

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

ADONIAS PINHEIRO PIRES

**UM ALGORITMO TEMPORIZADO DE
CLUSTERIZAÇÃO E ROTEAMENTO, COM
EFICIÊNCIA ENERGÉTICA PARA REDES DE
SENSORES SEM FIO**

UFPA/ICEN/PPGCC

Campus Universitário do Guamá

Belém-Pará-Brasil

2012

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

ADONIAS PINHEIRO PIRES

**UM ALGORITMO TEMPORIZADO DE CLUSTERIZAÇÃO
E ROTEAMENTO, COM EFICIÊNCIA ENERGÉTICA PARA
REDES DE SENSORES SEM FIO**

Dissertação submetida à Banca Examinadora do Programa de Pós-Graduação em Ciência da Computação da UFPA para a obtenção do Grau de Mestre em Ciência da Computação.

UFPA/ICEN/PPGCC

Campus Universitário do Guamá

Belém-Pará-Brasil

2012

dedicatoria

A minha esposa e minha filha

Agradecimentos

Dedico meus sinceros agradecimentos para:

- o professor Doutor Dionne, pela orientação e incentivo para o desenvolvimento de todo o trabalho científico realizado.
- ao professor Doutor Eduardo, pela orientação e incentivo para o desenvolvimento de todo o trabalho científico realizado.
- a equipe do laboratório GERCOM, em especial pelos colegas Claudio, Denis, Rodrigo, Vagner, Romulo e André pela ajuda em diversos momentos.
- aos meus familiares pelo apoio durante o curso.
- a todos aqueles que de alguma forma contribuíram para realização desse trabalho.

Resumo

Resumo da Dissertação apresentada à UFPA como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Um Algoritmo Temporizado de Clusterização e Roteamento, com Eficiência Energética para Redes de Sensores Sem Fio

Orientador: Dr. Dionne Cavalcante Monteiro

Co-orientador:

Palavras-chave: RSSF; clusters; hot spot; eficiência energética; fuzzy.

A arquitetura de uma Rede de Sensores Sem Fio (RSSF) em *clusters* provê eficiência energética para a rede. Entretanto, aspectos relacionados à eleição de *cluster heads*, formação de *clusters* e *hot spot* devem ser considerados no projeto de uma RSSF clusterizada. Esses aspectos se tornam problemas para a rede se não alinhados as premissas de eficiência energética das RSSF. Diversos protocolos foram propostos com o objetivo de prover eficiência energética para uma RSSF, porém, não constituem soluções completas, ou seja, que considerem todos os aspectos inerentes à eficiência energética em uma RSSF clusterizada. Nesse trabalho, é proposto um algoritmo denominado Algoritmo Temporizado de Clusterização e Roteamento (ATCR) que objetiva fornecer eficiência energética para uma RSSF através de uma solução híbrida de clusterização e roteamento integrados com mecanismos de construção de *clusters* desiguais e otimização multicritério utilizando lógica *fuzzy*. Modelos de simulação foram construídos com o objetivo de avaliar a desempenho do ATCR e os resultados demonstraram que o ATCR possui mais eficiência energética que os demais protocolos, sendo adequado para implantações de RSSF densas e de larga escala.

Abstract

Abstract of Dissertation presented to UFPA as a partial fulfillment of the requirements for the degree of Master in Computer Science.

A Timed Clustering and Routing Algorithm with Energy Efficiency for Wireless Sensor Networks

Advisor: Dr. Dionne Cavalcante Monteiro

Co-advisor:

Key words: WNS; clusters; hot spot; energy efficiency; fuzzy;

The architecture of a Wireless Sensor Network (WSN) in clusters provides more energy efficiency for the network compared with the flat architecture. However, issues related to election of cluster heads, clustering and hot spot should be considered when designing a WSN clustered. These aspects become problems for the network when not aligned the assumptions of energy efficiency of WSN. Several protocols have been proposed in order to provide energy efficiency for a WSN, but are not complete solutions, that consider all aspects of the energy efficiency in a clustered WSN. In this work, we propose an algorithm called Timed Clustering and Routing Algorithm (ATCR) that aims to provide energy efficiency for a WSN, through a hybrid solution of clustering and routing integrated with building mechanisms of unequal clusters and multi-criteria optimization using fuzzy logic. Simulation models were built in order to evaluate the performance of ATCR and the results showed that the TCRA has more energy efficiency than other protocols and is suitable for WSN deployments dense and large scale.

Sumário

Lista de Abreviaturas	p. viii
Lista de Figuras	p. x
Lista de Tabelas	p. 1
1 Introdução	p. 2
1.1 Visão geral	p. 2
1.2 Motivação e desafios	p. 2
1.3 Objetivos	p. 4
1.4 Justificativa.....	p. 4
1.5 Organização do texto	p. 4
2 Referencial Teórico	p. 6
2.1 Visão Geral	p. 6
2.2 Tipos e Aplicações de RSSF	p. 8
2.3 Restrições e Desafios em RSSF	p. 9
2.4 Arcabouço de Comunicação	p. 10
2.5 Padrões e Especificações para RSSFs	p. 11
2.6 Roteamento em RSSFs	p. 12
2.7 Clusterização e Roteamento Hierárquico em RSSFs	p. 14
3 Trabalhos Relacionados	p. 16
3.1 Clusterização hierárquica adaptativa com baixa energia	p. 16
3.2 Um protocolo de roteamento baseado em clusters desiguais	p. 18
3.3 Protocolo de roteamento e clusterização híbridos	p. 19
3.4 Formação de cluster com eficiência energética	p. 20
3.5 Algoritmo de clusterização desigual baseado em localização	p. 21

3.6	Esquema de clusterização e roteamento ciente de energia	p. 22
3.7	Algoritmo de clusterização hierárquico e temporizado com eficiência energética	p. 22
3.8	Algoritmo de clusterização LEACH-ERE	p. 23
3.9	Análise sobre os Trabalhos Relacionados	p. 24
4	Desafios e Proposta	p. 27
4.1	Descrição dos Desafios	p. 27
4.1.1	Eleição do Cluster Head	p. 27
4.1.2	Formação de Cluster	p. 28
4.1.3	Hot Spot	p. 29
4.2	Modelo da Rede	p. 30
4.3	Algoritmo Temporizado de Clusterização e Roteamento	p. 31
4.3.1	Algoritmo de Clusterização.....	p. 32
4.3.2	Otimização MultiCritério Utilizando Lógica Fuzzy.....	p. 35
4.3.3	Análise e Corretude do Algoritmo.....	p. 39
4.3.4	Considerações Finais	p. 40
5	Avaliação de desempenho	p. 41
5.1	Metodologia de Avaliação.....	p. 41
5.2	Descrição do Cenário e Análise dos Resultados	p. 43
6	Conclusões	p. 52
	Referências	p. 54
	Apêndice A – Dados de Simulação	p. 58

Lista de Abreviaturas

MEMS	Sistemas Microeletromecânicos
IOT	Internet of Things
RFID	Radio-Frequency Identification
BS	Base Station
WMSN	Wireless Multimedia Sensor Networks
IWSNs	Industrial Wireless Sensor Networks
WBANs	Wireless Body Area Networks
WSAN	Wireless Sensor and Actuator Network
MSNs	Mobile Sensor Networks
VSNs	Vehicular Sensor Networks
QoS	Quality of Service
WLAN	Wireless Local Area Network
FFD	full function device
RFD	reduced function device
MAC	Media Access Control
AODV	Ad hoc On Demand Distance Vector routing algorithm
TDMA	Time Division Multiple Access
CSMA/CA	Carrier sense multiple access with collision avoidance
HART	Highway Addressable Remote Transducer
CH	Cluster Head
CM	Cluster Members
LEACH	Low Energy Adaptive Clustering Hierarchy
UCR	Unequal Cluster-based Routing
EEUC	Energy-Efficient Unequal Clustering
HCR	Hybrid Clustering and Rounting
EECF	Energy-Efficient Cluster Formation
RCRA	Relative Cluster head Rank Advertisement
LUCA	Location-based Unequal Clustering Algorithm
PARC	Power Aware Routing and Clustering Scheme
RREQ	Route Request
RREP	Router Reply
EEBHC	Energy Efficient Backoff Hierarchical Clustering Algorithm
ERE	Expected residual energy
FIS	Fuzzy inference system

RSSI	Received Signal Strength Indicator
LQI	Link quality indicator
TSFS	Takagi and Sugeno fuzzy system
IC	Intervalo de Confiança
E_r	Energia Residual
GERCOM	Research Group on Computer Networks and Multimedia Communication
HNA	Half of the Nodes Alive
PDR	Packet delivery rate

Lista de Figuras

Figura 1	Organização interna de um nodo sensor.	6
Figura 2	Modelo de uma RSSF.	7
Figura 3	Pilha de protocolos de uma RSSF.	11
Figura 4	Estrutura de uma RSSF baseada em clusters.	15
Figura 5	Organização da rede com LEACH	17
Figura 6	Problema de eleição de CHs	28
Figura 7	Problema de otimização na formação de clusters	29
Figura 8	Problema de Hot Spot	30
Figura 9	Linha do tempo da operação do protocolo	31
Figura 10	Função de R_c	33
Figura 11	Operação do ATCR	34
Figura 12	Organização do sistema fuzzy	35
Figura 13	Funções de pertinência	37
Figura 14	Média da quantidade de clusters formados	44
Figura 15	Nodos vivos por rodadas	45

Figura 16	Desvio padrão da energia residual	46
Figura 17	Energia residual	47
Figura 18	Metade dos Nodos Vivos	47
Figura 19	Mapa de energia da rede	48
Figura 20	Latência	49
Figura 21	Overhead de controle	49
Figura 22	Taxa de pacotes entregues	51

Lista de Tabelas

Tabela 1	Comparação funcional dos protocolos	26
Tabela 2	Regras de inferência fuzzy	38
Tabela 3	Parâmetros do cenário e do algoritmo	43
Tabela 4	Intervalos de confiança para 10 repetições	58
Tabela 5	Intervalos de confiança para 50 repetições	59

1 Introdução

1.1 Visão geral

Através da evolução e rápida convergência da microeletrônica, comunicação sem fio, micro sistemas mecânicos (*Sistemas Microeletromecânicos* - MEMS), e a necessidade de monitorar e eventualmente controlar um ambiente eletromecânico, com baixo custo, levou ao surgimento das redes de sensores sem fio (RSSF), sendo ela considerada uma vertente da computação ubíqua e um tipo especial de rede *ad hoc* que impõe novos desafios e oportunidades de pesquisa.

Uma RSSF consiste de vários nodos sensores que executam aplicações individuais e se comunicam usando enlaces de comunicação sem fio. Um conjunto de nodos sensores pode ser distribuído em um ambiente para monitorar e coletar dados sobre grandezas físicas ou ambientais tais como temperatura, pressão, som, vibração, poluição e entre outros, e enviar esses dados a um nodo centralizador denominado estação base (BS) ou distribuir os dados permitindo assim uma tomada de decisão descentralizada.

Uma das vantagens das RSSFs é a sua habilidade de operar sozinha em ambientes inóspitos onde o monitoramento humano é arriscado, ineficiente e algumas vezes impraticável. Portanto, os nodos sensores podem ser implantados de uma maneira aleatória na área de interesse através de uma forma não controlada e, assim, formar uma rede de maneira *ad-hoc* [1]. As RSSFs podem ser implantadas em quase todos os tipos de ambiente, devido seu pequeno tamanho, facilidade de comunicação e baixo custo.

As RSSFs podem ser utilizadas em diversas aplicações: militares, industrial, transporte, agricultura, médica, monitoramento de áreas de difícil acesso, e monitoramento de integridade estrutural. Apesar do grande potencial das RSSFs, os nodos sensores são severamente restringidos em recursos de armazenamento, capacidades computacionais, largura de banda de comunicação e suprimento de energia. Essas restrições tornam o projeto de protocolos de comunicação para essas redes um desafio.

1.2 Motivação e desafios

A eficiência energética é um balanceamento entre a qualidade do serviço oferecido pela rede medida por taxa de envio de dados, conectividade e latência e a quantidade de

energia consumida pela rede. Portanto, para aumentar o tempo de vida da rede de maneira eficiente, os protocolos de comunicação empregados devem prover o equilíbrio entre esses dois aspectos.

A maior restrição associada ao nodo sensor é sua pouca capacidade energética. Tipicamente são usadas baterias de baixa tensão que podem ser trocadas ou recarregadas, porém realizar a troca ou a recarga das baterias de um nodo é um processo que pode envolver um grande custo operacional dependendo do tamanho da rede e em algumas situações é impraticável. Dessa forma, a eficiência energética é o principal fator a ser considerado no desenvolvimento de protocolos de comunicação para as RSSFs.

Em virtude do grande porte da rede e da impossibilidade de comunicação direta entre os nodos sensores e a BS, é de extrema importância que o protocolo de comunicação garanta conectividade entre os nodos sensores e a BS ao longo de sua operacionalidade, evitando a possibilidade de particionamento da rede.

Dessa forma, os protocolos de comunicação utilizados em RSSFs devem possuir características distintas em relação aos protocolos empregados em outros tipos de redes. Essas características impõem desafios, no projeto dos protocolos, relacionados à existência de limitações energética dos nodos, processamento, largura de banda, armazenamento e endereçamento, além das condições dinâmicas do ambiente.

A utilização de protocolos de roteamento do tipo plano em RSSFs pode não ser ótimo em relação à eficiência energética, pois não trata a redundância de dados e a diminuição do tráfego de repasse. A formação de grupos de nodos sensores na rede denominados *clusters* é um mecanismo de comunicação e organização que visa economia energética e pode ser utilizado em RSSFs para rotear dados dos nodos sensores para a BS ou agrupar os dados em uma rede distribuída.

O projeto de protocolos de comunicação baseados em *clusters* para RSSFs deve implementar mecanismos para o provimento de eficiência energética em todas as fases de sua operação, pois sua aplicabilidade resulta no surgimento de problemas estritamente inerentes ao processo de formação de *clusters*, tais como: *hot spot*, eleição de líderes de *clusters* denominados *cluster heads* (CH), conectividade e *overhead* de controle. Para solucionar esses problemas, nos últimos anos, muitos protocolos e algoritmos baseados em *clusters* tem sido propostos para RSSFs, no entanto, não se configuram como uma solução completa que vise atender todos os requisitos necessários para tornar uma RSSF eficiente, tais como:

- Eleição de CH e formação de grupos baseados em múltiplos critérios;
- Diminuição do *overhead* de controle;
- Formação de *clusters* com tamanhos desiguais;
- Conectividade entre os *clusters*;

1.3 Objetivos

Este trabalho tem como objetivo desenvolver um algoritmo de roteamento e clusterização que garanta eficiência energética a uma RSSF. Especificamente o algoritmo provê:

- Um mecanismo que permite a formação de *clusters* com tamanhos desiguais, com o objetivo de mitigar o problema de *hot spot*.
- Um algoritmo temporizado para eleição e comunicação de CHs com o objetivo de reduzir a sobrecarga de mensagens de controle no processo de reorganização dos *clusters*.
- Otimização por múltiplos critérios através do uso de lógica *fuzzy* a ser empregado no processo de formação de clusters e eleição de CH.
- Conectividade entre os *clusters* através de uma solução de roteamento e clusterização híbrida.

1.4 Justificativa

Atualmente a Amazônia é uma das maiores áreas verdes do mundo e sua biodiversidade influencia nas mudanças climáticas de todo o planeta. Assim, o Brasil tem um papel importante de preservá-la e explorá-la com sustentabilidade. Diante dessas situações fica claro que o grande desafio está em possibilitar o monitoramento ambiental em diversas áreas da Região Amazônica, com soluções de baixo custo e que cumpram com seu objeto de auxiliar na preservação das florestas tropicais.

As RSSFs apresentam-se como uma ótima solução para o monitoramento ambiental, em se tratando de um cenário amazônico, que possui áreas de grande extensão territorial e sem nenhuma infraestrutura. Além disso, o desenvolvimento de um algoritmo de roteamento e clusterização que se adeque as RSSFs de larga escala considerando todas as restrições de projeto no que tange a eficiência energética, permite as RSSFs serem implantadas em um cenário amazônico sem maiores problemas de disponibilidade e escalabilidade.

1.5 Organização do texto

Este trabalho está estruturado da seguinte forma: O Capítulo 2 aborda as principais questões referentes às RSSFs, incluindo uma visão geral, tipos e aplicações de RSSFs, restrições, desafios, a arquitetura de comunicação, alguns padrões e especificações comerciais de RSSFs e aspectos teóricos sobre roteamento em RSSFs, fornecendo uma visão geral sobre o assunto e tratando mais especificamente sobre clusterização e roteamento

hierárquico. O Capítulo 3 apresenta os trabalhos relacionados e uma análise funcional dos trabalhos em comparação com o algoritmo proposto.

O Capítulo 4 apresenta os pressupostos definidos, uma descrição detalhada sobre os problemas em RSSFs inerentes ao processo de clusterização e roteamento, o algoritmo de roteamento e clusterização proposto, descrevendo o funcionamento do algoritmo e o processo de inferência utilizando lógica *fuzzy* e uma análise de complexidade e correteude do algoritmo proposto, demonstrando o seu alinhamento aos objetivos que se pretende alcançar com o trabalho.

No capítulo 5 é apresentada a metodologia utilizada para avaliação de desempenho, os cenários utilizados para avaliação do algoritmo em um ambiente simulado e os resultados obtidos pelo algoritmo, demonstrando a viabilidade de implantação. Finalmente, o Capítulo 6 encerra o trabalho com as contribuições, uma avaliação geral da proposta e sugestões de alternativas de trabalhos futuros.

2 Referencial Teórico

2.1 Visão Geral

Atualmente é bem sabido que os avanços na tecnologia de rádio têm alcançado um baixo nível em termos de consumo de energia o que permitiu a criação de pequenos dispositivos de comunicação alimentados por bateria. Assim, quando integrados com microprocessadores de baixa potência e conectados em qualquer tipo de pequenos sensores - por exemplo, luminosidade, temperatura, batimentos cardíacos - um pequeno nodo sensor alimentado por bateria, capaz de processar dados e comunicar-se com outros dispositivos sem fio, é criado. [2]

Esses pequenos dispositivos autônomos, denominados nodos sensores, apresentam a capacidade de sensoriamento, processamento e comunicação, sendo seus principais componentes uma fonte de energia. Usualmente uma bateria de baixa tensão ou um gerador de energia, um sistema de processamento, que inclui um microcontrolador e uma memória, transceptor sem fio e a unidade de sensoriamento que compreende em um ou mais sensores e um conversor analógico-digital. Em alguns casos específicos o nodo sensor pode integrar um componente adicional tal como um sistema de localização, mobilidade entre outros, entretanto, estes componentes são opcionais. A organização interna de um nodo sensor é ilustrada na figura 1.

