



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

ALEXANDRE MÁRCIO MELO DA SILVA

**CONTROLE ENERGETICAMENTE EFICIENTE DE MÚLTIPLOS
SALTOS PARA REDES DE SENSORES SEM FIO HETEROGÊNEAS
UTILIZANDO LÓGICA FUZZY**

**UFPA/ICEN/PPGCC
Belém-Pará-Brasil
2014**

ALEXANDRE MÁRCIO MELO DA SILVA

**CONTROLE ENERGETICAMENTE EFICIENTE DE
MÚLTIPLOS SALTOS PARA REDES DE SENSORES SEM FIO
HETEROGÊNEAS UTILIZANDO LÓGICA FUZZY**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal do Pará - Mestrado Acadêmico—vinculado a Linha de Pesquisa Sistemas Inteligentes do Instituto de Ciências Exatas e Naturais, como exigência para obtenção do grau de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Antonio Morais da Silveira.

**Belém-Pará-Brasil
2014**

Dados Internacionais de Catalogação-na-Publicação (CIP) –

Biblioteca Profª Elcy Rodrigues Lacerda/ Instituto de Ciências da Computação /UFPA, Belém-PA

Silva, Alexandre Márcio Melo

Controle Energeticamente Eficiente de Múltiplos Saltos para Redes de Sensores Sem Fio Heterogêneas Utilizando Lógica Fuzzy / (Alexandre Márcio Melo da Silva); orientador, Antônio Morais da Silveira. – 2014.

57 f. il. 28 cm

Dissertação (Mestrado) – Universidade Federal do Pará. Instituto de Ciências Exatas e Naturais. Programa de Pós-Graduação em Ciência da Computação. Belém, 2014.

1. Sistemas inteligentes. 2. Lógica Fuzzy. 3. Redes de sensores sem fio. I. Silveira, Antonio Morais, orient. II. Universidade Federal do Pará, Instituto de Ciências Exatas e Naturais, Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDD 22.ed. 006.3

ALEXANDRE MÁRCIO MELO DA SILVA

**CONTROLE ENERGETICAMENTE EFICIENTE DE
MÚLTIPLOS SALTOS PARA REDES DE SENSORES SEM FIO
HETEROGÊNEAS UTILIZANDO LÓGICA FUZZY**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal do Pará – Mestrado Acadêmico – vinculado à linha de pesquisa Redes de Computadores, do Instituto de Ciências Exatas e Naturais, como exigência para a obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

Prof. Dr. Antonio Morais da Silveira
Orientador

Prof. Dr. Raimundo Viégas Junior
Membro Interno

Prof. Dr. Jefferson Magalhães de Moraes
Membro Interno

Prof. Dr. Roberto Célio Limão de Oliveira
Membro Externo

Visto:
Prof. Dr. Nelson Cruz Sampaio Neto
(COORDENADOR DO PPGCC/ICEN/UFPA)

CONCEITO: _____

AGRADECIMENTOS

Primeiramente, agradeço a Deus, por saúde e esta vitória.

Agradeço e dedico à minha família, em especial à minha esposa, Adriana Pessôa, que esteve ao meu lado no decorrer de todo esse trabalho com amor, compreensão e suporte psicológico. À minha mãe Dinete Regina, por tudo, infinita são as minhas palavras para agradecer-lhe todo seu amor.

Agradeço a toda minha família. Divaldo, pelo amor paterno; Débora, pelo apoio diário; Diná e Diesel, por sempre estarem ao meu lado em orações. A minha filha Larissa Melo, por compreender os momentos que não ficamos juntos pelo tempo dedicado ao trabalho.

Ao PPGCC - UFPA, agradeço a oportunidade.

Ao GERCOM e LPRAD, agradeço por todo apoio de seus coordenadores e membros, em especial ao André Miranda por indicações de trabalhos.

A todos amigos de pesquisa, em especial ao Christiano por horas de simulações dedicadas, trabalhos de disciplinas e artigos trabalhados em equipe.

Agradeço, em especial, ao meu orientador, Prof. Dr. Antônio Morais da Silveira, pela imensa compreensão e desafio na aceitação da orientação.

MÁQUINA...

*EU POSSO ALCANÇAR A IMORTALIDADE:
BASTA NÃO ME DESGASTAR.*

*VOCÊ TAMBÉM PODE ALCANÇAR A
IMORTALIDADE: BASTA FAZER APENAS
UMA COISA NOTÁVEL*

(JOHNIE WALKER - KEEP WALKING)

RESUMO

O presente trabalho objetiva demonstrar um controle centralizado para eleger Cluster Heads (CHs) mais adequados, admitindo três níveis de heterogeneidade e uma comunicação de múltiplos saltos entre Cluster Heads. O controle centralizado utiliza o algoritmo k-means, responsável pela divisão dos clusters e Lógica Fuzzy para eleição do Cluster Head e seleção da melhor rota de comunicação entre os eleitos. Os resultados indicam que a proposta apresentada oferece grandes vantagens comparado aos algoritmos anteriores de eleição, permitindo selecionar os nós mais adequados para líderes do grupo a cada round com base nos valores do Sistema Fuzzy, como também, a utilização da Lógica Fuzzy como ferramenta de decisão para implementação de múltiplos saltos entre CHs, uma vez que minimiza a dissipação de energia dos CHs selecionados mais afastados do ponto de coleta. A inserção de três níveis de heterogeneidade, correspondente aos sensores normais, avançados e super sensores, contribui consideravelmente para o aumento do período de estabilidade da rede. Outra grande contribuição obtida a partir dos resultados é a utilização de um controle central na estação base (EB) apresentando vantagens sobre o processamento local de informações em cada nó, processo este encontrado nos algoritmos tradicionais para eleição de CHs. A solução proposta comprovou que a eleição do CH mais eficiente, considerando sua localização e discrepâncias de níveis de energia, como também, na inclusão de novos níveis de heterogeneidade, permite aumentar o período de estabilidade da rede, ou seja, o período que a rede é totalmente funcional, aumentando consideravelmente o tempo de vida útil em Redes de Sensores Sem Fio (RSSF) heterogêneas.

Palavras-chave - Redes de Sensores Sem Fio, Lógica Fuzzy, Consumo de Energia, Cluster Head, Múltiplos Saltos.

ABSTRACT

This study presents a centralized control to elect appropriate Cluster Heads (CHs), assuming three levels of heterogeneity and multi-hop communication between Cluster Heads. The centralized control uses the k-means algorithm, responsible for the division of clusters and Fuzzy Logic to elect the Cluster Head and selecting the best route of communication between elected. The results indicate that the proposal offers great advantages, allowing us to select the most suitable nodes for group leaders at each round based on the Fuzzy System values, and also the use of Fuzzy Logic as a decision tool to implement multiple hops between CHs, since it minimizes the power dissipation of the selected CHs more distant from the collection point. The insertion of three levels of heterogeneity, corresponding to normal, advanced and super sensors, contributes considerably to increasing the period of network stability. Another great contribution obtained from the is the use of a central control in base station (BS) with advantages over local information processing in each node, a process usually found in traditional algorithms for electing CHs. The proposed solution proved that the election of the more efficient CH, considering its location and energy levels discrepancies, and also , the inclusion of new heterogeneity levels, allows to increase the network stability period, ie , the period that the network is fully functional, greatly increasing the useful lifetime in heterogeneous WSN

Keywords -Wireless Sensor Networks, Fuzzy Logic, Power Consumption, Cluster Head, Multiple Hops.

LISTA DE TABELAS

Tabela 1- Análise comparativa do K-MEANS e Fuzzy C-Means.....	32
Tabela 2 - Tempo de complexidade do K-MEANS e FCM com variação do numero de clusters.....	32
Tabela 3- Base de regras do sistema fuzzy.....	41
Tabela 4 - Parâmetros do Modelo de Rádio	46
Tabela 5 - Resultados da Seleção de Cluster Heads.....	48
Tabela 6 - Saída do Sistema Fuzzy.....	49

LISTA DE ABREVIATURAS E SIGLAS

ACHTHLEACH	Adaptive Cluster Head Election and Two-hop
LEACH ACK	Low-Energy Adaptive Clustering Hierarchy Acknowledgement
CH	Cluster Head
CABCF	Clustering algorithm based on communication facility
CF	Communication facility
DEEC	Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks
EDEEC	Enhanced Distributed Energy Efficient Clustering Scheme
EF-TREE	Earliest-First Tree
FND	Fist Node Dies
HERF	Hybrid Energy Efficient Routing using a Fuzzy Method in Wireless Sensor Networks
ISM	Industrial, Scientific and Medical
JPL	Jet Propulsion Laboratory
LEACH	Low Energy Adaptative Clustering Hierarchy
MANETS	Mobile Ad Hoc Networks
MIPS	Milhões de Instruções por Segundo
MEMS	Micro Electro Mecanical System
RSSF	Redes de Sensores Sem Fio
SID	Source-Initiated Dissemination
SEP	Stable Election Protocol for clustered heterogeneous wireless sensor networks
WINS	Wireless Integrated Network Sensor
WSN	Wireless Sensor Network

SUMÁRIO

1. INTRODUÇÃO	14
1.1. Motivação	14
1.2. Objetivos	16
1.2.1. Objetivo Geral	16
1.2.2. Objetivos Específicos	16
1.3. Contribuições	17
1.4. Organização da Dissertação	17
2. Trabalhos relacionados	18
3. Referencial Teórico.....	20
3.1. Redes de sensores sem fio	20
3.1.1. Contextualização a RSSF	20
3.1.2. Ponto de Coleta em RSSF.....	22
3.1.3. Energia em RSSF.....	22
3.2. Lógica Fuzzy	23
3.2.1. Contextualização a Lógica Fuzzy	23
3.2.2. Teoria de Conjuntos Fuzzy	24
3.2.3. Estrutura de um Sistema Fuzzy	27
3.3. Algoritmo K-Means	30
3.3.1. Contextualização ao Algoritmo K-Means	30
3.3.2. Aplicação do Algoritmo	30
3.3.3. Comparativo dos algoritmos K-Means e Fuzzy C-Means.....	32
3.3.4. Processo de Agrupamento K-means	33
3.3.5. Limitações no Processo de Agrupamento.....	35
4. Modelagem Utilizando Sistema Fuzzy e Algoritmo K-Means Aplicado ao Processo de SEleção do Cluster Head.....	36
4.1. Cenário para modelagem	36
4.2. Modelo de dissipação de energia	37

4.3. NÍVEIS de HETEROGENIDADE	37
4.4. Critérios para eleição do CH.....	38
4.5. Centralidade dos nós e distância para A EB	40
4.6. Múltiplos Saltos entre CHs.....	41
4.7. Modelo de Sistema Fuzzy.....	44
4.8. Modelagem da rede.....	44
4.9. Propriedades da rede.....	45
5. SIMULAÇÃO E SELEÇÃO DO CLUSTER HEAD IDEAL	46
6. CONCLUSÕES E TRABALHOS FUTUROS.....	52
REFERÊNCIAS	54
APÊNDICE	56
apêndice A - Publicação Relacionada à Dissertação	57
TÍTULO: Controle energeticamente eficiente de múltiplos saltos para Redes de Sensores sem Fio heterogêneas utilizando LógicaFuzzy	57

1. INTRODUÇÃO

1.1. MOTIVAÇÃO

Estudos referentes ao consumo de energia em Redes de Sensores Sem Fio (RSSF) são importantes para o melhor estado de funcionamento dos sensores envolvidos e de extrema importância para auxiliar no prolongamento do seu período de vida útil, necessitando da programação de ações preventivas e do monitoramento destas redes. Visto que são aplicadas em muitos cenários, cada um com características únicas em termos de conectividade, localização e tempo de vida útil. É necessário avaliar as especificidades de um ambiente de destino, geralmente complexo, o que leva à importância de estudos baseados na restrição de energia, visto que o consumo de energia é associado a várias atividades dos sensores, como a quantidade de dados coletados, posicionamento na rede, distância para os outros sensores e a comunicação entre eles, atividade estacada em termos de energia, o que implica em propostas de algoritmos para auxiliar em uma escolha preliminar no posicionamento destes sensores.

Representando uma subclasse das redes ad hoc sem fio, as RSSF são consideradas como uma nova geração de sistemas embarcados de tempo real com recursos computacionais, energia e memória limitados, tendo na restrição de energia um dos principais entraves apresentados devido à capacidade limitada das baterias internas dos nós sensores. Vários fatores são culminantes para o desgaste da bateria dos nós, sendo o módulo de rádio um dos principais consumidores de energia dos nós sensores no processo de transmissão de dados (POTTIE *et. al.*, 2000).

O consumo de energia pode ser reduzido, admitindo que apenas alguns nós possam enviar dados para a Estação Base (EB). RSSF hierárquicas organizam seus nós em agrupamentos (*clusters*) e elegem um nó líder do grupo, denominado cluster head (CH). O CH é responsável por coletar todos os dados dos nós de seu cluster, informações provenientes de sensoriamento, podendo agregá-los e posteriormente encaminhá-los à Estação Base (EB) (AKYILDIZ, 2002).

A estrutura hierárquica pode ser formada por dois tipos de redes, homogêneas ou heterogêneas. Em estruturas heterogêneas, tratado nesta proposta, alguns nós sensores podem apresentar requisitos de hardware diferenciados, como melhor capacidade energética. Estes nós sensores dão à rede um maior período de estabilidade (Saini *et. al.*, 2010; QUING *et. al.* 2006). Trabalhos que consideram a heterogeneidade dos nós podem ser encontrados em

(AKYILDIZ, 2002; QUING *et. al.*, 2006; MUBARAK *et. al.*, 2011; HAN, 2010; SMARAGDAKIS *et. al.*, 2004).

