

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

THALES RUANO BARROS DE SOUZA

EFICIÊNCIA ENERGÉTICA DE REDES DE SENSORES SEM FIO
APLICADA AO CONFORTO TÉRMICO EM AMBIENTES FECHADOS

MANAUS
2014

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

THALES RUANO BARROS DE SOUZA

EFICIÊNCIA ENERGÉTICA DE REDES DE SENSORES SEM FIO
APLICADA AO CONFORTO TÉRMICO EM AMBIENTES FECHADOS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção de título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Orientador: Prof. Dr. Rogério Caetano

MANAUS
2014

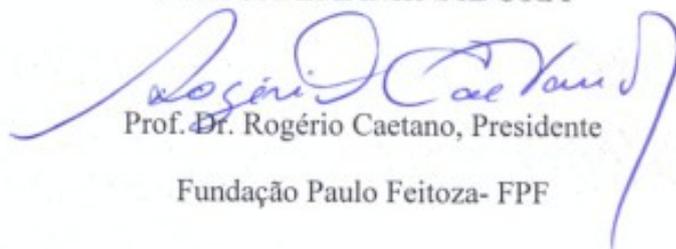
THALES RUANO BARROS DE SOUZA

EFICIÊNCIA ENERGÉTICA DE REDES DE SENSORES SEM FIO
APLICADA AO CONFORTO TÉRMICO EM AMBIENTES FECHADOS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovado em 21 de Fevereiro de 2014.

BANCA EXAMINADORA



Prof. Dr. Rogério Caetano, Presidente
Fundação Paulo Feitoza- FPF



Profa. Dra. Fabíola Guerra Nakamura, Membro

Universidade Federal do Amazonas- UFAM



Prof. Dr. Raimundo da Silva Barreto, Membro

Universidade Federal do Amazonas- UFAM

© 2014, Thales Ruano Barros de Souza

Todos os direitos reservados.

Ficha Catalográfica

(Catalogação realizada pela Biblioteca Central da UFAM)

Souza, Thales Ruano Barros de

Eficiência energética de redes de sensores sem fio aplicada ao conforto térmico em ambientes fechados / Thales Ruano Barros de Souza. - 2014.

65 f. : il. color. ; 31 cm.

Dissertação (mestre em Engenharia Elétrica) — Universidade Federal do Amazonas.

Orientador: Prof. Dr. Rogério Caetano.

1. Sistemas de comunicação sem fio 2. Sistemas de parâmetros distribuídos 3. Conforto humano 4. Engenharia elétrica I. Caetano, Rogério, orientador II. Universidade Federal do Amazonas III. Título

CDU (2007): 621.391(043.3)

RESUMO

As Redes de Sensores Sem Fio (RSSFs) podem ser utilizadas em aplicações de monitoramento e estudo do conforto térmico em ambientes fechados, devido a sua capacidade de sensoriamento, baixo custo e implantação rápida. Porém o custo de manutenção pode ultrapassar o custo do próprio sistema de monitoramento, isso porque os nós sensores são alimentados por bateria. Nesta dissertação é proposto e avaliado um protocolo de roteamento que leva em consideração as variáveis monitoradas na formação dos grupos de nós sensores, para redução da comunicação intragrupo. A solução proposta utiliza, não apenas a localização dos nós sensores, mas também a temperatura e a umidade relativa coletadas por cada nó sensor, para a eleição dos *cluster-heads* (CHs). Para agrupar os nós sensores com medidas correlacionadas, utilizamos o algoritmo k-means, que agrupa instâncias similares. Os experimentos mostram que, em comparação com outros algoritmos da literatura, o consumo de energia chegou a ser reduzido em 50%. Portanto há um ganho substancial quando se explora a estrutura de correlação das variáveis coletadas na formação dos grupos e na eleição dos CHs.

Palavras-chave: Redes de Sensores Sem Fio, Conforto Térmico, Aprendizado Não Supervisionado, k-means.

ABSTRACT

Wireless Sensor Networks can be used in several applications such as monitoring and study of thermal comfort in indoor environment. However, since the sensors nodes are battery powered, the cost of network maintenance can quickly exceed the cost of whole monitoring system. In this work, it is proposed and evaluated a routing protocol considering the monitored variables in sensors nodes to form groups of sensors in order to reduce the intra cluster communication. Our solution uses not only localization but also temperature and relative humidity collected by each sensor node to elect *cluster heads* (CHs). The k-means algorithm was used to group the correlated sensors nodes. Results show that the proposed scheme has a performance 50% higher than the state-of-the-art algorithms in relation to residual energy and network lifetime. Therefore there is a substantial gain when we exploit the correlation structure of the collected data to form groups and to elect CHs.

Keywords: Wireless Sensor Networks, Thermal Comfort, Unsupervised Learning, k-means

LISTA DE FIGURAS

Figura 1 Laboratório Intel-Berkeley (Intel Berkeley, 2004)	3
Figura 2 Volume de pacotes gerados pela RSSF	9
Figura 3 Modelo de roteamento de salto-único (esquerda) e de saltos-múltiplos (direita).....	10
Figura 4 Agrupamento com salto único (esquerda) e com saltos múltiplos (direita).....	18
Figura 5 Estrutura de uma rodada do protocolo LEACH (Park et al, 2013)	23
Figura 6 Topologia básica do LEACH	24
Figura 7 Algoritmo k-Means	28
Figura 8 Função silhueta para 3 grupos.....	29
Figura 9 Ordenação dos nós sensores com ID (Park et al, 2013).....	41
Figura 10 Quantidade de pacotes por nó sensor	45
Figura 11 Distribuição da temperatura no 1º e 2º dia de Março.....	46
Figura 12 Distribuição da temperatura no 3º e 4º dia de Março.....	46
Figura 13 Centróides dos 16 grupos	48
Figura 14 Valores da função silhueta para 5 grupos	50
Figura 15 Definição da quantidade de grupos.....	51
Figura 16 Protocolo de Roteamento Proposto	52
Figura 17 Roteamento single hop intragrupo e intergrupo.....	53
Figura 18 Algoritmo de agrupamento	54
Figura 19 Energia residual restante da rede por rodada	56
Figura 20 Número de nós vivos por rodada.....	57

LISTA DE QUADROS

Tabela 1 PMV (Predicted Mean Vote)	36
Tabela 2 Parâmetros de configuração da rede	49

SUMÁRIO

ABSTRACT.....	V
LISTA DE FIGURAS.....	VI
LISTA DE QUADROS.....	VII
INTRODUÇÃO.....	1
1.1 VISÃO GERAL.....	1
1.2 MOTIVAÇÃO.....	2
1.3 DEFINIÇÃO DO PROBLEMA.....	3
1.4 OBJETIVOS.....	4
1.4.1 GERAL.....	4
1.4.2 ESPECÍFICOS.....	4
1.5 CONTRIBUIÇÕES.....	5
1.6 ESTRUTURA DA DISSERTAÇÃO.....	5
FUNDAMENTOS TEÓRICOS.....	6
2.1 REDES DE SENSORES SEM FIO.....	6
2.1.1 CONSUMO DE ENERGIA EM RSSFS.....	7
2.1.2 ROTEAMENTO EM RSSFS.....	9
2.1.3 LEACH.....	22
2.2 APRENDIZADO NÃO SUPERVISIONADO.....	25
2.2.1 K-MEANS.....	26
2.2.2 FUNÇÃO SILHUETA.....	28
2.3 CONFORTO TÉRMICO.....	31
2.3.1 FATORES DETERMINANTES PARA O CONFORTO TÉRMICO.....	32
2.3.1.1 TAXA METABÓLICA.....	32
2.3.1.2 ISOLAMENTO TÉRMICO.....	33
2.3.1.3 TEMPERATURA DO AR.....	33
2.3.1.4 TEMPERATURA RADIANTE.....	33
2.3.1.5 VELOCIDADE DO AR.....	34
2.3.1.6 UMIDADE RELATIVA DO AR.....	35
2.3.2 PREDITORES DO CONFORTO TÉRMICO.....	36
TRABALHOS CORRELATOS.....	38
3.1 PROTOCOLOS DE ROTEAMENTO DETERMINÍSTICOS.....	38
METODOLOGIA.....	45
4.1 MINERAÇÃO DE DADOS.....	45
4.2 MODELAGEM DA REDE.....	48
4.3 PROTOCOLO DE ROTEAMENTO.....	51
4.3.1 AGRUPAMENTO INICIAL.....	52
4.3.2 REAGRUPAMENTO.....	52
RESULTADOS.....	55
5.1 CONSUMO DE ENERGIA DA REDE.....	55
5.2 NÚMERO DE NÓS VIVOS.....	56
CONCLUSÃO.....	58
REFERÊNCIAS BIBLIOGRÁFICAS.....	60
APÊNDICE.....	67
PUBLICAÇÕES.....	67

INTRODUÇÃO

1.1 VISÃO GERAL

As Redes de Sensores Sem Fio (RSSFs) têm recebido bastante atenção nos últimos anos. Essas redes consistem em um número de nós sensores com capacidade de sensoriamento, processamento de dados e comunicação sem fio (Akyildiz et al., 2002). Devido às suas características de auto-organização, baixo custo e implantação rápida, RSSFs são aplicadas em muitos sistemas de monitoramento em ambientes fechados para coletar e processar dados do ambiente (Li et. al., 2009).

Neste trabalho, tratamos o problema do tempo de vida de uma RSSF implantada em ambientes fechados para monitorar variáveis do ambiente para o estudo do conforto térmico dos usuários. Estamos interessados em estender o tempo de vida da rede como um todo, para evitar manutenções constantes que podem inclusive ultrapassar o custo do próprio sistema de monitoramento em um período curto de tempo.

Este trabalho usa a estrutura de correlação espacial das variáveis monitoradas no protocolo de roteamento para redução da comunicação dentro de cada grupo que compõe uma RSSF, conseqüentemente reduzindo o consumo de energia da rede. Dentre os protocolos de roteamento em RSSF destacam-se os protocolos planos e os hierárquicos. No protocolo plano, todos os nós executam a mesma tarefa e possuem as mesmas funcionalidades na rede. A transmissão de dados é realizada salto a salto geralmente usando a forma de inundação. Nos

protocolos hierárquicos, os nós sensores possuem funções diferentes na RSSFs e tipicamente são organizados em muitos grupos de acordo com requisitos ou métricas específicas.

Propomos utilizar as variáveis monitoradas no protocolo de roteamento. Nossa solução divide a RSSF em grupos de nós, por meio de roteamento hierárquico, com valores de temperatura e umidade relativa correlacionadas. Em seguida ocorre a eleição do nó sensor que melhor representa cada grupo. Para dividir a rede em grupos de nós sensores, utilizamos o algoritmo k-means que agrupa instâncias de dados similares.

1.2 MOTIVAÇÃO

Uma das aplicações onde RSSFs têm sido empregadas está diretamente relacionada ao conforto térmico em ambientes fechados, onde os nós sensores realizam o sensoriamento contínuo de variáveis como temperatura e umidade relativa do ar, para estudar a relação entre essas variáveis e o conforto térmico dos usuários em ambientes fechados controlados. A maioria das estratégias operam de modo centralizado, onde os dados são enviados para um dispositivo central no qual ocorre todo o seu processamento (Li et. al., 2009). Além disso, costuma-se trabalhar com valores médios dos dados coletados (Yun et. al., 2012), perdendo-se a estrutura de correlação destes dados, os quais podem revelar situações de desconforto local devido a condições microclimáticas.

Um dos grandes desafios na aplicação de RSSFs está relacionado ao consumo de energia, pois o custo de manutenção para substituir ou recarregar a bateria dos nós sensores de uma RSSF, quando possível, pode ultrapassar o custo de todo o sistema em um período de tempo relativamente pequeno. Para contornar esse problema, tem sido pesquisadas estratégias para adquirir energia de fontes do ambiente (por exemplo, luz, calor ou vibração) com o

objetivo de prolongar o tempo de vida da rede (Wang et. al., 2010). Além disso, também empregam-se algoritmos de roteamento para redução do consumo de energia (Akkaya et al., 2005), maximizando assim o tempo de vida da rede.

1.3 DEFINIÇÃO DO PROBLEMA

Em cenários reais, diversos nós sensores são dispostos em um ambiente fechado com o objetivo de monitorar o ambiente o maior tempo possível. A RSSF deve coletar dados do ambiente, temperatura e umidade relativa por exemplo, e posteriormente encaminhar a informação coletada para uma estação base, onde os dados serão processados.

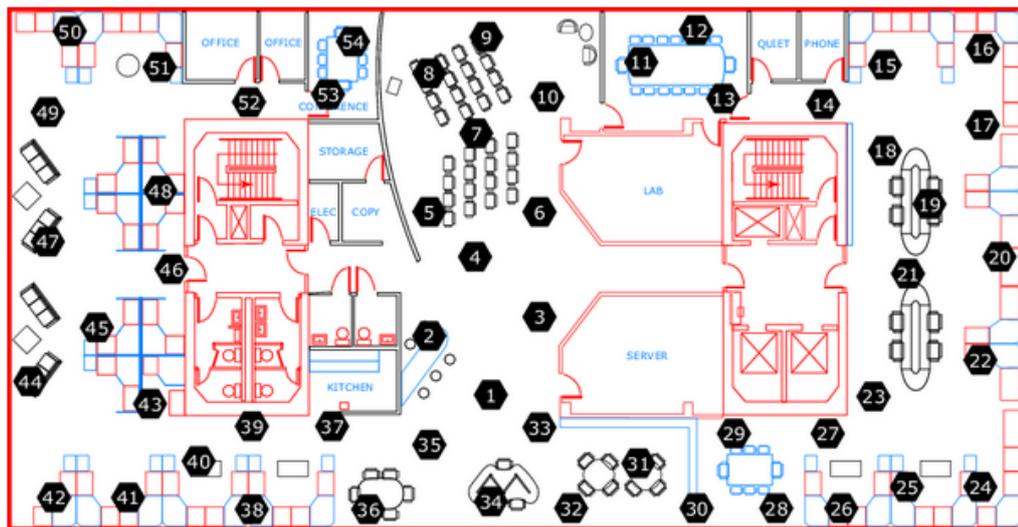


Figura 1 Laboratório Intel-Berkeley (Intel Berkeley, 2004)

Uma ilustração deste cenário é apresentada na Figura 1, onde 54 nós sensores foram dispostos em um laboratório. Desta forma, como os nós sensores próximos coletam dados similares, as soluções de roteamento por agrupamento visam reduzir essas redundâncias nos dados coletados, elegendo um nó sensor líder para cada grupo. Assim cada nó sensor

comunica-se apenas com seu respectivo líder, que por sua vez, tem o papel de agregar e encaminhar os dados do grupo para fora da rede.

Neste trabalho o cenário problema a ser atacado é a redução da comunicação intragrupo, que é a comunicação de cada nó sensor com seu respectivo líder, por meio da inserção das variáveis monitoradas no algoritmo de roteamento.

1.4 OBJETIVOS

1.4.1 GERAL

Desenvolver e avaliar algoritmos de roteamento para extensão do tempo de vida de redes de sensores sem fio no cenário de conforto térmico em ambientes fechados.

1.4.2 ESPECÍFICOS

Os objetivos específicos consistem em:

- Desenvolver e avaliar um algoritmo de aprendizado não supervisionado que encontre o conjunto de dados que melhor represente cada região microclimática;
- Desenvolver e avaliar um protocolo de roteamento usando um algoritmo de aprendizado não supervisionado como método de agrupamento e eleição de líderes que melhor representem cada região microclimática, baseado nos valores de temperatura e umidade relativa coletados pela RSSF.

1.5 CONTRIBUIÇÕES

As contribuições do trabalho são:

1. Um novo algoritmo de roteamento para RSSF aplicada ao cenário de conforto térmico em ambientes fechados;
2. Análise comparativa de desempenho entre o algoritmo proposto e outros algoritmos da literatura.

1.6 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está dividida em seis capítulos. Capítulo 1 é a introdução onde é apresentada a visão geral, a motivação e a definição do problema, os objetivos e as contribuições do trabalho. O Capítulo 2 apresenta os fundamentos relacionados à RSSF e os fundamentos relacionados ao conforto térmico. O Capítulo 3 apresenta os trabalhos correlatos. A solução proposta para o problema de redução da comunicação intragrupo é apresentada no Capítulo 4. O Capítulo 5 apresenta os resultados e discussões relacionadas. Por fim, temos as considerações finais no Capítulo 6.

FUNDAMENTOS TEÓRICOS

2.1 REDES DE SENSORES SEM FIO

Os avanços tecnológicos dos últimos anos possibilitaram o surgimento de nós sensores multifuncionais de baixo custo e consumo de energia que se comunicam em curtas distâncias. Esses pequenos nós sensores, que têm capacidade de sensoriamento, processamento de dados e comunicação, levaram à criação de redes de sensores baseadas em um esforço colaborativo de uma quantidade de nós.