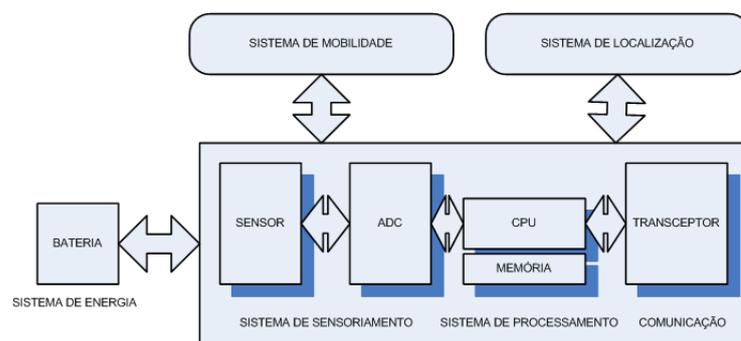


Figura 1: Organização interna de um nodo sensor.

A partir da implantação de vários nodos sensores em um determinado ambiente, o circuito de sensoriamento mede as condições naturais, tais como temperatura, umidade ou luminosidade e transforma-as em sinais elétricos. O processamento do sinal revela algumas propriedades sobre os objetos locais ou eventos acontecendo na vizinhança do sensor. Assim, o nodo sensor envia os dados coletados, através do transceptor, para a BS

que encaminha os dados para o usuário final.

Essa interconexão entre os nodos sensores resulta na criação das RSSFs que têm ganhado a atenção do mundo todo nos últimos anos, particularmente com a proliferação nos Sistemas Microeletromecânicos (*Microelectromechanical Systems* - MEMS) [3]. As RSSFs consistem em vários nodos sensores que executam aplicações individuais, se comunicam usando enlaces de comunicação sem fio, são distribuídos em um determinado local para monitorar condições físicas ou ambientais e enviam os dados coletados a BS que possui alta capacidade de energia e processamento, podendo essa comunicação acontecer diretamente ou através da retransmissão dos dados por múltiplos nodos.

As RSSFs são parte de um paradigma de novas redes, chamado de Internet das Coisas (*Internet of Things* - IOT) [4], que está se tornando muito popular em sistemas de comunicação sem fio. O princípio deste paradigma é que os objetos ou as “coisas” estão interagindo e cooperando uns com os outros para garantir as comunicações onipresentes. Esses objetos ou “coisas” poderiam ser *Radio-Frequency Identification* (RFID), RSSFs, atuadores, aplicações, telefones celulares entre outros.

Nesse trabalho será utilizado um modelo de RSSF conforme descrito na figura 2, que consiste em uma BS e um grande número de nodos sensores implantados sobre uma vasta área geográfica, onde os dados são transmitidos dos nodos sensores para a BS através de múltiplos saltos.

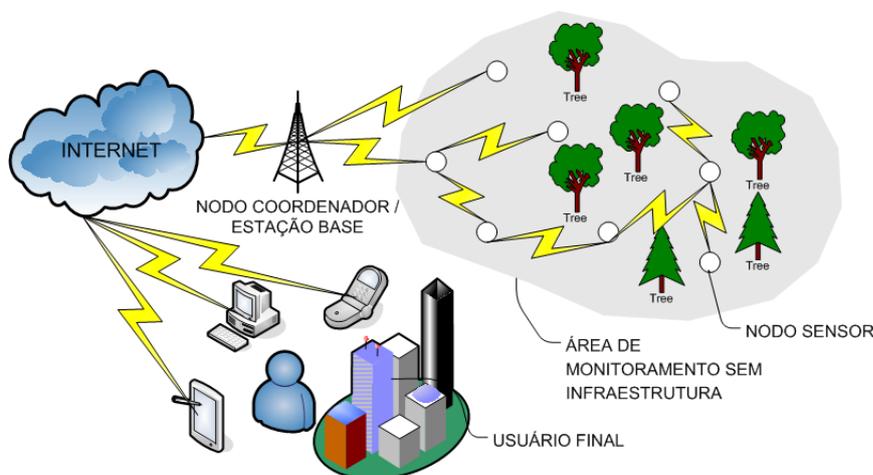


Figura 2: Modelo de uma RSSF.

Estes nodos sensores são em geral densamente implantados tornando frequentes os casos em que as áreas de sensoriamento de cada nó, por vezes, se sobrepõem. Além disso, as áreas de implantação, em muitos casos, inviabilizam o gerenciamento humano sobre os nodos, como consequência, os mesmos devem automaticamente colaborar uns com os outros para criar uma rede autônoma. A capacidade autônoma de organização dos nodos em redes funcionais sem a intervenção humana diferencia as RSSFs de outras redes convencionais. [5]

2.2 Tipos e Aplicações de RSSF

Devido as suas características, uma RSSF pode ser implantada em quase todos os tipos de ambientes e isto é justificado, entre outras razões, pelo pequeno tamanho dos nodos, facilidade de comunicação e baixo custo. Consequentemente existem diversas aplicações para RSSFs. São elas:

- Monitoramento e rastreamento de alvos militares; [6]
- Monitoramento da saúde humana; [7]
- Monitoramento e rastreamento de animais; [8]
- Monitoramento e rastreamento em vias públicas e indústrias; [9]
- Monitoramento ambiental; [10]

As RSSFs podem ser implantadas para diferentes propósitos. Em consequência disto recebem diferentes denominações na literatura quanto ao seu tipo. Seguem abaixo alguns dos principais tipos de RSSFs:

- Rede de sensor sem fio multimídia (*Wireless Multimedia Sensor Networks* - WMSN): Transmissão de áudio e ou vídeo capturados na área de monitoramento;
- Rede de sensor sem fio industrial (*Industrial Wireless Sensor Networks* - IWSNs): Implantadas dentro de fábricas;
- Redes de sensores sem fio corporais (*Wireless Body Area Networks* - WBANs): Implantadas no corpo humano para monitoramento e cuidados com a saúde;
- Redes de sensores e atuadores sem fio (*Wireless Sensor and Actuator Network* - WSAN): Permite o monitoramento e a atuação sobre as variáveis do ambiente;
- Redes de sensores sem fio móveis (*Mobile Sensor Networks* - MSNs): Rede que possui nodos com capacidades de mobilidade;
- Redes de sensores sem fio veiculares (*Vehicular Sensor Networks* - VSNs): Sensores veiculares são equipados em veículos com um padrão de mobilidade rápido.

Cada tipo de RSSF obedecerá a regras e arquiteturas de comunicação diferentes, em virtude de suas características imporem diferentes restrições, por exemplo, antes de projetar e analisar uma solução para WMSN devem ser considerados questões como qualidade de serviço (*Quality of Service* - QoS) e possíveis obstáculos no ambiente que possam prejudicar a qualidade do enlace na transmissão, resultando na perda da qualidade do vídeo transmitido.

2.3 Restrições e Desafios em RSSF

Ao contrário das tradicionais redes de computadores, como por exemplo, as redes *ad-hoc* sem fio (*wireless ad-hoc networks*), as RSSFs possuem suas próprias e específicas restrições de projeto e recursos. Embora inúmeros protocolos de comunicação tenham sido propostos para redes *ad-hoc*, eles não são bem adaptados para as características e requisitos das RSSFs pelas seguintes razões [11]:

- A topologia da RSSF muda com muita frequência;
- A quantidade de nodos sensores nas RSSFs pode ser diversas vezes maior do que a quantidade de nodos em redes *ad-hoc*;
- Os nodos sensores são densamente implantados no campo de monitoramento;
- Os nodos sensores normalmente utilizam um paradigma de comunicação de difusão de pacotes, enquanto a maioria das redes *ad-hoc* são baseadas em comunicação ponto a ponto;
- Os nodos sensores não podem ter um identificador global devido a grande quantidade de overhead gerado em função da grande quantidade de nodos;
- Os nodos sensores possuem restrições de energia, potência para comunicação, largura de banda, processamento e armazenamento;

As restrições de projeto dependem da aplicação e do ambiente de implantação das RSSFs. Isto determinará o tamanho e a topologia da rede. Esse tamanho pode variar de ambientes internos para externos, além disso, o esquema de implantação, *ad-hoc* ou pré-planejado, pode criar restrições de comunicação entre os nodos afetando a conectividade da rede.

A energia se configura como a principal restrição das RSSFs, pois a substituição ou recarga das baterias dos nodos sensores é cara e impraticável em alguns casos. Assim, a eficiente utilização dos recursos de energia dos nodos foi e permanece sendo o principal objetivo de muitos protocolos propostos para RSSFs. Dessa forma, consideráveis pesquisas tem sido realizadas com o objetivo de superar o problema de consumo de energia dos nodos sensores e prolongar o tempo de vida das RSSFs. Nesse contexto, define-se como tempo de vida da rede o tempo decorrido até o esgotamento das reservas de energia do primeiro ou último nodo da rede [12].

A implementação de protocolos em diferentes camadas na pilha de protocolos pode afetar significativamente o consumo de energia, o atraso fim-a-fim e a eficiência do sistema. Dessa forma, é importante otimizar e minimizar o uso de energia nas comunicações [3]. Assim, ao contrário das redes cabeadas ou outras redes sem fio (celula, WLAN, etc.), a conservação de energia em RSSFs é uma questão crítica que tem sido tratada por substanciais trabalhos de pesquisa. Geralmente, a conservação de energia é tratada em cinco níveis diferentes: [13]:

- Eficientes agendamentos de estados do transceptor dos nodos sensores para alternar entre os modos ativo e adormecido;
- Eficiente controle de potência de transmissão para assegurar um ótimo balanceamento entre consumo de energia e conectividade;
- Compressão de dados para reduzir a quantidade desnecessária de dados transmitidos;
- Roteamento, clusterização e agregação de dados com eficiência energética;
- Eficiente acesso ao canal e protocolos de retransmissão de pacotes na camada de enlace de dados;

Existem diversas abordagens que podem ser utilizadas como: um baixo ciclo de trabalho desligando o transceptor em determinados momentos, agregação de dados, fusão de dados, filtragem e entre outros. Outro fator importante é o protocolo de roteamento, que pode afetar de forma significativa o consumo de energia na rede. Assim, ele deve ser desenvolvido com o objetivo de prover um balanceamento de carga na rede e otimizar a escolha das rotas considerando múltiplas métricas que possam afetar diretamente o consumo de energia reduzindo o tempo de vida da rede. Este trabalho trata os quatro primeiros itens da lista acima, mais especificamente a técnica de clusterização e roteamento com eficiência energética.

Os desafios e propriedades das RSSFs não se restringem apenas ao que foi explanado aqui, existem outros desafios inerentes ao processo de implantação, localização dos nodos, agregação e fusão de dados, agendamentos de ciclos de trabalho, segurança, heterogeneidade, tolerância a falhas, cobertura e qualidade de serviço. A discussão sobre cada desafio não será explanado em detalhes neste trabalho, pois não compõe o escopo central do mesmo.

2.4 Arcabouço de Comunicação

As RSSFs consideram apenas as seguintes camadas [14]: Aplicação, Rede, Controle de Acesso ao meio e Física. Embora não exista a camada de transporte, desde que esta é complexa e isto poderia desperdiçar energia dos nodos [15]. Alguns protocolos têm sido projetados com controle de congestão e confiabilidade na comunicação fim a fim. A pilha de protocolos pode ser vista na figura 3.

A camada de rede é responsável por controlar o roteamento de pacotes na rede através de soluções escaláveis, auto organizáveis e que atenda outras restrições inerentes ao tipo de aplicação da RSSF. A camada de enlace é responsável pela comunicação entre dois nodos que compartilham o mesmo enlace de comunicação estabelecendo um controle e gerenciamento de acesso ao meio compartilhado, sendo que, as soluções devem considerar as restrições impostas. A camada física fornece uma interface para transmitir um fluxo

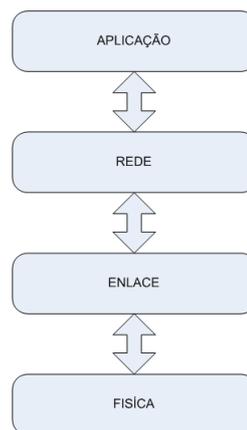


Figura 3: Pilha de protocolos de uma RSSF.

de bits sobre o meio de comunicação físico, interagindo constantemente com a camada de enlace para efetuar diversas funções, tais como transmissão e recepção.

Baseadas nessa arquitetura de comunicação, diversas especificações e protocolos comerciais foram desenvolvidos com o intuito de mitigar os problemas existentes e garantir confiabilidade na transmissão de dados.

2.5 Padrões e Especificações para RSSFs

Os padrões e especificações definem as funções e protocolos necessários para os nodos sensores se comunicarem. O padrão IEEE 802.15.4 é empregado em redes de comunicação de baixo custo compostas por dispositivos sem fio com limitada capacidade energética e pequenos requisitos para taxa de transmissão de dados. Apesar do padrão não ter sido projetado especificamente para as RSSFs ele é largamente utilizado nessas redes, em função das características das redes as quais se aplica se adequarem as características das RSSFs.

O padrão IEEE 802.15.4 define a especificação para camada física e a subcamada de controle de acesso ao meio para conectividade sem fio com baixa taxa de transmissão de dados entre dispositivos fixos, portáteis e móveis, sem bateria ou com requisitos limitados para o consumo de bateria, operando dentro de um espaço de operação pessoal de 10 m ou mais dependendo da aplicação. O padrão define dois tipos de dispositivos, sendo eles o *full function device* (FFD) e *reduced function device* (RFD). Um RFD pode apenas se comunicar com um FFD, enquanto um FFD pode se comunicar tanto com um FFD como um RFD. O padrão pode operar sobre uma topologia estrela ou *peer-to-peer*. Na topologia estrela cabe ao nodo coordenador todo o controle da rede assumindo um papel central e de comunicação direta com todos os dispositivos finais.

O padrão ZigBee [16] é um padrão baseado nas camadas físicas e MAC do padrão IEEE 802.15.4 e fornece protocolos para as camadas de rede e aplicação além de mecanismos de segurança. Os nodos sensores podem ser organizados em uma topologia estrela,

em árvore ou em grade. Além disso, o padrão inclui três diferentes tipos de nodos: coordenador ZigBee, roteador ZigBee e dispositivos finais ZigBee. O coordenador ZigBee é um nodo FFD que gerencia a rede e as chaves de segurança, o roteador ZigBee são nodos FFD com capacidades de rotear os pacotes e estabelecer a comunicação entre grupos de nodos, o dispositivo final ZigBee pode ser um FFD ou um RFD, ele transmite os dados coletados pelos sensores para um coordenador ou roteador.

A camada de rede especificada pelo padrão ZigBee é responsável pela formação, endereçamento e roteamento da rede. A formação e endereçamento funciona através de um mecanismo de descoberta e associação estabelecendo uma hierarquia entre os nodos e o roteamento adota o algoritmo baseado em vetor de distância para redes Ad Hoc (*Ad hoc On Demand Distance Vector routing algorithm - AODV*). A camada de aplicação propõe um *framework* para o desenvolvimento e comunicação de aplicações distribuídas, sendo unidades de softwares controlando dispositivos de hardware.

O padrão *Open Wireless HART* [17] foi lançado em 2007 e é adequado para aplicações de controle e medição e definem diferentes tipos de nodos, incluindo nodos *field*, *gateways*, *adapters* e *handheld*. Os nodos *field* podem ser organizados em uma topologia estrela ou em grade, os nodos *gateway* são uma ponte entre os dispositivos de campo e as aplicações do usuário, além disso eles gerenciam a rede e controlam a segurança, os nodos *adapters* e *handheld* são nodos opcionais e permitem a configuração e associação de nodos a rede. O padrão fornece protocolos para as camadas de enlace de dados, rede, transporte e aplicação, assim, na camada de enlace de dados o padrão utiliza TDMA e CSMA/CA, na camada de rede trabalha com dois protocolos de roteamento oferecendo transmissão *broadcast*, *multicast* and *unicast*, na camada de transporte suporta comunicações orientada e não orientada a conexões e a camada de aplicação é baseada em comandos HART.

O padrão ISA100.11a-2009 [18] pode ser organizado em diferentes topologias e são compostos de nodos *field*, *gateways* e *handheld*, os nodos *field* são responsáveis pela coleta e roteamento dos dados, os nodos *gateways* asseguram a conexão entre a RSSF e a aplicação do lado do usuário, além disso, permite a interoperabilidade com diferentes padrões, tais como o *WirelessHART*. Os nodos *handheld* suportam a instalação, configuração e manutenção na rede. O padrão trata todas as camadas do modelo OSI, assim, a camada física é baseada sobre o padrão IEEE 802.15.4, a camada de enlace de dados é responsável por gerenciar o emprego de esquemas TDMA configurando a duração dos slots de tempo e o gerenciamento de superframes, a camada de rede fornece esquemas de roteamento e QoS, além de ser compatível com o padrão IETF 6LoWPAN. A camada de transporte pode dar suporte a comunicações fim-a-fim com e sem reconhecimento de entrega de pacotes, controle de fluxo, segmentação e reconstrução de pacotes e segurança.

2.6 Roteamento em RSSFs

O paradigma de comunicação empregado pelas RSSFs se baseia na coleta de dados de determinado ambiente através dos sensores e envio dessas informações para a BS,

que posteriormente, envia para uma aplicação do lado do usuário final. A comunicação direta entre os nodos sensores e a BS impulsiona o nodo sensor a utilizar altos níveis de potência para transmitir os dados, conseqüentemente, os recursos de energia dos nodos serão esgotados rapidamente. É necessário o estabelecimento de um comportamento colaborativo entre os nodos através do uso de nodos intermediários que retransmitam os dados por múltiplos saltos até a BS, a fim de permitir que todos possam se comunicar com a BS sem grandes impactos sobre os recursos energéticos.

Uma simples solução seria o uso de um algoritmo de inundação, onde um nodo transmite os dados por difusão e os mesmos são retransmitidos por todos os nodos que recebem os dados até que este último chegue ao seu destino. Nesta abordagem são encontradas duas deficiências. Uma delas é a redundância de dados na rede em função da sobreposição de áreas afetadas pelo evento e a outra está relacionada à intensa transmissão dos nodos que pode ocasionar no rápido esgotamento de seus recursos energéticos.

Em virtude das deficiências encontradas por essas soluções é necessário o uso de protocolos de roteamento em RSSFs que realizem a configuração e o estabelecimento de rotas na rede considerando as restrições que as mesmas possuem. Assim, um protocolo de roteamento é um protocolo que especifica como os nodos sensores se comunicam um com o outro, disseminando informações que permitam que eles selecionem as rotas entre dois ou mais nodos na rede, sendo a escolha da rota realizada pelo algoritmo de roteamento [19].

Diferentes protocolos de roteamento são propostos para RSSF e a literatura classifica-os conforme alguns parâmetros operacionais, organizacionais e estruturais. Assim, do ponto de vista operacional os protocolos de roteamento pode ser [19]:

- Protocolos pró-ativos: Transmitem os dados coletados para a BS através de uma rota pré-definida;
- Protocolos reativos: Transmitem os dados coletados para a BS apenas na ocorrência de um evento, estabelecendo a rota durante o envio;
- Protocolos híbridos: Incorpora os conceitos do pró-ativo e reativo.