O algoritmo LEACH (*Low Energy Adaptive Clustering Hierarchy*) proposto por Mubarak *et. al.* (2011) é a referência de algoritmo hierárquico para formação de clusters mais citada encontrada no estado da arte e serve como base para vários artigos científicos que discutem melhorias para formação de clusters.

Igualmente ao algoritmo LEACH, o maior problema encontrado nos algoritmos que o tem como base é a utilização de informações locais com base em cálculos de probabilidade para escolha dos nós sensores principais, citados na literatura como *Clusters Heads*. Desta forma, este tipo de seleção leva a escolha dos líderes que encontram-se em posições não favoráveis ao consumo de energia, visto que os eleitos mais próximos da borda da rede, aumentam a dissipação de energia devido à distância de transmissão na comunicação dos outros sensores envolvidos para o CH. Assim como, o problema da proximidade dos CH da borda da rede, outros pontos foram identificados e tratados neste trabalho, com referência a forma de eleição utilizada pelo algoritmo LEACH. Um deles é a falta de tratamento discriminatório sobre as discrepâncias energéticas dos nós que constituem a rede, uma vez que os CHs eleitos devem possuir recursos energéticos suficientes para suportar as cargas de transmissão dos nós associados a ele, visto a perda e o consumo maior de energia. Outro problema encontrado, refere-se ao ponto de coleta dos nós sensores, sendo centralizados na EB, onde são executadas a função colaborativa de coleta e processamento de dados o que tendem ao maior consumo de energia envolvido. Segundo Loureiro (2003), os consumidores de energia são os modelos de rádio, processador e elementos que fazem o sensoriamento do ambiente.

Dentro das funcionalidades exercidas pelas redes de sensores sem fio – RSSF (RUIZ, 2002), neste trabalho está principalmente relacionado ao estabelecimento da Rede, onde são tratados alguns pontos como a distribuição dos nós sensores, a ativação dos nós pertencentes a rede, as dimensões envolvidas, densidades, tipos de sensores, área de cobertura, organização, topologia e conectividade (LOUREIRO, 2002), podendo ser interna e externa a rede.

Considerando os entraves relacionados à eleição de CHs com base em informações locais sem considerar critérios de posicionamento, este trabalho propõe uma estratégia para eleição do CH mais adequado em RSSF heterogêneas, utilizando a Lógica Fuzzy com base em informações centralizadas na EB. Este controle centralizado define o CH com base em informações adquiridas no momento de formação da rede. As informações coletadas são utilizadas para carregar o algoritmo k -means, responsável pela divisão dos clusters, como

também, alimentar o sistema fuzzy , que se encarrega de selecionar o líder de cada grupo formado pelo algoritmo k-means.

O sistema fuzzy também é responsável por determinar quais CHs passarão pelo processo de encaminhamento de dados dos CHs mais afastados. Nós eleitos que ultrapassam o limiar de comunicação com a EB, devem enviar seus dados para os nós mais próximos . Para esta seleção o Sistema Fuzzy utiliza como critérios : distância na comunicação e níveis de energia. Os critérios adotados para seleção de CHs são : nível de energia , centralidade e proximidade da EB. Como especificado em (MACIEL, 2013), a inserção de três níveis de heterogeneidade, a utilização de informações centralizadas na Estação Base (EB) e a comunicação através de múltiplos saltos , permitem eleger CHs bem posicionados e com níveis adequados de energia para suportar a carga de transmissão do referido cluster, aumentando o período de estabilidade e vida útil da rede.

1.2. OBJETIVOS

1.2.1. Objetivo Geral

Propor e analisar uma solução de controle centralizado no ponto de coleta para eleger *Cluster Heads* mais adequados, admitindo heterogeneidade quanto ao nível de energia e uma comunicação de múltiplos saltos para melhor eficiência energética em uma rede de sensores sem fio.

1.2.2. Objetivos Específicos

- Identificar e eleger com uso do algoritmo *k-means* e Lógica Fuzzy o *Cluster Head* e a melhor rota de comunicação entre os nós eleitos na rede;
- Aplicar a solução do sistema *fuzzy* como ferramenta de decisão para implementação de múltiplos saltos entre CHs;
- Avaliar a solução proposta, por meio de simulação dos resultados obtidos neste trabalho e, apresentar as melhorias agregadas em relação aos algoritmos usados para eleição do CH.
- Contribuir para o debate sobre o referido tema.

1.3. CONTRIBUIÇÕES

Dentre as contribuições que a solução proposta apresenta, pode-se destacar a diferenciação das abordagens propostas nos algoritmos tradicionais do referido tema e melhores resultados obtidos nas simulações em comparação com os algoritmos vistos no estado da arte, podendo-se destacar:

- Uso da abordagem *fuzzy* para tratar as incertezas inerentes ao processo de escolha dos Nós sensores;
- Uso do algoritmo *k-means* no processo de formação dos *clusters*, permitindo minimizar as distâncias entre os nós participantes.
- A possibilidade de uma simples implementação do algoritmo com a utilização do sistema *fuzzy*, buscando otimizar o consumo de energia dos sensores.

1.4. ORGANIZAÇÃO DA DISSERTAÇÃO

A presente dissertação está organizada em cinco capítulos, com os assuntos assim distribuídos:

Capítulo 1, Introdução, contextualização do problema investigado, uma visão geral da dissertação e os objetivos do trabalho. No **Capítulo 2** são apresentadas as técnicas utilizadas como referencial teórico, compreendendo uma introdução às Redes de Sensores sem Fio, Lógica Fuzzy, e Algoritmo K-Means. O **Capítulo 3** apresenta a modelagem aplicada ao processo de eleição do CH mais adequado, empregada para alcançar os objetivos do trabalho. O **Capítulo 4** apresenta a simulação e a eleição do referido CH na análise dos gráficos gerados e comparação de desempenho com outros algoritmos tradicionais. No **Capítulo 5** são apresentadas as conclusões da dissertação e as propostas de continuidade da pesquisa com possíveis trabalhos futuros.

Finalizando a dissertação temos o **Apêndice A** onde são apresentados os trabalhos publicados pelo autor durante o desenvolvimento de mestrado e no **Apêndice B** são apresentados os códigos de linguagem de programação gerados nas simulações e os dados de entrada e saída na ferramenta matlab.

2. TRABALHOS RELACIONADOS

Conforme já citado, o fornecimento de energia para as RSSF é o principal entrave para projetos referente ao tipo de organização da rede como o proposto neste trabalho. Selecionar o CH apropriado pode reduzir significativamente o consumo de energia dos nós sensores e aumentar a vida útil da rede. Este capítulo apresenta o estado da arte sobre o assunto explorado neste trabalho. Desse modo, são apresentados os principais algoritmos de seleção de CH, para este tipo de rede sob a ótica da eficácia.

Um dos algoritmos mais citados na literatura é o LEACH (*Low Energy Adaptative Clustering Hierarchy*), proposto por Mubarak *et. al.* (2011). Elege o CH com base em informações locais a cada novo ciclo (round). Para que o nó seja eleito, o número escolhido deve ser menor que o limiar T . Após a eleição, o novo CH envia mensagens de anúncio para todos os nodos da rede. Os demais nós associados decidem a que CH devem se conectar.

Embora o LEACH apresente uma estrutura hierárquica que permite reduzir o consumo de energia em RSSF, o mesmo não adota nenhum critério de posicionamento no momento de eleição do CH, podendo selecioná-lo próximo a borda da rede (MACIEL, 2012). Outro problema do algoritmo é descrito no trabalho apresentado em Smaragdakis *et. al.*, (2004), SEP (*Stable Election Protocol*), comprovando que o LEACH não é eficiente em estruturas heterogêneas, por não considerar a discrepância de energia dos nós que formam a rede.

O DEEC (*Distributed energy efficient clustering*) proposto por Quing *et. al.* (2006) utiliza informações locais para eleição do CH, entretanto, é capaz de tratar a heterogeneidade da rede, característica considerada na solução proposta neste trabalho, sendo que a heterogeneidade inserida pelos autores diz respeito apenas à capacidade energética diferenciada de um conjunto de nós que formam a rede.

Baseado no algoritmo DEEC, os algoritmos E-DEEC (*Enhanced Distributed Energy Efficient Clustering*) proposto por Saini *et. al.* (2010) e LEACH-HPR proposto por Han (2010), também consideram a energia residual dos nós no processo de eleição para o CH da rede. O diferencial está na inserção de três tipos de nós com diferentes níveis de energia: nós normais, nós avançados e super nós, permitindo prolongar o período totalmente funcional da rede com a adição dos super nós. Utilizam informações locais para eleição do CH. Para minimizar a dissipação de energia na fase de comunicação entre os CHs mais afastados e a EB. Os autores de LEACH-HPR propõem um algoritmo de múltiplos saltos entre os CHs

eleitos, semelhante ao ACHTLEACH (Adaptive Cluster Head Election and Two-hop LEACH) proposto por Saini *et. al.* (2010)

Em Yan *et. al.* (2010), os autores propõem um algoritmo que considera uma estratégia híbrida móvel com o objetivo de distribuir o consumo de energia por toda a rede , com o CH movimentando-se para um local de maior concentração de energia quando ocorre um evento e com possibilidade de controlar sua potência de transmissão . Entretanto, este deslocamento pode maximizar o consumo de energia gasto no envio de dados para a BS . A estratégia móvel híbrida do algoritmo BS -CH permite a movimentação da estação base para minimizar a distância para o CH e o consumo de energia na transmissão de dados . Desconsidera a energia consumida, supõe que a estação base é móvel e sem informação se a fonte de alimentação é contínua.

3. REFERENCIAL TEÓRICO

Para a consolidação do método e técnicas utilizadas neste trabalho foram realizadas pesquisas, em fontes seguras, tais como artigos científicos, livros e pesquisas de bases conceituais, com diversas vertentes, que serviram para a determinação do referencial teórico.

Foram verificados estudos envolvendo técnicas da Inteligência Computacional tais como: Lógica Fuzzy e Algoritmo *k-means*, e a validação através de técnicas de simulação aplicadas aos fatores que influenciam no processo de avaliação de Redes de Sensores Sem Fio Heterogêneas, constatando-se ainda ser uma linha de pesquisa muito discutida na academia. Por isso, propostas como a apresentada, inter-relacionando esses domínios do conhecimento podem ser extremamente importantes no processo de escolha do melhor sensor entre inúmeros de uma rede em estudo, tornando eficaz energeticamente o processo de formação de um grupo heterogêneo com seu respectivo líder. Desse modo, este trabalho é inovador, uma vez que aborda em sua proposta, os fatores da heterogeneidade, centralidade dos nós sensores, distância para a estação base e políticas que permitem evitar a escolha de nós com proximidade a borda da rede, que estão relacionados com as influências que podem ocorrer em relação à ineficiência energética da rede. Pontos não abordados em conjunto em análise de outros algoritmos propostos como solução para o problema.

As técnicas da Inteligência Computacional investigadas, mostraram-se alternativas viáveis para apoio ao processo de tomada de decisão em uma rede de sensores heterogênea.

3.1. REDES DE SENSORES SEM FIO

3.1.1. Contextualização a RSSF

Área de sensores está em constante avanço tecnológico. Sensores são usados em várias aplicações, como por exemplo, no monitoramento e controle das condições ambientais em diversos cenários, como florestas e rios. Rastreamento, roteamento e processamento em diferentes contextos, também são exemplos de suas aplicações (LOUREIRO et al., 2003). A necessidade de interconexão de sensores (nós), exige a construção de redes que os interliguem para prover comunicação entre os mesmos, formando as conhecidas Redes de sensores sem fio (RSSF), compostas de um grande número de nós de sensores, com a finalidade de executar uma tarefa de sensoriamento maior (YOUNIS, 2006).

Os nós sensores possuem dimensões reduzidas, o que causa restrições no *hardware* e limitações para cada particularidade da aplicação usada, além da exigência de processamento

e roteamento para prover comunicação na rede. O conjunto destas restrições, diferem as RSSF das redes sem fio em diversos aspectos e invalidam o uso de protocolos desenvolvidos para as tradicionais redes ad hoc sem fio, que não possuem limitações tão severas de energia (AKYILDIZ, 2002). Por outro lado, a eficiência energética das RSSF é um problema delicado e objeto de vários estudos na área. Além da restrição de energia, as RSSF também devem possuir mecanismos para autoconfiguração e adaptação, devido a ocorrência de possíveis problemas, como falhas de comunicação e perda de nodos. Uma RSSF tende a ser autônoma e requer um alto grau de cooperação para executar as tarefas definidas para a rede. Isto significa que algoritmos distribuídos tradicionais, como protocolos de comunicação e eleição de líder (Cluster -Head), devem ser revistos para esse tipo de ambiente antes de serem usados diretamente (LUIZ, 2005). A formação de clusters, a relação dos nós (*Node*) aos líderes (*Leader 1, 2, n*) de cada cluster e a estação base (EB) são mostrados respectivamente na figura 1 (SONG, 2013), onde a EB está representada pelo *sink*.