Uma das principais vantagens de RSSFs é que a posição dos nós sensores não precisa ser pré-definida. Isso possibilita deposição aleatória em terrenos inacessíveis. Por outro lado, isso também significa que o protocolo e os algoritmos da RSSF devem ter a capacidade de auto-organização. Outra característica importante de redes de sensores é o esforço cooperativo dos nós sensores. Os nós sensores são equipados com um processador que ao invés de enviar dados crus para os nós responsáveis pela fusão, realizam processamentos simples e transmitem apenas os dados necessários.

RSSFs têm sido utilizadas para estudar o conforto térmico em ambientes fechados. Porém a maioria dos trabalhos existentes na literatura utilizam redes de sensores apenas para o sensoriamento contínuo do ambiente, realizando todo o processamento em uma estação base, fora da rede (Wu et. al., 2007; Li et. al., 2009).

A natureza distribuída de uma RSSF pode revelar as condições microclimáticas que levam ao desconforto térmico em ambientes fechados. Por outro lado, o sensoriamento

contínuo pode ser proibitivo em aplicações práticas, devido ao consumo de energia na transmissão periódica dos dados sensorizados. Além disso, para cenários maiores onde é necessário o uso de uma rede de sensores densa, surgem problemas como ruído e colisão de pacotes, que levam a perdas significativas tanto em quantidade quanto em qualidade dos dados, passando a ser necessário o uso de um algoritmo de roteamento.

2.1.1 CONSUMO DE ENERGIA EM RSSFs

Apesar de RSSFs apresentarem muitas características vantajosas, uma desvantagem é que os nós sensores são alimentados por bateria, um recurso energético limitado, tornando necessária a sua manutenção com frequência (Wang et al., 2011). O custo de manutenção para substituir ou recarregar os nós sensores de uma RSSF, em aplicações típicas de ambiente fechado, pode ultrapassar o custo do sistema em um período de tempo relativamente pequeno. Muito esforço de pesquisa tem sido realizado para adquirir energia de fontes do ambiente (por exemplo, luz, calor ou vibração) para prolongar o tempo de vida da rede (Wang et al., 2010; Lhermet et al., 2008; Mitcheson et al., 2008). Além disso, muitos trabalhos investigaram protocolos de rede para economizar energia de RSSFs (Chamam et al, 2009; Marinkovic, 2009).

A principal tarefa de um nó sensor é detectar eventos, realizar processamento local de forma rápida e transmitir os dados. O consumo de energia pode, portanto, ser dividido em três domínios: sensoriamento, comunicação e processamento de dados (Akyildiz et al., 2002).

A energia consumida no sensoriamento depende da natureza da aplicação. Sensoriamento esporádico deve consumir menos energia que o monitoramento constante de eventos. A complexidade da detecção também desempenha um papel crucial na determinação

do consumo de energia. Além disso, ambientes altamente ruidosos causam erros significativos e aumentam a complexidade da detecção.

Entre os três domínios, aquele em que o nó sensor gasta mais energia é na comunicação de dados. Isso envolve transmissão e recepção de dados. O consumo de energia no processamento de dados é bem menor quando comparado à comunicação. O exemplo descrito em Pottie et al. (2000) ilustra essa disparidade. Os resultados mostram que o custo de energia para transmitir 1KB a uma distância de 100 metros é aproximadamente o mesmo para a execução de 3 milhões de instruções a 100 milhões de instruções por segundo. Portanto, processamento local é crucial para minimizar o consumo de energia em uma rede de sensores sem fio.

O gráfico da Figura 2 mostra o volume de pacotes gerados pela RSSF no experimento realizado no laboratório Intel Berkeley (Wu et. al., 2007). Pode-se observar que a rede inteira fica sem energia em menos de um mês de uso. Além disso, o trabalho de Wu et al. (2007) mostra que os dados coletados durante os últimos dias são *outliers*, ou seja, dados que apresentam valores muito diferentes do padrão dos dados coletados. Isso pode ter ocorrido devido ao baixo nível de energia dos nós sensores.

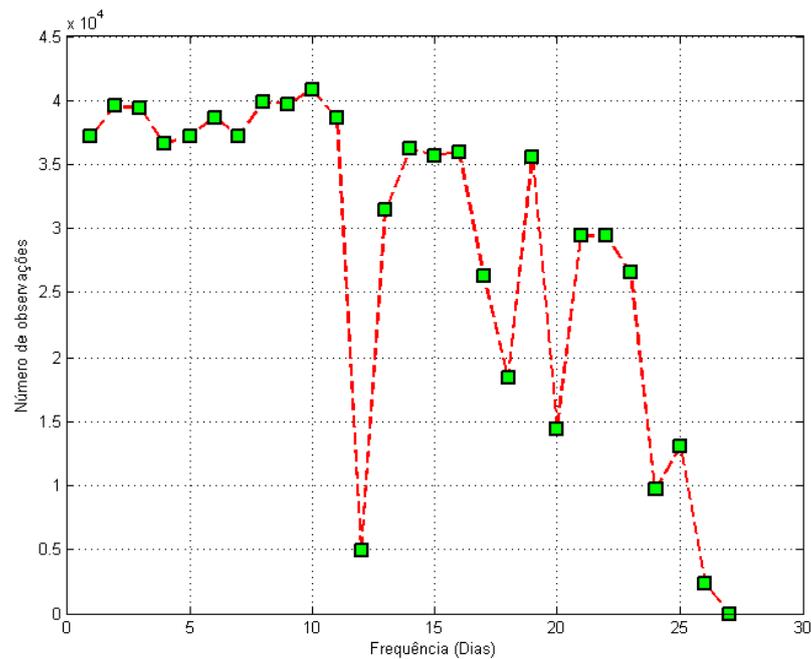


Figura 2 Volume de pacotes gerados pela RSSF

2.1.2 ROTEAMENTO EM RSSFs

Roteamento é uma tarefa crítica em RSSFs. Ao contrário das tradicionais redes *ad hoc*, roteamento em RSSFs é mais desafiador por causa das suas características inerentes (Seah et al., 2010; Li et al., 2011). Em primeiro lugar porque os recursos são muito limitados em termos de consumo de energia, processamento e largura de banda de transmissão. Em segundo lugar, é difícil projetar um esquema de endereçamento global como Protocolo de Internet (IP). Além disso, IP não pode ser aplicado a RSSFs, pois atualizar os endereços em larga escala em uma RSSF dinâmica pode resultar em sobrecarga. Em terceiro lugar, devido aos recursos limitados, é difícil para o algoritmo de roteamento lidar com mudanças frequentes e imprevisíveis da topologia da rede, especialmente em um cenário móvel. Em quarto lugar, a coleta de dados por muitos nós sensores geralmente resulta em uma alta

probabilidade de redundância de dados, que deve ser considerada pelo protocolo de roteamento. Em quinto lugar, a maioria das aplicações de RSSFs requerem um esquema simples de comunicação do tipo muitos-para-um, por exemplo, de múltiplas fontes para um particular nó sorvedouro (sink), ao invés de multipercurso e par-a-par, como mostra a Figura 3. Por fim, em aplicações de RSSFs com restrições temporais, a transmissão de dados deve ser realizada em um certo período de tempo. Portanto, uma latência limitada deve ser considerada para transmissão de dados para esse tipo de aplicação. Porém, conservação de energia geralmente é mais importante que a qualidade de serviço (QoS), na maioria das aplicações onde todos os nós sensores possuem restrições quanto a energia que está diretamente relacionada ao tempo de vida da rede.

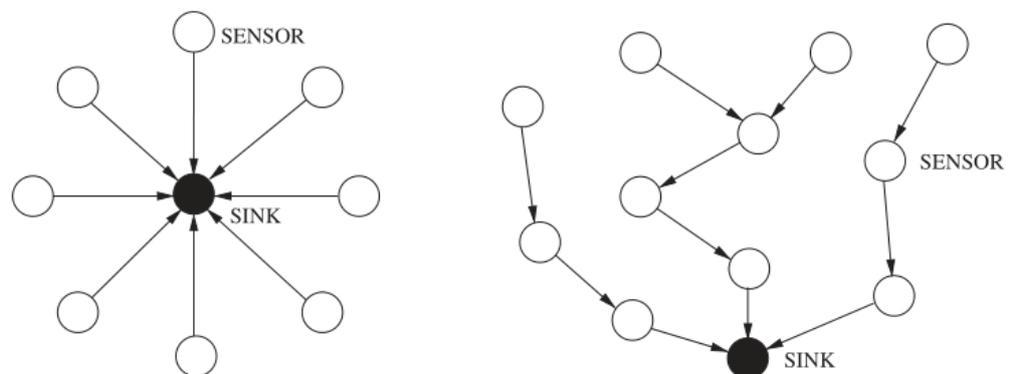


Figura 3 Modelo de roteamento de salto-único (esquerda) e de saltos-múltiplos (direita)

Baseado na estrutura da rede, protocolos de roteamento de RSSFs podem ser divididos em duas categorias: roteamento plano e roteamento hierárquico. Em uma topologia plana, todos os nós executam a mesma tarefa e possuem as mesmas funcionalidades na rede. A transmissão de dados é realizada salto a salto geralmente usando a forma de inundação. Os roteamentos típicos em RSSFs incluem Flooding and Gossiping (Haas et al., 2005), Sensor Protocols of Information via Negotiation (SPIN) (Kulik et al., 2002), Directed Diffusion (DD)

(Intanagonwiwat et al., 2006), Rumor (Braginsky et al., 2002), Greedy Perimeter Stateless Routing (GPSR) (Karp et al., 2000), Trajectory Based Forwarding (DD) (Niculescu et al., 2003), Energy-Aware Routing (EAR) (Shah et al., 2002), Gradient-Based Routing (GBR) (Schurges et al., 2001), Sequential Assignment Routing (SAR) (Sohrabi et al., 2000), entre outros.

Em redes de pequena escala protocolos de roteamento plano são relativamente efetivos. Porém, em redes de larga escala todos os sensores geram mais processamento de dados e uso de banda comprometendo a questão de escassez de recursos. Por outro lado, na topologia hierárquica, os nós sensores possuem funções diferentes na RSSFs e tipicamente são organizados em muitos grupos de acordo com requisitos ou métricas específicas. Geralmente, cada grupo inclui um líder chamado de “Cluster Head” (CH) e outros nós membros (MNs), os CHs podem ser organizados em outro nível hierárquico. Em geral, nós com alto nível de energia atuam como CH e realizam o processamento e a transmissão da informação, enquanto nós com pouca energia atuam como MNs fazendo o sensoriamento da informação.

Os protocolos de roteamento por agrupamento em RSSFs incluem Low-energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman et al., 2000), Hybrid Energy-Efficient Distributed Clustering (HEED) (Younis et al., 2004), Distributed Weight-based Energy Efficient Hierarchical Clustering protocol (DWEHC) (Ding et al., 2005), Position-based Aggregator Node Election protocol (PANEL) (Buttayan et al., 2007; Buttayan et al., 2010), Two-Level Hierarchy LEACH (TL-LEACH) (Loscri et al., 2005), Unequal Clustering Size (UCS) (Soro et al., 2005), Energy Efficient Clustering Scheme (EECS) (Ye et al., 2005; Ye et al., 2006), Energy-Efficient Uneven Clustering (EEUC) (Li et al., 2005), Algorithm for Cluster Establishment (ACE) (Chan et al., 2004), Base-Station Controlled Dynamic

Clustering Protocol (Murugunathan et al., 2005), Power-Efficient Gathering in Sensor Information Systems (PEGASIS) (Lindsey et al, 2002), Threshold sensitive Energy Efficient sensor Network protocol (TEEN) (Manjeshwar et al., 2001), The Adaptive Threshold sensitive Energy Efficient sensor Network protocol (APTEEN) (Manjeshwar et al., 2002), Two-Tier Data Dissemination (TTDD) (Luo et al., 2005), Concentric Clustering Scheme (CCS) (Jung et al., 2007) e Hierarchical Geographic Multicast Routing (HGMR) (Koutsonikola et al., 2010).

- **VANTAGENS E OBJETIVOS DO AGRUPAMENTO**

Comparado ao roteamento plano em RSSFs, protocolos de roteamento por agrupamento possuem várias vantagens, como escalabilidade, menor carga, menor consumo de energia e maior robustez. Nesta sessão, serão apresentadas as suas vantagens, bem como os objetivos do agrupamento em RSSFs.

Escalabilidade: Nos esquemas de roteamento por agrupamento, os nós sensores são divididos em uma variedade de grupos com níveis diferentes de atribuição. Os CHs são responsáveis pela agregação dos dados, disseminação da informação e gerenciamento da rede, e os MNs pelo sensoriamento dos eventos e coleta da informação do ambiente. Comparado à topologia plana, esse tipo de topologia de rede é mais fácil de gerenciar, e mais escalável para responder a eventos no ambiente (Seah et. al., 2010).

Fusão de Dados: É o processo de agregação de dados de múltiplos nós para eliminar transmissões redundantes e fornecer dados fundidos para a “Base Station” (BS), sendo eficaz para minimizar o consumo de energia em RSSFs (Rajagopalan et al, 2006). O método mais popular de fusão de dados é a agregação por agrupamento, em que cada CH agrega os dados

coletados pelos MNs e transmite o dado fundido para a BS (Yue et. al., 2012). Geralmente CHs formam uma estrutura de árvore para transmitir os dados agregados por saltos múltiplos por meio de outros CHs que resulta em economia significativa de energia (Ozdemir et al., 2009).

Menor Carga: Como os nós sensores geram dados significativamente redundantes a agregação de dados emergiu como sendo um princípio e objetivo importante em RSSFs. O principal objetivo da fusão de dados é combinar dados de fontes diferentes para eliminar transmissões redundantes, e fornecer uma visualização rica e multidimensional do que está sendo monitorado (Seah et. al.,2010; Li et. al., 2011). Muitos esquemas de roteamento com capacidade de agregação de dados requerem uma seleção cuidadosa para a abordagem de agrupamento. Para topologia de agrupamento, todos os membros do grupo enviam dados para o CH, e a agregação dos dados é realizada nos CHs, o que ajuda a reduzir drasticamente a transmissão de dados, e conseqüentemente, economizando energia. Além disso, as rotas são configuradas dentro de cada grupo, o que reduz o tamanho da tabela de roteamento armazenada nos nós sensores individuais (Abbasi et. al., 2007; Akkaya et. al., 2005).

Menor Consumo de Energia: No esquema de roteamento por agrupamento, a fusão de dados ajuda a reduzir as transmissões de dados, e conseqüentemente, economizando energia. Além disso, agrupamento com comunicação intragrupo e intergrupo pode reduzir o número de nós sensores realizando a tarefa de comunicação a longas distâncias, portanto permitindo menor consumo de energia para rede inteira. Dessa forma, apenas CHs tem a função de transmitir os dados no esquema de roteamento por agrupamento, o que pode economizar grande quantidade de energia.

Maior Robustez: A estratégia de agrupamento é mais conveniente para o controle da topologia e para responder a mudanças relacionadas ao aumento da quantidade de sensores,

mobilidade dos nós e falhas imprevisíveis. Um esquema de roteamento por agrupamento precisar lidar com essas mudanças dentro de cada grupo individualmente, portanto a rede inteira é mais robusta e mais gerenciável. Para compartilhar as responsabilidades, os CHs são geralmente alternados entre todos os sensores para evitar um único ponto de falha em algoritmos de roteamento por agrupamento.

Anticolisão: No modelo plano de múltiplos saltos, o meio sem fio é compartilhado e gerenciado por nós individuais, causando altas taxas de colisão de pacotes resultando em baixa eficiência nos recursos utilizados. Por outro lado, no modelo de agrupamento de múltiplos saltos, uma RSSF é dividida em grupos e a comunicação compreende dois modos, intragrupo e intergrupo, respectivamente para coleta de dados e para transmissão de dados. Os recursos podem ser alocados ortogonalmente para cada grupo para reduzir a colisão entre grupos (Lee et. al., 2011). Como resultado, o modelo de agrupamento de saltos múltiplos é apropriado para RSSFs de larga escala.

Redução de Latência: Quando uma RSSF é dividida em grupos, apenas os CHs executam a tarefa de transmissão de dados para fora do grupo. O modo de transmissão de dados para fora do grupo ajuda a evitar colisão de pacotes entre os nós. Consequentemente a latência é reduzida. Além disso, a transmissão de dados é realizada salto a salto e geralmente usando o formato de inundação no esquema plano, porém apenas os CHs executam a tarefa de transmissão de dados no esquema de roteamento por agrupamento, o que diminui quantidade de saltos da fonte de dados até a BS, reduzindo a latência.

Balanceamento de Carga: É essencial para prolongar o tempo de vida da rede em RSSF. A distribuição uniforme entre os nós sensores em cada grupo é geralmente considerada para construção dos grupos onde o CH executa a tarefa de processamento e gerenciamento intragrupo. Em geral, a construção dos grupos de tamanho igual é adotada para prolongar o

tempo de vida da rede, pois previne o esgotamento prematuro de energia dos CHs. Além disso, roteamento por múltiplos percursos é um método para balanceamento de carga.

