

Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Coordenação de Pós-Graduação em Ciência da Computação

Dissertação de Mestrado

Configuração das Posições dos Nós de uma Rede de Sensores Sem Fio Baseada em Algoritmos Genéticos

Alexandre Sales Vasconcelos

Campina Grande, Paraíba, Brasil

Agosto de 2013

Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Coordenação de Pós-Graduação em Ciência da Computação

Configuração das Posições dos Nós de uma Rede de Sensores Sem Fio Baseada em Algoritmos Genéticos

Alexandre Sales Vasconcelos

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande – Campus I como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Elmar Uwe Kurt Melcher

Joseana Macêdo Fechine Régis de Araújo

(Orientadores)

Campina Grande, Paraíba, Brasil

Agosto de 2013

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

V331c Vasconcelos, Alexandre Sales.
Configuração das posições dos nós de uma rede de sensores sem fio baseada em algoritmos genéticos / Alexandre Sales Vasconcelos. – 2013.
190 f. : il. color.

Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

"Orientação: Prof. Dr. Elmar Uwe Kurt Melcher, Prof^a. Dr^a. Joseana Macêdo Fechine Régis de Araújo".

Referências.

1. Rede de Sensores sem Fio. 2. Algoritmos Genéticos. I. Melcher, Elmar Uwe Kurt. II. Araújo, Joseana Macêdo Fechine Régis de. III. Título.

CDU 004.7(043)

**"CONFIGURAÇÃO DAS POSIÇÕES DOS NÓS DE UM REDE DE SENSORES SEM FIO
BASEADO EM ALGORITMOS GENÉTICOS"**

Alexandre Sales Vasconcelos
ALEXANDRE SALES VASCONCELOS

DISSERTAÇÃO APROVADA EM 21/08/2013

E. Melcher

ELMAR UWE KURT MELCHER, Dr., UFCG
Orientador(a)

Joseana Macêdo Fechine Régis de Araújo

JOSEANA MACÊDO FECHINE RÉGIS DE ARAÚJO, D.Sc, UFCG
Orientador(a)

Alisson Vasconcelos de Brito

ALISSON VASCONCELOS DE BRITO, D.Sc, UFPB
Examinador(a)

Andrey Elísio Monteiro Brito

ANDREY ELÍSIO MONTEIRO BRITO, Dr., UFCG
Examinador(a)

CAMPINA GRANDE - PB

A minha filha Amanda Gabriela e esposa Andréa, pelo tempo que deixamos de estar
juntos...

Aos meus pais, Joval e Maria Tereza, a eles todos os créditos.

Agradecimentos

Aos Professores Dr. Elmar Uwe Kurt Melcher e Dra. Joseana Macêdo Fachine Régis de Araújo, pela oportunidade, dedicação e paciência, que foram fundamentais em todos os momentos da realização do trabalho.

Aos meus colegas de pós-graduação, em especial à Maria Isabel, à Danielle Chaves e a Sérgio Espínola, que tornaram um período de longa dedicação menos cansativo.

Aos meus irmãos, Antônio Carlos, Carlos Alberto, Luiz Gonzaga, Jomar e João Batista, que sempre participaram direta ou indiretamente da minha formação acadêmica.

À Kalinka Walderea pela ajuda na revisão da dissertação.

Ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande (COPIN/UFCG), pela oportunidade que me foi dada.

Aos professores do mestrado que, a partir das disciplinas, contribuíram na construção do trabalho e aos servidores da COPIN pelo apoio.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Resumo

As Redes de Sensores Sem Fio (RSSF) estão sendo utilizadas nos mais variados tipos de aplicação e um dos aspectos importantes a considerar é o tempo de funcionamento, que está ligado diretamente ao consumo de energia. Além deste, existem outros tão importantes quanto, como os relacionados à qualidade de serviço (QoS), mais especificamente à Taxa de Cobertura da Área de interesse da aplicação, o Tempo de Atraso Fim-a-Fim (TAFF) e a Taxa de Perda de Pacotes (TPP) da RSSF. Neste contexto, esta dissertação propõe a utilização de um algoritmo de busca, baseado em Algoritmos Genéticos (AG), cujo objetivo é encontrar a melhor configuração para implantação dos nós sensores de uma RSSF, utilizando as métricas energia residual, taxa de cobertura, TAFF e TPP. Para que os resultados obtidos a partir da abordagem proposta sejam os mais realistas possíveis, nas simulações foram utilizados módulos que implementam a presença de obstáculos nos cenários. Os resultados obtidos, inicialmente, sem o uso de obstáculos, mostram que o uso de AG proporciona um ganho acima de 78,0% para o somatório da energia residual, uma redução do TAFF de, aproximadamente, 50,0% e uma redução da TPP de 44,0% em relação a uma abordagem que utiliza distribuição aleatória dos nós sensores, ora denominada abordagem tradicional. Outro resultado relevante mostra que a utilização da abordagem proposta apresentou taxa de sucesso de 95,0%, na busca da configuração de implantação dos nós sensores que proporcione a cobertura de 100% da área do cenário, contra uma taxa de 10% de sucesso para a abordagem tradicional. Observou-se, também, uma diferença significativa entre os resultados das simulações com o módulo de obstáculos do cenário ativo e com o módulo desativado. No experimento para verificação da influência do módulo de obstáculos sobre a TPP, a redução com esse módulo desativado foi de 12,0% enquanto que a redução com o módulo ativado foi superior a 39,0%. Os resultados dos experimentos confirmaram que o modelo implementado proporcionou uma melhoria nas métricas utilizadas na busca, aumentando a taxa de cobertura e o tempo de vida da RSSF e, ao mesmo tempo, reduzindo o TAFF e a TPP, quando comparado ao modelo tradicional.

Abstract

Wireless Sensor Networks (WSN) are being used in various kinds of application and one of the important aspects to consider is the running time, which is directly linked to energy consumption. Furthermore, there are other aspects just as important, like those related to Quality of Service (QoS), more specifically the Coverage Ratio of the region of interest in the application, the End-to-End Delay time (EED) and the Package Loss Rate (PLR) of a WSN. In this context, this paper proposes the use of a search algorithm, based on Genetic Algorithms (GA), whose objective is to find the best configuration for deployment of sensor nodes in a WSN, using the residual energy metrics, coverage rate, EED and PLR. In order for the results obtained from the proposed approach to be the most realistic as possible, it was used simulation modules that implement the presence of obstacles in the surroundings. The results obtained, initially without the use of obstacles, show that the use of GA provides a gain up to 78.0% for the sum of the residual energy, a reduction of almost 50.0% in the EED and 44.0% in the PLR, compared to an approach that uses random distribution of sensor nodes, called traditional approach. Another result shows that the use of the proposed approach had a success rate of 95.0% when searching for the deployment configuration of the sensor nodes to provide 100% coverage of the scene region, compared to a 10% success rate for traditional approach. There was also a significant difference between the results of the simulations with the module of the scenario obstacles activated and the ones with the module disabled. In the experiment to check the influence of the obstacles on the module over the PLR, it was observed a reduction of the PLR with the module of the scenario obstacles off of 12.0%, whereas the reduction with the module activated was of more than 39.0%. The results of the experiments confirmed that the implemented model gives a better gain in the metrics the search, maximizing the coverage rate and time of life the WSN and minimizing the TAFF and TPP compared with the traditional model.

Sumário

1	Introdução.....	19
1.1	Motivação e Relevância do Trabalho	22
1.2	Objetivo Geral	23
1.3	Objetivos Específicos	23
1.4	Estrutura do Trabalho.....	24
2	Rede de Sensores Sem Fio.....	25
2.1	Caracterização das RSSF.....	26
2.2	Arquitetura de uma Rede de Sensores Sem Fio.....	31
2.3	Modelo Funcional de uma RSSF	33
2.3.1	Estabelecimento	34
2.3.2	Manutenção.....	35
2.3.3	Sensoriamento.....	37
2.3.4	Processamento.....	37
2.3.5	Comunicação.....	38
2.4	Energia em Rede de Sensores Sem Fio	39
2.5	Protocolos de Acesso ao Meio para RSSF	40
2.5.1	Padrão IEEE 802.11 (Wi-Fi)	41
2.5.2	Padrão IEEE 802.15.1 (<i>Bluetooth</i>).....	42
2.5.3	Padrão IEEE 802.15.4 (<i>ZigBee</i>)	44
2.6	Discussão Geral.....	47
3	Algoritmos Genéticos.....	48
3.1	Técnicas de busca aleatórias guiadas	48
3.2	O algoritmo genético básico	52
3.2.1	Cromossomos e Genes	54
3.2.2	Representação Cromossomial.....	54
3.2.3	População Inicial.....	59
3.2.4	Função de Avaliação	60
3.2.5	Critério de Parada.....	62
3.2.6	Seleção.....	62
3.2.7	Elitismo	65
3.2.8	Operadores Genéticos.....	66
3.3	Pseudocódigo de um Algoritmo Genético	69

3.4	Discussão Geral	70
4	Trabalhos Relacionados	72
4.1	Discussão Geral	81
5	Descrição do Modelo RSSF-AG	85
5.1	Modelo RSSF-AG.....	85
5.1.1	Configuração do AG e do Simulador	86
5.1.2	Execução do Algoritmo	86
5.1.3	Configuração de Implantação dos Nós da RSSF	88
5.2	Algoritmo Genético Proposto	88
5.2.1	Inicialização	90
5.2.2	Função de Avaliação	93
5.2.3	Critério de Parada.....	95
5.2.4	Elitismo	95
5.2.5	Descarte	96
5.2.6	Seleção.....	96
5.2.7	Operadores genéticos	97
5.3	Discussão Geral.....	97
6	Apresentação e Análise dos Resultados	98
6.1	Ambiente de Desenvolvimento e Execução dos Experimentos	98
6.2	Parâmetros de Configuração para o AG do Modelo Proposto	99
6.2.1	Tamanho da População.....	100
6.2.2	Seleção.....	102
6.2.3	Cruzamento	103
6.2.4	Mutação.....	104
6.2.5	Parâmetros de Configuração.....	105
6.3	Experimento 1	108
6.4	Experimento 2	115
6.5	Experimento 3	121
6.6	Desempenho do RSSF-AG comparado com outros trabalhos.....	126
6.7	Resumo dos Resultados	128
6.8	Discussão Geral.....	129
7	Considerações Finais e Sugestões para Trabalhos Futuros	130
7.1	Sugestões para Trabalhos Futuros	133

Referências Bibliográficas	135
APÊNDICE A Análise Estatística	144
A.1 Parâmetro Tamanho da População	144
A.2 Parâmetro Método de Seleção.....	148
A.3 Parâmetro Método de Cruzamento	152
A.4 Parâmetro Probabilidade de Mutação	157
A.5 Energia Residual.....	161
A.6 Tempo de Atraso Fim-a-Fim	164
A.7 Taxa de Perda de Pacotes.....	168
A.8 Ganho do Algoritmo sobre a QoS	172
A.9 Análise inferencial do modelo RSSF-AG.....	175
APÊNDICE B Gerador do Arquivo de Configuração.....	178
B.1 Entrada dos Dados de Configuração	178
B.2 Execução do Algoritmo	183
B.3 Saída do Arquivo de Implantação	185
APÊNDICE C Script de Configuração do INET.....	187

Lista de Abreviaturas

AE	Algoritmos Evolutivos
AG	Algoritmos Genéticos
ANOVA	Análise de Variância
AODV	<i>Ad Hoc On-Demand Distance Vector Routing</i>
AP	<i>Access Point</i>
APL	<i>Application Layer</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDMA	<i>Code Division Multiple Access</i>
CPT	Controle de Potência de Transmissão
CSMA/CA	<i>Carrier Sense Multiple Access with Collision Avoidance</i>
CTS	<i>Clear to Send</i>
DAAD	Serviço Alemão de Intercâmbio Acadêmico
DSR	<i>Dynamic Source Routing</i>
DSSS	<i>Direct Sequence Spread Spectrum</i>
EE	Estratégias Evolucionárias
FDMA	<i>Frequency Division Multiple Access</i>
FFD	<i>Full Function Device</i>
GPSR	<i>Greedy Perimeter Stateless Routing</i>
HD	<i>Hard Disk</i>
IC	Intervalo de Confiança
ICMP	<i>Internet Control Message Protocol</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IP	<i>Internet Protocol</i>
ISM	<i>Industrial, Scientific and Medical bands</i>
LEACH	<i>Low-Energy Adaptive Clustering Hierarchy</i>
MAC	<i>Medium Access Control</i>
ns-2	<i>Networks Simulator 2</i>
NWK	<i>Network Layer</i>
OFDM	<i>Orthogonal Frequency Division Multiplexing</i>

<i>OLSR</i>	<i>Optimized Link State Routing</i>
<i>OSI</i>	<i>Open Systems Interconnection</i>
<i>PC</i>	<i>Personal Computer</i>
PE	Programação Evolutiva
<i>PHY</i>	<i>Physical Layer</i>
<i>PPP</i>	<i>Point to Point Protocol</i>
PROBRAL	Programa de apoio a projetos conjuntos de pesquisa entre grupos brasileiros e alemães
QoS	Qualidade de Serviço
<i>RFD</i>	<i>Reduced Function Device</i>
RSSF	Rede de Sensores Sem Fio
<i>RTS</i>	<i>Request to Send</i>
<i>SDMA</i>	<i>Space Division Multiple Access</i>
<i>SIG</i>	<i>Special Interest Group</i>
<i>S-MAC</i>	<i>Sensor-MAC</i>
<i>SoC</i>	<i>System on a Chip</i>
<i>SPIN</i>	<i>Sensor Protocols for Information via Negation</i>
TAFF	Tempo de Atraso Fim-a-Fim
<i>TCP</i>	<i>Transmission Control Protocol</i>
<i>TDMA</i>	<i>Time Division Multiple Access</i>
TPP	Taxa de Perda de Pacotes
<i>UDP</i>	<i>User Datagram Protocol</i>
<i>Wi-Fi</i>	<i>Wireless Fidelity</i>

Lista de Figuras

Figura 2.1 Exemplo da distribuição dos nós sensores por: (a) difusão simples, (b) constante; (c) <i>r-random</i> , adaptado de (ISHIZUKA e AIDA, 2004; NAMBA, 2011)	27
Figura 2.2 Estrutura básica do <i>hardware</i> de um nó sensor, adaptado de LOUREIRO et al. (2003)	32
Figura 2.3 Modelo genérico de uma RSSF, adaptado de LOUREIRO et al. (2003)	33
Figura 2.4 Ciclo de vida da RSSF, adaptada de LOUREIRO et al. (2003)	34
Figura 2.5 Modelos de posicionamento dos nós sensores em uma RSSF, adaptado de NAKAMURA (2003)	35
Figura 2.6 Fases de estabelecimento da rede de sensores, adaptado de LOUREIRO et al. (2003)	36
Figura 2.7 Limitações quanto a obstáculos e faixa de alcance, adaptada de LOUREIRO et al. (2003)	38
Figura 2.8 Representação gráfica do mapa de energia, adaptada de NAMBA (2011)	40
Figura 2.9 Exemplo de uma rede piconet, adaptada de NAMBA (2011)	44
Figura 2.10 Exemplo de uma scatternet, adaptada de NAMBA (2011)	44
Figura 2.11 Exemplos das topologias empregadas nas redes ZigBee, adaptada de NAMBA (2011)	45
Figura 2.12: Camadas do padrão IEEE 802.15.4 (ZigBee), adaptada de NAMBA (2011)	46
Figura 3.1 Algoritmos evolucionários como técnicas de busca, adaptada de BURIAN (2009)	50
Figura 3.2 Função hipotética com um máximo local e outro global, adaptada de LINDEN (2006)	51
Figura 3.3 Diagrama básico de um algoritmo genético básico, adaptada de LINDEN (2006)	53
Figura 3.4 Representação de um vetor cromossomo do tipo binário	57
Figura 3.5 Representação de um vetor cromossomo do tipo real	58
Figura 3.6 Método de seleção por roleta viciada, adaptada de LOBO (2005)	63
Figura 3.7 Método de seleção por torneios, adaptada de LINDEN (2006)	64
Figura 3.8 Método de amostragem universal estocástica, adaptada de (LOBO, 2005)	65
Figura 3.9 Exemplo do operador de cruzamento de um ponto, adaptada de LOBO (2005)	67
Figura 3.10 Exemplo do operador de cruzamento de dois ponto, adaptado de (LOBO, 2005)	68
Figura 3.11 Exemplo do operador de cruzamento uniforme, adaptado de (LOBO, 2005)	68
Figura 3.12 Exemplo do operador de mutação, adaptado de (LOBO, 2005)	69
Figura 5.1 Fluxograma das fases do modelo proposto	85

Figura 5.2 Fluxograma simplificado do modelo proposto	89
Figura 6.1 Boxplot do ganho do algoritmo em função do tamanho da população ...	101
Figura 6.2 Boxplot do ganho do algoritmo em função do método de seleção	103
Figura 6.3 Boxplot ganho do algoritmo em função da configuração de cruzamento	104
Figura 6.4 Boxplot ganho do algoritmo em função da probabilidade de mutação ...	105
Figura 6.5 Exemplo detalhado de um cenário com obstáculo	116
Figura 6.6 Boxplot do somatório da energia residual em relação a presença de obstáculos no cenário	118
Figura 6.7 Boxplot da média dos TAFF em relação aos obstáculos no cenário	119
Figura 6.8 Boxplot da média da TPP em relação aos obstáculos no cenário	121
Figura 6.9 Mosaicos Regulares, adaptado de (ALVES e DALCIN, 1999)	122
Figura 6.10 Configuração de implantação dos nós sensores para área de 100,0x100,0 m, utilizando a técnica de mosaicos planos	123
Figura 6.11 Configuração de implantação dos nós sensores para um cenário com área de 28,3x14,1 m, utilizando técnica de mosaicos planos	124
Figura 6.12 Representação grafica do resultado de busca do algoritmo proposto nesta dissertação, para um cenário de 100,0x100,0 metros e raio do nó sensor de 10,0 metros	125
Figura 6.13 Comparação do desempenho do modelo RSSF-AG com os resultados da taxa de cobertura do trabalho de (QUINTÃO, MATEUS e NAKAMURA, 2004)	127
Figura 7.1 Exemplo de um cromossomo para 2 (dois) nós sensores.....	132
Figura A.1 Boxplot da variável ganho do algoritmo para o tamanho da população.	145
Figura A.2 Boxplot do Ganho do Algoritmo em relação aos grupos do tamanho da população	148
Figura A.3 Boxplot da variável ganho do algoritmo para o método de seleção	150
Figura A.4 Boxplot do ganho do algoritmo em função dos grupos do método de seleção	152
Figura A.5 Boxplot da variável ganho do algoritmo para a configuração de cruzamento	154
Figura A.6 Boxplot do ganho do algoritmo em função das configurações de cruzamento	156
Figura A.7 Boxplot da variável ganho do algoritmo para a probabilidade de mutação	158
Figura A.8 Boxplot do ganho do algoritmo em função dos valores da probabilidade de mutação	160
Figura A.9 Boxplot do somatório da energia residual em relação a presença de obstáculos no cenário	163
Figura A.10 Histogramas dos dados originais e do logaritmo do TAFF	166
Figura A.11 Boxplot da média dos TAFF em relação aos obstáculos no cenário ...	167
Figura A.12 Boxplot da média da TPP em relação aos obstáculos	171
Figura B.1 Tela do aplicativo Terminal do Ubuntu.....	178

Figura B.2 Tela do aplicativo do gerador do arquivo de configuração do algoritmo proposto.....	180
Figura B.3 Exemplo do arquivo de configuração do algoritmo proposto	184
Figura B.4 Tela do terminal com o comando de execução do algoritmo	184
Figura B.5 Exemplo do arquivo gerado ao final da execução do algoritmo de busca	185
Figura B.6 Tela do terminal com o comando de geração do arquivo de saída da abordagem proposta.....	186
Figura B.7 Exemplo do arquivo de saída gerado pela abordagem proposta	186

Lista de Quadros

Quadro 2.1 Classificação de uma RSSF segundo a organização utilizada na rede (LOUREIRO <i>et al.</i> , 2003)	28
Quadro 2.2 Classificação de uma RSSF segundo o tipo de sensoriamento utilizado pelos nós (LOUREIRO <i>et al.</i> , 2003).....	28
Quadro 2.3 Classificação de uma RSSF segundo o tipo de processamento realizado nos nós (LOUREIRO <i>et al.</i> , 2003)	29
Quadro 2.4 Classificação de uma RSSF segundo o tipo de comunicação utilizada pelos nós (LOUREIRO <i>et al.</i> , 2003).....	30
Quadro 2.5: Consumo de energia de uma interface baseada no <i>IEEE 802.11b</i> , adaptado de CUNHA (2004).....	42
Quadro 2.6: Classes dos dispositivos <i>Bluetooth</i> , adaptado de NAMBA (2011)	43
Quadro 3.1 Resumo das vantagens e desvantagens das técnicas de buscas aleatórias guiadas.....	52
Quadro 4.1 Resumo dos trabalhos relacionados	82
Quadro 6.1 Descrição dos recursos computacionais de hardware e software do ambiente	98
Quadro 6.2 Parâmetros do AG utilizado nos experimentos propostos para este trabalho.....	99
Quadro 6.3 Parâmetros complementares de configuração do AG	106
Quadro 6.4 Parâmetros básicos de configuração do AG	107
Quadro 6.5 Resumo dos resultados dos experimentos.....	128
Quadro C.1 Configuração do simulador de RSSF.....	188

Lista de Tabelas

Tabela 3.1 Conversão de inteiro para binário e <i>Gray</i> , adaptada de CASTRO e ZUBEN (2012)	57
Tabela 3.2: Comparação da avaliação entre uma função pura e normalizada, adaptada de LINDEN (2006)	62
Tabela 6.1: Ganho do algoritmo em função do tamanho da população	100
Tabela 6.2: Ganho do algoritmo em função do método de seleção	102
Tabela 6.3: Ganho do algoritmo em função da probabilidade de mutação	104
Tabela 6.4 Médias do ganho do modelo e IC para cada metrica em comparação a uma configuração aleatória.....	114
Tabela 6.5 Médias do ganho do modelo e IC para cada metrica em comparação a uma configuração uniforme	114
Tabela 6.6 Média do somatório da energia residual em relação a presença de obstáculos no cenário	117
Tabela 6.7 Média do <i>TAFF</i> em relação aos obstáculos no cenário.....	119
Tabela 6.8 Resultado do teste para verificar se existem diferenças entre as médias da TPP.....	120
Tabela 6.9 Resultados do experimento com o algoritmo completamente aleatório	124
Tabela 6.10 Resultados do experimento com o algoritmo proposto.....	125
Tabela A.1 Medidas da análise descritivas dos testes para o tamanho da população	144
Tabela A.2 Coeficiente de correlação desempenho do algoritmo e tamanho da população	144
Tabela A.3 Resultado do teste de normalidade da distribuição para tamanho da população	146
Tabela A.4 Resultado do teste de homogeneidade da variância para o tamanho da população	146
Tabela A.5 Resultado do teste para verificar se existem diferenças entre as médias dos grupos	147
Tabela A.6 Ganho do algoritmo em função do tamanho da população.....	147
Tabela A.7 Medidas da análise descritivas dos testes para o método de seleção ..	149
Tabela A.8 Coeficiente de correlação ganho do algoritmo e o método de seleção.	149
Tabela A.9 Resultado do teste de normalidade da distribuição para método de seleção	150
Tabela A.10 Resultado do teste de homogeneidade da variância para o método de seleção	151
Tabela A.11 Resultado do teste para verificar se existem diferenças entre as médias dos grupos para o método de seleção	151
Tabela A.12 Ganho do algoritmo em função do método de seleção	152
Tabela A.13 Análise descritiva do ganho do algoritmo para o método de cruzamento	153

Tabela A.14 Coeficiente de correlação do Ganho do Algoritmo e a configuração de cruzamento	153
Tabela A.15 Resultado do teste de normalidade da distribuição para configuração de cruzamento	155
Tabela A.16 Resultado do teste de homogeneidade da variância para a configuração de cruzamento	155
Tabela A.17 Resultado do teste para verificar se existem diferenças entre as médias dos grupos para a configuração de cruzamento	155
Tabela A.18 Análise descritiva do ganho do algoritmo para a probabilidade de mutação	157
Tabela A.19 Coeficiente de correlação do ganho do algoritmo e a probabilidade de mutação	157
Tabela A.20 Resultado do teste de normalidade da distribuição para a probabilidade de mutação	159
Tabela A.21 Resultado do teste de homogeneidade da variância para a probabilidade de mutação.....	159
Tabela A.22 Resultado do teste para verificar diferenças entre as médias do ganho do algoritmo de cada valor de probabilidade de mutação	159
Tabela A.23 Ganho do algoritmo em função dos valores da probabilidade de mutação	160
Tabela A.24 Resultado do teste de normalidade da distribuição do somatório da energia residual	162
Tabela A.25 Resultado do teste de homogeneidade da variância do somatório da energia residual	162
Tabela A.26 Resultado do teste para verificar se existe diferenças entre o somatório da energia residual em relação a presença de obstáculos no cenário	163
Tabela A.27 Média do somatório da energia residual	164
Tabela A.28 Resultado do teste de normalidade da distribuição do TAFF.....	165
Tabela A.29 Resultado do teste de homogeneidade da variância do TAFF	165
Tabela A.30 Resultado do teste de normalidade da distribuição do logaritmo do TAFF.....	166
Tabela A.31 Resultado do teste de homogeneidade da variância do logaritmo do TAFF.....	166
Tabela A.32 Resultado do teste para verificar se existe diferenças entre as médias do TAFF	167
Tabela A.33 Média do TAFF	168
Tabela A.34 Resultado do teste de normalidade da distribuição da TPP.....	169
Tabela A.35 Resultado do teste de homogeneidade da variância da TPP	169
Tabela A.36 Resultado do teste de normalidade das transformações nos dados da TPP.....	170
Tabela A.37 Resultado do teste de homogeneidade da variância do logaritmo da TPP.....	170

Tabela A.38 Resultado do teste para verificar se existe diferenças entre as médias da TPP	170
Tabela A.39 Teste de Wilcoxon para determinar qual o grupo, com ou sem utilização de obstáculos no cenário, apresenta maior TPP	172
Tabela A.40 Resultado do teste de normalidade da distribuição com e sem obstáculos para o experimento do ganho no TAFF	173
Tabela A.41 Resultado do teste de normalidade da distribuição com e sem obstáculos para o experimento do ganho na TPP	173
Tabela A.42 Resultado do teste de homocedasticidade da variância da TPP para o experimento do ganho no TAFF e do ganho na TPP	174
Tabela A.43 Média do ganho do TAFF	174
Tabela A.44 Média do ganho da TPP	174
Tabela A.45 Valores dos parâmetro do modelo RSSF-AG	176
Tabela A.46 Estimativas finais para os parâmetros significativos no modelo	177
Tabela A.47 Valores dos erros dos parâmetro do modelo RSSF-AG	177

1 Introdução

A área de atuação das Redes de Sensores Sem Fios (RSSF) vem se expandindo rapidamente, aproveitando-se dos avanços ocorridos nos microprocessadores, nos materiais de sensoriamento, nos microsistemas eletromecânicos e na comunicação sem fio (CARVALHO et al., 2012; HENRIQUES, LOVISOLO e RUBINSTEIN, 2009).

As RSSF têm uma ampla gama de aplicações potenciais, dentre as quais se destacam: monitoramento ambiental, seja interno, como em prédios e residências, ou externo, como em plantações, encostas, morros ou oceanos; monitoramento industrial, como automação de refinarias e de siderúrgicas; detecção da presença de matérias, como gases venenosos ou radiação, movimentação de inimigos no campo de batalhas; sistemas médicos, responsáveis por monitorar o funcionamento de órgãos, dados fisiológicos humanos; tráfego urbano, para monitorar veículos, rodovias e malha viária (BERETTA et al., 2012; MIRANDA, 2011).

Normalmente, as RSSF são formadas por um grande número de nós sensores¹. Estes nós são compostos por 4 (quatro) módulos principais: módulo de processamento, módulo de comunicação, módulo de sensores e módulo de energia. Este último utiliza, geralmente, uma bateria como fonte de energia (ALVES et al., 2012; HENRIQUES, LOVISOLO e RUBINSTEIN, 2009).

Um dos problemas mais importantes dos nós sensores de uma RSSF está relacionado ao tempo de vida da rede, que é o tempo em que todos os nós indispensáveis de uma RSSF ainda estão ativos e enviando dados para o nó soverdouro², também conhecido como estação base ou *Access Point* (AP). Isto ocorre devido aos nós sensores possuírem uma fonte de energia limitada e com localização, frequentemente de difícil acesso, o que torna impraticável a substituição de suas baterias (HENRIQUES, LOVISOLO e RUBINSTEIN, 2009). Este tempo de vida está diretamente relacionado ao conceito de consumo de energia, que

¹ Nós sensores – dispositivos sem fio equipados com processador, rádio transceptor e uma variedade de sensores.

² Nó soverdouro – é o componente da RSSF responsável por receber as informações de todos os nós sensores e disponibilizá-los para o usuário.

caracteriza o quanto cada nó sensor gasta de energia para realizar suas funções de coleta, processamento e transmissão de dados. O consumo de energia dos nós de uma RSSF é, em grande parte, devido ao módulo de comunicação (SILVA *et al.*, 2010).

Outro problema das RSSF, que precisa ser levado em consideração, é a Qualidade de Serviço (QoS), que tem como objetivo medir a qualidade do serviço da rede, utilizando principalmente parâmetros como: taxa de cobertura da área de sensoriamento, atraso médio na transferência de dados entre os nós sensores e o nó soverdouro (atraso fim-a-fim) e a taxa de perda de pacotes na rede (SONG, WANG e PEI, 2012).

A distribuição dos nós sensores em uma RSSF, em muitos casos, é realizada de forma estocástica³, podendo ser por difusão simples, colocação constante ou colocação *r-random* (ISHIZUKA e AIDA, 2004). Esta forma é apropriada para a grande maioria das aplicações, pois está mais próxima da realidade. Outra forma de distribuição é a determinística, cujo uso é mais indicado para aplicações em plantas industriais ou quando é possível a colocação dos nós sensores de forma precisa nos locais previamente definidos. Uma maneira de determinar os locais de instalação dos nós sensores em uma distribuição determinística leva em consideração o aumento no tempo de vida da rede e também um ganho na QoS, utilizando principalmente parâmetros do tipo sensoriamento e disseminação (VERONA, 2010).

Nos parâmetros do tipo sensoriamento, são tratados assuntos relativos à forma como os dados são sensoreados. Um dos mais importantes dentre esses parâmetros é a densidade, ou seja, a quantidade de nós utilizados na RSSF. O controle sobre a densidade da rede evita dois problemas. O primeiro é o aumento do consumo de energia quando a rede tem um número excessivo de nós ativos, o que provoca o aumento do tráfego de dados na rede, congestionando-a e, conseqüentemente, aumentando o seu consumo de energia. O segundo problema é a deficiência de cobertura quando a quantidade de nós sensores é muito reduzida, o

³ Processos Estocásticos – São processos que têm origem em eventos não determinísticos que podem produzir resultados diferentes e imprevisíveis a cada nova execução.

que causa um monitoramento insatisfatório dos fenômenos que estão sendo observados pela aplicação da rede (VERONA, 2010).

Nos parâmetros do tipo disseminação, são tratados assuntos relativos à capacidade da rede de transmitir os dados coletados nos nós sensores até o AP, respeitando os requisitos de integridade e tempo necessários para a transferência.

A perda de pacotes nas RSSF é um problema recorrente. Isto ocorre principalmente devido à baixa qualidade dos canais sem fio, ao congestionamento de pacotes ao redor dos AP ou dos nós *cluster-header*⁴ (CH), nas redes hierárquicas, além das próprias falhas dos nós sensores devido a sua fragilidade. Apesar destes problemas, as RSSF devem garantir algum grau de confiabilidade quanto à entrega dos pacotes gerados nos nós sensores para manter a integridade da informação. Outro aspecto que torna o parâmetro da perda de pacotes importante é o fato de que, em algumas aplicações críticas, cada pacote gerado nos nós sensores deve, necessariamente, chegar à aplicação em execução no AP. Logo, neste caso, não pode haver nenhuma perda de pacote (JABOUR, 2009).

Os algoritmos de roteamento que não estão preparados para trabalhar com a instabilidade da topologia das RSSF tendem, em cenários realistas, a produzir um aumento elevado na latência da rede. Por exemplo, nos algoritmos baseados em árvore, quando um nó sensor, no momento do envio dos pacotes de dados, encontrar seu ancestral na fase de hibernação, provocará um atraso na transmissão. Em certos tipos de aplicação, cujo requisito de latência é extremamente rígido, o algoritmo não suprirá as suas necessidades (CUNHA *et al.*, 2010).

Numa aplicação hipotética, em que a RSSF é responsável por monitorar uma região de risco sujeita a deslizamento de terra, como as encontradas nas principais cidades do país, o sucesso desta aplicação será comprometido se o deslizamento ocorrer e a rede não conseguir alertar o fato a tempo. Uma das possíveis causas da falha da aplicação pode ser devido aos requisitos de QoS dessa rede não terem sido satisfeitos. Este fato pode ter ocorrido porque a área na qual se iniciou o deslizamento não estava coberta por nenhum dos nós sensores, caracterizando-se, assim, como um problema de cobertura ou densidade de nós.

⁴ *Cluster-header* – São nós de uma RSSF de arquitetura hierárquica que têm a função de coletar os dados gerados nos nós sensores e enviá-los à estação base.

Outras possíveis causas de falha na aplicação podem ocorrer devido a uma perda de informação durante a transferência dos dados do nó sensor para o nó AP ou ao aumento do tempo necessário para que esses dados cheguem ao nó AP, o que caracterizaria problemas com os parâmetros do tipo disseminação.

1.1 Motivação e Relevância do Trabalho

A evolução dos dispositivos físicos (*hardware*), bem como de protocolos e *firmwares* (*software*) desenvolvidos para as RSSF, possibilitou a expansão das áreas de aplicação para este tipo de serviço. Com essa expansão, os estudos relacionados ao aumento do tempo de vida das RSSF tornaram-se um pré-requisito importante para avaliar a viabilidade da implantação da rede.

Esses estudos têm como objetivo propor soluções para um dos principais problemas das RSSF: o consumo de energia dos nós sensores. Este problema e o tempo de vida útil da rede estão diretamente relacionados. Quanto maior o consumo de energia, menor será o tempo de vida útil da rede, mantendo-se a mesma quantidade de energia disponível (HENRIQUES, LOVISOLO e RUBINSTEIN, 2009). Isto explica a grande quantidade de métodos, modelos e algoritmos propostos para reduzir este consumo.

Muitos trabalhos que se propõem a solucionar o problema do tempo de vida das RSSF utilizam técnicas para melhorar os algoritmos de roteamento, conforme descrito em (BUTT *et al.*, 2012; DANTAS *et al.*, 2010; GONÇALVES *et al.*, 2007; INTANAGONWIWAT *et al.*, 2000; KOULALI *et al.*, 2012; MACEDO, 2006; SHAH; RABAEY, 2002; SILVA, 2009; YE *et al.*, 2002). Outros trabalhos tentam resolver o problema utilizando técnicas de formação de *cluster* de sensores (LI *et al.*, 2010), outros realizam uma redução na transmissão de dados, utilizando o processamento dos dados coletados nos nós sensores (BUDKE *et al.*, 2011) e o adormecimento de nós inativos (YE, HEIDEMANN e ESTRIN, 2002). No entanto, muitas destas propostas não consideram a importância de alguns parâmetros, como a taxa de cobertura, o atraso fim-a-fim e a taxa de perda de pacotes, ou não levam em consideração os efeitos dos obstáculos presentes nos cenários, quando da simulação das RSSF.

Portanto, neste trabalho, apresenta-se uma proposta para o aumento do tempo de vida das RSSF, sem penalizar outros parâmetros importantes de QoS.

Para tanto, são utilizadas técnicas para definição de cobertura e colocação de nós sensores, associadas aos algoritmos genéticos (AG), e levando em consideração a presença de obstáculos nos cenários.

1.2 Objetivo Geral

O objetivo deste trabalho é propor um modelo para determinar a melhor configuração de implantação dos nós sensores de uma RSSF utilizando AG como técnica de busca.

Essa configuração deve considerar as métricas: (i) aumento do tempo de vida da rede; (ii) a razão de cobertura; (iii) o tempo de propagação dos dados entre o nó sensor e o nó sorvedouro e (iv) as perdas de pacotes de informação, entre o nó sensor e o nó sorvedouro, devido a falhas na rede. Também será verificada a influência da presença de obstáculos nos cenários sobre as quatro métricas.

A melhor configuração de implantação é determinada a partir da maximização da função de avaliação do modelo. Essa maximização depende das quatro métricas que são consideradas importantes para o correto funcionamento da aplicação que utiliza as informações coletadas pela RSSF. Logo, a melhor configuração é obtida quando se tem: o maior somatório do nível de energia residual armazenado em cada nó sensor da RSSF; a maior área do cenário sendo coberta por, no mínimo, 1 (um) nó sensor da RSSF, que garante que o fenômeno em monitoramento seja detectado na maior parte da rede; a menor média de tempo gasto para que o fenômeno que é detectado por qualquer nó sensor chegue ao nó sorvedouro, portanto, a aplicação; e a menor quantidade de pacotes pedidos entre os nós sensores e o nó sorvedouro.

1.3 Objetivos Específicos

A partir do objetivo geral do trabalho, têm-se os seguintes objetivos específicos:

1. Propor um algoritmo de busca e otimização, para determinar a melhor configuração de implantação dos nós sensores em uma RSSF;
2. Identificar e configurar um simulador para RSSF com modelos para consumo de energia, parâmetros de QoS e obstáculos em cenários de simulação.