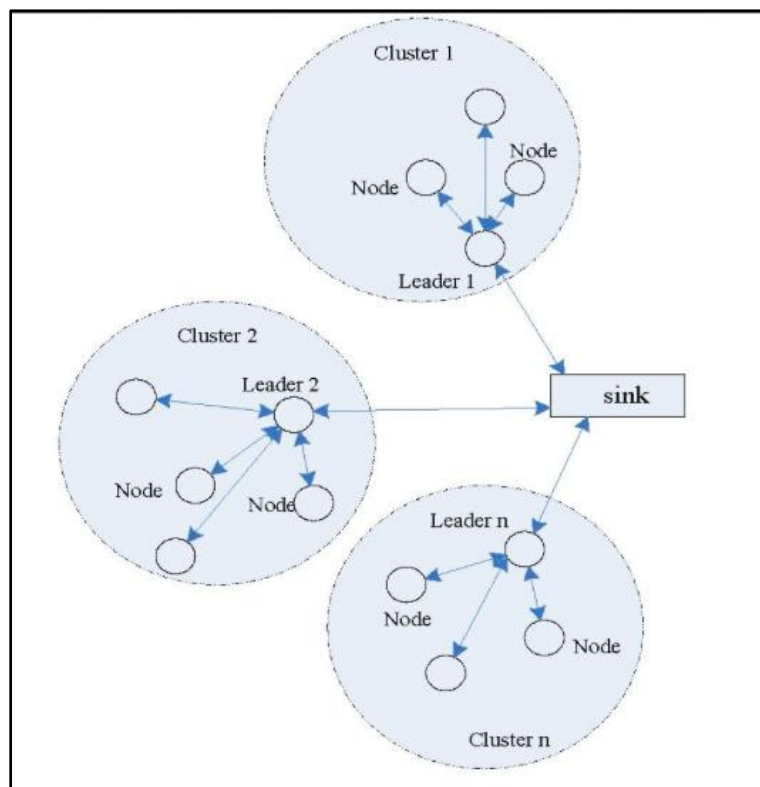


Figura 1 - Modelo de formação em RSSF. Fonte (SONG, 2013)

Como visto anteriormente nos cenários para RSSF, o tempo de vida da rede depende da quantidade de energia disponível nos nós sensores, sendo que estes devem balancear seus recursos limitados com o objetivo de aumentar o tempo de vida total da rede (MINI, 2012). Desta forma, a conservação de energia é um dos aspectos mais importantes a ser considerado no projeto das RSSF, assunto do trabalho proposto.

3.1.2. Ponto de Coleta em RSSF

O conhecimento de abordagens para o ponto de coleta requer uma avaliação caso-a-caso, de modo a respeitar os requisitos da aplicação utilizada nas redes de sensores. São discutidos vários pontos dentre os quais se pode citar: energia concentrada nos pontos de coletas, como na EB; distância deste ponto de coleta para o cluster formado; nível de energia concentrada no ponto de coleta e nos nós pertencentes ao cluster etc...

O ponto de coleta está relacionado à agregação dos dados. Essa característica indica a capacidade de uma RSSF de agregar ou sumarizar dados coletados pelos sensores. Caso a rede tenha essa funcionalidade, é possível reduzir o número de mensagens que precisam ser transmitidas por ela. Assim, os dados coletados podem ser combinados e sumarizados ainda na rede, antes de serem enviados à estação base.

Os modelos de sensores encontrados na literatura revelam que sua capacidade de detecção diminui com o aumento da distância do sensor à EB e cresce com o aumento do tempo de coleta de informações (MEGERIAN, 2002).

No presente trabalho referido, o ponto de coleta está concentrado na EB, responsável pelo recebimento, processamento e reenvio dos dados referentes à RSSF em questão.

3.1.3. Energia em RSSF

Como descrito anteriormente, os nós de uma RSSF possuem limitações quanto aos recursos, como baixa reserva de energia, capacidade computacional e memória, uma vez que algumas aplicações destas redes se localizam em regiões inóspitas, com pouca ou nenhuma manutenção, o que torna o tempo de vida útil da rede dependente do nível de energia usado pelos *clusters* formados. Neste cenário, torna-se necessário o estudo da melhor forma de balancear seus recursos limitados com o objetivo de aumentar o tempo de vida da rede.

O conhecimento sobre a quantidade de energia disponível em cada parte da rede é chamado de mapa de energia, assunto não tratado no escopo deste trabalho, mas vale ressaltar, como um dos exemplos de informação que pode auxiliar a prolongar o tempo de vida da rede (LOUREIRO, 2003). De posse do mapa de energia é possível determinar se alguma parte da rede está na eminência de falhar devido à falta de energia.

Outro ponto importante para o auxílio no tempo de vida da rede é a escolha da melhor localização para o nó principal, citado em algumas literaturas como nó sorvedouro (cluster head). Citado neste trabalho como o nó de maior capacidade de energia que faz comunicação

com o ponto de coleta referenciado neste trabalho como BS, sendo este responsável pela comunicação com a área externa da rede. Conseqüentemente, se o *cluster head* estiver localizado em áreas apropriadas para comunicação, otimizará a quantidade de energia disponível, tornando possível que o tempo de vida da rede seja prolongado.

Protocolos de roteamento também podem beneficiar-se da informação da melhor localização do cluster head. Um algoritmo de roteamento pode fazer um melhor uso das reservas de energia se este seletivamente escolher rotas que utilizam nodos com maior quantidade de energia disponível de tal forma que partes da rede com poucas reservas de energia possam ser preservadas (LOUREIRO, 2003). Portanto, as soluções propostas para otimização da vida útil das RSSF levam em consideração a localização dos nós sensores.

3.2. LÓGICA FUZZY

3.2.1. Contextualização a Lógica Fuzzy

Proposta por Zadeh (1965), a Lógica Fuzzy utiliza métodos com o objetivo de controlar a linguagem vaga e a imprecisão utilizada diretamente pelo homem, por meio de um conjunto de valores representados por variáveis linguísticas. Cada conjunto de valores tem um intervalo diretamente associado a regras semânticas. Ao contrário da Lógica booleana onde uma proposição pode compreender apenas valores exatos e excludentes como 0 e 1, verdadeiro e falso, na lógica fuzzy a avaliação de uma determinada proposição pode compreender valores e graus de pertinência que variam no intervalo de [0,1].

A lógica fuzzy pode mapear um espaço de entradas em um espaço de saída; diversos tipos de sistemas fazem processamentos desse tipo, como as redes neurais, as equações diferenciais entre outros. No entanto, segundo Zadeh, a lógica fuzzy pode realizar esse processamento com um menor custo e maior rapidez (JANG e SUN, 1995).

O controle por meio da lógica fuzzy visa gerar o espaço de saída por meio de processamento baseado em regras, com entradas imprecisas, ruidosas ou mesmo incompletas (WANG, 1997). Dessa forma, esse tipo de abordagem é adequado para tratar problemas que normalmente necessitam da experiência de especialistas humanos, que por natureza desenvolvem um raciocínio aproximativo em situações complexas e imprecisas; para isso a lógica fuzzy faz uso de linguagem natural (TANSCHKEIT, 2002).

3.2.2. Teoria de Conjuntos Fuzzy

3.2.2.1. Conjunto Fuzzy

Para Jang e Gulley (1997) um conjunto fuzzy A é útil para tratar situações onde os conjuntos não têm limites bem definidos, de modo que não podem ser tratadas pela lógica de conjuntos clássica. Assumem assim, valores que variam dentro de determinado intervalo conhecido como universo do discurso U , ao invés de apenas valores excludentes. Assim, uma entrada x assume um grau de pertinência $\mu(x)$ dentro do conjunto, sendo representado por: $A = \{(x, \mu(x) \mid x \in U)\}$.

Assim, os conjuntos fuzzy indicam qualquer afirmação pode ser fuzzy. A ferramenta que o raciocínio distorcido dá a capacidade para responder a uma pergunta sim-não com uma resposta não-muito, sim-ou-não, algo comum para seres humanos, mas ainda novo para computadores.

3.2.2.2. Variáveis Linguísticas

Segundo Gonçalves (2007), uma variável linguística é uma variável cujos valores são nomes de conjuntos fuzzy. A ideia de capacitar de modo impreciso as variáveis de um problema em termos de conceitos qualitativos ao invés de quantitativos, traduz a ideia de uma variável linguística (SOUZA, 2003).

As variáveis linguísticas podem ser definidas também como uma extensão das variáveis numéricas, pois elas são transformadas em conjuntos fuzzy com seus respectivos valores (WANG, 1997).

3.2.2.3. Funções de pertinência

Para Jang e Sun (1995) a função de pertinência (FP) é uma curva que define como cada ponto no espaço de entrada é mapeado para um valor de adesão (ou grau de associação) entre 0 e 1. Embora as funções de pertinência sirvam para representar situações imprecisas, elas são funções matematicamente precisas utilizadas para defuzzificar o mundo.

Os valores da função de pertinência devem variar entre 0 e 1. A função em si pode ser uma curva cuja forma arbitrária em função da conveniência, do ponto de vista da simplicidade, conveniência, velocidade e eficiência.

A seleção da função de pertinência mais adequada ao problema pode ser feita com o auxílio de especialistas ou através de dados coletados por sensores.

Para Miranda (2003), o número de funções de pertinência de conjuntos fuzzy e a forma escolhida dependem de alguns fatores, tais como: precisão requerida, a capacidade de dar a resposta desejada e a estabilidade do sistema, facilidade de implementação, manipulação, manutenção, e assim por diante. As funções de pertinência trapezoidal e triangular são as mais comumente utilizadas. Estas funções de pertinência de linha reta têm a vantagem da simplicidade.

3.2.2.4. Operadores lógicos (ou, E, não)

O raciocínio aproximativo fuzzy permite a utilização das operações lógicas clássicas como as operações AND (intersecção), OR (união) e NOT (negação), desde que mantidos os valores extremos do conjunto. Porém, para entender o resultados dos testes clássicos, já que os valores fuzzy são representados por graus de pertinência entre 0 e 1, pode-se utilizar a função $\min(A, B)$ como equivalente a operação A .AND. B, onde A e B são limitados ao intervalo (0,1). Usando o mesmo raciocínio, pode-se substituir a operação OR com a função $\max(A, B)$, de modo a que A .OR. B torna-se equivalente a $\max(A, B)$. Finalmente, a operação NOT torna-se equivalente à função $1 - A$.

A equação (14) representa a intersecção de dois conjuntos fuzzy A e B a função T agregado graus de adesão como segue.

$$\mu_{A \cap B}(x) = T(\mu_A(x), \mu_B(x)) = \mu_A(x) \otimes \mu_B(x) \quad (14)$$

onde \otimes é o operador binário para a função T. Esse operador é referido como operador T-norma (norma triangular), satisfaz os seguintes requisitos:

Elemento neutro: $T(0, 0) = 0, T(a, 1) = T(1, a) = a$

Monotonicidade: $T(a, b) \leq T(c, d)$ if $a \leq c$ and $b \leq d$

Comutatividade: $T(a, b) = T(b, a)$

Associatividade: $T(a, T(b, c)) = T(T(a, b), c)$

O primeiro requisito impõe a generalização correta de conjuntos nítidos. O segundo requisito implica que uma diminuição nos valores de adesão em A ou B não pode produzir um aumento no valor de adesão a intersecção A.B. O terceiro requisito indica que o operador é indiferente a ordem dos conjuntos difusos a serem combinados. Finalmente, o quarto requisito nos

permite tomar a intersecção de qualquer número de conjuntos em qualquer ordem de agrupamentos de pares.

Como na intersecção fuzzy, o operador de união fuzzy é especificado em geral por uma função S de acordo com a equação (15):

$$\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x)) = \mu_A(x) \oplus \mu_B(x) \quad (15)$$

onde \oplus é o operador binário para a função S . Esse operador geralmente é referido como operador T-conorma (ou S-norma), satisfaz os seguintes requisitos:

Elemento neutro: $S(1, 1) = 1, S(a, 0) = S(0, a) = a$

Monotonicidade: $S(a, b) \leq S(c, d)$ if $a \leq c$ and $b \leq d$

Comutatividade: $S(a, b) = S(b, a)$

Associatividade: $S(a, S(b, c)) = S(S(a, b), c)$

A justificativa destes requisitos básicos é semelhante a dos requisitos para o operador T-norma.

3.2.2.4. Regras Fuzzy

O sistema a ser tratado é representado por regras do tipo Se – Então. A parte Se representa os antecedentes e a parte Então representa os consequentes. O antecedente corresponde aos graus de pertinência calculados durante o processo de fuzzificação.

O número de regras que compõem a base de regras de um sistema é formado combinando-se os conjuntos fuzzy gerados pelas variáveis de entrada.

A regra fuzzyse-então assume a forma.

Se x é A então y é B

onde A e B são os valores linguísticos definidos por conjuntos fuzzy sobre os intervalos (universos de discurso) X e Y , respectivamente.

Assim, o antecedente ou entrada para uma regra é um valor real x entre 0 e 1 relacionado a um conjunto A , e o consequente ou conclusão é um conjunto fuzzy B .

O antecedente de uma regra pode ter múltiplas partes, assim como o consequente:

se x é A e y é B , então z é C e w é D

3.2.2.5. Método de implicação

A conclusão de uma regra fuzzy atribuído um conjunto fuzzy para a saída. Se o antecedente é apenas parcialmente verdadeiro, então o conjunto fuzzy de saída é truncado de acordo como método de implicação.

Duas ou mais regras são necessárias para ativar determinado conjunto fuzzy. Para gerar um número real na saída os conjuntos fuzzy de saída para cada regra são agregados em um único conjunto fuzzy de saída que depois é defuzzificado para um número crisp.

3.2.3. Estrutura de um Sistema Fuzzy

A Figura 2 apresenta a estrutura genérica de um sistema de inferência fuzzy, sendo composta, basicamente por entrada e saída, fuzzificador de entradas, base de regras, defuzzificador de saídas.