Do ponto de vista organizacional os protocolos de roteamento podem ser [19]:

- Protocolos com comunicação direta: Qualquer nodo pode enviar os dados diretamente para a BS;
- Protocolos planos: Quando existir a necessidade de transmitir os dados, este pesquisa por rotas válidas para a BS para então transmitir os dados;
- Protocolos com clusterização: A área total da rede é dividida em *clusters* e cada um contém um CH que pode se comunicar diretamente com a BS, assim, todos os demais nodos do grupo, denominados *cluster members* (CM), enviam os dados coletados por eles para seus respectivos CHs associados.

E finalmente do ponto de vista estrutural os protocolos de roteamento podem ser [19]:

- Protocolos hierárquicos: É um roteamento com eficiência energética, onde nodos com altos níveis de energia podem ser usados para processar e enviar as informações para a BS e os demais nodos com baixo nível de energia são usados para realizar o sensoriamento;
- Protocolos centrados em dados: São baseados em consultas, dependendo sobre as nomeações dos dados de interesse;
- Protocolos baseados em informações de localização: Utiliza informações de localização dos nodos para rotear o sinal.

Neste trabalho, o protocolo proposto pode ser classificado, segundo os critérios explanados, como um protocolo pró-ativo, hierárquico e com clusterização visando garantir a eficiência energética da rede.

2.7 Clusterização e Roteamento Hierárquico em RSSFs

Do ponto de vista da habilidade funcional, todos os nodos sensores são considerados iguais. Porém, o modelo de roteamento hierárquico estabelece duas classes de nodos que executam serviços diferentes: nodos com alto nível de energia que funcionam como concentradores para processamento e envio de informações e nodos com baixo nível de energia que efetuam apenas o sensoriamento e enviam informações para os nodos com alto nível de energia.

O roteamento plano como protocolo de disseminação de dados em RSSF pode não ser eficiente em relação ao consumo de energia [20]. Portanto, técnicas de clusterização podem ser empregadas em protocolos de comunicação hierárquicos visando a eficiência energética e pode ser utilizado em nodos sensores para rotear seus dados para a BS.

Os protocolos de clusterização para RSSFs permitem numerosas aplicações para o monitoramento ambiental. Essas redes além de prover eficiência energética possuem outros desejáveis objetivos, tais como balanceamento de carga, tolerância a falhas e aumento de conectividade [21], além disso, são particularmente úteis para aplicações que requeram escalabilidade para cem ou duzentos nodos.

Nos algoritmos de clusterização cada *cluster* é administrado por um CH que é responsável em coordenar a transmissão de dados de todos os nodos ativos pertencentes a seu *cluster* e transmitir os mesmos para a BS. Os demais nodos, CM, coletam e enviam as informações para seu respectivo CH associado. Cada CH recebe os dados coletados pelos CM, agrega os dados, utilizando alguma técnica de agregação de dados presente na

literatura [22], e então transmite os dados agregados para a BS diretamente ou através de outros CHs, estabelecendo assim, uma comunicação por múltiplos saltos. A figura 4 demonstra a estrutura da rede baseada em *clusters*. Os nodos envoltos de um quadrado vermelho são os CHs e os demais nodos são os CMs, sendo que cada *cluster* possui o seu raio de atuação.

A composição dos *clusters* pode acontecer de maneira centralizada e descentralizada, na abordagem centralizada um nodo controlador é responsável por periodicamente determinar o tamanho ótimo dos *clusters*, adaptando a rede baseado em seu estado atual. Por outro lado, na abordagem descentralizada cada nodo antes do início da operação da rede toma decisões sobre o seu papel a ser desempenhado na rede de forma autônoma, baseadas em determinados parâmetros acerca da distribuição dos nodos da rede.

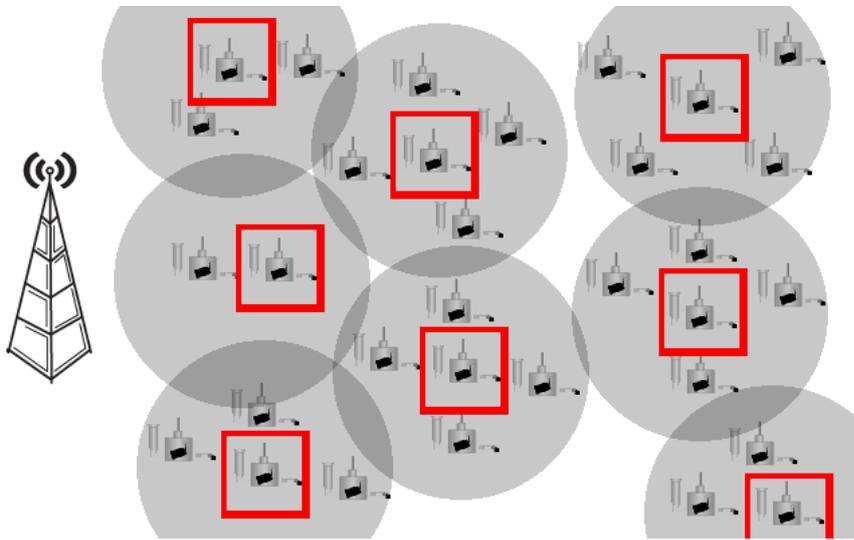


Figura 4: Estrutura de uma RSSF baseada em clusters.

A análise do consumo de energia sobre essa organização da rede revela que os CHs consomem mais energia que os demais nodos em função de suas contínuas tarefas de recepção, agregação e transmissão a longas distâncias, assim, esses nodos devem ser periodicamente alternados a fim de garantir a uniforme distribuição do consumo de energia na rede.

A redução da distância para transmissão dos nodos membros dentro dos *clusters* possibilita uma redução no consumo de energia, além disso, a agregação de dados no CH permite a redução da transmissão de dados redundantes, reduzindo assim a quantidade de pacotes a serem transmitidos na rede. Diversos estudos demonstraram que a clusterização aumenta o tempo de vida da rede [12] [23].

3 Trabalhos Relacionados

3.1 Clusterização hierárquica adaptativa com baixa energia

O *Low Energy Adaptive Clustering Hierarchy* (LEACH) [24] é o principal precursor dos protocolos de roteamento hierárquico para RSSF. Foi projetado para redes que tem características de envio de dados contínuo, com homogeneidade de funcionalidades e nodos sem mobilidade. Utiliza técnica de agregação de dados que combina ou agrega o dado original em uma única mensagem que contém uma informação resumida de todos os dados coletados pelos CMs [25]. Assim, o LEACH divide a rede em vários *clusters* que são construídos utilizando parâmetros locais.

Cada *cluster* possui um CH que é responsável em fazer comunicação com seus respectivos CMs, agregar os dados recebidos dos mesmos e transmiti-los para a BS. Os CMs possuem a função de coletar dados e enviá-los ao seu respectivo CH associado. Desse modo, não se reduz apenas a quantidade de dados a serem transmitidos para a BS, mas também realiza roteamento e disseminação de dados de forma mais escalável e robusta.

Dado que a energia dissipada pelo nodo sensor depende da potência utilizada para transmissão do sinal e da quantidade de dados a serem transmitidos, o LEACH tenta formar CHs que possam transmitir para a BS em pequenas distâncias reduzindo o número de operações de transmissão e recepção [25]. Os nodos escolhidos para atuarem como CHs não mantêm essa condição permanentemente, pois dessa forma sua energia rapidamente se esgotaria. Em vez disso, o LEACH utiliza distribuição aleatória para dar chances a todos os nodos sensores atuarem como CHs e evitar o gasto exacerbado de energia de algum nodo individual. A operação do LEACH é dividida em rodadas, onde cada rodada possui duas fases principais: fase de configuração, responsável em organizar a estrutura da rede, conforme a Figura 5, e a fase de operação onde ocorre transmissão de dados até a BS.

Cada CH é selecionado através de uma equação probabilística, e cada um divulga sua condição de CH para todos os nodos sensores da rede. Essa decisão de se tornar CH é feita independente, sem intervenção de alguma comunicação com algum outro nodo. Especificamente, cada nodo sensor decide ser CH baseado em uma porcentagem P de CHs predefinida em relação ao total de nodos da rede. Assim, cada nodo gera um número aleatório e se esse valor for menor que $T(n)$ (Equação 3.1), o nodo sensor n será CH na

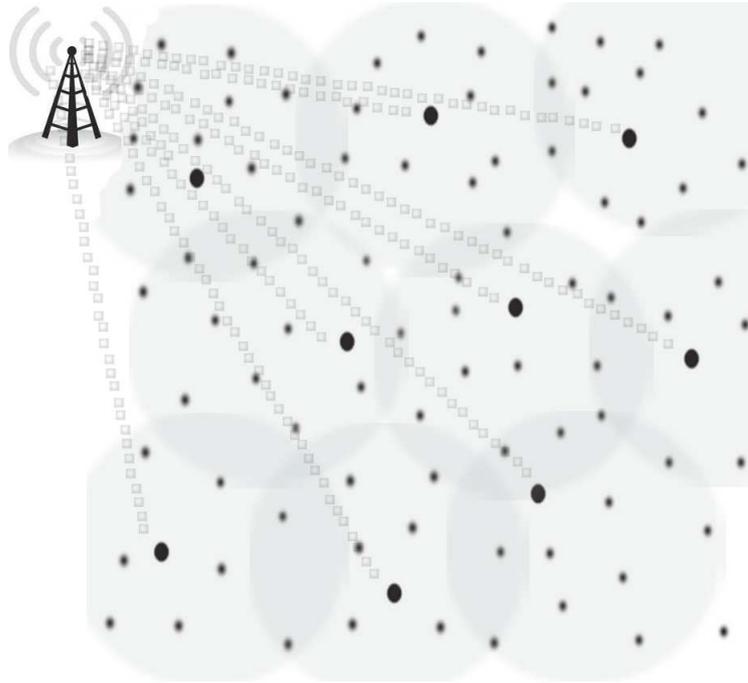


Figura 5: Organização da rede com LEACH

rodada atual.

$$T(n) = \begin{cases} \frac{P}{1 - P \times (r \bmod \frac{1}{P})} & \text{if } n \in G \\ 0 & \text{Caso contrário} \end{cases} \quad (3.1)$$

onde:

- $P \rightarrow$ Porcentagem de CHs a serem formados;
- $r \rightarrow$ Rodada atual;
- $n \rightarrow$ Identificação do nodo sensor;
- $G \rightarrow$ Conjunto de nodos que ainda não atuaram como CH.

O nodo não poderá se candidatar novamente a CH enquanto o número de rodadas não ultrapassar o valor de P . Entre todas as mensagens de controle enviadas por CHs e recebidas pelos CMs, os CMs selecionam o sinal menos atenuado recebido para minimizar o custo energético de comunicação e informa ao respectivo CH sobre sua decisão. Após o CH receber todas as mensagens de requisição de associação de cada CM, o mesmo cria um ciclo de trabalho baseado em TDMA para transmissão de cada nodo. Em seguida, cada CH reúne os dados recebidos de todos os nodos, agrega e transmite para a BS. Cada nodo não necessita manter em estado ativo o seu transceptor, por isso, será ativado somente no momento exato de realizar a transmissão de dados de acordo com seu tempo pré-determinado.

3.2 Um protocolo de roteamento baseado em clusters desiguais

Unequal Cluster-based Routing (UCR) [26] é baseado no protocolo LEACH e incorpora na sua fase de configuração um algoritmo de clusterização desigual com eficiência energética (*Energy-Efficient Unequal Clustering* - EEUC) para o gerenciamento de topologia e um protocolo de roteamento ciente de energia e baseado em localização geográfica (*greedy geographic and energy-aware routing protocol*) para a comunicação entre clusters.

EEUC é um algoritmo competitivo, onde os CHs são selecionados baseados sobre informações locais, assim, define-se uma probabilidade inicial para seleção de candidatos à CHs e cada um terá uma área de cobertura que estabelecerá uma competição entre os nodos. O raio de competição diminui conforme a distância entre o nodo sensor e a BS diminui. O raio de competição é determinado pela equação 3.2. Posteriormente o CH é eleito com base na sua energia residual.

$$R_i = 1 - c \times \frac{d_{max} - d(S_i, BS)}{d_{max} - d_{min}} \times R_0 \quad (3.2)$$

onde:

- d_{min} → Distância mínima entre o nodo sensor e a BS em metros;
- d_{max} → Distância máxima entre o nodo sensor e a BS em metros;
- $d(S_i, BS)$ → Distância entre o nodo sensor e a BS em metros;
- R_0 → Raio máximo de clusterização em metros;
- c → Fração que determina o raio mínimo de clusterização.

Em seguida, para o estabelecimento do roteamento por múltiplos saltos entre *clusters* cada CH difunde uma mensagem de controle na rede, essa mensagem permitirá a escolha do nodo retransmissor, assim, para os nodos vizinhos da BS é estabelecido um limiar, denominado TD_MAX, para que os nodos com distância menor que TD_MAX em relação a BS realizem a transmissão direta para está última. Para os demais nodos um conjunto de informações de nodos vizinhos é definido, segundo a expressão 3.3.

$$s_i.R_{CH} = s_j | d(s_i, s_j) < x.R_i, d(s_j, BS) < d(s_i, BS) \quad (3.3)$$

onde:

- $d(x_i, y_j)$ → Representa a distância entre os nodos x e y em metros;
- R_i → Representa o raio de clusterização de um determinado nodo em metros;
- x → Mínimo inteiro que considera que $s_i.R_{CH}$ contém pelo menos um item.

Para escolha do CH é definida uma métrica a partir da equação 3.4, que permite a composição do conjunto de CH elegíveis definido na expressão 3.5 para finalmente definir os melhores em relação a métrica, ou seja, que possuem a menor distância e então eger o nodo com maior nível de energia residual.

$$E_{relay}(s_i, s_j) = d^2(s_i, s_j) + d^2(s_j, BS) \quad (3.4)$$

$$s_i.S_{eligible} = s_j \in s_i.R_{CH}, Min_k(E_{relay}(s_i, s_j)) \quad (3.5)$$

3.3 Protocolo de roteamento e clusterização híbridos

Hybrid Clustering and Rounting (HCR) [27] é baseado no protocolo LEACH incorporando na fase de configuração um protocolo de roteamento e clusterização híbrido por tratar simultaneamente tanto a eleição de CH para formação dos *clusters* quanto a comunicação entre eles. Permite a formação de *clusters* com baixo *overhead* a partir de um algoritmo baseado em um temporizador e o estabelecimento de um campo de custo para garantir a conectividade na rede. Assim, o protocolo assegura conectividade e eficiência energética para a rede.

Inicialmente é feito o estabelecimento dos custos, onde a BS difunde uma mensagem com o custo $k=0$, o nodo que receber a mensagem deve retransmitir com custo $k=k+1$, assim, este processo repete-se até todos os nodos receberem a mensagem. Em seguida, a BS difunde outra mensagem de controle informando para todos os nodos da rede qual o custo máximo para comunicação com a BS.

O nodo configura o seu temporizador de acordo com o seu custo k e o seu nível de energia residual, conforme a equação 3.6, decidindo por ser CH ou CM com base na quantidade de mensagens de controle de requisição para ser um nodo de retransmissão. Este processo ocorre através do envio de mensagens de controle com um atraso definido pelo temporizador. Para os CMs o processo de escolha de nodos CH para composição dos clusters considera o menor custo K e o maior nível de energia residual. Os CHs devem utilizar o raio de transmissão para encaminhamento dos dados para os nodos de repasse ou diretamente para a BS.

$$T_b = (K - k) \times T_{slot} + \frac{E_{max} - E(r)}{E_{max}} \times T_{slot} \quad (3.6)$$

onde:

- $K \rightarrow$ Custo máximo na rede;
- $T_{slot} \rightarrow$ Tamanho dos slots usados para as partes geradas pela quebra do processo de eleição de CH;

- E_{MAX} → Energia máxima em Joules;
- E_r → Energia Residual em Joules;

3.4 Formação de cluster com eficiência energética

O *Energy-Efficient Cluster Formation* (EECF) [28] é um protocolo de clusteração distribuído, onde um conjunto de CHs são eleitos seguindo uma troca de mensagens em três vias entre cada nodo sensor e os seus vizinhos a um salto de distância, com o objetivo de maximizar o tempo de vida da rede e garantir a conectividade entre os CHs.

Primeiramente cada nodo difunde uma mensagem de controle contendo uma pontuação estabelecida por uma combinação linear entre o grau de adjacência de cada nodo em relação aos demais e sua energia residual, determinada pela equação 3.7, que promoverá cada nodo a ser CH.

$$P(i) = \alpha \times \frac{\text{degree}(i)}{N - 1} + (1 - \alpha) \times \frac{Er(i)}{E_0} \quad (3.7)$$

onde:

- $\text{degree}(i)$ → Grau de adjacência do nodo i ;
- N → Quantidade de nodos na rede;
- $Er(i)$ → Energia residual do nodo i em Joules;
- E_0 → Energia inicial do nodo i em Joules.

Após o recebimento da mensagem de controle cada nodo divide o conjunto de vizinhos conhecidos em dois *clusters* cuja pontuação é maior ou menor que a sua, em seguida determina a preferência relativa a cada nodo vizinho através da equação 3.8.

$$RCR_i(v) = \frac{N^+(i) + 1 - \text{Rank}(v)}{\frac{N^+(i) \times N^+(i) + 1}{2}} \quad (3.8)$$

onde:

- $\text{Rank}(v)$ → Ordenamento dos nodos sensores dentro de subconjuntos;
- N^+ → Subconjunto de nodos com pontuação maior;

Posteriormente, os nodos difundem a mensagem *Relative Cluster head Rank Advertisement* (RCRA) contendo os valores de $RCR_i(v)$ de todos os vizinhos e reciprocamente cada nodo recebe os mesmos valores de seus vizinhos. Em seguida, cada nodo calcula uma função denominada RCN(i), determinada pela equação 3.9, que representa a

contribuição relativa do nodo para rede representando a probabilidade do nodo se tornar CH, baseado sobre as informações recebidas dos vizinhos.

$$RCN(i) = 1 - \prod (1 - RCR_v(i)) \quad (3.9)$$

Uma mensagem de controle denominada *Contribution to the Network Advertisement* (RCNA) é difundida contendo o RCN(i) e após todos os nodos receberem a informação os mesmos decidem por qual CH se associar ou se o nodo em questão será um CH. Em seguida, é difundida uma mensagem denominada *Election Message* (EM) contendo a identificação dos nodos escolhidos como CH para associação ou a própria identificação do nodo informando que o mesmo será um CH. Este processo decisório considera a pontuação e a quantidade recente de vezes que o nodo atuou como CH.

3.5 Algoritmo de clusterização desigual baseado em localização

Location-based Unequal Clustering Algorithm (LUCA) [29] é um algoritmo de clusterização que tem como objetivo minimizar o consumo total de energia quando todos os nodos sensores na rede enviam os dados coletados para a BS. O protocolo emprega a configuração de *clusters* de tamanho desigual considerando a distância entre o CH e a BS e utiliza um temporizador aleatório para eleger os CHs de maneira uniforme.

