slots*, o número de descobrimentos não apresenta aumento significativo à medida que o número de *slots* é incrementado. Portanto, foi definido um número de 2000 *slots* para ser utilizado no protocolo de descobrimento. A Tabela 10 apresenta um resumo dos parâmetros configurados nessa etapa.

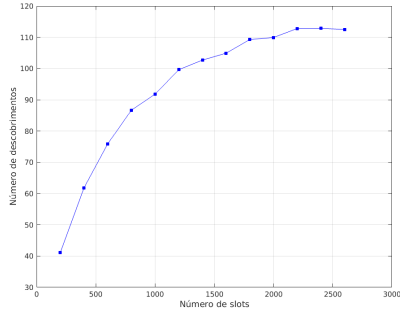


Figura 21: Experimento para definição do número de *slots* utilizado na etapa de Descobrimento de Vizinhança.

Tabela 10: Parâmetros de descobrimento de vizinhança empregados pela ACArch.

Parâmetro	Descrição	Valor
$p_t$	Probabilidade de transmitir pacote de descobrimento.	0,015
$p_e$	Probabilidade de escutar o meio.	0,985
$p_d$	Probabilidade de dormir.	0
$nSlots$	Número de <i>slots</i> de descobrimento	2000

### 5.3.1.2 Definição dos parâmetros de formação

Cada método analisado, ACArch, SPH e LEACH, possui uma série de parâmetros específicos para configuração da formação da rede. A Tabela 11 apresenta esses parâmetros, com uma breve descrição dos mesmos e seus respectivos valores definidos.

Conforme mostrado na Tabela 11, as simulações são realizadas para um conjunto de densidades diferentes, no qual, o número de nodos varia de 51 a 201 em passos de 50. Os cenários utilizados para as simulações possuem dimensão de  $50 \times 50 \text{ metros}^2$  e  $100 \times 100 \text{ metros}^2$ , tendo a disposição dos nodos baseada em uma distribuição aleatória uniforme. Essas duas dimensões foram escolhidas para possibilitar uma análise em cenários de baixa e alta conectividade, sendo que, quanto maior as dimensões, menor a conectividade.

Assim como foi assumido no Capítulo 4, a distância entre os nodos é proporcional à qualidade de sinal da comunicação entre eles. Dessa forma, conforme apresentado na Seção 4.3.1, o parâmetro  $d\_limit$

Tabela 11: Parâmetros utilizados na formação da rede.

Parâmetro	Descrição	Valores		
		ACArch	SPH	LEACH
<i>numNodes</i>	Número de nodos presentes no cenário (incluindo coordenador PAN).	51-201	51-201	51-201
<i>deployment</i>	Tipo de implantação disponibilizada pelo Castalia.	<i>uniform</i>	<i>uniform</i>	<i>uniform</i>
<i>field_x</i>	Tamanho do eixo x do cenário (metros).	50 e 100	50 e 100	50 e 100
<i>field_y</i>	Tamanho do eixo y do cenário (metros).	50 e 100	50 e 100	50 e 100
<i>d_limit</i>	Distância máxima permitida para relação entre nodos (metros).	NA	20	NA
<i>RSSI_limit</i>	Limite mínimo permitido do valor RSSI para associação.	-91 dBm	NA	NA
<i>maxChilden</i>	Limite máximo permitido de filhos por CH.	50	NA	NA
<i>percentage</i>	Probabilidade inicial de um nodo ser CH.	NA	NA	5 %

é definido através do experimento representado na Figura 9.

Para definição do parâmetro *RSSI\_limit*, foi realizado um experimento similar ao realizado para definição do *d\_limit*, onde mediu-se a taxa de sucesso para diversas localizações diferentes, as quais correspondem respectivamente a valores de RSSI. Portanto, conforme apresentado na Figura 22, o valor de *RSSI\_limit* foi definido para *-91 dBm*, sendo que, quanto maior o valor, melhor a conexão.

A definição do parâmetro *maxChildren* foi realizada baseada no experimento da Figura 23. Através desse experimento, se pôde observar a variação no desempenho da rede conforme o número máximo de filhos permitidos era incrementado. Enquanto que a taxa de sucesso e o atraso fim-a-fim médios das mensagens caem conforme o número de filhos cresce, a energia consumida sobe. Esse comportamento deve-se pelo fato da utilização do mecanismo CSMA-CA, o qual consome maior energia e sobrecarrega a rede à medida que o número de nodos disputando acesso ao meio aumenta. Com base nessa análise, se conclui que a definição do parâmetro *maxChildren* depende dos requisitos da aplicação. Portanto, para comparação da ACArch com os outros métodos, optou-se por definir um valor conservador, que preserve a taxa

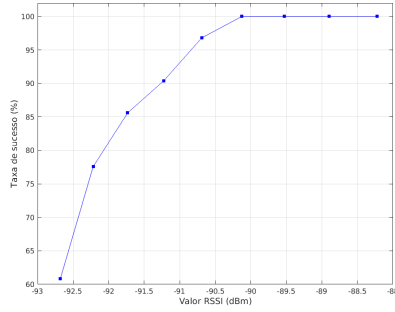


Figura 22: Experimento para definição do parâmetro  $RSSI\_limit$ .

de sucesso e a energia, permitindo um maior atraso das mensagens.

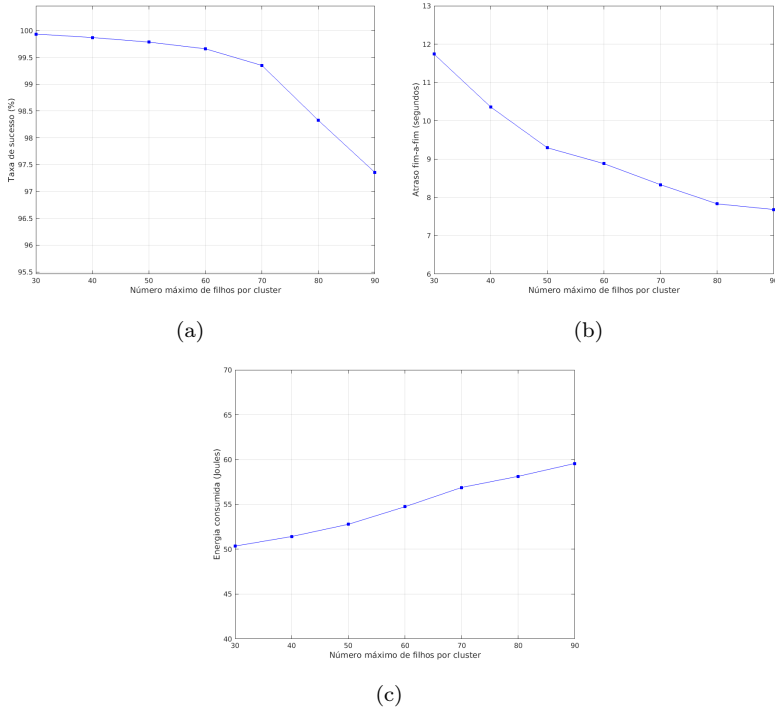


Figura 23: Análise da variação do parâmetro  $maxChildren$  para (a) Taxa de sucesso, (b) Atraso fim-a-fim e (c) Consumo de energia.