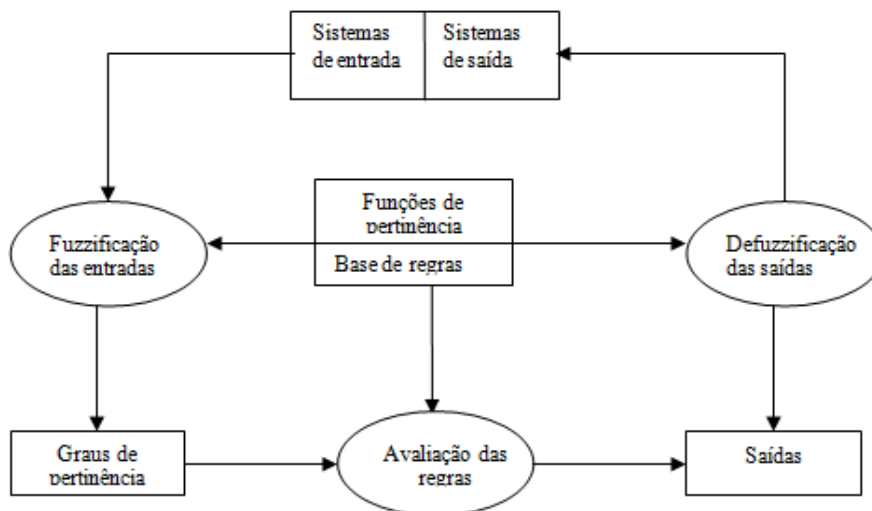


Figura 2 - Organização de um sistema fuzzy

3.2.3.1. Processo de Inferência Fuzzy

Para Jang e Sun (1995) inferência fuzzy é o processo de mapeamento a partir de uma dada entrada a uma saída usando a lógica fuzzy, envolvendo funções de pertinência, operadores de lógica fuzzy, e regras se-então.

A figura 3 apresenta simplificada o fluxo de informações no sistema:

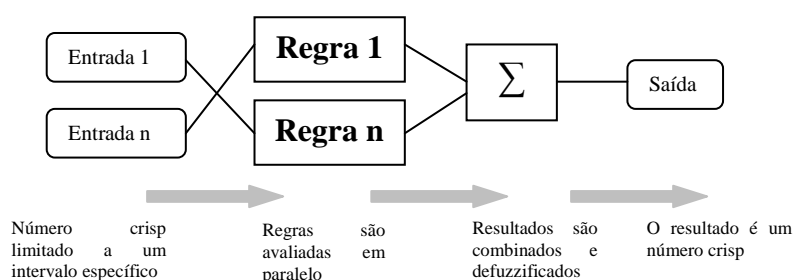


Figura 3 - Fluxo de informações em um sistema fuzzy

3.2.3.2. Fuzzificação das entradas

A fuzzificação das entradas é o processo de associação de um valor numérico de entrada x a um determinado conjunto fuzzy, definindo o seu grau de pertinência ao conjunto através de funções de pertinência. A valor numérico x é limitado ao universo de discurso da variável de entrada e a saída um grau de adesão fuzzy entre 0 e 1.

3.2.3.3. Aplicação do Operador Fuzzy

A fuzzificação das entradas fornece o grau dos antecedentes para cada regra. Se o antecedente de uma determinada regra tem mais de uma parte, o operador fuzzy é aplicado para obter um único número que representa o resultado do antecedente para aquela regra. Este número será então aplicado à função de saída. A entrada para o operador fuzzy são dois ou mais valores de pertinência das variáveis de entrada fuzzificadas. A saída é um valor verdade único.

3.2.3.4. Aplicação do método de implicação

A implicação é a formação de um conseqüente, ou seja, a obtenção de um conjunto fuzzy de saída, conjunto truncado com base no valor único obtido na fase de aplicação do operador fuzzy descrita anteriormente e é realizada para cada uma das regras que formam a base (JANG e GULLEY, 1997).

3.2.3.5. Agregação das saídas

A agregação é o processo de combinar todos os conjuntos truncados resultantes da aplicação do método de implicação de cada regra, gerando um conjunto único para cada variável de saída; esse conjunto será usado no processo de defuzzificação da saída.

3.1.3.6. Defuzzificação

A defuzzificação é o processo de obtenção de um número crisp (real) a partir do conjunto resultante do processo de agregação. Um dos métodos mais utilizados para realizar esse processo é o cálculo do centróide, que retorna o centro de gravidade da área do conjunto fuzzy de saída, conforme a Figura 4. Existem outros métodos como: bissetriz, média do valor máximo e maior dos máximos, e menor dos máximos, etc.

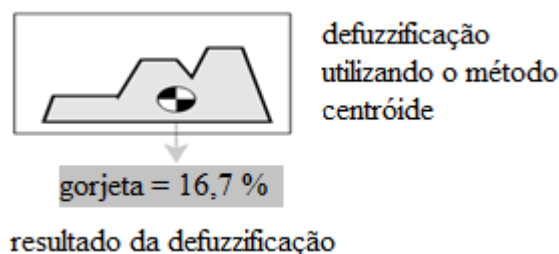


Figura 4 - Exemplo de defuzzificação utilizando o método centróide

3.3. ALGORITMO K-MEANS

O algoritmo k-means é uma técnica de agrupamento de dados por k -médias muito popular por sua facilidade de implementação . Normalmente os algoritmos de clusterização são amplamente utilizados em aplicações que necessitem gerar padrões , dividindo os objetos em grupos úteis ou significativos (SMARAGADAKIS *et. al.*, 2004). Para a proposta apresentada neste trabalho, o algoritmo k-means é carregado com as coordenadas de todos os nós que formam a rede . Desta forma, o algoritmo gera agrupamentos de nós, formando clusters por meio das proximidades dos nós.

3.3.1. Contextualização ao Algoritmo K-Means

O conceito de agrupamento visto neste trabalho, como formação de *Clusters*, foi introduzido a bastante tempo mas, ainda continua sendo amplamente utilizado em soluções nos diferentes domínios do conhecimento. *Clusterização* é um processo em que são classificados objetos que são de uma forma ou de outra comuns em propriedades dentro de um mesmo cenário. O principal objetivo do Cluster é oferecer uma combinação de registros semelhantes. O termo classificação e agrupamento é confuso, mas eles têm uma diferença de que na classificação de objetos aloca em classes pré-definidas, enquanto nas classes de agrupamento ele é criado. Tendo como exemplo a gestão de banco de dados, cluster é um processo onde, as informações armazenadas fisicamente são semelhantes às informações lógicas.

Inúmeras propostas usando grafos, algoritmos genéticos, etc. tem sido apresentadas, buscando otimizar a formação de clusters. Sendo que quando simulado algumas destas técnicas tornam a implementação impraticável (MURTY, 1999). Foi abordado na pesquisa os mais utilizados em RSSF, sendo um deles o algoritmo K-means, método desenvolvido por Dunn (1973) e melhorado por Bezdek (1981).

3.3.2. Aplicação do Algoritmo

A solução do K-means se dá com os seguintes passos, primeiramente elementos de dados K são selecionados como centros iniciais, então a distância de todos os elementos de dados é deliberada pela fórmula da distância euclidiana. Os elementos de dados com

menos distância dos centróides são estimulados ao grupo apropriado. O processo é repetido até não haver mais alteração. A Figura 5 mostra os passos do algoritmo K-means (WANG, 2011).

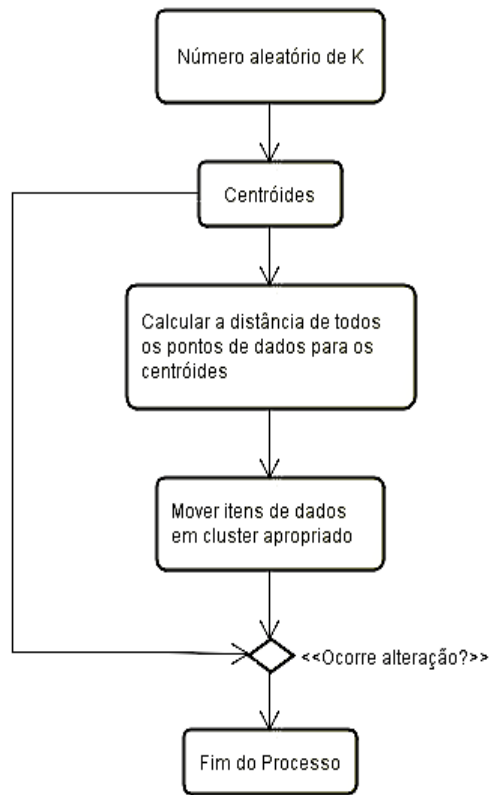


Figura 5–Processo de agrupamento do algoritmo K-means (AGRAWAL, 2013)

Estudos revelam a existência de dois problemas relacionados ao uso desse algoritmo. Primeiro seus resultados dependem muito das posições iniciais dos centros de agrupamento (LOZANO, 1999), cujo resultado está preso às soluções ótimas locais, e o segundo problema é que apenas agrupamentos linearmente separáveis podem ser descobertos por este algoritmo. Em vista destes problemas, na presente proposta utilizamos um sistema fuzzy para alcançar a solução final de clusterização.

Realizou-se um extensivo estudo bibliográfico de variações do K-means, como o caso do algoritmo GKM. Segundo Agrawal (2013), foram feitos alguns trabalhos anteriores para aumentar a eficiência na formação de clusters, tendo como vantagem desta proposta a melhoria na forma de criar o próximo cluster, mais centralizado em relação aos já formados. Definido uma nova função para selecionar o candidato mais adequado para o conjunto seguinte. Porém, tem o problema de maior espaço de tempo entre a escolha do primeiro nó

eleito e o próximo, outro problema é referente ao armazenamento das informações paracalcularesta eleição, que requer grandes conjuntos de dados exigindo grande capacidade computacional, o que inviabilizou seu uso nesta solução, pelo fato de sensores não possuírem a capacidade computacional e energética exigida.

3.3.3. Comparativo dos algoritmos K-Means e Fuzzy C-Means

Para escolha do melhor algoritmo de formação de cluster, considerando a baixa complexidade computacional e o tempo de resposta, pontos imprescindíveis em RSSF, comparou-se os dois principais algoritmos que atendem esses requisitos, os algoritmos K-Means e Fuzzy C-Means (FCM).

Os algoritmos K-Means e Fuzzy C-Means (FCM), também têm problemas com o espaço de memória enúmero desconhecido de etapas de iteração que são necessários para formação dos clusters. A complexidade de tempo do algoritmo K-means é $O(ncdi)$ e do FCM é $O(ndc^2i)$, analisados a partir dos resultados obtidos em (SOUMI, 2013), onde mantendo o número de pontos de dados constante, podemos assumir que $n = 100, d = 3, i = 20$ e do número de aglomerados variando em que $n =$ número de pontos de dados, $c =$ número de cluster, $d =$ número de dimensão e $i =$ Número de iterações. As tabelas 1 e 2 e a figura 6 apresentam a comparação em detalhes.

Tabela 1- Análise comparativa do K-MEANS e Fuzzy C-Means

Algoritmo	Tempo de Complexidade	Tempo decorrido (Segundos)
K-Means	$O(ncdi)$	0.443755
Fuzzy C-Means	$O(ndc^2i)$	0.781679

Tabela 2 - Tempo de complexidade do K-MEANS e FCM com variação do numero de clusters

S.No.	Numero de clusters	K-Means tempo de complexidade	FCM tempo de complexidade
1	1	6000	6000
2	2	12000	24000
3	3	18000	54000
4	4	24000	96000

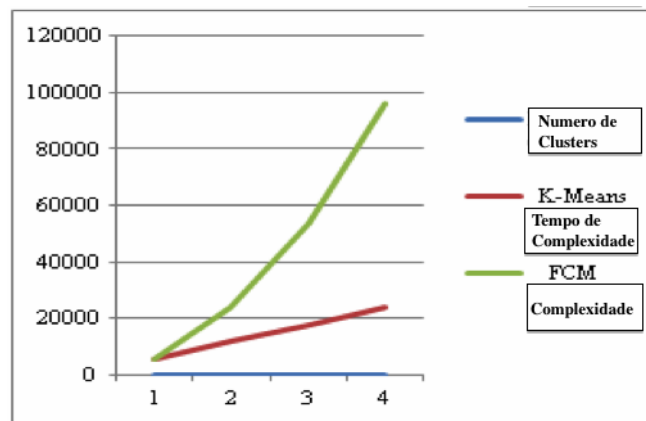


Figura 6 - Tempo de complexidade do K-Means e FCM por variação de números de clusters(SOUMI, 2013)

Pode-se concluir que o algoritmo K-Means é melhor do que o algoritmo FCM nas medidas esperadas para esta solução. O FCM produz resultados aproximados ao K-Means, porém requer mais tempo de computação do que o K-Means devido aos cálculos de medidas fuzzy incluídos no algoritmo. Assim, a conclusão pela escolha do algoritmo K-Means neste trabalho, é que o mesmo mostrou-se superior ao algoritmo fuzzy C-Means, além de ser uma técnica de agrupamento de dados muito popular por sua facilidade de implementação.

Normalmente, os algoritmos de clusterização são muito utilizados em aplicações que necessitam gerar padrões, dividindo os objetos em grupos úteis ou significativos (SMARAGADAKIS *et. al.*, 2004), como o caso da proposta deste trabalho, onde o algoritmo k-means é carregado com as coordenadas de todos os nós que foram à rede, gerando os clusters por meio do critério da proximidade dos nós.

3.3.4. Processo de Agrupamento K-means

O processo de agrupamento K-means, segue um modo simples e fácil de classificar um determinado conjunto de dados num certo número K de agrupamentos. A priori, a ideia principal é definir k centroides (ponto central do Cluster), um para cada cluster. Estes centroides devem ser colocados de uma forma inteligente, por causa de diferentes localizações, o que faz com que o resultado seja diferente a cada processamento. Portanto, a melhor escolha é colocá-los tanto quanto possível longe um do outro. O próximo passo, é levar cada ponto pertencente a um determinado conjunto de dados e associá-la ao centroide mais próximo. Quando nenhum ponto está pendente, a primeira etapa está concluída e um

agrupamento prévio é feito. Neste ponto, precisamos recalculamos um novo K-centroide como possíveis centros dos clusters resultantes da etapa anterior. Depois temos esses novos k-centroides, e uma nova ligação deve ser feita entre os dados dos mesmos pontos de ajuste e o novo centroide mais próximo. Com isso, um ciclo foi gerado. Como resultado deste ciclo, podemos notar que os k-centroides mudam a sua localização passo a passo até que não haja mais mudanças a serem feitas.