Tolerância a Falhas: Devido à aplicabilidade de RSSFs em um bom número de cenários dinâmicos, nós sensores podem sofrer esgotamento de energia, erro de transmissão, funcionamento defeituoso de hardware, ataques maliciosos e assim por diante. Tolerância a falhas é um importante desafio em RSSFs (Chitnis et al., 2009). Para evitar a perda significativa de dados de nós sensores chaves, os CHs devem ser tolerantes a falhas, tornando necessário o projeto de abordagens efetivas para tolerância a falhas. Reagrupamento é o método mais intuitivo de recuperação de um grupo defeituoso, embora geralmente desarranje a operação que esteja ocorrendo.

Garantia de Conectividade: Nós sensores geralmente transmitem dados para uma ou mais BSs via salto único ou múltiplos saltos em RSSFs, portanto a conectividade de cada nó sensor ao longo do caminho para a BS é o fator determinante para a entrega dos dados. Além disso, nós sensores que não podem comunicar-se com outro nó sensor serão isolados e seus dados podem nunca ser transmitidos para a BS. Portanto, garantir a conectividade é um objetivo essencial para protocolos de roteamento por agrupamento em RSSFs (Seah et al, 2010; Li et al, 2011). Um exemplo importante é quando alguma informação a respeito de todos os nós sensores precisa ser coletada por determinado nó de fusão em protocolos de roteamento por agrupamento (Freris et al, 2010).

Evitar Buraco de Energia: Geralmente, o roteamento de múltiplos saltos é usado para entregar o dado coletado para o nó sorvedouro ou para a BS. Nessas redes, o tráfego de dados de cada nó inclui o tráfego autogerado e o tráfego de retransmissões. Independentemente do protocolo MAC, os nós sensores próximos da BS têm que transmitir mais pacotes que aqueles distantes da BS (Li et al., 2007). Como resultado, os nós próximos da BS descarregam

primeiro, levando a um buraco próximo a BS, separando a rede inteira, impedindo que os nós sensores enviem informações para BS, enquanto os muitos nós restantes ainda possuem bastante energia. Esse fenômeno é conhecido como buraco de energia (Tran-Quang et al., 2010). Mecanismos para evitar o buraco de energia podem ser classificados em três grupos: deposição dos nós, balanceamento de carga e mapeamento de energia (Ishmanov et al., 2011). Em especial, agrupamento desigual é um dos métodos de balanceamento de carga. Nesse método, um grupo de raio menor, próximo ao sorvedouro e um grupo de raio maior longe do sorvedouro são definidos respectivamente, então o consumo de energia no processamento de dados intergrupo é menor para grupos com raio menor e, portanto, mais energia pode ser usada para retransmitir os dados dos nós remotos (Liu et al., 2011).

Maximizando o Tempo de Vida da Rede: O tempo de vida da rede é uma consideração inevitável em RSSFs, porque os nós sensores são limitados em energia, capacidade de processamento e banda de transmissão, especialmente para aplicações em ambientes inóspitos. É indispensável minimizar a energia consumida na comunicação intragrupo por meio dos CHs que são ricos em recursos energéticos. Além disso, os nós sensores que estão mais próximos da maioria dos nós sensores devem ser propensos a serem CHs. Adicionalmente, com o objetivo de economizar energia, uma boa abordagem é selecionar rotas que possam prolongar o tempo de vida da rede na comunicação intergrupo, e rotas compostas por nós com mais recursos energéticos devem ser preferidas.

Qualidade de Serviço: Aplicações e funcionalidade de RSSFs necessitam de qualidade de serviço (QoS). Geralmente, amostra eficaz, menor atraso e precisão temporal são necessárias. É difícil para todos os protocolos de roteamento satisfazer todos os requisitos de QoS, porque algumas demandas podem violar um ou mais dos princípios do protocolo. Abordagens existentes de agrupamento em RSSFs são principalmente focadas no aumento da

eficiência energética ao invés da QoS. Métricas de QoS devem ser levadas em consideração em muitas aplicações de tempo real, como rastreamento de campo de batalha, monitoramento de eventos, entre outros.

- **ATRIBUTOS DO SISTEMA DE AGRUPAMENTO**

Os atributos de agrupamento em RSSFs, geralmente, podem ser classificados em características do agrupamento, características do líder do grupo, processo de agrupamento e o processo inteiro do algoritmo. Nessa seção, serão discutidos alguns detalhes a respeito dos atributos do agrupamento em RSSFs.

- **CARACTERÍSTICAS DO GRUPO**

Variabilidade da contagem dos grupos: Baseado na variabilidade da contagem dos grupos, esquemas de agrupamento podem ser classificados em dois tipos: fixos e variáveis. No primeiro, o conjunto de CHs é predeterminado e o número de grupos é fixo. Já no último caso, o número de grupos é variável, onde os CHs são selecionados, aleatoriamente ou baseado em um conjunto de regras, dos nós sensores depositados.

Uniformidade do tamanho dos grupos: À luz da uniformidade do tamanho dos grupos, protocolos de roteamento por agrupamento em RSSFs podem ser classificados da seguinte forma: uniforme e não uniformes, respectivamente com grupos de mesmo tamanho e com grupos de tamanhos diferentes na rede. Em geral, o agrupamento com grupos de tamanhos diferentes é usado para alcançar um consumo de energia mais uniforme e evitar buracos de energia.

Roteamento Intragrupo: No que tange os métodos de roteamento intragrupo, as maneiras de roteamento por agrupamento em RSSFs incluem duas classes: métodos de roteamento com um único salto dentro do grupo e os de saltos múltiplos. Para a forma de salto simples, todos os MNs do grupo transmitem os dados para seu CH correspondente diretamente. Ao contrário, a retransmissão de dados é usada quando os MNs comunicam-se com o correspondente CH do grupo.

Roteamento Intergrupo: Baseado na forma de roteamento intergrupo, protocolos de roteamento por agrupamento em RSSFs incluem duas classes: roteamento de salto simples e de saltos múltiplos. Para a forma de saltos simples, todos os CHs comunicam-se diretamente com a BS. Já no esquema de roteamento intergrupo com saltos múltiplos a retransmissão de dados é usada, como pode ser visto na Figura 4.

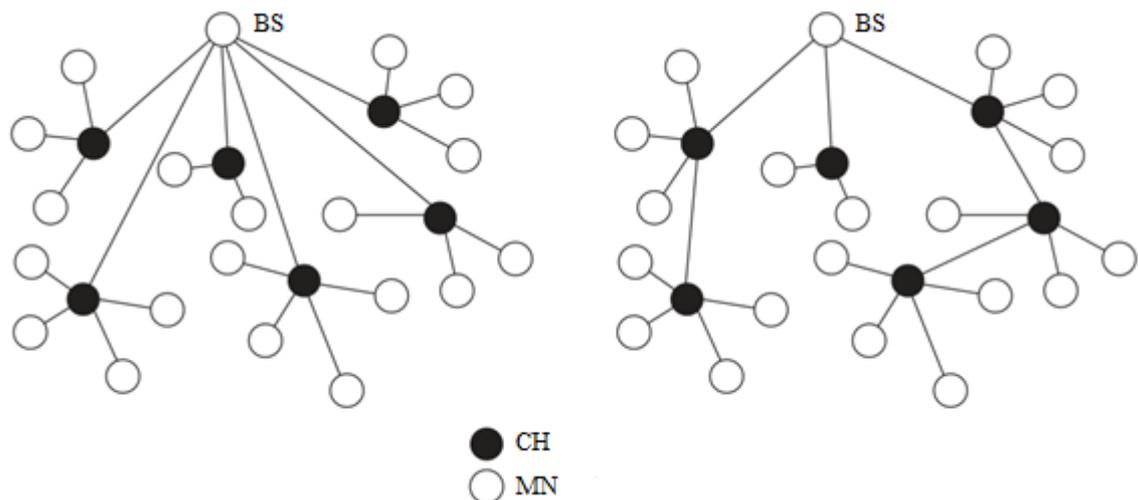


Figura 4 Agrupamento com salto único (esquerda) e com saltos múltiplos (direita)

- **CARACTERÍSTICAS DO LÍDER DO GRUPO**

Existência: Baseado na existência de líder de grupo dentro de um grupo, os esquemas de agrupamento podem ser divididos em dois tipos: baseados em líder de grupo e agrupamento que não é baseado em líder do grupo. No primeiro, existe pelo menos um CH dentro do grupo, enquanto no outro esquema não há nenhum CH dentro do grupo.

Diferença de Capacidades: Baseado na uniformidade de recursos de cada nó sensor, esquemas de agrupamento em RSSFs podem ser classificados em homogêneos e heterógenos. No esquema homogêneo, todos os nós sensores possuem recursos iguais de energia, processamento e comunicação e os CHs são escolhidos de forma aleatória ou seguindo algum critério específico. Já em ambientes heterogêneos, os nós sensores possuem capacidades diferentes, onde a regra é escolher como CH os nós sensores com maior capacidade.

Mobilidade: De acordo com os atributos de mobilidade dos CHs, abordagens de agrupamento em RSSFs são definidas de duas formas: móvel ou estacionária. Na primeira forma, CHs são móveis e o conjunto de membros muda dinamicamente. O contrário acontece na maneira estacionária, onde CHs são fixos e podem manter grupos estáveis, o que é mais fácil de gerenciar. Em alguns casos, os CHs podem mover-se em distâncias limitadas para um melhor desempenho da rede (Abbasi et al., 2007).

Função: Um CH pode atuar simplesmente como um retransmissor do tráfego gerado pelos nós sensores em seu grupo ou realizar agregação/fusão de dados das informações coletadas pelos nós sensores em seu grupo. Em alguns casos, o líder do grupo atua como sorvedouro/BS tomando decisões baseadas no fenômeno detectado ou objetivo (Abbasi et al, 2007).

- **PROCESSO DE AGRUPAMENTO**

Formas de Controle: Baseado nas formas de controle do agrupamento, métodos de roteamento por agrupamento em RSSFs podem ser agrupados de forma centralizada, distribuída e híbrida. Na forma centralizada, o sorvedouro ou CH requer informação global da rede ou do grupo para controlar a rede ou o grupo. Na abordagem distribuída, um nó sensor pode se tornar um CH ou entrar em um grupo formado por iniciativa própria sem a informação global da rede ou do grupo. Esquemas híbridos são compostos pelas abordagens centralizada e distribuída. Dessa forma a abordagem distribuída é usada para a coordenação entre CHs, e a forma centralizada é usada pelos CHs para construir os grupos individuais.

Natureza de Execução: Considerando a natureza de execução da formação dos grupos, modos de agrupamento em RSSF podem ser classificados da seguinte forma: probabilística e iterativa. No agrupamento probabilístico, uma probabilidade atribuída a todos os nós sensores é usada para determinar a função dos nós sensores. Em outras palavras, cada nó sensor pode decidir independentemente o seu próprio papel. Já na abordagem iterativa, cada nó sensor deve esperar até que certo número de iterações seja alcançada ou determinada quantidade de nós decidam suas funções antes de tomar a decisão.

Tempo de convergência: Considerando o tempo de convergência, os métodos de agrupamento em RSSFs podem ser classificados em algoritmos de tempo de convergência variável e constante. O tempo de convergência depende da quantidade de nós sensores na rede nos algoritmos de convergência variáveis, que é um método que se adapta bem a redes pequenas. Já para os algoritmos de convergência de tempo constante, o tempo de convergência é independente do tamanho da rede.

Parâmetros para Eleição do CH: Baseado nos parâmetros de eleição do CH, as abordagens de agrupamento podem ser categorizadas como determinística, adaptativa e aleatória. Nos esquemas determinísticos, atributos especiais inerentes aos nós sensores são considerados, tais como identificador (ID) e número de vizinhos. Na forma adaptativa, os nós sensores eleitos como CHs são aqueles que apresentam maior peso, que é calculado a partir de métricas que incluem energia residual, custo de comunicação, etc. Na maneira aleatória, a mais utilizada em algoritmos de segurança, CHs são eleitos aleatoriamente sem levar em conta métricas como energia residual, custo de comunicação, entre outros fatores.

Proatividade: Conforme a proatividade do roteamento por agrupamento, métodos de agrupamento podem ser classificados em proativo, reativo e híbrido. Nas redes proativas, todas as rotas entre as fontes e a BS são calculadas e armazenadas antes que sejam realmente necessárias, independentemente do tráfego de dados. Dessa forma uma mensagem viaja através de uma rota predeterminada para a BS. Nas redes reativas, não há rota predeterminada da fonte para BS. Abordagens híbridas usam uma combinação das duas ideias acima. Para esse tipo de roteamento por agrupamento, às vezes o agrupamento proativo é adotado, porém em outros momentos o modo reativo é usado. Por exemplo, APTEEM (Manjeshwar et al, 2002) é uma abordagem híbrida clássica. De acordo com a necessidade do usuário, esse protocolo ajusta algum parâmetro e chaveia entre os modos proativo e reativo para transmitir dados.

- **PROCESSO INTEIRO DO ALGORITMO**

Estágios do Algoritmo: Em geral, um algoritmo de roteamento por agrupamento completo compreende dois estágios, construção e transmissão de dados. Em relação ao

procedimento completo de formação dos grupos nos algoritmos de agrupamento, protocolos de roteamento em RSSFs podem ser classificados em baseados na construção dos grupos e nos baseados na transmissão. No primeiro, trata-se principalmente a construção dos grupos, enquanto há menos preocupação com a transmissão dos dados, que é executada por um método simples. No último caso o foco é dado para a transmissão dos dados, e dá-se menos importância para formação dos grupos.

2.1.3 LEACH

Low-Energy Adaptive Clustering Hierarchy (LEACH), proposto por (Heinzelman et. al., 2002) é uma das abordagens pioneiras de roteamento por agrupamento para RSSFs. A ideia básica do LEACH tem sido uma inspiração para protocolos de roteamento subsequentes. O principal objetivo do LEACH é selecionar os CHs por rotação, dessa forma a energia dissipada na comunicação para a BS é espalhada para todos os sensores na rede.

A operação do LEACH ocorre em várias rodadas, onde cada rodada é separada em duas fases, fase de construção e fase estacionária. Na fase de construção os grupos são organizados, enquanto na fase estacionária os dados são entregues para a BS, como pode visto na Figura 5. A fase de construção pode ser dividida em três etapas: anúncio, formação dos grupos e criação do cronograma.

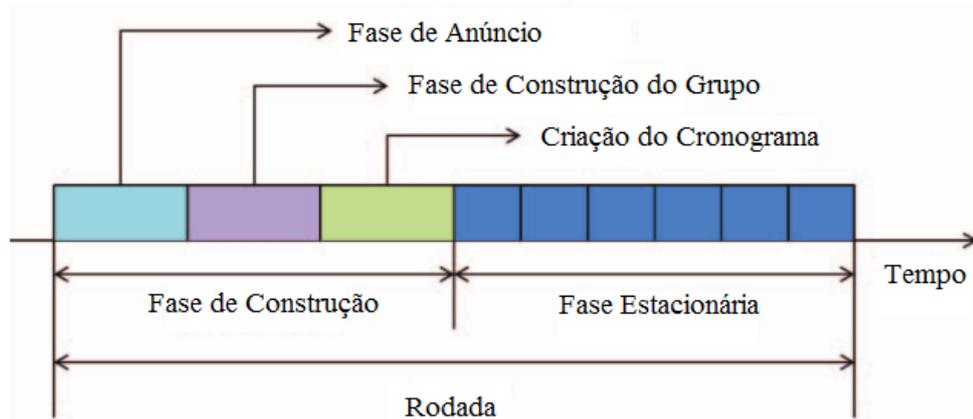


Figura 5 Estrutura de uma rodada do protocolo LEACH (Park et al, 2013)

Durante a fase de construção, cada nó sensor decide se será um CH ou não na rodada atual. Essa decisão é baseada na porcentagem de grupos sugerida de CHs para a rede e o número de vezes que cada nó foi CH até aquele dado momento. Para que essa decisão seja tomada cada nó escolhe aleatoriamente um número entre 0 e 1. Os nós sensores tornam-se CH para a rodada atual se o número é menor que o seguinte limiar:

$$T(n) = \begin{cases} \frac{P}{1 - P \left(r \bmod \frac{1}{P} \right)}, & \text{se } n \in G \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

Onde P é a porcentagem de CHs desejada, r é a rodada atual, e G é o conjunto de nós que não foram eleitos como CHs nas últimas $1/P$ rodadas. Quando um nó é eleito CH com sucesso, difundi-se uma mensagem de aviso para os outros nós. De acordo com a força do sinal recebido, outros nós sensores decidem em que grupo entrar para uma dada rodada e enviam uma mensagem de adesão para seu CH. Para distribuir uniformemente o consumo de energia entre os nós sensores, ocorre a rotação dos CHs em cada rodada pela geração de uma nova fase de anúncio baseada na equação (5). Durante a fase estacionária, os nós sensores coletam e transmitem dados para os CHs. Os CHs realizam a compressão dos dados que

chegaram dos nós que pertencem ao respectivo grupo, e enviam um pacote agregado diretamente para a BS. Além disso, o LEACH utiliza TDMA/CDMA MAC para reduzir as colisões intergrupo e intragrupo. Após certo período de tempo, que é determinado a priori, a rede volta para a fase de construção novamente e acontece uma nova rodada de seleção de CH. A Figura 6 mostra a topologia básica do LEACH.