Assim, considere r_i o tamanho do *cluster* em termos de números de saltos calculado pela equação 3.10:

$$r = \frac{1}{r_0} \sqrt[3]{\frac{3D}{\lambda\pi}} \quad (3.10)$$

onde:

- $D \rightarrow$ Distância entre o CH e a BS em metros;
- $r \rightarrow$ Tamanho do cluster em metros;
- $\lambda \rightarrow$ Intensidade da distribuição de Poisson, para distribuição dos nodos.

Cada nodo sensor configura o seu temporizador usando um gerador de números aleatórios, dessa forma, nos casos onde o tempo expirar e o nodo não recebeu qualquer mensagem de controle este se promoverá um CH e enviará uma mensagem informando o seu estado e os nodos que não foram eleitos tomarão a decisão sobre o CH a se associar com base na sua distância em relação ao CH.

3.6 Esquema de clusterização e roteamento ciente de energia

Power Aware Routing and Clustering Scheme (PARC) [30] é um algoritmo de clusterização e roteamento que integra as fases de eleição CH e construção de rotas para a BS com o objetivo de reduzir o *overhead* de controle na rede e permitir a eleição de uma grande quantidade de CHs ao redor da BS para mitigar o problema de *hot spot*.

No início de cada rodada a BS transmite em baixa potência um mensagem denominada *slave control packet* (SCP) com informações do identificador da BS. Os nodos que receberam o SCP tornam-se escravos da origem. A BS difunde uma mensagem de controle denominada *Route Request* (RREQ) com os seguintes dados: identificador do nodo, identificador do nodo pai, número de saltos e número de sequência. Os nodos que receberam a mensagem configuram um temporizador conforme a equação 3.11.

$$T = \frac{1}{E} \times \alpha. \quad (3.11)$$

onde:

- $T \rightarrow$ Valor do temporizador em segundos;
- $E \rightarrow$ Energia residual em joules;
- $\alpha \rightarrow$ Valor aleatório.

O nodo se torna um CH caso seu temporizador expire e o mesmo não tenha recebido qualquer RREQ, em seguida, o mesmo envia uma mensagem de controle denominada *slave control*(SC) e um RREQ. Caso o temporizador esteja ativo e receba RREQ com o mesmo número de sequência o nodo será um nodo escravo. Caso o nodo torne-se um CH, o mesmo envia uma mensagem de controle *router reply* (RREP) para a BS. Se um CH receber um RREP não endereçado para este, ele julga que existe outro CH dentro da sua área de clusterização e muda sua condição para CM.

3.7 Algoritmo de clusterização hierárquico e temporizado com eficiência energética

Energy Efficient Backoff Hierarchical Clustering Algorithm (EEBHC) [31] é um algoritmo de clusterização baseado em temporização ciente dos níveis de energia residual dos nodos que objetiva melhorar o balanceamento do consumo de energia da rede através de uma distribuição uniforme de CHs.

A operação do algoritmo é dividida em rodadas, cada rodada inicia com a fase de configuração seguida pela fase de transmissão. Na fase de configuração cada nodo espera por um tempo aleatório calculado pelas equações 3.12 e 3.13:

$$t_i = -\frac{1}{\lambda_i} \ln(1 - x_i) \quad (3.12)$$

onde:

- $x_i \rightarrow$ Número aleatório distribuído sobre o intervalo $[0,1]$.

$$\lambda_i = \alpha \frac{E_{res}}{E_{max}} \quad (3.13)$$

onde:

- $\alpha \rightarrow$ Constante.

Após esse tempo, se o nodo não receber qualquer mensagem o mesmo se elege automaticamente um CH e transmite uma mensagem de controle que é difundida na rede até uma determinada quantidade de saltos de distância. Se o nodo receber uma mensagem antes do término do tempo de espera o mesmo se tornará um CM. Após a definição de todos os papéis, os CMs escolherão dentre as opções, a qual *cluster* os mesmos devem se associar e enviam uma mensagem de controle para informar suas escolhas para o CH, o mesmo processo é realizado pelos CHs para a escolha dos nodos de repasse. Os nodos que não enviaram ou receberam qualquer mensagem de controle, estes serão considerados obrigatoriamente CHs.

O algoritmo pode atuar de maneira hierárquica de uma forma *bottom-up*, onde, os CHs de 1º nível são eleitos por toda a rede e posteriormente os de 2º nível são eleitos a partir dos de 1º nível e assim sucessivamente até o nível mais alto, ou seja, o processo de temporização é agendado conforme os diferentes níveis de hierarquia da rede.

3.8 Algoritmo de clusterização LEACH-ERE

O LEACH-ERE [32] é um algoritmo de clusterização baseado em lógica *fuzzy* com o propósito de prolongar o tempo de vida de uma RSSF através da distribuição de carga na rede. No algoritmo é utilizada uma métrica denominada *expected residual energy (ERE)* que atua como um descritor *fuzzy* durante a eleição de CHs. A métrica ERE é estimada com base em outra métrica denominada *expected energy consumption (EEC)*.

O algoritmo proposto adota a arquitetura do LEACH com uma extensão para a predição de energia baseada sobre ERE, por isso a abordagem proposta denomina-se LEACH-ERE. Para o cálculo da métrica ERE inicialmente é obtido o número de *frames* por rodada após a formação dos *clusters*, a equação 3.14 demonstra como obter a quantidade de *frames*.

$$N_{frame} = \frac{t_{ssPhase}}{n \times t_{slot} + t_{CHtoBS}} \quad (3.14)$$

onde:

- $t_{ssPhase}$ → Tempo de operação da fase de estado estável;
- t_{slot} → Tempo alocado requerido para transmissão de um CM para um CH em segundos;

A energia a ser consumida por um nodo para tornar-se CH depois da fase de estado estável é calculada pela equação 3.15.

$$N_{expConsumed}(l, d_{toBS}, n) = N_{frame} \times (E_{Tx}(l, d_{toBS}) + n \times E_{Rx}(l)) \quad (3.15)$$

onde:

- $E_{Tx}(l, d_{toBS})$ → Energia gasta na transmissão para BS em Joules;
- $E_{Rx}(l)$ → Energia gasta na recepção de dados em Joules;

Com base nas equações acima a energia residual esperada de um nodo para tornar-se CH após a fase de estado estável pode ser obtida pela equação 3.16.

$$E_{expResidual}(l, d_{toBS}, n) = E_{residual} - E_{expConsumed} \quad (3.16)$$

onde

- $E_{residual}$ → Energia residual do nodo sensor antes da eleição de CH.

Para controlar as incertezas, a abordagem utiliza um sistema de inferência *fuzzy* (FIS) para calcular a chance de cada nodo. O FIS utiliza duas variáveis de entrada que são, $E_{residual}$ e $E_{expResidual}$ e uma saída que é a probabilidade de um nodo tornar-se CH.

O FIS utiliza funções de pertinência do tipo trapezoidal e triangular para todas as variáveis linguísticas do sistema e utiliza o método de centro de área (COA) como método de *defuzzification*. O LEACH-ERE utiliza heurísticas para determinação das regras que compõem a base de regras do FIS.

3.9 Análise sobre os Trabalhos Relacionados

O LEACH pode ser visto como uma proposta híbrida de pequenas e longas distâncias de repasse de dados. O nodo sensor transmite seus dados em pequenas distâncias ao CH, entretanto, cada CH comunica-se diretamente com a BS que pode estar relativamente distante. Dessa forma, o LEACH ajuda os nodos sensores de um *cluster* a dissipar sua energia lentamente, enquanto os CHs consomem considerável quantidade de energia por seus transceptores estarem sempre ativos e por estarem distantes da estação base.

Apesar das grandes vantagens, o LEACH possui algumas deficiências. A principal deficiência do protocolo é a eleição aleatória de CHs sem considerar os níveis de energia residual dos candidatos, resultando em casos onde os CHs param de funcionar por não terem energia suficiente e os CMs desperdiçam suas energias tentando se comunicar com o CH.

O protocolo UCR elabora um adequado mecanismo para mitigar o problema de *hot spot* e um algoritmo de clusterização com eficiência energética. No entanto, o *overhead* de controle gerado pelo protocolo nos processos de reorganização de *clusters* é elevado, levando a um gasto de energia na rede, além de tratar o processo de eleição de CH e formação de *clusters* de forma separada, o que pode resultar em um estabelecimento de comunicação entre *clusters* pouco eficiente.

O protocolo HCR utiliza um algoritmo de clusterização com baixo *overhead* de controle e possui eficiência energética. No entanto, para garantir a conectividade entre os *clusters* o protocolo gera uma quantidade grande de CHs, o que acarreta em um desequilibrado e excessivo consumo de energia na rede. Além disso, o protocolo não trata o problema de *hot spot* e múltiplos critérios para formação dos *clusters*.

O protocolo EECF consegue reduzir o *overhead* de controle durante a fase de configuração dos *clusters* de forma eficiente, garantir conectividade entre os clusters a partir da métrica de grau de adjacência do nodo e uma eleição de CH com eficiência energética através do uso da métrica de energia residual no processo de eleição de CH. No entanto, o algoritmo não considera o uso de uma solução para mitigar o problema de *hot spot* na rede a fim de garantir um balanceamento do consumo de energia e evitar a partição da mesma. Além disso, não considera diversas variáveis de significância global e local para o processo de formação de cluster.

O protocolo LUCA utiliza um mecanismo de *backoff* que garante baixo *overhead* de controle na rede e estabelece *clusters* de tamanho desiguais para mitigar o problema de *hot spot*, no entanto, o mesmo desconsidera múltiplas variáveis para o processo de formação de clusters, tais como a energia residual e as distâncias entre os CMs e os CHs e os CHs em relação a BS, e não estabelece um mecanismo para garantia de conectividade na rede que impossibilite o seu particionamento.

O protocolo PARC reduz o *overhead* de controle através da integração das fases de formação dos *clusters* e roteamento entre *clusters* e elimina o problema de *hot spot* por empregar uma técnica que elege uma quantidade alta de CHs para retransmissão de dados para a BS, no entanto, o mesmo não garante conectividade e não melhora o processo de composição de *cluster* a partir de múltiplos critérios.

O protocolo EECBHC reduz o *overhead* de controle através do uso de um mecanismo de temporização que considera os níveis de energia residual dos nodos, todavia, o mesmo não garante conectividade e não melhora o processo de composição de *cluster* a partir de múltiplos critérios, além de não tratar o problema de *hot spot*.

O algoritmo LEACH-ERE prolonga o tempo de vida da rede através da métrica

ERE e do sistema *fuzzy* integrado, porém seus resultados mostraram um desbalanceado consumo de energia na rede, além de a proposta não expandir a capacidade do protocolo LEACH para atuar em redes de múltiplos saltos.

O ATCR visa prover eficiência energética na rede a partir do uso de mecanismos que possam mitigar os problemas encontrados e melhorar a comunicação em termos de consumo de energia em grandes RSSFs, que utilizam um paradigma de comunicação por múltiplos saltos e um fluxo de comunicação de muitos para um. Dessa forma, a proposta provê otimização utilizando múltiplos critério na formação de *clusters* e eleição de CHs, tamanhos desiguais de *clusters* para tratar o problema de *hot spot*, e por fim o emprego de um algoritmo baseado em temporização com o objetivo de diminuir o *overhead* de controle na rede. A tabela 1 apresenta as características de cada proposta permitindo uma análise comparativa funcional das mesmas.

Tabela 1: Comparação funcional dos protocolos

Protocolos	Conectividade	Eficiência Energética	Baixo Overhead	Hot Spot	Otimização multi critério
LEACH	Não	Sim	Não	Não	Não
UCR	Sim	Sim	Não	Sim	Não
HCR	Sim	Sim	Sim	Não	Não
EECF	Sim	Sim	Sim	Não	Não
LUCA	Não	Sim	Sim	Sim	Não
PARC	Não	Sim	Sim	Sim	Não
EEBHC	Não	Sim	Sim	Não	Não
LEACH-ERE	Não	Sim	Sim	Não	Não
ATCR	Sim	Sim	Sim	Sim	Sim

4 Desafios e Proposta

4.1 Descrição dos Desafios

4.1.1 Eleição do Cluster Head

Em virtude de os CHs consumirem mais energia na agregação e roteamento dos dados, é importante a existência de um mecanismo com eficiência energética para a escolha desses nodos [13]. Assim, a eleição de CHs é o núcleo dos algoritmos de clusterização. Em geral, os algoritmos de clusterização propostos na literatura tencionam selecionar um ótimo número de nodos que tenham um maior nível de energia residual para ser eleito CH [33].

Muitos protocolos de clusterização usam uma abordagem probabilística para escolha do CH e utilizam aleatoriedade para garantir que o papel de CH seja compartilhado entre todos os nodos da rede com o objetivo de balancear o consumo de energia. Entretanto, os algoritmos que selecionam os CHs aleatoriamente utilizando-se de uma pré-determinada probabilidade devem assegurar que os nodos estão bem distribuídos, cobrindo todo o campo de sensoriamento de maneira uniforme e que os *clusters* estejam adequadamente conexos.

Outro problema inerente a este processo é o tratamento isolado do processo de escolha de CHs em relação ao processo de descoberta de rotas para uma rede de múltiplos saltos, assim, os nodos são eleitos sem considerar a construção das rotas de comunicação entre *clusters* até a BS. Nesse caso, a eficiência da comunicação entre *clusters* dificilmente pode ser alcançada [27]. As figuras 6a e 6b demonstram o problema enfrentado.

Nesse exemplo, a figura 6a apresenta o caso onde um CH é eleito sem realizar simultaneamente o processo de estabelecimento de rotas entre *clusters*, o que acarreta em uma distância maior e conseqüentemente um maior consumo de energia, na figura 6b o processo é realizado de forma simultânea o que garante uma eficiente comunicação entre os *clusters*.

Outro problema decorrente do processo de clusterização é o *overhead* de controle gerado pelo processo, pois é exigida uma grande quantidade de mensagens a serem transmitidas para que a formação dos *clusters* seja consolidada. Assim, uma das desvantagens dos protocolos de clusterização é o *overhead* adicional de pacotes de controle gerado pelo processo de eleição de CHs e formação de *clusters* [34]. Portanto, *overhead* decorrente do

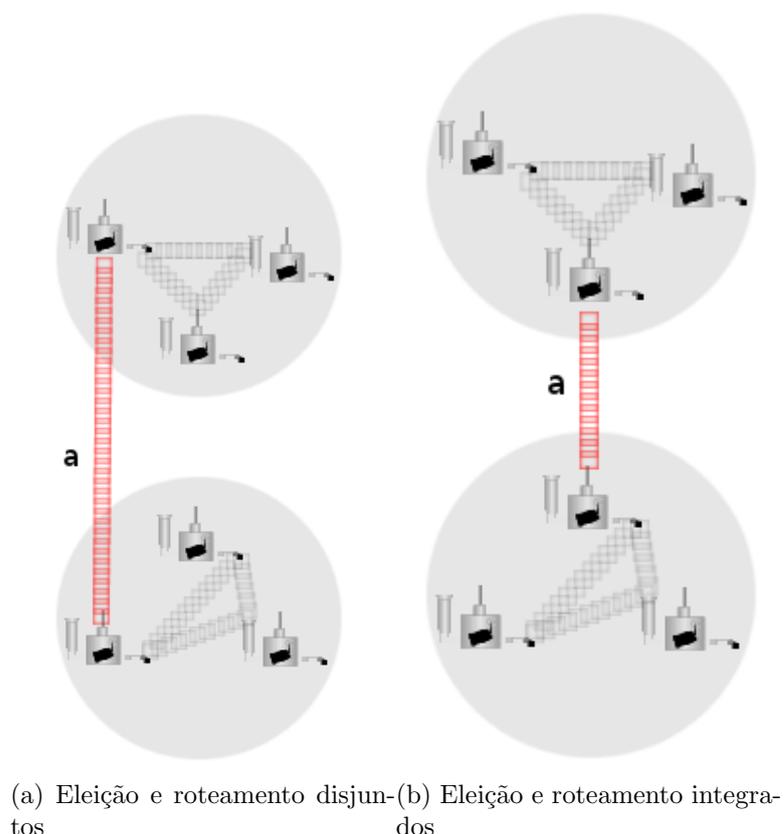


Figura 6: Problema de eleição de CHs

processo de clusterização deve ser umas das questões mais importantes a ser considerada durante o desenvolvimento de um protocolo de clusterização para RSSFs [35].

4.1.2 Formação de Cluster

A formação do *cluster* é um processo de tomada de decisão onde os CMs decidem para qual CH eles pretendem se associar para realizar a transmissão dos dados. Essa decisão normalmente se baseia em um critério que vise otimizar o uso de energia.

A escolha de um CH baseada apenas sobre um simples critério pode levar a um pobre uso de energia, isto acontece pois, um nodo sensor ao escolher um CH dentre outros candidatos para se associar pode, por exemplo, escolher um CH próximo dele porém distante da BS, assim, para determinados CMs isto pode não ser a melhor escolha. Além disso, fatores como nível de energia residual do CH e seu grau de adjacência com outros nodos pode também ser de extrema importância para a tomada de decisão [36].

O processo de tomada de decisão do CM para associar-se a um CH baseado sobre um único critério resulta em uma péssima escolha em um contexto geral da rede. A figura 7 demonstra o problema na formação de *clusters*.

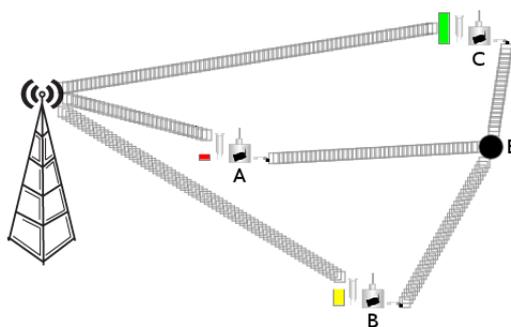


Figura 7: Problema de otimização na formação de clusters

Nesse exemplo, caso o nodo sensor E decida por escolher o nodo C para se associar, baseado apenas sobre o nível de energia residual, esta não seria uma escolha adequada, já que o mesmo está mais distante da BS que os demais nodos. Se a escolha se basear na distância para a BS o mesmo problema aconteceria já que dessa vez o nodo A teria um nível de energia residual menor que os demais. Assim, o processo de tomada de decisão para formação do *cluster* deve considerar parâmetros de local e global significância.

4.1.3 Hot Spot

Os dados transmitidos na rede seguem um padrão de difusão muitos para um, ou seja, um nodo CM coleta os dados e envia para o CH do *cluster* que transmite os dados diretamente para a BS ou através de outros CHs. Este processo de retransmissão por múltiplos saltos acontece porque em um cenário real um nodo não teria potência suficiente para alcançar a BS e mesmo que tenha a comunicação não teria eficiência energética.

Após a formação dos *clusters* os CHs possuem um consumo de energia relacionado a comunicação dentro do *cluster* e outro relativo a comunicação entre CHs. Assim, o consumo de energia do CH nas comunicações dentro do *cluster* é inversamente proporcional a quantidade de nodos que compõem o *cluster*. Entretanto, na comunicação entre CHs os nodos mais próximos da BS possuem uma carga alta de tráfego de repasse. Conseqüentemente, a energia desses nodos termina mais rápido em relação aos demais nodos causando a redução da área de cobertura do sensoriamento e o particionamento da rede próximo a BS. Dessa forma, nenhum dado poderá ser entregue para a BS, pois essa não conterà ao seu redor nenhum nodo ativo que conecte ela a rede. A figura 8 demonstra o problema de *hot spot*.