3. Utilizar o Algoritmo Genético (AG) como técnica de busca e otimização para determinar a melhor configuração de implantação dos nós da rede.
4. Verificar se a presença de obstáculos no cenário influencia no resultado da busca da melhor configuração de implantação dos nós sensores de uma RSSF.
5. Verificar a influência da quantidade de nós sensores utilizados na composição do cromossomo sobre o desempenho do AG.

1.4 Estrutura do Trabalho

Esta dissertação foi estruturada em sete capítulos, apresentadas da seguinte forma: (i) são apresentados, no Capítulo 2, conceitos sobre as RSSF, suas características, modelo funcional e arquitetura, bem como são abordados as principais limitações que afetam o funcionamento da rede; (ii) no Capítulo 3, é descrita a teoria sobre os AG, abordando os fundamentos da técnica de busca baseada nos algoritmos evolutivos, bem como os principais componentes e um esquema básico do algoritmo; (iii) o Capítulo 4 contém uma revisão dos trabalhos relacionados aos dois principais temas abordados neste trabalho, RSSF e AG; (iv) no Capítulo 5, é apresentado o modelo RSSF-AG desenvolvido neste trabalho; (v) o Capítulo 6 traz a análise dos resultados, obtidos a partir dos experimentos propostos para demonstrar o modelo RSSF-AG; e (vi) no Capítulo 7, são descritas as considerações finais deste trabalho, trazendo também algumas sugestões para trabalhos futuros.

2 Rede de Sensores Sem Fio

As RSSF podem ser consideradas como um tipo especial de redes móveis *ad hoc*, dado que são redes em que os nós se comunicam entre si, sem necessitar de uma infraestrutura (roteadores). Os nós de uma RSSF são responsáveis por coletar informações do ambiente no qual estão instalados, processar essas informações e transmiti-las para o nó sorvedouro da rede. No entanto, as RSSF possuem algumas características que as diferenciam das tradicionais redes *ad hoc*, tais como: possuir um grande número de nós distribuídos; realizar operações, normalmente, sem a interferência humana; possuir fonte de energia com a capacidade limitada; apresentar restrições quanto ao poder de processamento, memória e largura de banda de comunicação. Além disto, em muitos casos, a sua localização é de difícil acesso (LOUREIRO *et al.*, 2003).

RSSF também necessitam de mecanismo para autoconfiguração, automanutenção, auto-organização e autoproteção, pois, em algumas aplicações, os elementos (nós) são implantados em áreas remotas e de difícil acesso (GOMES *et al.*, 2012). Assim, essas restrições tornam inadequada a utilização de algoritmos e protocolos de comunicação desenvolvidos para redes de computadores tradicionais (LOUREIRO *et al.*, 2003).

Segundo Chiang *et al.* (2004), os principais requisitos para as RSSF são:

- Baixo consumo de energia, uma vez que, normalmente, os nós sensores utilizam baterias como fonte de energia e a substituição das baterias geralmente não são possíveis. Assim, a busca pela minimização do consumo de energia torna-se necessária;
- Escalabilidade. Uma RSSF pode comportar de dezenas a milhares de nós sensores. Logo, os protocolos devem estar preparados para lidar com essa variação na quantidade de nós sensores;
- Densidade da distribuição dos nós na rede, que está relacionada à quantidade de nós e ao tamanho da área a ser monitorada. As redes podem se apresentar bastante densa, quando existem muitos nós agrupados em um espaço reduzido, ou dispersas, quando são usados poucos nós.

- Auto-organização, que tem por objetivo tornar a rede mais robusta. As RSSF podem ser grandes em tamanho e estar em ambientes hostis, o que leva ao aumento do número de falhas individuais dos nós sensores. Assim, se torna essencial usar mecanismos para reorganizar a rede depois dessas falhas.

2.1 Caracterização das RSSF

Uma RSSF, basicamente, tem sua caracterização dependente da aplicação que ela suportará, dado que a aplicação influenciará diretamente nas funções exercidas pelos elementos da rede (LOUREIRO *et al.*, 2003). Pode-se classificar uma RSSF segundo os tipos de sensoriamento, processamento e comunicação realizados por cada nó na rede. A configuração também caracteriza uma possível classificação das RSSF e será detalhada no Quadro 2.1.

Os nós de uma RSSF são normalmente distribuídos de forma estocástica, podendo ser por difusão simples, colocação constante ou colocação *r-random*. Esta forma é apropriada para a grande maioria das aplicações, pois está mais próxima da realidade. Outra forma de distribuição é a determinística, cujo uso é mais indicado para aplicações em plantas industriais (ISHIZUKA e AIDA, 2004).

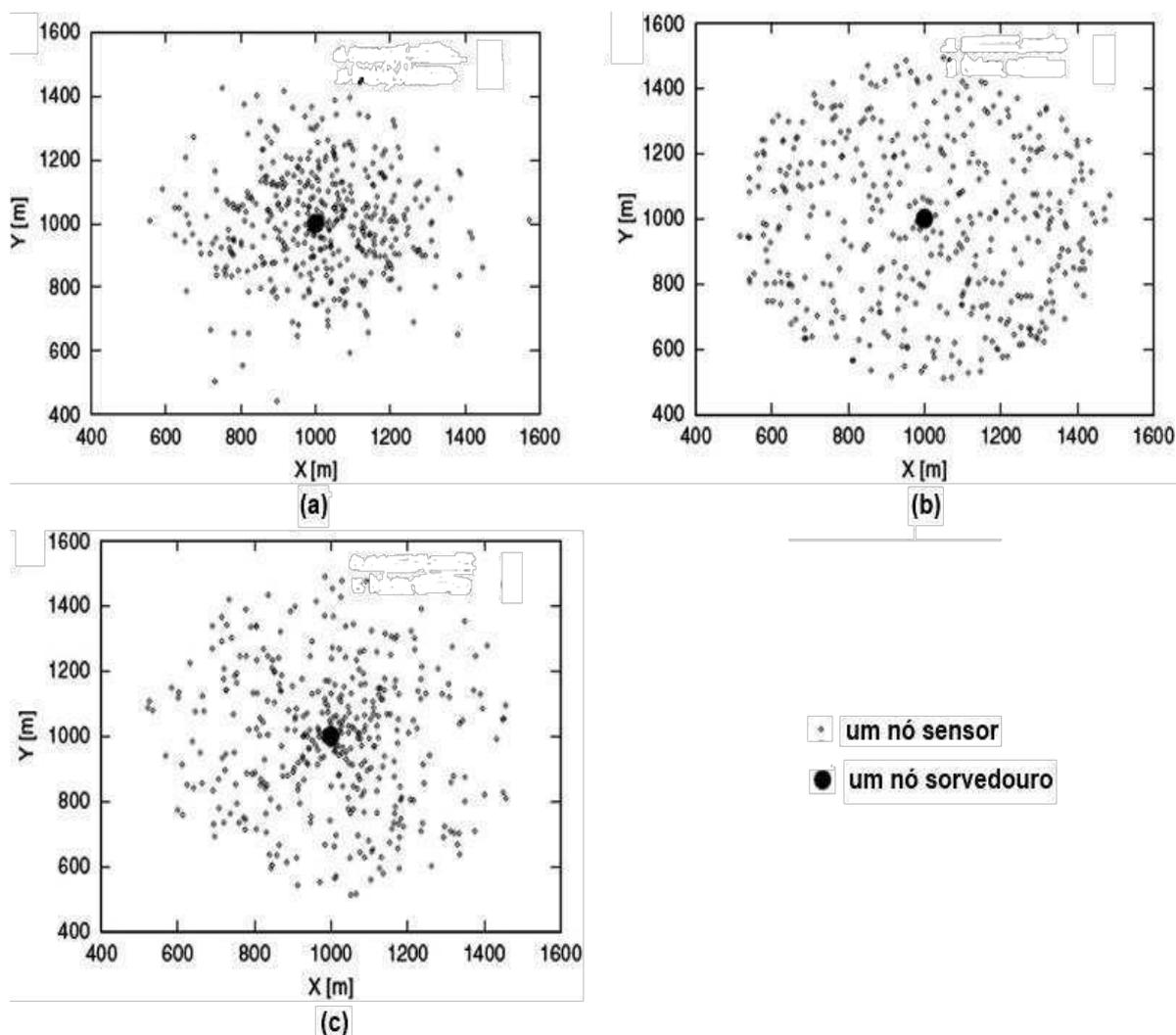
Na distribuição por difusão simples, os nós da RSSF são lançados no cenário de forma aleatória e, dependendo do seu peso, a distribuição pode ser descrita com base na solução da equação de uma distribuição normal bidimensional (ISHIZUKA e AIDA, 2004).

Na distribuição por colocação constante, a implantação é realizada de forma a manter constante a densidade dos nós sensores no cenário. Na implantação utilizando *r-random*, os nós sensores são uniformemente lançados obedecendo as direções radiais e angulares, a partir do nó sorvedouro (Ishizuka e Aida, 2004; Younis e Akkaya, 2008). Na Figura 2.1, são ilustradas as distribuições por: (a) difusão simples, (b) colocação constante e (c) *r-random*.

A classificação segundo o tipo de sensoriamento varia de acordo com as necessidades da aplicação, já que existem aplicações que necessitam de informações em tempo real, como no controle de processos industriais. Outras aplicações têm a coleta de dados programada em tempos pré-determinados

(OLIVEIRA, 2011). Portanto, é possível classificar as RSSF segundo o tipo de sensoriamento conforme o Quadro 2.2.

Figura 2.1 Exemplo da distribuição dos nós sensores por: (a) difusão simples, (b) constante; (c) *r-random*, adaptado de (ISHIZUKA e AIDA, 2004; NAMBA, 2011)



Para os parâmetros do tipo sensoriamento, são tratados assuntos relativos à forma como os dados são coletados. Um dos mais importantes desses parâmetros é a densidade, ou seja, a quantidade de nós utilizada na RSSF. O controle sobre a densidade da rede evita dois problemas. O primeiro, o aumento do consumo de energia quando a rede tem um número excessivo de nós ativos, o que provoca o aumento do tráfego de dados na rede, congestionando-a e, conseqüentemente, aumentando o seu consumo de energia. O segundo problema é a deficiência de cobertura quando a quantidade de nós sensores é muito reduzida, o que causa um

monitoramento insatisfatório dos fenômenos que estão sendo observados pela aplicação da rede (LIN e LIAO, 2010).

Quadro 2.1 Classificação de uma RSSF segundo a organização utilizada na rede (LOUREIRO *et al.*, 2003)

Classificação	Tipo	Descrição
Composição	Homogênea	São redes formadas por nós que tem a mesma configuração de <i>hardware</i> .
	Heterogênea	São redes formadas por dois ou mais tipos diferentes de configuração de <i>hardware</i> .
Distribuição	Irregular	Quando a distribuição dos nós no campo a ser monitorado é realizada de forma não uniforme.
	Regular	Quando a distribuição dos nós no campo a ser monitorado é realizada de forma uniforme.
Organização	Hierárquica	Quando os nós de uma rede são divididos em grupos para formar os <i>clusters</i> e em cada grupo é eleito um líder (<i>cluster-head</i>), que será responsável por estabelecer comunicação com o nó sorvedouro, formando uma hierarquia entre os grupos.
	Plana	Quando todos os nós têm acesso ao nó sorvedouro, sem intermediários.
Mobilidade	Estacionária	Quando todos os nós que compõem a rede não têm capacidade de locomoção e estarão fixos durante todo o funcionamento da rede.
	Móvel	Quando os nós têm capacidade de locomoção e durante a vida útil os mesmos são deslocados do seu local inicial.
Densidade	Balanceada	Quando a concentração e distribuição dos nós por unidade de área em uma rede estiverem próximo do ideal.
	Densa	Quando a concentração e distribuição dos nós por unidade de área estiverem acima do valor ideal.
	Esparsa	Quando a concentração e distribuição dos nós por unidade de área estiverem abaixo do valor ideal.

Quadro 2.2 Classificação de uma RSSF segundo o tipo de sensoriamento utilizado pelos nós (LOUREIRO *et al.*, 2003)

Classificação	Tipo	Descrição
Coleta	Periódica	Os dados são coletados em intervalos de tempo pré-configurados. Apresenta a vantagem de poder desligar o nó durante o intervalo de cada coleta.
	Contínua	Os dados são coletados pelos nós sensores ininterruptamente e enviado para o nó sorvedouro.
	Reativa	Os dados são adquiridos no momento da ocorrência dos eventos de interesse da aplicação. Têm-se como exemplo um sistema de segurança.
	Tempo Real	Os dados são capturados no momento da ocorrência do evento, mas o seu envio para a aplicação precisa ser feito imediatamente após a coleta.

O processamento que o nó realiza sobre os dados coletados antes de enviá-los é mais uma forma de se classificar as RSSF. Esta classificação leva em

consideração a forma de cooperação utilizada no processamento, conforme o Quadro 2.3 (LOUREIRO *et al.*, 2003).

Quadro 2.3 Classificação de uma RSSF segundo o tipo de processamento realizado nos nós (LOUREIRO *et al.*, 2003)

Classificação	Tipo	Descrição
Cooperação	Infraestrutura	Os nós sensores executam funções de roteamento, algoritmo de controle de acesso ao meio, escolha de líder de grupo, criptográfica dos dados coletados e determinação de localização.
	Localizada	Os nós sensores executam, além das funções relacionadas ao modo infraestrutura, algum tipo de processamento local, como a tradução dos dados coletados pelos sensores para uma determinada unidade física.
	Correlação	Os nós sensores podem executar vários tipos de tratamentos de dados, como supressão seletiva, contagem, fusão, compressão, agregação e multi-resolução.

A classificação de uma RSSF baseada no tipo de comunicação é dividida de acordo com o tipo de disseminação, o tipo de conexão, o tipo de transmissão, a alocação do canal e o fluxo de informação (LOUREIRO *et al.*, 2003). No Quadro 2.4, é possível observar esta classificação de forma mais detalhada.

Nos parâmetros do tipo disseminação, são tratados assuntos relativos à capacidade da rede transmitir os dados coletados nos nós sensores até o AP, respeitando os requisitos de integridade e tempo que são necessários para a transferência.

Quadro 2.4 Classificação de uma RSSF segundo o tipo de comunicação utilizada pelos nós (LOUREIRO *et al.*, 2003)

Classificação	Tipo	Descrição
Disseminação	Programada	Os nós estão configurados para disseminar os dados em intervalos regulares programados na aplicação.
	Contínua	Os nós estão configurados para disseminar os dados de forma contínua.
	Sob Demanda	Os nós estão configurados para disseminar os dados de acordo com a ocorrência de algum evento pré-definido na aplicação.
Tipo de Conexão	Simétrica	Todos os nós possuem o mesmo raio de alcance da conexão.
	Assimétrica	O raio de alcance da conexão dos nós é diferente.
Transmissão	Simplex	Utilizam transceptores do tipo simplex, que só possuem capacidade de transmitir informação, não recebendo informação de outros nós.
	<i>Half-duplex</i>	Utilizam transceptores do tipo <i>half-duplex</i> , que podem transmitir ou receber informações, mas limitada a uma função por vez.
Alocação de canal	Estática	Os recursos de largura de banda do canal são divididos em iguais partes para todos os elementos da rede. Os recursos são: Frequência - <i>Frequency Division Multiple Access (FDMA)</i> ; Tempo - <i>Time Division Multiple Access (TDMA)</i> ; Código - <i>Code Division Multiple Access (CDMA)</i> ; Espaço - <i>Space Division Multiple Access (SDMA)</i> ; Ortogonal - <i>Orthogonal Frequency Division Multiplexing (OFDM)</i> .
	Dinâmica	Os recursos da largura de banda do canal são disputados por todos os elementos da rede.
Fluxo de Informação	<i>Flooding</i>	Os nós utilizam a técnica de transmissão em broadcast, cooperativamente, até que a informação alcance o nó sorvedouro da rede.
	<i>Multicast</i>	Neste tipo de rede os nós formam grupos e utilizamos a técnica <i>multicast</i> para distribuir as informações para múltiplos nós simultaneamente.
	<i>Unicast</i>	Os nós estabelecem comunicação diretamente com o ponto de acesso (nó sorvedouro) utilizando protocolos de roteamento do tipo multi-saltos.
	<i>Gossiping</i>	Os nós escolhem previamente para quais nós serão transmitidos as informações.
	<i>Bargaining</i>	Os nós somente enviam as informações para nós que as tenham solicitado, necessitando previamente de um processo de negociação.

2.2 Arquitetura de uma Rede de Sensores Sem Fio

Os dispositivos utilizados em uma RSSF possuem uma arquitetura comum composta por: sensoriamento, processamento e comunicação, que são essenciais para a implementação da rede.

A RSSF basicamente é composta por um dispositivo, conhecido como nós ou *nodes*. Os nós podem ser do tipo sensores ou do tipo sorvedouros (*sink node*) (CHEN e YU, 2010).

Os nós sensores são dispositivos com capacidade de sensoriamento, processamento e comunicação. Segundo Loureiro *et al.*, (2003), eles são distribuídos na área de interesse e ficam responsáveis por:

- Coletar os dados utilizando seus sensores;
- Processar esses dados localmente ou coordenadamente com outros nós sensores vizinhos;
- Enviar a informação para a aplicação utilizando geralmente o nó sorvedouro.

No mercado, existem vários fabricantes e modelos de nós sensores, que geralmente são projetados visando sua utilização para uma determinada aplicação. No entanto, eles compartilham uma estrutura básica que é formada por um transceptor, um processador, uma memória, um sensor e uma fonte de energia (OLIVEIRA, 2011). Na Figura 2.2 é apresentada a estrutura básica de um nó sensor.

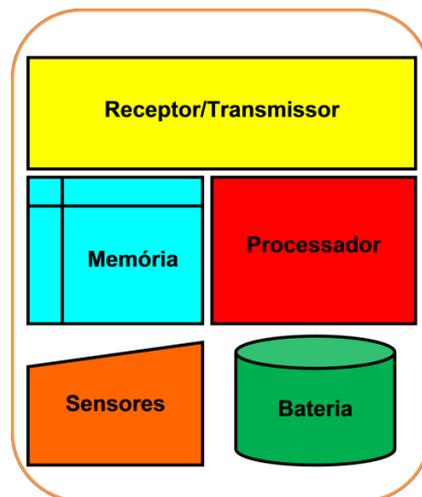
A estrutura de um nó sensor é composta pelos seguintes blocos:

- *Transceptor*: unidade formada por um transmissor e um receptor. Sua função no dispositivo é fornecer um meio de comunicação entre os nós para receber ou enviar os dados coletados pela unidade de sensor do dispositivo;
- *Processador*: unidade com a função de controlar o funcionamento de todo o dispositivo. Ele é o responsável por ativar a coleta de dados da unidade do sensor, armazenar os dados coletados na unidade de memória e enviar esses dados a partir da unidade de rádio (transceptor). Alguns dispositivos têm uma unidade de

processamento com maior capacidade que é capaz de realizar um tratamento nos dados antes de enviá-los;

- *Memória*: unidade com a função de guardar os dados coletados pela unidade de sensor. Os dados armazenados ficarão disponíveis para serem processados e/ou disseminados pela rede;
- *Sensor*: unidade responsável pela coleta dos dados, transformando a grandeza física (temperatura, umidade do ar, pressão atmosférica, gases, radiação luminosa, etc.) em sinais elétricos, que serão codificados e armazenados na memória;
- *Fonte de energia*: unidade responsável por fornecer energia para as demais unidades do dispositivo. Na grande maioria das redes, essa unidade utiliza uma bateria para fornecer a energia. No entanto, outros tipos de fonte de energia podem ser utilizados.

Figura 2.2 Estrutura básica do *hardware* de um nó sensor, adaptado de LOUREIRO et al. (2003)

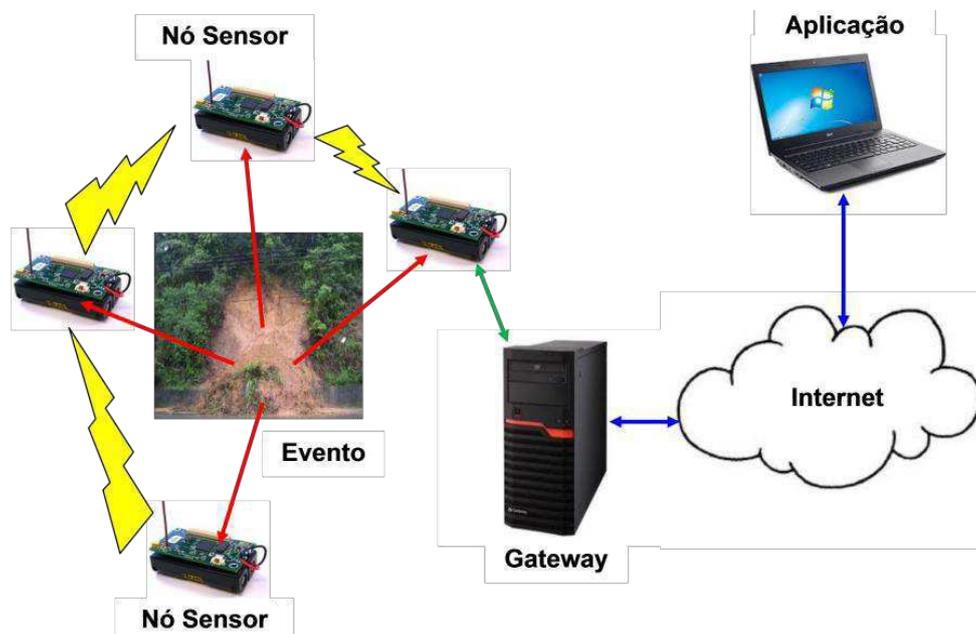


O nó sorvedouro em uma RSSF é responsável por concentrar os dados enviados por todos os nós sensores e repassá-los para a aplicação. Ele funciona como um *gateway*⁵ entre a rede de sensores e a rede de computadores que está executando a aplicação, podendo ser a *internet* ou qualquer outro tipo de *link* de dados (OLIVEIRA, 2011). Na Figura 2.3, é

⁵ *Gateway* – Um computador que executa conversão de protocolo entre tipos diferentes de redes ou aplicativos. Um *gateway* pode conectar uma rede local de computadores pessoais a uma RSSF (FREEDMAN, 1995).

apresentado um modelo genérico de uma RSSF interligada com uma rede do tipo *internet*, por meio do nó sorvedouro.

Figura 2.3 Modelo genérico de uma RSSF, adaptado de LOUREIRO et al. (2003)

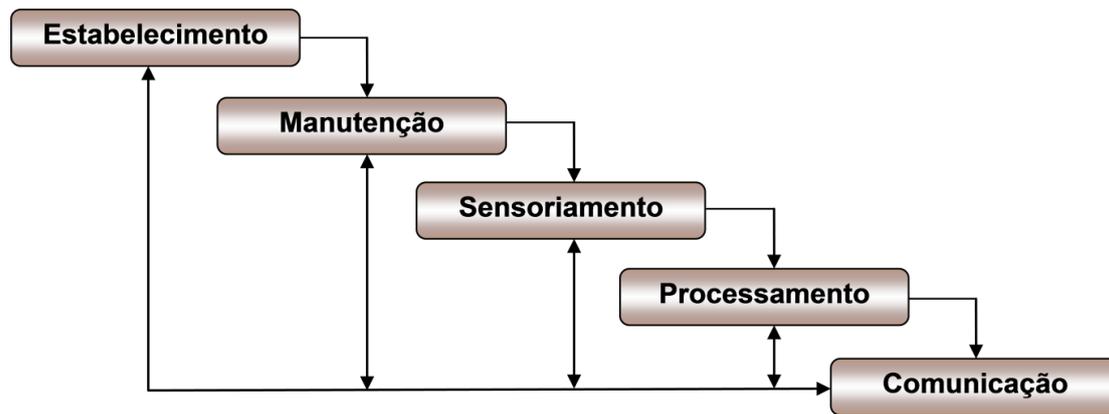


Os nós sensores trocam informações de forma colaborativa, fazendo o processamento dessas informações antes da disseminação, que pode ocorrer a partir de vários saltos até que se chegue ao nó sorvedouro. A partir do nó sorvedouro, a informação chega ao computador que executa a aplicação utilizando a internet. Essa é apenas uma das formas de se fazer a interconexão dos nós sensores com a aplicação, mas existem várias outras formas, como a utilização do padrão *IEEE 1451* (JONES, 2006) que define uma interface de comunicação para transdutores (LOUREIRO *et al.*, 2003).

2.3 Modelo Funcional de uma RSSF

Ruiz (2003) propôs as principais funcionalidades de RSSF, divididas em 5 (cinco) grupos de atividades realizadas durante o ciclo de vida de um nó sensor: estabelecimento da rede; manutenção; sensoriamento; processamento; e comunicação. Na Figura 2.4 é apresentada a divisão das funcionalidades propostas.

Figura 2.4 Ciclo de vida da RSSF, adaptada de LOUREIRO et al. (2003)



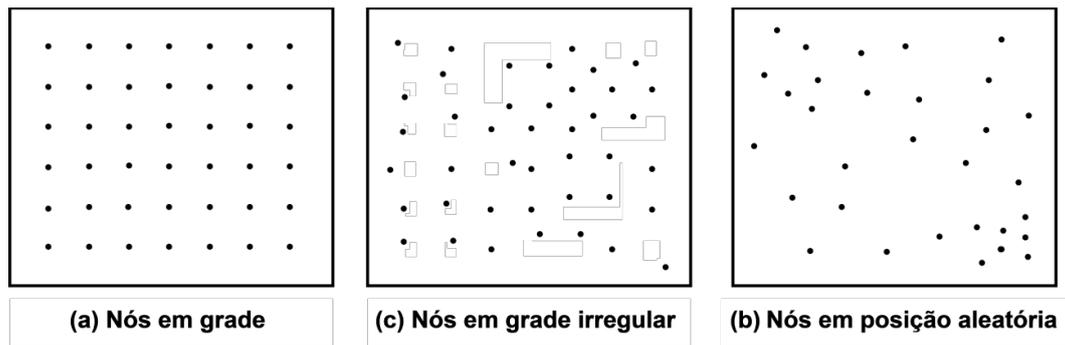
2.3.1 Estabelecimento

A fase de estabelecimento de uma RSSF envolve atividades de implantação dos nós sensores na área a ser monitorada e o estabelecimento da rede. Na implantação da rede, podem ser utilizadas basicamente duas técnicas: distribuição determinística ou distribuição não determinística (NAKAMURA, 2003).

Na técnica de distribuição determinística, os nós sensores são implantados seguindo uma disposição pré-determinada, que visa à obtenção do aumento do tempo de vida da rede, bem como a melhoria na QoS. Essa disposição pode ser regular, em forma de uma grade, conforme ilustrado na Figura 2.5a (NAKAMURA, 2003), ou utilizando os resultados de um algoritmo de busca, como proposto neste trabalho.

Na técnica de distribuição não determinística, é possível utilizar 2 (duas) estratégias para o lançamento dos nós sensores. Na primeira, os nós sensores podem ser lançados de um avião aleatoriamente. Nesta estratégia, não se tem controle sobre a densidade dos nós na área monitorada, o que pode aumentar, em algumas áreas, a concentração de nós sensores, enquanto outras ficarão sem cobertura, conforme ilustrado na Figura 2.5b. Na segunda estratégia, os nós sensores são lançados de forma controlada, seguindo um padrão de cobertura determinística. No entanto, devido à influência de fatores, como obstáculos, velocidade do avião e do vento, e danos aos sensores quando atingem o chão, a formação da rede de sensores se dará de forma irregular, como ilustrado na Figura 2.5c (NAKAMURA, 2003).

Figura 2.5 Modelos de posicionamento dos nós sensores em uma RSSF, adaptado de NAKAMURA (2003)



Antes de iniciarem o sensoriamento, os nós sensores precisam realizar a atividade de estabelecimento da rede, que ocorre logo após eles serem posicionados na área a ser monitorada. As RSSF são formadas por dispositivos com capacidade de auto-organização, criando espontaneamente uma rede não pré-defina, que pode agrupar-se e adaptar-se dinamicamente para explorar a redundância resultante da alta densidade de nós sensores em uma região (LOUREIRO *et al.*, 2003).

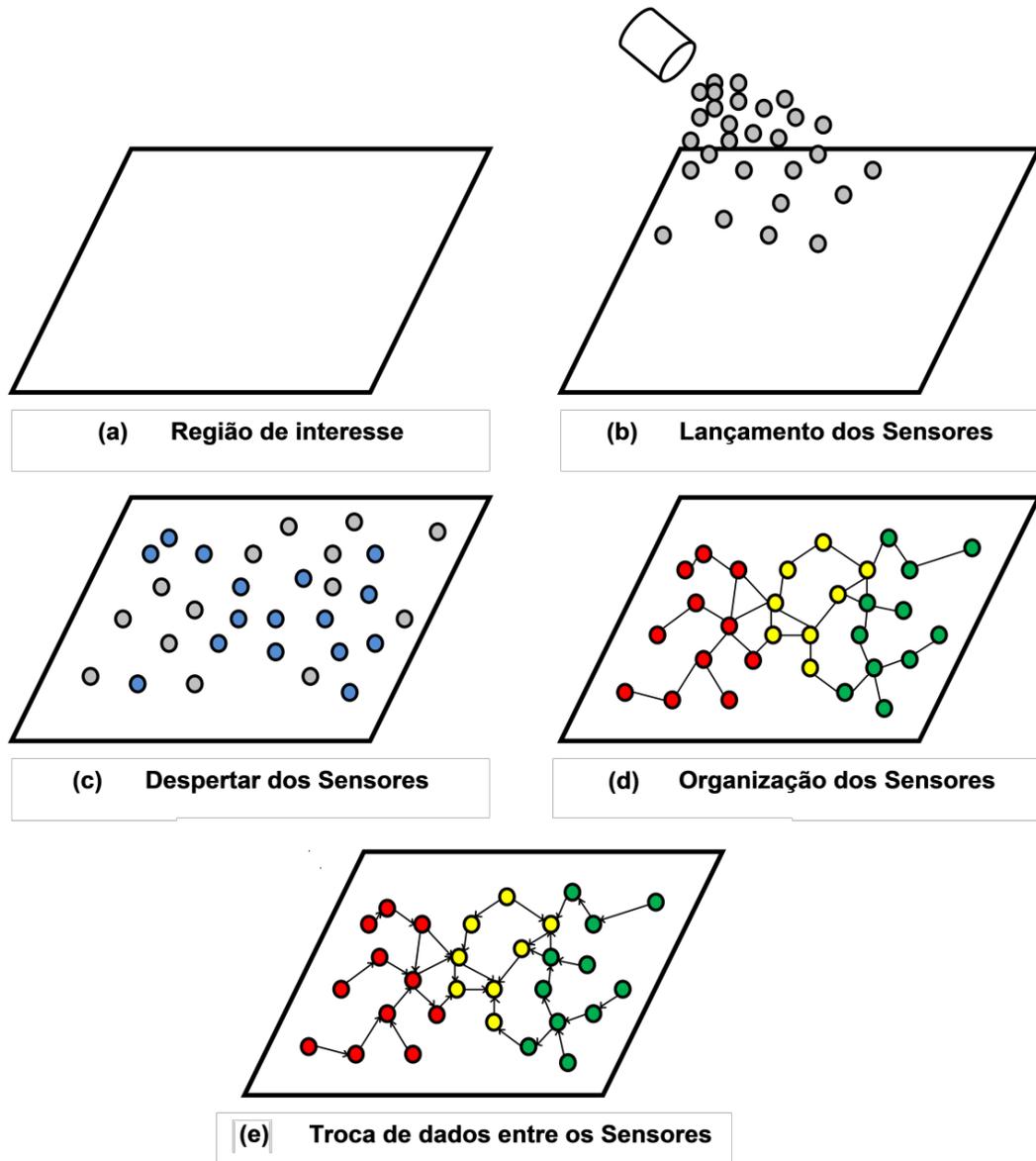
Na Figura 2.6, ilustram-se as 5 (cinco) fases do estabelecimento da rede de sensores. Inicialmente, tem-se a região de interesse a ser monitorada que precisa ser povoada com os nós sensores (Figura 2.6a). Na segunda fase, os nós sensores são lançados na região de interesse, na qual podem ser utilizadas técnicas de distribuição determinísticas ou não determinísticas (Figura 2.6b). Na terceira fase, os nós sensores são despertados (Figura 2.6c). Na quarta fase, após os nós sensores terem sido ativados, se dá início à organização da rede, na qual ocorre a descoberta de sua localização (Figura 2.6d). Por fim, na última fase, os nós sensores dão início à troca de dados (Figura 2.6e) (LOUREIRO *et al.*, 2003).

2.3.2 Manutenção

A manutenção é a fase que tem como principal objetivo prolongar o tempo de vida útil da RSSF, além da redução de ocorrências de problemas que possam comprometer, de forma parcial ou total, o funcionamento da aplicação. Um desses problemas é a baixa densidade de nós sensores, que provoca o esgotamento das fontes de energia prematuramente. A fase de manutenção pode atuar de forma preventiva, adaptativa, corretiva ou reativa, dependendo

do estado em que se encontra a rede e dos seus nós sensores (NAMBA, 2011).

Figura 2.6 Fases de estabelecimento da rede de sensores, adaptado de LOUREIRO et al. (2003)



A fase de manutenção também está presentes nas demais fases do ciclo de vida das RSSF: estabelecimento, sensoriamento, processamento e comunicação. Ela é funcional durante toda a vida da rede, com o objetivo de proporcionar a recuperação de falhas, que ocorrem com frequência devido à fragilidade e capacidade reduzida de energia dos nós sensores. Esta ocorrência de falhas pode levar a rede à realização de uma redistribuição dos nós sensores e a uma nova organização (LOUREIRO *et al.*, 2003).

2.3.3 Sensoriamento

A fase de sensoriamento no ciclo de vida das RSSF é responsável por reconhecer o ambiente no qual a rede está instalada e por realizar a coleta de dados. Levando em consideração os parâmetros definidos pela aplicação e os tipos de nós sensores utilizados, é determinada a distância do alvo, os tipos de obstáculos, os ruídos do ambiente que possam interferir na coleta dos dados, o tipo de dado a ser coletado, o volume de dados e os períodos de amostragem necessários para o correto funcionamento da aplicação (LOUREIRO *et al.*, 2003).

A determinação das áreas de sobreposição dos nós sensores também é outra atividade importante realizada na fase de sensoriamento. A localização dessas áreas poderá alterar o estado dos nós sensores que estejam numa mesma área, fazendo com que um deles seja colocado no estado desativado (*idle*), pois só se faz necessária a coleta de dados de apenas um deles para que a aplicação funcione corretamente. Essa atividade resultará na economia de energia e, conseqüentemente, no aumento do tempo de vida da rede.

Outra atividade realizada pela fase de sensoriamento é a verificação constante da quantidade e distribuição dos nós sensores ativos na área monitorada e se eles atendem aos requisitos da aplicação, uma vez que fatores como a fragilidade do dispositivo, a capacidade reduzida de comunicação, o processamento e a energia dos nós, podem levar os nós à inoperância temporária ou definitiva. Neste caso, a rede realiza rearranjo dos nós sensores para se manter ativa até que não seja mais possível manter a cobertura para qual foi projetada (NAMBA, 2011).

2.3.4 Processamento

A fase de processamento nas RSSF é dividida basicamente em 2 (dois) grupos: o processamento de suporte da rede e o processamento de informações coletadas pela rede.

O processamento de suporte abrange todo o processamento funcional dos sensores, que inclui a execução de procedimentos relacionados à infraestrutura da rede, tais como: gerenciamento por meio dos protocolos de roteamento; eleição de líderes de *cluster* e descoberta de localização;

comunicação com os algoritmos de controle de acesso ao meio; e manutenção da rede de sensores.

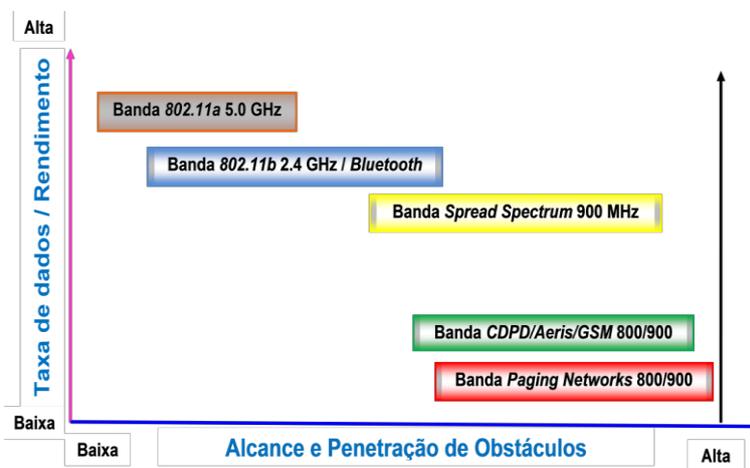
O segundo grupo de processamento tem como foco o tratamento dos dados coletados pelo sensor. Estes dados podem ser submetidos a procedimentos de correlação de dados, como fusão, compressão, agregação, supressão seletiva, criptografias, dentre outros processamentos. Os dados também podem ser submetidos a processamentos básicos, como a tradução dos dados para uma unidade física diferente.

2.3.5 Comunicação

A fase de comunicação no modelo funcional das RSSF é responsável por estabelecer o elo entre os dispositivos (nó sensor) com a rede responsável por conectá-los à *internet*. A RSSF apresenta algumas características que a difere das demais redes sem fio, como a grande quantidade de elementos de rede que, na maioria das aplicações, são estacionários. No entanto, a topologia das redes de sensores é bastante instável, devido aos recursos limitados de energia e à fragilidade dos elementos.