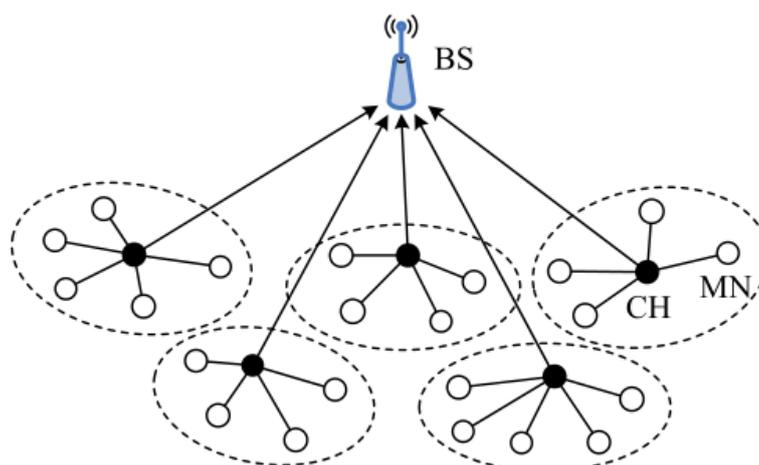


Figura 6 Topologia básica do LEACH

LEACH é uma abordagem completamente distribuída e não requer informação global da rede. Na literatura, várias modificações têm sido feitas no protocolo LEACH, tais como: TL-LEACH (Loscri et. al., 2005), E-LEACH (Fan et. al., 2007), M-LEACH (Mo Xiaoyan, 2006), LEACH-C (Heinzelman et. al., 2002), V-LEACH (Yassein et. al., 2002), LEACH-FL (Ran et. al., 2010), W-LEACH (Abdulsalam et. al., 2010), T-LEACH (Hong et. al., 2009), etc. As vantagens do LEACH incluem o seguinte (Lai et. al., 2012): (1) Qualquer nó eleito como CH em uma certa rodada não pode ser selecionado como CH novamente, então cada nó pode compartilhar igualmente a carga imposta aos CHs até certo ponto; (2) Utilizar TDMA evita colisões desnecessárias aos CHs; (3) Os membros dos grupos podem

abrir ou fechar a interface de comunicação em conformidade com seu espaço de tempo alocado para evitar gasto excessivo de energia.

Porém, LEACH possui algumas desvantagens: (1) Executa o método de roteamento intergrupo em um salto, diretamente dos CHs para a BS, o que não é aplicável para redes de larga escala. Nem sempre é uma suposição realística, o roteamento intergrupo de salto único para longas distâncias de comunicação. Além disso, a comunicação direta dos CHs para a BS pode gerar muito consumo de energia; (2) Apesar do fato da rotação dos CHs ser realizada em cada rodada para alcançar o balanceamento de carga, o LEACH não pode garantir real balanceamento de carga nos casos dos nós sensores com quantidade inicial de energia diferente, porque CHs são eleitos em termos da probabilidade sem a considerar energia. Nós sensores, com menos energia inicial, morrerão prematuramente. Isso pode ocasionar problemas como buracos de energia e cobertura; (3) Desde que a eleição de CH é executada em termos de probabilidade, é difícil obter uma distribuição uniforme de CHs na rede inteira. Portanto pode haver CHs em uma parte concentrada da rede, porém alguns nós podem nem mesmo ter CHs em sua vizinhança; (4) A ideia de agrupamento dinâmico traz sobrecarga extra. Por exemplo, CH muda e os avisos podem diminuir o ganho em consumo de energia.

2.2 APRENDIZADO NÃO SUPERVISIONADO

Os algoritmos de aprendizado supervisionado fazem o uso de um conjunto de treinamento que consiste em uma coleção de alvos rotulados. Dessa forma pode-se mostrar ao algoritmo a resposta correta para diversas situações, porém em muitas circunstâncias isso é algo difícil de se conseguir.

O objetivo do aprendizado não supervisionado é explorar as similaridades entre as entradas para formar grupos, já que não existem informações sobre a saída correta disponível e o algoritmo tem que detectar, sozinho, alguma similaridade entre as diferentes entradas.

2.2.1 K-MEANS

Entre as técnicas de aprendizado não supervisionado a mais popular é um algoritmo conhecido como k-Means, por sua simplicidade e flexibilidade. Exemplos de aplicações onde as técnicas de agrupamento são utilizadas abrangem segmentação de mercado, análise de redes sociais, organização de grupos de servidores e análise de dados astronômicos.

Suponha que queremos dividir os dados de entrada em k classes, onde sabemos o valor de k . Alocamos k centros de grupos para o espaço de entrada, e queremos posicionar esses centros para que resulte em um centro de grupo que melhor represente cada grupo. Porém, não sabemos onde os grupos estão, muito menos onde estão seus centros, portanto precisamos de um algoritmo para encontrá-los.

Dessa forma podemos descrever o algoritmo k-Means da seguinte forma, suponha que temos um conjunto de dados $x^{(i)}$, $i = 1, \dots, m$, e queremos organiza-los em k grupos onde $c^{(i)}$, $i = 1, \dots, K$ representa o índice do grupo onde a amostra $x^{(i)}$ está atualmente, μ_k representa o centróide do grupo k ($\mu_k \in \mathbb{R}^n$) e $\mu_{c^{(i)}}$ representa o centróide do grupo a que a amostra $x^{(i)}$ foi atribuído. Portanto o objetivo é, por meio da uma função de custo representada na Equação 2, minimizar a soma das distâncias entre cada centróide e os dados pertencentes ao seu grupo, como pode ser visto na Equação 3:

$$J(c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_k) = \frac{1}{m} \sum_{i=1}^m \|x^i - \mu_{c^{(i)}}\|^2 \quad (2)$$

$$\min_{c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_K} J(c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_K) \quad (3)$$

O algoritmo é composto pelos seguintes passos (Xu, 2005):

1. Inicializar k centróides aleatoriamente, que representam os centróides dos grupos inicialmente.
2. Atribuir as amostras ao grupo com centróide mais próxima.
3. Quando todos os pontos forem atribuídos, recalculer a posição do centróide de cada grupo, utilizando a informação dos pontos pertencentes a cada grupo.
4. Repetir os passos 2 e 3 até que os grupos não mudem mais.

Na Figura 7 pode-se observar um exemplo de aplicação do algoritmo k-Means. Inicialmente os centróides, representados por 'x', foram atribuídos a pontos aleatórios. No processo de execução do algoritmo cada centróide percorre o caminho, representado pelas linhas, até que os grupos não mudem, ou seja, os centróides parem. No final, podemos perceber a formação de três grupos, representados pelos pontos vermelhos, verdes e azuis.

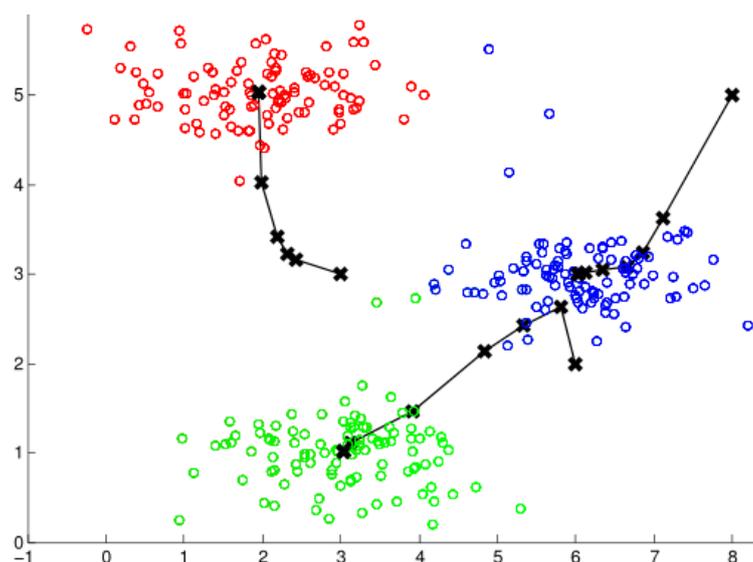


Figura 7 Algoritmo k-Means

2.2.2 FUNÇÃO SILHUETA

Um dos problemas que surgem no uso de técnicas de aprendizado não supervisionado é como avaliar tais técnicas em determinada aplicação, pois nesse caso, não temos uma coleção de alvos rotulados. Uma das questões levantadas pelo algoritmo k-means, apresentado anteriormente, é a quantidade de grupos k escolhida. Geralmente o número de grupos k é escolhido por meio da experiência do usuário ou o valor k deve ser estimado exclusivamente para os dados. A qualidade dos grupos resultantes depende fortemente do k estimado. Uma divisão com muitos grupos complica o resultado, portanto, tornando difícil de interpretar e analisar, enquanto uma divisão com poucos grupos causa perda de informação induzindo a decisão final ao erro. O problema de determinar o número de grupos é conhecido como “o problema fundamental da validade dos grupos” (Chen et. al., 1993).

A função silhueta é um dos métodos usados para estimar a quantidade de grupos k . O valor da função silhueta para cada ponto é uma medida do grau de similaridade entre um ponto e os outros pontos pertencentes ao mesmo grupo comparado à similaridade em relação aos pontos pertencentes a outros grupos, esse valor varia de -1 a 1. O valor da função silhueta para um ponto i , S_i , é definida como:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (4)$$

Onde a_i é a distância média do ponto i aos outros pontos no mesmo grupo i , e b_i é a distância média mínima do ponto i aos pontos em diferentes grupos, minimizada ao longo do grupo. A Figura 8 ilustra um exemplo da função silhueta aplicada a três grupos. Pode-se observar que uma grande quantidade dos dados está localizada na região próxima de zero, indicando que os pontos estão localizados na periferia dos grupos.

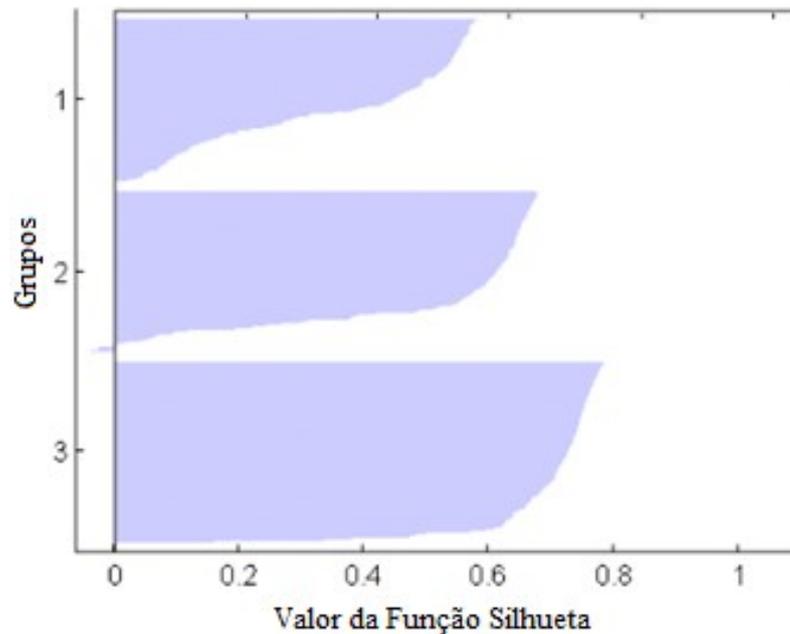


Figura 8 Função silhueta para 3 grupos

Dessa forma, a quantidade de grupos k que apresentar a maior quantidade de valores da função silhueta próximos de 1 e, ao mesmo tempo, a menor quantidade de valores menores que 0, é selecionada.

2.3 CONFORTO TÉRMICO

Conforto térmico é a condição da mente que expressa satisfação com a temperatura ambiente e é avaliada de forma subjetiva (ASHARAE, 2010). A manutenção desse padrão de conforto térmico para ocupantes de ambientes fechados é um dos principais objetivos para o projeto de engenharia de um HVAC (heating, ventilation and air conditioning).

O modelo PMV (Predicted Mean Vote) está entre os modelos de conforto térmico mais reconhecidos. Foi desenvolvido utilizando princípios de equilíbrio térmico e dados coletados de forma experimental em uma câmara de controle sob condições de estado estacionário. Como o modelo PMV foi derivado de dados coletados em um ambiente controlado, não é adequado para um ambiente onde há espaços com ventilação natural onde geralmente existem níveis altos de movimento do ar. Essa deficiência motivou o desenvolvimento de uma família de modelos estatísticos empíricos do conforto térmico em espaços ventilados conhecidos como modelos adaptativos (Djamila et. al., 2012). Modelos adaptativos do conforto térmico consideram os ocupantes interagindo dinamicamente com seu ambiente, e controlando seu conforto térmico por meio de roupas ou abrindo-fechando uma janela. Isso está em contraste com conforto térmico em ambiente fechados, em que os ocupantes experimentam o ambiente passivamente (Djalmila et. al., 2013).

2.3.1 FATORES DETERMINANTES PARA O CONFORTO TÉRMICO

Uma vez que há grande variação de pessoa para pessoa em termos de satisfação fisiológica e psicológica, é difícil encontrar uma temperatura ótima para todos em um dado espaço. Dados de laboratório e campo foram coletados para definir condições que são estabelecidas como confortáveis para uma porcentagem específica de ocupantes (ASHARAE, 2010).

Há seis fatores primários que afetam diretamente o conforto térmico e que podem ser agrupados em duas categorias: **fatores pessoais** – porque são característicos dos ocupantes e **fatores ambientais** – que são condições do ambiente. O primeiro refere-se a taxa metabólica e o nível de isolamento térmica devido a roupa. O último refere-se a temperatura do ar, temperatura radiante, velocidade do ar e umidade relativa do ar.

Mesmo que todos esses fatores possam variar com o tempo, os padrões geralmente referem-se ao estado estacionário para estudar conforto térmico, permitindo apenas variações de temperatura limitadas.

2.3.1.1 TAXA METABÓLICA

Pessoas têm taxas metabólicas diferentes que pode flutuar conforme o nível de atividade e condições ambientais (Toftem et. al., 2005; Smolder et. al., 2002; Khodakami et. al., 2009).

O padrão ASHARE 55-2010 define taxa metabólica como sendo o nível de transformação de energia química em calor e trabalho mecânico pela atividade metabólica dentro de um organismo, geralmente expressa em termos de unidade de área de toda a

superfície corporal. No padrão, taxa metabólica é expressa na unidade MET, que é definida como: $1 \text{ MET} = 58,3 \text{ W/m}^2$.

2.3.1.2 ISOLAMENTO TÉRMICO

A quantidade de isolamento térmico causado pela roupa usada por uma pessoa possui um impacto substancial no conforto térmico, pois influencia a perda de calor e conseqüentemente no equilíbrio térmico. A quantidade padrão de isolação necessária para manter uma pessoa em repouso aquecida em uma sala sem vento à $21,1 \text{ }^\circ\text{C}$ é igual a 1 clo.

2.3.1.3 TEMPERATURA DO AR

A temperatura do ar é a temperatura média do ar em volta do ocupante, com respeito a localização e o tempo. De acordo com o padrão ASHARE 55, a média espacial leva em consideração o tornozelo, a cintura e a cabeça, que varia para ocupantes sentados e em pé. A média temporal é baseada em intervalos de três minutos com pelo menos 18 pontos igualmente espaçados no tempo.

2.3.1.4 TEMPERATURA RADIANTE

A temperatura radiante está relacionada à quantidade de calor irradiada por uma superfície, e depende da emissividade do material – ou seja, a habilidade de absorver ou emitir calor. A temperatura média radiante é definida como a temperatura uniforme de um ambiente fechado imaginário onde a transferência de calor radiante do corpo humano é igual a

transferência de calor radiante no ambiente fechado atual não uniforme, é uma variável chave para o cálculo do conforto térmico para o corpo humano.

MRT (Mean Radiant Temperature) é o parâmetro que mais influencia o equilíbrio da energia humana, especialmente em dias quentes e ensolarados. De fato, em climas quentes é cerca de duas vezes mais significativo que a temperatura do ar devido às roupas mais leves, enquanto em climas frios possui a mesma influência que a temperatura do ar (Szokolay et. al., 2008). MRT influencia fortemente os índices de conforto termo-fisiológicos como PET (Physiological Equivalent Temperature) ou PMV (Predicted Mean Vote) (Fanger et al., 1970).

O que experimentados e sentimos em relação ao conforto térmico em ambientes fechados está relacionado a influência de ambos, temperatura do ar e a temperatura das superfícies desse espaço. A manutenção do equilíbrio entre a temperatura prática e a temperatura média radiante pode criar espaços mais confortáveis.