Uma BS localizada em um determinado ponto da rede possui uma área ao redor caracterizada como um gargalo, pois, todas as informações originadas fora dessa área sempre serão retransmitidas por nodos localizados dentro dessa área. Assim, um nodo dentro da área de gargalo impõe restrições sobre o tempo de vida da rede. Dessa forma, evitar a falha de nodos causada por esgotamento de energia em áreas de gargalo é fundamental para garantir um prolongado tempo de vida para rede [37].

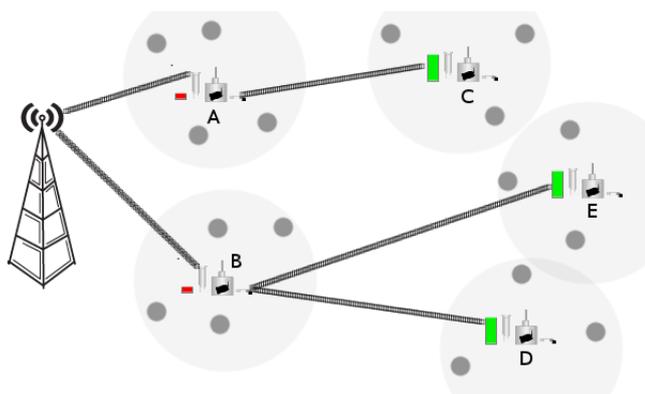


Figura 8: Problema de Hot Spot

4.2 Modelo da Rede

Este trabalho considera uma RSSF utilizada no monitoramento ambiental para coleta periódica de dados meteorológicos, em decorrência da necessidade de criação de um histórico sobre esses dados. A RSSF é constituída de n nodos sensores aleatoriamente distribuídos de acordo com uma distribuição uniforme em uma vasta área geográfica e uma BS. Para a análise em questão são considerados os seguintes pressupostos a respeito dos nodos e da rede:

- Os nodos sensores utilizam como fonte de energia uma bateria e a BS não possui restrições energéticas e está localizada dentro do campo de sensoriamento;
- Os nodos são deixados no campo de sensoriamento depois da implantação. Nesse caso, não é possível realizar a substituição das baterias;
- Os nodos sensores são uniformes em relação a capacidade de seus recursos;
- Os nodos sensores não têm capacidades de mobilidade;
- Os enlaces de comunicação são simétricos. Assim, dois nodos podem se comunicar usando a mesma potência para transmissão;
- Os nodos sensores não possuem informações de sua localização;
- Os nodos sensores podem mudar os níveis de potência dinamicamente;
- Cada nodo sensor possui um identificador único.

Cada nodo possui um transceptor com uma antena omnidirecional que cobre uma área definida por um raio de transmissão R_t e cada nodo n_i pode estender ou diminuir seu raio de cobertura ajustando a potência de transmissão do transceptor. Apesar dos nodos não possuírem informações sobre sua localização, os mesmos podem estimar sua distância em relação a BS baseado no indicador de intensidade do sinal recebido (*Received Signal Strength Indicator* - RSSI).

Os valores de RSSI muito próximos do limiar de sensibilidade do transceptor, os fatores locais do ambiente, como o ruído, desempenham um papel importante. Assim, normalmente é usado um ponto de corte para os valores de RSSI, onde para valores abaixo do limiar torna-se questionável a correlação entre RSSI e a distância. Entretanto, a necessidade de ágeis e menos custosos estimadores de qualidade de enlace tornam o RSSI melhor do que o LQI (*link quality indicator*) [38].

4.3 Algoritmo Temporizado de Clusterização e Roteamento

O ATCR, tal como o protocolo LEACH, é dividido em rodadas sendo cada rodada composta por uma fase de configuração e outra de transmissão. Na fase de configuração são escolhidos os CHs, formados os *clusters* e estabelecidos os *slots* de tempo para transmissão dos dados de cada CM do *cluster*. Na fase seguinte inicia-se a transmissão dos dados coletados pelos sensores. Cada CM irá transmitir seus dados dentro do seu espaço de tempo pré-alocado. A figura 9 descreve a estrutura de funcionamento do protocolo.

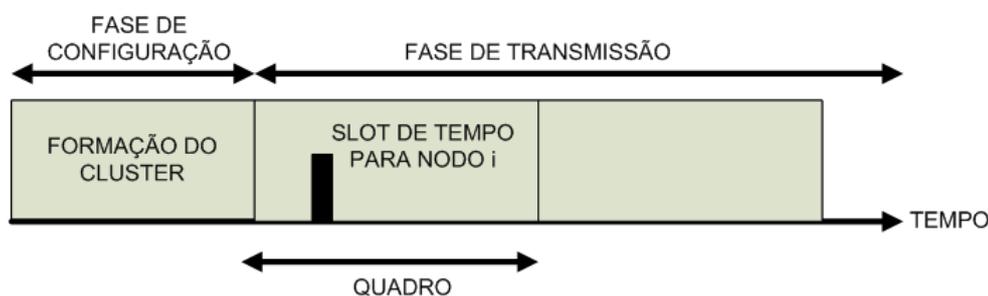


Figura 9: Linha do tempo da operação do protocolo

No início de cada rodada existe a fase de configuração e em seguida a fase de transmissão dividida em quadros, onde cada quadro é subdividido em pequenos *slots* de tempo que correspondem aos períodos em que cada CM deve transmitir, assim, uma rodada pode conter n quadros de tamanho t , sendo que, para cada quadro todos os nodos devem possuir dados para serem transmitidos para o CH, que será o responsável por criar e divulgar essa estrutura para todos os nodos de seu *cluster*.

Cada nodo possui duas faixas de transmissão: o raio de clusterização (R_c) que corresponde ao tamanho dos *clusters* e o raio de transmissão (R_t) que corresponde a distância máxima para comunicação entre os grupos. Em um ambiente real esses raios poderiam ser alcançados através da capacidade que os nodos possuem de dinamicamente alterar a potência do sinal do seu transceptor.

As características e mecanismos empregados pelo ATCR são descritos abaixo:

- Seleciona um conjunto de CHs de maneira distribuída na rede e garante simultaneamente a formação e a interconexão dos *clusters*. A seleção dos CHs é efetuada

com base em um algoritmo temporizado [39] considerando a energia residual e a distância para a BS como critérios de escolha.

- Emprega a formação de *clusters* de tamanhos desiguais, dessa forma o tamanho de R_c diminui conforme a proximidade do nodo em relação a BS, isso permitirá que os CHs mais próximos da BS reservem mais energia para transmissão do tráfego de repasse, assim não sofrerá sobrecarga permitindo um balanceado consumo de energia na rede.
- Utiliza a lógica fuzzy para realizar tomadas de decisão multicritério no processo de formação de *clusters* e escolha de CHs. O objetivo é selecionar os melhores CHs, ou seja, os que exijam um custo mínimo para comunicação, considerando como critérios a distância do nodo membro para o CH, distância entre o CH e a BS e a energia residual do CH.

Considerando os desafios explanados na seção anterior e as estratégias empregadas pelo ATCR, é possível obter eficiência energética na rede através da otimização do processo de formação dos *clusters* com o objetivo de reduzir o consumo de energia na fase de transmissão. A parte do que foi proposto, os cenários e critérios usados nos trabalhos relacionados podem ser utilizados no ATCR com pequenos ajustes. Nas próximas subseções será explicado com mais detalhes todas as estratégias empregadas pelo ATCR.

4.3.1 Algoritmo de Clusterização

O algoritmo de clusterização, presente na fase de configuração da rede, inicia com o envio pela BS de duas mensagens de controle, denominadas ADV_MESSAGE. A primeira é enviada com uma potência capaz de atingir todos os nodos da rede permitindo que os mesmos estimem sua distância em relação a BS com base no RSSI e a segunda é enviada com a mesma potência máxima dos nodos sensores da rede com o objetivo de permitir que determinados nodos saibam que os mesmos são capazes de se comunicar diretamente com a BS.

Inicialmente, cada nodo decide se será um CH de acordo com um tempo de atraso para o envio da mensagem de controle, calculado a partir de um temporizador denominado (T_b), para transmissão do sinal de controle denominado ADV_CH_ELECTED que terá como funções:

- Inibir os temporizadores de *backoff* dos nodos que estejam ainda com os mesmos ativos e dentro do R_c do CH de origem da mensagem;
- Anunciar a sua condição de CH para os nodos dentro do R_c ;
- Alertar para todos os demais nodos que estão fora do seu R_c , dentro de R_t e mais distante da BS que o nodo poderá atuar como um CH retransmissor;

O valor desse tempo varia dentro de um intervalo fechado de 0 a 1 determinado pelo sistema *fuzzy* que será descrito na próxima seção.

A determinação do R_c pelos CHs é calculado pela equação 4.1:

$$R_c = (R_c - (F_r \times R_c)) + p - p \times e^{-\frac{(D_{bs})^2}{2\sigma^2}} \quad (4.1)$$

onde:

- R_c → Raio de clusterização em metros;
- F_r → Fração sobre o raio de clusterização para determinação do raio mínimo;
- D_{bs} → Distância entre o nodo sensor e a BS em metros;
- σ → A variação crescente ou decrescente de cada valor da função;
- p → Valor de ponderação para equação gaussiana;

O comportamento da equação que determina R_c é demonstrado na figura 10, considerando $R_c=80$ e $F_r=0.3$, os valores de p e σ são determinados heurísticamente conforme observações sobre o gráfico até se atingir a curva desejada. O gráfico demonstra que o R_c cresce linearmente conforme a distância entre o CH e a BS aumenta, formando *cluster* desiguais que permitem a diminuição do problema de *hot spot* descrito na seção 4.1.3.

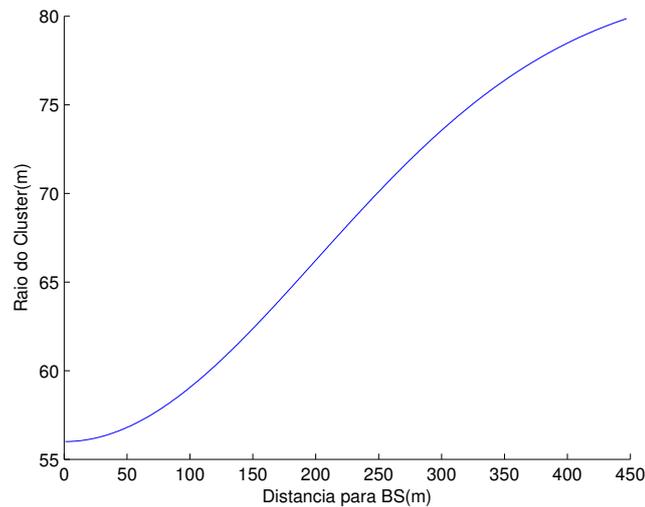
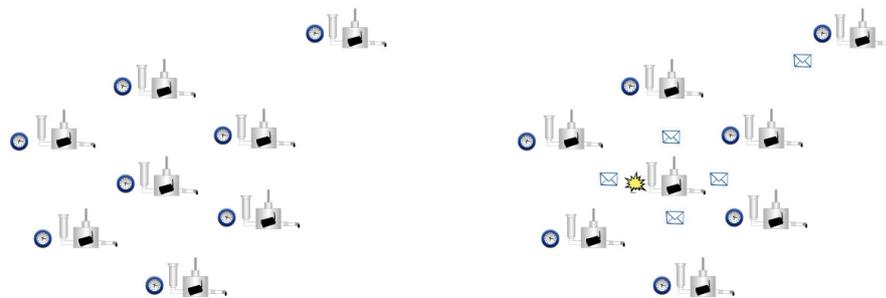


Figura 10: Função de R_c

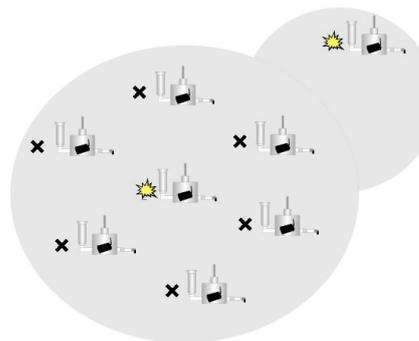
Após a escolha de todos os CHs da rede, cada nodo possuirá informações sobre os outros nodos vizinhos que se encontram dentro do seu R_c e R_t , essas informações serão úteis para decidir sobre a escolha de um deles para atuar como seus nodos retransmissores para o caso de CHs na comunicação entre *clusters* e para decidir sobre a escolha de um

CH para composição de um *cluster* para o caso de CMs. Essa tomada de decisão será realizada utilizando lógica *fuzzy* e será explicado na próxima seção.

A figura 11 demonstra, através de um pequeno exemplo, o funcionamento do ATCR em sua operação de configuração em uma rodada. Na figura 11a inicialmente o temporizador de cada nodo sensor é inicializado com o valor determinado pelo sistema *fuzzy* integrado a cada nodo, em seguida na figura 11b um dos nodos tem o seu temporizador finalizado, tornando-se CH, enquanto os demais nodos ainda permanecem com seus temporizadores ativos. Em virtude da finalização do temporizador do nodo o mesmo envia, com potência máxima, uma mensagem de controle na rede com o objetivo principal de desativar o temporizador ativo dos demais nodos que estejam dentro de R_c conforme a figura 11c. O pseudocódigo 1 denominado ATCR descreve a operação do algoritmo.



(a) Inicialização dos temporizadores (b) Finalização de um temporizador e envio da mensagem de controle



(c) Inibição dos temporizadores e eleição de CH e CM

Figura 11: Operação do ATCR

O processo de formação de *clusters* depende da escolha pelos nodos CMs de nodos CHs para se associarem e formarem um *cluster*. Esse processo depende dos custos calculados a partir do mesmo sistema *fuzzy* utilizado para eleição de CH, porém com diferentes finalidades, pois nesse processo o custo associado não é utilizado como um temporizador, mas como um valor de custo para o enlace. Dessa forma, quanto maior o custo calculado melhor será o enlace de comunicação entre o nodo CM e o CH.

Algorithm 1 ATCR

StartUp

- 1: **if** BS **then**
- 2: Broadcast *ADV_MESSAGE* message (*ID*)
- 3: **end if**
- On receiving a *ADV_MESSAGE* from BS**
- 4: estimateDistance(*ADV_MESSAGE.RSSI*)
- 5: $N_i.isNeighbor \leftarrow True$
- T_b setup**
- 6: $T \leftarrow FuzzySystem(ResidualEnergy, Distance)$
- 7: $N_i.setTimer(T_1, T_b)$
- On expire timer T_1**
- 8: $N_i.isCH \leftarrow True$
- 9: Broadcast *ADV_MESSAGE_ELECTED* message ($N_i.ID, N_i.R_c$)
- On receiving a *ADV_MESSAGE_ELECTED* from node N_j**
- 10: **if** $d(N_i, N_j) < N_j.R_c$ & *isNotCH* **then**
- 11: $N_i.cancelTimer(T_1)$
- 12: $N_i.isCM \leftarrow TRUE$
- 13: $N_i.S_{join} \leftarrow FuzzySystem(ResidualEnergy, Distance)$
- 14: **end if**
- 15: **if** $d(N_i, N_j) < N_i.R_t$ *isCH* **then**
- 16: $N_i.S_{relay} \leftarrow ID$
- 17: **end if**

4.3.2 Otimização MultiCritério Utilizando Lógica Fuzzy

Para tratar o processo de tomada de decisão para formação de *clusters* e eleição de CHs permitindo a otimização multicritério, utilizou-se um sistema *fuzzy* de Takagi e Sugeno (TSFS) [40] que aplica uma lógica difusa sobre um raciocínio multidimensional. O sistema *fuzzy* utilizado realiza o mapeamento da entrada de dados *crisp* para a saída de dados *crisp*. O sistema é constituído de quatro módulos: uma base de regras, uma máquina de inferência, um módulo fuzzificador e defuzzificador, a organização do sistema é apresentada na figura 12.

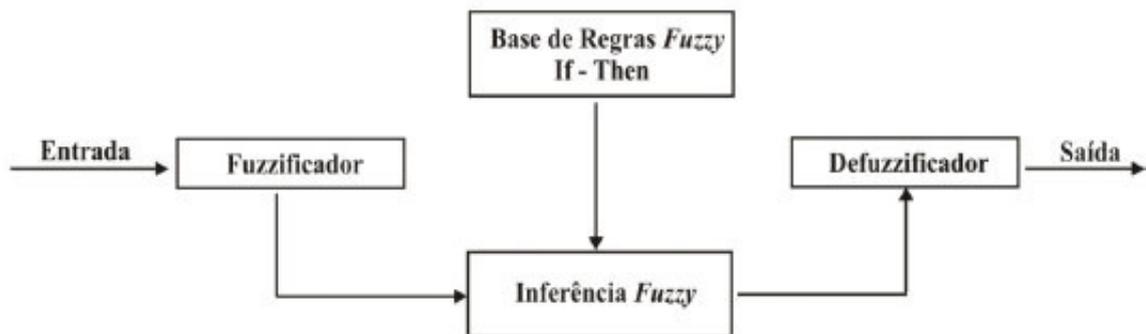


Figura 12: Organização do sistema fuzzy

No TSFS, a implicação fuzzy é baseada sobre a partição do espaço de entrada, assim em cada sub-espço, uma relação linear de entrada e saída é formada. A saída do raciocínio *fuzzy* é dada pela agregação dos valores inferidos por algum mecanismo de implicação que foi aplicado para a entrada [40]. As variáveis linguísticas do sistema *fuzzy* são a energia residual e distância, expressos na forma de porcentagem para normalizar os valores de energia em diferentes nodos sensores e expandir o propósito do algoritmo para outros nodos sensores com diferentes tipos de transeptores.

A capacidade para determinar os apropriados tipos de função de pertinência e operações *fuzzy* dentro do contexto de cada aplicação em particular é crucial para tornar a teoria de conjuntos *fuzzy* útil. As características de um sistema *fuzzy* dependerão do tipo de aplicação em que será empregado. Diferentes aplicações irão requerer diferentes operações *T-norm*, funções de pertinência ou algoritmos de *defuzzification*. Assim, o mesmo problema pode ser resolvido por diferentes projetistas usando sistemas *fuzzy* com diferentes características.