De acordo com os autores Heinzelman et al. (2000), o protocolo LEACH apresenta seu melhor desempenho quando 5% de seus nodos têm a possibilidade de se tornar CHs. Com base nisso, o valor do parâmetro *percentage* foi definido.

### 5.3.1.3 Análise dos resultados

Para uma melhor compreensão da formação da rede, a Figura 24 ilustra as três formações obtidas pelos métodos ACArch, SPH e LEACH, respectivamente, em um cenário específico de  $50 \times 50$  metros<sup>2</sup> com 100 nodos.

Como se pode observar nas figuras 24(a) e 24(b), a formação para ACArch e SPH são similares, o que significa que a ACArch consegue reduzir o número de saltos mesmo sem conhecimento global da rede. Porém, devido a ACArch empregar o Algoritmo 3 para resolver o problema de cobertura de conjuntos, o número de agrupamentos formados é menor que o método SPH.

Na formação realizada pelo protocolo LEACH, representada pela Figura 24(c), pode-se notar a presença de alguns nodos sem nenhuma ligação, ou seja, nodos órfãos. Essa presença é resultado da aleatoriedade do método de seleção dos CHs, ocasionando má distribuição dos agrupamentos no ambiente. Porém, devido ao mecanismo de reconfiguração empregado pelo protocolo LEACH, novos CHs são selecionados periodicamente em cada *round*, alternando as relações entre os nodos. Assim, um nodo que ficou órfão em determinado *round*, pode não ficar em outro.

Para avaliar os métodos em relação aos diferentes cenários e densidades, uma série de simulações foi realizada. As figuras 25, 26 e 27 apresentam uma comparação dos resultados obtidos para o número de agrupamentos, profundidade média da rede e número de nodos órfãos, respectivamente. Para cada um dos resultados, foram realizadas simulações com variação na densidade e na disposição dos nodos, assim como nas dimensões do cenário. Enquanto que as linhas representam os resultados médios para as diferentes densidades, os desvios padrões anexados a cada ponto representam a variação desses resultados para as diferentes disposições dos nodos.

Conforme a Figura 25(a), para o cenário de  $50 \times 50$  m<sup>2</sup>, o método SPH apresenta um maior número de agrupamentos para todas as densidades e disposições de nodos. Isso deve-se ao fato que a formação SPH não é baseada em agrupamentos, ou seja, seu objetivo não é for-

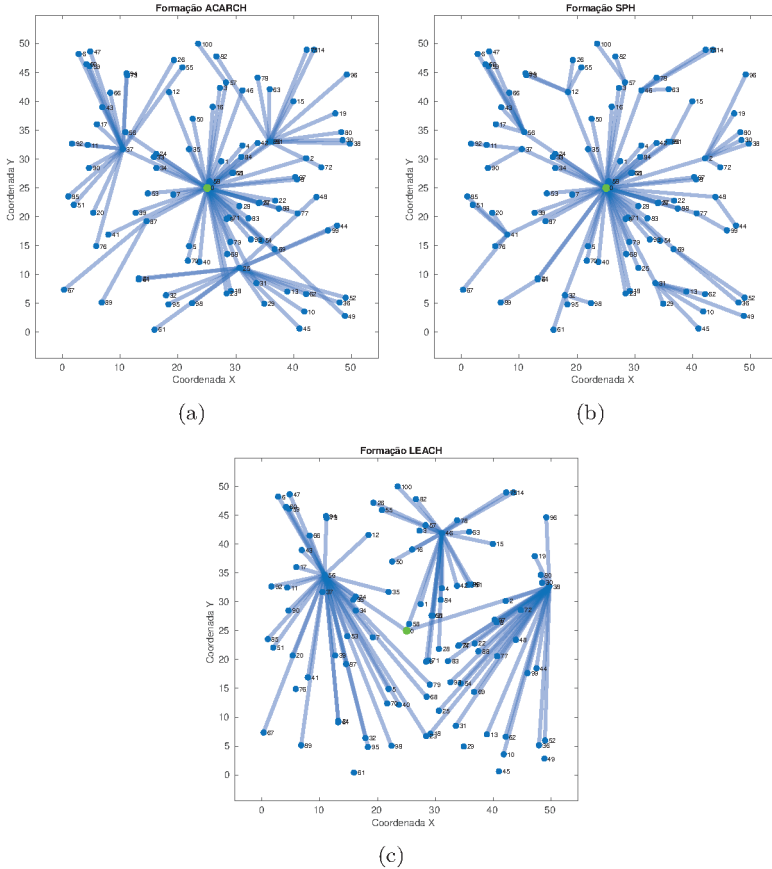


Figura 24: Representação da formação para os métodos (a) ACArch, (b) SPH e (c) LEACH.

mar agrupamentos, e sim uma árvore de menor profundidade. Para o cenário da Figura 25(b), os métodos ACArch e SPH apresentam comportamento similar ao do cenário anterior, porém com maior número de agrupamentos, pois a conectividade nesse cenário é menor. O fator que contribui para que o protocolo LEACH tenha o menor número de agrupamentos para os dois cenários é o valor atribuído ao parâmetro *percentage*, o qual estabelece o número de CHs baseado em uma probabilidade de 5% do número total de nodos.

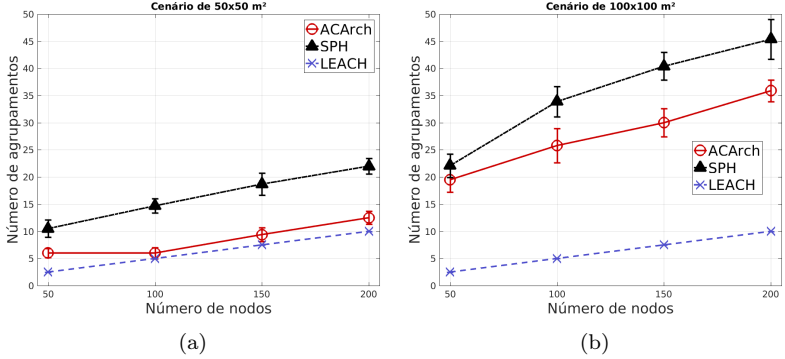


Figura 25: Número médio de agrupamentos formados para os cenários de (a) 50x50 m<sup>2</sup> e (b) 100x100 m<sup>2</sup>.