2.3.1.5 VELOCIDADE DO AR

Velocidade do ar é definida como a taxa de movimento do ar em um ponto, desconsiderando a direção. De acordo com o padrão ASHRAE 55, é a velocidade média do ar a que o corpo está exposto, com respeito ao local e ao tempo. A média temporal é a mesma que a temperatura do ar, enquanto a média espacial é baseada na suposição que o corpo está exposto ao ar com velocidade uniforme. Porém, em alguns espaços pode-se encontrar ar com velocidade não uniforme e, conseqüentemente, a perda de calor da pele não pode ser considerada uniforme. Portanto, deve-se decidir a velocidade média adequada, especialmente incluindo incidência da velocidade do ar em partes do corpo descoberta, que possui grande efeito de resfriamento e potencial desconforto local (ASHRAE, 2010).

2.3.1.6 UMIDADE RELATIVA DO AR

Enquanto o corpo humano tem sensores na pele muito eficientes na sensação de quente e frio, umidade relativa (RH) é difícil de detectar. A influência da umidade na percepção do ambiente fechado pode desempenhar um papel importante na temperatura percebida e seu conforto térmico. Quanto menor a RH mais suor evapora do corpo, enquanto para valores elevado de RH é mais difícil esse processo acontecer, porque a umidade do ar já é elevada (La Roche et. al., 2011). Portanto, ambientes muito úmidos ($RH > 70-80\%$) são geralmente desconfortáveis, porque o ar está próximo do nível de saturação, reduzindo fortemente a possibilidade de perda de calor por evaporação. Por outro lado, ambientes muito secos ($RH < 20-30\%$) também são desconfortáveis por causa do seu efeito nas membranas mucosas. O recomendado nível de umidade em ambientes fechados está na faixa de 30%-60% (Balaras et. al., 2007; Wolkoff et. al., 2007).

2.3.2 PREDITORES DO CONFORTO TÉRMICO

De longe o preditor de conforto térmico mais utilizado, o PMV (Predicted Mean Vote) é um preditor do valor médio de votos de um número grande de pessoas em uma escala térmica de 7 pontos baseado no equilíbrio térmico do corpo humano (ISO 7730, 2005), como pode ser observado na Tabela 1. O equilíbrio térmico é alcançado quando o calor produzido no corpo é igual a perda de calor para o ambiente. Diz-se que um ambiente fechado, como uma sala, está em condições térmicas ótimas quando o valor do PMV é igual a zero.

Tabela 1 PMV (Predicted Mean Vote)

+3	Quente
+2	Morno
+1	Levemente Morno
0	Neutro
-1	Levemente Fresco
-2	Fresco
-3	Frio

Porém a principal desvantagem do PMV é que para seu cálculo, são necessárias variáveis difíceis de obter, como a taxa metabólica e o nível de isolamento térmico no que diz respeito aos fatores pessoais e em relação aos fatores ambientais, velocidade do ar e radiação térmica devido ao custo dos sensores para uma medição precisa. Dessa forma a maioria das variáveis do PMV costuma ser estimada para casos específicos.

Outro preditor de conforto térmico que tem sido estudado nos últimos anos é conhecido como Humidex (Materson et al., 1979). O Humidex é um índice numérico

utilizado pelos meteorologistas canadenses para descrever a sensação térmica percebida por uma pessoa, combinando os efeitos de calor e umidade. O Humidex é um valor numérico sem unidade, porém, geralmente é referido como a temperatura em graus Celsius. Por exemplo, se a temperatura é de 30°C, mas o Humidex calculado é de 35, então a temperatura percebida pelas pessoas é de 35°C.

O Humidex é usado principalmente em ambientes abertos, porém em (Rajib Rana et. al., 2013) os autores o avaliam como um preditor de conforto térmico para ambientes fechados. Os resultados mostram que o Humidex é até 20% melhor na predição do conforto térmico em relação à temperatura e umidade relativa para regiões úmidas.

A principal vantagem do Humidex é sua simplicidade, pois para o seu cálculo basta obter os valores de temperatura e umidade de um ambiente fechado. Porém esse preditor é utilizado apenas para prever situações de desconforto devido a altos valores de umidade relativa do ar e temperatura, não indicando desconforto térmico para ambientes muito secos, devido à umidade relativa, ou muito frios, devido a baixas temperaturas.

TRABALHOS CORRELATOS

Os protocolos de roteamento tem como um dos objetivos a extensão do tempo de vida da rede. Nesse contexto uma das abordagens mais disseminadas é a probabilística onde há a rotação dos CHs para o balanceamento do consumo de energia dos nós sensores da rede. O LEACH e seus derivados adotam esse mecanismo de seleção de CHs. Porém, nos últimos anos têm surgido abordagens determinísticas que visam reduzir a distância entre os nós sensores pertencentes ao mesmo grupo. Reduzindo o consumo de energia por meio da redução da distância das transmissões e recepções dentro de cada grupo. Neste capítulo trataremos as abordagens determinísticas que têm surgido nos últimos anos.

3.1 PROTOCOLOS DE ROTEAMENTO DETERMINÍSTICOS

Li et al. (2009) desenvolveu um algoritmo de roteamento por agrupamento de nós sensores para o monitoramento de ambientes fechados utilizando o algoritmo k-Means. O principal argumento dos autores é que os algoritmos de agrupamento em RSSFs não são flexíveis para o monitoramento de ambientes fechados por possuírem natureza distribuída. Os autores listam algumas vantagens em utilizar agrupamento em RSSFs como a criação de uma estrutura organizacional e de gerenciamento entre os sensores, bom desempenho na extensão do tempo de vida da rede, pois evita-se o desperdício de energia nas transmissões frequentes para a base central e a redução de tráfego na rede, quando se realiza fusão de dados no líder do grupo, evitando perdas devido a colisão de pacotes. Além disso, o agrupamento depende

da aplicação, ou seja, quando se projeta um algoritmo de agrupamento, as características da aplicação devem ser levadas em consideração.

Para resolver a questão da escolha da quantidade de grupos k , pode-se utilizar a Equação 5 (Heinzelman et al, 2002) que determina o número ótimo de grupos, quando há N nós distribuídos uniformemente em uma região com dimensões $M \times M$, levando em consideração o alcance de cada nó sensor:

$$k_{opt} = \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}} \frac{M}{d_{toBS}^2}} \quad (5)$$

Onde ϵ_{fs} e ϵ_{mp} são, respectivamente, a energia amplificada em unidades de distância para modelos diferentes de canal, f_s é o espaço livre e mp é o desvanecimento por multipercurso. d_{toBS} denota a distância entre o líder do grupo e a estação base (sorvedouro) e não é um valor fixo, mas uma faixa de valores que de 0 a M . Portanto o número de grupos é um inteiro k , $k \in \left(1, \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}} \frac{M}{d_{toBS}^2}}\right)$.

O algoritmo é dividido em duas partes, seleção do líder do grupo e formação dos grupos. Na seleção do líder do grupo inicialmente cada nó envia suas coordenadas à estação base que por sua vez calcula a quantidade de grupos k , utilizando, além da equação (1), a silhueta, que é uma função que mede a similaridade entre os nós sensores dentro e fora do grupo, variando de -1 a +1. Essa medida varia de +1, indicando pontos que são muito distantes dos grupos vizinhos, passando por 0, indicando pontos que estão na divisa dos grupos, até -1, indicando pontos que provavelmente foram atribuídos a grupos errados. Uma vez que k é confirmado, a informação do agrupamento pode ser determinada. Os nós mais próximo dos centróides são escolhidos como líderes dos grupos.

Na etapa de formação dos grupos a estação base envia as informações a respeito do agrupamento para a rede e os nós sensores entram no grupo correspondente. A estação base envia também um limiar de energia que determina quando deverá ocorrer o re-agrupamento para mudança do líder do grupo.

Por fim os autores argumentam que o algoritmo centralizado seria a melhor opção para reduzir o consumo de energia nos nós sensores, visto que a estação base está ligada diretamente à energia. Porém nenhum estudo comparativo é realizado para comprovar o desempenho do algoritmo de roteamento proposto.

Em (Park et al, 2013) os autores propõem um algoritmo de seleção dos líderes de grupos para melhorar a questão do consumo de energia de RSSF. Ao contrário de abordagens aleatórias e probabilísticas como LEACH e suas variações, emprega-se uma abordagem determinística usando o algoritmo k-Means. O algoritmo é usado para formar grupos de modo que a distância entre cada nó sensor e seu respectivo CH seja mínima. A abordagem proposta portanto permite minimizar o consumo de energia para os nós sensores no envio de dados para o CH do seu grupo. Como consequência, o tempo de vida da RSSF pode ser prolongado.

O modelo de consumo de energia adotado, o qual está baseado em transmissões e recepções, é apresentado na Equação 6.

$$E_{Tx}(k, d) = \begin{cases} E_{elec} * k + \varepsilon_{fs} * d^2, & d < d_0 \\ E_{elec} * k + \varepsilon_{mp} * d^4, & d \geq d_0 \end{cases} \quad (6)$$

Onde E_{elec} é a energia necessária para o processamento de 1 bit de dados com o circuito eletrônico. ε_{fs} e ε_{mp} são a energia consumida na transmissão de 1 bit de dados para alcançar uma taxa de erro aceitável no caso do modelo de espaço livre e modelo multipercurso, respectivamente, e dependem da distância de transmissão. A energia dissipada

para os modelos de espaço livre e multipercursos são proporcionais a d^2 e d^4 , respectivamente. O limiar d_o , é calculado pela Equação 7:

$$d_o = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \quad (7)$$

A energia consumida para receber uma mensagem de k bits é calculada pela Equação 8:

$$E_{Rx} = E_{elec} * k \quad (8)$$

O esquema proposto usa o algoritmo k-Means para formar os grupos baseados na distância Euclidiana entre os nós sensores. Após a formação dos grupos, um ID é atribuído a cada nó sensor de um grupo de acordo com a distância do centróide, atribuindo números menores aos mais próximos. A figura 9 mostra a forma como os IDs são atribuídos. O ID indica a ordem para escolha do líder do grupo (CH). Portanto, o ID tem um papel importante na seleção do CH.

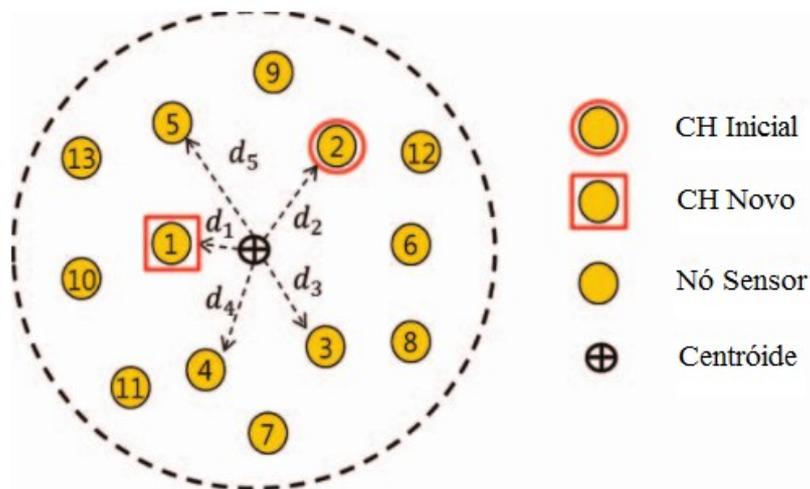


Figura 9 Ordenação dos nós sensores com ID (Park et al, 2013)

Os resultados da simulação revelam que o esquema proposto supera os protocolos de roteamento existentes, como LEACH, em termos de energia residual da rede e número de nós sensores mortos por rodada. Além disso, o CH consome menos energia que nos outros esquemas.

No trabalho de Sasikumar et al. (2012) o algoritmo k-Means é utilizado na sua forma centralizada e distribuída para formar grupos de nós sensores de uma rede de sensores sem fio. Os resultados obtidos mostram que o agrupamento utilizando o k-Means distribuído, além de fornecer à rede uma maior estabilidade e independência em relação ao algoritmo de roteamento, também é mais rápido na formação dos grupos da rede e apresenta consumo de energia similar a forma centralizada.

Agrupamento é feito para relacionar nós similares e economizar a energia desperdiçada na transmissão direta de dados para o nó sorvedouro. Os nós da rede organizam-se em um nível hierárquico. Dentro de um grupo particular são realizadas agregação e roteamento pelo líder do grupo para reduzir a quantidade de transmissões de dados à base central. A formação do grupo é geralmente baseada na energia restante dos nós sensores e na proximidade dos sensores ao líder do grupo (Lin et al, 1997). Os nós, exceto o líder, escolhem o seu líder logo após a deposição e transmitem a informação coletada ao líder do grupo. O papel do líder do grupo, sendo ele também um nó sensor, é encaminhar esses dados e seus próprios dados a estação base após executar a agregação de dados.

A estratégia de simulação, do trabalho de Sasikumar et al. (2012), para o agrupamento centralizado e distribuído seguiu os seguintes passos:

Envio da posição e a energia de cada nó para o nó central: no agrupamento centralizado o nó sorvedouro (nó central) atua como autoridade na tomada de decisão. Então a posição e a energia de todos os nós devem estar disponíveis ao nó sensor central. No

agrupamento distribuído todos os nós participam no processo de tomada de decisão. Portanto a posição e a energia de todos os nós devem estar disponíveis para cada nó sensor.

Seleção do líder do grupo: após a parada dos centróides do algoritmo k-means no processo de agrupamento, considerou-se os nós sensores mais próximos e também os próximos nós sensores mais próximos. O nó com maior energia é eleito líder do grupo. Se mais de um nó sensor em níveis diferentes possuem a maior energia, então o nó sensor mais próximo do centróide é selecionado como líder do grupo. Se mais que um nó sensor possui a maior energia no mesmo nível de distância, o nó sensor com menor ID é eleito líder do grupo.

Declaração do líder do grupo: no agrupamento centralizado, após o fim do processo de agrupamento e seleção do líder do grupo, cada nó deve obter a informação sobre o grupo a que pertence e seu respectivo líder. Essa informação é fornecida a cada nó pelo nó central. No agrupamento distribuído, após a parada dos centróides no processo de agrupamento, consideram-se os nós que estão mais próximos e também os próximos mais próximos do centróide. O nó com energia mais elevada é eleito como líder do grupo.

Para avaliar o desempenho das duas estratégias de agrupamento, centralizada e distribuída, tanto o tempo para formação dos grupos quanto o consumo de energia foram levados em consideração. Na forma distribuída o tempo consumido inclui o tempo para troca de mensagens de controle (com informações da posição e energia de todos os nós) e o tempo de agrupamento (tempo de execução do algoritmo). Para o agrupamento centralizado o tempo consumido inclui o tempo de envio (envio de posição e energia da rede para a BS), tempo de agrupamento (tempo de execução do algoritmo) e reenvio (tempo para a BS enviar de volta para a rede as informações sobre o agrupamento).

Os resultados mostram que variando apenas o número de nós sensores, o tempo para criar grupos para a forma centralizada é maior que o tempo para o agrupamento distribuído,

devido à necessidade de reenviar as informações sobre o agrupamento, que não é necessário na forma distribuída, pois todos os nós executam o processo de agrupamento individualmente. Também mostram que não há muita diferença, entre a forma centralizada e distribuída, quanto à energia consumida, variando apenas o número de nós. Isso se deve à quantidade de energia gasta no agrupamento distribuído relacionada à troca de mensagens de controle (contendo posição e energia) entre todos os nós sensores ser praticamente igual à energia consumida em ambos, envio (dos MNs) para o nó central e reenvio (da BS) para todos os nós, na forma centralizada.

No trabalho de Sasikumar et al. (2012) avalia-se o algoritmo de agrupamento k-Means na sua forma centralizada e distribuída comparando o tempo necessário para a formação dos grupos e a quantidade de energia gasta. Porém esse trabalho desconsidera o tempo de execução do algoritmo k-Means por considerar desprezível.

Nenhum dos trabalhos abordados anteriormente utilizam as variáveis coletadas do ambiente no processo de formação dos grupos e de seleção de CHs. A novidade do nosso trabalho, em relação aos anteriores, é explorar a estrutura de correlação existente nas variáveis coletadas para formação de grupos de nós sensores com medidas correlacionadas e seleção do nó sensor que melhor represente cada grupo como CH. Dessa forma minimizamos tanto a distância entre os nós sensores pertencentes ao mesmo grupo quanto a quantidade de transmissões dentro de cada grupo.

METODOLOGIA

4.1 MINERAÇÃO DE DADOS

A base de dados usada na pesquisa desenvolvida nesta dissertação (Intel Berkeley, 2004) é o resultado de um experimento realizado no laboratório de Berkeley, ilustrado na Figura 1 na página 3. Desse experimento resultou uma base de dados com cerca de 2,3 milhões de leituras dos 54 nós sensores dispostos no laboratório. A Figura 10 mostra o volume de dados que cada nó sensor gerou no período de 28 de fevereiro a 5 de abril de 2004.