Diversos experimentos foram realizados para determinar quais tipos de funções de pertinência melhor representam as variáveis linguísticas. Foram examinados os resultados usando as funções trapezoidal, triangular, forma de sino, gaussiana, forma de Z e S. Para representação das variáveis linguísticas de entrada baixa, alta, pequena e grande, os graus de pertinência para esses conjuntos devem permanecer constantes a partir de determinados limiares máximos da função. Assim, foram escolhidas as funções em forma de *S* e *Z*, pois demonstram de forma mais adequada o comportamento desejado para as variáveis linguísticas. As equações 4.2 e 4.3 descrevem as funções.

$$f_s(x) = \begin{cases} 0, & x < a \\ 2x^2, & a \leq x \leq \frac{a+b}{2} \\ 1 - 2(1-x)^2, & \frac{a+b}{2} \leq x \leq b \\ 1, & x \geq b \end{cases} \quad (4.2)$$

$$f_z(x) = \begin{cases} 1, & x \leq a \\ 1 - 2 \left(\frac{x-a}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ \left(\frac{b-x}{b-a} \right), & \frac{a+b}{2} \leq x \leq b \\ 0, & x > b \end{cases} \quad (4.3)$$

Para representação das variáveis linguísticas intermediárias de entrada do sistema e da função de custo de saída foram realizados experimentos e os resultados foram insatisfatórios utilizando funções trapezoidal e triangular e pouco satisfatórios com funções em forma de sino. Com a função gaussiana foram alcançados bons resultados, pois essas descrevem melhor as não-linearidades e incertezas sobre as variáveis. A equação 4.4 descreve a função gaussiana.

$$f(x, \sigma, c) = e^{-\frac{(x-c)^2}{2\sigma^2}} \quad (4.4)$$

As figuras 13a e 13b apresentam as funções de pertinência de entrada do sistema.

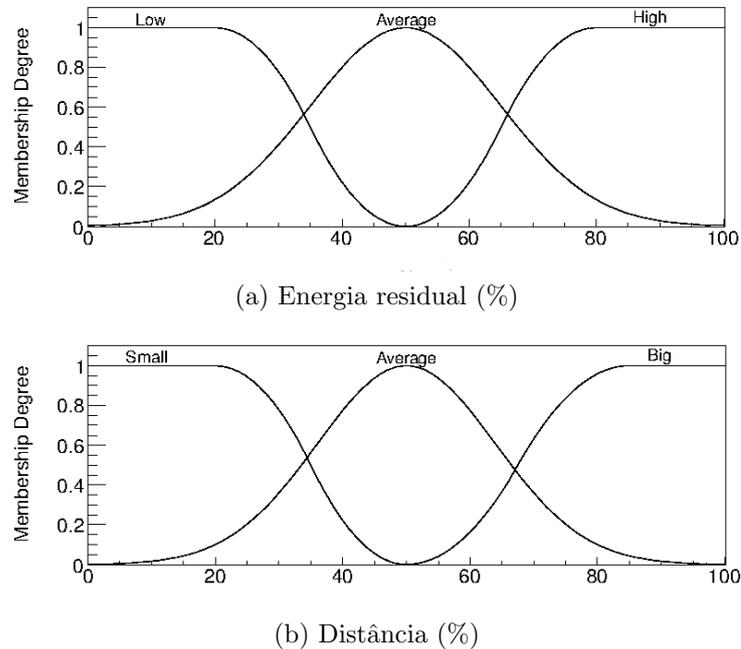


Figura 13: Funções de pertinência

As regras são expressas como implicações lógicas na forma *IF-THEN* dentro de um mapeamento de conjuntos fuzzy de entrada para funções de saída. As regras foram determinadas com base em uma análise sobre todo comportamento da rede realizando extensivas simulações, assim as regras que resultam em um custo muito baixo asseguram uma excelente chance para eleição de nodos com altos níveis de energia e mais distantes da BS. Os nodos eleitos CH com níveis de energia médios e baixo podem tornar-se inoperantes rapidamente em função do término da bateria, assim nodos com essas características devem ser evitados.

A distância do nodo para BS é uma variável muito importante, pois os nodos próximos para BS requerem menor quantidade de energia para comunicação, porém estão sujeitos ao tráfego de repasse. Dessa forma, os nodos distantes em relação a BS são preferíveis para atuar como CH para evitar problemas de desconexão na rede em virtude do desativamento por falta de energia dos nodos próximos a BS, além de permitir o balanceamento do consumo de energia na rede.

A saída das regras são funções de ordem zero e as regras são expressas conforme a equação 4.5.

$$IF f(x_1 is A_1, \dots, x_k is A_k) THEN y = p_0 \quad (4.5)$$

A tabela 2 demonstra as regras de inferência fuzzy usadas no sistema.

A máquina de inferência foi projetada de acordo com a abordagem dirigida a dados, assim, os dados disponíveis são alimentados no sistema especialista que se utiliza

Tabela 2: Regras de inferência fuzzy

ENERGY	DISTANCE	COST	COST VALUES
Alto	Grande	Muito Baixo	y=0
Alto	Médio	Baixo	y=0.1
Alto	Pequeno	Moderadamente Baixo	y=0.2
Médio	Grande	Ligeiramente Baixo	y=0.4
Médio	Médio	Médio	y=0.5
Médio	Pequeno	Ligeiramente Alto	y=0.6
Baixo	Grande	Moderadamente Alto	y=0.8
Baixo	Médio	Alto	y=0.9
Baixo	Pequeno	Muito Alto	y=1

deles para avaliar a relevância nas regras produzidas e projetar todas as possíveis conclusões. Dessa forma, a implicação *fuzzy* R_i é definida por um conjunto de regras, onde cada regra possui o seguinte formato:

$$IF f(x_1 is A_1, \dots, x_k is A_k) THEN z = g(x_1, \dots, x_k) \quad (4.6)$$

onde:

- $\{x_0, \dots, x_k\} \rightarrow$ É o conjunto de entradas do sistema;
- $\{A_1, \dots, A_k\} \rightarrow$ É o conjunto de funções de pertinência que definem a regra;
- $z \rightarrow$ A saída da regra;
- $f \rightarrow$ Uma função lógica que conecta as proposições na premissa;
- $g \rightarrow$ A função que implica sobre o valor de y ;

Para cada implicação R_i , z_i é calculado pela função g_i no consequente, conforme a equação 4.7.

$$z_i = p_0 + p_1x_1, \dots, +p_kx_k. \quad (4.7)$$

Onde p é a ordem da função de saída. Assim, o grau de verdade da proposição é calculado pela equação 4.8.

$$w_i = \prod_{i=1}^n A(x_i) \quad (4.8)$$

onde:

- $A(x_i) \rightarrow$ É o grau de pertinência de x_0 ;

- $w_i \rightarrow \acute{E}$ o grau de verdade da regra R_i ;

A saída z inferida de n implicações é dada pela média ponderada de $\{z_1, \dots, z_n\}$, conforme a equação 4.9.

$$z_i = \frac{\sum_{i=1}^n w_i z_i}{\sum_{i=1}^n w_i} \quad (4.9)$$

4.3.3 Análise e Corretude do Algoritmo

Esta seção apresenta a análise do algoritmo ATCR de acordo com o pseudocódigo 1. Inicialmente é analisada a complexidade do algoritmo proposto e em seguida serão demonstrados a partir dos lemas como o algoritmo consegue convergir para a formação de uma rede conexa.

Proposição 1. *O protocolo ATCR possui uma complexidade algorítmica em troca de mensagens de $O(n)$ na rede.*

Demonstração. Considerando a rede composta de N nodos sensores, durante a operação do protocolo x na fase de eleição de CH, os nodos CH eleitos difundem uma mensagem de controle, denominada *ADV_MESSAGE_ELECTED*, assim, se $k \times$ CH nodos são eleitos, eles enviam $k \times$ *ADV_MESSAGE_ELECTED*, seguido de $(N - K) \times$ *JOIN_CH_MESSAGE* transmitidos dos nodos CM. Portanto, o numero de mensagens trocadas na rede é limitada por $N \times C + K + (N - K) = N \times T$, onde C é o tempo associado ao custo para eleição de CHs, assim, o protocolo ATCR possui complexidade de $O(n)$. \square

Proposição 2. *O protocolo ATCR é completamente distribuído.*

Demonstração. Em cada rodada do algoritmo, um nodo pode se eleger para tornar-se um CH de acordo com o seu temporizador T_b , cada nodo CH pode eleger o seu nodo de repasse através do overhead gerado pelo envio das mensagens *ADV_MESSAGE_ELECTED* em R_t e um nodo pode se associar a um cluster através do overhead gerado pelo envio das mensagens *ADV_MESSAGE_ELECTED* em R_c , assim, mostra-se que o algoritmo é completamente distribuído. \square

Lema 1. *No final do algoritmo de eleição de CH, cada nodo é um CH ou um CM.*

Demonstração. Não há como assumir no algoritmo que ao seu término um nodo não seja um CH ou um CM, pois, isto implicaria na existência de uma contradição nas condições a serem satisfeitas nas linhas 8 e 10 do algoritmo, o que não é permitido em função da existência do temporizador que quando finalizado impede que a condição da linha 10 seja satisfeita e caso contrário, satisfeita a condição da linha 10, o mesmo é inibido antes de sua finalização, dessa forma, sempre uma proposição será uma negação e outra uma afirmação. \square

Lema 2. *No final do algoritmo de eleição de CH, cada nodo CH possuirá um nodo de retransmissão ou transmitirá diretamente para a BS.*

Demonstração. No início do algoritmo a BS difunde uma mensagem de controle usando a mesma potência máxima que os nodos sensores podem utilizar permitindo que os nodos a um salto de distância da BS saibam que possuem comunicação direta com o mesmo. Para os demais CHs, quando ocorre a finalização do temporizador, o mesmo difunde a mensagem *ADV_MESSAGE_ELECTED* que quando recebida por algum nodo em R_t e $D_{bs}(i) < D_{bs}$ o mesmo registra o CH como nodo candidato de repasse. \square

Teorema 1. *O algoritmo de eleição de CH termina com a geração de um grafo conexo.*

Demonstração. Satisfeitos os lemas 1 e 2 a comunicação ocorre através da transmissão dos dados do CH para a BS diretamente (um salto) ou através dos nodos de repasse (múltiplos saltos) e a transmissão dos dados entre os nodos sensores e a BS ocorre através dos CHs por um único salto, portanto, a rede (grafo) é conexa. \square

4.3.4 Considerações Finais

Nessa seção foi apresentado os problemas relacionados a formação de clusters em RSSF e proposto uma solução na forma de um algoritmo. Esta solução teve como objetivo primordial, encontrar um equilíbrio entre as necessidades e características das RSSF, que comportam-se como uma espécie de balança de dois pratos, onde a extrema otimização de uma determinada característica pode tornar outras ineficientes. Finalmente é realizada a análise do algoritmo afim de comprovar a sua operacionalidade e escalabilidade na rede.

5 Avaliação de desempenho

5.1 Metodologia de Avaliação

Com o objetivo de avaliar o desempenho do algoritmo proposto e realizar uma análise comparativa com outros algoritmos relacionados foi desenvolvido um modelo de simulação que permite uma ampla e flexível variedade de testes considerando diversos cenários e parâmetros.

O tipo de simulação utilizada é a simulação de eventos discretos, que trabalha segundo uma sucessão de ações condicionadas a ocorrência de um evento que provoque mudanças de estado no sistema, bem como a geração de novos alertas de eventos no futuro [41]. Existem dois tipos de simulação de eventos discretos dependendo do tipo de medições que estão sendo estimadas: simulação de terminação e de estado estável. Dessa forma, a coleta e análise estatística sobre os dados de ambos os tipos de simulações são diferentes.

A simulação de terminação é uma simulação em que o modelo tem específicas condições de início e término como uma reflexão natural de como o sistema realmente opera. Como o próprio nome sugere, a simulação terminará de acordo com uma regra ou condição especificada pelo modelo. O espaço temporal da simulação tem um fim natural e bem definido, bem como uma clara e definida forma para iniciar [42]. Para exemplificar considere uma loja que abre suas portas sem clientes presentes às 9 horas da manhã e fecha as mesmas às 9 da noite, mas continua operando por pouco tempo até que todos os clientes deixem a loja.

A simulação de estado estável é uma simulação em que a quantidade a ser estimada é definida em longas execuções, ou seja, sobre um espaço de tempo infinito. Para este tipo de simulação, o dado da performance após o sistema ter alcançado o estado estável da simulação, é o que interessa para fins de resultados finais, para tanto é necessário o emprego de heurísticas, tais como longas execuções, truncamento, exclusão de dados iniciais e média móvel de repetições independentes [43]. Estas heurísticas possibilitam determinar o fim do estado transitório, que constitui a parcela inicial da simulação que é desprezada para os resultados finais.

O modelo de simulação empregado nesse trabalho caracteriza-se como de terminação, pois o estado operacional de uma RSSF possui a condição de permanecer ativa em função do nível de energia dos nodos sensores, dessa forma a partir do término de

toda a carga de energia de todos os nodos da rede, esta última torna-se inoperante.

A coleta dos dados é conceitualmente simples, sendo necessária apenas a execução de n repetições independentes e a coleta das observações através das repetições das variáveis aleatórias escolhidas. Com o conjunto inicial de dados coletados calcula-se a média amostral, o desvio padrão e o intervalo de confiança. Inicialmente realiza-se uma quantidade arbitrária de simulações onde é extraído o intervalo de confiança. A partir disso objetiva-se a redução deste último a partir do aumento da quantidade de repetições a serem realizadas, porém deve-se determinar essa quantidade a partir da equação 5.1.

$$n \cong n_0 \times \frac{h_0^2}{h^2} \quad (5.1)$$

onde:

- n_0 → Número inicial de simulações realizadas;
- h_0 → Valor do intervalo de confiança (IC) calculado sobre os resultados da variável escolhida.
- h → Valor desejado para o intervalo de confiança.

Foi desenvolvida uma heurística para determinar o valor desejável para h , dessa forma, inicialmente definiu-se um quantitativo de 10 simulações iniciais, assim, $n_0=10$. A variável utilizada é a média da energia remanescente (E_r) de cada nodo da rede em função da quantidade de rodadas (r) realizadas pelo algoritmo de clusterização. A fórmula 5.2 estabelece que h é a média dos valores de IC de E_r .

$$h = \frac{\sum_{i=1}^r IC(E_r)}{r} \quad (5.2)$$

A variável h_0 é definida como o menor valor de IC calculado sobre E_r .

Nesta avaliação de resultados realizou-se inicialmente 10 simulações, utilizando a energia residual do ATCR como variável experimental e IC de 95%. Os dados correspondentes aos valores de E_r extraídos das simulações encontram-se no apêndice A. Calcula-se h e n a partir das equações 5.3 e 5.4 respectivamente.

$$h = \frac{\sum_{i=1}^r IC(E_r)}{r} \quad h = \frac{0.892404}{25} \quad h = 0.03569616 \quad (5.3)$$

$$n \cong n_0 \times \frac{h_0^2}{h^2} \quad n \cong 10 \times \frac{0.113141^2}{0.03569616^2} \quad n \cong 100 \quad (5.4)$$

Para simplificar a carga de trabalho com simulações mantendo o nível de IC aceitável, optou-se por aumentar o valor de h para 0.0508337, assim a quantidade de simulações necessárias é dada pela equação 5.5.

$$n \cong n_0 \times \frac{h_0^2}{h^2} \quad n \cong 10 \times \frac{0.113141^2}{0.0508337^2} \quad n \cong 49.5 \quad n \cong 50 \quad (5.5)$$

Com o objetivo de assegurar que o modelo de simulação desenvolvido foi implementado de maneira correta e que representa o sistema em seu ambiente real, foram realizados diversos testes de verificação e validação do modelo.

As técnicas de verificação utilizadas foram a *Run Simplified Cases* e *Trace* [43]. A técnica de validação utilizada foi a *Expert Intuition* [43], onde para realização desta, publicou-se o modelo desenvolvido no sitio do grupo de pesquisa GERCOM e foi divulgado no grupo de discussão do simulador Castalia a disponibilização do modelo para uso sob licença GPL. Dessa forma, foi possível a realização de um *brainstorming* entre especialistas conhecedores do sistema que utilizaram o modelo em seus trabalhos e avaliaram os resultados gerados pelo modelo. Foram obtidos diversos apontamentos que culminaram no lançamento de versões corretivas e evolutivas do modelo de simulação desenvolvido.

5.2 Descrição do Cenário e Análise dos Resultados

Para avaliarmos o desempenho do ATCR foi desenvolvido um modelo de simulação no simulador de eventos discretos Castalia [44]. A tabela 3 descreve os parâmetros de simulação utilizados para o cenário proposto e para o ATCR.

Tabela 3: Parâmetros do cenário e do algoritmo

Parâmetro	Valor
Área de Sensoriamento	400m X 200m
Localização da BS	(0m,0m)
Quantidade de Nodos	400
Energia Inicial	20J
Tamanho do pacote de dados	26bytes
Modelo do Transceptor	CC2530 ZigBee-ready
Expoente de Path Loss	2.0
Modelo do Canal	Simétrico
Raio de Clusterização	80m
Índice de Raio Mínimo	0.3
Tempo da Rodada	20s
Tamanho do Slot de Transmissão	0.2s

A avaliação de performance do ATCR considerou de forma geral os seguintes aspectos: eficiência energética, entrega de pacotes e *overhead* de controle. A avaliação da eficiência energética considera de forma específica diversos parâmetros a serem avaliados, esses serão detalhados no decorrer desta seção através do estudo de gráficos que apresentam de forma resumida os resultados obtidos a partir das simulações.

A performance do ATCR foi comparada apenas com os protocolos HCR [27] e UCR [26], pois ambos os protocolos tratam todas restrições, citadas nesse trabalho, relacionadas com eficiência energética em RSSFs, o que torna dispensável a comparação com os demais protocolos citados nos trabalhos relacionados para evitar a redundância.

A figura 14 apresenta o gráfico que demonstra a média da quantidade de *clusters* formados na rede em porcentagem. A figura mostra que o ATCR e o UCR geram em média quantidades de *clusters* próximas, sendo que o UCR consegue formar em média menos *clusters* que o ATCR em função do seu algoritmo competitivo eliminar qualquer mínima possibilidade de sobreposição de *clusters* na rede. No entanto, o ATCR gera uma quantidade de *clusters* suficiente para dar cobertura a toda a rede sem perder a eficiência energética. O HCR apresenta uma quantidade elevada e exagerada de *clusters* formados na rede em função do seu algoritmo garantir sobre qualquer circunstância a conectividade da rede, isso melhora o QoS da rede porém prejudica a eficiência energética, como será demonstrado nos próximos gráficos.