As RSSF utilizam vários tipos de tecnologia de comunicação, como *IEEE 802.11 (Wi-Fi)*, *IEEE 802.15.1 (Bluetooth)* e *IEEE 802.15.4 (ZigBee)*. Cada uma destas tecnologias tem suas vantagens e possuem limitações quanto a obstáculos e faixas de alcance (Figura 2.7), sendo o tipo de aplicação e o ambiente no qual a rede está instalada, o fator determinante na escolha de qual tecnologia será utilizada.

Figura 2.7 Limitações quanto a obstáculos e faixa de alcance, adaptada de LOUREIRO et al. (2003)



2.4 Energia em Rede de Sensores Sem Fio

Nos últimos anos, os dispositivos das RSSF sofreram uma grande evolução. No entanto, a preocupação com o consumo de energia ainda é um dos problemas que os pesquisadores buscam resolver. Na grande maioria das aplicações, os nós sensores utilizados têm como fonte de energia as baterias, que são ainda bastante limitadas. Além disto, em algumas aplicações, os nós sensores são lançados em locais de difícil acesso, o que torna a manutenção desses dispositivos praticamente inviáveis, seja do ponto vista de restrição de acesso ao dispositivo ou do custo financeiro para realização da manutenção (NAMBA, 2011).

Nesse cenário, o tempo de vida da rede depende de vários fatores, tais como: a densidade de nós sensores da rede (MELO, 2009); a quantidade de energia disponível em cada nó sensor da rede (LOUREIRO et al., 2003); o tipo de tecnologia de rádio utilizada, entre vários outros fatores. No entanto, os estudos sobre os protocolos de roteamento e controle de acesso ao meio tem sido alvo de muitas pesquisas com o objetivo de aumentar o tempo de vida das RSSF (CORREIA et al., 2005).

Uma das técnicas utilizadas para reduzir o consumo de energia das RSSF é baseada no aprimoramento de suas características de configuração, comunicação e processamento. No aprimoramento da configuração, foram propostas mudanças nos protocolos de roteamento, dando origem aos seguintes protocolos (YE, HEIDEMANN e ESTRIN, 2002) (DANTAS et al., 2010):

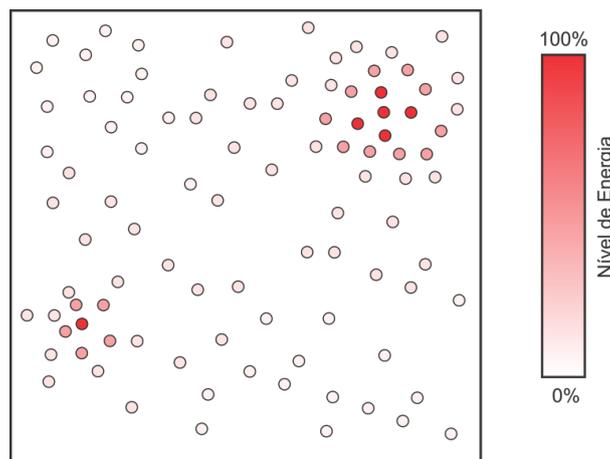
- *Sensor – Medium Access Control (S-MAC)*, que modifica a camada de controle de acesso ao meio *Medium Access Control (MAC)*;
- *Sensor Protocols for Information via Negotiation (SPIN)*, para redes com organização plana;
- *Low-Energy Adaptive Clustering Hierarchy (LEACH)*, para organização hierárquica;
- *Greedy Perimeter Stateless Routing (GPSR)*, que utiliza informações sobre a localização do nó sensor.

O Controle de Potência de Transmissão (CPT) em RSSF é umas das técnicas utilizadas para reduzir o consumo de energia. Kostin e Amorim (2007) propuseram um método que permite o CPT em cada nó sensor. A forma como os dados são processados nas RSSF também foi modificada para se obter a

redução do consumo de energia. No trabalho realizado por Pinto *et al.* (2007), é proposta a agregação ou fusão de dados em tempo real, que consiste em receber as informações de vários nós sensores próximos e processá-las para realizar apenas uma transmissão. Matos *et al.* (2008) sugerem a agregação com predição de dados, ou seja, uma extensão dos mecanismos de agregação que incorporam uma função de predição de dados, baseada em regressão linear aplicada aos dados que já foram recebidos.

Outras técnicas importantes que visam o aumento do tempo de vida da rede são baseadas no balanceamento da quantidade de energia e fazem uso do mapa de energia da RSSF, que é construído a partir do levantamento da quantidade de energia disponível em cada região da RSSF. Na Figura 2.8, tem-se um exemplo de um mapa de energia de uma RSSF. As regiões mais escuras indicam que o nível de energia é maior, ou seja, a concentração de nós com energia próxima a 100% é maior. Nas regiões mais claras, o nível de energia é menor, estando mais próximo do esgotamento da fonte de energia.

Figura 2.8 Representação gráfica do mapa de energia, adaptada de NAMBA (2011)



2.5 Protocolos de Acesso ao Meio para RSSF

Devido às limitações de processamento, memória e energia impostas pelos nós sensores, se faz necessário que os componentes de *hardware* e de *software* busquem minimizar a quantidade de energia consumida. Nos dispositivos sensores atuais, o módulo de memória, que emprega a tecnologia *flash*, e o módulo de rádio transceptor são os de maior consumo de energia (MACEDO, 2006). Com o objetivo de controlar e minimizar o uso destes módulos, foi proposto a utilização de protocolos de acesso ao meio, dentre os quais

(NAMBA, 2011): *IEEE 802.11 (Wi-Fi)*; o *IEEE 802.15.1 (Bluetooth)*; e o *IEEE 802.15.4 (ZigBee)*.

As técnicas de alocação de canal utilizadas nos protocolos de acesso ao meio são (MACEDO, 2010):

- Acesso Múltiplo por Divisão de Frequência ou *Frequency Division Multiple Access (FDMA)*;
- Acesso Múltiplo por Divisão de Tempo ou *Time Division Multiple Access (TDMA)*;
- Acesso Múltiplo por Divisão de Código ou *Code Division Multiple Access (CDMA)*;
- Acesso Múltiplo com Verificação de Portadora com Prevenção de Colisão ou *Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)*.

2.5.1 Padrão IEEE 802.11 (Wi-Fi)

O padrão *IEEE 802.11*, também conhecido como *Wi-Fi*, *Wireless* ou *Ethernet* sem fio, foi proposto com o objetivo de estabelecer um padrão de rede local sem fio para comunicação de dados. Inicialmente, com taxas de transferências de 1 Mbits/s ou 2 Mbits/s, o padrão *IEEE 802.11* foi proposto pelo *Institute of Electrical and Electronics Engineers (IEEE)* durante a década de 90. Em 1990, o *IEEE* criou o comitê para definir um padrão de conectividade sem fio, que só após 7 (sete) anos, em 1997, aprovou o padrão *IEEE 802.11*. Atualmente, é o padrão mais utilizado por dispositivos móveis, como *notebooks*, *tablets*, *smartphones*, dentre outros. Foi também o responsável por tornar possível a comunicação entre equipamentos e dispositivos de diferentes fabricantes.

A utilização do *IEEE 802.11* em RSSF se deve à quantidade de fabricantes e de dispositivos de comunicação de dados disponíveis no mercado que são baseados neste padrão. Ele utiliza a técnica de acesso ao meio *CSMA/CA* com 2 (dois) modos de operação: o *CSMA/CA* comum e o *CSMA/CA* com *Request to send/Clear to Send (RTS/CTS)*. Este último modo de operação foi proposto com o objetivo de resolver o problema do terminal escondido, que é comum em redes com um grande número de nós sensores e nas quais existem obstáculos.

A banda de frequência utilizada pelo *IEEE 802.11*, denominada *ISM*, é de livre acesso, que trabalha na faixa de 2.4 GHz ou 5 GHz e aplica a técnica *Direct Sequence Spread Spectrum (DSSS)* para minimizar os efeitos dos ruídos. Com a adoção desse padrão por toda indústria, o modelo foi modificado para atender as demandas de velocidade e qualidade, surgindo os padrões: *IEEE 802.11b*; *IEEE 802.11g*; *IEEE 802.11a*; *IEEE 802.11n* (LOUREIRO *et al.*, 2003).

A interface de rede baseada no padrão *IEEE 802.11* é, dentre todas as outras interfaces que fazem uso dos padrões de interfaces utilizados em RSSF, a que apresenta o maior consumo de energia e o maior preço. Portanto, seu emprego é indicado para aplicações que necessitem transmissão de áudio ou vídeo. No Quadro 2.5, é apresentado o consumo de energia de uma interface *IEEE 802.11b*.

Quadro 2.5: Consumo de energia de uma interface baseada no *IEEE 802.11b*, adaptado de CUNHA (2004)

Estado	Potência (W)
Dormência	0,050
Ocioso	0,740
Recepção	0,900
Transmissão	1,350

2.5.2 Padrão IEEE 802.15.1 (*Bluetooth*)

O padrão *IEEE 802.15.1*, mais conhecido como *bluetooth*, é muito utilizado em redes pessoais, devido à praticidade no estabelecimento de conexão e troca de dados. Sua abrangência fica na faixa de 10 a 100 metros, o que o torna ideal para ser utilizados em dispositivos móveis como smartphones, fones de ouvido, aparelhos de som automotivos, dentre outros (NAMBA, 2011).

O padrão *bluetooth* foi proposto inicialmente pela *Ericsson Mobile Communications*, em 1994, com o objetivo de unir as tecnologias da computação com as das telecomunicações, utilizando uma interface de baixo custo e de baixo consumo. Com o padrão *bluetooth*, foi possível conectar acessórios, como fone de ouvido, aos telefones celulares. Em 1998, foi criado o grupo *Special Interest Group (SIG)* e, logo depois, em 2003, o grupo foi expandido chegando a um total de 1.790 fabricantes (NAMBA, 2011).

O desenvolvimento do padrão *bluetooth* tinha como objetivo a otimização dos dispositivos usando alguns princípios, como: a manipulação de

dados e voz; a utilização de banda aberta para eliminar interferências; o baixo custo de produção; e o baixo consumo de energia. Isso foi obtido diminuindo o número de estágios analógicos, fazendo uso, sempre que possível, de circuitos digitais. Esse objetivo também se enquadra aos adotados pelas RSSF (NAMBA, 2011).

Atualmente, o padrão *bluetooth* está dividido em 3 (três) classes apresentadas no Quadro 2.6. As classes se diferenciam pelo raio de alcance e pelo nível de potência da cobertura. A conexão entre os dispositivos ocorre de forma transparente e é estabelecida no momento em que dois ou mais dispositivos com tecnologia *bluetooth* estão dentro de uma mesma área de cobertura. Essa característica é extremamente importante, já que facilita o seu manuseio pelos usuários dos dispositivos (NAMBA, 2011).

Quadro 2.6: Classes dos dispositivos *Bluetooth*, adaptado de NAMBA (2011)

Classe	Potência	Nível de Potência (mW)	Cobertura (m)
1	Alta	100,0	até 100,0
2	Média	2,5	até 10,0
3	Baixa	1,0	0,1 a 10,0

Quando 2 (dois) ou, no máximo, 8 (oito) dispositivos com tecnologia *bluetooth* estão numa mesma área de cobertura, ocorre o estabelecimento de comunicação entre os dispositivos formando uma rede, que recebe o nome de *piconet* e é ilustrada na Figura 2.9. Elas são formadas por um dispositivo que assume o papel de dispositivo principal ou mestre e por vários dispositivos secundários ou escravos. Para tornar possível conectar além dos 8 (oito) dispositivos iniciais numa mesma rede, foi proposta a conexão de várias *piconets* formando uma grande rede, a qual foi nomeada como *scatternet* (NAMBA, 2011).

Na Figura 2.10, é possível observar que um mesmo dispositivo pode fazer parte de diferentes *piconets*. No entanto, o papel de mestre ou dispositivo principal só será assumido em apenas uma das *piconets*. Assim, o protocolo *bluetooth* forma uma rede *ad hoc* espontânea (NAMBA, 2011).

Figura 2.9 Exemplo de uma rede piconet, adaptada de NAMBA (2011)

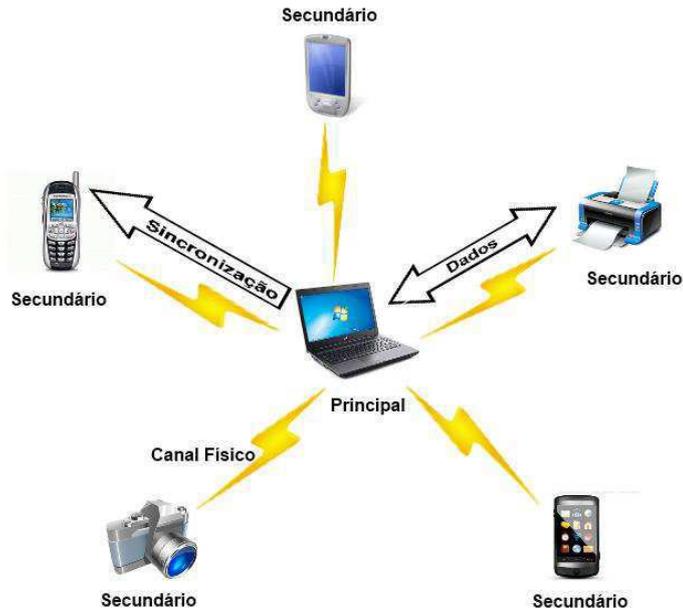
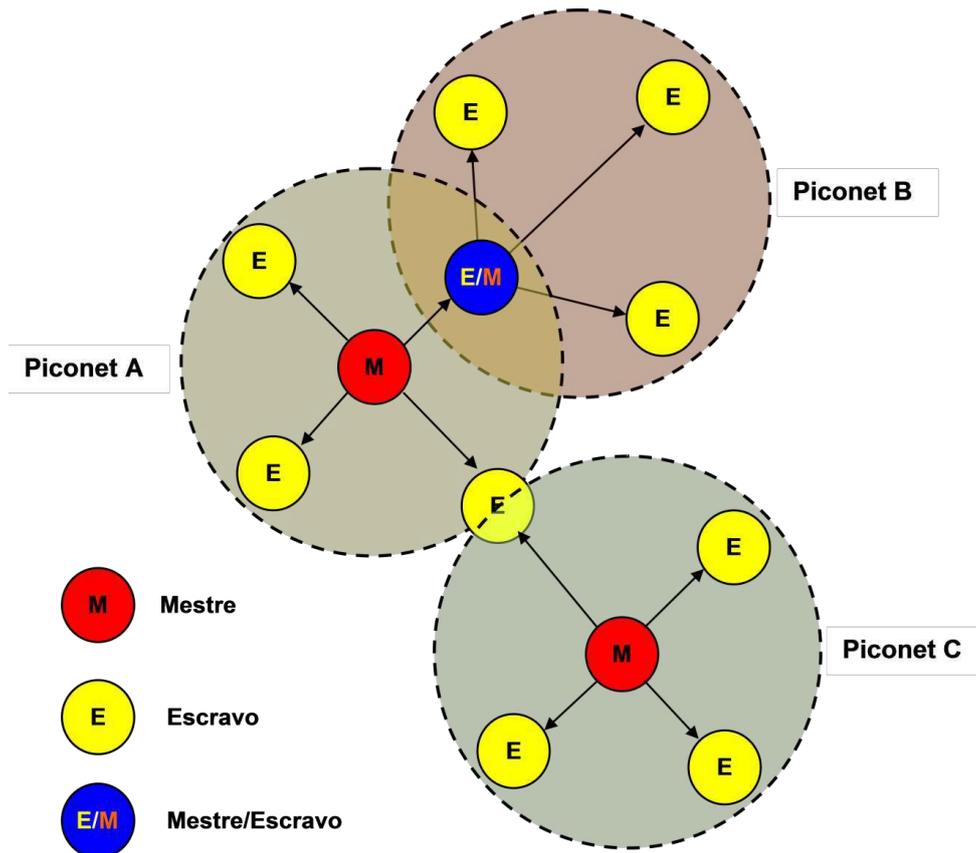


Figura 2.10 Exemplo de uma scatternet, adaptada de NAMBA (2011)



2.5.3 Padrão IEEE 802.15.4 (ZigBee)

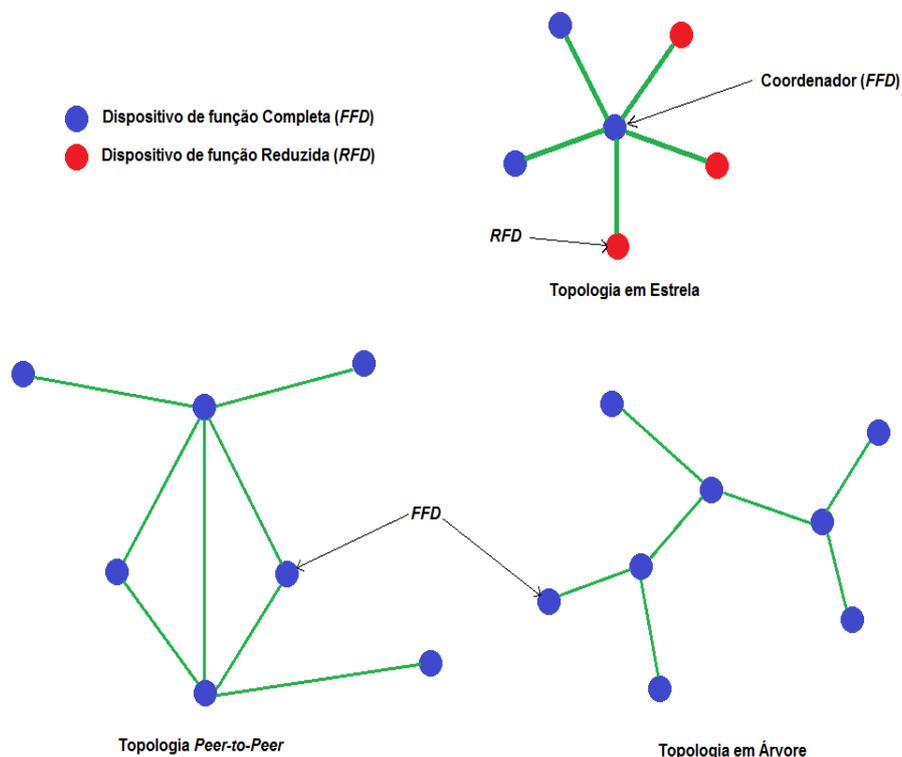
O padrão *IEEE 802.15.4*, também conhecido como *ZigBee*, foi proposto como uma alternativa para a comunicação de dados sem fio em redes de baixa

complexidade de gerenciamento. Esta tecnologia faz uso de protocolo específico, que oferece flexibilidade, sem custo de licenciamento para uso da faixa de espectro *Industrial Scientific Medical (ISM)* e custo de aquisição e manutenção reduzido (NAMBA, 2011).

A estrutura das redes *ZigBee* é formada por 2 (dois) tipos de dispositivos. O *Full Function Device (FFD)*, que possui a função de coordenar a rede e capacidade de acesso a qualquer outro dispositivo, podendo funcionar em toda a topologia do padrão. Entretanto, a construção do *hardware* é mais complexa. O segundo, o *Reduced Function Device (RFD)* possui dispositivos com a construção de *hardware* mais simples e barata. No entanto, ele não pode executar a função de coordenar a rede e é limitado a uma configuração com topologia estrela (NAMBA, 2011).

Na Figura 2.11, pode-se observar que em uma rede com topologia em estrela é necessária a presença de, pelo menos, 1 (um) dispositivo do tipo FFD, que ficará responsável pela coordenação da rede. Nas topologias ponto-a-ponto, só é possível sua implementação quando todos os dispositivos da rede são do tipo FFD.

Figura 2.11 Exemplos das topologias empregadas nas redes *ZigBee*, adaptada de NAMBA (2011)



Outra característica do padrão IEEE 802.15.4 é sua arquitetura baseada no modelo OSI de 7 (sete) camadas, mas que apenas define as camadas de interesse: camada física (*Physical Layer* - PHY); camada de controle de acesso ao meio (*Media Access Control* - MAC); camada de rede (*Network Layer* - NWK); e camada de aplicação (*Application Layer* - APL). Na Figura 2.12, são ilustradas as camadas da arquitetura do padrão ZigBee.

A PHY do padrão *ZigBee* utiliza técnicas de transmissão de sequência direta, também conhecida como *Direct Sequence Spread Spectrum* (DSSS), que torna possível a diminuição do custo dos dispositivos dessa rede, dado que o custo de implementação da interface do tipo *DSSS* é menor em comparação ao custo de implementação de uma interface *Frequency Hopping Spread Spectrum* (FHSS).

A MAC permite a utilização de dispositivos com baixa capacidade de memória em redes com uma grande quantidade de dispositivos, sem afetar a qualidade do serviço da rede, como ocorre em outros padrões.

Figura 2.12: Camadas do padrão IEEE 802.15.4 (ZigBee), adaptada de NAMBA (2011)



A NWK torna possível a expansão da rede sem demandar a utilização de dispositivos de transmissão mais potentes. A camada NWK é responsável por estabelecer e finalizar a conexão dos dispositivos na rede, detectar novos membros e atribuir endereços. Nessa camada estão presentes os mecanismos de busca de rotas e, posteriormente, o encaminhamento dos pacotes de informação, além da implementação da pilha de protocolo que disponibiliza informação para que a aplicação possa controlar o consumo de energia.

Por fim, tem-se a APL que é a interface entre o dispositivo *ZigBee*, o usuário e a camada dos perfis dos dispositivos ou *device profiles*, que são agrupamentos de dispositivos com protocolos de aplicações comuns (NAMBA, 2011).

2.6 Discussão Geral

Neste capítulo, foram apresentados os principais conceitos sobre a RSSF, sua arquitetura e modelo funcional, os problemas fundamentais que afetam o funcionamento destas redes e as principais formas de minimizar esses problemas. Em seguida, foram apresentados os conceitos necessários ao entendimento do problema abordado no trabalho ora descrito.

Com o exposto, foi possível compreender o funcionamento da RSSF e compreender a importância de se buscar soluções para problemas, tais como: tempo de vida da rede; cobertura da área a ser monitorada; TAFF; e TPP.

3 Algoritmos Genéticos

Este capítulo tem por objetivo apresentar informações acerca das teorias que fundamentam as técnicas de busca baseadas em algoritmos evolutivos e, em especial, os Algoritmos Genéticos (AG).

3.1 Técnicas de busca aleatórias guiadas

A classe de técnicas de busca aleatórias guiadas, assim como as técnicas de busca enumerativas, realiza a procura das soluções pesquisando cada ponto do espaço de busca de forma sequencial. No entanto, ela também utiliza informações extraídas nas interações passadas para dirigir a busca por melhores soluções. As duas principais técnicas da classe de buscas aleatórias guiadas são: Algoritmos Evolutivos (AE) e Resfriamento Simulado, ou também conhecido como *Simulated Annealing* (SA).

O SA é uma meta-heurística para otimização, que consiste em uma técnica de busca probabilística, simular ao processo de fundição de metais ou vidros, no qual o metal é aquecido em altas temperaturas e, em seguida, é resfriado lentamente, obtendo-se assim uma massa homogênea. O processo consiste de duas fases: a primeira consiste em elevar a temperatura do sólido a um valor que o funde; na segunda fase, o resfriamento do sólido fundido, deve ser realizado lentamente, até que ele volte ao seu estado sólido. O resfriamento é realizado de forma controlada, na segunda fase, para que os átomos que compõem o material organizem-se numa estrutura uniforme com menor energia interna, reduzindo os defeitos presente no material (HAESER e GOMES RUGGIERO, 2008).

Quando um problema do tipo múltiplo objetivo é submetido a uma otimização utilizando técnicas de busca exata ou até aleatórias guiadas como AS, faz-se necessário transformar as várias funções objetivos do problema em uma mono função objetivo. No entanto, existem outras técnicas de busca aleatórias guiadas como AE que trabalham diretamente com problemas de múltiplos objetivos, utilizando suas múltiplas funções objetivos. Portanto, os AE são mais indicados para

serem utilizado em problemas com múltiplas funções objetivos em comparação com as demais técnicas de otimização (ZINI, 2009).

A técnica conhecida como AE, na computação evolucionária, utilizam os mesmos princípios biológicos da Teoria da Evolução proposta por Darwin. Nessa teoria, os indivíduos que estão mais bem adaptados, têm mais chances de sobreviver e reproduzir-se. Portanto, sua carga genética é passada para as gerações futuras (HOLLAND, 1975). O mesmo ocorre com os AE, em que as melhores soluções para um problema são mantidas para as iterações futuras do algoritmo.

Os AE também são classificados como técnicas de busca que pertencem à classe das técnicas aleatórias guiadas, isto é, fazem uso de técnicas de busca aleatórias, mas também usam as informações do estado corrente para direcionar os métodos puramente aleatórios, como a técnica de *Random Walk* (BURIAN, 2009; LINDEN, 2006).

Três ramos distintos de pesquisa foram utilizados para construir a base dos AE, a saber: As Estratégias Evolucionárias (EE); a Programação Evolutiva (PE); e o AG (GABRIEL e DELBEM, 2008).

As EE envolvem técnicas de busca que tratam de problemas de otimização e são principalmente utilizadas nas áreas de engenharia, como alternativas mais eficientes que as técnicas exatas. Uma das principais características das EE é que elas, originalmente, só consideram o uso do operador genético de mutação na geração das novas soluções. Outras características importantes dos EE são: auto ajuste dos parâmetros do algoritmo; mutação, que se altera de acordo com as variações das características dos dados que estão sendo usados na busca; são bons otimizadores de valores reais; e o fato de possuírem baixo tempo de *takeover*, fazendo com que a população se ajuste mais rapidamente às características dos dados do momento (MINKU, LUDERMIR e ARAÚJO, 2005).

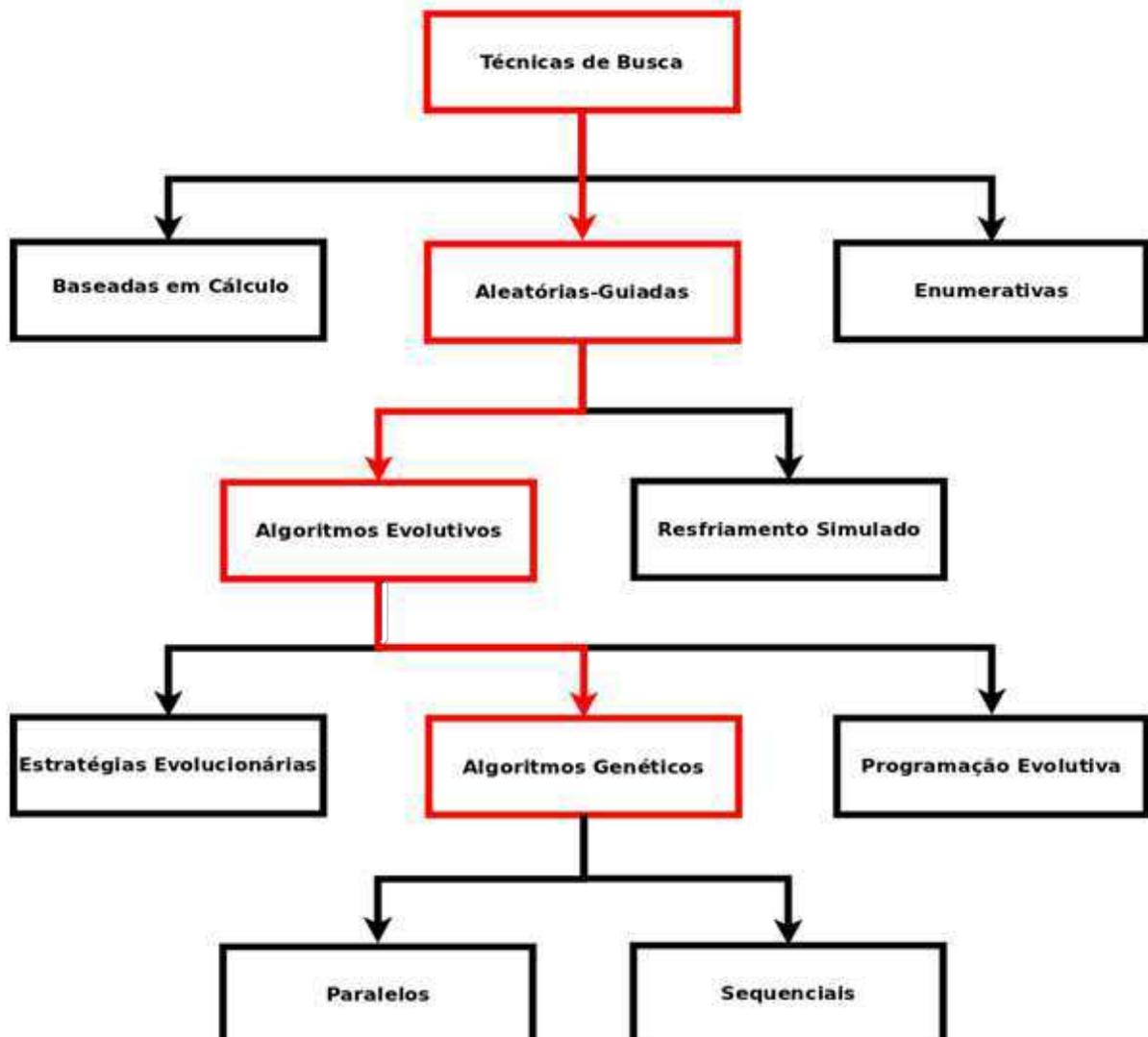
A PE foi apresentada por (FOGEL, OWENS e WALSH, 1966), originalmente como uma técnica para criar inteligência artificial utilizando-se da evolução de máquinas de estado finitos. No entanto, com o passar do tempo, a PE começou a ser aplicada a problemas de otimização, se tornando mais um ramo das buscas por algoritmos evolutivos. A PE, assim como as demais técnicas baseadas em AE, utiliza os conceitos de evolução para obter progressivamente novas soluções. No entanto,

a PE difere dos AG, pois são técnicas que dão mais importância à ligação comportamental entre as populações geradas, do que à ligação genética.

Na PE, a população inicial é gerada aleatoriamente utilizando uma função densidade e, logo após, a população é classificada segundo um objetivo. As próximas gerações são geradas utilizando o operador de mutação aplicado a um indivíduo da atual geração (COELHO e MARIANI, 2003).

Na Figura 3.1, é apresentada a estrutura hierárquica das técnicas de busca nas quais os AG estão em destaque.

Figura 3.1 Algoritmos evolucionários como técnicas de busca, adaptada de BURIAN (2009)

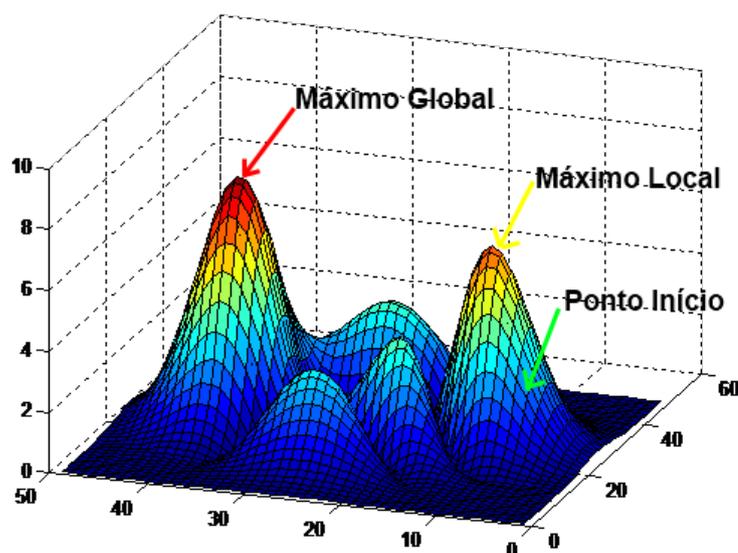


Os AG são métodos adaptativos utilizados em técnicas de busca e otimização de problemas (FERNANDES, 2003). Seus princípios básicos foram propostos no início da década de 70, por *John Holland*, ao pesquisar as características da evolução natural. Ele acreditava ser possível utilizar essa teoria da mesma forma que a natureza utiliza, visando solucionar seus problemas e produzir um novo método para solucionar problemas, para os quais não existiam algoritmos exatos e eficientes.

Os AG fazem parte de um campo de pesquisas sobre os AE. Portanto, podem ser classificados como uma técnica de busca baseada numa metáfora do processo biológico de evolução natural, proposta por Darwin, operando sobre o espaço de solução codificado (GOLDBERG, 1989; LINDEN, 2006). Os métodos adaptativos dos AG podem ser usados para solucionar problemas de busca e otimização que apresentem um alto custo de processamento, em que a utilização do método de busca exato não é recomendada.

As técnicas heurísticas de otimização global usadas nos AG são utilizadas com objetivo de solucionar deficiências de outros métodos de busca, como subida de encosta ou *Hill Climbing*, que faz uso do cálculo da derivada de uma função para encontrar o máximo desta função, ficando facilmente retido em máximos locais, como ilustrado na Figura 3.2 (LINDEN, 2006).

Figura 3.2 Função hipotética com um máximo local e outro global, adaptada de LINDEN (2006)



No Quadro 3.1 é apresentado um resumo das vantagens e desvantagens das técnicas de buscas aleatórias guiadas.

Quadro 3.1 Resumo das vantagens e desvantagens das técnicas de buscas aleatórias guiadas.

Técnica aleatória guiada	Vantagens	Desvantagens
Resfriamento simulado	O AS tem a características de não se prender a mínimos ou máximos locais. Portanto garantindo a convergência da busca.	Não foi projetada para trabalhar com múltiplos objetivos e a convergência é muito lenta para a maioria das aplicações.
Estratégia Evolucionária	Capacidade de auto ajustar seus parâmetros de configuração o que facilita a sua utilização.	Convergência lenta além da busca ponto a ponto ser susceptível a estagnar em mínimos locais.
Algoritmo Genético	Facilmente adaptável a muitas classes de problemas, sejam mono ou multiobjectivos e trabalha com grandes populações de soluções sem comprometer o desempenho da busca.	O ajuste dos parâmetros utilizado para configurar AG requer conhecimento prévio ou será necessário testa-los por tentativa-e-erro. Também não é possível garantir que a solução apresentada seja a melhor solução para o problema.
Programação Evolutiva	O uso do conceito de elitismo torna a convergência mais rápida. A representação da solução é mais simples e diretas, pois não precisar ter um comprimento fixo como nos AG.	Tem uma menor aleatoriedade tornando-o mais propensos a ficarem estagnados a mínimos locais.

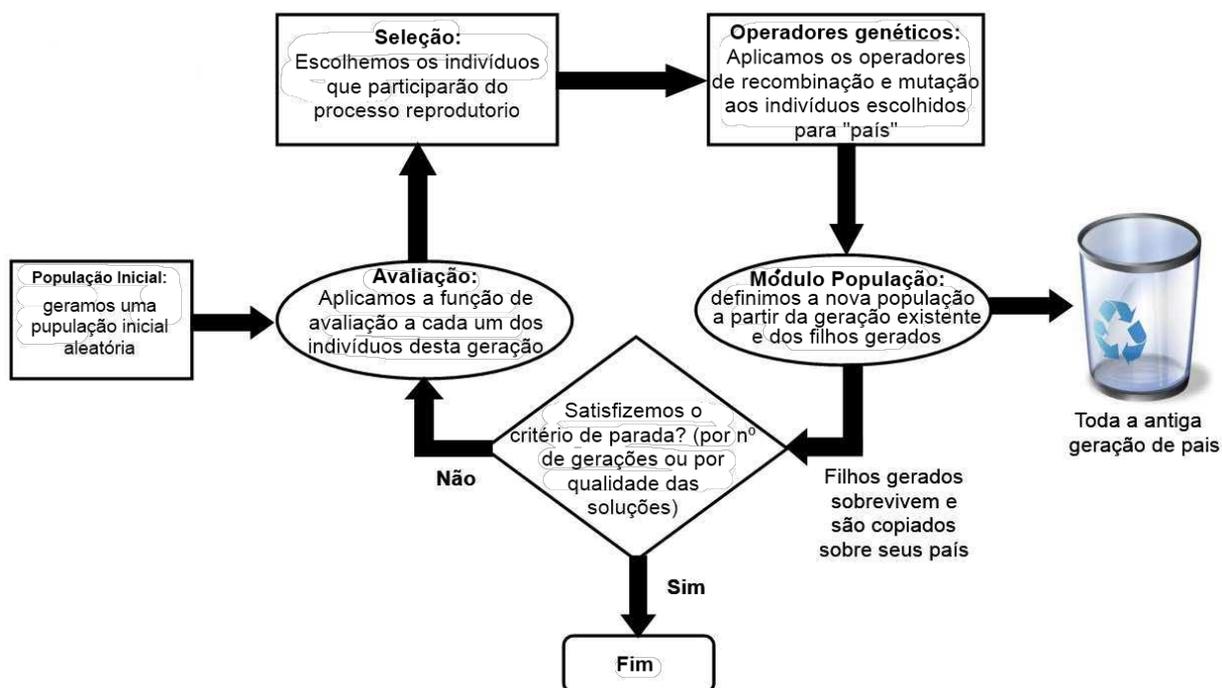
Existem, ainda, várias outras técnicas de busca exatas, algumas delas são: programação linear; algoritmos gulosos; e GRASP. Essas técnicas quando utilizadas para realizar buscas com múltiplos objetivos tem um alto custo computacional que torna seu uso impraticável.

3.2 O algoritmo genético básico

As pesquisas sobre AG baseiam-se na teoria da evolução proposta por Charles Darwin, na década de 50, no século passado. Em suas viagens pelo mundo, Darwin percebeu que animais de uma determinada espécie eram ligeiramente diferentes de seus parentes quando permaneciam em ecossistemas diferentes. Eles apresentavam adaptações que lhes garantiam tirar um melhor proveito do ecossistema no qual viviam. Essa e outras observações induziram à formulação da teoria da evolução, detalhada e descrita no seu livro, intitulado “A Origem das Espécies”, em 1859 (LINDEN, 2006).