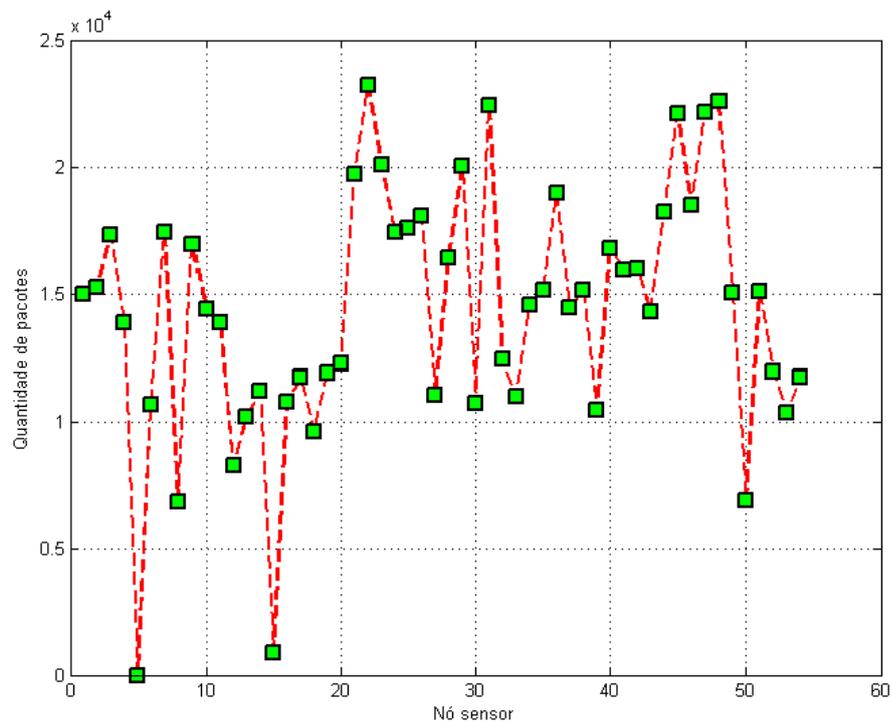


Figura 10 Quantidade de pacotes por nó sensor

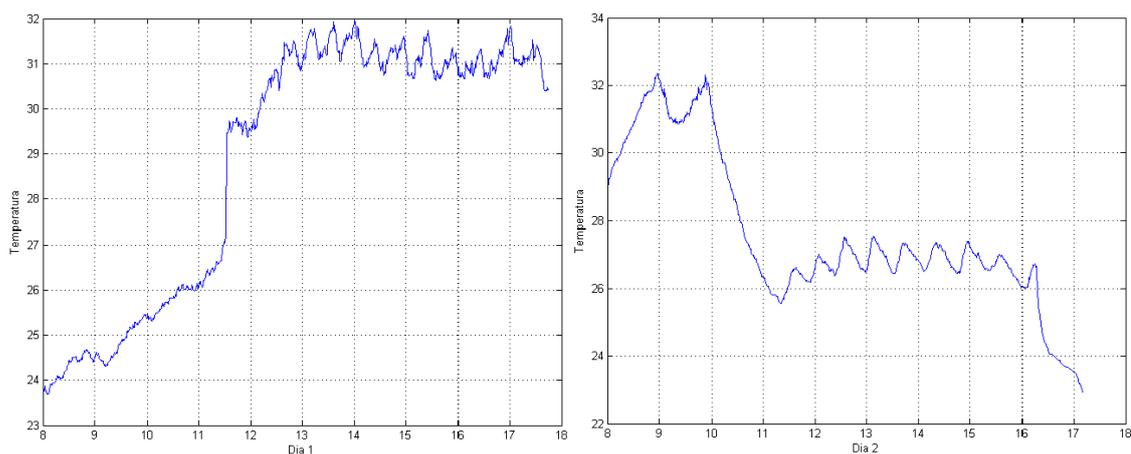


Figura 11 Distribuição da temperatura no 1º e 2º dia de Março

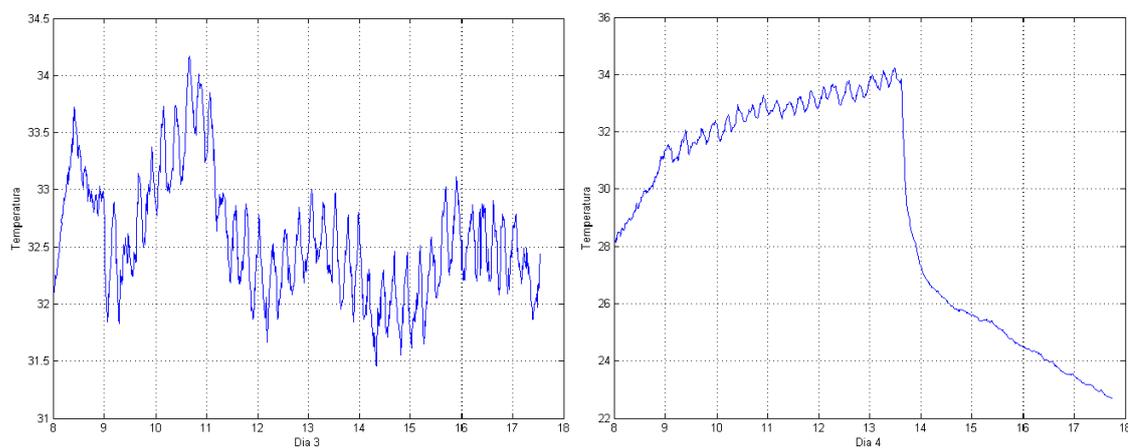


Figura 12 Distribuição da temperatura no 3º e 4º dia de Março

É difícil estimar a distribuição da temperatura de um lugar em um dia. Por exemplo, dado quatro dias, 1-4 de Março, podemos perceber que as distribuições da temperatura coletada pelo nó sensor 22 são diferentes (Figura 11 e 12). Elas mostraram padrão geral diferente. Por exemplo, a temperatura no primeiro de Março aumenta dramaticamente entre 8:00 até 14:00, enquanto no segundo dia de Março a temperatura apresenta um padrão completamente diferente quando comparado do primeiro dia de Março.

Como o objetivo da pesquisa realizada nesta dissertação é utilizar a correlação espacial presente nos dados coletados, é necessário o tratamento prévio dos dados para

remoção de valores inválidos e *outliers* (Wu et al., 2007), os quais são valores coletados que apresentam comportamento muito distante daquele apresentado pela maioria dos dados.

Dessa forma a metodologia de tratamento dos dados adotada foi a seguinte:

Compreensão dos dados: os dados de cada uma das 2.313.682 observações foram armazenados da seguinte forma: data, hora, época, ID do nó sensor, temperatura, umidade, luminosidade e tensão da bateria. Duas leituras da mesma época foram produzidas por nós diferentes ao mesmo tempo. Os IDs, que variam de 1 a 54, são as identidades dos nós sensores. A temperatura está em graus Celsius, variando de 10 a 40°C. A umidade relativa varia de 0 a 100%. A luminosidade é dada em Lux. A tensão na bateria é expressa em volts, variando de 2 a 3V.

Remoção de observações inválidas: nessa etapa toda medida que está fora do intervalo definido na etapa de compreensão dos dados é considerada inválida e removida. Deste modo, houve a remoção dos dados coletados de 26 de março a 5 de abril.

Remoção de outliers: foi utilizado o algoritmo k-Means para remoção dos outliers, que são medidas que estão dentro do intervalo de variação, porém estão bastante desconectados da maioria das outras medidas. Após variar a quantidade de grupos de 15 a 50, quando $k=16$, ambos, temperatura e umidade, apresentaram valores anormalmente elevados em um grupo pequeno, como pode ser visto na Figura 13. As leituras pertencentes a esse grupo foram removidas, restando apenas 750.148 observações.

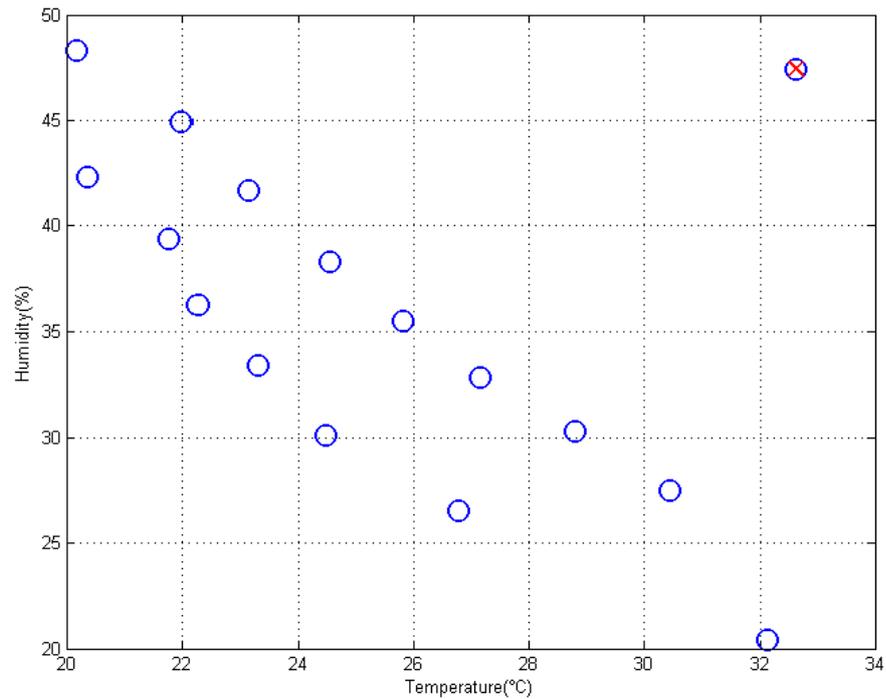


Figura 13 Centróides dos 16 grupos

4.2 MODELAGEM DA REDE

Na modelagem de uma RSSF define-se tanto a topologia da rede quanto o modelo de consumo de energia a ser adotado para os nós sensores que compõem a rede. Além disso, um dos pontos cruciais é a quantidade de grupos usada no protocolo de roteamento hierárquico.

A disposição dos nós sensores adotada neste trabalho, foi a do experimento realizado no laboratório de Berkeley, como mostra a Figura 1 na página 3. Como nosso cenário é de uma aplicação *indoor*, o modelo de propagação utilizado foi o multipercurso, pois em ambientes fechados há muita degradação no sinal transmitido devido aos obstáculos e o ruído do meio. Dessa forma o modelo de consumo de energia adotado é baseado em transmissões e recepções como pode ser observado nas equações a seguir:

$$E_{Tx}(k, d) = E_{elec} * k + \varepsilon_{mp} * d^4 \quad (9)$$

$$E_{Rx}(k) = E_{elec} * k \quad (10)$$

Onde E_{elec} é a energia necessária para o processamento de 1 bit de dados com o circuito eletrônico. ε_{mp} é a energia consumida na transmissão de 1 bit de dados para alcançar uma taxa de erro aceitável no modelo multipercurso. A Tabela 2 mostra os valores dos parâmetros adotados para simulação.

Tabela 2 Parâmetros de configuração da rede

Parâmetros	Valores
Tamanho da rede	(40,5m x 31m)
Localização da Estação Base	(24,5m:12m)
Número de nós sensores	53
Energia inicial	0,5J
E_{elec}	50nJ/bit/m ²
ε_{mp}	0,0013pJ/bit/m ²
Tamanho do pacote de dados	6400 Bytes
Energia consumida na agregação (E_{DA})	5nJ/bit

Para a definição da quantidade de grupos ótima, neste trabalho utilizamos como ferramenta a função silhueta que mede, para uma quantidade de grupos k , a similaridade entre os dados coletados e a localização dos nós sensores dentro e fora do grupo, variando de -1 a +1. Essa medida varia de +1, indicando pontos que são muito distantes dos grupos vizinhos, passando por 0, indicando pontos que estão na periferia dos grupos, até -1, indicando pontos que provavelmente foram atribuídos a grupos errados.

Dessa forma para a escolha da quantidade de grupos, variamos k de 3 a 10 e definimos um limiar descrito na Equação 11:

$$y = \text{sum_positive} - \text{sum_negative} \quad (11)$$

Onde sum_positive representa a soma dos valores cujo valor atribuído pela função silhueta é maior que 0,5 e sum_negative representa a soma dos valores cujo valor atribuído

pela função silhueta é menor que 0. A Figura 14 mostra valores da função silhueta para 5 grupos. Como pode ser visto há muito mais valores da função silhueta maiores que 0,5 do que valores da função silhueta menores que 0.

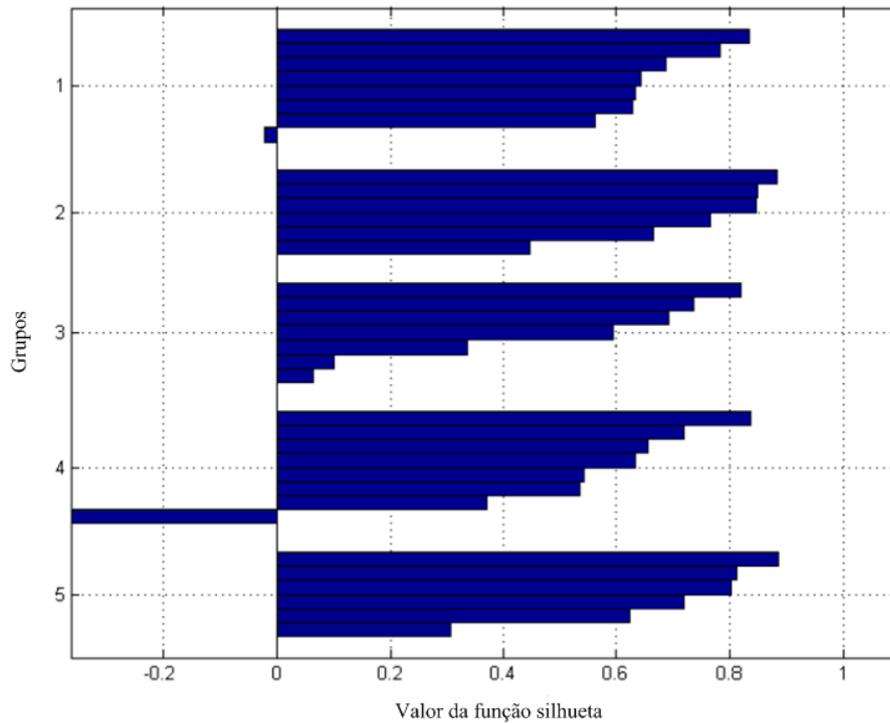


Figura 14 Valores da função silhueta para 5 grupos

A Figura 15 mostra o resultado da Equação 4 para k variando de 3 a 10. Pode-se perceber que para $k=5$ o resultado apresenta valor máximo, indicando que essa é a melhor escolha para esse cenário.

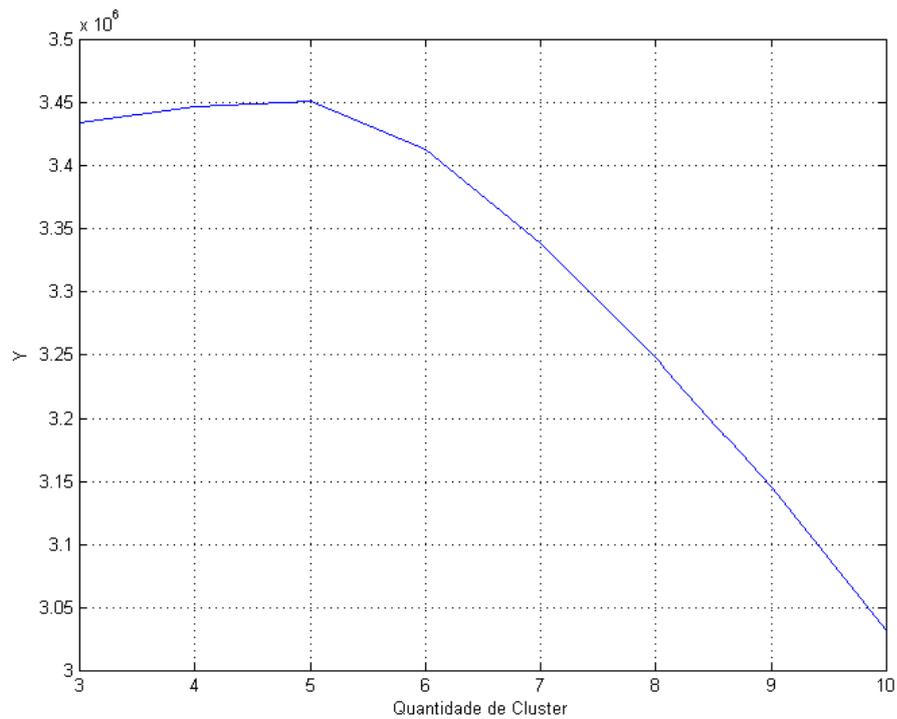


Figura 15 Definição da quantidade de grupos

4.3 PROTOCOLO DE ROTEAMENTO

Os protocolos de roteamento em RSSF são compostos pela fase de construção e fase estacionária, como pode ser observado na Figura 16. Na fase de construção há a troca de mensagens de controle para a seleção dos líderes dos grupos, o anúncio dos líderes para a rede e a entrada de cada nó sensor em um grupo. Na fase estacionária cada nó sensor coleta informação do ambiente periodicamente e transmite essa informação para o líder do grupo, que por sua vez, transmite a informação para a estação base.