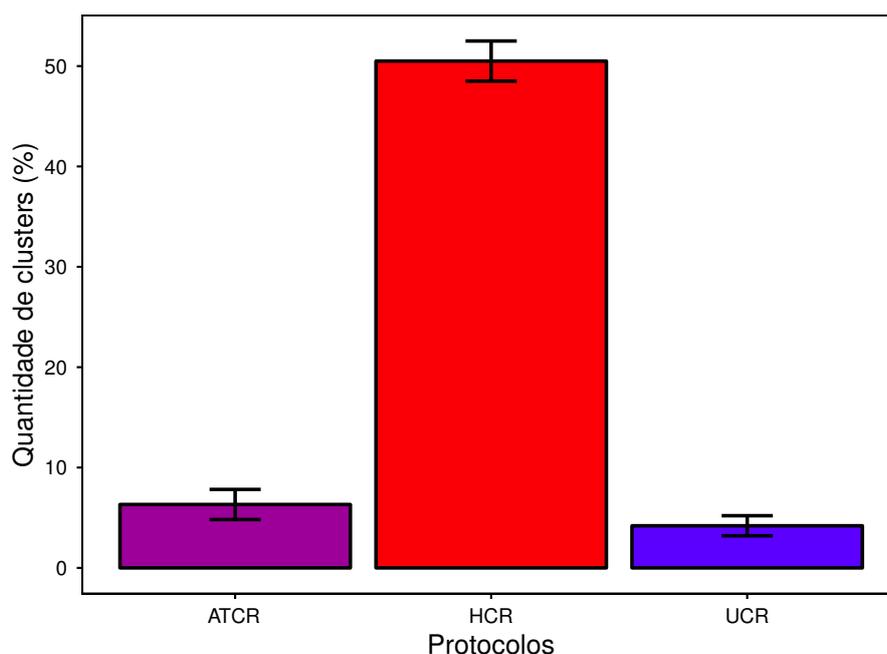


Figura 14: Média da quantidade de clusters formados

A figura 15 apresenta o gráfico que demonstra o tempo de vida da rede. A partir desse gráfico é possível obter informações sobre a duração em que a rede permaneceu ativa, a partir da distribuição de nodos sensores vivos em relação ao número de rodadas, e o nível de balanceamento do consumo de energia na rede, a partir do grau de curvatura da linha do gráfico. A figura claramente mostra que o ATCR consegue prolongar o tempo de vida da rede em relação aos demais algoritmos, pois os primeiros nodos sensores da rede utilizando o ATCR começam a morrer posteriormente em relação aos demais protocolos e continua linearmente até todos os nodos estarem mortos.

A partir do gráfico é possível observar também que o ATCR e o UCR possuem

um grau de curvatura pequeno no gráfico demonstrando que eles possuem um consumo balanceado de energia na rede, ao contrário do protocolo HCR que possui diversas irregularidades na sua linha correspondente no gráfico, o que denota um consumo de energia desbalanceado na rede. Essa observação será detalhada e visualizada com mais clareza em outros gráficos que serão apresentados posteriormente.

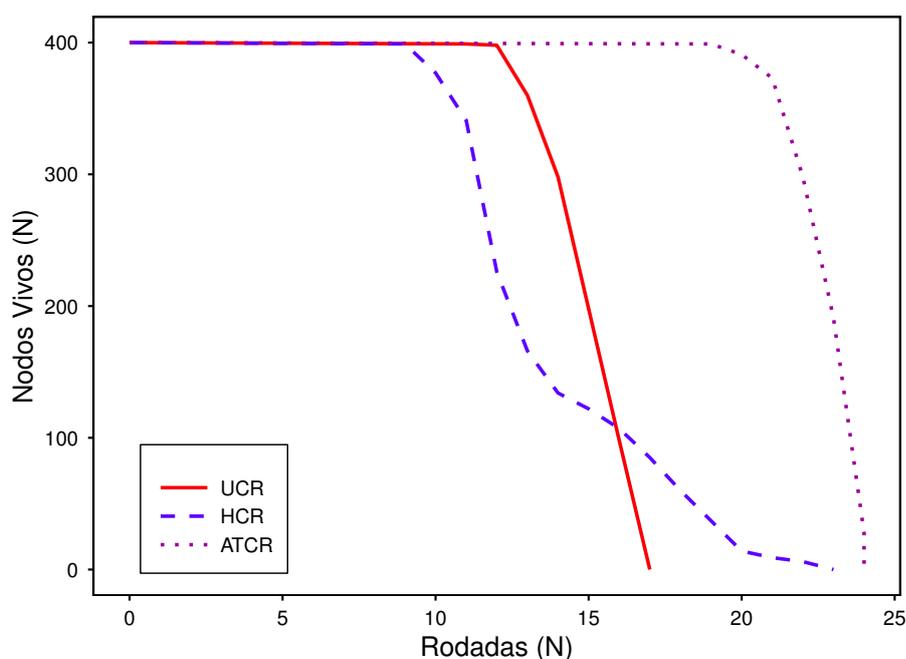


Figura 15: Nós vivos por rodadas

A figura 16 apresenta o gráfico que demonstra o desvio padrão da energia residual em relação ao número de rodadas. A partir desse gráfico é possível obter informações sobre o balanceamento do consumo de energia na rede. Assim, quanto maior o desvio sobre a energia residual maior será o desbalanceamento do consumo de energia entre os nós sensores, com esse indicador é possível demonstrar o quanto os algoritmos conseguem minimizar o problema de *hot spot* descrito no capítulo 4, seção 4.1.3.

Uma análise sobre o gráfico aponta que o ATCR consegue ser mediano em relação a mitigação do problema de *hotspot*, enquanto o protocolo UCR, que tem como foco principal a diminuição desse problema, consegue ótimos resultados. Isso decorre dos requisitos aplicados ao processo de formação de *clusters* visto que o UCR foca em um algoritmo competitivo que elimina sobreposições entre clusters, além da composição de clusters desiguais.

O ATCR tenta minimizar o *hotspot* a partir da composição de *clusters* desiguais, porém o algoritmo prioriza a composição de clusters distantes da BS, afim de evitar a desconexão da rede. Esse fator prejudica o balanceamento do consumo de energia, porém melhora o protocolo em outros pontos, como por exemplo o tempo de vida da rede e a entrega de pacotes para BS. O protocolo HCR possui os piores resultados por não estabelecer nenhuma técnica para mitigar o problema de *hotspot*.

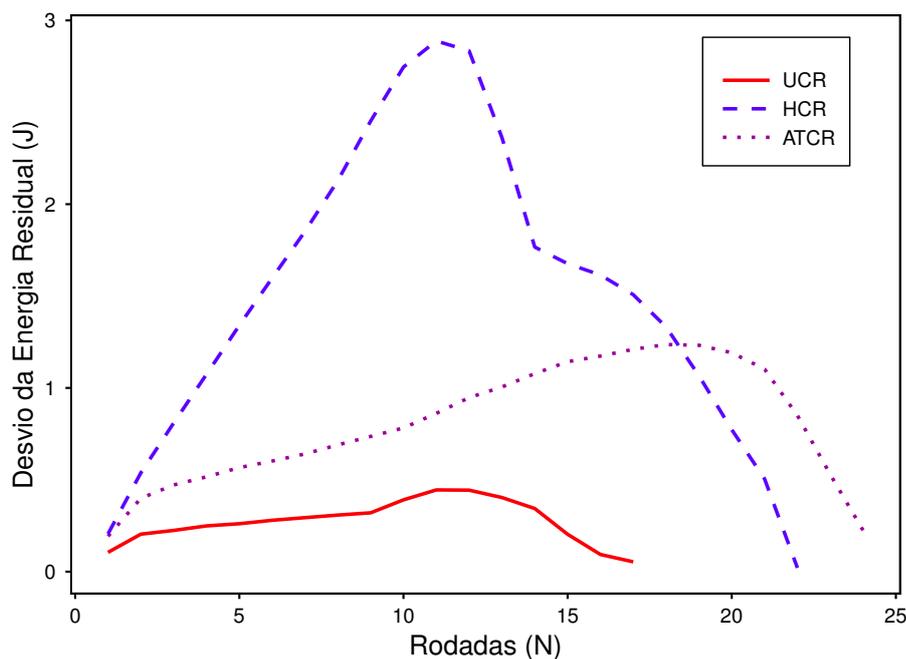


Figura 16: Desvio padrão da energia residual

A figura 17 apresenta o gráfico que demonstra a energia residual em relação ao número de rodadas. A partir desse gráfico é possível obter informações sobre o consumo de energia na rede. Os resultados demonstram que o ATCR consegue economizar mais energia através das rodadas que os demais protocolos, em função dos mecanismos de otimização multicritério empregados no processo de eleição de CH e formação de *cluster* que visam a tomada de decisão baseada em variáveis que afetam o consumo de energia na rede.

A figura 18 apresenta o gráfico que demonstra uma métrica denominada Metade dos Novos Vivos - *Half of the Nodes Alive (HNA)* [45]. A partir desse gráfico é possível obter um valor estimado para a rodada em que metade dos nodos sensores estão mortos. Esta métrica é útil para redes densamente implantadas, que é uma característica do cenário utilizado nas simulações. Os resultados demonstram que o ATCR com metade dos nodos vivos é superior aos demais protocolos, com relação a rodadas decorridas, em mais de 50%.

As figuras 19a, 19b e 19c apresentam o gráfico de densidade que demonstra a energia residual dos nodos na metade do total das rodadas que compõem o tempo de vida da rede em relação a cada protocolo. Através desse gráfico é possível visualizar o grau de distribuição do consumo de energia na rede relacionado a distância do nodo para BS. Os pontos representam os nodos sensores a uma dada distância da BS com determinado nível de energia e a mancha em azul representa a densidade de nodos em uma dada região do plano cartesiano.

Através desses gráficos pode-se visualizar que o HCR possui um grau de dispersão

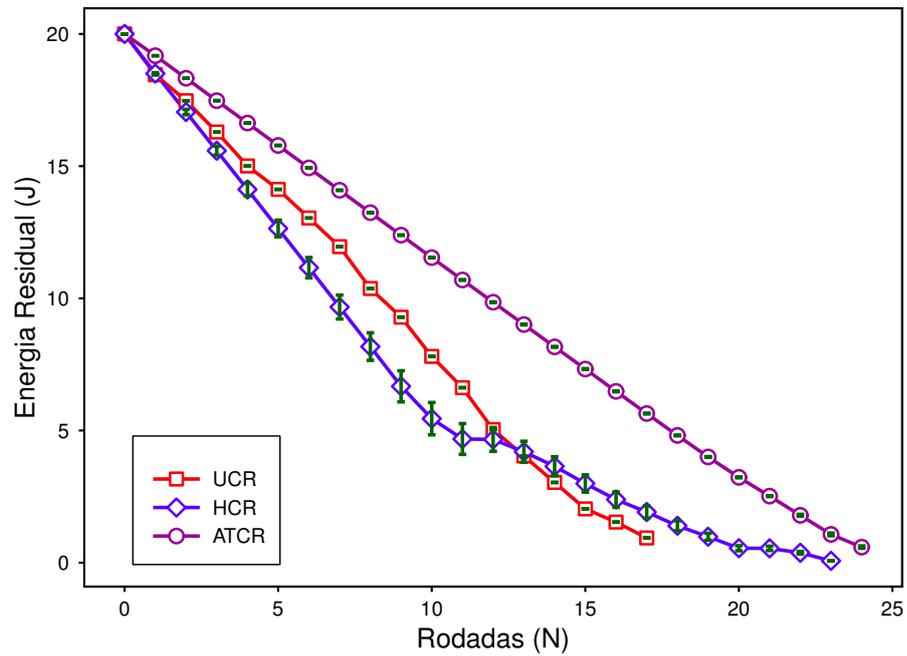


Figura 17: Energia residual

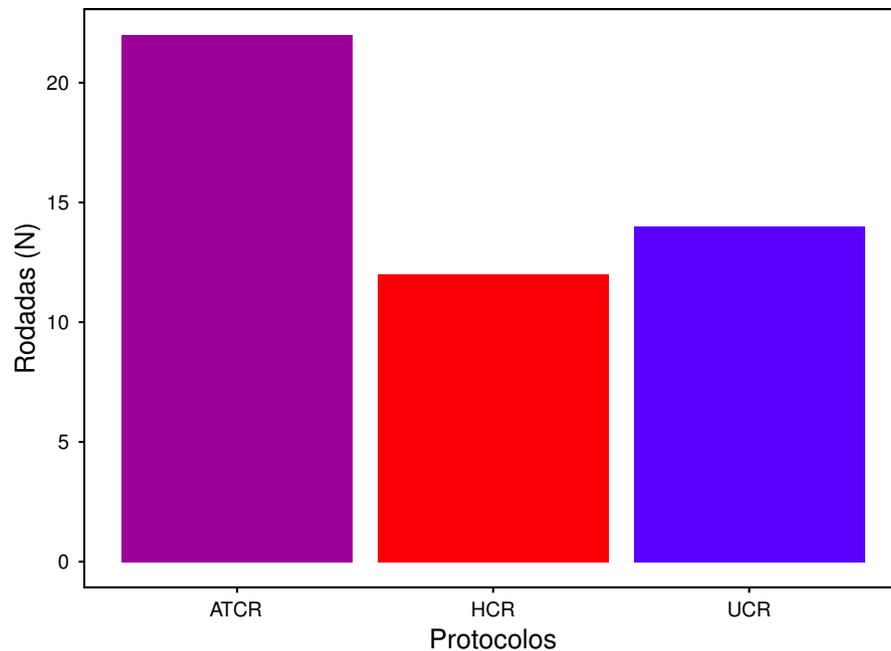


Figura 18: Metade dos Nodos Vivos

muito grande quanto ao consumo de energia na rede o que caracteriza um desbalanceado consumo de energia na rede, por outro lado o UCR consegue agregar uma concentração de nodos maior em determinado nível de energia, o que caracteriza um consumo uniforme de energia na rede. O ATCR consegue minimizar o grau de dispersão da energia

residual, porém não consegue ser melhor que o UCR, como já demonstrado e explicado anteriormente na análise de desvio padrão da energia.

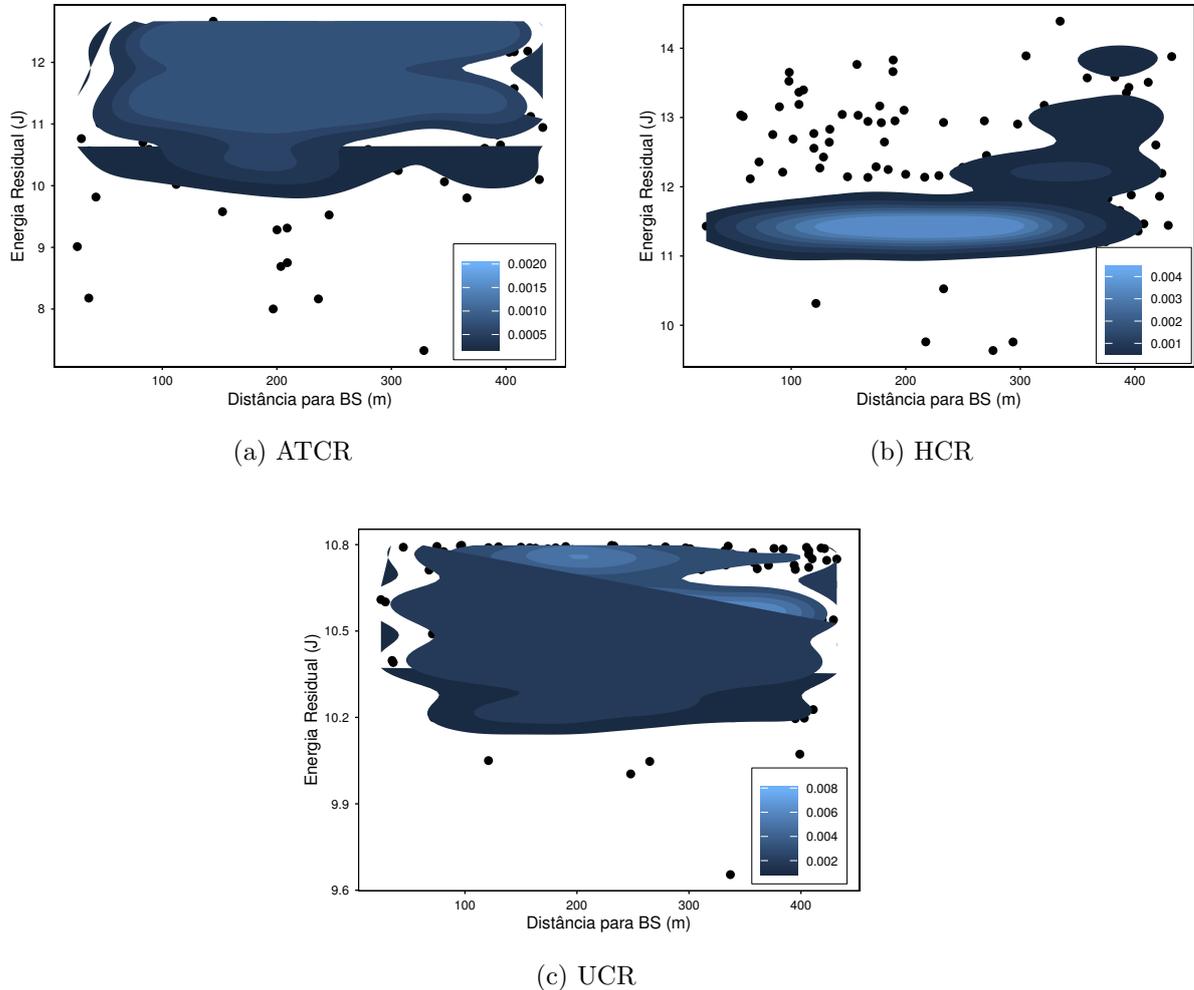


Figura 19: Mapa de energia da rede

A figura 20 apresenta o gráfico que demonstra a latência na entrega dos pacotes para BS. A partir desse gráfico é possível obter informações sobre a média e a variabilidade do atraso na entrega dos pacotes para BS. Os resultados demonstram que o UCR possui a menor variabilidade apesar de apresentar muitos pontos discrepantes, o HCR apresenta muita variabilidade e uma média de atraso maior, pois enfatiza na construção de uma malha densa de *clusters* afim de garantir a conectividade da rede, logo um pacote para chegar a BS é retransmitido por muitos *clusters*. O ATCR consegue equilibrar a latência na rede obtendo uma variabilidade superior a média com menos pontos discrepantes em relação ao UCR e uma média igual.

A figura 21 apresenta o gráfico que demonstra o *overhead* de controle gerado por rodada para cada protocolo. O mecanismo integrado de clusterização e roteamento empregado pelo ATCR é superior ao HCR e UCR em relação aos pacotes de controle necessários para organização da rede em *clusters*. O ATCR consegue criar *clusters* na

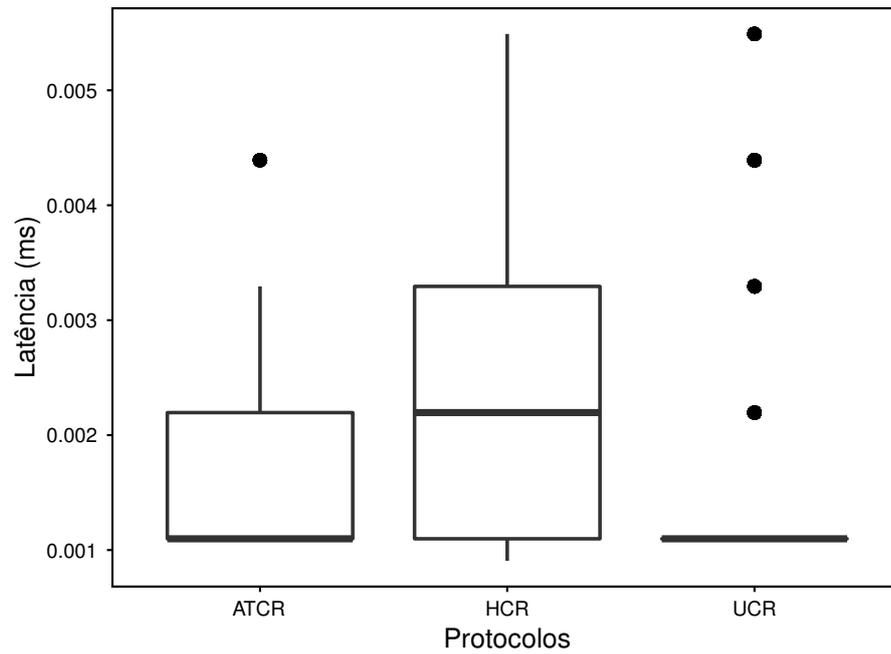


Figura 20: Latência

rede com o mínimo de pacotes de controle, dessa forma o mesmo consegue minimizar o consumo de energia na fase de formação dos clusters, contribuindo assim para prolongar o tempo de vida da rede.