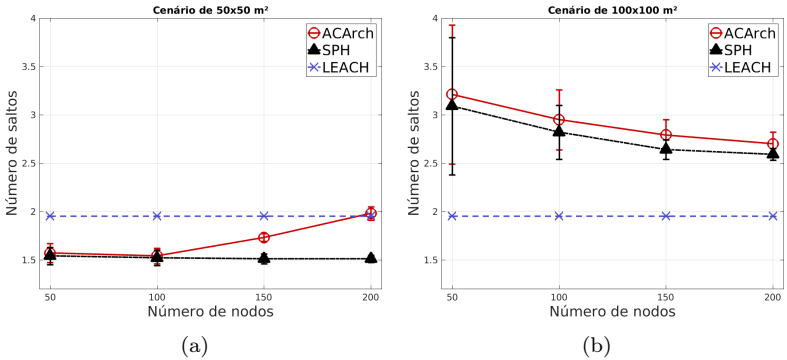


Figura 26: Número médio de saltos entre todos nodos para os cenários de (a) 50x50 m<sup>2</sup> e (b) 100x100 m<sup>2</sup>.

Conforme apresentado na Figura 26(a), enquanto que a profundidade média das árvores formadas pelo método SPH diminui à medida que a densidade de nodos aumenta, as árvores formadas pela ACArch aumentam quando chega na densidade de 101 nodos, pois o limite do tamanho dos agrupamentos é atingido. O comportamento do SPH é justificado pelo limite de distância definido, pois assim, conforme a densidade aumenta, a conectividade da rede também aumenta, permitindo que caminhos menos profundos sejam formados. Para o cenário apresentado na Figura 26(b), a profundidade média das árvores formadas pela ACArch e pelo SPH são maiores que no cenário anterior,

diminuindo conforme o aumento da densidade dos nodos.

Além disso, como é possível observar em ambos cenários, os desvios padrões são maiores quando o ambiente possui baixa densidade, representando alta sensibilidade dos métodos ACArch e SPH à disposição dos nodos. Devido ao fato do protocolo LEACH empregar uma topologia de no máximo dois saltos até a estação base, ele mantém uma profundidade média constante.

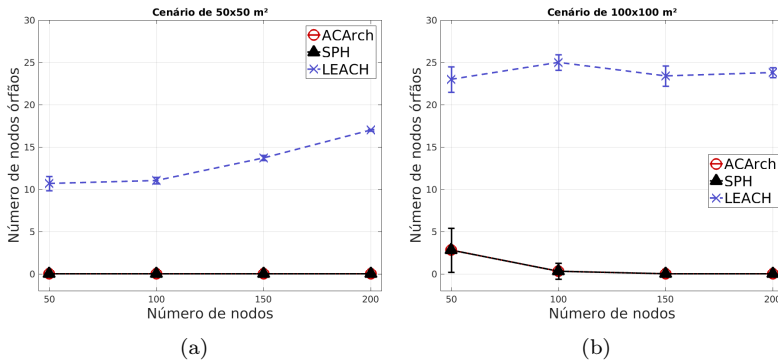


Figura 27: Número médio de nodos órfãos para os cenários de (a) 50x50  $m^2$  e (b) 100x100  $m^2$ .

Como se pode observar na Figura 27, enquanto que as árvores formadas pelos métodos ACArch e SPH não apresentam um número significativo de nodos órfãos, as formadas pelo protocolo LEACH possuem um alto número de nodos órfãos para todas densidades. Esse comportamento é explicado pela aleatoriedade na escolha dos CHs, pois não há como garantir que essas escolhas vão ser bem distribuídas. Como existe rotação dos CHs no protocolo LEACH, os resultados apresentados são as médias obtidas a partir de todos os *rounds* de cada simulação.

### 5.3.2 Análise de Escalabilidade

Nesta seção, os métodos ACArch e SPH são comparados e avaliados em relação às suas respectivas escalabilidades. Conforme o modelo de comunicação CT-SIM, empregado pelos métodos, existem restrições que devem ser satisfeitas para que a rede opere de maneira adequada (Equação 2.5). Basicamente, as restrições garantem que, dentro do intervalo BI definido, todos os nodos consigam transmitir seus paco-

tes à uma determinada taxa de dados. Devido ao protocolo LEACH empregar os esquemas TDMA e CDMA para escalonar e codificar a comunicação, respectivamente, esta seção não faz nenhuma comparação com os outros métodos, porém, uma breve análise de sua escalabilidade é realizada.

Para analisar a escalabilidade dos métodos ACArch e SPH, são considerados os mesmos parâmetros apresentados na Tabela 8 da Seção 4.2. A partir desses parâmetros, foram realizadas diversas simulações com variações na densidade e disposição dos nodos no cenário. A Figura 28 ilustra os somatórios resultantes dos SDs para cada método, considerando as variações impostas. O desvio padrão anexado a cada ponto representa a variação do somatório em relação às diferentes disposições. Também é ilustrado o valor, em segundos, da variável BI que limita a escalabilidade da rede.

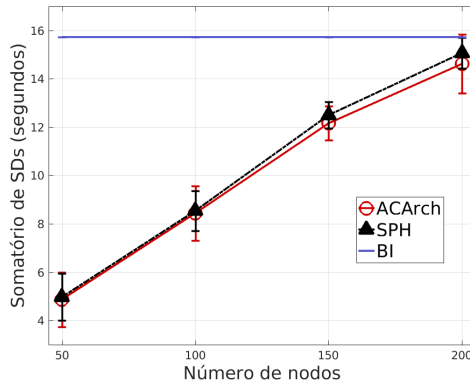


Figura 28: Análise de escalabilidade dos métodos ACArch e SPH.

Conforme se pode observar na Figura 28, o somatório de SDs cresce para ambos os métodos à medida que o número de nodos aumenta, pois uma quantidade maior de tempo é necessária para o escalonamento das mensagens. Os valores médios encontrados são similares, porém, conforme o desvio padrão anexado, o método ACArch apresenta uma maior variação em relação à disposição dos nodos. Isso se deve pelo fato que a ACArch não utiliza o conhecimento global da rede para redução da profundidade, principal fator relacionado à escalabilidade. Mesmo assim, em média, os somatórios resultantes para ACArch são menores que os encontrados pelo método SPH.

Em relação ao protocolo LEACH, o uso da combinação dos es-

quem as TDMA e CDMA contribui para que o número de nodos da rede possa ser maior que os métodos que empregam o modelo CT-SIM. Por exemplo, assumindo os valores dos parâmetros da Tabela 12, temos que, em cada *round*, 100 nodos podem ser escalonados pelo esquema TDMA para cada CH. Com isso, considerando que 5% do total de nodos são CHs, temos que, a cada nodo adicionado na rede, 5 ( $0,05 \times 100$ ) nodos a mais podem ser escalonados.

Tabela 12: Parâmetros de comunicação do protocolo LEACH.