Finalmente, este algoritmo visa minimizar uma *função objetivo*, neste caso, uma função de erro quadrado. A função objetivo

$$J = \sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - c_j|^2, \text{ onde } |x_i^{(j)} - c_j|^2 \text{ é uma medida da distância entre}$$

um ponto escolhido de dados $x_i^{(j)}$ e o agrupamento c_j . É um indicador da distância entre os n pontos de dados dos seus respectivos centros de fragmentação.

Nas Figuras 7.a, 7.b e 8 são apresentadas as etapas de atribuição dos nós na rede e o modo de atualização para formação dos *clusters* em relação ao *cluster head*, respectivamente.

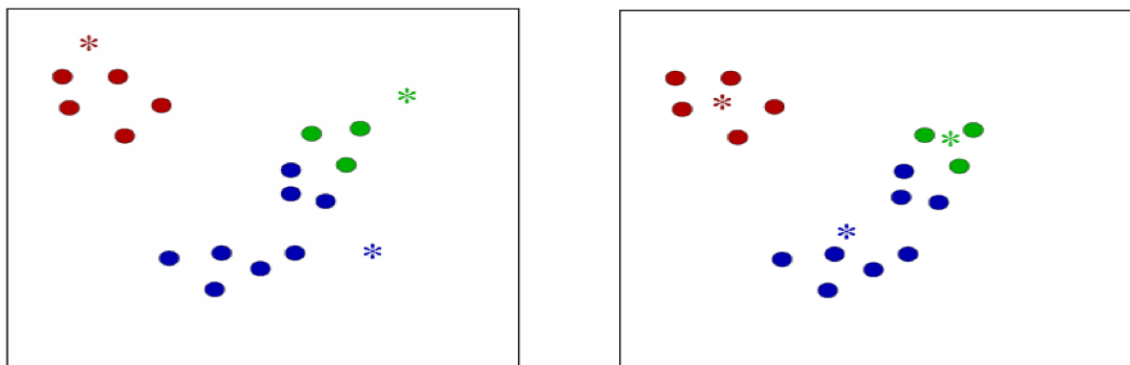


Figura 7 - a) Atribuição ao representante mais próximo b) Reapresentação da Atribuição

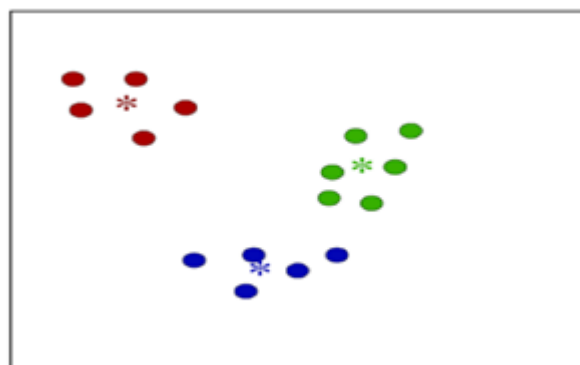


Figura 8 - Convergência após outra iteração

3.3.5. Limitações no Processo de Agrupamento

É importante citar as limitações do algoritmo K-means. As figuras 9.a e 9.b mostram limitações no agrupamento com diferentes densidades, onde a associação falha quando existe um centróide com baixa densidade para nós mais potentes energeticamente. Neste caso, determinamos a suficiência energética do nó central tratados nos próximos capítulos como super nós (sensores mais avançados), em cada *round* (ciclo de formação do cluster), dividindo a formação do *cluster* em rounds. O que permite a utilização deste algoritmo onde cada nó só pode pertencer a um grupo, visto que o K-Means faz um agrupamento *hard* (*Hard clustering*) (TAN, 2006).

Outro problema pode ocorrer durante o cálculo de similaridade e atribuição, podendo não ter nenhum nó associado, desta forma gerando um *cluster* vazio, devendo substituir seu respectivo centróide, caso a substituição não ocorra é necessário a eliminação do *cluster* (TAN 2006).

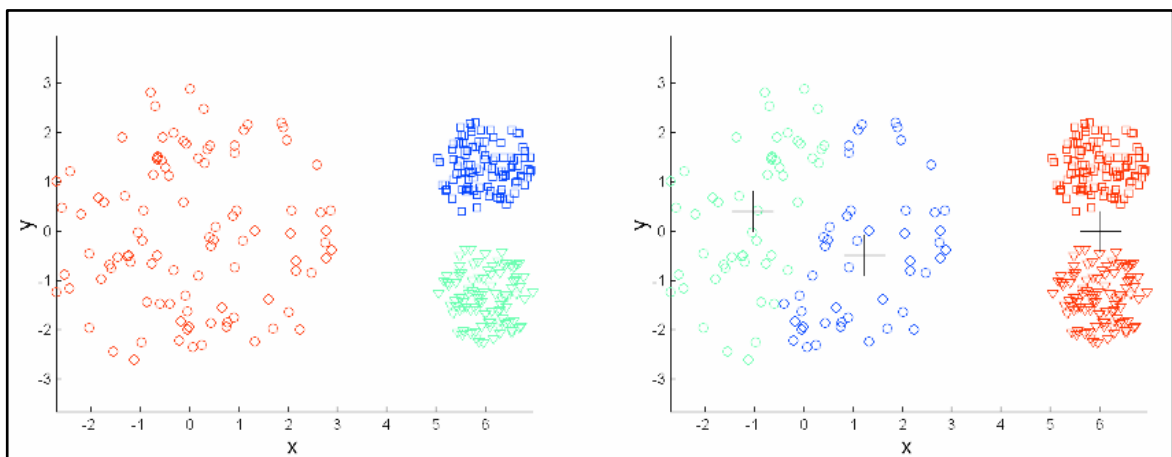


Figura 9 - a) Formação original

b). K-means (3 Clusters)

As técnicas da Inteligência Computacional investigadas mostraram-se alternativas viáveis para apoio ao processo de tomada de decisão em uma rede de sensores heterogênea.

4. MODELAGEM UTILIZANDO SISTEMA FUZZY E ALGORITMO K-MEANS APLICADO AO PROCESSO DE SELEÇÃO DO CLUSTER HEAD

Um dos objetivos deste trabalho é identificar e selecionar com uso do algoritmo K-means e LF o CH ideal no processo de comunicação entre os nós pertencentes a rede em cada *round*. Nesse processo, é levado em conta que após a formação da rede, os nós enviam seus dados, sua localização atual e respectivos níveis de energia para o ponto de coleta no caso, a EB. A cada round, os dados atualizados são usados como critério para carregar o sistema, que podem influenciar o processo de seleção dos nós e a formação da rede como um todo. O sistema fuzzy proposto é utilizado na escolha do CH levando em conta o grau de influência na localização e nível de energia (variáveis nível de bateria e centralidade do *cluster*) para o posicionamento dos nós na rede.

4.1. CENÁRIO PARA MODELAGEM

Na solução proposta, a escolha do *cluster head* se dá a cada round. Um round termina no final do processo de agregação e envio de dados para a EB. Basicamente o processo é dividido em três etapas: (i) A primeira consiste no início do processo de formação dos clusters. A divisão dos clusters é feita utilizando o algoritmo K-means; (ii) A segunda etapa é dividida em duas fases. A primeira fase consiste na seleção do CH para cada cluster formado pelo K-means. Os dados de cada cluster são carregados no Sistema Fuzzy, e este, baseado nos critérios adotados seleciona os CHs mais adequados para cada cluster.

Na segunda fase, após o resultado da seleção de líderes, a EB calcula a distância dos CHs que ultrapassaram o limiar de comunicação para todos os Cluster Heads selecionados. Este cálculo é utilizado para definir para qual líder o CH mais afastado deve transmitir seus dados já agregados. Após a conclusão das duas fases, a estação base envia mensagens em broadcast para os nós da rede informando o ID do líder do grupo para que os nós possam enviar dados para seu respectivo CH; (iii) A terceira etapa compreende o processo de agregação dos dados pelo CH. Este processo consiste em comprimir os dados e enviá-los a EB.

No processo de associação dos nós de um determinado cluster para seu respectivo CH, os nós que formam o cluster recebem a mensagem de anúncio da BS informando a que líder deve enviar pedidos de associação. O líder reserva um slot TDMA (*Time Division Multiple*

Access) para cada nó associado , para que possam transmitir seus dados . Os líderes responsáveis por propagar para o ponto de coleta os dados já agregados de CHs mais afastados, recebem da EB, juntamente com a mensagem de anúncio , o pedido de reserva de um slot TDMA para este processo.

4.2. MODELO DE DISSIPACÃO DE ENERGIA

O modelo de dissipação de energia adotado é semelhante ao modelo utilizado pelo LEACH. O modelo consiste na energia dissipada na transmissão e recepção de k -bit de mensagem em uma distância d, o radio consome:

$$E_{Tx}(k, d) = E_{elec} * k + \varepsilon_{fs} * k * d^2 \text{ para } d < d_0 \quad (1)$$

$$E_{Tx}(k, d) = E_{elec} * k + \varepsilon_{amp} * k * d^4 \text{ para } d \geq d_0 \quad (2)$$

$$E_{Rx}(k, d) = E_{elec} * k \quad (3)$$

A energia dissipada na transmissão e recepção do radio é representada por $E_{elec} = 50 \text{ nJ/bit}$. (eq. 1). Dois modelos são utilizados , *Free space*(eq. 2) e *Multipath* (eq. 3), para que o amplificador de transmissão alcance um nível aceitável , dependendo da distância entre o transmissor e o receptor . Se esta distância não ultrapassar o limiar d_0 , o modelo de *Free space* é utilizado . A energia dissipada pelo amplificador de transmissão é dada por $E_{fs} = 10 \text{ pJ/bit/m}^2$, caso o limiar seja ultrapassado o modelo *Multipath* é utilizado $E_{amp} = 0.0013 \text{ pJ/bit/m}^4$. Onde $d_0 = \sqrt{E_{fs}/E_{amp}}$, (MOUBARAK, 2011; SAINI, 2010).

O modelo de *Multipath* gera maior dissipação de energia no processo de comunicação. Estratégias energeticamente eficientes devem considerar oposicionamento dos nós no processo de seleção de CHs para que o modelo de *Free space* seja mantido.

4.3. NÍVEIS DE HETEROGENIDADE

Na proposta apresentada, a RSSF é formada com três níveis de heterogeneidade. A heterogeneidade apresentada neste trabalho se dá a níveis de energia. As discrepâncias energéticas dos nós sensores são tratadas de forma discriminatória, uma vez que os LIDERES

são selecionados com base em sua energia residual. Para os nós com maiores níveis energéticos, maior a possibilidade de serem selecionados.

A centralidade para o centro do cluster e proximidade com o ponto de coleta são pontos importantes que também são utilizados no processo de seleção. O objetivo é minimizar o consumo de energia gerado na comunicação entre os nós que fazem parte de um determinado *cluster* e seu líder. Para evitar que o limiar de comunicação seja ultrapassado, na fase de propagação de dados, é implementada uma política de múltiplos saltos entre *cluster heads* selecionados para rodada atual, uma vez que além de sobrecarregados devido a quantidade de transmissões de seu cluster, o líder deve tentar minimizar o consumo de energia na fase final, de propagação de dados para o ponto de coleta.

Desta forma, a proposta apresentada garante a seleção adequada de líderes dando a rede um maior período de estabilidade. O acréscimo no período de estabilidade permite um aumento considerável no tempo de vida útil da RSSF.

4.4. CRITÉRIOS PARA ELEIÇÃO DO CH

Os critérios utilizados para carregar o sistema fuzzy são : (i) Energia – O nível de energia de cada nó da rede é representado pela variável Bateria e possui os termos linguísticos Baixa, Moderada e Alta. Consideramos que a rede possui características heterogêneas no que diz respeito à energia dos nós. Logo, os super nós e nós avançados representam maior chance para seleção do CH , Figura 1. (ii) Centralidade – A variável centralidade diz respeito ao posicionamento do nó em relação ao centro do cluster e , possui os seguintes termos linguísticos: Perto, Moderado e Longe. Quanto menor o valor de centralidade mais próximo o nó está do centro do cluster, Figuras 10 e 11.

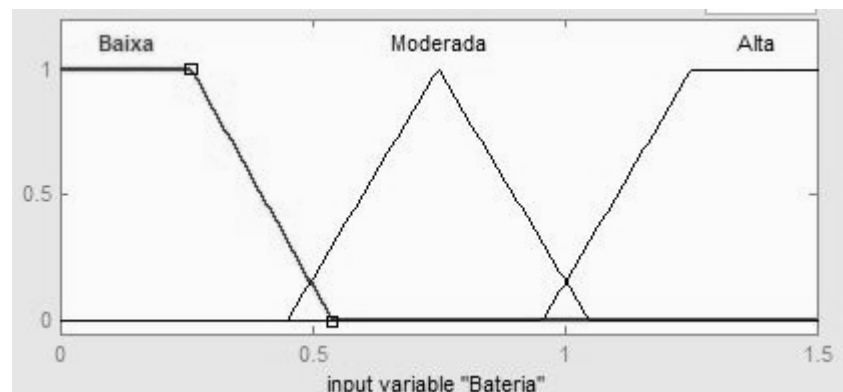


Figura 10 - Variável bateria com os valores linguísticos correspondentes aos níveis de energia dos três tipos de nós sensores disponíveis na rede

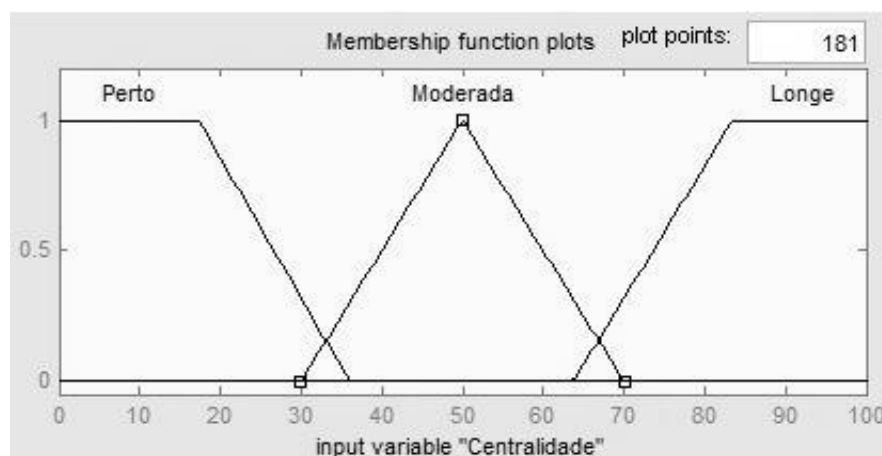


Figura 11 - Variável linguística centralidade

(iii) Distância para a BS – A variável é representada no sistema como DistBS, com seu universo de discurso mapeado pelos seguintes termos linguísticos: Perto, Moderado e Longe. O critério de distância para a BS é utilizado para gerar uma aproximação do CH com a BS, objetivando minimizar o consumo de energia do CH na fase de transmissão de dados, Figura 12.