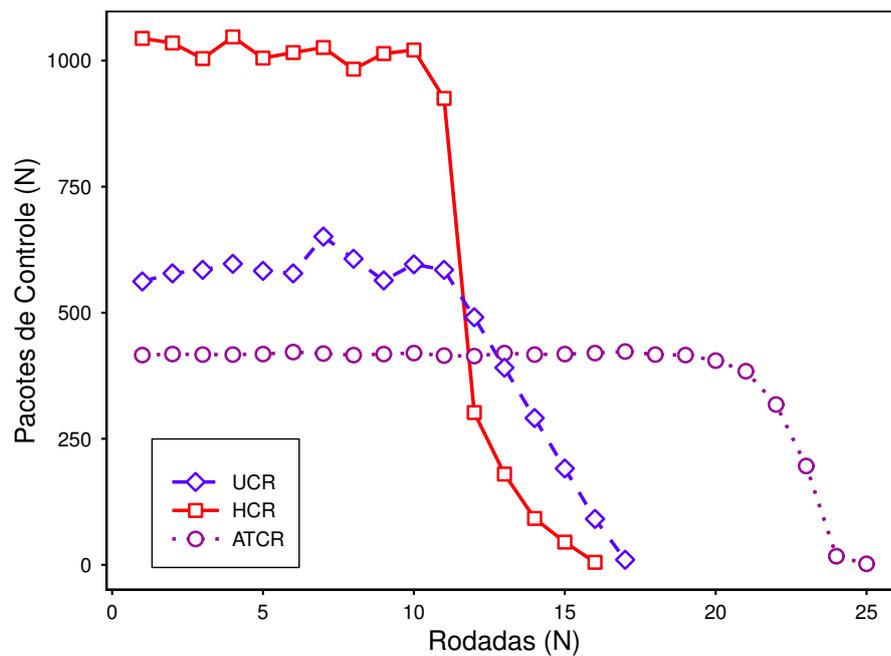


Figura 21: Overhead de controle

A figura 22 apresenta o gráfico que demonstra a taxa de pacotes recebidos pela

BS (packet delivery rate - PDR). A métrica PDR é apresentada pela equação 5.6.

$$PDR = \frac{Pkts_{received}}{Pkts_{sent}} \quad (5.6)$$

onde:

- $Pkts_{received}$ → Pacotes recebidos pela BS;
- $Pkts_{sent}$ → Pacotes enviados pelos CH;

A métrica PDR é apresentada sobre um cenário ideal que corresponde aos parâmetros descritos na tabela 3 e outro cenário cujo modelo do canal é assimétrico e o índice de *path loss* corresponde a 2.4, que corresponde a um cenário mais realístico. A intenção dessa avaliação é analisar o comportamento dos protocolos em relação ao PDR sobre situações adversas de comunicação.

O gráfico demonstra que o ATCR e UCR possuem taxas de entrega equivalentes no cenário ideal, ao contrário do HCR que possui uma baixa taxa, a explicação para isso é que o ATCR e UCR formam cluster de maneira mais distribuída, enquanto o HCR de forma mais densa, como consequência a perda de pacotes por interferência é quase que inevitável. A avaliação do PDR no cenário assimétrico demonstra que o ATCR sofreu menos impacto que os demais protocolos. Isso decorre do processo de formação de *cluster* organizar a rede com *clusters* bem dispersos em função do processo de eleição de CH não probabilístico e a otimização multicritério realizada através de lógica *fuzzy* que evita a desconexão da rede conforme a perda de nodos sensores ocorre.

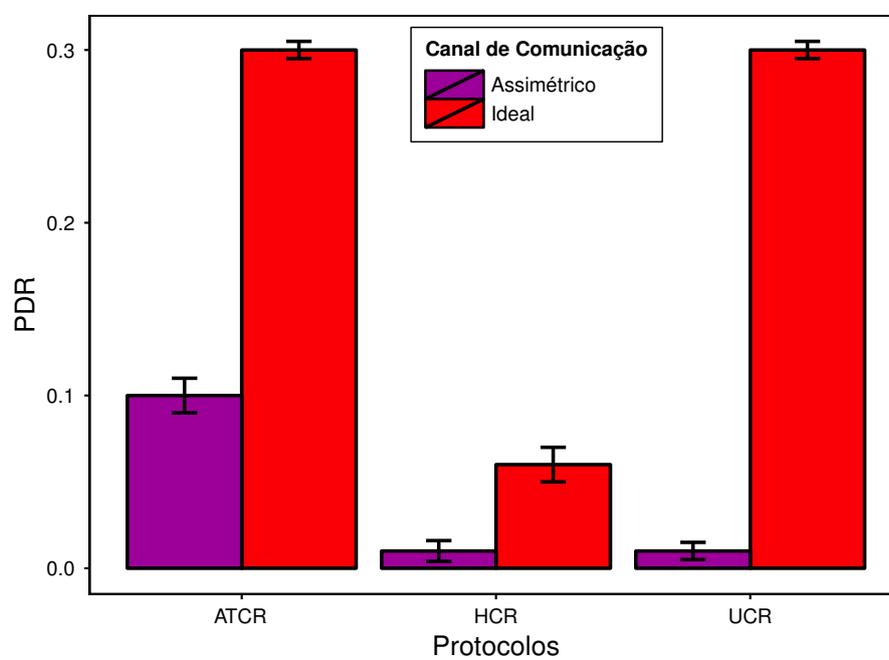


Figura 22: Taxa de pacotes entregues

6 Conclusões

Em uma RSSF organizada em *clusters* um dos principais desafios inerentes ao processo de formação de *clusters* é a eleição de CHs, pois fatores como eficiência energética, baixo *overhead* e conectividade estarão presentes nessa fase. O ATCR em seu projeto considerou esses fatores e forneceu uma solução híbrida de roteamento e formação de *clusters* que garantiu a conectividade da rede e a redução do *overhead*. Além disso, integra um mecanismo de temporização para eleição de CH que reduz ainda mais o *overhead* de controle dentro da solução híbrida e permite a distribuição adequada de *clusters* na rede por não se caracterizar como uma solução probabilística de eleição, comum em outros protocolos.

Outro ponto importante é a utilização de lógica *fuzzy* nos processos de tomadas de decisão, pois ela permitiu a utilização de múltiplos critérios importantes para garantir a eficiência energética e o balanceamento do consumo de energia na rede. E por fim o ATCR promove a estruturação desigual de *clusters* com o objetivo de mitigar o problema de *hot spot*.

O ATCR tratou os mais relevantes aspectos inerentes a eficiência energética em uma RSSF de múltiplos saltos, caracterizando-se assim como uma excelente solução para RSSFs de implantação densa e de larga escala. Dessa forma, o ATCR se destaca em relação ao estado da arte por figurar como uma abordagem eficiente e eficaz no contexto das RSSFs.

Considerando os resultados apresentados e as pesquisas realizadas pretende-se como trabalhos futuros:

- Promover melhorias no algoritmo afim de que o mesmo possa tornar mais eficiente o balanceamento do consumo na energia;
- Fornecer melhores garantias na qualidade de serviço sem prejudicar a eficiência energética da solução;
- Estruturar o algoritmo na forma de um protocolo com o objetivo de implantar em ambientes de teste real para avaliar o seu desempenho;
- Avaliar o desempenho da solução utilizando outros mecanismos de tomada de decisão multicritério.

- Modificar o ATCR para tornar-se uma solução *cross-layer*, afim de promover outras formas de divisão do canal de comunicação dentro do *cluster*, como por exemplo, a seleção dentro de uma determinada frequência de diferentes canais de comunicação.

Referências

- [1] A. A. Abbasi and M. Younis, “A survey on clustering algorithms for wireless sensor networks,” *Computer Communications*, vol. 30, pp. 2026–2841, jun 2007.
- [2] L. Mendes and J. Rodrigues, “A survey on cross-layer solutions for wireless sensor networks,” *Journal of Network and Computer Applications*, vol. 34, pp. 523–534, mar 2011.
- [3] Y. Jennifer, M. Biswanath, and G. Dipak, “Wireless sensor network survey,” *Computer Networks*, vol. 52, pp. 2292–2330, april 2008.
- [4] L. Atzori, A. Lera, and G. Morabito, “The internet of things: A survey,” *Computer Networks*, vol. 54, pp. 2787–2805, oct 2010.
- [5] S. Fang, S. M. Berber, and A. K. Swain, “Energy distribution-aware clustering algorithm for dense wireless sensor networks,” *International Journal of Communication Systems*, vol. 23, pp. 1223–1251, feb 2010.
- [6] D. Kim, D. Kim, H. Park, and S. Yoo, “Performance evaluation of routing protocols for wireless sensor networks in military scenarios,” in *Proc. IEEE Third International Conference on Ubiquitous and Future Networks (ICUFN’2011)*, jun 2011, pp. 101–106.
- [7] H. J. Lee, D. O. Kim, B. J. Kang, and S. W. Ban, “Mobile embedded health-care system working on wireless sensor network,” in *Proc. IEEE Third International Conference on Communications and Mobile Computing (CMC’2011)*, april 2011, pp. 161–164.
- [8] A. J. G. Sanchez *et al.*, “Wireless sensor network deployment for monitoring wildlife passages,” *sensors*, vol. 10, pp. 7236–7262, aug 2010.
- [9] B. Athanassios, B. Rodney, A. Saeed, and T. Yuri;, “A wireless sensor network test-bed for structural health monitoring of bridges,” in *Proc. IEEE 36th Conference on Local Computer Networks (LNC’2011)*, oct 2011, pp. 1040–1043.
- [10] J. Lloret, M. Garcia, D. Bri, and S. Sendra, “A wireless sensor network deployment for rural and forest fire detection and verification,” *sensors*, vol. 9, pp. 8722–8747, oct 2009.
- [11] N. A. Pantazis and D. D. Vergados, “A survey on power control issues in wireless sensor networks,” *IEEE Communications Surveys*, vol. 9, pp. 86–107, april 2007.
- [12] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “Heed: A hybrid, energy efficient distributed clustering approach for ad hoc sensor networks,” *IEEE Transaction on Mobile Computing*, vol. 3, pp. 366–379, oct 2004.

-
- [13] A. Chamam and S. Pierre, “A distributed energy-efficient clustering protocol for wireless sensor networks,” *Computers and Electrical Engineering*, vol. 36, pp. 303–312, may 2010.
- [14] P. Baronti *et al.*, “Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards,” *sensors*, vol. 26, pp. 1655–1695, may 2007.
- [15] S. Misra, M. Reisslein, and G. Xue, “A survey of multimedia streaming in wireless sensor networks,” *IEEE Communications Surveys and Tutorials*, vol. 10, pp. 1553–877X, jan 2008.
- [16] P. Baronti, P. Pillai, V. Chook, S. Chessa, A. Gotta, and Y. Hu, “Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards,” *Computer Communications*, vol. 30, pp. 1655–1695, Mar 2007.
- [17] J. Griessmann. (2012) Wirelesshart, an overview. [Online]. Available: <http://www.hartcomm.org>
- [18] I. S. of Automation. (2012) Isa-100.11a-2009,wireless systems for industrial automation: Process control and related applications. [Online]. Available: <http://www.isa.org>
- [19] D. Bhattacharyya, T. Kim, and S. Pal, “A comparative study of wireless sensor networks and their routing protocols,” *sensors*, vol. 10, pp. 10 506–10 523, nov 2010.
- [20] J. Zheng and A. Jamalipour, *Wireless sensor networks: a networking perspective*. 111 River Street, Hoboken: Wiley-IEEE Press, 2009.
- [21] N. Amini, A. Vahdatpour, W. Xu, M. Gerla, and M. Sarrafzadeh, “Cluster size optimization in sensor networks with decentralized cluster-based protocols,” *Computer Communications*, vol. 35, pp. 207–220, sept 2011.
- [22] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “In-network aggregation techniques for wireless sensor networks: A survey,” *IEEE Wireless Communications*, vol. 14, pp. 70–87, april 2007.
- [23] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.
- [24] —, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proc. IEEE 33rd Annual Hawaii International Conference on System Sciences ('2000)*, jan 2000, pp. 1–10.
- [25] —, “An applicationspecific protocol architecture for wireless microsensor networks,” *IEEE Transactions on wireless communications*, vol. 1, pp. 660–670, oct 2002.
- [26] G. Chen, C. Li, M. Ye, and J. Wu, “An unequal cluster-based routing protocol in wireless sensor networks,” *Wireless Networks*, vol. 15, pp. 193–207, feb 2009.

- [27] Z. Xu, C. Long, C. Chen, and X. Guan, "Hybrid clustering and routing strategy with low overhead for wireless sensor networks," in *Proc. IEEE International Conference on Communications, (ICC'2010)*, may 2010, pp. 1–5.
- [28] A. Chamam and S. Pierre, "A distributed energy-efficient clustering protocol for wireless sensor networks," *Computer and Electrical Engineering*, vol. 36, pp. 303–312, May 2010.
- [29] S. Lee, H. Choe, B. Park, Y. Song, and C. Kim, "Luca: An energy-efficient unequal clustering algorithm using location information for wireless sensor networks," *Wireless Personal Communications*, vol. 56, pp. 715–731, Oct 2011.
- [30] S. Momma, T. Mikoshi, and T. Takenaka, "Power aware routing and clustering scheme for wireless sensor networks," in *Proc. IEEE 8th Asia-Pacific Symposium on Information and Telecommunication Technologies, (APSITT'2010)*, Jun 2010, pp. 1–6.
- [31] J. Wang, Y. Cao, J. Xie, and S. Chen, "Energy efficient backoff hierarchical clustering algorithms for multi-hop wireless sensor networks," *Journal of Computer Science and Technology*, vol. 26, pp. 283–291, Mar 2011.
- [32] J.-S. Lee and W.-L. Cheng, "Fuzzy-logic-based clustering approach wireless sensor networks using energy predication," *IEEE Sensor Journal*, vol. 12, pp. 2891–2896, Sept 2012.
- [33] S. Fang, S. M. Berber, and A. K. Swain, "Energy distribution-aware clustering algorithm for dense wireless sensor networks," *International Journal of Communication Systems*, vol. 23, pp. 1223–1251, feb 2010.
- [34] O. Boyinbode, H. Le, A. Mbogho, M. Takizawa, and R. Poliah, "A survey on clustering algorithms for wireless sensor networks," in *Proc. IEEE 13th International Conference on Network-Based Information Systems, (NBIS'2010)*, Sept 2010, pp. 358–364.
- [35] S. Yi, J. Heo, Y. Cho, and J. Hong, "Peach: Power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks," *Computer communications*, vol. 30, pp. 2842–2852, oct 2007.
- [36] N. Aslam, W. Phillips, W. Robertson, and S. Sivakumar, "A multi-criterion optimization technique for energy efficient cluster formation in wireless sensor networks," *Information Fusion*, vol. 12, pp. 202–212, july 2011.
- [37] D. Wei, Yichao, and K. Moessner, "An energy-efficient clustering solution for wireless sensor networks," *IEEE Transaction on Wireless Communication*, vol. 10, pp. 3973–2983, nov 2011.
- [38] K. Srinivasan and P. Levis, "Rssi is under appreciated," in *Proc. IEEE Third Workshop on Embedded Networked Sensors, (EmNets'2006)*, may 2006, pp. 40–45.
- [39] Y. Cao and C. He, "A distributed clustering algorithm with an adaptative backoff strategy for wireless sensor networks," *IEICE Transactions on Communication*, vol. E89-B, pp. 609–613, feb 2006.

-
- [40] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, pp. 116–132, May 1985.
- [41] M. G. K. Wehrle and J. Gross, *Modeling and Tools for Network Simulation*. New York/Heidelberg: Springer-Verlag, 2010.
- [42] A. Law and W. D. Kelton, *Simulation Modeling and Analysis*. Columbus, OH 43272: McGraw-Hill Science, 1999.
- [43] R. Jain, *Art of Computer Systems Performance Analysis Techniques For Experimental Design Measurements Simulation And Modeling*. New York/Heidelberg: Wiley Computer Publishing, 1991.
- [44] NICTA. (2011) Castalia a simulator for wsns. [Online]. Available: <http://castalia.npc.nicta.com.au/>
- [45] M. H. M. J. Handy and D. Timmermann, “Low energy adaptive clustering hierarchy with deterministic cluster-head selection,” in *Proc. IEEE International Workshop Mobile Wireless Communication Network, ('2002)*, Jun 2002, pp. 368–372.

APÊNDICE A – Dados de Simulação

Tabela 4: Intervalos de confiança para 10 repetições

Rodadas	Energia Residual (E_r)	Intervalo de Confiança de E_r
0	20	0
1	19.174	0.00592635
2	18.3257	0.0113122
3	17.4759	0.0154092
4	16.6263	0.0178579
5	15.7751	0.0207647
6	14.9298	0.0215145
7	14.0863	0.0208846
8	13.2449	0.0194451
9	12.395	0.0266766
10	11.5438	0.032306
11	10.7038	0.0332417
12	9.86625	0.0368752
13	9.02642	0.0386221
14	8.18154	0.0395347
15	7.33671	0.0355224
16	6.4857	0.0348036
17	5.6429	0.0336283
18	4.81037	0.0298469
19	3.99247	0.0271041
20	3.22637	0.0256102
21	2.51778	0.0270324
22	1.69395	0.0864188
23	1.02142	0.113141
24	0.5531	0.0906312
25	0.109172	0.048294

Tabela 5: Intervalos de confiança para 50 repetições

Rodadas	Energia Residual (E_r)	Intervalo de Confiança de E_r
0	20	0
1	19.1768	0.00295148
2	18.3253	0.00557053
3	17.4768	0.00774509
4	16.6297	0.00811301
5	15.7826	0.0094857
6	14.9347	0.0103876
7	14.088	0.0101429
8	13.2404	0.010409
9	12.3901	0.0111702
10	11.54	0.0120372
11	10.695	0.013218
12	9.85267	0.0144449
13	9.00849	0.0150252
14	8.16398	0.0161111
15	7.32495	0.0163602
16	6.48012	0.0161179
17	5.64207	0.0166773
18	4.81444	0.0156602
19	3.99997	0.0152022
20	3.22949	0.017958
21	2.51938	0.0167901
22	1.79316	0.0370476
23	1.07109	0.0508337
24	0.589476	0.0450201
25	0.220885	0.0477
26	0.151302	0.0315078