Na Figura 3.3, é ilustrado o diagrama básico de um AG, com os principais componentes. Este diagrama será utilizado como base para a elaboração do algoritmo de busca adotado neste trabalho.

Figura 3.3 Diagrama básico de um algoritmo genético básico, adaptada de LINDEN (2006)



O primeiro passo do AG tem a função de criar os indivíduos da primeira geração de forma aleatória. Na etapa de avaliação, a função de avaliação é utilizada para determinar o nível de aptidão de cada um dos indivíduos criados e, logo após, os indivíduos são ordenados. Na terceira etapa do algoritmo, são realizadas as seleções dos indivíduos mais aptos para utilizar os seus genes na etapa de cruzamento e mutação. Ao final da quarta etapa, em que os operadores genéticos são aplicados, tem-se uma nova geração, que poderá ser acrescentada ou substituir a atual. Na sexta etapa, caso algum dos critérios de parada seja satisfeito (número de gerações; objetivo alcançado; ou convergência), o algoritmo será finalizado. No caso do nenhum critério de parada ser satisfeito um novo ciclo do algoritmo será iniciado.

3.2.1 Cromossomos e Genes

A estrutura dos AG é baseada no modelo utilizado na biologia. Nela, a figura do cromossomo representa as possíveis soluções geradas pelo algoritmo ao longo das gerações e a figura dos genes forma uma coleção de valores correspondente a um conjunto de parâmetros, que compõe a solução do problema (FERNANDES, 2003).

O cromossomo é um dos componentes mais importantes utilizados nos AG. Ele possui a função de representar, por meio de um vetor, uma das soluções obtidas na execução do algoritmo, que contém os valores que serão avaliados na função custo (também conhecidos como função objetivo ou *fitness*), além de ser modificado pelos operadores genéticos de cruzamento e mutação.

Estes cromossomos são compostos de genes que, ao serem manipulados pelo AG, dão origem a novos cromossomos com características diferentes. Em um AG, os genes que representam os parâmetros individuais de cada solução são separados de forma a permitir aos operadores genéticos manipulá-los e efetivamente buscar a solução para o problema (HEATON, 2005). Para isso, a representação cromossomial é fundamental nos AG.

3.2.2 Representação Cromossomial

Os primeiros passos a serem realizados com o objetivo de aplicar os AG em qualquer tipo de problema, seja de busca ou de otimização de soluções, é a definição do tipo de representação a ser utilizada no algoritmo. A representação precisa traduzir da maneira mais fiel possível o problema, além de possibilitar que o algoritmo consiga manipulá-la adequadamente. Quanto maior a adequação da representação ao problema, melhor será a qualidade dos resultados obtidos (LINDEN, 2006).

Segundo Linden (2006), a definição da representação cromossomial é completamente arbitrária, e a sua definição fica a critério das escolhas realizadas pelo projetista com base nas características do problema. Linden (2006) também sugeriu que fossem adotadas algumas regras:

- Buscar uma representação que seja o mais simples possível;
- Quando o problema possuir soluções proibidas, sua representação não deverá ser possível de ser gerada;

- Quando o problema impõe condições de algum tipo, estas deverão estar implícitas na representação.

Os AG podem utilizar qualquer base numérica para representação de soluções, as mais utilizadas são (LINDEN, 2006): a representação binária, a representação real e a representação por permutação de símbolos.

3.2.2.1 Representação Binária

A estrutura de dados mais utilizada para representar as soluções candidatas ou cromossomos, em um AG, são as cadeias binárias. A representação binária foi inicialmente utilizada por Holland em seu livro sobre AG (HOLLAND, 1975). Por esse aspecto histórico e pela representação binária ser muito simples de implementar, ela ainda é amplamente adotada pelos pesquisadores da área, mesmo quando as variáveis do problema que o AG pretende resolver são do tipo inteiro ou real (LINDEN, 2006).

A Equação (3.1) apresenta a variável do tipo real x , com intervalo entre a e b , pode ser codificada utilizando um vetor binário de comprimento l , tendo uma precisão numérica dada pela Equação (3.2).

$$x \in (a, b), \quad (3.1)$$

$$\frac{(a - b)}{(2^l - 1)}. \quad (3.2)$$

em que:

$$x \in \mathbb{R},$$

$$l \in \mathbb{N},$$

$$a, b \in \mathbb{Z}.$$

Outra motivação para o uso da representação binária vem da teoria dos esquemas. Nela, é sugerida a utilização de alfabetos de baixa ordem com poucos símbolos, já que eles são mais eficientes que os alfabetos de alta ordem. Segundo Golderg (1989), alfabetos pequenos maximizam o número de esquemas disponíveis para o processamento do AG. O cálculo do número de esquemas é simples e pode ser obtido a partir da Equação (3.3).

$$n_s = (k + 1) \frac{1}{\log_2 k}. \quad (3.3)$$

Considere um AG que utiliza um alfabeto de cardinalidade k . Uma vez que existem $(k + 1)$ esquemas por posição e cada posição pode ter $\log_2 k$ bits, tem-se n_s esquemas por bit de informação. Como nenhum alfabeto pode ter menos de 2 (dois) símbolos, conclui-se que a representação binária é a forma mais eficiente de manipular esquemas.

Para representar um número do tipo real por meio da representação binária, torna-se necessário definir primeiro a faixa de operação da variável e qual a precisão desejada. Esses dois parâmetros, faixa de operação da variável e a precisão, definem quantos bits serão usados na codificação (LINDEN, 2006).

Um exemplo seria uma aplicação hipotética, em que se deseja codificar um par ordenado (x, y) de um ponto no plano cartesiano utilizando a representação binária. Para cada ponto, tem-se 2 (duas) variáveis do tipo real x_i e y_i . Os valores destas variáveis são os seguintes:

$$x_1 = 10 ,$$

$$x_2 = 55 ,$$

$$y_1 = 100 ,$$

$$y_2 = 125 .$$

Para cada uma das variáveis serão usados k bits, em uma faixa $[\text{inf}_i, \text{sup}_i]$, em que:

$$k = 8 ,$$

$$\text{inf}_i = 0 ,$$

$$\text{sup}_i = 255 .$$

A precisão máxima é dada pela Equação (3.4), cujo resultado para o exemplo é apresentado na Equação (3.5).

$$\frac{\text{sup}_i - \text{inf}_i}{2^k - 1} , \quad (3.4)$$

$$\frac{255 - 0}{2^8 - 1} = 1,000 . \quad (3.5)$$

O vetor cromossomo binário é ilustrado na Figura 3.4:

Figura 3.4 Representação de um vetor cromossomo do tipo binário

00001010	00110111	01100100	01111101
x_1	y_1	x_2	y_2

Embora a representação binária seja a mais simples de implementar e a forma mais eficiente de manipular, ela apresenta alguns problemas. O principal deles é o efeito chamado abismo de *Hamming*, que ocorre quando é preciso efetuar um acréscimo simples de uma unidade. Na representação inteira, para mudar o valor de 7 (sete) para 8 (oito), só é necessário o acréscimo de uma unidade. Já na representação binária, para a mesma mudança de 7 (0111) para 8 (1000) se faz necessária a mudança de todos os bits (LINDEN, 2006).

Objetivando contornar este problema, foi proposta a utilização do código *Gray* na representação binária. Com esta alteração, a distância de *Hamming* entre 2 (duas) cadeias consecutivas quaisquer é sempre 1 (um). Na Tabela 3.1, é apresentada a conversão de inteiro para os códigos binários e *Gray*.

Tabela 3.1 Conversão de inteiro para binário e *Gray*, adaptada de CASTRO e ZUBEN (2012)

Inteiro	Binário	Gray
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

A representação binária também apresenta dificuldades na representação de variáveis contínuas de múltiplas dimensões e alta precisão, pois este tipo de uso exige uma quantidade grande de *bits*, tornando o tamanho do vetor cromossomo muito grande e de difícil manipulação para o AG. Portanto, quando o problema possui esta característica, é necessária a utilização de outras formas de representação do cromossomo (LINDEN, 2006).

3.2.2.2 Representação Real

O bom desempenho dos AG na busca de soluções não tem relação direta com a utilização do tipo de representação binária. Em muitos problemas que lidam com otimização de parâmetros do tipo contínuo, o uso de uma representação real é mais natural e adequado, pois é possível representar diretamente os parâmetros sendo otimizados como números reais, de tal forma que o espaço de busca também pertença ao domínio dos reais. Logo, tem-se uma busca mais precisa e eficiente quando os parâmetros que serão otimizados são do tipo contínuo.

Algumas implementações dos AG utilizam cadeias de *string* como vetores. Neste caso, o uso de cromossomos com representação real para parâmetros contínuos reduz o tamanho do vetor utilizado pelo cromossomo e também elimina a necessidade do uso da função de mapeamento. Com isso, espera-se um ganho no tempo de processamento.

Existem outros fatores que a utilização da representação real proporciona quando o problema apresenta parâmetros do tipo real a serem otimizados: a precisão da representação do problema é mantida, já que não existem perdas decorrentes da utilização de funções de mapeamento nas transformações; e o tamanho do cromossomo só dependerá do número de parâmetros que precisa ser otimizado, mantendo a precisão mesmo para parâmetros que pertençam a grandes domínios (LINDEN, 2006).

O cromossomo real também pode ser representado por um vetor de pontos flutuantes utilizando a norma IEEE 754 para sua representação interna. A principal vantagem da utilização da norma IEEE 754 é permitir uma maior portabilidade para a implementação do AG, pois a maioria dos fabricantes de computadores e desenvolvedores de compiladores já possui este tipo de representação de dados implementado (ELEFTHERIOS, FUKAZAWA e OTA, 2012; VIANA, 1999). A Figura 3.5 ilustra um exemplo de um vetor cromossomo real.

Figura 3.5 Representação de um vetor cromossomo do tipo real

10,000	55,000	100,000	125,000
x_1	y_1	x_2	y_2

3.2.2.3 Representação por Permutação de Símbolos

Quando o problema que se pretende resolver utilizando AG não pertence à classe dos problemas de otimização numérica, como os baseados em uma ordem, a representação cromossomial mais indicada é a permutação de símbolos.

Um exemplo dessa classe é o problema do caixeiro viajante, que precisa percorrer várias cidades, mas, para que a distância percorrida seja a menor possível, ele não pode visitar mais de uma vez a mesma cidade. Portanto, para resolver este problema utilizando AG, é necessário escolher 1 (um) tipo de codificação, 1 (um) conjunto de operadores genéticos e 1 (uma) função custo, que seja apropriada à representação.

Os vetores do Exemplo 3.1 são possíveis codificações para o caso do caixeiro viajante, considerando que o problema tem 7 (sete) cidades a serem visitadas.

Exemplo 3.1 Vetores com Representação por Permutação de Símbolos

Vetor A = (1, 5, 3, 2, 7, 6, 4) ,
Vetor B = (1, 5, 3, 7, 6, 2, 4, 1) ,
Vetor C = (1, 5, 3, 7, 6, 4) .

Observa-se que o vetor A é uma codificação válida para uma representação por permutação e símbolos, pois o vetor possui exatamente 7 (sete) símbolos diferentes. Os vetores B e C não são válidos, dado que infringem uma das restrições. O Vetor B repete a cidade 1 (um) e o vetor C não visita a cidade 2 (dois) (LINDEN, 2006).

3.2.3 População Inicial

A população de um AG é formada por um conjunto de possíveis candidatos à solução do problema e é utilizada para o fornecimento de material genético às futuras gerações (LOBO, 2005). As implementações mais simples de um AG geralmente eliminam todos os indivíduos da geração atual, substituindo-os pelos seus filhos (LINDEN, 2006). No entanto, existem outros comportamentos que podem

ser utilizados e que proporcionam ganhos no desempenho do AG. Dentre as alternativas mais conhecidas, tem-se o elitismo e o *Steady State*.

O *Steady State* ou regime permanente é uma alternativa à utilização do elitismo, que também proporciona uma melhoria no desempenho do AG. No *Steady State*, os indivíduos da geração atual são preservados para formar a nova geração. Apenas os indivíduos com menor aptidão serão substituídos por filhos gerados a partir dos cruzamentos de indivíduos de melhor aptidão (LINDEN, 2006).

Um parâmetro de grande influência no desempenho do AG é o tamanho da população. Quando a população escolhida possui um número de indivíduos reduzidos, a variedade genética da população torna-se insuficiente, uma vez que os poucos indivíduos não conseguem caracterizar todo o espaço de busca do problema. Logo, o algoritmo é incapaz de achar boas soluções. No entanto, quando o número de indivíduos da população é muito grande, o algoritmo demandará mais tempo de processamento e se assemelhará ao comportamento de um algoritmo de busca exaustiva (LINDEN, 2006).

Os indivíduos ou soluções da população inicial são gerados de forma aleatória, formando a primeira geração que será submetida ao AG. O uso de técnicas heurísticas para a determinação da população inicial, em conjunto com um tamanho adequado da população, são importantes para evitar uma convergência prematura do algoritmo que poderá levar o AG a convergir para um ótimo local (FERNANDES, 2003).

3.2.4 Função de Avaliação

A função de avaliação ou *fitness* é outro componente importante dos AG que interfere no seu desempenho. Ela é utilizada para a escolha dos indivíduos (cromossomos) que serão utilizados como reprodutores (pais) da próxima geração. Esta característica do AG, que consiste em usar os melhores cromossomos como sementes para as próximas gerações, torna esse processo uma busca direcionada no espaço de soluções (LINDEN, 2006).

A partir do uso dessa função, é possível verificar quão próximo os indivíduos se encontram da solução. Portanto, todas as restrições do problema devem estar de alguma forma, presentes na função de avaliação. Geralmente, a forma mais comum de inserir as restrições nas funções de avaliação consiste em punições apropriadas.

O seu uso permite que a avaliação obtida para um determinado indivíduo, que infringiu alguma restrição, seja reduzida de forma proporcional à sua gravidade.

Outra boa característica que deve ser buscada nas funções de avaliação é que elas sejam suaves e regulares, de maneira que as boas soluções estejam próximas uma das outras e que não haja discrepâncias nos valores dessas soluções. Ao se somar as características de suavidade e regularidade com funções que tenham poucos máximos locais, tem-se a função de avaliação que proporcionará o melhor desempenho para o AG. No entanto, essas características não são de fácil obtenção (LINDEN, 2006).

Nos AG mais simples é usada como medida de qualidade do indivíduo o valor retornado pela função de avaliação. Apesar de essa estratégia apresentar bons resultados na maioria dos problemas, ela não é apropriada para os problemas que tenham a presença de super indivíduo. Este tipo de problema ocorre quando a diferença entre os valores retornados pela função de avaliação de um indivíduo de baixa qualidade é próxima ao valor retornado de um indivíduo de boa qualidade.

Quando o problema apresenta a ocorrência de um ou mais indivíduos, cuja avaliação é consideravelmente superior a dos demais, o módulo de seleção do AG quase sempre selecionará esses super indivíduos. Como consequência, a diversidade genética das próximas gerações será reduzida significativamente.

Quando a diferença da avaliação para todos os indivíduos está muito próxima, o módulo de seleção criará espaços muito semelhantes na roleta viciada⁶ para todos os indivíduos, independentemente se ele é uma boa solução ou não. Isso pode levar o algoritmo a descartar uma solução ótima, além de desperdiçar recursos em um espaço de busca pouco promissor (LOBO, 2005).

Quando a função que retorna o valor exato da avaliação não for apropriada para o problema, outros tipos de funções podem ser utilizados. Linden (2006) propõe algumas alternativas que podem contornar este problema. Dentre as mais utilizadas estão: a normalização; *windowing*; e o escalonamento sigma. Na Tabela 3.2 é

⁶ Roleta Viciada – Neste método, cria-se uma roleta virtual no qual o espaço para cada cromossomo na roleta é proporcional ao grau de aptidão do cromossomo ao problema. A soma de todos os espaços não pode exceder ao espaço total disponível na roleta virtual (LINDEN, 2006).

apresentado o resultado da utilização da técnica de normalização, com o objetivo de eliminar o problema do super indivíduo.

Tabela 3.2: Comparação da avaliação entre uma função pura e normalizada, adaptada de LINDEN (2006)

Indivíduo	Avaliação ($f(x)$)	Avaliação ($\log_{10}f(x)$)
10000	256	2,410
00100	16	1,200
00001	1	0,000
00011	9	0,950
00010	4	0,600

3.2.5 Critério de Parada

O módulo critério de parada é o responsável por determinar quando o algoritmo deve parar. Existem vários critérios que podem ser utilizados para determinar o melhor momento de parada da execução do algoritmo, que normalmente depende do resultado da função de avaliação. Os critérios mais mencionados na literatura são (BECKMANN, 2010; LOBO, 2005):

- Número de Gerações: definição prévia de quantas interações serão usadas na busca do melhor indivíduo;
- Desempenho ótimo: quando se conhece qual o melhor valor que a função de avaliação pode alcançar;
- Grau de convergência: quando a diferença da aptidão do melhor indivíduo é muito próxima a do pior indivíduo de uma mesma geração;
- Perda de diversidade: quando os indivíduos gerados são muito parecidos, uns com os outros.

3.2.6 Seleção

A seleção é mais um dos módulos que compõem um AG. Em comparação com a biologia, este módulo baseia-se no processo de seleção natural que atua sobre as espécies biológicas. Na biologia, os indivíduos mais aptos a se sobressair sobre os demais logo desenvolvem o potencial de produzir descendentes mais adaptados ao

meio. Por este motivo, se reproduzem com mais facilidade que os demais. No entanto, os indivíduos menos aptos também podem gerar descendentes.

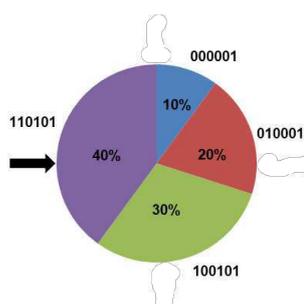
Um aspecto importante do módulo de seleção é que ele possibilita que os indivíduos menos aptos possam se reproduzir, evitando, assim, a convergência genética prematura para um máximo local. Esses indivíduos, mesmo com péssima avaliação, podem ter material genético, que, em combinação com outros indivíduos, pode gerar um descendente melhor para a solução do problema que está sendo tratado. Este material pode não ser encontrado em nenhum dos demais indivíduos (LINDEN, 2006).

O módulo de seleção irá escolher 2 (dois) indivíduos para fornecer material para os novos indivíduos da próxima geração. Existem várias estratégias para realização desta escolha, sendo as mais usadas: o método da roleta viciada; o método do torneio; e o método de amostragem estocástica uniforme (LOBO, 2005).

3.2.6.1 Método de Seleção da Roleta Viciada

O método mais utilizado pelo módulo de seleção na grande parte dos trabalhos é o da roleta viciada. Este método consiste em particionar um disco, denominado roleta, em fatias de tamanhos proporcionais ao valor da aptidão de cada indivíduo da população. Portanto, os indivíduos mais bem avaliados terão uma área maior que os indivíduos menos aptos. A roleta é girada para apontar um indivíduo da população que irá fornecer parte do material genético aos indivíduos da próxima geração. A quantidade de vezes que a roleta precisará ser acionada dependerá do número de indivíduos necessários para compor uma nova população (geração) e do tipo de cruzamento que será utilizado no algoritmo (LOBO, 2005). Na Figura 3.6 é ilustrado o método da roleta viciada.

Figura 3.6 Método de seleção por roleta viciada, adaptada de LOBO (2005)



3.2.6.2 Método de Seleção por Torneio (*Rank*)

O método de seleção por torneio é caracterizado por estabelecer uma competição entre os indivíduos da população. Inicialmente, são selecionados aleatoriamente k indivíduos de uma população, com um total de θ indivíduos (tamanho da população), que serão usados para formar uma subpopulação. Em seguida, esta subpopulação é colocada para competir entre si. O indivíduo que apresentar a melhor aptidão sairá vencedor desse torneio e seu material genético será utilizado no operador de cruzamento do algoritmo. Este processo de escolha da subpopulação e, logo após, a realização do torneio, se repetirá pela quantidade de vezes necessária para completar a nova população.

Neste método, o indivíduo com pior avaliação só poderá ser selecionado para participar do cruzamento se ele for o único competidor do torneio. Esta condição tem uma probabilidade muito reduzida, o que pode, em certos casos, levar à perda de informações importantes, que só existiriam nesse indivíduo de pior avaliação.

Na Figura 3.7 é ilustrado um exemplo de aplicação do método de seleção por torneio, com os parâmetros $k = 3$ e $\theta = 8$. No lado esquerdo da Figura 3.7 tem-se um quadro que representa todos os indivíduos da população com seus respectivos valores de aptidão. No lado direito da Figura 3.7, são ilustrados os 8 (oito) torneios necessários para formar a nova população. Para cada torneio, são mostrados os 3 (três) elementos participantes. Os indivíduos marcados com fundo cinza na Figura 3.7 são os vencedores de cada torneio realizado (LINDEN, 2006).

Figura 3.7 Método de seleção por torneios, adaptada de LINDEN (2006)

Indivíduo	Aptidão
X ₁	200
X ₂	100
X ₃	9500
X ₄	100
X ₅	100
X ₆	10000
X ₇	1
X ₈	40

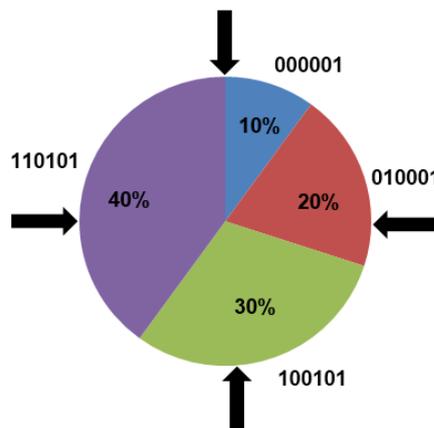
Torneios		
X ₁	X ₇	X ₈
X ₂	X ₃	X ₅
X ₆	X ₄	X ₄
X ₂	X ₇	X ₁
X ₅	X ₅	X ₅
X ₃	X ₄	X ₂
X ₄	X ₂	X ₆
X ₄	X ₆	X ₅

3.2.6.3 Método de Amostragem Universal Estocástica

O método de amostragem universal estocástica (*Stochastic Universal Sampling - SUS*), assim como o método da roleta viciada, pode ser implementado com a utilização de uma roleta, na qual os indivíduos terão uma fatia do disco proporcional ao valor de sua avaliação. A diferença entre os 2 (dois) métodos é que, no caso da amostragem universal estocástica, ao invés de uma única agulha, são usadas n agulhas igualmente espaçadas, que irão escolher os indivíduos que formarão a próxima geração. O parâmetro n representa o número de indivíduos necessários para formar uma nova população, portanto só é necessário girar a roleta uma única vez (LOBO, 2005).

Este método tem como característica a determinação de probabilidades proporcionais ao valor da avaliação de cada indivíduo (LINDEN, 2006). O que, conseqüentemente, pode levar os indivíduos mais bem avaliados a serem escolhidos mais de uma vez e os menos aptos a nunca serem escolhidos (LOBO, 2005). Na Figura 3.8, é apresentado o método de seleção por amostragem universal estocástica, com o parâmetro $n = 4$.

Figura 3.8 Método de amostragem universal estocástica, adaptada de (LOBO, 2005)



3.2.7 Elitismo

O módulo de elitismo normalmente está associado ao módulo de população e tem o objetivo de aumentar a velocidade de convergência do algoritmo (LOBO, 2005). Essa associação praticamente não altera o tempo de processamento, mas elimina o

problema de perda de desempenho do algoritmo durante a criação de novas gerações (LINDEN, 2006).

O módulo de elitismo basicamente copia os n melhores indivíduos da atual população para a próxima geração. Como a grande maioria dos métodos de avaliação de desempenho dos AG faz uso do valor de adequação do indivíduo ao problema para determinar qual o melhor, a manutenção do melhor indivíduo da atual geração garante que o melhor da próxima seja, pelo menos, igual ao melhor indivíduo da geração atual.

Apesar da simplicidade do módulo de elitismo, o efeito sobre o ganho de desempenho do algoritmo é nítido. Isso ocorre devido à manutenção dos melhores esquemas que estão presentes nos indivíduos mais aptos da população, o que vem de encontro ao princípio de *exploitation*, sem interferir no princípio de *exploration* (LINDEN, 2006). O princípio de *exploitation* é caracterizado por utilizar os melhores esquemas do indivíduo para refinar a busca, e o princípio de *exploration* é responsável pela busca de novos esquemas, evitando que o algoritmo fique preso a máximos locais.

3.2.8 Operadores Genéticos

Os operadores compõem mais um dos elementos utilizados nos AG que foram inspirados na natureza. Eles fazem parte da etapa de reprodução do algoritmo e tem o objetivo criar novos indivíduos para a população. Os operadores genéticos utilizados nos primeiros AG foram inspirados na reprodução humana, como os operadores de cruzamento e mutação (LOBO, 2005).

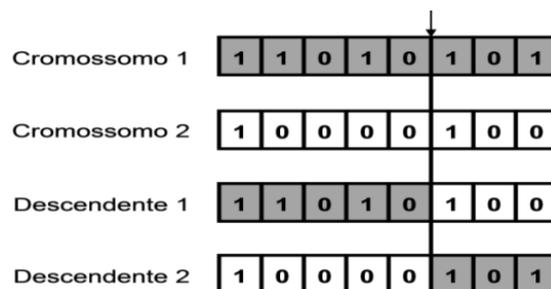
A utilização dos operadores genéticos no algoritmo tem papel fundamental na busca pela melhor solução. São eles os responsáveis por estender a busca até que a melhor solução seja alcançada, por meio das transformações que os indivíduos sofrem após sucessivas gerações. Os operadores genéticos possibilitam manter a diversidade das boas características dos indivíduos das gerações anteriores (*exploitation*).

3.2.8.1 Cruzamento

Existem vários operadores de cruzamento, também chamados de operadores *crossover*. O mais simples deles é o operador de cruzamento de um ponto. Com todos os indivíduos já avaliados, esses serão submetidos ao módulo de seleção, para que 2 (dois) deles sejam escolhidos como os pais de um novo indivíduo (cromossomo). Após a seleção, um ponto de corte é determinado estocasticamente. Esse ponto determina a posição no cromossomo, na qual ele será dividido. A primeira parte do primeiro cromossomo é cruzada com a segunda parte do segundo cromossomo, gerando um novo cromossomo filho. O segundo filho é gerado utilizando as partes restantes. Ao final do processo de cruzamento, tem-se 2 (dois) novos indivíduos (LINDEN, 2006).

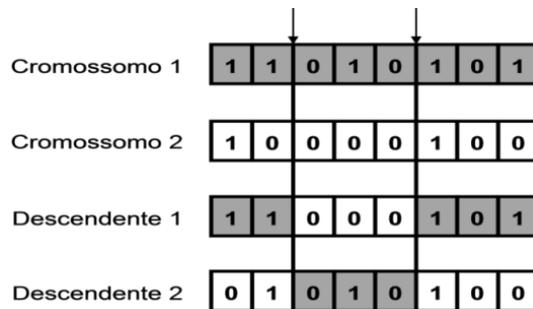
Na Figura 3.9, é apresentado um exemplo da utilização de um operador de cruzamento de um ponto. No exemplo, são utilizados 2 (dois) cromossomos de 8 (oito) *bits* para a ocorrência do cruzamento. Após o sorteio aleatório, os 2 (dois) cromossomos darão origem a 2 (dois) novos cromossomos, que serão completamente diferentes dos seus pais e, ainda assim, irão manter diversas características em comum.

Figura 3.9 Exemplo do operador de cruzamento de um ponto, adaptada de LOBO (2005)



No operador de cruzamento de 2 (dois) pontos, o processo é semelhante. A diferença é que nele serão sorteados 2 (dois) pontos de corte que irão dividir o cromossomo em 3 (três) partes. O primeiro cromossomo terá a sua segunda parte, ou parte central, substituída pela segunda parte do segundo cromossomo, formando o primeiro filho. Enquanto que as partes restantes serão usadas para compor o segundo filho, seguindo o mesmo processo. Na Figura 3.10, é ilustrado o cruzamento de 2 (dois) pontos.

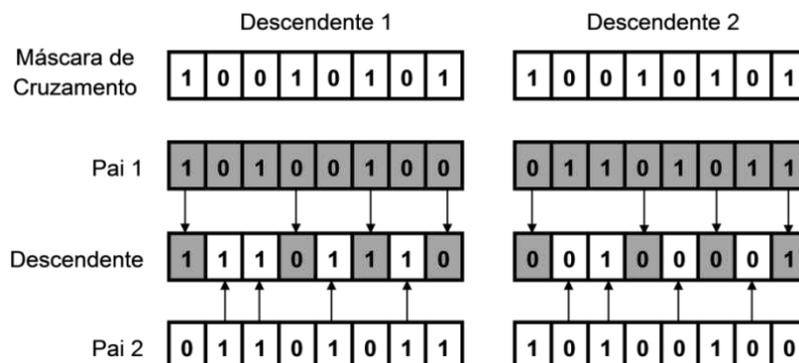
Figura 3.10 Exemplo do operador de cruzamento de dois ponto, adaptado de (LOBO, 2005)



O terceiro tipo é o cruzamento uniforme. O seu funcionamento é completamente diferente dos 2 (dois) primeiros tipos de cruzamento, explanados anteriormente. No cruzamento uniforme, existe um novo componente denominado de máscara de cruzamento, que é criada de forma aleatória. Esta máscara irá controlar qual o cromossomo a ser retirado do gene para formar o seu descendente, de modo que, quando o *bit* da máscara for 1 (um), o gene selecionado virá do cromossomo 1 (um), e caso seja 0 (zero), o gene virá do cromossomo 2 (dois). Logo após, o processo será repetido invertendo os cromossomos pais, para se obter o segundo descendente (LOBO, 2005). Na Figura 3.11, é apresentado o cruzamento uniforme.

Frequentemente, nos trabalhos que utilizam a taxa de probabilidade de cruzamento para determinar quando o operador de cruzamento deve ser aplicado, encontram-se valores entre 50,0% e 100,0% (BECKMANN, 2010).

Figura 3.11 Exemplo do operador de cruzamento uniforme, adaptado de (LOBO, 2005)



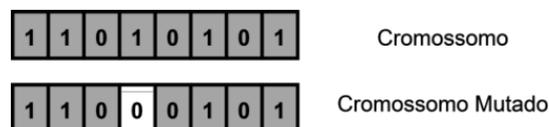
3.2.8.2 Mutação

Operador de mutação é um operador básico, que realiza uma pequena modificação de forma aleatória, nos indivíduos que resultaram da operação de cruzamento. Este operador tem a função de proporcionar mudança na direção de convergência, evitando que o algoritmo fique preso a mínimos ou máximos locais (FERNANDES, 2003).

A mutação permite que o algoritmo utilize novos esquemas do espaço de busca. É sabido que a evolução das populações a partir do operador de cruzamento tende a originar indivíduos, em uma mesma geração, muito parecidos entre si, causando o problema de perda de diversidade. Portanto, o número de indivíduos que sofreram mutação deve ser reduzido. Os estudos realizados por Beckmann (2010) indicam que a probabilidade de ocorrência de mutação não deve ultrapassar 1,0 %. No entanto, é preciso garantir que esse valor seja suficiente para assegurar a diversidade dos indivíduos.

O parâmetro μ , que define a probabilidade de mutação no algoritmo, decide a quantidade de vezes que o operador será aplicado durante a execução do algoritmo. O valor de μ adotado na maioria dos trabalhos varia entre 0,1% e 5,0%. Quanto maior for essa percentagem, maior será a diversidade dos indivíduos, e o comportamento do algoritmo se aproximará ao de uma busca aleatória (BECKMANN, 2010). A Figura 3.12 traz um exemplo de aplicação do operador de mutação. Nela, é possível observar que o quarto *bit* do cromossomo sofreu uma mutação de 1 (um) para 0 (zero) (LOBO, 2005).

Figura 3.12 Exemplo do operador de mutação, adaptado de (LOBO, 2005)



3.3 Pseudocódigo de um Algoritmo Genético

Conforme pode ser observado no pseudocódigo apresentado no Código 3.1 e no esquema da Figura 3.3, o algoritmo começa com a criação da primeira população, que será avaliada pelo módulo de avaliação. Terminada a avaliação, o módulo de

população assume o controle do algoritmo e faz a seleção dos 2 (dois) cromossomos pais que irão gerar os 2 (dois) novos descendentes. Após a seleção, os cromossomos escolhidos são submetidos aos operadores de cruzamento e o processo de mutação e o módulo de população serão repetidos até que a nova população esteja completa. No próximo passo, o algoritmo descarta os antigos cromossomos e os novos cromossomos são enviados para a avaliação. No último passo, o critério de parada é aplicado nos cromossomos para decidir se o algoritmo volta para o módulo de população ou é finalizado (LINDEN, 2006).

Código 3.1 Pseudocódigo de um algoritmo genético básico, adaptado de LINDEN (2006).

```
inicializar a população de cromossomos
avaliar cada cromossomo na população
repetir enquanto não é verificada a condição de fim
    selecionar os pais para gerar novos cromossomos
    aplicar o operador de recombinação (cruzamento)
    aplicar o operador de mutação com probabilidade  $\mu$ 
    apagar os velhos membros da população
    avaliar os novos cromossomos e insira-os na população
fim da repetição
fim do algoritmo
```

3.4 Discussão Geral

Neste capítulo, foram abordados conceitos sobre o algoritmo de busca utilizado no trabalho, o AG, detalhando seus principais componentes, tais como: os tipos de representação; a importância da população inicial; a função de avaliação; os tipos de operadores básicos utilizados na evolução das gerações; os possíveis critérios de parada; e, por fim, um esquema básico do algoritmo que servirá de base para o AG proposto neste trabalho.

O exposto neste capítulo e no Capítulo 2, possibilitou a compreensão do funcionamento da RSSF e da importância por se buscar soluções para os problemas, a saber: tempo de vida da rede; a cobertura da área a ser monitorada; o

TAFF; e a TPP. Para tanto, será utilizado um AG adaptado de forma a proporcionar o melhor desempenho da RSSF.

No próximo capítulo, serão apresentados os trabalhos relacionados ao proposto, que tratam da: maximização do tempo de vida da RSSF e da taxa de cobertura da área do cenário, visando minimizar também o TAFF e o TPP.

4 Trabalhos Relacionados

As RSSF, frequentemente, são compostas por um grande número de nós sensores densamente implantados, o que está tornando essa tecnologia muito promissora para áreas como: Vigilância e monitoramento, Rastreamento de alvos em ambientes hostis e monitoramento de tráfego (CHENG, CHUAH e LIU, 2004).

Apesar desse desenvolvimento, as RSSF ainda precisam conviver com vários problemas e limitações, destacando-se as questões que envolvem a baixa capacidade de energia disponível nos nós sensores, pois esses utilizam baterias ou pilhas comuns. Logo, a busca por alternativas para diminuir o consumo de energia dos nós de uma RSSF, em outras palavras, o aumento do tempo de vida útil da rede, é uma área de pesquisa que vem atraindo muitos pesquisadores (MIRANDA, 2011).

Atualmente, existem várias linhas de pesquisa que objetivam resolver o problema do tempo de vida das RSSF, dentre os quais se destacam o melhoramento nos algoritmos de roteamento, a redução na transmissão de dados e o controle do estado dos nós na rede. Este trabalho visa solucionar o problema a partir da busca do melhor posicionamento dos nós sensores na área a ser monitorada, a partir do uso de AG.

O problema abordado em Ganesan *et al.* (2004) leva em consideração a implantação de um número finito de nós sensores em uma área geográfica e recomenda uma estrutura de comunicação entre os nós da RSSF. Ganesan *et al.* (2004) procura mostrar a importância de se implantar os nós nas melhores posições possíveis, para diminuir ao máximo a quantidade de nós utilizados, pois nesse caso o custo do dispositivo era elevado. Outro problema abordado no trabalho é o consumo de energia.

O trabalho de Ganesan *et al.* (2004) propõe que seja simplificado o problema de determinação do melhor posicionamento dos nós sensores para um caso unidimensional, que possibilita a aplicação de uma solução analítica quando se tem um caso de regime de agregação simples. Após o resultado obtido na análise unidimensional, Ganesan *et al.* (2004) propõe que os resultados sejam estendidos para o caso bidimensional.

A metodologia utilizada por Ganesan *et al.* (2004), para obter a redução no consumo de energia, está baseada na utilização de uma árvore de roteamento que tem como raiz a estação base e as folhas são representados pelos nós sensores. Os dados são transmitidos utilizando encaminhamento multi-hop e os nós são implantados utilizando uma estratégia de posicionamento bidimensional numa estrutura circular. Outra técnica utilizada é a agregação de dados que ocorre quando um nó de junção recebe dados dos seus nós filhos, decodifica os dados recebidos e agrega-os aos seus próprios dados e encaminha todos ao seu nó pai na árvore.

A métrica utilizada em Ganesan *et al.* (2004) é baseada em uma função de avaliação que está relacionada à energia gasta nos nós sensores. Os resultados obtidos por Ganesan *et al.* (2004) mostraram ganhos significativos na redução do consumo de energia quando são utilizadas as estratégias de posicionamento dos nós propostas no trabalho em relação à implantação dos nós por processos aleatórios comumente utilizados. A métrica utilizada em Ganesan *et al.* (2004) utiliza o impacto da escolha do melhor posicionamento para implantação dos nós sensores na RSSF em relação ao ganho no consumo de energia e também proporciona a redução do número de nós utilizados.

Não foi utilizada simulação em Ganesan *et al.* (2004). No entanto, ele faz uso do MATLAB na análise dos processos de dados correlacionados e apresenta soluções analíticas para obter os seus resultados. A proposta apresentada em Ganesan *et al.* (2004) é importante para o trabalho ora proposto, pois mostra a importância de se utilizar a quantidade mínima de nós sensores numa RSSF. Apesar de o trabalho também propor objetivo semelhante com relação à quantidade de nós, serão utilizados algoritmos genéticos para obtenção da configuração de distribuição dos nós com a menor quantidade de nós que não afeta a área de cobertura.