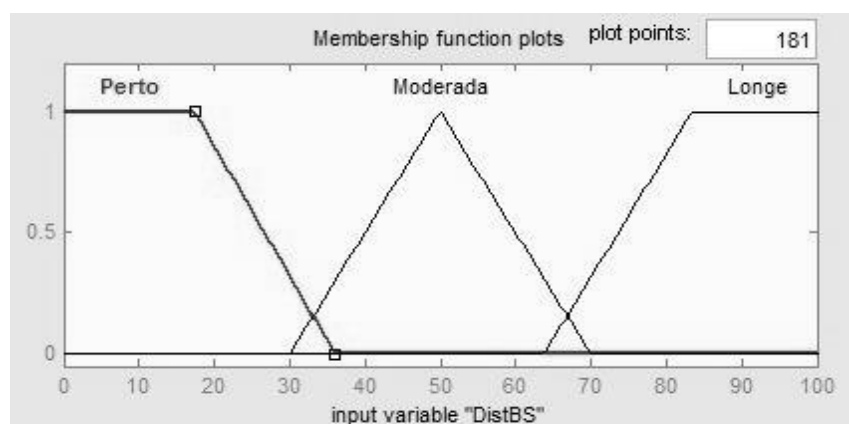


Figura 12 - Distância para BS

As variáveis descritas acima correspondem às variáveis de entrada do sistema. A cada valor de entrada, é mapeado para um conjunto fuzzy de entrada com seu determinado grau de pertinência e, os consequentes conjuntos fuzzy de saída do sistema determinam o CH ideal em um cluster x , usando os termos linguísticos: muito fraco, fraco, médio, forte e muito forte da variável de saída Eleição-CH de acordo com o especificado na Figura. 13.

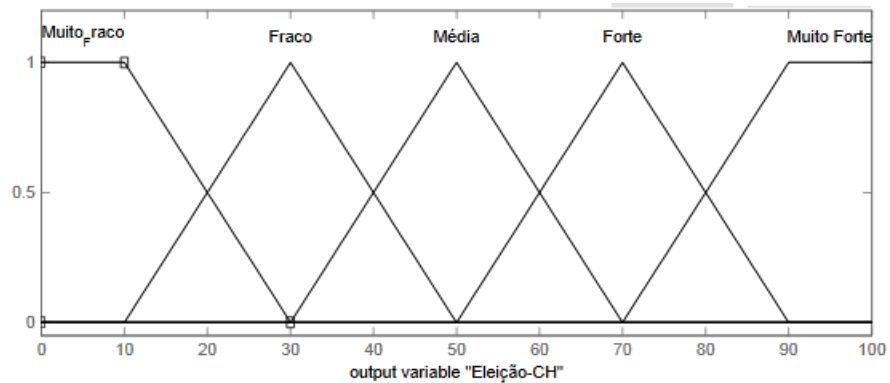


Figura 13 - Saída do Sistema Fuzzy

No processo de mapeamento dos conjuntos fuzzy de saída para valores *crisp* (defuzzificação), cada nó sensor apresenta seu respectivo valor de saída do sistema fuzzy. Os maiores valores de saída indicam os cluster heads selecionados para o round atual. A Base de Regras do sistema é composta por vinte e sete regras, conforme especificado na Tabela 3. A melhor condição para a eleição do cluster head é dada pela seguinte regra: Se Centralidade é PERTO e Bateria é ALTA e DistBS é PERTO então Eleição cluster head é MUITO FORTE.

4.5. CENTRALIDADE DOS NÓS E DISTÂNCIA PARA A EB

Assume-se que a BS detém o conhecimento do nível de energia e posicionamento dos nós. Estas informações são enviadas no início de formação da rede e os nós dissipam energia neste processo. Para determinar os valores de centralidade, a BS seleciona cada nó e calcula a distância euclidiana destes nós para o centro dos seus respectivos clusters, definidos pelo algoritmo k-means, Eq. (4). O nó que apresentar maior centralidade, como CH, permitirá que a dissipação de energia na comunicação ocorra de forma mais homogênea em relação aos demais nós associados.

$$d(P_i, P_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4)$$

Tabela 3- Base de regras do sistema fuzzy

Regras	Entradas			Saída
	Centralidade	Bateria	DistBS	Eleição
1	Longe	Alta	Longe	Fraco
2	Longe	Alta	Moderado	Fraco
3	Longe	Alta	Perto	Média
4	Longe	Baixa	Longe	Muito Fraco
5	Longe	Baixa	Moderado	Muito Fraco
6	Longe	Baixa	Perto	Muito Fraco
7	Longe	Moderada	Longe	Muito Fraco
8	Longe	Moderada	Moderado	Fraco
9	Longe	Moderada	Perto	Muito Fraco
10	Moderado	Alta	Longe	Média
11	Moderado	Alta	Moderado	Forte
12	Moderado	Alta	Perto	Forte
13	Moderado	Baixa	Longe	Muito Fraco
14	Moderado	Baixa	Moderado	Fraco
15	Moderado	Baixa	Perto	Muito Fraco
16	Moderado	Moderada	Longe	Fraco
17	Moderado	Moderada	Moderado	Forte
18	Moderado	Moderada	Perto	Média
19	Perto	Alta	Perto	Muito Forte
20	Perto	Alta	Longe	Forte
21	Perto	Alta	Moderado	Muito Forte
22	Perto	Baixa	Longe	Muito Fraco
23	Perto	Baixa	Moderado	Muito_Fraco
24	Perto	Baixa	Perto	Muito Fraco
225	Perto	Moderada	Longe	Média
26	Perto	Moderada	Moderado	Forte
27	Perto	Moderada	Perto	Forte

O mesmo processo se repete no cálculo da distância de cada nó para a BS . Entretanto, vale resaltar que este critério é muito importante quando a EB não está localizada demasiadamente longe do cluster em questão . Logo, a disposição da rede deve ser considerada para que o critério distância para a EB seja válido . Com isso, definimos regras seguras para não gerar CHs próximo da borda da rede , já que a centralidade é mais importante.

4.6. MÚLTIPLOS SALTOS ENTRE CHS

Esta estratégia é utilizada para Cluster Heads que ultrapassam o limiar de comunicação, $d > d_0$, para o ponto de coleta . A aproximação para a BS , apresentada anteriormente, nem sempre se aplicará , uma vez que dependendo da disposição dos nós , clusters podem ser formados demasiadamente longe. A estratégia de múltiplos saltos é implementada para minimizar a dissipação de energia com comunicação.

Nesta proposta, para determinar que líderes passarão pelo processo de propagação de dados dos CHs mais afastados, o Sistema Fuzzy adota critérios de níveis de energia e distância. Primeiramente é utilizado o cálculo pela Eq. (3) para determinar as distâncias de comunicação entre os CHs mais afastados e o restante dos líderes eleitos. Com o resultado do cálculo de distância e os níveis de energia de cada líder, o Sistema Fuzzy é alimentado. As variáveis linguísticas que representam o nível de Bateria dos líderes e distância de todos os CHs eleitos para os CHs mais afastados são representadas, respectivamente, pelas Figuras 14 e 15.

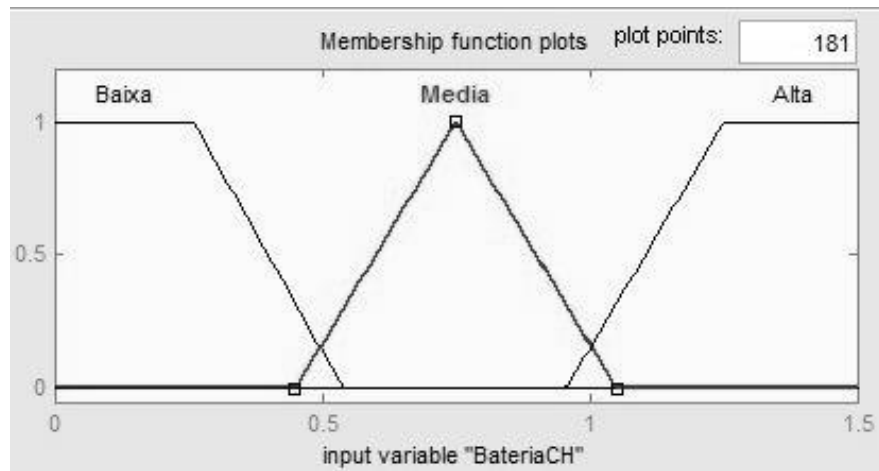


Figura 14 - Variável Bateria CH

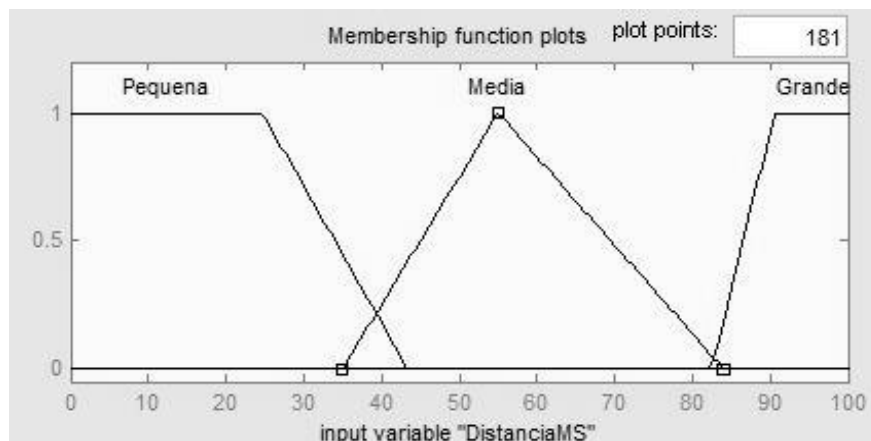


Figura 15 - Distância para múltiplos saltos

Ao longo do tempo de simulação, é natural que os nós apresentem um declínio de nível de energia. Logo, para evitar que nós líderes com baixo nível de energia passem pelo processo de encaminhamento de dados dos nós mais afastados, o critério energia é utilizado de forma discriminatória a fim de excluir estes nós do processo. Os sensores avançados e super nós tem maior chance de participarem do processo de propagação, uma vez que seus recursos energéticos são aumentados.

A variável Distância MS é utilizada para selecionar os nós líderes que propagarão os dados, já agregados, dos CHs cujas distâncias de comunicação para a BS ultrapassam o limiar d_0 . A variável possui os valores linguísticos: *Pequena, Moderada e Grande*, e o universo de discurso varia entre 0 e 100 m. A função de pertinência que representa o conjunto fuzzy *Grande* é definida a partir de 87,05m. Este limiar de comunicação foi estabelecido com base no quociente dos amplificadores de potência de sinal, $d_0 = \sqrt{E_{fs}/E_{amp}}$.

A base de regras é formada por 9 regras, uma vez que temos dois antecedentes como entrada, 3^2 . A melhor condição utilizada na seleção do CH ideal para propagação é dada por: *Se DistânciaMS é Pequena e BateriaCH é Alta Então a Saída é Forte*. Sendo a distância fator agravante, a base de regras exclui os nós que apresentam distância significativa para o CH mais afastado. A Figura 16 mostra o gráfico de superfície do sistema que indica os intervalos de valores para a localização das combinações ideais de saída do sistema na escolha entre os dois conjuntos de dados: *DistânciaMS* e *BateriaCH* do nó sensor.