Parâmetros	Descrição	Valor
<i>slotLength</i>	Tempo de duração de um <i>slot</i> TDMA (segundos).	0,2
<i>roundLength</i>	Tempo de duração de cada <i>round</i> (segundos).	20
<i>packet_rate</i>	Taxa de geração de pacotes (pacotes/segundo).	0,05
<i>percentage</i>	Probabilidade inicial de um nodo qualquer se tornar CH.	5%

### 5.3.3 Análise de Desempenho

Após as análises da formação e da escalabilidade da rede, essa seção avalia cada método em relação ao seu comportamento e desempenho durante o monitoramento da rede. Além disso, diversos cenários, empregando densidades e disposições diferentes para os nodos da rede, são utilizados. Para isso, emprega-se as seguintes métricas de avaliação: consumo de energia, atraso fim-a-fim e taxa de sucesso. A justificativa para o emprego dessas métricas é a mesma utilizada na Seção 4.3.3.

Para simulação do desempenho da rede, os parâmetros mostrados na Tabela 13 são utilizados. Como o desempenho da rede é medido na etapa de comunicação, esses parâmetros se referem ela.

Enquanto as figuras 29, 32 e 34 ilustram o comportamento dos métodos ACArch, SPH e LEACH para cada uma das métricas de avaliação, respectivamente, as figuras 30, 33 e 35 apresentam uma comparação entre os métodos para cada métrica. Para as figuras que ilustram o comportamento, são apresentados os resultados de cada um dos 100 nodos, dispostos em um cenário específico de  $50 \times 50 \text{ m}^2$ . O objetivo dessa análise é apresentar as principais características do funcionamento de cada um dos esquemas utilizados.

Conforme a Figura 29, os métodos ACArch e SPH apresentam comportamentos similares, concentrando um maior consumo de ener-

Tabela 13: Parâmetros utilizados na comunicação da rede.

Parâmetro	Descrição	Valores		
		ACArch	SPH	LEACH
<i>sim-time-limit</i>	Tempo de duração da simulação (segundos).	7200	7200	7200
<i>packet_rate</i>	Taxa de geração de pacotes (pacotes/segundo).	0,05	0,05	0,05
<i>beaconOrder</i>	Parâmetro que define a duração de BI através da Equação 2.1.	10	10	NA
<i>macMinBE</i>	Expoente de <i>backoff</i> mínimo, que determina a duração dos <i>backoffs</i> do mecanismo CSMA-CA.	7	7	NA
<i>macMaxBE</i>	Expoente de <i>backoff</i> máximo, que determina a duração dos <i>backoffs</i> do mecanismo CSMA-CA.	8	8	NA
<i>slotLength</i>	Tempo de duração de um <i>slot</i> TDMA (segundos).	NA	NA	0,2
<i>roundLenght</i>	Tempo de duração de um <i>round</i> (segundos).	NA	NA	20

gia nos nodos mais próximos do coordenador PAN, localizado no centro do cenário. Esse comportamento deve-se ao emprego da topologia em agrupamentos em árvore e do protocolo de comunicação CT-SIM, fazendo com que exista um maior tráfego de mensagens nessas proximidades. Além disso, como consequência desse alto tráfego, a disputa pelo acesso ao meio aumenta, contribuindo também no aumento do consumo de energia. A principal diferença entre esses dois métodos está relacionada ao número máximo de CHs selecionados, os quais representam os picos formados nas figuras 29(a) e 29(b).

Ao contrário dos métodos anteriores, o comportamento do protocolo LEACH ilustrado na Figura 29(c) não apresenta variações significativas de consumo energético entre os nodos. Isso deve-se ao fato do LEACH empregar um mecanismo de rotação de CHs, distribuindo melhor esse consumo entre os nodos da rede.

Contudo, para avaliar a quantidade total de energia gasta pela rede em diversos cenários, uma análise do consumo médio entre todos nodos é feita considerando densidades, disposições e tamanhos do ambiente diferentes. A Figura 30 mostra uma análise comparativa da energia consumida entre os três métodos para dois ambientes de tamanhos distintos, um de  $50 \times 50 \text{ m}^2$  e outro de  $100 \times 100 \text{ m}^2$ .

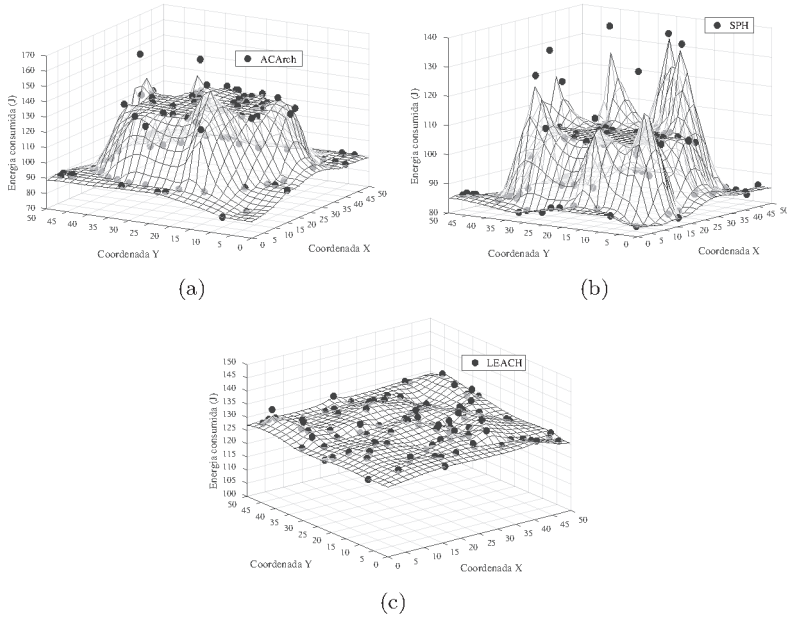


Figura 29: Energia consumida pelos nodos para os métodos (a) ACArch, (b) SPH e (c) LEACH.

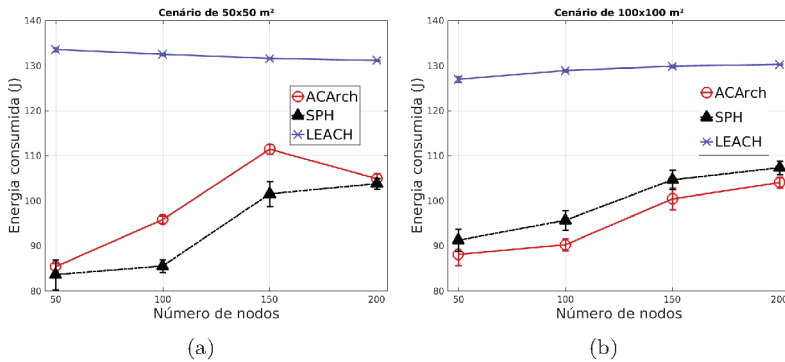


Figura 30: Análise do consumo médio de energia com variação de nodos para um cenário de (a)  $50 \times 50 \text{ m}^2$  e (b)  $100 \times 100 \text{ m}^2$ .