O problema abordado em Chen e Chuah (2005) faz referência às formas de implantação dos nós sensores, que pode ser de forma aleatória ou determinística. Geralmente, são necessários menos nós sensores para realizar a mesma cobertura de área quando a forma de implantação é determinística. Chen e Chuah (2005) afirmam que o aumento na quantidade de nós sensores implantado numa RSSF provoca um ganho no tempo de vida útil da rede. No entanto, esse ganho tende a zerar à medida que a quantidade de nós aumenta muito.

O trabalho de Chen e Chuah (2005) propõe uma forma de determinar quantos nós sensores devem ser implantados e como implantá-los para maximizar a eficiência de utilização dos nós sensores em uma RSSF. A metodologia utilizada em Chen e Chuah (2005) foi baseada em duas estratégias. A primeira estratégia utiliza um algoritmo guloso para otimizar o posicionamento do nó sensor. Na segunda estratégia é utilizada uma aproximação numérica para determinar o número ideal de nós sensores para a RSSF.

A métrica utilizada em Chen e Chuah (2005) é baseada no custo por unidade de cada sensor. Para tanto, foi utilizada uma relação entre o custo para implantar cada novo nó sensor com o tempo de vida ganho com essa implantação. Assim, foi possível verificar a eficiência da utilização de mais um nó sensor na RSSF. Os resultados obtidos por Chen e Chuah (2005) mostram que a implantação de uma quantidade extremamente pequena, ou extremamente grande, de nós sensores é ineficiente em termos de aumento da vida útil dos nós sensores da RSSF.

Não foi utilizado ambiente de simulação em Chen e Chuah (2005), os resultados foram obtidos utilizando diretamente os algoritmos propostos no trabalho. Além de simulações Chen e Chuah (2005), também foram utilizados métodos numéricos para calcular o ganho obtido com a proposta.

A proposta apresentada em Chen e Chuah (2005), que determina a quantidade ideal de nós sensores utilizando como métrica o ganho no tempo de vida da rede em relação o custo de instalação, se mostrou como uma boa estratégia, a qual será utilizada no trabalho ora proposto. No entanto, serão utilizados algoritmos genéticos em substituição aos algoritmos gulosos.

O problema abordado em Dhillon e Chakrabarty (2003), faz referência à otimização da cobertura de uma RSSF levando em consideração restrições de detecção e as propriedades do terreno na qual a RSSF será implantada, como obstáculos entre os nós da rede que obstruem o campo de visão de sensoriamento. O trabalho de Dhillon e Chakrabarty (2003), propõe um modelo de detecção para o nó sensor e um algoritmo de colocação de nós sensores com uma visão “minimalista” de uma RSSF. Nessa rede, uma quantidade mínima de nós sensores é implantada. Para diminuir o custo de energia foi proposta a utilização de nós sensores inteligentes, que realizam um pré-tratamento nos dados coletados antes de transmiti-los.

Outra proposta do trabalho de Dhillon e Chakrabarty (2003), é a representação da área a ser coberta pela RSSF como uma grade com duas ou três dimensões. A metodologia utilizada em Dhillon e Chakrabarty (2003), utiliza dois algoritmos para resolver o problema de colocação de nós sensores, que otimiza a cobertura da área a ser monitorada pela RSSF. O primeiro algoritmo utiliza programação linear inteira (PLI) para otimizar o modelo de colocação dos nós sensores em grades de duas ou três dimensões. Resumidamente, os problemas de PLI consistem em um problema de Programação Linear (PL) em que todas ou algumas das suas variáveis são do tipo inteiro. A PL trata de problemas de otimização nos quais a função objetivo e as restrições são lineares.

O Segundo algoritmo é baseado nas soluções do problema da galeria de arte (AGP) para determinar a quantidade mínima de nós sensores. O problema AGP pode ser informalmente definido como o da determinação do número mínimo de guardas necessários para cobrir o interior de uma galeria de arte.

A métrica utilizada em Dhillon e Chakrabarty (2003), no modelo de detecção, foi a probabilidade de detecção de um alvo por um nó sensor que varia exponencialmente com a distância entre o alvo e o nó sensor. Para o algoritmo de colocação de nós sensores, a métrica utilizada foi a determinação do número mínimo de nós sensores e a sua localização na representação em grade da área.

Os resultados obtidos em Dhillon e Chakrabarty (2003), mostram que se não há obstáculos na área a ser monitorada e os nós sensores utilizados são idênticos, a colocação aleatória dos nós sensores mostra-se tão eficaz quanto à colocação baseada no modelo proposto. No entanto, quando a área apresenta obstáculos, a colocação aleatória apresentou desempenho significativamente inferior. Não foi utilizado ambiente de simulação em Dhillon e Chakrabarty (2003), os resultados foram obtidos utilizando diretamente os algoritmos propostos no trabalho.

A proposta apresentada em Dhillon e Chakrabarty (2003), será útil, pois a proposta ora apresentada pretende utilizar o modelo de detecção do nó sensor e a representa em grade do terreno. No entanto, pretende-se substituir o algoritmo utilizado por Dhillon e Chakrabarty (2003) devido seu alto custo computacional, que é baseado em programação linear.

O problema abordado em Hou et al (2006), também trata da cobertura de uma RSSF. No entanto, diferente dos trabalhos anteriores, Hou *et al.* (2006) têm

como objetivo melhorar a qualidade dos serviços das RSSF. A preocupação com a qualidade do serviço é relevante, pois aplicações como o monitoramento de risco necessitam de uma maior cobertura de informação para minimizar atrasos e perdas de dados.

O trabalho de Hou et al (2006), propõe a implantação de novos nós sensores para otimizar a cobertura da RSSF e um algoritmo em tempo polinomial para determinar a melhor posição para a implantação do novo nó. A metodologia utilizada em Hou *et al.* (2006) para implementar o algoritmo em tempo polinomial combina técnicas de geometria computacional e algoritmos gráficos, como o diagrama de *Voronoi*, triangulação de *Delaunay* e busca gráfica. Uma métrica utilizada em Hou et al (2006) foi o custo do caminho, que foi usada para determinar em qual local o novo nó será implantado. O resultado obtido em Hou et al (2006), utilizando o algoritmo de implantação proposto, se mostrou significativamente mais eficiente em comparação com os algoritmos de implantação heurísticos.

Não foi utilizado ambiente de simulação em Hou et al (2006), os resultados foram obtidos utilizando diretamente os algoritmos propostos no trabalho. A proposta apresentada em Hou et al (2006) será útil, pois aborda o problema da cobertura de uma RSSF utilizando como parâmetro a ser otimizado a qualidade do serviço, que é também uma dos parâmetros de estudo da proposta. No entanto, ressalta-se que se pretende utilizar algoritmos genéticos para determinar a melhor configuração para colocação dos nós.

Em Kim *et al.* (2011), é abordado o problema do tempo necessário para que os dados gerados nos nós sensores cheguem à estação base da RSSF. Existem muitas aplicações de RSSF, como a vigilância de ambientes e detecção de intrusos no campo de batalha, em que é extremamente importante saber qual o pior caso de latência na transmissão dos dados para se verificar se as restrições da aplicação serão atendidas.

O trabalho de Kim *et al.* (2011), propõe dois algoritmos gulosos de aproximação, o primeiro tem como objetivo resolver o problema de *k-center*, que na teoria dos grafos significa encontrar um conjunto de vértices *k* para o qual a maior distância de qualquer ponto para o seu vértice mais próximo é mínima. O segundo tenta resolver o problema de *k-Sink Placement Problem* (*k-SSP*) que foi formalmente

definido por Kim et al. (2011) como um problema de colocação de *k-sink* (nós sumidouros).

A metodologia utilizada por Kim *et al.* (2011), para resolver o problema de latência na transmissão dos pacotes de dados, consiste no uso de múltiplos pontos de coleta, também chamados de nós sorvedouro, que limitam a distância de transmissão dos nós sensores e possibilita a implementação da comunicação multi-hop. Também foi utilizado o algoritmo para buscar a melhor posição de implantação da estação base.

Algumas das métricas utilizadas em Kim *et al.* (2011) foram a distância euclidiana entre dois nós u e v ($\text{Euclidis}(u,v)$) e a distância *hop* entre dois nós u e v ao longo do caminho mais curto entre eles ($\text{Hopdist}(u,v)$). Os resultados obtidos por Kim *et al.* (2011) mostraram que os custos dos dois algoritmos são menores quando o número de nós sensores é grande. No entanto, o algoritmo *k-SSP* supera o *k-Center* sob qualquer quantidade de nós sensores principalmente quando essa quantidade é pequena. Portanto, foi concluído que o algoritmo *k-SSP* é melhor que o *k-Center*. Outro benefício que recentemente vem sendo constatado com a utilização de múltiplos sumidouros é o aumento na escalabilidade das RSSF.

Kim *et al.* (2011) não utilizaram nenhum ambiente de simulação para realizar suas simulações, executando os algoritmos propostos diretamente no computador. As propostas apresentadas em Kim *et al.* (2011) serão úteis ao trabalho proposto, pois além de mostrar a importância do atraso para várias aplicações das RSSF, também propuseram algoritmos que se mostraram eficientes na redução da latência. No ora trabalho proposto será implementado um algoritmo genético para reduzir a latência da transmissão na RSSF.

A porcentagem de área coberta é umas das principais métricas utilizadas para determinar a qualidade de serviço de uma RSSF. Segundo Quintão *et al.* (2004), as RSSF com alta densidade de nós estão sujeitas a problemas como congestionamento do meio e colisões de pacotes, o que provoca o aumento da latência e do consumo de energia com retransmissões de pacotes perdidos. No entanto, uma rede com alta densidade também conta com mais energia disponível, dispõe de melhores mecanismos de tolerância a falhas, uma maior precisão no sensoriamento e possibilita múltiplos caminhos para a transmissão dos dados.

Em Quintão *et al.* (2004), são propostos dois algoritmos, um exato e outro baseado em computação evolutiva para tratar o problema de cobertura. O algoritmo baseado em computação evolutiva é usado para auxiliar o gerente ou observador da rede a determinar quais nós devem ser utilizados naquele momento. O algoritmo exato, que é baseado em programação linear, é usado para determinar o número mínimo de nós sensores que deverão ser utilizados. A metodologia em Quintão *et al.* (2004) utiliza um mecanismo de gerenciamento de densidade para controlar quantos nós serão ativos e quantos permanecerão inativos de acordo com as necessidades da aplicação. A métrica utilizada no trabalho de Quintão *et al.* (2004), para o escalonamento dos nós sensores que serão ativados na RSSF, visa obter a maior cobertura possível com a menor quantidade de nós sensores, utilizando sempre os nós sensores que tenham a maior quantidade de energia armazenada. Ou seja, tem-se uma função de aptidão que maximiza área de cobertura e a energia armazenada.

Os resultados obtidos em Quintão *et al.* (2004), mostram vantagens significativas no uso de algoritmos genéticos na implementação de um mecanismo de gerenciamento de densidade de nós em comparação com o uso de algoritmos exatos, como os baseados em programação linear. Os mesmos resultados foram obtidos utilizando as duas abordagens. No entanto, o custo de processamento dos algoritmos genéticos é muito inferior a dos baseados em programação linear.

Os resultados em Quintão *et al.* (2004) foram obtidos utilizando o CPLEX 7.0 para executar a solução baseada em programação linear e a solução que usa algoritmos genéticos foi executada diretamente no computador. O trabalho de Quintão *et al.* (2004) será útil, pois apresenta uma solução para o problema de cobertura e escalonamento utilizando algoritmos genéticos. No entanto, a função de avaliação adotada para o algoritmo faz referência à quantidade de energia e à área de cobertura, o que se diferencia da proposta ora apresentada, que também leva em consideração a qualidade do serviço.

Em Rocha e Silva (2011), é abordado o problema para determinação do melhor posicionamento de diversos tipos de nós sensores, de modo a cobrir toda a área especificada pela aplicação com o menor custo possível. Em outras palavras, é um problema de cobertura.

Rocha e Silva (2011), propôs um modelo baseada em algoritmos genéticos (AG) para encontrar uma configuração que apresente a melhor cobertura da região

de interesse com o menor custo possível. Foi proposta também a utilização da API do Google Maps de modo a facilitar o uso de mapas eletrônicos na determinação da área a ser monitorada e também para a visualização da solução fornecida pelo AG.

Na metodologia utilizada em Rocha e Silva (2011), os possíveis locais de posicionamento dos nós sensores são determinados calculando-se o retângulo que envolve a área a ser monitorada e o intervalo entre os nós sensores utilizando informações fornecidas pelo usuário. Também são utilizados três tipos de nós sensores com diferentes alcance e custo. Todas essas informações são utilizadas na determinação da função de avaliação que será usada pelo algoritmo genético na determinação da melhor configuração de ativação.

A métrica utilizada em Rocha e Silva (2011), é baseada nos custos de cada nó utilizado na implantação e na qualidade da cobertura. Nos resultados obtidos em Rocha e Silva (2011) verificou-se que o algoritmo genético proposto atingiu seu objetivo de determinar a melhor configuração, que propicia a solução de menor custo. Outro resultado importante foi o baixo tempo computacional demandado pelo algoritmo genético.

Em Rocha e Silva (2011) foi utilizada a linguagem Java para implementar as técnicas propostas em seu trabalho e também usou a API do Google Maps para entrada dos dados e sua apresentação. O trabalho de Rocha e Silva (2011) possibilitou verificar a eficiência dos algoritmos genéticos, bem como da API do Google Maps que facilitou a entrada de dados para o AG e a visualização dos resultados. Neste trabalho, foram considerados outros parâmetros além dos que foram usados em Rocha e Silva (2011).

Em Tuna *et al.* (2012), é abordado o problema de realizar o levantamento do cenário em que serão implantados os nós da RSSF, quando estes cenários estão contaminados com altos níveis de radiação. O objetivo é de determinar previamente a configuração de implantação dos nós sensores, de forma a prolongar a vida útil da rede.

Na metodologia adotada em Tuna *et al.* (2012), o levantamento do cenário e, posteriormente, a implantação dos nós sensores, são realizados com auxílio de robôs móveis. Estes robôs móveis são controlados remotamente por um operador humano ou de forma autônoma. No modo de controle autônomo, foram utilizadas duas estratégias. Na primeira estratégia, *Frontier Exploration* (FE), foi utilizado um

grupo de robôs móveis que são direcionados para as regiões de fronteira do cenário, a partir de um mapa de exploração. Nesta estratégia, os robôs compartilham as informações capturadas sobre o cenário, mas suas decisões continuam sendo independentes e baseadas no mapa de exploração.

Na segunda estratégia, *Market Driven Exploration* (MDE), foi adotado o princípio da divisão e conquista, cuja tarefa de exploração é dividida em subtarefas e cada robô móvel é responsável pela exploração de uma parte do cenário. Nessa estratégia, os robôs evitam deslocar-se além do raio de alcance de comunicação.

A métrica utilizada por eles é baseada no tempo total gasto para a exploração total do cenário. Os resultados apontam que a estratégia FE é mais rápida do que a estratégia MDE. Outra métrica utilizada é o erro de posicionamento do robô. Os resultados mostraram que os erros de localização do robô são inferiores a 0,2 metros nas coordenadas de longitude e latitude.

Tuna *et al.* (2012) propõem, em seu trabalho, que sejam usados robôs para realizar a exploração do cenário e a implantação dos nós sensores da RSSF. A estratégia utilizada por eles é uma solução a ser adotada para a implantação dos nós sensores nas posições determinadas pelo arquivo de configuração resultante da utilização do modelo proposto deste trabalho.

Em Elmagid *et al.* (2012), é proposto uma solução para tratar o problema da cobertura da área do cenário que está sendo monitorada, levando em consideração, além da cobertura, outros fatores, como: conectividade entre os nós sensores; confiabilidade do dispositivo; consumo de energia do nó sensor; e a divisão da área do cenário em subáreas.

Elmagid *et al.* (2012) sugerem uma abordagem diferente para a implantação dos sensores utilizados na detecção de radiação. Existem muitos parâmetros que não são normalmente considerados, tais como: a importância das subáreas; as distâncias de comunicação e do alcance de detecção dos sensores; e a energia residual do nó sensor. O tratamento de todos esses parâmetros aumenta consideravelmente a complexidade do problema. Portanto, foi utilizado um algoritmo de busca baseado na heurística dos AG.

Na metodologia empregada em Elmagid *et al.* (2012), o AG utilizado na busca da configuração de implantação dos sensores foi combinado com a lógica *fuzzy*, para melhorar o seu desempenho. A taxa de cruzamento e a taxa de mutação

são parâmetros que desempenham uma função crucial na diversificação dos genes do cromossomo de um AG. Na implementação do módulo responsável por esses dois parâmetros, foram utilizadas técnicas de lógica *fuzzy*, o que proporcionou um considerável ganho de desempenho.

A métrica adotada em Elmagid *et al.* (2012) é baseada na taxa de cobertura da área monitorada. Outra métrica utilizada é a qualidade dos dispositivos usados, determinada pela maior capacidade de energia, robustez, alcance de comunicação e de detecção, e pela sua localização nas áreas mais importantes para fins de detecção da radiação.

Nos resultados apresentados por Elmagid *et al.* (2012), pode-se destacar a comparação realizada entre o desempenho do algoritmo, quando se utilizou um AG padrão em relação ao AG modificado por lógica *fuzzy*. Os resultados apontam que a utilização do AG modificado proporcionou a convergência da busca em apenas 200 gerações, e utilizando o AG padrão, foram necessárias 300 gerações.

O trabalho de Elmagid *et al.* (2012) proporcionou um possível caminho para o melhoramento do modelo proposto neste trabalho, pois a modificação dos parâmetros de taxa de mutação e cruzamento pode aumentar o desempenho do algoritmo, principalmente quando o modelo é submetido a busca por uma melhor taxa de cobertura.

No Quadro 4.1, é apresentado um resumo dos trabalhos supracitados, com destaque para a solução proposta.

4.1 Discussão Geral

Neste capítulo, foram apresentados os trabalhos relacionados a esta pesquisa. O levantamento destes trabalhos serviu de base para a construção do modelo proposto no Capítulo 5. Em especial, pode-se destacar o trabalho de Quintão *et al.* (2004) que foi fundamental na elaboração da função de avaliação do AG proposto neste trabalho.

No próximo capítulo, será apresentado o modelo RSSF-AG que foi elaborado a partir do estudo dos trabalhos apresentados neste capítulo.

Quadro 4.1 Resumo dos trabalhos relacionados

Artigo	Problema	Proposta	Metodologia
Ganesan <i>et al.</i> (2004)	Otimização da colocação dos nós sensores em uma RSSF.	Uma solução analítica para otimização de colocação de nós sensores em casos unidimensionais simplificados.	Árvore de roteamento para representar a estrutura da rede, Técnicas de agregação de dados para reduzir a quantidade de transmissões.
Chen e Chuah (2005)	Determina a quantidade de nós a ser implantada.	Maximizar a eficiência de utilização dos nós sensores.	Algoritmo guloso e aproximação numérica.
Dhillon e Chakrabarty (2003)	Otimização da cobertura da RSSF.	Modelo de detecção para o nó sensor, algoritmo de colocação de nós sensores e uma representação em grade para a área de cobertura.	Uso de programação linear inteira e das soluções para o problema da galeria de arte (AGP).
Hou <i>et al.</i> (2006)	De cobertura da RSSF com o objetivo de melhorar a qualidade dos serviços.	Implantação de novos nós sensores e um algoritmo para determinar a melhor posição para colocação.	Combinação de técnicas de geometria computacional e algoritmos gráficos.
Kim <i>et al.</i> (2011)	Latência da transmissão dos dados entre os nós sensores e a estação base da RSSF.	Algoritmo guloso de aproximação para problema de <i>k-Center</i> e <i>k-SSP</i> .	Uso de comunicação <i>multi-hop</i> e busca da melhor posição para implantação da estação base.
Quintão <i>et al.</i> (2004)	Otimização da cobertura da RSSF.	Algoritmo evolutivo para definir a melhor configuração a ser utilizada. Algoritmo programação linear para determinar o número mínimo de nós sensores a serem usado.	Utiliza um mecanismo de gerenciamento para controlar os nós sensores que serão ativados.
Rocha e Silva (2011)	Problema de cobertura. Determinação da melhor configuração.	Algoritmo genético para determinar a configuração ideal e uso da <i>API</i> do <i>Google Maps</i> .	A função de avaliação do algoritmo genético usa informações da área a ser monitorada e dos tipos de sensores disponíveis.
Tuna <i>et al.</i> (2012)	Levantamento da topologia do cenário a ser monitorado para reduzir o consumo de energia.	Utilização de robôs para fazer o levantamento da topologia e, depois, estes mesmos robôs foram responsáveis pela instalação dos nós sensores nas posições previamente determinadas.	Foi proposto o método de controle autônomo, em que é utilizado um mapa de exploração <i>Frontier Exploration</i> (FE) e <i>Market Driven Exploration</i> (MDE) e o método de controle tele operado.
Elmagid <i>et al.</i> (2012)	Problema de cobertura.	Utilização de um algoritmo genético modificado por meio de lógica <i>fuzzy</i> para obter a melhor configuração de implantação dos nós sensores.	A lógica <i>fuzzy</i> foi empregada para alterar dinamicamente os valores dos parâmetros de probabilidade de mutação e cruzamento do AG.
Solução Proposta	Otimização da configuração de implantação inicial dos nós sensores com objetivo de aumentar a taxa de cobertura e o tempo de vida da RSSF e reduzir o TAFF e a TPP.	Um modelo de busca baseado em AG em conjunto com simulações para encontrar a melhor configuração de implantação dos nós sensores de RSSF.	Um algoritmo de busca baseado em AG utilizando as métricas taxa de cobertura, tempo de vida, TAFF e TPP.

Quadro 4.1 Resumo dos trabalhos relacionados (Continuação)

Artigo	Métricas	Resultados	Simuladores
Ganesan <i>et al.</i> (2004)	Função de custo da energia gasta nos nós sensores. Números de nós utilizados.	Ganhos na redução do consumo de energia e a redução do número de nós sensores.	Foi utilizado o MATLAB para rodar os algoritmos proposto.
Chen e Chuah (2005)	Custo do nó sensor por tempo de vida ganho.	Mostra que existe uma quantidade ideal de nós sensores a ser implantada que proporciona o melhor ganho tanto de vida útil como de custo da RSSF.	Os resultados foram obtidos rodando os algoritmos proposto sem uso de simuladores.
Dhillon e Chakrabarty (2003)	Probabilidade de detecção de um alvo por um sensor. Número mínimo de sensores e sua localização.	Quando a área a ser monitorada contiver obstáculos, o modelo proposto apresentou ganhos.	Os resultados foram obtidos rodando os algoritmos proposto sem uso de simuladores.
Hou <i>et al.</i> (2006)	Custo do caminho.	O algoritmo proposto se mostrou mais eficientes que os algoritmos heurísticos.	Os resultados foram obtidos rodando os algoritmos proposto sem uso de simuladores.
Kim <i>et al.</i> (2011)	Distância euclidiana e <i>hop</i> entre dois nós u e v.	O algoritmo k-SSP se mostrou mais eficientes que o <i>k-Center</i> .	Os resultados foram obtidos rodando os algoritmos proposto sem uso de simuladores.
Quintão <i>et al.</i> (2004)	Área de cobertura e Energia acumulada no nó sensor.	Os resultados mostram que o algoritmo genético chegou à solução semelhante à proposta pelo algoritmo exato, com uma necessidade de processamento bem menor.	Os resultados foram obtidos utilizando o CPLEX 7.0 para o algoritmo exato e diretamente no computador para o algoritmo evolutivo.
Rocha e Silva (2011)	Qualidade da cobertura e o custo de cada nó utilizado.	Foi verificada a eficiência dos algoritmos genéticos para encontrar a melhor configuração dos nós sensores.	Java para implementar o algoritmo genético e a API do Google Maps para entrada e apresentação dos dados.
Tuna <i>et al.</i> (2012)	Tempo utilizado na exploração da área do cenário e o erro de posicionamento.	O método de controle FE se mostrou mais rápido na exploração em relação ao MDE e o erro médio de posicionamento foi de apenas 0,2 metros.	Os resultados foram obtidos utilizando o simulador Pioneer 3-DX robots.
Elmagid <i>et al.</i> (2012)	Taxa de cobertura e a qualidade dos dispositivos utilizados na rede.	Os resultados demonstraram um ganho de 33,0% no desempenho do AG modificado em relação ao AG padrão.	Os resultados foram obtidos executando os algoritmos propostos sem uso de simuladores.
Solução Proposta	Taxa de cobertura, Tempo de vida, TAFF e TPP	Os resultados demonstraram que a utilização do modelo proporcionou uma melhora das métricas de 16,0%.	Simulador OMNeT 4++ 4.2.2 e framework INET 2.0.0

Quadro 4.1 Resumo dos trabalhos relacionados (Continuação)

Artigo	Contribuição	Diferenças
Ganesan <i>et al.</i> (2004)	Conhecimento sobre a quantidade ideal de nós a ser utilizado em uma RSSF.	Substituição da solução analítica por um algoritmo genético.
Chen e Chuah (2005)	A estratégia utilizada como parâmetro para determinar a quantidade de nós sensores.	Substituição do algoritmo guloso por um genético.
Dhillon e Chakrabarty (2003)	Modelo de detecção e a representação em grande da área a ser monitorada.	Substituição do algoritmo baseado em programação linear.
Hou <i>et al.</i> (2006)	Uso do parâmetro de QoS para determinar a melhor cobertura para a RSSF.	Serão usados algoritmos genéticos.
Kim <i>et al.</i> (2011)	Um modelo para reduzir a latência.	Serão usados algoritmos genéticos.
Quintão <i>et al.</i> (2004)	Uma solução para o problema de cobertura e escalonamento baseado em algoritmo genético.	Será usada outra função avaliação para garantir a qualidade do serviço.
Rocha e Silva (2011)	Foi verificada a eficiência dos algoritmos genéticos e a utilização da API do Google Maps.	Serão considerados outros parâmetros para se determinar a melhor configuração.
Tuna <i>et al.</i> (2012)	Uma estratégia para obter a topografia da área a ser monitorada e posteriormente a automação na implantação dos nós sensores em seus locais.	Apesar da proposta também incluir a redução do consumo de energia obtida com o posicionamento dos nós sensores, ela não leva em consideração todas as métricas que são tratadas neste trabalho.
Elmagid <i>et al.</i> (2012)	Apresentação de um método para aumentar o desempenho dos AG a partir da utilização de parâmetros que são dinamicamente modificados.	O objetivo do artigo era melhorar o nível da taxa de cobertura e ao mesmo tempo a qualidade da rede, utilizando os melhores dispositivos nas subáreas mais importantes. No entanto, eles não buscaram diretamente o aumento da QoS e tempo de vida da RSSF.
Solução Proposta	Um modelo para buscar a configuração de implantação dos nós sensores de uma RSSF que proporcionou um ganho nas métricas taxa de cobertura, tempo de vida, TAFF e TPP.	Neste modelo a função de avaliação foi configurada para realizar buscas com múltiplas métricas, considerando o aumento do tempo de vida e o QoS da RSSF.

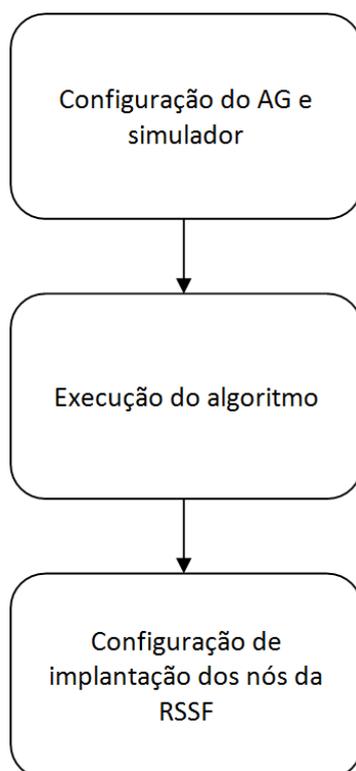
5 Descrição do Modelo RSSF-AG

Este capítulo tem como objetivo apresentar o modelo proposto neste trabalho, que tem por fim a realização da busca de uma configuração que proporcione o aumento da taxa de cobertura, do tempo de vida da RSSF e reduzindo o TAFF e a TPP empregando AG. Nesta busca, são utilizadas as 4 (quatro) métricas: taxa de cobertura, tempo de vida, TAFF e TPP.

5.1 Modelo RSSF-AG

O modelo é composto por 3 (três) fases. Na primeira fase, são definidos os valores para os parâmetros de configuração do AG e do simulador. Na segunda, o algoritmo proposto é executado para realizar a busca da melhor configuração de implantação dos nós sensores da RSSF. Na terceira e última fase, é gerado um arquivo com as coordenadas dos nós sensores que compõem a RSSF. Na Figura 5.1 é ilustrado o fluxograma com as 3 (três) fases do modelo descrito.

Figura 5.1 Fluxograma das fases do modelo proposto



5.1.1 Configuração do AG e do Simulador

Na primeira fase do modelo apresentado neste trabalho, utilizou-se o aplicativo desenvolvido para criação do arquivo de configuração do algoritmo proposto. Nessa etapa, são fornecidas as informações sobre os parâmetros: da área do cenário; de configuração do AG; e da configuração do simulador.

Para gerar o arquivo de configuração é necessário que usuário já tenha definido as informações que caracterizam o problema a ser tratado pelo modelo proposto. As principais informações requeridas nessa etapa são divididas pelos seguintes grupos:

- Parâmetros de configuração das métricas: determina o escopo a ser alcançado com a busca. Para cada uma das 4 (quatro) métricas é definido um valor que quantifica o quão importante a métrica é para a busca;
- Parâmetros do cenário: para caracterizar o cenário é preciso informação sobre as dimensões de área, o relevo e o posicionamento de obstáculos, quando houver;
- Parâmetro do AG: usado para configurar o comportamento do AG durante a execução da busca;
- Parâmetros do simulador: utilizado para configurar o simulador com informações sobre as características dos nós sensores utilizados na rede, o tempo de simulação, o tipo de protocolos de roteamento e o tipo de aplicação a ser usada na simulação.

Após o fornecimento das informações ao aplicativo, é gerado o arquivo de configuração a ser utilizado na execução da busca. No APÊNDICE B, é apresentado com mais detalhes o aplicativo de geração do arquivo de configuração e os comandos necessários para sua utilização.

5.1.2 Execução do Algoritmo

A segunda fase é responsável por realizar a busca de uma configuração de implantação dos nós sensores da RSSF que proporcione o valor máximo para Equação (5.5). Com esse intuito, é utilizado um algoritmo de busca baseado nos AG. As configurações de implantação dos nós resultantes dessa busca são avaliadas por

meio de simulações, utilizando o simulador de evento discreto OMNeT++, em conjunto com o *framework* INET.

Todos os indivíduos (uma configuração de implantação dos nós sensores) gerados pelo algoritmo de busca são submetidos a uma simulação, para que seja possível quantificar: a taxa de cobertura; o somatório da energia residual de cada nó sensor; o TAFF da rede; e o TPP da rede.

Os valores das métricas que forem obtidos da extração dos dados gerada pelo simulador são utilizados no cálculo da função de avaliação proposto neste modelo. Utilizando o valor da função de avaliação, os indivíduos são classificados do maior valor para o menor. Sendo que, o indivíduo com maior valor é aquele que apresenta a melhor configuração de implantação e o indivíduo com menor valor, aquele com a pior configuração. Essa classificação será utilizada pelo módulo de elitismo para determinar quais indivíduos deverão ser preservados para a próxima geração.

No modelo proposto nesse trabalho, ficou definido, com a utilização de experimentos, que o uso do elitismo irá proporcionar melhores resultados nas buscas realizadas. O valor obtido da função de avaliação também é utilizado no módulo de seleção para calcular a probabilidade do indivíduo ser selecionado.

O módulo de seleção realiza um sorteio para escolher quais indivíduos serão utilizados pelo módulo de cruzamento na geração dos indivíduos da nova geração. Após a criação dos novos indivíduos, a geração anterior é descartada. As exceções desse descarte são para os indivíduos selecionados pelo módulo de elitismo, que serão preservados para a nova geração.

Os novos indivíduos também são submetidos ao módulo de mutação, responsável pela modificação de um bit do vetor binário que representa o indivíduo. A aplicação do módulo de mutação depende do parâmetro de probabilidade de mutação, definido na primeira fase.

Ao final da execução da busca, o modelo proposto gera um arquivo contendo o vetor binário com as informações de configuração de implantação dos nós sensores que proporcionou o maior valor para a função de avaliação utilizada nesse modelo. Além do vetor binário, também são disponibilizados: o número de indivíduos criados; o número de ocorrência de mutações; o número de gerações; os valores para as 4 (quatro) métricas; e o tempo de execução do algoritmo.

O comando utilizado para executar o algoritmo de busca do modelo proposto neste trabalho é mostrado com detalhes no APÊNDICE B.

5.1.3 Configuração de Implantação dos Nós da RSSF

A terceira e última fase do modelo é responsável por gerar o arquivo com os dados de posicionamento de cada nó sensor da RSSF. Para sua criação, foi desenvolvido um aplicativo que consegue abrir o arquivo gerado na segunda fase e extrair o vetor binário com as informações de posicionamento de cada nó sensor no cenário proposto.

Ao final da execução do aplicativo de extração de dados, ele gera um arquivo contendo as informações das coordenadas de localização dos nós sensores no cenário. Esse arquivo deve ser utilizado na implantação dos nós sensores da RSSF em estudo.

A apresentação do resultado da busca realizada pelo modelo sugerido nesse trabalho é um arquivo com as informações das coordenadas de cada nó da rede. No entanto, essas informações podem ser utilizadas para gerar outras formas de apresentação dos resultados, como o mostrado na Figura 6.12.

O comando utilizado para executar aplicativo de extração de dados é mostrado com detalhes no APÊNDICE B.

5.2 Algoritmo Genético Proposto

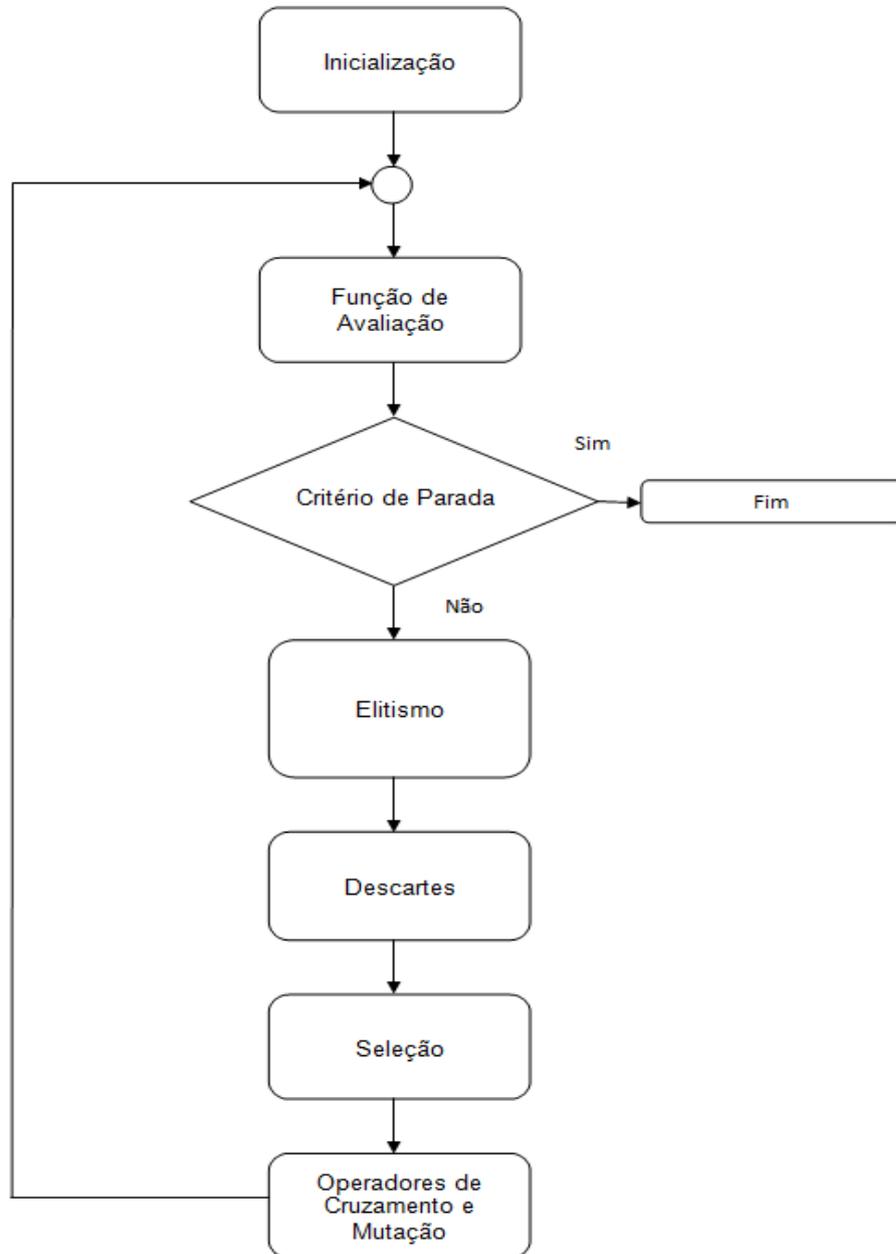
Com o propósito de buscar a melhor configuração de implantação dos nós sensores, propõe-se a utilização de um AG (Figura 5.2).