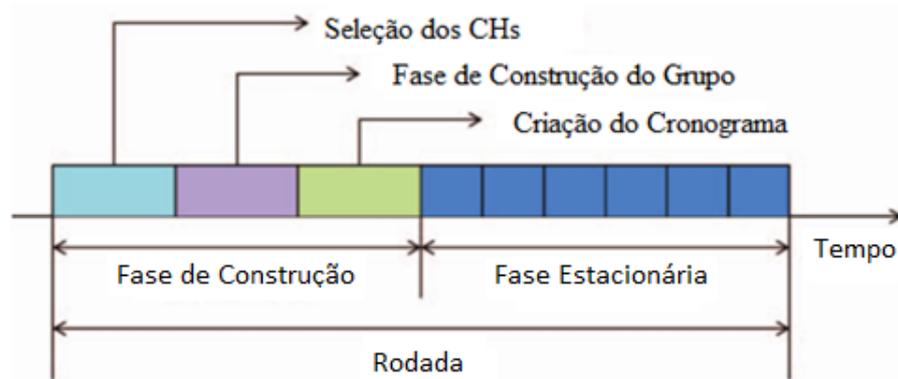


Figura 16 Protocolo de Roteamento Proposto

4.3.1 AGRUPAMENTO INICIAL

Inicialmente usa-se a localização de cada nó sensor como entrada do algoritmo *k-means*. O nó sensor mais próximo do centróide do grupo é selecionado como líder do grupo. A estação base realiza um *broadcasting* para que cada líder fique ciente de sua função na rede, estes por sua vez executam um *broadcasting* para que cada nó sensor possa se associar ao líder mais próximo. Posteriormente, cada nó sensor coleta informação do ambiente, neste caso temperatura e umidade relativa, e envia para o seu respectivo líder, que por sua vez encaminha a informação para a estação base.

4.3.2 REAGRUPAMENTO

Nesta etapa ocorre o agrupamento por meio do algoritmo *k-means*, porém as entradas do algoritmo são: localização de cada nó sensor, temperatura e umidade relativa. Com isso, a seleção do líder de cada grupo ocorre por meio da distância mínima em relação à localização,

temperatura e umidade relativa. Dessa forma, na fase estacionária, não há comunicação de cada nó sensor com seu respectivo líder, economizando a energia consumida na comunicação intragrupo. O líder selecionado, informa aos outros nós sensores a mudança de líder.

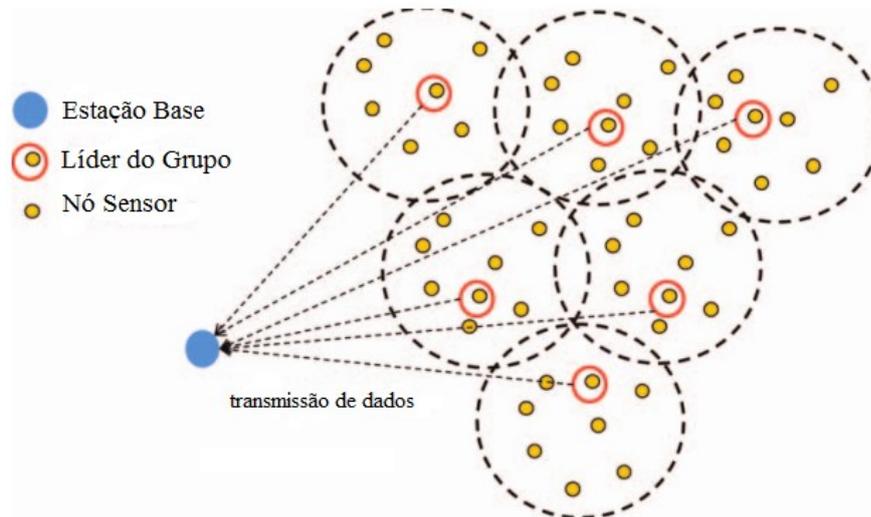


Figura 17 Roteamento single hop intragrupo e intergrupo

Para cada rodada, o esquema proposto adota protocolo de roteamento *single hop*, tanto na comunicação intragrupo quanto na comunicação intergrupo, como pode ser observado na Figura 17. As métricas a serem avaliadas são: tempo de vida da rede e eficiência energética.

Algoritmo de Seleção de CHs

Procedimento

1. Entrada: número de grupos, leituras de temperatura e umidade de cada nó sensor
 2. Formação dos grupos por meio do algoritmo k-means
 3. Líder atual = MinDist()
 4. Todos os nós <- InformMsg()
 5. Enviar dados para a BS
-

O processo de seleção dos CHs está descrito na Algoritmo de seleção de CHs, onde a função `MinDist()` encontra o nó sensor com leitura mais próxima do centróide para cada grupo. A função `InformMsg()` é responsável pelo aviso aos outros nós sensores, sobre a mudança de CH. Depois disso apenas os CHs enviam os dados coletados à BS.

Na Figura 18 pode ser observado o fluxograma do algoritmo completo. Podemos perceber que o algoritmo é composto por duas fases, a fase de controle onde há a formação dos grupos e a seleção dos CHs e a fase estacionária onde os dados dos CHs são encaminhados para a BS.

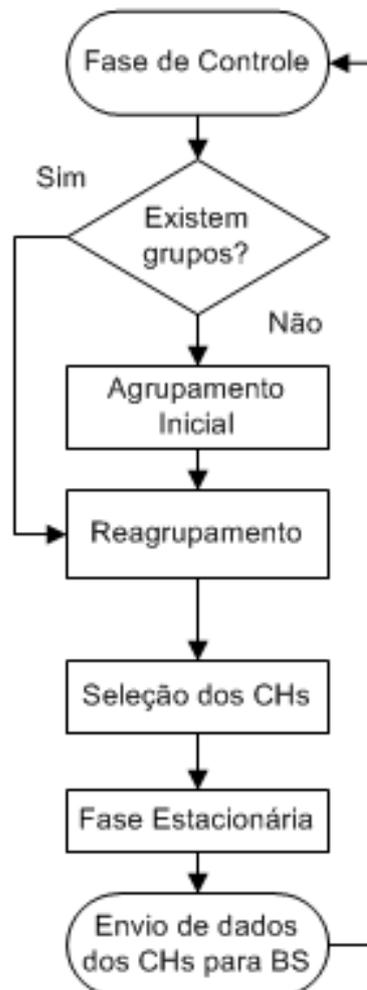


Figura 18 Algoritmo de agrupamento

RESULTADOS

Neste capítulo, apresentamos a avaliação da solução proposta através de experimentos de simulações implementadas em Matlab, e posteriormente, comparamos com outros dois esquemas. Estamos utilizando o conjunto de amostras coletados no experimento realizado no laboratório de Berkeley (Intel Berkeley, 2004). Os protocolos de roteamento selecionados para comparação foram o LEACH, que é descrito a seção 2.1.3 e o proposto no trabalho de Park et. al. (2013).

5.1 CONSUMO DE ENERGIA DA REDE

O tempo de vida da rede é definido como o número de rodadas em que a energia de todos os nós sensores acaba. A simulação foi configurada com os parâmetros da Tabela 2. A Figura 19 ilustra o consumo de energia da rede tanto para os protocolos de comunicação selecionados para comparação quanto o esquema proposto nesta dissertação. Podemos perceber que no esquema proposto neste trabalho, a energia da rede chega ao fim no dobro de rodadas, quando comparado ao LEACH. Dessa forma, o LEACH apresenta desempenho aproximadamente 50% inferior ao esquema proposto. Já o trabalho de Park et. al. (2013) apresenta desempenho 30% inferior ao esquema proposto. Isso ocorre porque estamos explorando não só a correlação espacial mas também a correlação dos dados para formar grupos e escolher os CHs.

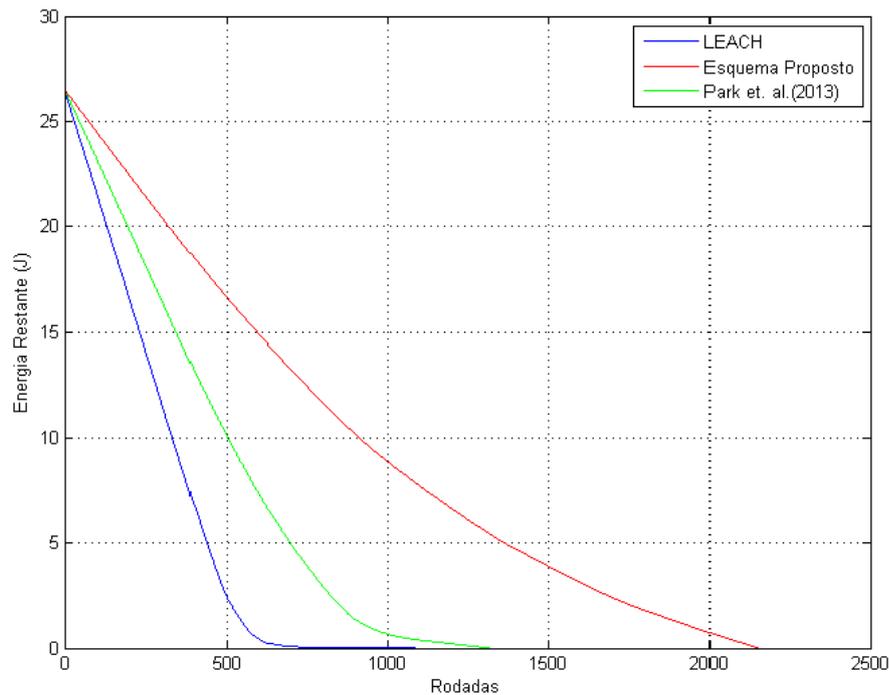


Figura 19 Energia residual restante da rede por rodada

5.2 NÚMERO DE NÓS VIVOS

Em seguida analisamos o número de nós vivos. A Figura 20 ilustra o número de nós vivos por rodada para os dois protocolos selecionados para efeito de comparação bem como para o esquema proposto. Percebe-se que no esquema proposto neste trabalho os nós sensores começam a morrer antes dos outros dois algoritmos de roteamento. Isso se deve a hipótese de seleção do nó sensor que melhor representa cada grupo, enquanto nos outros esquemas há a alternância de CHs para extensão do tempo de vida da rede. Porém, na rodada 1323, a energia de todos os nós sensores já chegou ao fim tanto para o LEACH quanto para o trabalho de Park et. al. (2013), já o esquema proposto neste trabalho ainda possui metade a rede viva. Quando comparado ao algoritmo proposto, o LEACH apresenta desempenho aproximadamente 50%

inferior. Já o trabalho de Park et. al. (2013) apresenta desempenho 30% inferior ao esquema proposto neste trabalho.

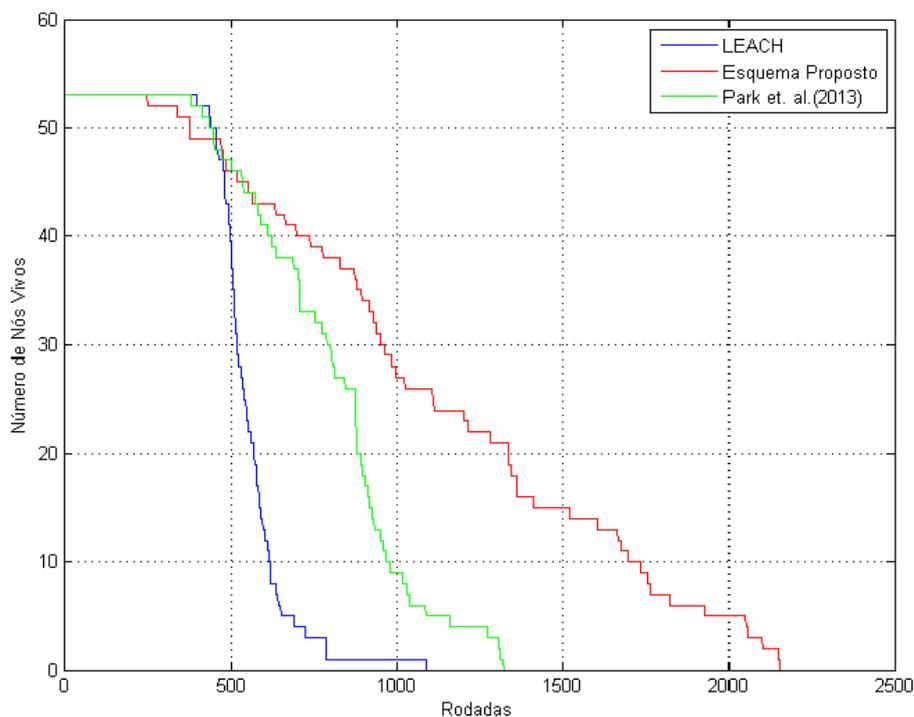


Figura 20 Número de nós vivos por rodada

Observando a rodada 1000 nas Figuras 19 e 20 podemos perceber que em relação a energia residual da rede, tanto o LEACH quanto o trabalho de Park et. al. (2013) praticamente não possuem mais energia, enquanto o esquema proposto neste trabalho ainda possui aproximadamente 8J de energia residual distribuída nos nós sensores da rede. Já em relação ao número de nós vivos, na rodada 1000 o LEACH possui apenas um nó sensor vivo e o proposto por Park et. al. (2013) possui menos que 10 nós sensores vivos, enquanto o esquema proposto neste trabalho ainda possui cerca de 27 nós sensores vivos. Dessa forma mostramos os ganhos possíveis na exploração da estrutura de correlação das variáveis coletadas tanto na formação dos grupos quanto na seleção do nó sensor que melhor representa seu grupo como CH.

CONCLUSÃO

Uma das principais preocupações na aplicação de RSSF, é o tempo de vida da rede, mesmo em aplicações *indoor*, pois o custo de manutenção pode superar rapidamente o próprio custo de implantação do sistema de sensoriamento. Para resolver esse problema, protocolos de roteamento têm sido propostos.

Atualmente, em sua maioria, os esquemas propostos buscam reduzir a distância dos nós sensores para seus respectivos *cluster-heads* como o trabalho de Park et. al. (2013), por meio de algoritmos de aprendizado não supervisionado, como o k-means. Dessa forma o consumo de energia, no envio e recepção de informação de cada nó sensor para o *cluster-head* diminui.

O esquema proposto nesta dissertação busca, não somente reduzir a distância da comunicação intragrupo, mas também eleger o nó sensor que melhor represente cada grupo como CHs. Há portanto, redução da comunicação intragrupo, e conseqüentemente, redução do consumo de energia da rede. Os resultados das simulações mostram que o esquema proposto reduz em até 50% o consumo de energia da rede quando comparado a outros protocolos de roteamento da literatura, como o LEACH. Para trabalhos que representam o estado-da-arte, como o proposto por Park et. al. (2013), os resultados demonstram uma redução no consumo de energia de cerca de 30%. Portanto há um ganho substancial quando se explora a estrutura de correlação das variáveis coletadas na formação dos grupos e na eleição dos CHs.

Este trabalho pode ser continuado e estendido de diferentes formas. Primeiro podemos pensar em um modelo orientado a eventos, onde o objetivo seria a detecção de variações bruscas na temperatura (Figuras 11 e 12) e umidade relativa que influenciariam

diretamente o conforto térmico dos usuários do ambiente fechado. Essa detecção poderia ser feita por meio de filtros implementados nos nós sensores.

Também poderíamos pensar na eleição de mais nós sensores no mesmo grupo, para estudar o compromisso com a precisão dos dados coletados. Assim, poderia ser feita uma análise comparativa dos dados enviados apenas pelos nós eleitos com os dados coletados por toda a rede.

Além disso, outra abordagem possível seria estudar o atraso na entrega dos dados para estação base. Um estudo comparativo em relação aos protocolos da literatura seria necessário. Nesse sentido, alguns trabalhos mostram que o algoritmo k-means, quando implementado na forma distribuída, apresenta menor tempo de agrupamento.

REFERÊNCIAS BIBLIOGRÁFICAS

AKYILDIZ, I. F.; SU, W.; SANKARASUBRAMANIAM, Y.; CYIRCI, E. **Wireless sensor networks: A survey**. *Computer Networks*, 38(4):393—422, 2001

YUN, J.; WON, K. **Building Environment Analysis based on Temperature and Humidity for Smart Energy Systems**. *Sensors*, 2012.

LINA, C.R.; LIN, M. **Adaptive clustering for mobile wireless networks**. *IEEE Journal on Selected Areas in Communications* 15 (7) (1997) 1265–1275.