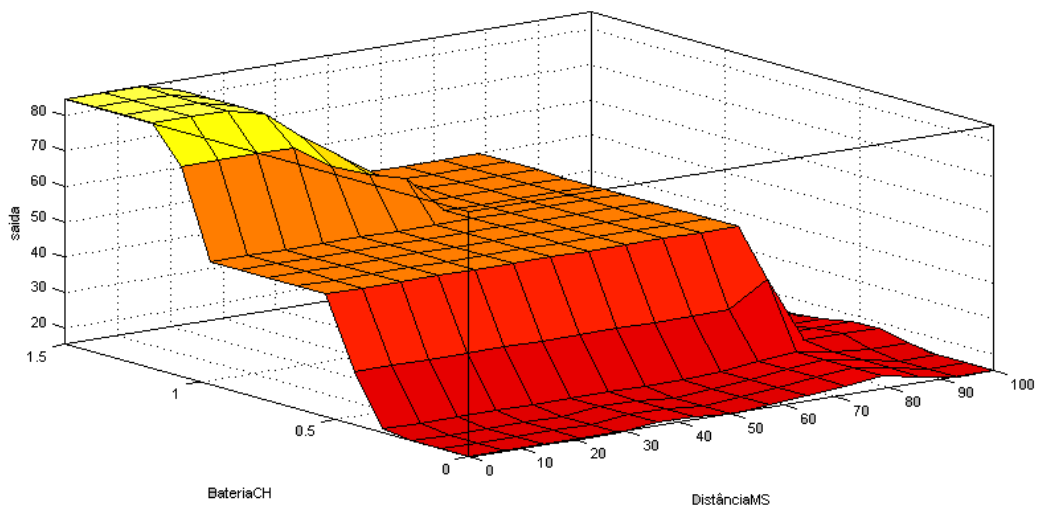


Figura 16 - Gráfico de superfície

4.7. MODELO DE SISTEMA FUZZY

Para o modelo de Lógica *Fuzzy*, foram utilizadas funções de pertinência triangulares e trapezoidais, máquina de inferência de *Mamdani* e Defuzzificador Centro Ponderado. No processo de seleção de CH, para cada entrada (x_1, x_2, x_3) , a saída do sistema é calculada, como mostra a Eq. (4).

$$y(x_1, x_2, x_3) = \frac{\sum_{l=1}^{27} \mu_{F_l^1}(x_1) \mu_{F_l^2}(x_2) \mu_{F_l^3}(x_3) C_{avg}^l}{\sum_{l=1}^{27} \mu_{F_l^1}(x_1) \mu_{F_l^2}(x_2) \mu_{F_l^3}(x_3)} \quad (5)$$

Para a seleção dos nós líderes que participarão do processo de propagação de dados de CHs mais afastados, a saída é calculada como mostra a Eq. (6). Onde (x, x_2) compreendem as entradas *BateriaCH* e *DistânciaMS*.

$$y(x_1, x_2) = \frac{\sum_{l=1}^9 \mu_{F_l^1}(x_1) \mu_{F_l^2}(x_2) C_{avg}^l}{\sum_{l=1}^9 \mu_{F_l^1}(x_1) \mu_{F_l^2}(x_2)} \quad (6)$$

4.8. MODELAGEM DA REDE

Para o modelo de rede assumimos que N (número de sensores) estão distribuídos em uma área $N \times M$, com três tipos de nós sensores, sensores normais, sensores avançados e super sensores, apresentando diferentes níveis de energia inicial, e representando a heterogeneidade da rede. O cálculo que determina a quantidade de sensores normais, sensores avançados e super sensores na rede, é semelhante ao utilizado por E-DEEC, calculados respectivamente pelas equações 7, 8 e 9.

$$N \cdot (1 - mf) \quad (7)$$

$$N \cdot mf \cdot (1 - mp) \quad (8)$$

$$N \cdot mf \cdot mp \quad (9)$$

Onde mf é a fração do número total de N nós sensores e mp a percentagem para o número total de nós sensores que apresentam mais energia que o nó sensor normal na rede.

Paracalcular-se a energia inicial total da rede, adotou-se a mesma equação apresentada em (SAINI, 2010), Eq.(9)

$$E_{total} = N. (1 - mf). E_o + N. mf(1 - mp). (2). E_o + N. mf. mp. E_o(1 + e) = N. E_o(1 + mf(2 + mp. e)) \quad (10)$$

Para a proposta apresentando três níveis de heterogeneidade, a energia total da rede é acrescida, considerando a maior capacidade energética dos sensores avançados e *super* sensores. Esta diferença é dada pelo fator $1 + mf(2 + mp. e)$ (SAINI, 2010)

4.9. PROPRIEDADES DA REDE

Para o cenário de rede proposto, assumimos algumas propriedades: (i) No que concerne à distribuição de N nós da rede, esta é feita de forma aleatória e não possuem nenhuma mobilidade; (ii) Os nós enviam sua localização para a BS utilizando GPS; (iii) Os nós dissipam energia para o envio de informação; (iv) Todos os nós detêm a mesma capacidade de transmissão e processamento, a heterogeneidade é aplicada a níveis de energia, já que alguns nós possuem recurso energético aumentado, o que difere dos sensores normais. (v) A BS é fixa e sua localização é pré-definida no algoritmo; (vi) Os nós sempre têm dados para transmitir para o CH.

5. SIMULAÇÃO E SELEÇÃO DO CLUSTER HEAD IDEAL

Conforme já citado, a seleção do CH é a base para escolha da melhor rota de comunicação entre os nós selecionados, a simulação com o MatLab (BIRAN, 2002), foi a ferramenta utilizada para atingir os resultados da pesquisa para referida eleição. As fases do processo de simulação, os valores ou requisitos da solução são apresentadas a seguir, por meio de uma breve descrição das mesmas.

A simulação no trabalho foi dividida em rounds. A cada round obtém-se um valor de saída com base nos parâmetros de entrada do sistema e um novo CH é selecionado para cada k cluster. Vale ressaltar que se torna inviável a apresentação de 5000 rounds neste trabalho, sendo descrito apenas a coleta de dados sem a apresentação de gráfico já que se trata dos valores inicialmente do round 0 (zero), visando a melhor apresentação do cenário proposto. Os valores são atualizados para entrada do round seguinte, onde executa as fases descritas no capítulo 3. Cada nó é distribuído aleatoriamente em uma área de 100 m², onde a classificação do número de nós normais, avançados e super, com seus respectivos níveis de energia, são calculados pelas equações 5, 6 e 7, sendo $e = 1$ e $mp = 0.6$. O nível de energia inicial depende de cada tipo de sensor especificado anteriormente com determinado percentual de distribuição e dispostos de forma aleatória, vale ressaltar que devido esta distribuição aleatória é possível que clusters sejam formados com maiores requisitos energéticos. Sendo que para os nós normais é de 0.5J, para os nós avançados e de 1.0J e para super nós 1.5J. A EB é previamente definida com as coordenadas $x = 5$ e $y = 95$. O modelo de rádio utilizado é geral e o mais analisado na literatura, segue a descrição da Tabela 4.

Tabela 4 - Parâmetros do Modelo de Rádio

Parâmetros	Valores
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/bit/m ²
E_{amp}	0.0013 pJ/bit/m ⁴
E_{DA} (Energia dissipada na agregação de dados)	5 nJ/bit/sinal
K -bit mensagem (dados)	4000 bits
K -bit mensagem (info)	100 bits
d_0 (Limiar de distância)	87m

Estão distribuídos 100 nós, divididos em k clusters, onde a quantidade de clusters assim como o modelo de rádio mencionado também está relacionada com vários trabalhos analisados no estado da arte [1], citados anteriormente no capítulo 2, esta quantidade é proporcionalmente 0.05 da quantidade de nós pertencentes a rede. Cada nó envia 4000 Bits de mensagem por round para o cluster head da rede. A taxa de compressão dos dados é de 5%. A simulação é feita com 5000 rounds. Neste cenário, é aplicado a métrica FND (*First Node Dies*) para determinar o período de estabilidade da rede.

Na primeira fase da simulação é obtida a coordenada e nível de energia de cada nó que compõe a rede. A energia dissipada no envio das informações de cada nó é calculada utilizando a Eq. (1). A energia inicial de cada nó é decrementada neste processo. Visto não ser energia residual e tratando-se de uma rede com 3 níveis de heterogeneidade e ainda com distribuição aleatória dos nós no plano como mencionado anteriormente, alguns clusters em particular, como no caso do inicial é formado pelo K-means com a maioria dos sensores sendo super nós. Após o envio das coordenadas, o algoritmo k-means, implementado na BS, estipula um padrão com base na coordenada dos nós, calculando o posicionamento de cada nó e dividindo os que formam a rede em k clusters. O algoritmo também calcula o centro de cada cluster, informação utilizada posteriormente para o processo de eleição do CH, descrito na seção 3.3. O número de clusters utilizados para simulação é $k=5$.

Cada nó, com seu respectivo valor de centralidade, proximidade para a EB e nível de energia, terá um consequente, pela aplicação do sistema *fuzzy*, um valor com grau de pertinência y , determinando a chance deste nó se tornar *cluster head*. No processo de defuzzyficação, o nó que apresentar maior valor de saída crisp, será eleito como CH ideal no round atual. Após este processo, a fase 2 (dois) da segunda etapa é iniciada, verificando se algum nó ultrapassa o limiar de comunicação. No caso do limiar ultrapassado, o Sistema Fuzzy determina o nó líder mais adequado para propagar os dados do líder mais afastado, considerando seu nível de energia e distância para o referido nó que ultrapassa o limiar.

A Figura 17, exibe a divisão dos clusters, em uma área de 100mx 100m, e o final da eleição de CHs no round 0, sendo representado por ‘ \diamond ’ vermelho os clusters head selecionados pelo sistema e o \square localizado próximo ao ponto 95 no eixo y, representa a EB. Os nós circulos representam os CHs eleitos que ultrapassam o limiar de comunicação com a BS.

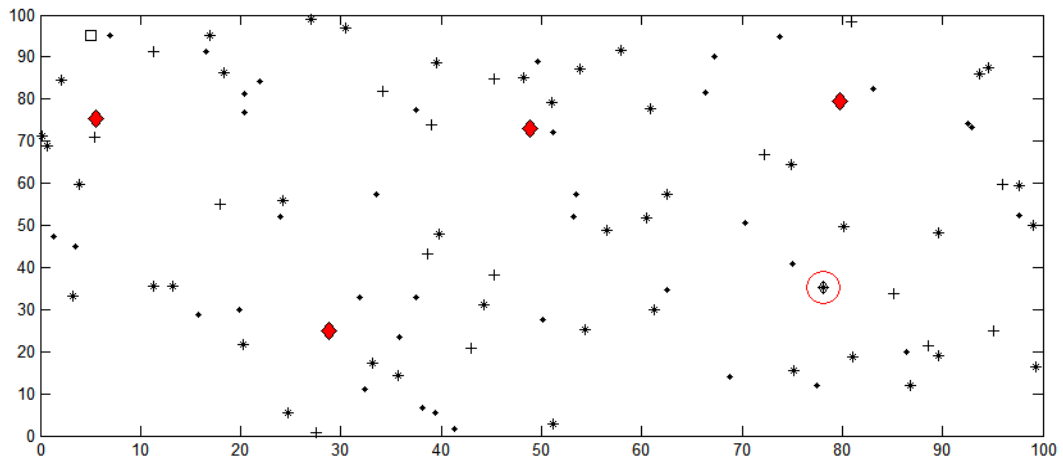


Figura 17 - CHs Selecionados pelo Sistema Fuzzy no Round 0

A Tabela 5 exibe o resultado da seleção de líderes para o round 0. Os valores de energia dos nós, exibidos na Tabela 5 já apresentam a dissipação de energia gerada na transmissão de informações para o ponto de coleta e dissipação na comunicação gerada no processo de associação ao CH eleito. O CH 5 apresentou maior dissipação de energia devido a quantidade de nós associados.

Tabela 5 - Resultados da Seleção de Cluster Heads

CH	Nível de energia	Centralidade No Cluster	Número de nós p/ cluster	Distância para BS
1	1.4965	7.7574	21	19,6 m < d_0
2	1.4965	2.4817	21	49,0 m < d_0
3	1.4972	12.2433	17	76,3 m < d_0
4	1.4975	12.1966	15	94,5 m > d_0
5	1.4957	4.6183	26	73,9 m < d_0

O CH 4, selecionado pelo Sistema Fuzzy, ultrapassa o limiar de comunicação. Em uma transmissão direta para o ponto de coleta, este nó, precisaria utilizar o modelo Multipath, gerando maior dissipação de energia no processo de comunicação, uma vez que o expoente de path loss seria d^4 . O consumo de energia gerado neste tipo de transmissão seria de aproximadamente 2.3935 J, levando o nó à inatividade de forma prematura e quebrando o período de estabilidade da rede. Entretanto, com a estratégia de múltiplos saltos adotada, o CH comunica-se com líder mais adequado, mantendo o modelo *Free space* de rádio. A Tabela 6 exibe a saída do Sistema Fuzzy, no round 0, para seleção do nó líder adequado para

propagação dos dados do CH mais afastado. O maior valor de saída *crisp* corresponde ao nó líder selecionado, CH 3.

Tabela 6 - Saída do Sistema Fuzzy

<i>CH</i>	<i>Nível de energia</i>	<i>Distância para o CH > d₀</i>	<i>Saída Fuzzy</i>
1	1.4965	83.1566	18.2613
2	1.4965	47.8956	52.9779
3	1.4972	44.4752	57.3361
5	1.4957	50.3652	50.0000

Para melhor validação da simulação, a proposta deste trabalho foi comparada com os algoritmos LEACH e E-DEEC. Para avaliação de desempenho é utilizado o final do período de estabilidade da rede e o tempo de vida útil. A escolha dos algoritmos para comparação se dá principalmente pela utilização de informações locais sem considerar critérios de posicionamento, para eleição dos CHs. Além do método de escolha do líder, o algoritmo LEACH, não trata as discrepâncias de energia dos nós que compõem a rede.

Diferente de LEACH, o algoritmo E-DEEC, considera a heterogeneidade dos nós para eleição do CH. Entretanto, utiliza informações locais para eleição do líder e insere três níveis de heterogeneidade, semelhante a proposta apresentada neste trabalho.

A Figura 18 exibe a quantidade de nós sensores ativos no tempo de vida útil da rede. Esta medida reflete o número total de nós que ainda não esgotaram sua energia.