Como mostrado na Figura 30, o protocolo LEACH possui um maior gasto de energia que os outros métodos. Isso porque, em cada



*round*, a rede precisa ser reconfigurada, exigindo maior consumo energético. Em relação a ACArch e ao SPH, o consumo aumenta à medida em que a rede se torna mais densa, sendo que, para o cenário de  $50 \times 50 m^2$  o SPH consome menos energia e para o cenário de  $100 \times 100 m^2$ , consome mais. No primeiro caso, isso se deve pelo fato da ACArch limitar os tamanhos dos agrupamentos e, conseqüentemente, gerar redes mais profundas. No segundo cenário, apesar de formar redes de menor profundidade, o método SPH gera mais agrupamentos com um número desbalanceado de nodos, ou seja, alguns agrupamentos possuem muito mais nodos que outros, causando uma certa sobrecarga na rede, representada pelos picos da Figura 29(b).

Para ilustrar melhor como a energia média global da rede é consumida pelos três métodos durante o período simulado, considere o cenário específico da Figura 31, com dimensão de  $100 \times 100 m^2$  e 100 nodos.

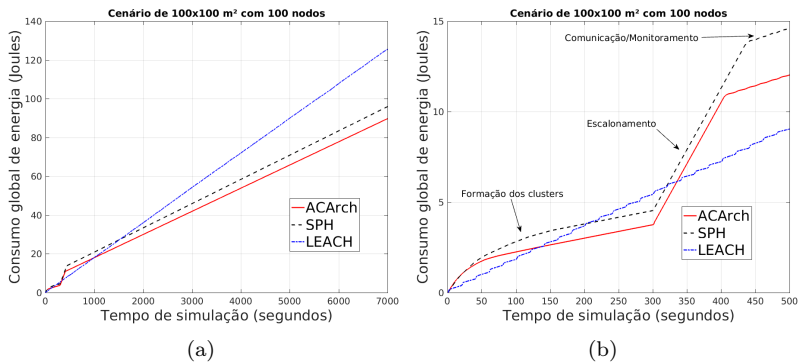


Figura 31: Análise do consumo global médio de energia pelo tempo decorrido de um cenário específico (a) para toda simulação e (b) para o período de configuração.

Como se pode notar na Figura 31(a), mesmo iniciando com menor consumo de energia, o protocolo LEACH ultrapassa os outros métodos durante a simulação devido a sua reconfiguração periódica. A Figura 31(b) ilustra as etapas que as redes formadas pela ACArch e pelo SPH passam até começar a monitoração do ambiente. Durante a formação e escalonamento dos agrupamentos, o método SPH consome mais energia, pois possui mais agrupamentos a serem formados e escalonados, respectivamente. As reconfigurações periódicas do protocolo LEACH estão representadas pelos degraus da Figura 31(b).

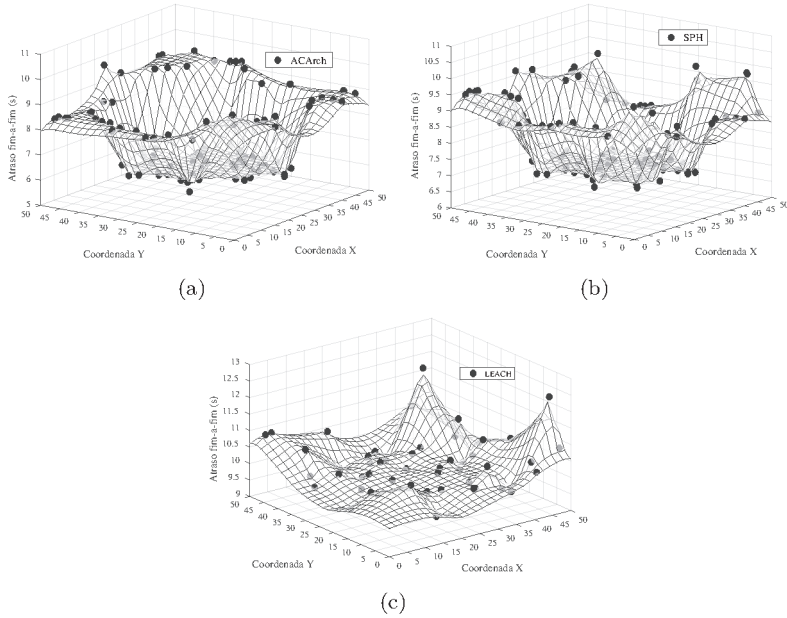


Figura 32: Atraso fim-a-fim médio das mensagens de cada nodo para os métodos (a) ACArch, (b) SPH e (c) LEACH.

A Figura 32 apresenta, para um cenário específico, o atraso médio sofrido entre todas as mensagens enviadas ao coordenador PAN por cada nodo. Como se pode notar, a ACArch e o método SPH apresentam maiores atrasos na mensagens para os nodos mais distantes do coordenador PAN, pois os mesmos se encontram em profundidades maiores. Embora seja menos perceptivo devido à rotação de CHs, o protocolo LEACH também apresenta um comportamento similar aos dos métodos anteriores, pois os nodos mais distantes do centro possuem maiores chances de ficarem órfãos e não conseguirem enviar as mensagens no *round* corrente.

Na Figura 33 são apresentados os atrasos médios sofridos entre todos os nodos para cenários com variação nas dimensões, na densidade e na disposição dos nodos. Como se pode ver, para o cenário de  $50 \times 50 \text{ m}^2$ , enquanto que no protocolo LEACH o atraso médio diminui à medida que a densidade aumenta, a ACArch aumenta na medida que a densidade também aumenta. No LEACH, isso se deve ao aumento no número de CHs, sendo que, devido ao emprego da técnica CDMA,

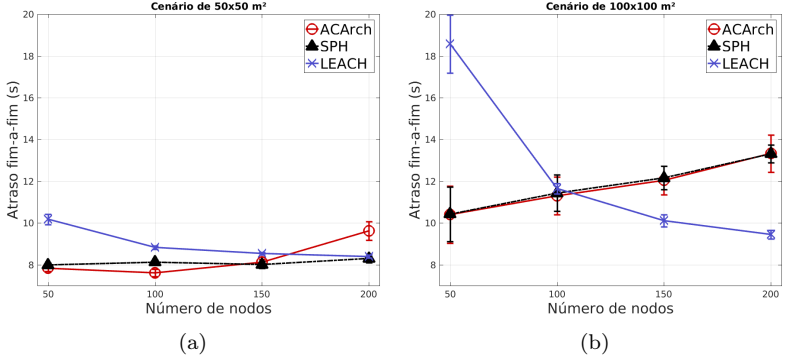


Figura 33: Análise do atraso fim-a-fim com variação de nodos para um cenário de (a)  $50 \times 50 \text{ m}^2$  e (b)  $100 \times 100 \text{ m}^2$ .

a comunicação dos agrupamentos acontece em paralelo. Na ACArch, o aumento do atraso ocorre devido ao aumento na profundidade da árvore (Figura 26), a qual é consequência na limitação do tamanho do agrupamento. Assim como a profundidade das árvores formadas pelo método SPH, o atraso médio sofrido pelas mensagens também se mantém constante.