Intel Berkeley Research Lab, <http://db.lcs.mit.edu/labdata/>, 2004. Acesso em: 10 jun. 2013.

WARD, J.H. **Hierarchical grouping to optimize an objective function**. *Journal of American Statistical Association* 58 (301) (1963) 236–244.

WANG, W.S.; O'KEEFFE, R.; WANG, N.; HAYES, M.; O'FLYNN, B.; O'MATHUMA, C.; **Practical Wireless Sensor Networks Power Consumption Metrics for Building Energy Management Applications**, CORA, 2011.

WANG, W. S.; O'DONNELL, T.; WANG, N.; HAYES, M.; O'FLYNN, B.; O'MATHUNA, C. **Design considerations of sub-mW indoor light energy harvesting for wireless sensor systems**. *ACM J. Emerg. Technol. Comput. Syst*, vol. 6, no. 2, pp. 1-26, 2010

LHERMET, H.; CONDEMINE, C.; PLISSONIER, M.; SALOT, R.; AUDEBERT, P.; ROSSET, M. **Efficient Power Management Circuit: From Thermal Energy Harvesting to Above-IC Microbattery Energy Storage**. *Solid-State Circuits, IEEE Journal of* , vol.43, no.1, pp.246-255, Jan. 2008

MITCHESON, P.D.; YEATMAN, E.M.; RAO, G.K.; HOLMES, A.S.; GREEN, T.C. **Energy Harvesting From Human and Machine Motion for Wireless Electronic Devices**. *Proceedings of the IEEE* , vol.96, no.9, pp.1457-1486, Sept. 2008

CHAMAM, A.; PIERRE, S. **On the Planning of Wireless Sensor Networks: Energy-Efficient Clustering under the Joint Routing and Coverage Constraint**. *Mobile Computing, IEEE Transactions on* , vol.8, no.8, pp.1077-1086, Aug.2009

MARINKOVIC, S.J.; POPOVICI, E.M.; SPAGNOL, C.; FAUL, S.; MARNANE, W.P. **Energy- Efficient Low Duty Cycle MAC Protocol for Wireless Body Area Networks**. *Information Technology in Biomedicine, IEEE Transactions on*, vol.13, no.6, pp.915-925, Nov. 2009

POTTIE, G.J.; KAISER, W.J. **Wireless integrated network sensors**. *Communications of the ACM* 43 (5) (2000) 551–558.

AL-KARAKI, J. N.; KAMAL, A. E. **Routing techniques in wireless sensor networks: a survey**. *IEEE Wireless Communication*, vol. 1, pp. 6–28, 2004.

SEAH, W., TAN, Y. Eds. **Sustainable Wireless Sensor Networks**. InTech Open Access Publisher: Rijeka, Croatia, 2010.

LI, C.; ZHANG, H.X.; HAO, B.B.; LI, J.D. **A survey on routing protocols for large-scale wireless sensor networks**. *Sensors* 2011, 11, 3498–3526.

HAAS, Z.J.; HALPERN, J.Y.; LI, L. **Gossip-Based Ad Hoc Routing**. In Proceedings of the 19th Conference of the IEEE Communications Society (INFOCOM), New York, NY, USA, 23–27 June 2002; pp. 1707–1716.

KULIK, J.; HEINZELMAN, W.R.; BALAKRISHNAN, H. **Negotiation based protocols for disseminating information in wireless sensor networks**. *Wirel. Netw.* 2002, 8, 169–185.

INTANAGONWIWAT, C.; GOVINDAN, R.; ESTRIN, D.; HEIDEMANN, J. **Directed diffusion for wireless sensor networking**. *IEEE/ACM Trans. Netw.* 2003, 11, 2–16.

BRAGINSKY, D.; ESTRIN, D. **Rumor Routing Algorithm for Sensor Networks**. In Proceedings of the First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA), Atlanta, GA, USA, 28 September 2002; pp. 22–31.

KARP, B.; KUNG, H. **GPSR: Greedy Perimeter Stateless Routing for Wireless Networks**. In Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking (MOBICOM), Boston, MA, USA, 6–11 August 2000; pp. 243–254.

NICULESCU, D.; NATH, B. **Trajectory Based Forwarding and Its Applications**. In Proceedings of the Ninth Annual International Conference on Mobile Computing and Networking (MOBICOM), San Diego, CA, USA, 14–19 September 2003; pp. 260–272.

SHAH, R.C.; RABAEY, J.M. **Energy Aware Routing for Low Energy Ad Hoc Sensor Networks**. In Proceedings of the Wireless Communications and Networking Conference (WCNC), Orlando, FL, USA, 17–21 March 2002; pp. 350–355.

SCHURGERS, C.; SRIVASTAVA, M.B. **Energy Efficient Routing in Wireless Sensor Networks**. In Proceedings of Military Communications Conference on Communications for Network-Centric Operations: Creating the Information Force, McLean, VA, USA, 28–31 October 2001; pp. 357–361.

SOHRABI, K.; GAO, J.; AILAWADHI, V.; POTTIE, G.J. **Protocols for self-organization of a wireless sensor network**. *IEEE Pers. Commun.* 2000, 7, 16–27.

HEINZELMAN, W.R.; CHANDRAKASAN, A.; BALAKRISHNAN, H. **Energy-Efficient Communication Protocol for Wireless Microsensor Networks**. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2000; pp. 10–19.

YOUNIS, O.; FAHMY, S. **HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks**. IEEE Trans. Mobile Comput. 2004, 3, 366–379.

DING, P.; HOLLIDAY, J.; CELIK, A. **Distributed Energy Efficient Hierarchical Clustering for Wireless Sensor Networks**. In Proceedings of the 8th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), Marina Del Rey, CA, USA, 8–10 June 2005; pp. 322–339.

BUTTYAN, L.; SCHAFFER, P. **PANEL: Position-Based Aggregator Node Election in Wireless Sensor Networks**. In Proceedings of the 4th IEEE International Conference on Mobile Ad-hoc and Sensor Systems Conference (MASS), Pisa, Italy, 8–11 October 2007; pp. 1–9.

BUTTYAN, L.; SCHAFFER, P. **PANEL: Position-based aggregator node election in wireless sensor networks**. Int. J. Distrib. Sens. Netw. 2010, 2010, 1–16.

LOSCRI, V.; MORABITO, G.; MARANO, S. **A Two-Level Hierarchy for Low-Energy Adaptive Clustering Hierarchy**. In Proceedings of the 2nd IEEE Semiannual Vehicular Technology Conference, Dallas, TX, USA, 25–28 September 2005; pp. 1809–1813.

SORO, S.; HEINZELMAN, W. **Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering**. In Proceedings of the 5th IEEE International Workshop on Algorithms for Wireless, Mobile, Ad Hoc and Sensor Networks (WMAN), Denver, CO, USA, 4–8 April 2005; pp. 236–243.

YE, M.; LI, C.; CHEN, G.; WU, J. **EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks**. In Proceedings of the 24th IEEE International Performance, Computing, and Communications Conference (IPCCC), Phoenix, AZ, USA, 7–9 April 2005; pp. 535–540.

YE, M.; LI, C.; CHEN, G.; WU, J. **An energy efficient clustering scheme in wireless sensor networks**. Ad Hoc Sens. Wirel. Netw. 2006, 3, 99–119.

LI, C.F.; YE, M.; CHEN, G.H.; WU, J. **An Energy-Efficient Unequal Clustering Mechanism for Wireless Sensor Networks**. In Proceedings of the 2nd IEEE International Conference on Mobile Ad-hoc and Sensor Systems Conference (MASS), Washington, DC, 7–10 November 2005; pp. 596–604.

CHAN, H.; PERRIG, A. **ACE: An Emergent Algorithm for Highly Uniform Cluster Formation**. In Proceedings of the 1st European Workshop on Sensor Networks (EWSN), Berlin, Germany, 19–21 January 2004; pp. 154–171.

MURUGUNATHAN, S.D.; MA, D.C.F.; BHASIN, R.I.; FAPAJUWO, A.O. **A Centralized Energy-Efficient Routing Protocol for Wireless Sensor Networks**. IEEE Radio Commun. 2005, 43, S8–S13.

LINDSEY, S.; RAGHAVENDRA, C.; SILVALINGAM, K.M. **Data gathering algorithms in sensor networks using energy metrics**. IEEE Trans. Parallel Distrib. Syst. 2002, 13, 924–935.

MANJESHWAR, A.; AGRAWAL, D. P. **TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks**. In Proceedings of the 15th International Parallel and Distributed Processing Symposium (IPDPS), San Francisco, CA, USA, 23–27 April 2001; pp. 2009–2015.

MANJESHWAR, A.; AGRAWAL, D. P. **APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks**. In Proceedings of the 2nd International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile computing, Lauderdale, FL, USA, 15–19 April 2002; pp. 195–202.

LUO, H.; YE, F.; CHENG, J.; LU, S.; ZHANG, L. **TTDD: Two-tier data dissemination in large-scale wireless sensor networks**. Wirel. Netw. 2005, doi:10.1007/s11276-004-4753-x.

JUNG, S.; HAN, Y.; CHUNG, T. **The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS**. In Proceedings of the 9th International Conference on Advanced Communication Technology, Gangwon-Do, Korea, 12–14 February 2007; pp. 260–265.

KOUTSONIKOLA, D.; DAS, S.; CHARLIE, H.Y.; STOJMENOVIC, I. **Hierarchical geographic multicast routing for wireless sensor networks**. Wirel. Netw. 2010, 16, 449–466.

RAJAGOPALAN, R.; VARSHNEY, P.K. **Data-aggregation techniques in sensor networks: A survey**. IEEE Commun. Surv. Tutor, 2006, 8, 48–63.

YUE, J.; ZHANG, W.; XIAO, W.; TANG, D.; TANGG, J. **Energy efficient and balanced cluster-based data aggregation algorithm for wireless sensor networks**. Procedia Eng. 2012, 19, 2009–2015.

OZDEMIR, S.; XIAO, Y. **Secure data aggregation in wireless sensor networks: A comprehensive overview**. Comput. Netw. 2009, 53, 2022–2037.

LEE, S.H.; LEE, S.; SONG, H.; LEE, H.S. **Gradual Cluster Head Election for High Network Connectivity in Large-Scale Sensor Networks**. In Proceedings of 13th International Conference on Advanced Communication Technology, Phoenix Park, Korea, 13–16 February 2011; pp. 168–172.

CHITNIS, L.; DOBRA, A.; RANKA, S. **Fault tolerant aggregation in heterogeneous sensor networks**. *J. Parallel Distrib. Comput.* 2009, 69, 210–219.

FRERIS, N.M.; KOWSHIK, H.; KUMAR, P.R. **Fundamentals of large sensor networks: Connectivity, capacity, clocks, and computation**. *Proc. IEEE* 2010, 98, 1828–1846.

LI, J.; MOHAPATRA, P. **Analytical modeling and mitigation techniques for the energy hole problem in sensor networks**. *Pervasive Mobile Comput.* 2007, 3, 233–254.

AKKAYA, K.; YOUNIS, M. **A survey on routing protocols for wireless sensor networks**. *Ad Hoc Netw.* 2005, 3, 325–349.

HEINZELMAN, W.B.; CHANDRAKASAN, A.P.; BALAKRISHAN, H. **An application-specific protocol architecture for wireless microsensor networks**. *IEEE Trans. Wirel. Commun.* 2002, 1, 660–670.

FAN, X.; SONG, Y. **Improvement on LEACH Protocol of Wireless Sensor Network**. In *Proceedings of International Conference on Sensor Technologies and Applications*, Valencia, Spain, 14–20 October 2007, pp. 260–264.

XIAOYAN, M. **Study and Design on Cluster Routing Protocols of Wireless Sensor Networks**. Ph.D. Dissertation. Zhejiang University, Hangzhou, China, 2006.

YASSEIN, M.B.; AL-ZOU'BI, A.; KHAMAYSEH, Y.; MARDINI, W. **Improvement on LEACH protocol of wireless sensor network (VLEACH)**. *Int. J. Digit. Content Technol. Appl.* 2009, 3, 132–136.

RAN, G.; ZHANG, H.; GONG, S. **Improving on LEACH protocol of wireless sensor networks using fuzzy logic**. *J. Inf. Comput. Sci.* 2010, 7, 767–775.

ABDULSALAM, H.M.; KAMEL, L.K. **W-LEACH: Weighted Low Energy Adaptive Clustering Hierarchy aggregation algorithm for data streams in wireless sensor networks**. In *Proceedings of IEEE International Conference on Data Mining Workshops (ICDMW)*, Sydney, Australia, 14 December 2010, pp. 1–8.

HONG, J.; KOOK, J.; LEE, S.; KWON, D.; YI, S. **T-LEACH: The method of threshold-based cluster head replacement for wireless sensor networks**. *Inf. Syst. Front.* 2009, 11, 513–521.

LAI, W.K.; FAN, C.S.; LIN, L.Y. **Arranging cluster sizes and transmission ranges for wireless sensor networks**. *Inf. Sci.* 2012, 183, 117–131.

ABBASI, A.A.; YOUNIS, M. **A survey on clustering algorithms for wireless sensor networks**. *Comput. Commun.* 2007, 30, 2826–2841.

MARSLAND, S. **Machine Learning: An Algorithmic Perspective**. Ed. Chapman & Hall 2009.

XU, R.; WUNSCH, D. **Survey of clustering algorithms**. IEEE Transactions on Neural Networks, vol. 16, pp. 645–678, 2005.

CHEN, C.; PAU, L.; WANG, P. **Cluster analysis and related issue**, Handbook of Pattern Recognition and Computer Vision Eds., World Scientific, Singapore, 1993, pp. 3–32. R. Dubes,.

ANSI/ASHRAE Standard 55-2010, **Thermal Environmental Conditions for Human Occupancy**.

DJAMLA, H.; CHU, C., SIVAKUMAR, K. **A Conceptual Review on Residential Thermal Comfort in the Humid Tropics**, International Journal of Engineering Innovation & Research, 2012, ISSN: 2277-5668 (6):539-544.

DJAMLA, H.; CHU, C., SIVAKUMAR, K. **Field Study of Thermal Comfort in Residential Buildings in the Equatorial Hot-Humid climate of Malaya**, Building and Environment & Research 62: 133-142.

TOFTUM, J. **Handbook of Human Factors and Ergonomics Methods**. Boca Raton, 2005, FL, USA: 63.CRC Press.

SMOLANDER, J. **Effect of cold exposure on older humans**. International Journal of Sports Medicine, 2002.

KHODAKARAMI, J. **Achieving thermal comfort**. VDM Verlag, 2009. ISBN: 978-3-639-18292-7.

SZOKOLAY, S. **Introduction to architectural science: the basis of sustainable design**. Architectural Press.

FANGER, P. **Thermal Comfort: Analysis and Applications in Environmental Engineering**. New York: McGraw Hill. 1970

ISO/FDIS 7730:2005, **International Standard, Ergonomics of the thermal environment — Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria**.

MATERSON, J.; RICHARDSON, A. **Humidex, a method of quantifying human discomfort due to excessive heat and humidity**. Environment Canada, p.45, 1979

RANA, R.; KUSY, B.; JURDAK, R.; WALL, J.; HU, W. **Feasibility Analysis of using Humidex as an Indoor Thermal Comfort Predictor**, Energy & Buildings, 2013.

WU, S.; CLEMENTS, D. **Understanding the indoor environment through mining sensory data—A case study**, Energy and Buildings, 2007.

LI, X. H.; FANG, K. L.; HE, L. Z. **A Clustering Algorithm Based on K-means for Wireless Indoor Monitoring System**, International Conference Information Technology and Computer Science, 2009.

MEHR, M. A. **Design and Implementation a New Energy Efficient Clustering Algorithm using Genetic Algorithm for Wireless Sensor Networks**, World Academy of Science, Engineering and Technology, 52, 2011.

PARK, G.; KIM, H.; JEONG, H. W.; YOUN, H. L. **A Novel Cluster Head Selection Method based on K-Means Algorithm for Energy Efficient Wireless Sensor Network**. 27th International Conference on Advanced Information Networking and Applications Workshops, 2013.

SASIKUMAR, P.; KHARA, S., **k-MEANS Clustering in Wireless Sensor Networks**, Fourth International Conference on Computational Intelligence and Communication Networks, 2012.

APÊNDICE

PUBLICAÇÕES

Durante o mestrado, alguns trabalhos e publicações surgiram como fruto de colaboração e esforço de pesquisa como grupo:

- Levy, P.C.; Antonio, N.S.; Souza, T.R.B.; Caetano, R.; Souza, P.G. ActiveIris: Uma solução para comunicação alternativa e autonomia de pessoas com deficiência motora severa. In Proc. of IHC'13 Brazilian Symposium on Human Factors in Computing Systems. October 8-11, 2013, Manaus, AM, Brazil. Copyright 2013 SBC. ISSN 2316-5138 (pendrive). ISBN 978-85-7669-278-2 (online).