A busca pela melhor configuração de implantação dos nós sensores na RSSF apresenta as seguintes etapas:

- Inicialização: inicializa a primeira geração;
- Função de avaliação: calcula o valor de aptidão de cada indivíduo;
- Critério de parada: verifica se algum critério de parada foi alcançado;
- Elitismo: mantém os melhores indivíduos de cada geração para a próxima;
- Descarte: descarta os indivíduos menos aptos;
- Seleção: seleciona, dentro da população, os indivíduos para reprodução;

- Operadores de cruzamento e mutação.

Figura 5.2 Fluxograma simplificado do modelo proposto



A função de avaliação proposta para o modelo RSSF-AG tem como base a função de avaliação utilizada em Quintão et al. (2004). A função de Quintão et al. (2004) utiliza duas métricas (área de cobertura e energia residual) para calcular o valor da função de avaliação. No modelo RSSF-AG, além das duas métricas utilizadas por Quintão et al. (2004), foram acrescentadas outras duas (TAFF e TPP).

As duas novas métricas têm como objetivo considerar a importância da qualidade do serviço da RSSF na busca da configuração de implantação, pois em muitas aplicações, o TAFF e a TPP podem causar falha na aplicação que estiver fazendo uso da RSSF. Para determinar se todas as métricas utilizadas no modelo RSSF-AG são representativas, foi realizado um estudo estatístico a partir de uma análise inferencial.

A análise inferencial mostrou que as seguintes métricas: taxa de cobertura; energia residual; e TAFF; são representativas para o modelo. Apesar de a análise inferencial ter indicado que a métrica TPP não é representativa, ainda assim o seu uso é importante, pois em caso específicos, onde a TPP é a única métrica de interesse da aplicação, se faz necessário que o modelo tenha a flexibilidade de realizar uma busca apenas para métrica TPP. O estudo estatístico pode ser verificado na seção A.9 do APÊNDICE A.

5.2.1 Inicialização

Na etapa de inicialização, são gerados os indivíduos (cromossomos) que irão compor a primeira geração a ser submetida à função de avaliação. A técnica utilizada neste trabalho é baseada em escolhas aleatórias. Neste tipo de inicialização, os indivíduos que compõem a população são selecionados aleatoriamente dentro do espaço de soluções possíveis. Este método apresenta como vantagem uma maior variabilidade dos membros escolhidos na primeira geração, o que possibilita que novas soluções sejam encontradas a cada nova execução do modelo (GUIMARÃES e RAMALHO, 2001). No Exemplo 5.1 é apresentado um vetor binário de um cromossomo.

Exemplo 5.1 Vetor binário da representação de um cromossomo

```
00100001100101101011111011000001101011101000001110010011000001110101001100
01000011101101101100111101000000111111100001011110010111010000010111001110
1010111001101001100100110010110010111110010111101101011010010111000111
100000000011110101101011101100110100101000100110110110010110000000100111100
1101111110111010101101000001111000000001010000000001010010010010111000111
110000011110111001000110000010011001100001001001010100111000010100111000000
011110111001111100111110110101100100110000110101101001011000110011101010001
1001000100100100110011111011110101011001101101000
```

Neste exemplo, o vetor binário é formado por 36 nós sensores. Cada nó utiliza 16 (dezesseis) bits para representar o par de coordenadas (x, y) , sendo 8 (oito) bits para a coordenada x , e 8 (oito) bits para a coordenada y , totalizando 576 bits para cada vetor binário (indivíduo).

A quantidade de bits utilizados para representar o par de coordenadas no presente modelo pode ser alterada a critério do usuário. No entanto, o valor de 8 (oito) bits foi utilizado nos experimentos realizados nesse trabalho, já que ele representa as posições de instalação de forma satisfatória, sem provocar uma carga excessiva de processamento para sua manipulação.

Utilizando a Equação (5.1), é possível calcular a resolução ou a menor distância entre cada possível posição de implantação de um nó sensor no cenário.

$$r = \frac{C}{2^n} \cdot \tag{5.1}$$

em que:

$C \in \mathbb{R}^+$, é o comprimento ou largura da área do cenário.

$n \in \mathbb{N}^+$, é o número de bits utilizados na representação binária.

Neste trabalho, foram utilizados 8 (oito) bits para representar a posição de cada coordenada do nó sensor e um cenário com 1000,0 metros de comprimento e largura. Aplicando-se na Equação (5.1), obtém-se uma resolução de aproximadamente 4 (quatro) metros que é a distância mínima entre cada possível posição de instalação de um nó sensor.

$$\begin{aligned} C &= 1000,0 \text{ m}, \\ n &= 8, \\ r &= \frac{1000,0 \text{ m}}{2^8} = \frac{1000,0 \text{ m}}{256} = 3,9 \text{ m}. \end{aligned}$$

A quantidade de nós sensores foi definida levando-se em consideração a distância do alcance do rádio de comunicação utilizado no dispositivo, que faz uso da tecnologia WLAN (IEEE802.11). Esta tecnologia tem um alcance que varia entre 100,0 e 500,0 metros (TANENBAUM, 2003).

A Equação (5.2) foi utilizada para calcular a área de uma circunferência, que é uma aproximação adequada para área de cobertura de comunicação de um nó sensor. A Equação (5.3) foi utilizada para calcular a área do cenário.

Com as áreas do cenário e da cobertura de cada nó sensor é possível calcular, a partir da Equação (5.4), um valor aproximado do número de nós sensores que serão necessários para obter uma taxa de cobertura de 100,0%.

$$A_c = \pi r^2, \quad (5.2)$$

$$A_r = l * h. \quad (5.3)$$

$$N = \frac{A_r}{A_c} + Err \quad (5.4)$$

em que:

$r \in \mathbb{R}^+$, é o comprimento do raio de alcance de comunicação do nó sensor.

$l \in \mathbb{R}^+$, é o comprimento da área do cenário.

$h \in \mathbb{R}^+$, é a largura da área do cenário.

$N \in \mathbb{N}^+$, é o número total de nós sensores da RSSF.

$Err \in \mathbb{N}^+$, é o número de nós sensores a ser acrescentado para compensar as perdas decorrentes das aproximações utilizadas para o calculo de cobertura.

Utilizando a menor distância de alcance dos dispositivos *WLAN* (100,0 metros) e as dimensões do cenário (1000,0 x 1000,0 metros), foi possível calcular a quantidade aproximada de nós sensores necessários para cobrir a área do cenário. O valor do número de nós sensores a ser acrescentado foi obtido empiricamente e o resultado que proporcionou os melhores ganhos para o algoritmo foi uma quantidade igual a 4 (quatro). Estes 4 (quatro) nós serão somados ao valor obtido da divisão da área do retângulo pela área da circunferência.

$$r = 100,0 \text{ m},$$

$$l = h = 1000,0 \text{ m},$$

$$A_c = \pi(100,0 \text{ m})^2 = 31415,9 \text{ m}^2,$$

$$A_r = 1000,0 \text{ m} * 1000,0 \text{ m} = 1000000,0 \text{ m}^2.$$

$$N = \frac{A_r}{A_c} + Err = \frac{1000000,0 \text{ m}^2}{31415,9 \text{ m}^2} + 4 = 35.8$$

Foram utilizados 36 nós sensores para cobrir a área do cenário. Este valor é uma aproximação obtida a partir da Equação (5.4).

5.2.2 Função de Avaliação

A função de avaliação precisa representar matematicamente a melhor forma de avaliar o grau de aptidão do cromossomo ao problema, refletindo as necessidades do problema de forma simples e direta (LINDEN, 2006), pois ela indicará qual indivíduo sobreviverá para as próximas gerações (elitismo) e influenciará na escolha dos pais (seleção), que irão fornecer material genético para os descendentes (cruzamento).

Neste trabalho, a função de avaliação foi moldada para atender os 4 (quatro) requisitos de interesse do trabalho. O principal deles foi aumentar o total de energia residual⁷ de toda a rede, sem penalizar a QoS e, também, ter a maior área de cobertura. Com estes objetivos, foi proposta uma função de avaliação que visa buscar um valor máximo para esta função, e que depende das seguintes variáveis: energia; taxa de cobertura; TAFF; e TPP.

Em Quintão *et al.* (2004), foi utilizada uma função de avaliação que atende, em parte, ao objetivo deste trabalho e, por este motivo, na Equação (5.5) proposta foram acrescentados outros 2 (dois) termos referentes às métricas TAFF e TPP. Esses termos foram propostos neste trabalho para minimizar os problemas com o atraso na transmissão dos dados e com a perda de pacotes de dados. A Equação (5.5) tem a restrição de que $\alpha + \beta + \gamma + \delta = 1,0$.

$$\begin{aligned}
 & f(A, ER, TAFF, TPP) \\
 &= \alpha \left(\frac{1}{A_{Total}} \sum_{k=1}^n A_k \right) + \beta \left(\frac{1}{ER_{Inicial}} \sum_{k=1}^n ER_k \right) \\
 &+ \gamma \left(1 - \frac{1}{n} \sum_{k=1}^n TAFF_k \right) + \delta \left(1 - \frac{1}{TPP_{Total}} \sum_{k=1}^n TPP_k \right)
 \end{aligned} \tag{5.5}$$

⁷ Energia resultante da diferença entre a energia inicial existente antes da execução da simulação e a consumida durante a simulação.

em que:

- $\alpha \in \mathbb{R}^+$, é o fator que determina o peso da variável de taxa de cobertura na função de avaliação, variando entre 0 (zero) e 1 (um);
- $A_{Total} \in \mathbb{R}^+$, é a área total a ser coberta pelos nós sensores;
- $A_k \in \mathbb{R}^+$, é a área efetivamente coberta por cada nó sensor;
- $\beta \in \mathbb{R}^+$, é o fator que determina o peso da variável de energia na função de avaliação, variando entre 0 (zero) e 1 (um);
- $ER_{Inicial} \in \mathbb{R}^+$, é o somatório da energia residual inicial de todos os nós sensores;
- $ER_k \in \mathbb{R}^+$, é a energia residual de cada nó sensor após executar a simulação;
- $\gamma \in \mathbb{R}^+$, é o fator que determina o peso da variável latência na função de avaliação, variando entre 0 (zero) e 1 (um);
- $n \in \mathbb{N}^+$, é o número de nós sensores utilizados na RSSF;
- $TAFf_k \in \mathbb{R}^+$, é o somatório dos tempos de atraso da rede;
- $\delta \in \mathbb{R}^+$, é o fator que determina o peso da variável TPP na função de avaliação, variando entre 0 (zero) e 1 (um);
- $TPP_{Total} \in \mathbb{R}^+$, é o somatório de todos os pacotes enviados pelos nós sensores;
- $TPP_k \in \mathbb{R}^+$, é o número de pacotes enviados por cada nó sensor efetivamente recebidos pelo nó sorvedouro.

Os dados utilizados pela função de avaliação são gerados a partir de simulações. Para tanto, foi utilizado o simulador de eventos discretos, OMNeT++, em conjunto com o *framework* INET.

O ambiente de simulação de eventos discreto OMNeT++ foi inicialmente desenvolvido para simulações de redes de comunicação, multiprocessadores e outros sistemas distribuídos (VARGA e HORNIG, 2008). No entanto, desde a sua concepção, OMNeT++ foi projetado para ser o mais genérico possível, o que proporcionou o desenvolvimento de vários *frameworks* para os mais diversos fins, entre os principais, tem-se o Castalia, MiXiM e INET.

O *framework* INET é formado por uma variedade de implementações de protocolos: TCP e UDP para a camada de transporte; ICMP e IP para a camada de rede; PPP e *Ethernet* para a camada de enlace; e alguns protocolos de roteamento

(TÜXEN, RÜNGELER e RATHGEB, 2008). Uma extensão do *framework* INET, que atende as peculiaridades das RSSF, foi disponibilizada pelo INETMANET. O INETMANET inclui vários modelos de propagação de ondas de rádio (*Log Normal Shadowing, Nakagami, Rayleigh, Rice e Two Ray*), modelos de bateria simples e suporte à comunicação multi-rádio.

5.2.3 Critério de Parada

O critério de parada do AG adotado neste trabalho foi baseado no número de gerações, em conjunto com a ocorrência de convergência.

O critério de número de gerações é utilizado em vários trabalhos na área, por ser de simples implementação e apresentar bons resultados (SWIECH, 2005; GAULKE, 2007; BRITO, *et al.*, 2011). Ele consiste na predefinição de um número máximo de gerações em que o algoritmo será executado.

A convergência, que ocorre quando não há melhoria significativa no resultado da função de avaliação, foi usada como critério de parada (BAIOCO, 2011), em conjunto com o número de gerações. O uso da convergência garante que, mesmo quando o critério de parada de número de gerações for alcançado, se a avaliação do melhor indivíduo da nova geração for melhor que o melhor indivíduo da geração anterior, o algoritmo não será finalizado.

5.2.4 Elitismo

O módulo de elitismo implementado pode ser configurado para salvar de 0 (zero) a n indivíduos, o que proporciona maior flexibilidade ao modelo proposto. Esta flexibilidade normalmente garante uma melhoria significativa no desempenho do AG, pois, quando se mantém os melhores cromossomos na população, o seu material genético é utilizado na geração dos novos cromossomos.

Neste trabalho, será utilizado o valor de 4 (quatro) para o parâmetro elitismo, obtido empiricamente através de experimentos utilizando o modelo na busca de uma configuração de implantação dos nós sensores que apresentasse taxas de cobertura próxima a 100,0 %.

5.2.5 Descarte

A estratégia de evolução baseada no descarte de indivíduos (cromossomos) de uma geração para outra é utilizada em muitos modelos de busca e tende a funcionar razoavelmente bem. No entanto, ela depende da natureza do problema que está sendo tratado. Assim, para este trabalho, foi implementada uma etapa de descarte, que pode ser ativada ou não (BDACK *et al.*, 1991; MICHALEWICZ e SCHOENAUER, 1996; LUGO e TEIXEIRA JR., 2012). Nos experimentos expostos neste trabalho foi utilizado o descarte.

5.2.6 Seleção

O desenvolvimento de métodos diferentes para realizar a seleção dos cromossomos que irão fornecer o material para a próxima geração é necessário, uma vez que, dependendo do problema que o AG está tentando resolver, um determinado método trará melhores resultados do que os demais.

Foram implementados dois métodos de seleção, dos vários propostos em Linden (2006): o método da roleta viciada (RUYAN *et al.*, 2010; SENGUPTA *et al.*, 2012) e o método de torneios (SCHONFELD e LOUIS, 2007; BARBOSA *et al.*, 2013). Um exemplo de utilização do método de torneios é ilustrado na Figura 3.7.

Foi realizado um experimento com o objetivo de encontrar o método de seleção que proporcionasse o melhor desempenho do algoritmo para as especificidades do cenário adotado nos demais experimentos. O método que utiliza o torneio com 6 (seis) indivíduos foi o escolhido para este trabalho.

Uma possível explicação para o fato de o método de torneios ter apresentado melhores resultados nos testes realizados para este trabalho, pode estar relacionado à forma como o método de torneios realiza a seleção dos indivíduos. No método de torneios, todo indivíduo, independente do seu grau de aptidão, tem as mesmas chances de ser sorteado na primeira etapa do método. Essa característica pode levar à seleção de um indivíduo menos apto para a reprodução, e nesse indivíduo, que tem no geral uma aptidão ruim, conter um gene que seja excelente. Logo, quando esse gene for utilizado para gerar um novo indivíduo, que também tem outros bons genes, o nível de aptidão resultante será maior.

5.2.7 Operadores genéticos

Linden (2006) propôs alguns métodos de cruzamento. No modelo, utilizou-se: o método de *crossover* de 1 (um) ponto (DEB *et al.*, 2002; RUYAN *et al.*, 2010); o *crossover* de 2 (dois) pontos (LI *et al.*, 2010); e o *crossover* uniforme (SENGUPTA *et al.*, 2012). No experimento realizado para determinar qual o método de cruzamento e qual a probabilidade dele ser utilizado, de forma a proporcionar o melhor ganho, os resultados obtidos para a probabilidade foi de 100,0% e o método escolhido foi o cruzamento em 2 (dois) pontos.

Em relação ao operador de mutação, alguns trabalhos baseados na técnica de busca por AG utilizam valores baixos para a probabilidade de se aplicar o operador de mutação. Os valores mais usados são de 0,5% a 1,0%, que proporcionam bons resultados quando o tipo de representação adotada para o vetor cromossomo é binário, que é a representação adotada pelo presente trabalho (LI *et al.*, 2010; RUYAN *et al.*, 2010).

Após a realização do experimento com o objetivo de determinar qual seria a probabilidade que proporcionaria o melhor desempenho para o modelo, obteve-se a probabilidade de 0,5%.

5.3 Discussão Geral

Neste capítulo, foi apresentado o modelo proposto para a busca da melhor configuração de implantação dos nós sensores em uma RSSF, utilizando AG. Também foi feita a descrição de cada etapa do modelo.

No próximo capítulo, serão apresentados os resultados dos experimentos realizados para a busca dos melhores valores para os parâmetros: tamanho da população; tipo de seleção; tipo de cruzamento e sua probabilidade de aplicação; e probabilidade de aplicação da mutação. Outros experimentos também foram realizados em busca das respostas para os questionamentos levantados no trabalho.

6 Apresentação e Análise dos Resultados

Este capítulo tem por objetivo apresentar os resultados obtidos a partir dos três experimentos realizados para demonstrar a eficiência do modelo proposto.

Os experimentos foram divididos em:

- Experimento 1: verificar o desempenho do modelo proposto utilizando 7 (sete) métricas: taxa de cobertura; energia residual; TAFF; TPP; taxa de cobertura e TAFF; energia residual, TAFF e TPP; e taxa de cobertura, energia residual, TAFF e TPP;
- Experimento 2: verificar a influência da presença de obstáculos nos cenários em relação às métricas: redução do consumo de energia (Hipótese A); TAFF (Hipótese B); e TPP (Hipótese C);
- Experimento 3: verificar a influência do número de nós sensores para um cenário de: 30,0 x 30,0 metros; e 38,0 x 11,0 metros.

6.1 Ambiente de Desenvolvimento e Execução dos Experimentos

Os recursos computacionais de *hardware* e *software* utilizados nas etapas de desenvolvimento, execução e análise dos dados resultantes dos experimentos encontram-se detalhados no Quadro 6.1.

Quadro 6.1 Descrição dos recursos computacionais de hardware e software do ambiente

Recurso	Especificação
Processador	Intel Core i7-2600K 3,40 GHz
Núcleos Físicos	4
Memória Principal	8 GB
Memória Secundária	2 TB
Sistema Operacional	Linux Ubuntu Desktop 11.10 (<i>oneiric</i>)
<i>Kernel</i>	Linux 3.0.0-26- <i>generic</i>
Simulador	OMNeT++ <i>Version 4.2.2</i>
<i>Framework</i>	INET 2.0.0
<i>Software</i> de Análise de dados	R <i>Version 2.13.0</i>
Ambiente de desenvolvimento do algoritmo	Eclipse <i>Version 3.7.0 Indico</i>
Compiladores	GCC 4.6.1 e Python 2.7.2+

6.2 Parâmetros de Configuração para o AG do Modelo Proposto

O modelo apresentado no Capítulo 5 e ilustrado na Figura 5.1, foi submetido a vários testes para a determinação dos principais parâmetros de configuração do AG. Nestes testes, foram utilizados os valores do Quadro 6.2 para os parâmetros do AG.

Nos vários experimentos realizados utilizando o presente modelo, verificou-se que os valores maiores que 10 (dez) para o número de gerações, não proporcionava mudanças significativas nos resultados obtidos, mas provocava o aumento no tempo de execução do algoritmo. De forma semelhante, o parâmetro do número de repetições do melhor indivíduo obteve a melhor relação entre desempenho do algoritmo e custo de processamento quando o valor estava configurado em 4 (quatro).

Quadro 6.2 Parâmetros do AG utilizado nos experimentos propostos para este trabalho

Parâmetro	Valor
Número mínimo de gerações	10
Número máximo de repetições do valor de avaliação do melhor indivíduo	4
Tamanho da área (m)	1000x1000
Número de nós em cada indivíduo	36
Descarte das gerações anteriores	Sim
Número de rodadas do algoritmo	30

O valor para o parâmetro área do cenário, utilizado para a realização dos experimentos, foi determinado após testes de desempenho do modelo proposto. Os testes apontaram que, para áreas menores que 1000,0 x 1000,0 metros, os resultados apresentam um baixo desempenho. Esse desempenho insatisfatório pode ser explicado pelo parâmetro do raio de alcance de comunicação do nó sensor utilizado nos experimentos (100,0 metros) que influencia diretamente no número de nós sensores necessários para cobrir toda a área do cenário.

Quando a área do cenário diminui, a quantidade de nós sensores também diminui, o que gera perdas na diversidade do cromossomo. Nessas condições, os AG não apresentam um bom desempenho. Essa deficiência é demonstrada pelo Experimento 3.

6.2.1 Tamanho da População

O primeiro passo foi a criação da população inicial de indivíduos do AG (BECKMANN, 2010). A estratégia adotada para formar a primeira população utilizou a geração aleatória dos indivíduos. O tamanho da população foi determinado, de forma empírica, a partir de testes.

Os testes foram realizados com o objetivo de verificar qual o tamanho da população inicial que melhor se adaptava aos objetivos do trabalho. Para tanto, foram fixados os valores de alguns parâmetros, conforme Quadro 6.2. Linden (2006), em seu trabalho, propôs uma forma de determinar o tamanho da população inicial, utilizando a quantidade de características do cromossomo multiplicada por 40. No entanto, Linden (2006) afirma que essa definição não é necessariamente a melhor estratégia para determinar o tamanho da população, mas que eventualmente tem gerado bons resultados. Portanto, a estratégia sugerida por Linden (2006) serviu como ponto de partida para os valores adotados nos testes.

Na Tabela 6.1, são apresentados os resultados dos testes realizados com o uso de 5 (cinco) tamanhos de população diferentes (20, 40, 60, 80 e 100 indivíduos), os demais parâmetros de configuração do AG foram fixados para observar apenas a influência dos diferentes tamanhos de população. Os dados do experimento foram avaliados estatisticamente, a partir da utilização do teste *t* de *Student* para obter o intervalo de confiança (IC) da média do ganho do algoritmo, com um grau de confiança de 95,0%.

Na Figura 6.1, é apresentado o gráfico *boxplot* do ganho do algoritmo em função do tamanho da população.

Tabela 6.1: Ganho do algoritmo em função do tamanho da população

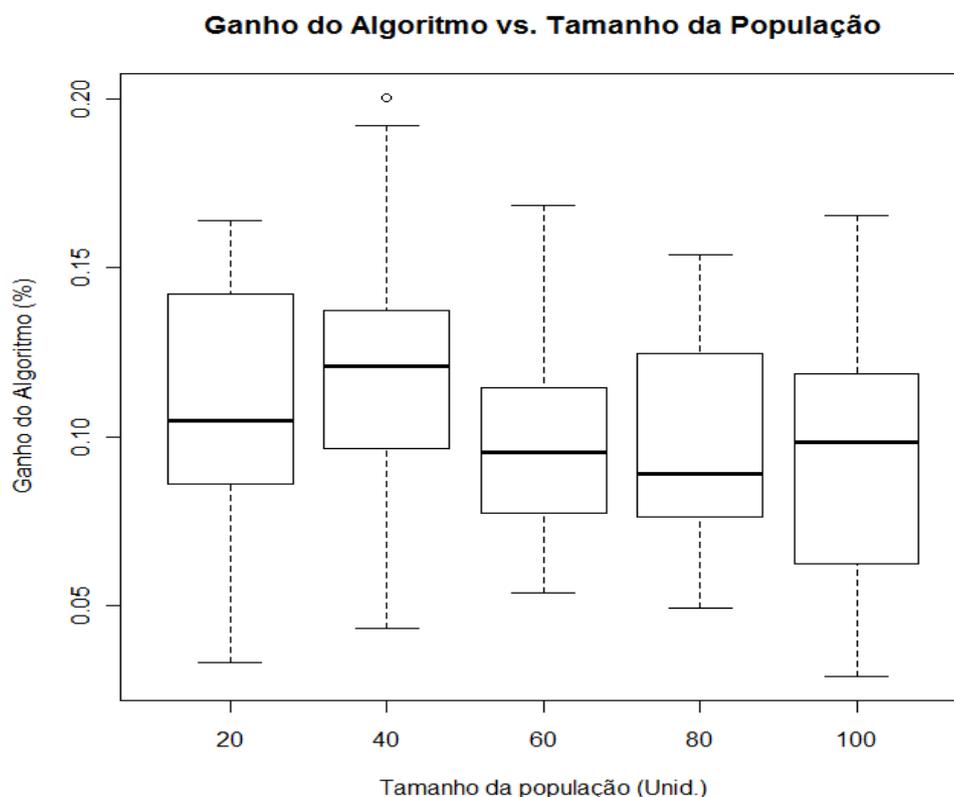
Tamanho da População	Média	IC
20	0,106	0,092 0,120
40	0,118	0,105 0,132
60	0,100	0,088 0,111
80	0,098	0,087 0,109
100	0,096	0,082 0,110

Para a análise dos dados coletados no experimento, foram utilizados testes estatísticos de normalidade dos dados (*Shapiro-Wilk*), de análise de variância

(ANOVA de uma via) e o teste de IC (*t* de *Student*), apresentados no APÊNDICE A. Concluiu-se, pelo resultado do teste ANOVA, com um grau de confiança de 95,0%, não há evidências da existência de diferença significativa entre os grupos testados.

Apesar de os testes mostrarem que não existem diferenças significativas entre os valores utilizados para o parâmetro do tamanho da população, ainda é preciso definir qual o tamanho da população que será utilizado nos demais experimentos. Portanto, o critério utilizado para a escolha do valor para este parâmetro foi baseado no teste que apresentou o maior ganho médio e na análise do gráfico do *boxplot* da Figura 6.1. Ao se observar os resultados apresentados na Figura 6.1, é possível perceber que o grupo que contém 40 indivíduos (cromossomos) apresentou uma pequena superioridade em relação aos demais, e, portanto, foi utilizado nos demais experimentos.

Figura 6.1 Boxplot do ganho do algoritmo em função do tamanho da população



6.2.2 Seleção

Para a realização dos testes, com o objetivo de verificar qual o método de seleção mais adequado às características do problema tratado neste trabalho, foram utilizados os valores do Quadro 6.2 acrescido do tamanho da população que foi definido na Seção 6.1.

Na Tabela 6.2, são apresentados os resultados dos testes realizados, utilizando-se o método da roleta viciada e o método do torneio com tamanho 2 (dois), 4 (quatro), 6 (seis) e 8 (oito). Os dados do experimento foram avaliados estatisticamente, a partir da utilização do teste t de *Student* para obtenção do IC da média do ganho do algoritmo. Para o teste, foi utilizado um grau de confiança de 95,0%.

Para a análise dos dados coletados neste experimento, também foram utilizados os mesmos testes estatísticos aplicados no experimento para determinar o tamanho da população, que também podem ser consultados em mais detalhes no APÊNDICE A.

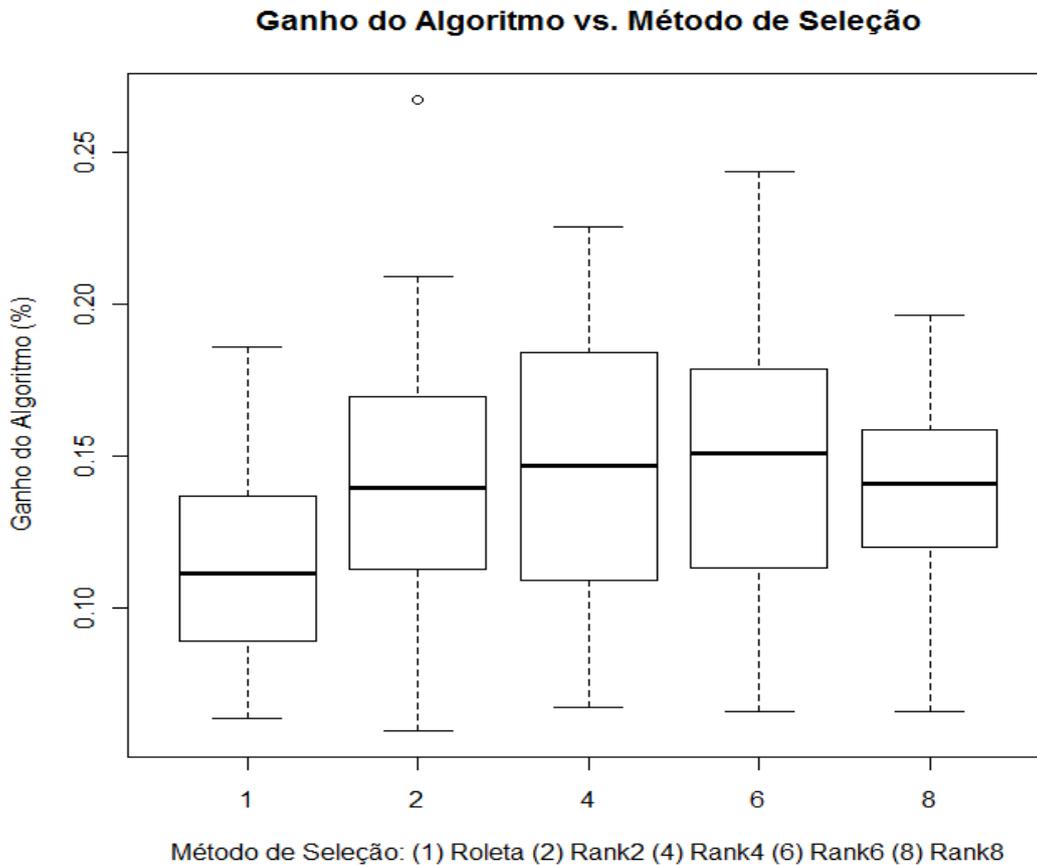
A partir da observação dos resultados apresentados na Tabela 6.2 e do *boxplot* da Figura 6.2, verifica-se que, tanto a variação *rank4*, com tamanho do torneio igual a 4 (quatro), e a variação *rank6*, com tamanho 6, são estatisticamente melhores que o método da roleta viciada. Apoiando-se nos resultados dos testes, o método de seleção escolhido foi o *rank6*, pois além de ser estatisticamente mais eficiente, foi o que apresentou a maior ganho médio.

Tabela 6.2: Ganho do algoritmo em função do método de seleção

Método de Seleção	Média	IC
Roleta	0,116	0,103
<i>Rank 2</i>	0,143	0,127
<i>Rank 4</i>	0,146	0,129
<i>Rank 6</i>	0,150	0,134
<i>Rank 8</i>	0,137	0,124

Na Figura 6.2 é apresentado o gráfico *boxplot* do ganho do algoritmo em função do método de seleção.

Figura 6.2 Boxplot do ganho do algoritmo em função do método de seleção



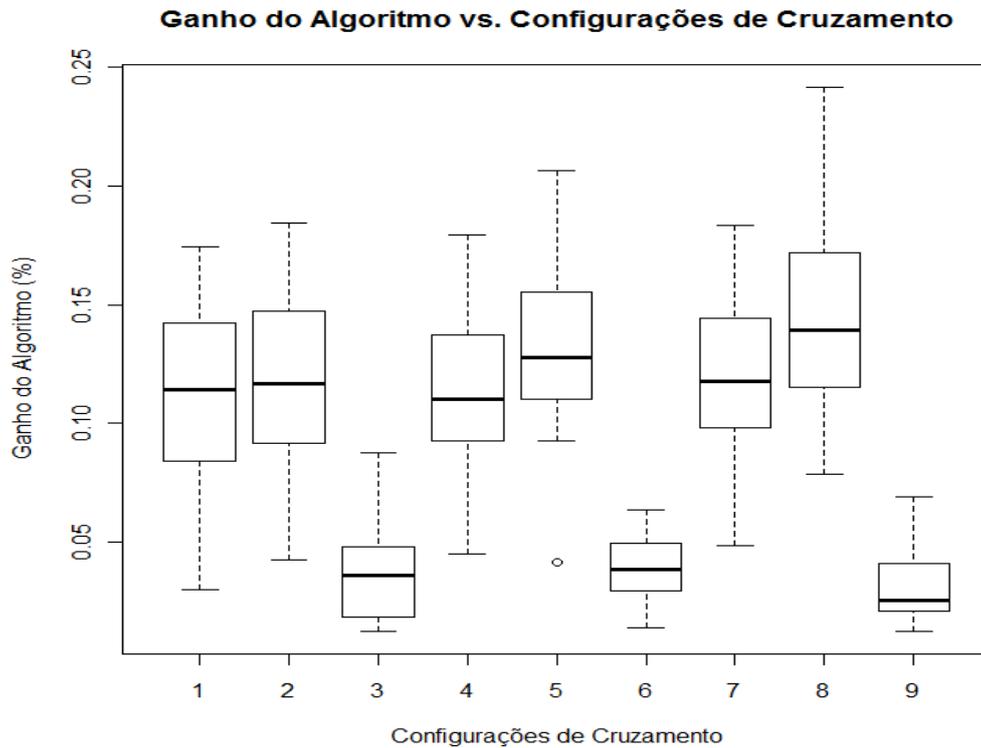
6.2.3 Cruzamento

No experimento realizado para verificar qual o método de cruzamento e qual a probabilidade de aplicação mais adequada às características do problema tratado neste trabalho, mais uma vez foram adotados os valores do Quadro 6.2.

Na Figura 6.3, é apresentado o gráfico *boxplot* do ganho do algoritmo em função do método de cruzamento e a probabilidade de aplicação. Foram utilizados os métodos de cruzamento *crossover* de 1 (um) ponto, *crossover* de 2 (dois) pontos e uniforme, e as probabilidades utilizadas foram 0,600, 0,800 e 1,000 (RUYAN *et al.*, 2010). Detalhes sobre os testes estatísticos realizados estão no APÊNDICE A.

Observando-se o *boxplot* da Figura 6.3, verificou-se que a configuração 8 (oito), que faz uso do método de cruzamento *crossover* de 2 (dois) pontos e utiliza uma probabilidade igual a 1,0 ou de 100,0 %, foi a que apresentou maior ganho médio. Portanto, a configuração 8 (oito) foi a escolhida para ser utilizada neste trabalho.

Figura 6.3 Boxplot ganho do algoritmo em função da configuração de cruzamento



6.2.4 Mutação

Para definir qual a probabilidade de mutação mais adequada ao problema tratado neste trabalho, foram fixados alguns valores dos parâmetros do AG, conforme Quadro 6.2.

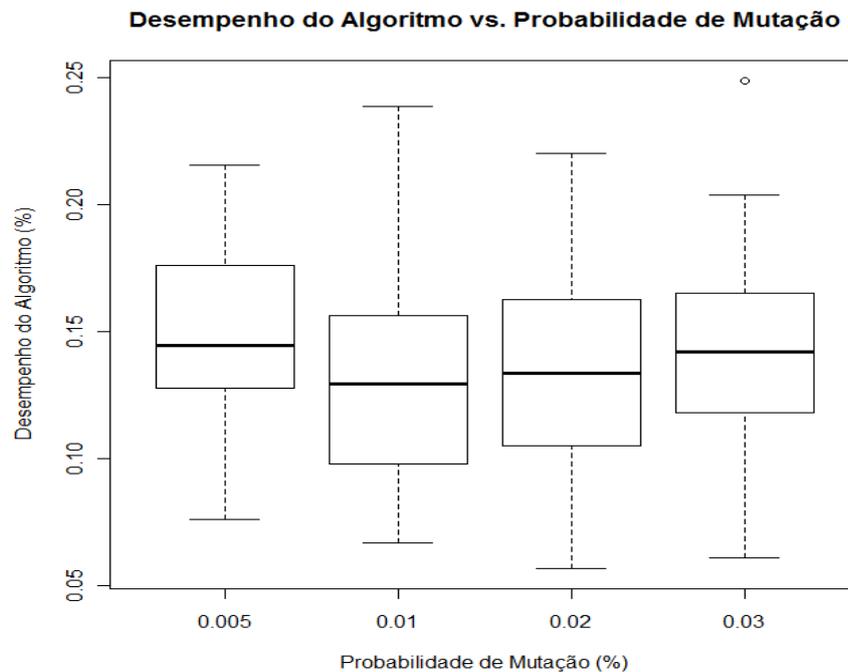
Na Tabela 6.3 são apresentados os resultados dos testes realizados, utilizando-se as probabilidades: 0,005; 0,010; 0,020; e 0,030. Os dados do experimento foram avaliados estatisticamente (mais detalhes dos testes são apresentados no APÊNDICE A).

Tabela 6.3: Ganho do algoritmo em função da probabilidade de mutação

Probabilidade de Mutação	IC	Média
0,005	0,139	0,152
0,010	0,116	0,131
0,020	0,120	0,135
0,030	0,124	0,140

Na Figura 6.4, é apresentado o gráfico *boxplot* do ganho do algoritmo em função do método de cruzamento e a probabilidade de aplicação.

Figura 6.4 Boxplot ganho do algoritmo em função da probabilidade de mutação



Concluiu-se, pelo resultado do teste *ANOVA*, com um grau de confiança de 95,0%, que não há evidência da existência de diferença significativa entre os 4 (quatro) valores de probabilidades testados. No entanto, ainda fez-se necessário determinar o valor para a probabilidade de mutação utilizada nos demais experimentos.

Baseado nos resultados apresentados na Tabela 6.7 e no *boxplot* da Figura 6.4, o valor para a probabilidade de mutação escolhido foi de 0,5%, levando em consideração que a maior média do ganho obtido foi alcançada com este valor. Esta pequena superioridade também pode ser observada no *boxplot* da Figura 6.4.

6.2.5 Parâmetros de Configuração

No Quadro 6.3, são apresentados os valores dos parâmetros de configuração do AG para os experimentos realizados neste trabalho.