Para o cenário da Figura 33(b), tem-se um comportamento parecido com o do cenário anterior, com as redes do protocolo LEACH diminuindo seu atraso médio conforme a densidade aumenta e as redes dos métodos ACArch e SPH aumentando seus atrasos médios conforme a densidade e a profundidade da rede.

Como mostrado na Figura 34, todos os métodos apresentam taxas de sucesso próximas de 100%. Isso é justificado pela adoção do modelo de disco unitário, que garante uma transmissão bem sucedida toda vez que os nodos estão dentro de alcance. Dessa forma, as principais fontes de perda de mensagens são consequência dos nodos órfãos e do mecanismo CSMA-CA que descarta as mensagens que não consegue transmitir.

Neste contexto, conforme as figuras 34(a) e (b), existem alguns vales na taxa de sucesso entre os nodos. Esses vales representam os CHs selecionados, os quais sofrem maiores perdas devido à disputa de acesso ao meio através do CSMA-CA para transmissão das suas mensagens e de seus respectivos filhos. No LEACH, pelo motivo de empregar uma esquema aleatório, o comportamento da taxa de sucesso não mantém um padrão.

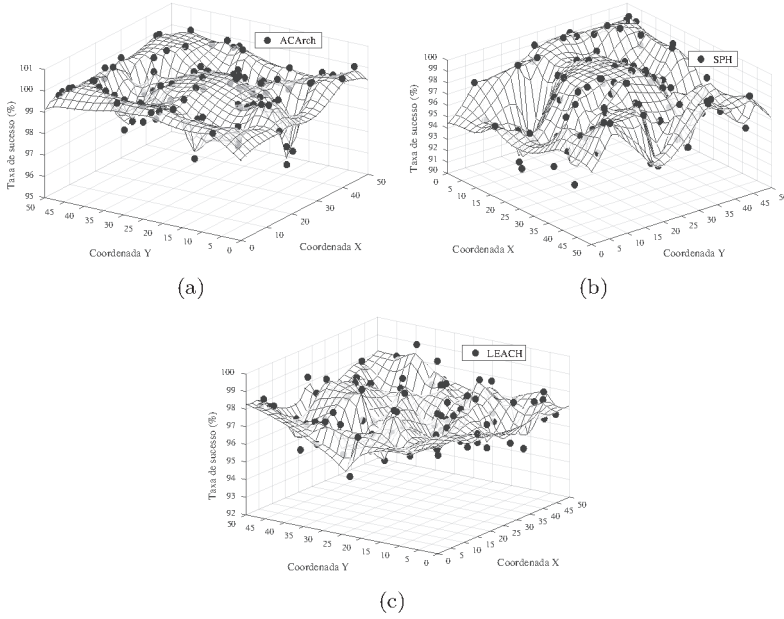


Figura 34: Taxa de sucesso média de cada nodo para os métodos (a) ACArch, (b) SPH e (c) LEACH.

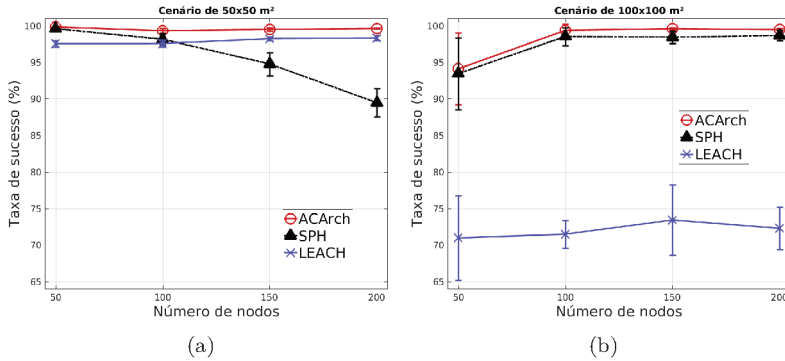


Figura 35: Análise da taxa de sucesso com variação de nodos para um cenário de (a)  $50 \times 50 \text{ m}^2$  e (b)  $100 \times 100 \text{ m}^2$ .

A Figura 35 faz uma comparação da taxa de sucesso para diferentes cenários, densidades e disposições de nodos. Conforme os resultados

apresentados no cenário da Figura 35(a), enquanto a taxa de sucesso da ACArch e do LEACH permanecem quase constantes, o método SPH apresenta uma queda significativa conforme a densidade de nodos aumenta. Essa queda se deve pelo fato de que o SPH não controla o tamanho dos agrupamentos, sobrecarregando o mecanismo CSMA-CA e gerando perdas.

De maneira diferente do cenário anterior, o cenário dos resultados apresentados pela Figura 35(b) não possui densidades que sobrecarreguem o método SPH. Por esse motivo, a taxa de sucesso permanece constante para todas densidades, exceto para densidade de 50 nodos, pois, devido à baixa conectividade desse cenário, diversos nodos ficam órfãos (27). Pelo mesmo motivo anterior, a seleção aleatória de CHs faz com que o protocolo LEACH apresente maiores perdas de mensagens para esse cenário.

## 5.4 CONSIDERAÇÕES FINAIS

Este capítulo descreveu a ACArch, uma arquitetura para RSSF baseada na topologia em agrupamentos em árvore. Foram apresentados todos os métodos, conceitos e configurações utilizados pela ACArch para formação, escalonamento e comunicação da rede. Além disso, uma avaliação da mesma foi realizada através da sua comparação com o método SPH, descrito anteriormente, e o protocolo LEACH, um dos protocolos de RSSF mais conhecido da literatura.

A avaliação mostrou que, embora não tenha conhecimento global e mecanismos de reconfiguração da rede, o desempenho apresentado pelas redes que empregaram a ACArch foi satisfatório para diversas condições, mostrando que a formação dos agrupamentos de forma distribuída e baseada em critérios de vizinhança e força de sinal possui potencial para diversas aplicações de RSSF. Entretanto, algumas melhorias são necessárias para aplicação em cenários reais, como, por exemplo, implementação de uma reconfiguração periódica da rede e otimização do mecanismo CSMA-CA.

Por fim, além de apresentar uma nova solução para RSSF, o estudo realizado mostrou uma relação entre diversos aspectos de formação e funcionamento das redes com o desempenho atingido por elas, possibilitando uma melhor compreensão do assunto.



## 6 CONCLUSÕES E TRABALHOS FUTUROS

Este capítulo apresenta uma revisão geral do trabalho, expondo os principais objetivos e contribuições, assim como os métodos desenvolvidos e os resultados que foram obtidos. Além disso, também são sugeridas algumas propostas para trabalhos futuros.