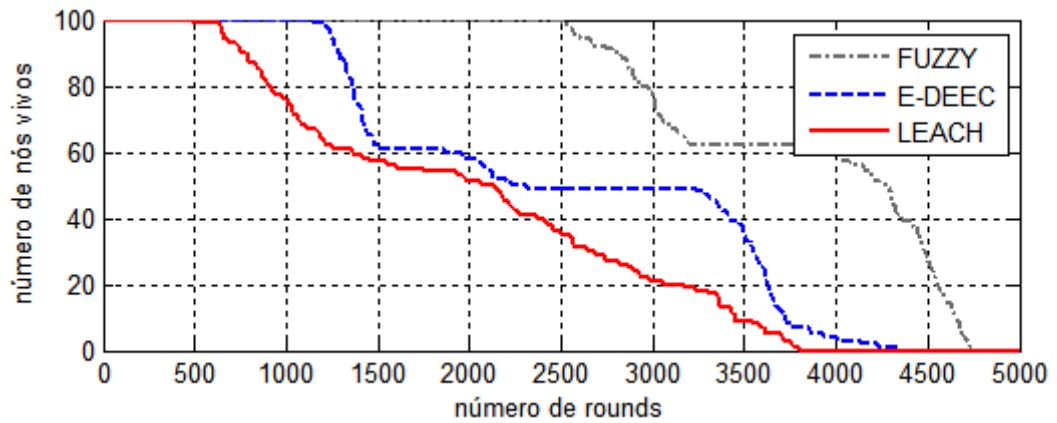


Figura 18 - Número de Sensores Ativos no Período de Simulação

Os resultados indicam que a inserção de novos níveis de heterogeneidade, a abordagem da lógica fuzzy como ferramenta de seleção, a utilização de múltiplos saltos entre os CHs e informações centralizadas na BS, permitem eleger líderes mais eficientes, aumentando o período de estabilidade e o tempo de vida da rede.

A proposta tende a apresentar melhores resultados, quando comparado com os algoritmos LEACH e E-DEEC. Pela Figura 18, pode-se perceber que o algoritmo LEACH apresentou o menor período de estabilidade, ocorrendo por volta de 500 rounds. O algoritmo E-DEEC apresentou desempenho superior ao do LEACH, com o período de instabilidade da rede iniciado por volta 1148 rounds. Já o algoritmo da proposta, mostra um melhor desempenho sobre os algoritmos comparados, aumentando o período de estabilidade da rede até aproximadamente 2500 rounds, quando ocorre a primeira inatividade de um nó por falta de energia.

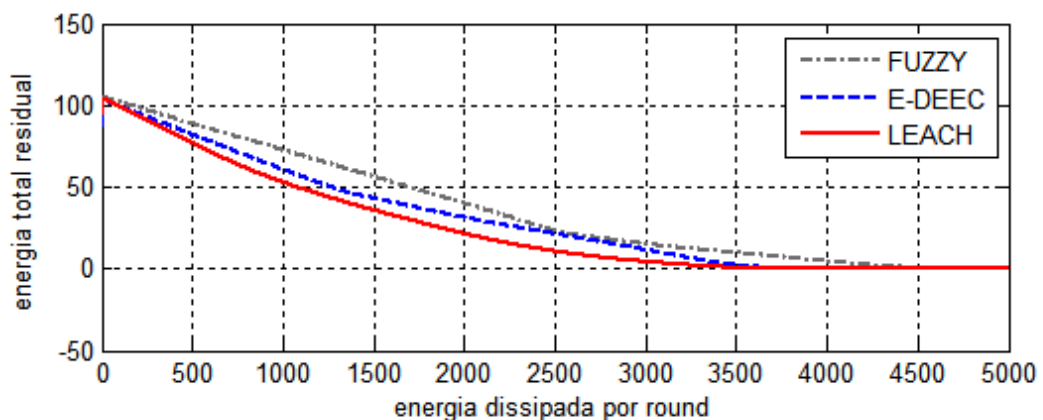


Figura 19 - Energia Total Residual de LEACH, E-DEEC e FUZZY

A Figura 19 exibe a energia dissipada pela rede ao longo dos 5000 rounds de simulação para cada algoritmo comparado. O total de energia para cada rede é de 104.5J. A dissipação de energia apresenta um declínio linear ao longo de 2500 rounds, para a proposta apresentada e para o algoritmo E-DEEC por volta de 1500 rounds, mudando a partir do momento em que o primeiro nó na rede fica inativo, quebrando o período de estabilidade. Ambos os algoritmos, *Fuzzy* e E-DEEC, permitem tratar as discrepâncias energéticas de cada nó na rede para eleição do CH, enquanto o algoritmo LEACH apresenta maior dissipação de energia por round, consequência dos problemas descritos anteriormente sobre a forma de eleição de CH pelo algoritmo e a falta de tratamento discriminatório das discrepâncias energéticas dos nós que formam a rede.

6. CONCLUSÕES E TRABALHOS FUTUROS

Os resultados indicam que a proposta apresentada oferece grandes vantagens, permitindo selecionar os nós mais adequados para líderes do grupo a cada *round* com base nos valores de defuzificação do Sistema Fuzzy. Vantagens também decorrentes da implementação de múltiplos saltos entre CHs por meio da lógica fuzzy, uma vez que minimiza a dissipação de energia dos CHs selecionados mais afastados do ponto de coleta. A inserção de três níveis de heterogeneidade, correspondente aos sensores normais, avançados e super sensores, contribui consideravelmente para o aumento do período de estabilidade da rede, uma vez que esta inserção dá a rede maior recurso energético. Entretanto, os resultados indicam que se esta discrepância não for considerada no momento de seleção do CH ela não influencia de forma considerável no aumento do período estável.

Outra grande vantagem que contribui para os resultados obtidos neste trabalho é a utilização de um controle central na EB. Por não possuir severas limitações de energia, processamento e armazenamento como os nós que formam a rede, a EB apresenta vantagens sobre o processamento local de informações em cada nó, quando envia informações de atualização do nível de energia e posicionamento em cada round, processo este encontrado nos algoritmos tradicionais usados para seleção de CHs. Entretanto, mesmo com esta atualização a proposta ainda apresenta melhorias sobre os outros modelos apresentados. Outra vantagem do controle central na EB está no momento de seleção do CH, por ter o papel de informar a rede sobre os líderes selecionados para cada cluster, previamente divididos pelo algoritmo k-means.

Este processo difere-se do processo encontrado nos algoritmos que utilizam informações locais para seleção de seu líder, cabendo ao próprio CH selecionado enviar mensagens em *broadcast* para a rede, gerando dissipação de energia no momento da propagação. Finalmente, o trabalho apresentado tem a principal contribuição na seleção do CH mais eficiente, considerando sua localização e discrepâncias de níveis de energia, como também, na inclusão de novos níveis de heterogeneidade, permitindo aumentar o período de estabilidade da rede, ou seja, o período que a rede é totalmente funcional, aumentando consideravelmente o tempo de vida útil em RSSF heterogêneas.

Observou-se que a seleção do CH mais eficiente compreende várias coletas de dados sobre o posicionamento do nó em relação a rede e ao nível de energia da bateria. Em vista da complexidade computacional envolvida no processamento destas informações, a comunicação dos nós sensores entre si e com a EB ainda constitui-se um processo delicado, porém é

possível. Considerando-se os resultados aqui obtidos e discutidos, sugerem-se como trabalhos futuros:

- A inclusão de um sistema *Neuro-Fuzzy* bem como melhorias das variáveis de entrada do sistema, através do aumento do conjunto de funções fuzzy, visando aumentar a precisão na seleção em cada *round*;
- A realização de testes aplicando um modelo de reorganização dos clusters com base na energia total residual da rede associado ao modelo proposto.
- A aplicação desta proposta em *testbed*, comparando os resultados obtidos com os usados em ferramentas de Simulação para RSSF.

Por fim, com o estudo foi possível constatar que o uso desta proposta utilizando ferramentas da inteligência computacional é válido para o controle energético em RSSF e podem ser utilizadas com outras formas de validação, mesmo em ambientes reais em trabalhos futuros.

REFERÊNCIAS

AKYILDIZ, I. F. W; SU, Y. ; SANKARASUBRAMANIAM, E. **A Survey on Sensor Networks**. IEEE Communications, Vol. 40, No 8:102–114, 2002.

BEZDEK, J. C. **Reconhecimento de Padrões com algoritmos fuzzy função objetivo**, Plenum Press, New York, 1984.

BIRAN A. e Breiner M. **Matlab for engineers**. Prentice Hall. 2002

DUNN, J. C. **A relativa fuzzy do Processo ISODATA e seu uso na detecção de compactos Clusters bem separados**, Journal of Cibernética, Vol. 3: 32-57. 1973.

GONÇALVES, A. P. **Aplicação de Lógica Fuzzy em Guerra Eletrônica**. IX Simpósio de Guerra Eletrônica, São Paulo, 2007.

HAN, L. **LEACH-HPR: An energy efficient routing algorithm for Heterogeneous WSN**, Intelligent Computing and Intelligent Systems (ICIS), IEEE International Conference on , vol.2, pp. 507-511, 2010.

JANG, J. R., SUN, C. **Neuro-Fuzzy Modeling and Control**, Proceedings of the IEEE, v. 83, n. 3, pp. 378-406, March . 1995.

JANG, J. R.; GULLEY, N. **Fuzzy Logic Toolbox User's Guide**. The Mathworks, Inc. 1997.
LOUREIRO, A. A. F. et al. **Redes de sensores sem fio**. 21º Simpósio Brasileiro de Redes de Computadores, Natal, RN, Brasil, 2003.

LOZANO, J.A.; PENA, J.M.; LARRANAGA, P. **An empirical comparison of four initialization methods for the k- means algorithm**, Pattern Recognition Lett. Vol. 20, 1027–1040, 1999.

LUIZ, H. A. Correia, Daniel F. Macedo, Aldri L. dos Santos, José M. Nogueira, Antonio A. F. Loureiro **A taxonomy for medium access control protocols in wireless sensor networks**.. Annales des Télécommunications (Annals of telecommunications), 2005

MACIEL, Christiano, C. Tassio, J. Jailton, M. Alexandre e T. Carlos. **Estratégia centralizada energeticamente eficiente para RSSF heterogêneas utilizando Lógica Fuzzy**. CRC: 12ª Conferência sobre Redes de Computadores, Aveiro, Portugal. Sessão 5. pp 126 - 132, 2012.

MIRANDA, P.; JUNIOR, M. B. V.; KRONBAUER, D. **Sistema de Controle Difuso de Mamdani Aplicações: Pêndulo Invertido e outras**. Monografia de graduação, Universidade Federal do Mato Grosso do Sul, MS, 2003.

MUBARAK, T.M., SATTAR, S. A., RAO, G. A., SAJITHA, M. **Intrusion detection: An energy efficient approach in heterogeneous WSN**, Emerging Trends in Electrical and Computer Technology (ICETECT), International Conference on vol., no., pp.1092- 1096, 23-24, 2011.

- POTTIE, G. J., KAISER, W. J. **Wireless integrated network sensors (WINS)**. Communications of the ACM. 43, 5, 51-58, 2000.
- QUING, L., ZHU, Q., Wang, M. **Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks**. ELSEVIER, Computer Communications 29, pp 2230- 2237, 2006.
- RUIZ, L. BEATRYS. **Uma arquitetura para o gerenciamento de redes de sensores sem fio**. Technical Report DCC/UFGM RT.005/2002, Departamento de Ciência da Computação da Universidade Federal de Minas Gerais, 2002.
- SAINI, P., SHARMA, A. K., **E-DEEC- Enhanced Distributed Energy Efficient Clustering scheme for heterogeneous WSN**, Parallel Distributed and Grid Computing (PDGC), 1st International Conference on , vol., no., pp.205-210, 28-30, 2010.
- SMARAGDAKIS, G., I. M., BESTAVROS, A. **SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks**, in: Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA), 2004.
- SOUZA, R. A. de.: **Sistemas Inteligentes de Apoio a Decisão Utilizando Lógica Fuzzy**. Dissertação (Mestrado em Engenharia Elétrica), DEEC, UFPA – Pará, 2003.
- TAN, P. **Introduction to Data mining**. Boston: Addison-Wesley, 769 p, 2006.
- TANSCHKEIT, R. **Sistemas Fuzzy**. Monografia, Graduação, Departamento de Engenharia Elétrica - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro. 2002.
- WANG, J. and SU, X. **An improved K-means clustering algorithm**, in 3rd International Conference on Communication Software and Networks (ICCSN), Xi'an, 2011.
- WANG, L. *A Course in Fuzzy Systems and Control*. PrenticeHall, Englewood Cliffs, NJ, 1997.
- YAN, B., WU, X., ZHOU, X. **A Improved Base Station Cooperative Mobile Strategy for WSN with Finite Powered Cluster Heads**. Wireless Communications Networking and Mobile Computing (WiCOM), 6th International Conference, vol., no., p.1-4, 23-25, 2010.
- ZADEH, L. A. **Fuzzy Sets. Information and Control**, vol. 8, pp 338- 353, 1965.

APÊNDICE

APÊNDICE A - PUBLICAÇÃO RELACIONADA À DISSERTAÇÃO

ANAIS: 31º Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos – SBRC 2013 – Trilha principal - Sociedade Brasileira de Computação (SBC). pp. 355- Universidade de Brasília (UNB) - Brasília-DF, Brasil - 2013.

TÍTULO:CONTROLE ENERGETICAMENTE EFICIENTE DE MÚLTIPLOS SALTOS PARA REDES DE SENSORES SEM FIO HETEROGÊNEAS UTILIZANDO LÓGICA

FUZZY

Abstract. *This paper, we propose a centralized control to elect more appropriate Cluster Heads, assuming three levels of heterogeneity and multi-hop communication between Cluster Heads. The centralized control uses the k-means algorithm, responsible for the division of clusters and Fuzzy Logic to elect the Cluster Head and selecting the best route of communication between elected. The simulations indicate that the centralized control, the inclusion of three levels of heterogeneity and multi-hop communication to farthest Cluster Heads can increase the period of stability and lifetime in WSN.*

Keywords: *RSSF, Fuzzy Logic, Cluster Heads, heterogeneity*