Quadro 6.3 Parâmetros complementares de configuração do AG

Parâmetro	Experimento 1							Experimento 2			Experimento 3	
	A	B	C	D	E	F	G	A	B	C	A	B
α	1,000	0,000	0,000	0,000	0,500	0,000	0,250	0,000	0,000	0,000	1,000	1,000
β	0,000	1,000	0,000	0,000	0,000	0,330	0,250	1,000	0,000	0,000	0,000	0,000
γ	0,000	0,000	1,000	0,000	0,500	0,330	0,250	0,000	1,000	0,000	0,000	0,000
δ	0,000	0,000	0,000	1,000	0,000	0,330	0,250	0,000	0,000	1,000	0,000	0,000
Número mínimo de gerações	10	10	10	10	10	10	10	10	10	10	300	300
Número máximo de repetições	4	4	4	4	4	4	4	4	4	4	2	2
Número de nós	36	36	36	36	36	36	36	36	36	36	5	2
Largura da área (m)	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	30,000	38,000
Comprimento da área (m)	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	1000,000	30,000	11,000
Raio de cobertura (m)	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	11,000	11,000
Tempo de simulação (s)	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	NA	NA
Somatório da energia da RSSF (mW-s)	$7,200 \times 10^7$	NA	NA									
Módulo de obstáculos ativos	Sim	Ambos	Ambos	Ambos	Não	Não						
Número de repetições	10	10	10	10	10	10	10	16	16	16	40	30

Os valores dos parâmetros do AG que proporcionaram o melhor desempenho do modelo proposto são apresentados no Quadro 6.4.

Quadro 6.4 Parâmetros básicos de configuração do AG

Parâmetro	Valor
Probabilidade de mutação (%)	0,5
Probabilidade de cruzamento (%)	100,0
Tamanho da população	40
Número de indivíduos salvos (Elitismo)	4
Método de cruzamento	<i>Crossover 2,0 Pontos</i>
Método de seleção	Torneio com 6 (seis) indivíduos
Descarte das gerações anteriores	Sim
Script de configuração do simulador	APÊNDICE C

6.3 Experimento 1

O objetivo do experimento é demonstrar o ganho proporcionado pelo uso do modelo. Para tanto, foram utilizados 7 (sete) planos de busca diferentes. Cada uma delas tem como escopo obter ganho de qualidade em uma métrica ou em um conjunto de métricas.

Os planos têm como escopo a busca por uma configuração de implantação dos nós sensores que proporcione:

- plano A: a maior taxa de cobertura da área do cenário;
- plano B: o maior somatório do nível da energia residual da rede;
- plano C: o menor TAFF;
- plano D: a menor TPP;
- plano E: a maior taxa de cobertura da área do cenário e o menor TAFF;
- plano F: o maior somatório do nível da energia residual da rede, o menor TAFF e a menor TPP;
- plano G: a maior taxa de cobertura da área do cenário, o maior somatório do nível da energia residual da rede, o menor TAFF e a menor TPP.

Para todos os planos, foram utilizados os mesmos parâmetros de configuração do AG e também do simulador (a taxa de cobertura; o somatório da energia residual; o TAFF; e a TPP), com exceção dos parâmetros alfa (α), beta (β), gama (γ) e delta (δ), que são responsáveis por determinar os escopos da busca. Alguns desses parâmetros foram retirados do Quadro 6.4 e do Quadro 6.3.

O teste realizado para o plano A tem o objetivo de verificar o ganho na taxa de cobertura da área do cenário proporcionado pelo modelo em comparação a uma configuração aleatória e outra uniforme. Para determinar o valor médio do ganho e o IC foram realizados 15 repetições do teste para o plano A.

As Equações (6.1) e (6.2) caracterizam o ganho na métrica taxa de cobertura em relação à configuração aleatório e uniforme respectivamente.

$$G_{f(TC)_A} = \frac{f(A)_M - f(A)_A}{f(A)_A}. \quad (6.1)$$

$$G_{f(TC)_U} = \frac{f(A)_M - f(A)_U}{f(A)_U} . \quad (6.2)$$

em que:

- $G_{f(TC)_A} \in \mathbb{R}^+$, é o ganho na função para a métrica taxa de cobertura da área do cenário resultante da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(TC)_U} \in \mathbb{R}^+$, é o ganho na função para a métrica taxa de cobertura da área do cenário resultante da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(A)_A \in \mathbb{R}^+$, é a taxa de cobertura da área do cenário resultante da utilização de uma configuração de implantação dos nós sensores obtida de forma aleatória;
- $f(A)_U \in \mathbb{R}^+$, é a taxa de cobertura da área do cenário resultante da utilização de uma configuração uniforme para implantação dos nós sensores;
- $f(A)_M \in \mathbb{R}^+$, é a taxa de cobertura da área do cenário resultante da utilização do modelo proposto.

O resumo do resultado da conformação que trata da métrica da taxa de cobertura da área do cenário encontra-se exposto na Tabela 6.4.

Na segunda parte do experimento, foi realizado o teste para estudar o plano B, com a finalidade de verificar o ganho no somatório da energia residual dos nós sensores da RSSF decorrente da aplicação do modelo proposto. Para obter o valor médio do ganho e seu IC, foram realizadas 10 (dez) repetições do teste para o plano B.

As Equações (6.3) e (6.4) caracteriza o ganho na métrica energia residual em relação à configuração aleatório e uniforme respectivamente.

$$G_{f(ER)_A} = \frac{f(ER)_M - f(ER)_A}{f(ER)_A} . \quad (6.3)$$

$$G_{f(ER)_U} = \frac{f(ER)_M - f(ER)_U}{f(ER)_U} . \quad (6.4)$$

em que:

- $G_{f(ER)_A} \in \mathbb{R}^+$, é o ganho na função para a métrica somatório da energia residual dos nós sensores resultantes da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(ER)_U} \in \mathbb{R}^+$, é o ganho na função para a métrica somatório da energia residual dos nós sensores resultantes da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(ER)_A \in \mathbb{R}^+$, é o somatório de energia residual dos nós sensores resultante da utilização de uma configuração de implantação dos nós sensores obtida de forma aleatória.
- $f(ER)_U \in \mathbb{R}^+$, é o somatório de energia residual dos nós sensores resultante da utilização de uma configuração uniforme para implantação dos nós sensores;
- $f(ER)_M \in \mathbb{R}^+$, é o somatório de energia residual dos nós sensores resultante da utilização do modelo proposto.

Nos testes realizados para os planos C e E deste experimento, o objetivo foi verificar o ganho nas métricas TAFF e TPP após a busca da melhor configuração de implantação realizada pelo presente modelo. Foram realizadas 10 (dez) repetições do teste para o plano C para aumentar a representatividade estatística das médias do TAFF e da TPP.

O ganho na métrica do TAFF, proporcionado pelo modelo proposto neste experimento em relação à configuração aleatório e uniforme, são caracterizados pelas Equações (6.5) e (6.11).

$$G_{f(TAFF)_A} = \frac{f(TAFF)_M - f(TAFF)_A}{f(TAFF)_A}. \quad (6.5)$$

$$G_{f(TAFF)_U} = \frac{f(TAFF)_M - f(TAFF)_U}{f(TAFF)_U}. \quad (6.6)$$

em que:

- $G_{f(TAFF)_A} \in \mathbb{R}^+$, é ganho na função para a métrica TAFF resultante da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(TAFF)_U} \in \mathbb{R}^+$, é ganho na função para a métrica TAFF resultante da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(TAFF)_A \in \mathbb{R}^+$, é o TAFF resultante da utilização de uma configuração de implantação dos nós sensores obtidos de forma aleatória;

- $f(TAFF)_U \in \mathbb{R}^+$, é o TAFF resultante da utilização de uma configuração de implantação uniforme dos nós sensores;
- $f(TAFF)_M \in \mathbb{R}^+$, é o TAFF resultante da utilização do modelo proposto.

E o ganho na métrica do TPP em relação às configurações aleatória e uniforme são caracterizados pelas Equações (6.7) e (6.15) respectivamente.

$$G_{f(TPP)_A} = \frac{f(TPP)_M - f(TPP)_A}{f(TPP)_A}. \quad (6.7)$$

$$G_{f(TPP)_U} = \frac{f(TPP)_M - f(TPP)_U}{f(TPP)_U}. \quad (6.8)$$

em que:

- $G_{f(TPP)_A} \in \mathbb{R}^+$, é o ganho na função para a métrica TPP resultante da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(TPP)_U} \in \mathbb{R}^+$, é o ganho na função para a métrica TPP resultante da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(TPP)_A \in \mathbb{R}^+$, é o TPP resultante da utilização de uma configuração de implantação dos nós sensores obtida de forma aleatória;
- $f(TPP)_U \in \mathbb{R}^+$, é o TPP resultante da utilização de uma configuração uniforme de implantação dos nós sensores;
- $f(TPP)_M \in \mathbb{R}^+$, é o TPP resultante da utilização de uma configuração de implantação dos nós sensores obtidos do modelo proposto.

O objetivo, neste caso, é mostrar o ganho proporcionado pelo modelo proposto quando a busca utiliza mais de uma métrica. Foram propostos outros 3 (três) planos, para mostrar o ganho quando forem utilizadas 2 (duas) métricas, 3 (três) métricas ou todas as 4 (quatro) métricas.

O ganho da função, utilizando as métricas taxa de cobertura da área do cenário e TAFF em relação às configurações aleatória e uniforme, proporcionado pelo modelo proposto neste experimento, são caracterizados pelas Equações (6.9) e (6.10) respectivamente.

$$G_{f(A,TAFF)_A} = \frac{f(A,TAFF)_M - f(A,TAFF)_A}{f(A,TAFF)_A}. \quad (6.9)$$

$$G_{f(A,TAFF)_U} = \frac{f(A,TAFF)_M - f(A,TAFF)_U}{f(A,TAFF)_U}. \quad (6.10)$$

em que:

- $G_{f(A,TAFF)_A} \in \mathbb{R}^+$, é o ganho na função para as métricas taxa de cobertura da área do cenário e o TAFF resultante da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(A,TAFF)_U} \in \mathbb{R}^+$, é o ganho na função para as métricas taxa de cobertura da área do cenário e o TAFF resultante da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(A,TAFF)_A \in \mathbb{R}^+$, é o somatório da taxa de cobertura da área do cenário com o TAFF resultante de uma configuração de implantação dos nós sensores obtidos de forma aleatória;
- $f(A,TAFF)_U \in \mathbb{R}^+$, é o somatório da taxa de cobertura da área do cenário com o TAFF resultante de uma configuração de implantação uniforme dos nós sensores;
- $f(A,TAFF)_M \in \mathbb{R}^+$, somatório da taxa de cobertura da área do cenário com o TAFF resultante de uma configuração de implantação dos nós sensores obtido por meio do modelo proposto.

O ganho da função utilizando as métricas do somatório da energia residual dos nós sensores, do TAFF e da TPP em relação às configurações aleatória e uniforme, são definidos a partir das Equações (6.11) e (6.12) respectivamente.

$$G_{f(ER,TAFF,TPP)_A} = \frac{f(ER,TAFF,TPP)_M - f(ER,TAFF,TPP)_A}{f(ER,TAFF,TPP)_A}. \quad (6.11)$$

$$G_{f(ER,TAFF,TPP)_U} = \frac{f(ER,TAFF,TPP)_M - f(ER,TAFF,TPP)_U}{f(ER,TAFF,TPP)_U}. \quad (6.12)$$

em que:

- $G_{f(ER,TAFF,TPP)_A} \in \mathbb{R}^+$, é o ganho na função para as métricas somatório da energia residual, do TAFF e do TPP, resultante da aplicação do modelo proposto em relação a uma configuração aleatória;

- $G_{f(ER,TAFF,TPP)_U} \in \mathbb{R}^+$, é o ganho na função para as métricas somatório da energia residual, do TAFF e do TPP, resultante da aplicação do modelo proposto em relação a uma configuração uniforme;
- $f(ER,TAFF,TPP)_A \in \mathbb{R}^+$, é o somatório da energia residual, do TAFF e do TPP, resultante de uma configuração de implantação dos nós sensores obtida de forma aleatória;
- $f(ER,TAFF,TPP)_A \in \mathbb{R}^+$, é o somatório da energia residual, do TAFF e do TPP, resultante de uma configuração de implantação uniforme dos nós sensores;
- $f(ER,TAFF,TPP)_M \in \mathbb{R}^+$, é o somatório da energia residual, do TAFF, e do TPP resultante de uma configuração de implantação dos nós sensores obtidos a partir do modelo proposto.

Na última conformação do experimento, o objetivo foi verificar o ganho da função utilizando as métricas da taxa de cobertura da área do cenário, do somatório da energia residual dos nós sensores, do TAFF e da TPP, proporcionado pelo modelo proposto em relação às configurações aleatória e uniforme. Nas Equações (6.13) e (6.14) encontram-se definidos os ganhos das funções em relação a configuração aleatória e uniforme respectivamente.

$$G_{f(A,ER,TAFF,TPP)_A} = \frac{f(A,ER,TAFF,TPP)_M - f(A,ER,TAFF,TPP)_A}{f(A,ER,TAFF,TPP)_A}. \quad (6.13)$$

$$G_{f(A,ER,TAFF,TPP)_U} = \frac{f(A,ER,TAFF,TPP)_M - f(A,ER,TAFF,TPP)_U}{f(A,ER,TAFF,TPP)_U}. \quad (6.14)$$

em que:

- $G_{f(A,ER,TAFF,TPP)_A} \in \mathbb{R}^+$, é o ganho na função para as métricas taxa de cobertura, somatório da energia residual, do TAFF e do TPP, resultante da aplicação do modelo proposto em relação a uma configuração aleatória;
- $G_{f(A,ER,TAFF,TPP)_U} \in \mathbb{R}^+$, é o ganho na função para as métricas taxa de cobertura, somatório da energia residual, do TAFF e do TPP, resultante da aplicação do modelo proposto em relação a uma configuração uniforme;

- $f(A, ER, TAFF, TPP)_A \in \mathbb{R}^+$, é o somatório da taxa de cobertura, da energia residual, do TAFF e do TPP resultante de uma configuração de implantação dos nós sensores obtidos de forma aleatória;
- $f(A, ER, TAFF, TPP)_U \in \mathbb{R}^+$, é o somatório da taxa de cobertura, da energia residual, do TAFF e do TPP resultante de uma configuração de implantação uniforme dos nós sensores;
- $f(A, ER, TAFF, TPP)_M \in \mathbb{R}^+$, é o somatório da taxa de cobertura, da energia residual, do TAFF e do TPP, resultante de uma configuração de implantação dos nós sensores obtido a partir do modelo proposto.

Utilizou-se o teste *t* de *student* para estimar o ganho médio e o IC de todos os planos do experimento.

Nas Tabela 6.4 e Tabela 6.5, são apresentados os resultados de todos os planos propostos para o experimento em relação às configurações aleatória e uniforme respectivamente.

Tabela 6.4 Médias do ganho do modelo e IC para cada metrica em comparação a uma configuração aleatória

Métrica	Média	CV(%)	IC	
$G_{f(A)}_A$	0,282	3,044	0,277	0,287
$G_{f(ER)}_A$	1,882	15,760	1,607	2,156
$G_{f(TAFF)}_A$	0,728	6,581	0,684	0,772
$G_{f(TPP)}_A$	0,832	2,947	0,806	0,858
$G_{f(A,TAFF)}_A$	0,106	8,867	0,099	0,113
$G_{f(ER,TAFF,TPP)}_A$	0,189	4,042	0,184	0,194
$G_{f(A,ER,TAFF,TPP)}_A$	0,160	5,217	0,154	0,166

Tabela 6.5 Médias do ganho do modelo e IC para cada metrica em comparação a uma configuração uniforme

Métrica	Média	CV(%)	IC	
$G_{f(A)}_U$	0,198	4,048	0,194	0,203
$G_{f(ER)}_U$	3,341	13,372	2,927	3,754
$G_{f(TAFF)}_U$	0,611	11,193	0,548	0,675
$G_{f(TPP)}_U$	0,840	2,781	0,815	0,865
$G_{f(A,TAFF)}_U$	0,063	14,337	0,057	0,069
$G_{f(ER,TAFF,TPP)}_U$	0,204	3,798	0,198	0,209
$G_{f(A,ER,TAFF,TPP)}_U$	0,141	5,801	0,136	0,147

Observa-se que os ganhos médios proporcionados pelo modelo, quando adotada apenas uma das métricas por busca, são superiores aos ganhos quando se adota um conjunto de métricas. Uma possível explicação para esse comportamento

pode estar relacionada ao fato de que, quando a busca da melhor configuração é realizada com apenas uma métrica, o ganho desta métrica é maximizado. No entanto, como consequência, as demais 3 (três) métricas serão penalizadas. Este resultado mostra a importância da utilização das múltiplas métricas do modelo. Um exemplo de sua relevância é que não adianta apenas aumentar a energia residual, degradando a QoS da rede, uma vez que isto pode causar falhas inaceitáveis para aplicação.

Também é possível observar que, das 4 (quatro) métricas adotadas no trabalho, a métrica da taxa de cobertura é a que apresenta o menor desempenho. Este fato pode ser observado no plano A, cujo objetivo é o de buscar apenas a melhor taxa de cobertura, e no plano E, que tem o objetivo de encontrar o maior ganho para as métricas taxa de cobertura em conjunto com a métrica TAFF.

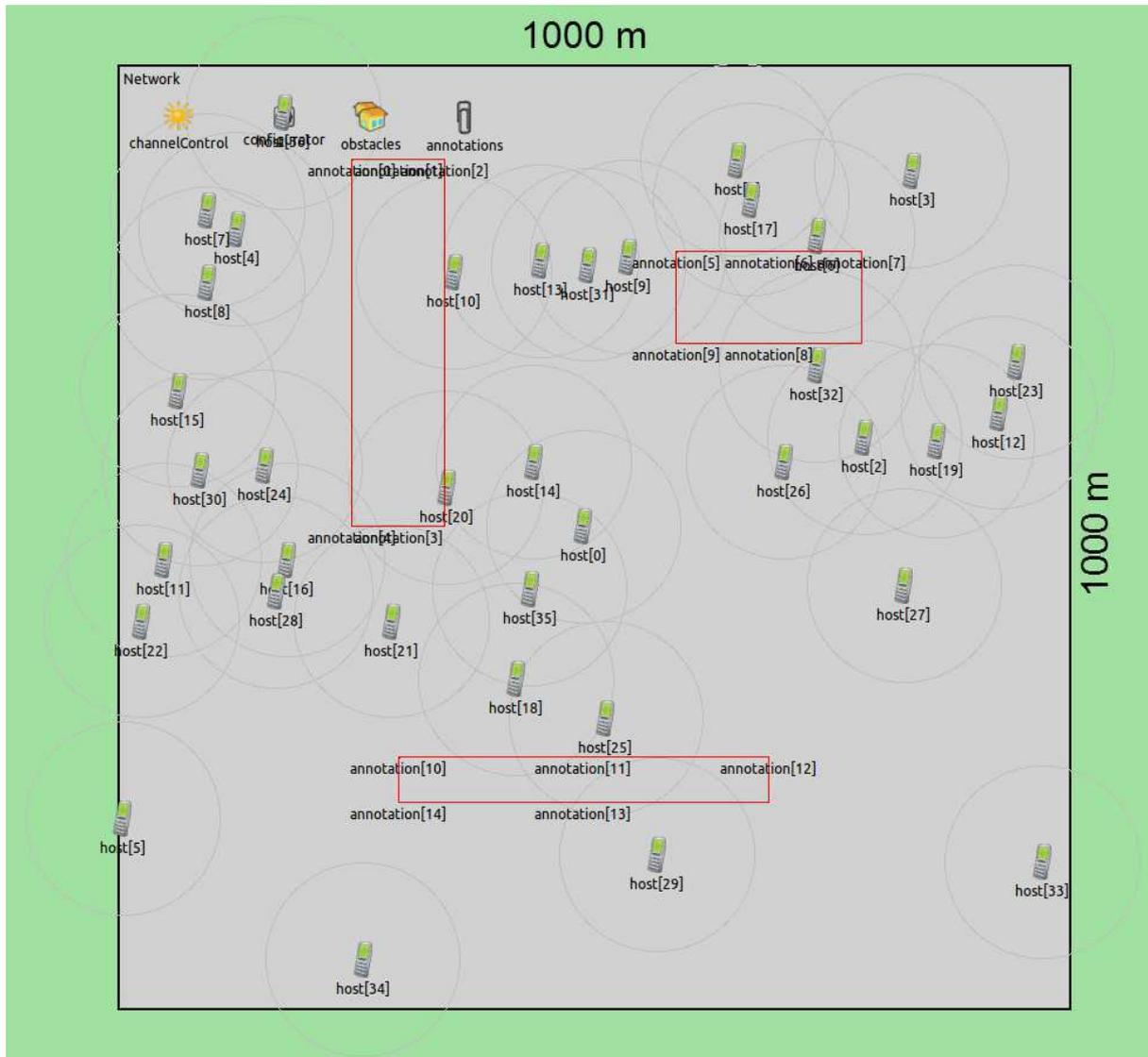
No modelo proposto, o cálculo da função de aptidão para a métrica taxa de cobertura se difere das demais, pois para o seu cálculo não é necessário realizar uma simulação. Outra característica do cálculo da função de aptidão para a métrica taxa de cobertura é a de que os valores de avaliação calculados para diferentes configurações são próximos, causando uma perda na diversidade e, conseqüentemente, uma diminuição do desempenho do AG.

6.4 Experimento 2

A maioria dos simuladores de RSSF não desenvolve módulos que possibilitem acrescentar obstáculos nos cenários de simulação. Em grande parte, porque esses módulos provocam um aumento considerável de processamento. Portanto, neste trabalho, foi necessário verificar se a presença de obstáculos nos cenários influencia na TPP, no TAFF e no consumo de energia da rede RSSF.

Para a realização dos testes propostos nesse experimento, foi criado um cenário com a presença de obstáculos para verificar a influência dos mesmos nos resultados das métricas de energia residual, TAFF e TPP. Na Figura 6.5, o cenário utilizado no experimento é caracterizado detalhadamente. O polígono em vermelho, no centro do cenário, é a representação do obstáculo. Nessa região, nenhum nó sensor pode ser instalado e também não é possível nenhuma comunicação entre os nós sensores nesta área.

Figura 6.5 Exemplo detalhado de um cenário com obstáculo



A. Energia Residual

A Hipótese A levantada neste trabalho tem o objetivo de verificar a influência dos obstáculos no consumo de energia da RSSF. Para testar esta hipótese, foi realizado um experimento. Neste experimento, o algoritmo foi configurado para verificar apenas a energia residual, dispensando os dados de TAFF, TPP e taxa de cobertura. Além disto, também foram utilizados os parâmetros do Quadro 6.4 e do Quadro 6.3. O simulador foi configurado para não considerar a influência dos obstáculos no cenário. O tempo de simulação foi de 200 segundos e a energia total de todos os nós sensores foi de, aproximadamente, $(7,232 \times 10^7 \text{ mW.s})$. Para esta hipótese foram realizadas 16 repetições do experimento.

Para cada repetição do algoritmo proposto, foi gerada, como resultado da busca, a configuração de implantação que apresenta o maior nível de energia residual da rede. Então, adicionou-se esta configuração ao simulador, inicialmente configurado para não considerar a presença dos obstáculos, e o resultado da energia residual total da rede foi salvo em uma tabela. Logo após, o simulador executou novamente a simulação, considerando agora a influência dos obstáculos no cenário. Ao final do experimento, obteve-se uma tabela com todos os valores da energia residual, considerando ou não a influência dos obstáculos.

Hipótese A (H_0): A presença de obstáculos nas simulações afeta o consumo de energia da rede, reduzindo o somatório da energia residual dos nós sensores de uma RSSF.

- H_0 : A presença de obstáculos no cenário de simulação não influencia no somatório da energia residual dos nós sensores. Formalmente.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- $\mu_0 \in \mathbb{R}^+$ é a média do ganho do algoritmo sem considerar os obstáculos;
- $\mu_1 \in \mathbb{R}^+$ é a média do ganho do algoritmo considerando os obstáculos.

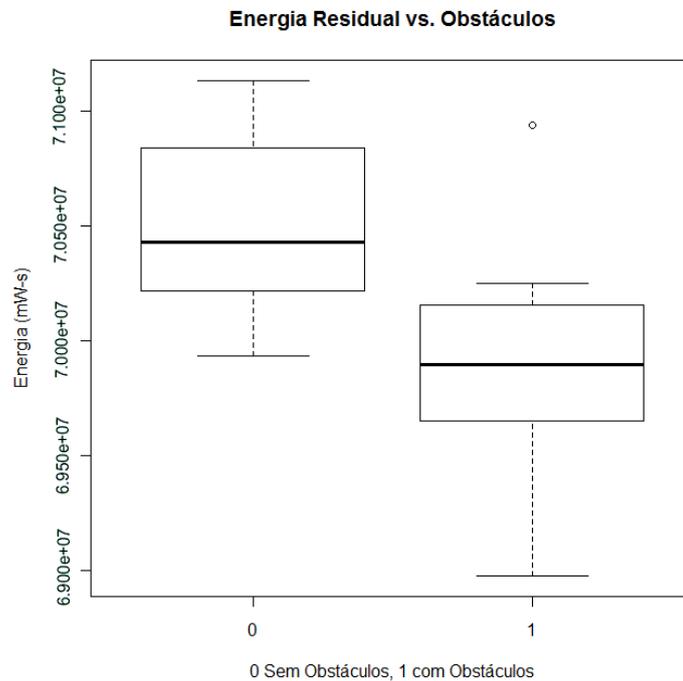
Na Figura 6.6, é ilustrado o *boxplot* e, na Tabela 6.6, a média do somatório da energia residual dos resultados dos experimentos com e sem a influência dos obstáculos. A partir dos resultados obtidos, é possível observar que a mediana do somatório da energia residual sem obstáculos apresenta um nível de energia residual ($7,043 \times 10^7$ mW-s) maior que o do nível com obstáculos ($6,990 \times 10^7$ mW-s). Vale ressaltar que, quanto maior for o nível de energia residual, maior será o tempo de vida da rede.

Na Tabela 6.6, são apresentados a média e o IC da energia consumida no experimento.

Tabela 6.6 Média do somatório da energia residual em relação a presença de obstáculos no cenário

Obstáculo	Média (mW-s)	IC (mW-s)	
0	$7,053 \times 10^7$	$7,032 \times 10^7$	$7,073 \times 10^7$
1	$6,988 \times 10^7$	$6,964 \times 10^7$	$7,013 \times 10^7$

Figura 6.6 Boxplot do somatório da energia residual em relação a presença de obstáculos no cenário



Após análise dos resultados do experimento, pode-se concluir que existe diferença significativa entre os somatórios da energia residual dos nós sensores em relação à presença ou não de obstáculos no cenário de simulação, com um nível de confiança de 95,0%. O estudo estatístico encontra-se detalhado no APÊNDICE A.

B. Tempo de Atraso Fim-a-Fim

A Hipótese B tem como objetivo verificar a influência dos obstáculos sobre os resultados do TAFF. Para verificar a Hipótese B, o algoritmo foi configurado com os valores semelhantes aos utilizados na verificação da energia residual (Hipótese A), com exceção dos valores dos parâmetros β e γ .

Hipótese B (H_0): A presença de obstáculos nas simulações afeta a média do TAFF dos pacotes entre os nós sensores e o nós sorvedouros.

- H_0 : A média dos TAFF dos pacotes na rede é igual para as simulações com e sem obstáculos no cenário. Formalmente.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- μ_0 é a média do ganho do algoritmo sem considerar os obstáculos;
- μ_1 é a média do ganho do algoritmo considerando os obstáculos.

Dos testes estatísticos realizados com os dados do experimento, pode-se observar, no *boxplot* ilustrado na Figura 6.7 e na Tabela 6.7, que a mediana do TAFF do grupo sem obstáculos obteve o tempo de (0,018) segundos, menor que o grupo com obstáculos, que foi de (0,034) segundos. Para essa métrica, o objetivo do algoritmo é encontrar a configuração que apresente o menor TAFF.

Figura 6.7 Boxplot da média dos TAFF em relação aos obstáculos no cenário

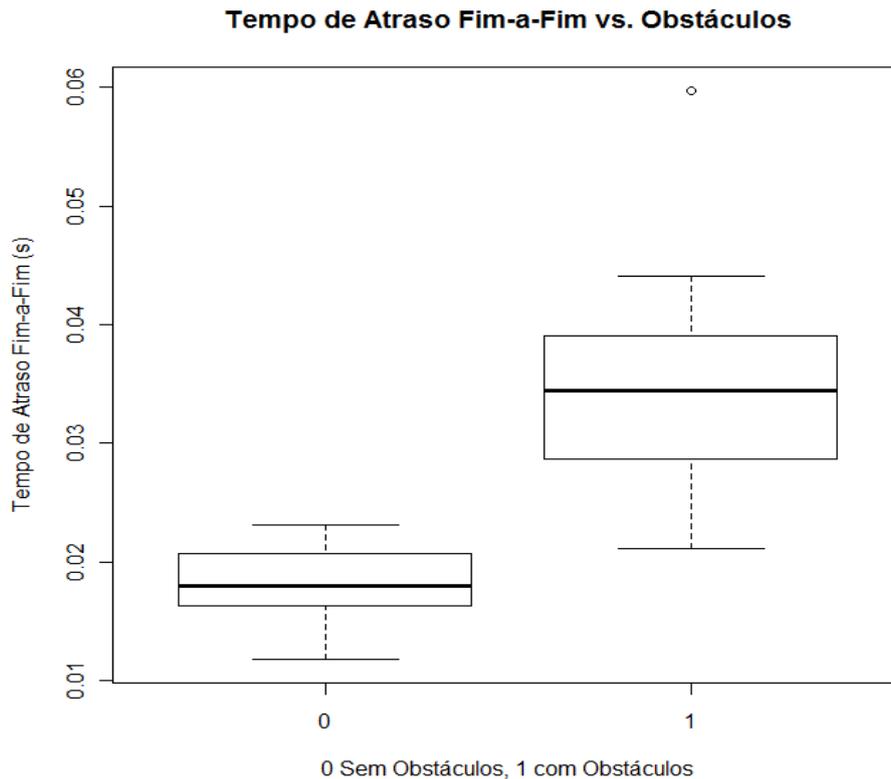


Tabela 6.7 Média do TAF em relação aos obstáculos no cenário

Obstáculo	Média (s)	IC (log(s))	
0	0,018	0,016	0,020
1	0,035	0,030	0,040

Assim, após a análise dos resultados dos testes estatísticos, que podem ser consultados no APÊNDICE A, é possível rejeitar, com o nível de significância de 5,0%, a hipótese B de que os dados do TAFF apresentam a mesma média quando aplicados a cenários com ou sem a presença de obstáculos.

C. Taxa de Perda de Pacotes

Na verificação da influência dos obstáculos sobre a TPP, foi formulada a Hipótese C e, novamente, utilizou-se um experimento semelhante ao da Hipótese A, modificando-se os parâmetros β e δ para configurar o algoritmo de busca da melhor TPP.

Hipótese C (H_0): A presença de obstáculos nas simulações afeta a TPP da RSSF.

- H_0 : A presença de obstáculos no cenário de simulação não influencia na média da TPP da rede.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- μ_0 é a média do ganho do algoritmo sem considerar os obstáculos;
- μ_1 é a média do ganho do algoritmo considerando os obstáculos.

Na Tabela 6.8, encontram-se os resultados do teste *Wilcoxon* para verificar se existem diferenças significativas nas médias da TPP em relação à presença ou não de obstáculos no cenário de simulação.

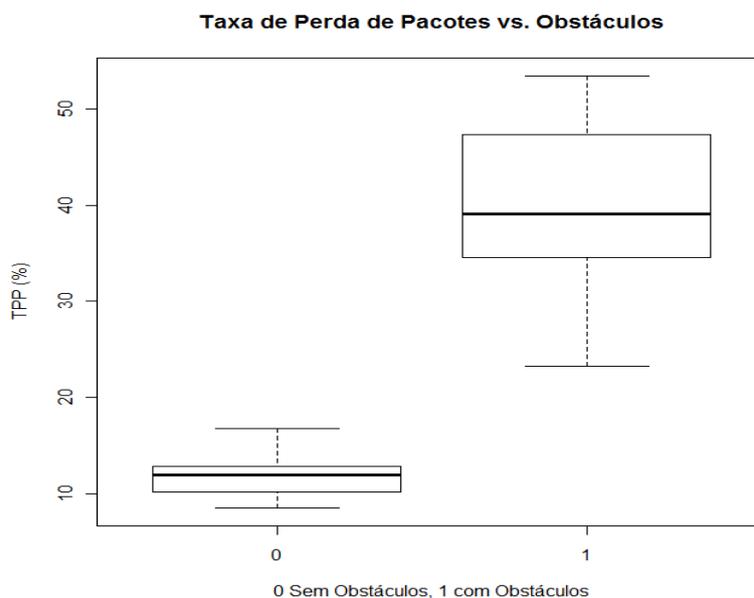
Tabela 6.8 Resultado do teste para verificar se existem diferenças entre as médias da TPP

<i>Wilcoxon</i>	Valores
V	0,000
p -valor	3,815e-06
α	0,050

Na Figura 6.8, é ilustrado o *boxplot* resultante dos dados do experimento, em que é possível verificar que a mediana da TPP do grupo sem obstáculos (12,0%)

apresenta melhor desempenho em relação ao grupo com obstáculos (39,1%). É importante lembrar que, quanto menor for a TPP, melhor a qualidade da rede.

Figura 6.8 Boxplot da média da TPP em relação aos obstáculos no cenário



Assim, a partir da análise estatística dos resultados obtidos no experimento, pode-se rejeitar, com um nível de confiança de 95,0%, a hipótese H_0 de que as médias da TPP são iguais para os testes realizados com e sem presença de obstáculos no cenário.

6.5 Experimento 3

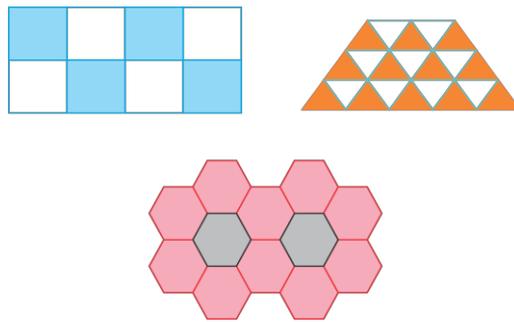
O presente experimento teve por objetivo verificar a influência do número de nós sensores utilizados na RSSF sobre o desempenho do AG utilizado no modelo proposto neste trabalho. Para tanto, o experimento foi modelado para realizar uma busca da configuração de implantação que apresente a maior taxa de cobertura da área a ser monitorada e a menor quantidade de nós sensores utilizados, fazendo uso de 3 (três) abordagens diferentes.

Inicialmente, utilizou-se um cenário simples de área igual a 100,0 x 100,0 metros, com um relevo plano, em que, para todos os pontos, o valor da coordenada Z é 0 (zero). Os nós sensores tinham raio de cobertura de sensoriamento de 10,0 metros. O segundo cenário tinha 38,0 x 10,0 metros e foi utilizada a mesma

configuração de nós sensores. Tinha-se como objetivo, para este experimento, encontrar a configuração que apresentasse a melhor cobertura de sensoriamento. Utilizou-se, para isso, o raio de sensoriamento de 10,0 metros.

A primeira abordagem utilizada para determinar a posição de instalação e a quantidade de nós sensores necessária para cobrir a área dos cenários de 100,0 x 100,0 metros e 28,3 x 14,1 metros, com o menor custo, foi a técnica gráfica baseada na teoria dos mosaicos planos. Segundo o teorema de Kepler dos Mosaicos no Plano, existem 11 maneiras de se cobrir o plano utilizando-se exclusivamente polígonos regulares. Sendo que, dessas 11, apenas 3 (três) fornecem mosaicos regulares e as demais são chamadas de mosaicos semirregulares (ALVES e DALCIN, 1999). Na Figura 6.9, são apresentados os mosaicos regulares.

Figura 6.9 Mosaicos Regulares, adaptado de (ALVES e DALCIN, 1999)



Escolheu-se o polígono regular hexágono para representar a área de cobertura do nó sensor, pois é a que mais se aproxima de um círculo. A Equação (6.17) foi utilizada para determinar o número aproximado de nós sensores necessário para cobrir toda a área de simulação. Nesse cálculo, as Equação (6.15) e Equação (6.16) são utilizadas para determinar a área de um quadrado e de um hexágono, respectivamente.

$$A_{Cenário} = l^2, \quad (6.15)$$

$$A_{Sensor} = 2(r)^2, \quad (6.16)$$

$$N_{nós} = A_{Cenário} / A_{Sensor} . \quad (6.17)$$

em que:

- $A_{Cenário} \in \mathbb{R}^+$, área total do cenário, em metros quadrados (m^2);
- $A_{Sensor} \in \mathbb{R}^+$, área coberta por um sensor, em metros quadrados (m^2);
- $l \in \mathbb{R}^+$, tamanho do lado de um quadrado, em metros (m);

- $r \in \mathbb{R}^+$, raio do sensor, em metros (m);
- $N_{nós} \in \mathbb{N}^+$, número total de nós sensores da RSSF.

O resultado para o cenário 100,0 x 100,0 metros assinalou que, inicialmente, são necessários, no mínimo, 3 (três) nós sensores.