### 6.1 VISÃO GERAL DO TRABALHO

Devido à constante evolução tecnológica, as RSSF vem sendo uma alternativa de baixo custo para aplicação nas mais variadas áreas, como por exemplo: aplicações industriais, aplicações militares, monitoramento ambiental, saúde e etc. Entre essas aplicações, estão aquelas de larga escala, que exigem que uma área relativamente grande seja monitorada por diversos dispositivos sensores. Neste contexto, a topologia mais adequada para RSSF em larga escala é em agrupamentos em árvore.

Pelo motivo de possuir um caráter multidisciplinar, as RSSF apresentam uma série de desafios. Entre eles estão aqueles relacionados a formação da rede, escalonamento de mensagens e comunicação. Com base nisso, o Capítulo 2 apresentou as principais especificações dos padrões mais relevantes para as RSSF, o IEEE 802.15.4 e o ZigBee. Além disso, também foi apresentado o modelo CT-SIM, o qual especifica esquemas de escalonamento e comunicação baseados nos padrões IEEE 802.15.4/ZigBee.

No Capítulo 3, foi realizado um levantamento bibliográfico sobre os protocolos de formação de topologias em agrupamentos em árvore e alguns outros relacionados à comunicação baseada no padrão IEEE 802.15.4. Com base nisso, no Capítulo 4, foram propostos três esquemas de formação integrados com o modelo CT-SIM, os quais assumem uma execução centralizada e com conhecimento global da rede. Esses esquemas foram comparados e analisados sobre diversos aspectos de formação e desempenho.

Devido à natureza intrínseca das RSSF, tanto a formação centralizada quanto a obtenção do conhecimento global da rede, são praticamente inviáveis, principalmente em redes de larga escala. Por esse motivo, a partir das análises realizadas no Capítulo 4, uma arquitetura de formação e gerenciamento de topologia é proposta no Capítulo 5. Essa arquitetura é composta por métodos de descobrimento de vi-

zinhança, associação de filhos e seleção de CHs, os quais operam de forma distribuída e visam a construção de uma rede adaptável a diferentes condições em relação à densidade de nodos e dimensões do cenário.

Os resultados obtidos nos capítulos 4 e 5 mostraram que formação da rede possui papel fundamental no desempenho de sua operação. Por exemplo, enquanto a profundidade da rede interfere diretamente nos atrasos fim-a-fim sofridos pelas mensagens, o número e o tamanho dos agrupamentos interferem no consumo de energia. Além disso, devido ao emprego do modelo de disco unitário, o número de nodos órfãos é o principal fator de perda de dados, diminuindo a taxa de sucesso e a cobertura da rede.

Um dos principais fatores que possibilitou a ACArch possuir um melhor desempenho em diversos cenários em relação aos métodos avaliados, foi o esquema empregado para seleção dos CHs. Mesmo sem possuir conhecimento global da rede, a arquitetura se mostrou capaz de formar redes de baixa profundidade, poucos agrupamentos e probabilidade reduzida de deixar nodos órfãos.

## 6.2 PRINCIPAIS CONTRIBUIÇÕES

Este trabalho apresentou diversos métodos de formação relacionados a formação de topologia em árvores em RSSF e baseados nos padrões IEEE 802.15.4/ZigBee. Os métodos foram implementados em um simulador e analisados sobre diferentes aspectos, possibilitando uma melhor compreensão do comportamento das redes.

Além disso, foi proposta a arquitetura ACArch, a qual integra métodos de descobrimento de vizinhança, formação de agrupamentos, escalonamento de mensagens e comunicação. O método foi avaliado através do simulador Castalia e os resultados obtidos mostraram que a arquitetura mantém um bom desempenho para diversas condições.

## 6.3 TRABALHOS FUTUROS

Através do estudo realizado neste trabalho, foi possível perceber o potencial que a ACArch possui. Entretanto, existem diversos ajustes que podem melhorar ainda mais o seu desempenho, entre eles, estão os seguintes:

- Implementar um esquema determinístico para formação dos agru-



pamentos, pois nos métodos propostos neste trabalho é necessário superestimar um tempo limite para formação da rede;

- Implementar a reconfiguração periódica da rede, pois devido aos CHs escolhidos consumirem mais energia que os outros nodos, uma reconfiguração adequada pode prolongar significativamente o tempo de vida da rede;
- Implementar um mecanismo CSMA-CA independente para cada agrupamento, ou seja, que os CHs possam configurar automaticamente seus parâmetros conforme a necessidade;
- Empregar mecanismos de agregação ou fusão de dados.

A fim de simplificar as implementações e análises realizadas, no Castalia foi assumido um modelo de rádio (modelo de propagação de sinal) denominado disco unitário, o qual pode ser considerado simples e que não modela adequadamente todas as características de um cenário real. Por esse motivo, como forma de ampliar a análise realizada e garantir o funcionamento adequado dos métodos para cenários reais, sugere-se assumir um modelo de rádio real, que considere canais de comunicação não simétricos e que perdas de pacotes possam ocorrer mesmo dentro do alcance do rádio.



## REFERÊNCIAS

- AKYILDIZ, I.; SU, W.; SANKARASUBRAMANIAM, Y.; CAYIRCI, E. Wireless sensor networks: a survey. *Computer Networks*, v. 38, p. 393–422, 2002.
- AKYILDIZ, I. F.; VURAN, M. C. *Wireless sensor networks*. [S.l.]: John Wiley & Sons, 2010.
- ALBEROLA, R. de P.; PESCH, D. Duty cycle learning algorithm (dcla) for ieee 802.15. 4 beacon-enabled wireless sensor networks. *Ad Hoc Networks*, Elsevier, v. 10, n. 4, p. 664–679, 2012.
- ANASTASI, G.; CONTI, M.; FRANCESCO, M. D. A comprehensive analysis of the mac unreliability problem in ieee 802.15. 4 wireless sensor networks. *IEEE Transactions on Industrial Informatics*, IEEE, v. 7, n. 1, p. 52–65, 2011.
- BECHKIT, W.; KOUDIL, M.; CHALLAL, Y.; BOUABDALLAH, A.; SOUICI, B.; BENATCHBA, K. A new weighted shortest path tree for convergecast traffic routing in wsn. In: IEEE. *Computers and Communications (ISCC), 2012 IEEE Symposium on*. [S.l.], 2012. p. 000187–000192.
- BHOLOWALIA, P.; KUMAR, A. Ebk-means: A clustering technique based on elbow method and k-means in wsn. *International Journal of Computer Applications*, Citeseer, v. 105, n. 9, 2014.
- BOULIS, A. et al. Castalia: A simulator for wireless sensor networks and body area networks. *NICTA: National ICT Australia*, v. 83, 2011.
- CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. *Introduction to algorithms*. [S.l.]: MIT press, 2009.
- DARGIE, W.; POELLABAUER, C. *Fundamentals of wireless sensor networks: theory and practice*. [S.l.]: John Wiley & Sons, 2010.
- DING, X.; TIAN, Y.; YU, Y. A real-time big data gathering algorithm based on indoor wireless sensor networks for risk analysis of industrial operations. *IEEE transactions on industrial informatics*, IEEE, v. 12, n. 3, p. 1232–1242, 2016.
- FARAHANI, S. *ZigBee wireless networks and transceivers*. [S.l.]: Newnes, 2011.

FELSKE, M. S.; MONTEZ, C.; PINTO, A. S. R.; VASQUES, F.; PORTUGAL, P. GLHOVE: A framework for uniform coverage monitoring using cluster-tree wireless sensor networks. In: *2013 IEEE 18th Conference on Emerging Technologies Factory Automation (ETFA)*. [S.l.: s.n.], 2013. p. 1–8. ISSN 1946-0740.

GAGARIN, A.; HUSSAIN, S.; YANG, L. T. Distributed hierarchical search for balanced energy consumption routing spanning trees in wireless sensor networks. *Journal of Parallel and Distributed Computing*, Elsevier, v. 70, n. 9, p. 975–982, 2010.

GILANI, M. H. S.; SARRAFI, I.; ABBASPOUR, M. An adaptive csma/tdma hybrid mac for energy and throughput improvement of wireless sensor networks. *Ad Hoc Networks*, Elsevier, v. 11, n. 4, p. 1297–1304, 2013.

HEINZELMAN, W. R.; CHANDRAKASAN, A.; BALAKRISHNAN, H. Energy-efficient communication protocol for wireless microsensor networks. In: *IEEE. System sciences, 2000. Proc. of the 33rd annual Hawaii international conference on*. [S.l.], 2000. p. 10–pp.

HUANG, G.; LI, X.; HE, J. Dynamic minimal spanning tree routing protocol for large wireless sensor networks. In: *IEEE. 2006 1ST IEEE Conference on Industrial Electronics and Applications*. [S.l.], 2006. p. 1–5.

IEEE, L. S. C. IEEE standard for local and metropolitan area networks–part 15.4: Low-rate wireless personal area networks (LR-WPANs). *IEEE Computer Society*, 2011.

KHANAFAER, M.; GUENNOUN, M.; MOUFTAH, H. T. A survey of beacon-enabled ieee 802.15. 4 mac protocols in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, IEEE, v. 16, n. 2, p. 856–876, 2014.

KUROSE, J. F.; ROSS, K. W. *Computer networking: a top-down approach: international edition*. [S.l.]: Pearson Higher Ed, 2013.

LACHOWSKI, R.; PELLENZ, M. E.; PENNA, M. C.; JAMHOUR, E.; SOUZA, R. D. An efficient distributed algorithm for constructing spanning trees in wireless sensor networks. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 15, n. 1, p. 1518–1536, 2015.

LEAO, E.; MONTEZ, C.; MORAES, R.; PORTUGAL, P.; VASQUES, F. Superframe duration allocation schemes to improve the throughput

of cluster-tree wireless sensor networks. *Sensors*, v. 17, n. 2, 2017. ISSN 1424-8220.

LEAO, E.; MORAES, R.; MONTEZ, C.; PORTUGAL, P.; VASQUES, F. CT-SIM: A simulation model for wide-scale cluster-tree networks based on the IEEE 802.15. 4 and ZigBee standards. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 13, n. 3, p. 1550147717698471, 2017.

LEÃO, E.; VASQUES, F.; PORTUGAL, P.; MORAES, R.; MONTEZ, C. An allocation scheme for ieee 802.15. 4-zigbee cluster-tree networks. In: IEEE. *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*. [S.l.], 2016. p. 4639–4644.

LEVIS, K. et al. RSSI is under appreciated. In: *Proceedings of the Third Workshop on Embedded Networked Sensors, Cambridge, MA, USA*. [S.l.: s.n.], 2006. v. 3031, p. 239242.

LI, M.; LI, Z.; VASILAKOS, A. V. A survey on topology control in wireless sensor networks: Taxonomy, comparative study, and open issues. *Proceedings of the IEEE*, IEEE, v. 101, n. 12, p. 2538–2557, 2013.

MCGLYNN, M. J.; BORBASH, S. A. Birthday protocols for low energy deployment and flexible neighbor discovery in ad hoc wireless networks. In: ACM. *Proc. of the 2nd ACM Int. Symp. on Mobile ad hoc networking & computing*. [S.l.], 2001. p. 137–145.

MURUGANATHAN, S. D.; MA, D. C.; BHASIN, R. I.; FAPOJUWO, A. O. A centralized energy-efficient routing protocol for wireless sensor networks. *IEEE Communications Magazine*, IEEE, v. 43, n. 3, p. S8–13, 2005.

OLIVEIRA, L. M.; RODRIGUES, J. J. Wireless sensor networks: A survey on environmental monitoring. *JCM*, v. 6, n. 2, p. 143–151, 2011.

PATEL, N.; PATEL, D. K. A survey on: Enhancement of minimum spanning tree. *Journal of Research and Application*, v. 5, n. 1, p. 06–10, 2015.

SARAVANAN, M.; MADHESWARAN, M. A spanning tree for enhanced cluster based routing in wireless sensor network. *World Academy of Science, Engineering and Technology, International*

*Journal of Computer, Electrical, Automation, Control and Information Engineering*, v. 9, n. 9, p. 2102–2111, 2016.

SHU, L.; MUKHERJEE, M.; WU, X. Toxic gas boundary area detection in large-scale petrochemical plants with industrial wireless sensor networks. *IEEE Communications Magazine*, IEEE, v. 54, n. 10, p. 22–28, 2016.

SORO, S.; HEINZELMAN, W. B. Cluster head election techniques for coverage preservation in wireless sensor networks. *Ad Hoc Networks*, Elsevier, v. 7, n. 5, p. 955–972, 2009.

SPECIFICATION, Z. Document 053474r17. *ZigBee Alliance*, 2008.

SUN, W.; YANG, Z.; ZHANG, X.; LIU, Y. Energy-efficient neighbor discovery in mobile ad hoc and wireless sensor networks: A survey. *IEEE Communications Surveys & Tutorials*, IEEE, v. 16, n. 3, p. 1448–1459, 2014.

XU, G.; SHEN, W.; WANG, X. Applications of wireless sensor networks in marine environment monitoring: A survey. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 14, n. 9, p. 16932–16954, 2014.

YOUNIS, O.; FAHMY, S. Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach. In: IEEE. *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. [S.l.], 2004. v. 1.

ZHAO, G. Wireless sensor networks for industrial process monitoring and control: A survey. *Network Protocols and Algorithms*, v. 3, n. 1, p. 46–63, 2011.