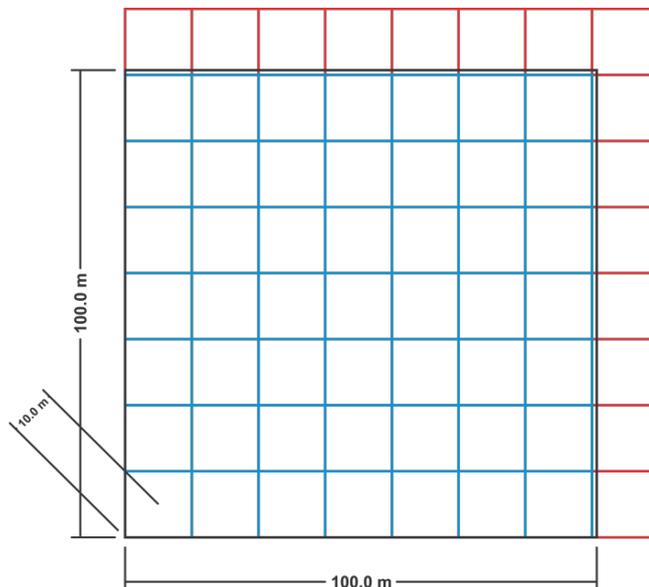
$$A_{Cenário} = (100,0 \text{ m})^2 = 10.000,0 \text{ m}^2,$$

$$A_{Sensor} = 2(10,0 \text{ m})^2 = 400,0 \text{ m}^2,$$

$$N_{nós} = \frac{10.000,0 \text{ m}^2}{200,0 \text{ m}^2} = 50.$$

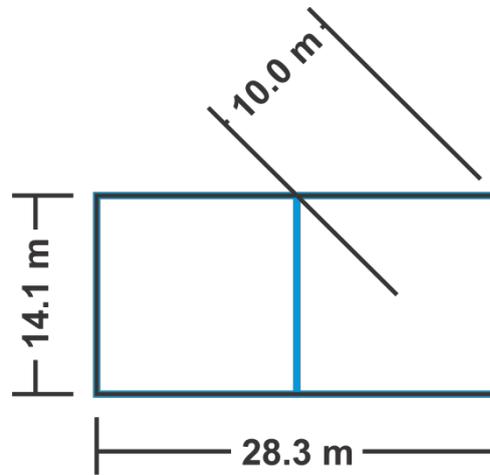
Apesar dos resultados apontarem a necessidade de apenas 50 nós sensores, foi necessária a utilização de 64 para cobrir 100% da área do cenário, ao se utilizar a técnica dos mosaicos planos. Esta diferença tem como causa a necessidade da utilização de 14 nós adicionais (em vermelho), para completar a cobertura do cenário. Na Figura 6.10 é ilustrada a resposta da técnica.

Figura 6.10 Configuração de implantação dos nós sensores para área de 100,0x100,0 m, utilizando a técnica de mosaicos planos



Nos cálculos para determinar a quantidade de nós sensores necessários para o cenário 28,3 x 14,1 metros, obteve-se um valor igual 2 (dois). Este mesmo resultado foi obtido quando se aplicou a técnica do mosaico plano. Na Figura 6.11, é representada a resposta da técnica.

Figura 6.11 Configuração de implantação dos nós sensores para um cenário com área de 28,3x14,1 m, utilizando técnica de mosaicos planos



Na segunda abordagem, foi utilizado um algoritmo de busca, baseado em técnicas aleatórias para encontrar uma configuração de implantação dos nós sensores para os dois cenários propostos, utilizando a menor quantidade de nós possível, mas mantendo a taxa de cobertura em 100,0%.

Utilizando as mesmas especificações para o raio de cobertura do nó sensor e os mesmos dois cenários já utilizados na técnica anterior, o algoritmo foi submetido a 30 rodadas de busca para o cenário de 28,3 x 14,1 metros e 40 rodadas de busca, para o cenário de 100,0 x 100,0 metros.

Em relação à quantidade de nós sensores utilizados para a busca do cenário 100,0 x 100,0 metros, primeiramente, foram utilizados 50 nós, mas nenhuma das buscas obteve sucesso. Assim, foram refeitas as buscas com 58 nós e os resultados deste experimento estão expostos na Tabela 6.9. Para o cenário 28,3 x 14,1 metros, foram utilizados apenas 2 (dois) nós sensores, já que a busca realizada com método gráfico obteve sucesso com apenas 2 (dois) nós. O resultado da busca é apresentado na Tabela 6.9.

Tabela 6.9 Resultados do experimento com o algoritmo completamente aleatório

	Cenário 100,0x100,0 metros	Cenário 28,3x14,1 metros
Rodadas	40	30
Sucessos	0	0
Tempo médio da busca (s)	500,2	29,3

Na última abordagem, foi utilizado o modelo proposto neste trabalho, que tem como base os AG. As condições do experimento foram as mesmas, sendo realizadas 30 repetições para o experimento do cenário 28,3 x 14,1 metros e 40 repetições para o experimento do cenário 100,0 x 100,0 metros. Novamente, as buscas foram iniciadas com a utilização de 50 nós sensores, e nenhum sucesso foi alcançado. Assim, passou-se a realizar as buscas com 58 nós. Para o cenário 28,3 x 14,1 metros, também só foram utilizados 2 (dois) nós sensores. Na Tabela 6.10, são apresentados os resultados dos dois experimentos. Na Figura 6.12, é ilustrado o resultado gráfico de uma saída do algoritmo de busca apresentado.

Figura 6.12 Representação gráfica do resultado de busca do algoritmo proposto nesta dissertação, para um cenário de 100,0x100,0 metros e raio do nó sensor de 10,0 metros

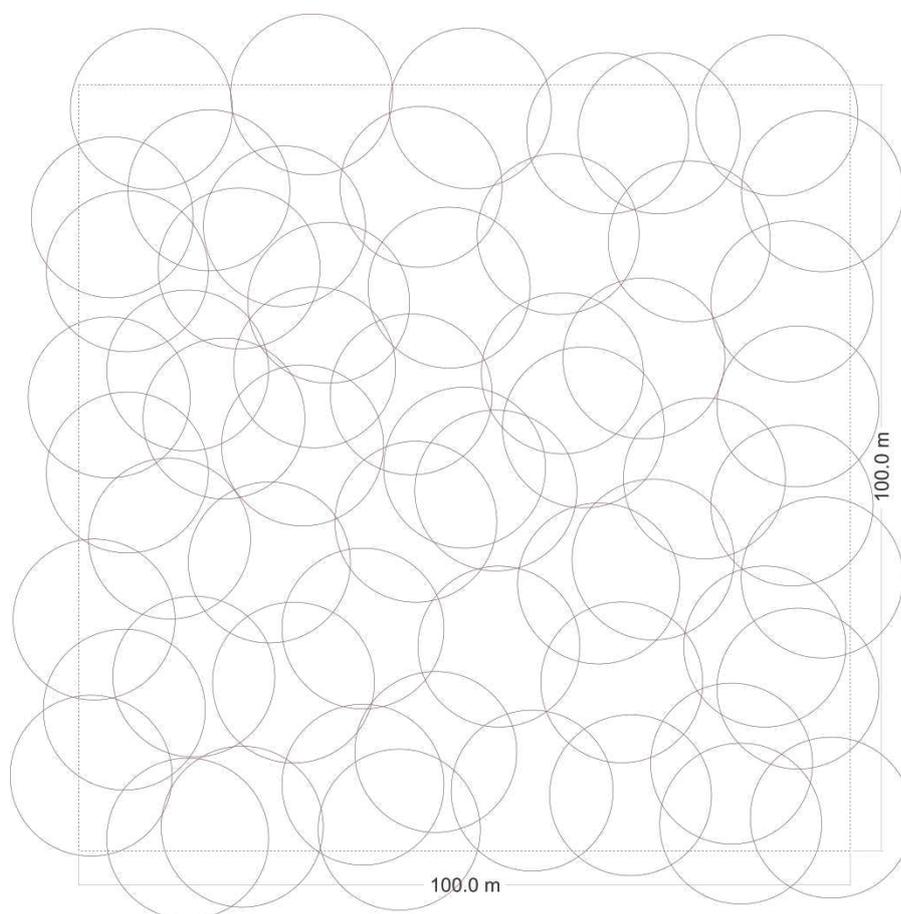


Tabela 6.10 Resultados do experimento com o algoritmo proposto

	Cenário 100,0x100,0 metros	Cenário 28,3x14,1 metros
Rodadas	40	30
Sucessos	36	0
Tempo médio da busca (s)	267,5	33,2

Os resultados apresentados na Tabela 6.9 e na Tabela 6.10 mostram a superioridade do algoritmo proposto neste trabalho, frente ao algoritmo baseado em buscas aleatórias. Outra constatação que foi possível obter é a de que, para cenários simples e que necessitem de poucos nós sensores, o uso do algoritmo proposto se equivale à utilização do algoritmo de busca aleatória e os dois não apresentam um bom desempenho. Portanto, nestas condições, não se deve utilizar o modelo proposto.

6.6 Desempenho do RSSF-AG comparado com outros trabalhos

Quintão et al. (2004) no seu trabalho propõem verificar a convergência do seu modelo, por meio de um experimento que tinha como objetivo encontrar a configuração de implantação dos nós sensores que proporcionasse a total cobertura do cenário. A rede e o cenário utilizados no experimento apresentavam as seguintes características:

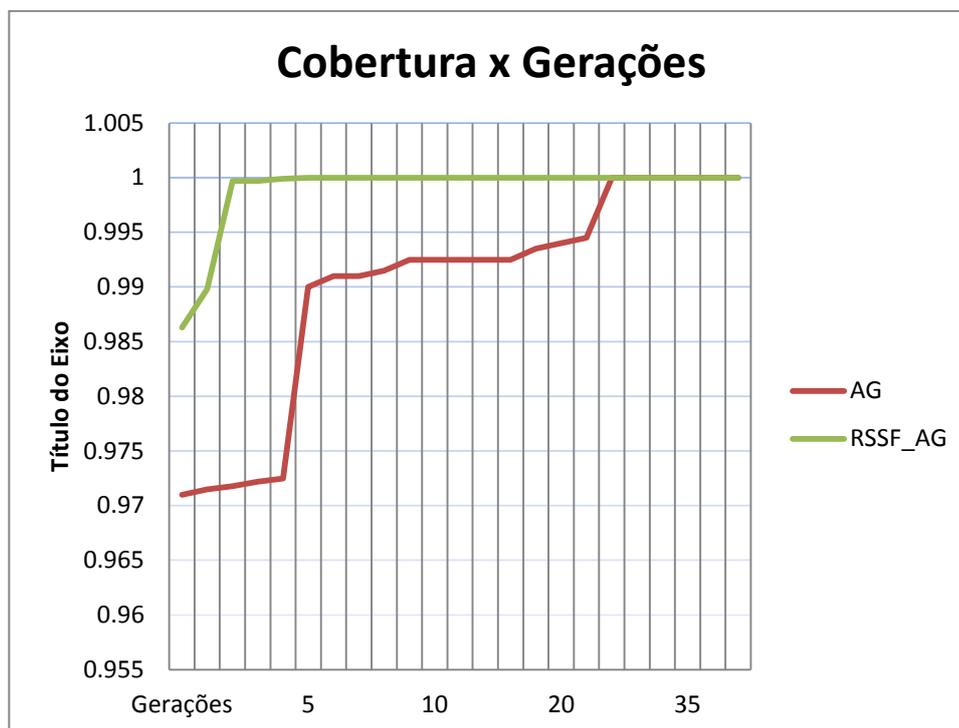
- Raio de sensoriamento dos nós sensores de 15 unidades de área;
- Cenário 50 x 50 unidades de área;
- Tamanho da população de 50 nós sensores;
- Probabilidade de mutação de 10%.

Utilizando os resultados obtidos no trabalho de Quintão et al. (2004) foi realizado um experimento semelhante ao proposto, como o objetivo de comparar o modelo proposto neste trabalho com o apresentado em Quintão et al. (2004).

Observa-se que o modelo proposto nesse trabalho (RSSF-AG) apresentou um desempenho pouco melhor do que o modelo utilizado em Quintão et al. (2004) (AG), pois o modelo RSSF-AG nos teste realizado consegue atingir o objetivo quando são utilizadas aproximadamente 6 gerações e o modelo AG com 25 gerações.

Na Figura 6.13 é ilustrado o gráfico da taxa de cobertura obtida em relação ao número de gerações utilizada na busca.

Figura 6.13 Comparação do desempenho do modelo RSSF-AG com os resultados da taxa de cobertura do trabalho de (QUINTÃO, MATEUS e NAKAMURA, 2004)



O melhor desempenho alcançado pelo modelo RSSF-AG pode ser explicado pelo fato que o modelo proposto por Quintão et al. (2004) utiliza um processo de seleção natural para criar sua população inicial. Esse processo consiste em gerar vários indivíduos utilizando um método aleatório com base em distribuição uniforme, logo após a geração desses indivíduos é aplicada a função de avaliação sobre cada um dos indivíduos da população para verificar o grau de aptidão de cada um, então os indivíduos melhor adaptados são selecionados e os demais descartados. No entanto, esse processo provoca um efeito não desejado, no qual os melhores genes para a solução do problema podem estar presentes em algum dos indivíduos descartados e com isso a informação valiosa que eles carregam não poderá mais ser usada.

Outras características que podem ter influenciado são a utilização do operador de cruzamento de 1 ponto e também do processo de seleção aleatório que foram aplicados no trabalho de Quintão et al. (2004).

6.7 Resumo dos Resultados

No Quadro 6.5, é apresentado o resumo dos resultados dos experimentos propostos neste trabalho, com o objetivo de demonstrar o funcionamento do modelo proposto.

Quadro 6.5 Resumo dos resultados dos experimentos

Experimento	Resultado
Verificar o ganho proporcionado pelo modelo quando é utilizada a métrica de taxa de cobertura.	O modelo proposto proporcionou um ganho médio de 28,2%.
Verificar o ganho proporcionado pelo modelo quando é utilizada a métrica de energia residual.	O modelo proposto proporcionou um ganho médio de 188,2%.
Verificar o ganho proporcionado pelo modelo quando é utilizada a métrica de TAFF.	O modelo proposto proporcionou um ganho médio de 72,8%.
Verificar o ganho proporcionado pelo modelo quando é utilizada a métrica de TPP.	O modelo proposto proporcionou um ganho médio de 83,2%.
Verificar o ganho proporcionado pelo modelo quando são utilizadas as métricas de taxa de cobertura e TAFF.	O modelo proposto proporcionou um ganho médio de 10,6%.
Verificar o ganho proporcionado pelo modelo quando são utilizadas as métricas de energia residual, TAFF e TPP.	O modelo proposto proporcionou um ganho médio de 18,4%.
Verificar o ganho proporcionado pelo modelo quando são utilizadas as métricas de taxa de cobertura, energia residual, TAFF e a TPP.	O modelo proposto proporcionou um ganho médio de 15,4%.
Verificar se a presença de obstáculos nos cenários de simulação influencia nos resultados do ganho do algoritmo, em relação à redução do consumo de energia.	A presença de obstáculos no cenário de simulação influencia no consumo de energia dos nós sensores.
Verificar se a presença de obstáculos nos cenários de simulação influencia nos resultados do TAFF da RSSF.	A presença de obstáculos no cenário de simulação influencia no TAFF da RSSF.
Verificar se a presença de obstáculos nos cenários de simulação influencia nos resultados do TPP da RSSF.	A presença de obstáculos no cenário de simulação influencia no TPP da RSSF.
Verificar a influência do número de nós sensores utilizados para forma um indivíduo sobre o desempenho do modelo proposto para um cenário de 100,0 x 100,0 metros.	O modelo proposto apresentou o melhor desempenho, pois conseguiu encontrar uma configuração de implantação que atendia os requisitos, em 90% das vezes que foi executado, utilizando 58 nós sensores.
Verificar a influência do número de nós sensores utilizados para forma um indivíduo sobre o desempenho do modelo proposto para um cenário de 28,3 x 14,1 metros.	O modelo proposto não conseguiu encontrar nenhuma configuração de implantação com os requisitos requeridos. No entanto, o método gráfico obteve sucesso, utilizando apenas 2 (dois) nós sensores. Esse resultado deixa explícito que a utilização do modelo proposto não é recomendada para problemas muito simples, como o proposto nesse experimento.

6.8 Discussão Geral

O objetivo principal deste trabalho foi o desenvolvimento de um modelo capaz de encontrar a melhor configuração de implantação dos nós sensores de uma RSSF, levando em consideração as métricas taxa de cobertura, energia residual, TAFF e TPP. A melhor configuração de implantação é caracterizada como sendo a que proporciona a maior taxa de cobertura, o maior tempo de vida, o menor TAFF e a menor TPP.

Ao final da execução do modelo proposto, é gerado um arquivo com a lista de coordenadas de localização dos nós sensores. Essas coordenadas serão utilizadas para implantar os nós sensores na área do cenário. Na Figura B.7 do Apêndice B, é ilustrado um exemplo do arquivo de saída.

Neste capítulo, foram apresentados os experimentos para teste do modelo proposto. Os resultados obtidos demonstram que a utilização do modelo proporcionou ganhos nas métricas de interesse.

7 Considerações Finais e Sugestões para Trabalhos Futuros

No trabalho ora apresentado, avaliou-se o desempenho de um algoritmo de busca para determinar a configuração de implantação dos nós sensores de uma RSSF, considerando as métricas de taxa de cobertura do cenário, energia residual dos nós sensores, TAFF na transmissão dos pacotes entre os nós sensores e o nó sorvedouro e a TPP. O algoritmo de busca proposto utilizou técnicas evolutivas baseadas nos mecanismos da evolução biológica da teoria de Darwin, mais especificamente um AG.

A utilização do AG foi fundamental para o bom desempenho do algoritmo de busca proposto, pois foi possível tratar, ao mesmo tempo, 4 (quatro) métricas diferentes (taxa de cobertura, energia residual, TAFF e TPP). Estas métricas tratadas pelo AG não são independentes, pois, em certas situações, uma influencia no resultado da outra, o que torna mais complicada a busca por máximos. Apesar destas dificuldades, os AG conseguem contornar esses problemas de forma simples e sem a necessidade de grande poder de processamento.

O tempo médio gasto pelo AG utilizado no modelo RSSF-AG para executar a busca de uma configuração de implantação é da ordem de 10^3 segundos. Enquanto que, a mesma busca utilizando um algoritmo exato (força bruta), no qual todas as combinações possíveis serão verificadas, o tempo dessa busca foi da ordem de 10^{131} segundos. Portanto, é impraticável a utilização do algoritmo de força bruta para encontrar a configuração de implantação dos nós sensores.

Além da utilização das 4 (quatro) métricas na busca da melhor configuração de implantação dos nós sensores de uma RSSF, este trabalho diferencia-se também dos demais, pois utiliza simuladores de rede com suporte à colocação de obstáculos nos cenários de simulação. Muitos trabalhos não levam em consideração a sua influência nos resultados da simulação. Este trabalho mostrou que, para a configuração de ambiente e características dos nós sensores utilizados, foram encontradas diferenças significativas nos resultados quando o módulo de obstáculos foi habilitado na simulação.

Uma característica importante do algoritmo proposto, que pôde ser observada durante a realização dos testes, foi a sua baixa eficiência quando utilizado em problemas muito simples, com poucos nós sensores a serem manipulados. Isto é uma característica dos AG, dado que precisam de diversidade para funcionar corretamente. Este problema pode ser observado nos resultados apresentados na Tabela 6.10, em que se tem apenas 2 (dois) nós sensores por cromossomo, o que torna muito difícil obter uma boa diversidade.

Conforme mencionada, uma característica dos AG é a de que sua convergência depende da diversidade entre os valores retornados pela função de avaliação. Quanto mais diferentes forem os valores de aptidão para cada indivíduo da população, mais diversa é a população, o que proporciona uma melhor convergência do algoritmo. No contexto deste trabalho, os valores retornados pela função de avaliação irão depender dos resultados gerados pelo simulador OMNeT++ 4.2.2, em conjunto com o *framework* INET 2.0.0. Especificamente, em relação à energia residual, os valores retornados pelo simulador não apresentaram uma boa diversidade, o que provocou uma deficiência no ganho obtido pelo modelo proposto. Em parte, esses valores de retorno com pouca diversidade podem ser explicados pelo modelo linear de bateria utilizado no ambiente de simulação. Esse modelo de bateria não é uma boa aproximação para uma bateria real.

Outra dificuldade encontrada neste trabalho, também causada pelo ambiente de simulação, foi o tempo de processamento. Durante os testes do algoritmo proposto, constatou-se que o tempo de processamento da simulação aumentava consideravelmente quando se aumentava o número de nós sensores e, também a quantidade de pacotes enviados. Assim, com o módulo de obstáculos desativado, o tempo médio para o algoritmo proposto encontrar a configuração foi de 3.238,6 segundos e, com o módulo ativado, foi de 4.941,3 segundos, o que representa uma diferença de, aproximadamente, 52%.

Apesar das dificuldades supracitadas, o algoritmo de busca implementado se mostrou eficiente, principalmente em se tratando do ganho da QoS. Foram obtidos ganhos consideráveis na redução do TAFF e também na TPP entre os nós sensores e o nó sorvedouro. Mesmo quando era habilitado o módulo de obstáculos nos cenários de simulação, o algoritmo conseguiu boas soluções para contornar essas dificuldades.

Outra característica relevante, que merece destaque, foi a utilização do arquivo de configuração para determinar o grau de importância de cada um dos fatores, Energia Residual, Taxa de Cobertura, TAFF e TPP, na busca da configuração de implantação dos nós sensores da RSSF. A utilização destes fatores proporcionou uma grande flexibilidade para o algoritmo, pois, devido a estes fatores, é possível realizar buscas direcionadas exclusivamente para uma das métricas ou definir um valor para cada uma das métricas na aplicação, ajustando adequadamente os fatores para refletir a importância de cada uma delas.

As principais contribuições apresentadas nesse trabalho foram:

- Formalização de um cromossomo para representar as soluções da melhor configuração de implantação dos nós sensores. Na Figura 7.1, é ilustrado um exemplo de um cromossomo com dois nós sensores;

Figura 7.1 Exemplo de um cromossomo para 2 (dois) nós sensores

0	1	1	0	0	1	0	0	0	0	1	1	0	0	1	0	1	0	0	1	0	1	1	0	1	0	0	1	0	1	1	0
X1								Y1								X2								Y2							
Nó 1																Nó 2															

- Formalização de uma função de avaliação utilizada pelo AG para quantificar a aptidão de cada indivíduo gerado, durante a busca da melhor configuração de implantação dos nós sensores. A função de avaliação está ilustrada na Equação (5.5);
- Padronização de valores dos parâmetros de configuração do AG, que foram ajustados com o objetivo de proporcionar o maior ganho para as buscas realizadas nos experimentos apresentados nesse trabalho. Os valores dos parâmetros estão disponibilizados no Quadro 6.4.

A utilização das 4 (quatro) métricas se faz necessária, a fim de permitir que a qualidade do serviço da rede não seja penalizada quando se busca apenas a maximização da energia residual. Um exemplo dessa perda de qualidade do serviço ocorre quando alguns nós sensores, que estão na periferia da rede, estão desconectados da rede. Todos os dados que eles geram não chegam ao nó sorvedouro. Em consequência, haverá um aumento na taxa de perda de pacotes,

que, para algumas aplicações, não é aceitável. Por outro lado, devido aos pacotes gerados pelo nó sensor desconectado não trafegarem pela rede, o nível da energia residual, ao final da simulação, será maior. Assim, a utilização das 4 (quatro) métricas proposta nesse trabalho é importante para determinar a melhor configuração de implantação dos nós sensores de uma RSSF.

7.1 Sugestões para Trabalhos Futuros

Com base na análise dos resultados apresentados nesta dissertação, foi possível observar a possibilidade de se desenvolver novos estudos, descritos a seguir.

Nos módulos de energia, os simuladores de RSSF geralmente utilizam modelos lineares para determinar a descarga de energia das baterias. Esses modelos não conseguem descrever com realismo o comportamento das baterias. Portanto, existe a necessidade de implementar novos modelos, como o proposto por Neto (2011), que é baseado em Rakhmatov-Vrudhula para o framework INET 2.0.0.

O módulo de obstáculos foi um dos principais problemas encontrados no desenvolvimento deste trabalho, pois a maioria dos simuladores não implementa modelos de obstáculos para as simulações de rede *ad hoc*. Assim, faz-se necessário tanto o desenvolvimento de novos modelos, como também um aprimoramento dos atuais, no tocante à carga adicional de processamento e na variação da sensibilidade de interferência dos obstáculos, em relação ao tipo de matéria que os constitui.

O modelo de propagação é outro problema que os simuladores desconsideram, principalmente os *opens source*, pois necessitam de mais ajustes às características dos cenários reais. Os modelos de propagação adotados pelos simuladores são, geralmente, simples e não consideram as atenuações provocadas pelas características do ambiente (como tipo de vegetação, materiais de construção de edifícios, dentre outros) e a altura do sistema de irradiação em relação ao solo (os modelos utilizados nos simuladores foram desenvolvidos para rede de celulares que estão a 1,5 metro do solo enquanto os nós de uma RSSF estão no nível do solo).

A integração do algoritmo proposto neste trabalho ao Google Maps proporcionaria uma melhoria considerável na qualidade da apresentação das respostas e também na sua usabilidade. A melhoria na apresentação das respostas

seria proporcionada pelos recursos disponibilizados na API do Google Maps, que permitiria a apresentação dos resultados por meio dos mapas, e o ganho na usabilidade seria resultante do uso de uma interface gráfica.

Há a necessidade de aprimorar o modelo de cobertura, fazendo com que este trate da influência da presença de obstáculos no cenário de simulação.

A utilização de um *cluster* para executar o modelo proposto reduziria consideravelmente o tempo de processamento. Para tanto, seria necessário modificar o algoritmo para que seja possível utilizar o poder de processamento do *cluster* nas simulações. Nesta dissertação, inicialmente o algoritmo proposto utilizava apenas um dos 4 (quatro) núcleos do processador core i7 para realizar a busca. A modificação realizada no algoritmo, para que fizesse uso dos 4 (quatro) núcleos, proporcionou a diminuição do tempo de busca. A alteração realizada no algoritmo consistiu em alterar a interface que executa o simulador, que, após a mudança, passou a executar 4 (quatro) instâncias do simulador simultaneamente, utilizando o comando *opp_runall* do Linux.

O algoritmo proposto precisa ser testado com outros protocolos de roteamento, protocolos de acesso ao meio e por outras aplicações.

Todas as sugestões apresentadas visam ao aprimoramento do modelo proposto neste trabalho, principalmente em relação ao melhoramento dos módulos de bateria e obstáculos, e, também, ao provável ganho de eficiência proporcionado pela utilização de *clusters* na execução da busca.

Referências Bibliográficas

ALVES, R. C. A. et al. Redes de Sensores sem Fio Aplicadas à Fisioterapia: Implementação e Validação de um Sistema de Monitoramento de Amplitude de Movimento. **Revista Brasileira de Sistemas de Informação - iSys**, p. 19, 2012.

ALVES, S. e DALCIN, M. Mosaicos do Plano. **Revista do Professor de Matemática**, p. 3–12, 1999.

BACK, T. HOFFMEISTER, F. e SCHWEFEL, H. P. A survey of evolution strategies. In: PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS. San Francisco, CA, USA: Morgan Kaufmann, 1991.

BAIOCO, J. S. **SÍNTESE E OTIMIZAÇÃO DE ROTAS DE DUTOS SUBMARINOS CONSIDERANDO CRITÉRIOS DE ESTABILIDADE HIDRODINÂMICA**. UFRJ - Rio de Janeiro, RJ, Brasil. 2011.

BARBOSA, F. LOPES, L. C. R. S. e CRUCIOL, M. Preditor de números para lotofácil: uma abordagem usando algoritmos evolutivos. In: ENACOMP. Catalão, GO, Brasil: UFG. Disponível em: <<http://www.enacomp.com.br/anais/pdf/49.pdf>>. Acesso em: 17 jul. 2013, 2013.

BECKMANN, M. **Algoritmos Genéticos Como Estratégia de Pré-Processamento em Conjunto de Dados Desbalanceados**. UFRJ - Rio de Janeiro, RJ, Brasil. 2010.

BERETTA, I. et al. Design exploration of energy-performance trade-offs for wireless sensor networks. In: PROCEEDINGS OF THE 49TH ANNUAL DESIGN AUTOMATION CONFERENCE ON - DAC '12. New York, USA: ACM Press. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2228360.2228549>>. Acesso em: 2 jun. 2013, 2012.

BRITO, W. De SEMAAN, G. e BRITO, J. de M. UM ALGORITMO GENÉTICO PARA O PROBLEMA DOS K-MÉDOIDES. **ic.uff.br**, p. 1–8, 2011.

BUDKE, G. F. et al. Uma Abordagem de Comunicação para Fusão de Dados Paralela em Redes de Sensores sem Fio IEEE 802 . 15 . 4. In: XXVI SIMPÓSIO BRASILEIRO DE ENGENHARIA DE SISTEMAS COMPUTACIONAIS. Florianópolis, SC, Brasil: publisher. Disponível em: <http://www.fe.up.pt/si/publs_pesquisa.formview?p_id=30615>. Acesso em: 22 jul. 2013, 2011.

BURIAN, R. **Algoritmos Genéticos na Alocação de Dispositivos de Proteção de Distribuição de Energia Elétrica**. USP - São Paulo, SP, Brasil. 2009.

BUTT, T. A. et al. TRENDY. In: PROCEEDINGS OF THE THIRD INTERNATIONAL WORKSHOP ON THE WEB OF THINGS - WOT '12. New York, New York, USA:

ACM Press. Disponível em: <<http://dl.acm.org/citation.cfm?id=2379758>>. Acesso em: 19 jul. 2013, 2012.

CARVALHO, S. et al. Uma Proposta para Gerenciamento de Energia em Redes de Sensores utilizando a metodologia Seis Sigma. In: XXX SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS. Ouro Preto, MG, Brasil: UFMG. Disponível em: <<http://sbrc2012.dcc.ufmg.br/app/pdfs/p-05/wgrs/WGRS-ST4-3.pdf>>. Acesso em: 2 jun. 2013, 2012.

CASTRO, L. De e ZUBEN, F. Vom. Representação e Operadores Evolutivos. . Campinas, SP, Brasil: DCA FEEC Unicamp. Disponível em: <ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia707_02/topico14_02.pdf>. Acesso em: 22 jul. 2013, 2012.

CHEN, X. e YU, P. Research on hierarchical mobile wireless sensor network architecture with mobile sensor nodes. In: 2010 3RD INTERNATIONAL CONFERENCE ON BIOMEDICAL ENGINEERING AND INFORMATICS. Yantai, China: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5639549>>. Acesso em: 19 jul. 2013, 2010.

CHEN, Y. CHUAH, C.-N. e ZHAO, Q. Sensor Placement for Maximizing Lifetime per Unit Cost in Wireless Sensor Networks. In: MILCOM 2005 - 2005 IEEE MILITARY COMMUNICATIONS CONFERENCE. Atlantic City, NJ, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1605825>. Acesso em: 19 jul. 2013, 2005.

CHENG, P. CHUAH, C. N. e LIU, X. Energy-aware node placement in wireless sensor networks. In: IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE, 2004. GLOBECOM '04. Davis, CA, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1378943>. Acesso em: 19 jul. 2013, 2004.

COELHO, L. dos S. e MARIANI, V. C. RESOLUÇÃO DO PROBLEMA DE WEBER USANDO PROGRAMAÇÃO EVOLUTIVA BASEADA EM DISTRIBUIÇÃO DE CAUCHY E ALGORITMO CULTURAL. In: XXXV SBPO. Natal, RN, Brasil: SBPO. Disponível em: <<http://www.din.uem.br/sbpo/sbpo2003/pdf/arq0154.pdf>>. Acesso em: 22 ago. 2013, 2003.

CORREIA, L. H. A. et al. Uma taxonomia para protocolos de controle de acesso ao meio em redes de sensores sem fio. In: ANNALES DES TELECOMMUNICATIONS. Paris, France: Departamento de Ciência da Computação, Universidade Federal de Minas Gerais. Disponível em: <<http://algot.dcc.ufma.br/~lcorreia/publicacoes/rt0505.pdf>>. Acesso em: 19 jul. 2013, 2005.

CUNHA, D. de O. COSTA, L. H. M. K. e DUARTE, O. C. M. B. Uma Análise do Consumo de Energia em Redes Ad Hoc. **Telecomunicações**, v. 07, n. 1, p. 39 – 47, 2004.

CUNHA, F. D. Da et al. Roteamento Oportunístico em Redes de Sensores Tolerantes a Atrasos e Desconexões. In: XXVIII SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS - SBRC 10. Gramado, RS, Brasil: LARC - SBC, 2010.

DANTAS, A. et al. DESRO: Um Protocolo de Roteamento com Gerenciamento Dinâmico de Energia para Redes de Sensores Sem Fio. In: XXVIII SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS - SBRC 10. Gramado, RS, Brasil: LARC - SBC. Disponível em: <http://ns1.dpi.ufv.br/arquivos/ppgcc/dissertacoes/2010-ms-Arthur_Vinicius_Maciel_Dantas.pdf>. Acesso em: 22 jul. 2013, 2010.

DEB, K. et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. **IEEE Transactions on Evolutionary Computation**, v. 6, n. 2, p. 182–197, doi:10.1109/4235.996017, 2002.

DHILLON, S. S. e CHAKRABARTY, K. Sensor placement for effective coverage and surveillance in distributed sensor networks. In: 2003 IEEE WIRELESS COMMUNICATIONS AND NETWORKING, 2003. WCNC 2003. New Orleans, LA, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1200627>. Acesso em: 25 jun. 2013, 2003.

ELEFATHERIOS, K. FUKAZAWA, Y. e OTA, J. Genetically optimizing query expansion for retrieving activities from the web. In: PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, MINING AND SEMANTICS - WIMS '12. New York, New York, USA: ACM Press. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2254129.2254190>>. Acesso em: 4 jun. 2013, 2012.

ELMAGID, A. a. a. et al. Radiation detection based heterogeneous Wireless Sensor Network. In: 2012 INTERNATIONAL CONFERENCE ON ENGINEERING AND TECHNOLOGY (ICET). Cairo, Egito: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6396122>>. Acesso em: 19 jul. 2013, 2012.

FERNANDES, A. M. da R. **Inteligência Artificial - Noções Gerais**. 1. ed. Florianópolis, SC, Brasil: Visual Books, 2003. p. 160

FOGEL, L. J. OWENS, A. J. e WALSH, M. J. **Artificial Intelligence Through Simulated Evolution**. [S.l.]: John Wiley & Sons, 1966. p. 170

FREEDMAN, A. **Dicionário de Informática: O Guia Ilustrado Completo**. 1. ed. São Paulo, SP, Brasil: Makron Books, 1995. p. 596

GABRIEL, P. H. R. e DELBEM, A. C. B. **Fundamentos de algoritmos evolutivos**. . São Carlos, SP, Brasil: ICMC-USP. Disponível em: <http://www.icmc.usp.br/~biblio/BIBLIOTECA/not_did/ND_075.pdf.pdf>. Acesso em: 19 jul. 2013, 2008.

GANESAN, D. CRISTESCU, R. e BEFERULL-LOZANO, B. Power-efficient sensor placement and transmission structure for data gathering under distortion constraints. **ACM Transactions on Sensor Networks**, v. 2, n. 2, p. 155–181, doi:10.1145/1149283.1149284, 2006.

GAULKE, E. **Metodologia para determinação do corte de carga em sistemas de distribuição utilizando inteligência artificial e múltiplos critérios**. UFSC - Florianópolis, SC, Brasil. 2007.

GOLDERG, D. E. **Genetic Algorithms in Search, Optimization and Machine Learning**. 1. ed. Boston, MA, USA: Addison-Wesley Longman, 1989. p. 372

GOMES, R. D. et al. Correlation between Spectral Occupancy and Packet Error Rate in IEEE 802.15.4-based Industrial Wireless Sensor Networks. **IEEE Latin America Transactions**, v. 10, n. 1, p. 1312–1318, doi:10.1109/TLA.2012.6142478, 2012.

GONÇALVES, R. T. et al. Multi-K: um protocolo de roteamento para redes de sensores sem fio usando árvores de espalhamento parciais. In: XXVI SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS - SBRC. Rio de Janeiro, RJ, Brasil: LARC - SBC. Disponível em: <<http://www.lbd.dcc.ufmg.br:8080/colecoes/sbrc/2008/055.pdf>>. Acesso em: 22 jul. 2013, 2007.

GUIMARÃES, F. G. e RAMALHO, M. C. **Implementação de um Algoritmo Genético**. . Belo Horizonte, MG, Brazil: UFMG. Disponível em: <<http://www.cpdee.ufmg.br/~lusoba/disciplinas/ele037/t1.pdf>>. Acesso em: 20 jul. 2013, 2001.

HAESER, G. e GOMES RUGGIERO, M. Aspectos Teóricos de Simulated Annealing e um Algoritmo duas Fases em Otimização Global. **TEMA - Tendências em Matemática Aplicada e Computacional**, v. 9, n. 3, doi:10.5540/tema.2008.09.03.0395, 2008.

HEATON, J. **Introduction to neural networks with Java**. 2. ed. St. Louis, Missouri, USA: Heaton Research, Inc., 2005. p. 380

HENRIQUES, F. da R. LOVISOLO, L. e RUBINSTEIN, M. G. Algoritmos para Aumentar o Tempo de Vida de Redes de Sensores Sem Fio Utilizando Inovação. In: XXVII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES (SBRT'09). Blumenau, SC, Brasil: SBrT. Disponível em: <<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Algoritmos+para+Aumentar+o+Tempo+de+Vida+de+Redes+de+Sensores+Sem+Fio+Utilizando+Inovação#0>>. Acesso em: 22 jul. 2013, 2009.

HOLLAND, J. H. **Adaptation in natural and artificial systems**. 1. ed. Cambridge, Massac, USA: 1st MIT Press, 1975. p. 211

HOU, Y.-T. et al. Node Placement for Optimal Coverage in Sensor Networks. In: IEEE INTERNATIONAL CONFERENCE ON SENSOR NETWORKS, UBIQUITOUS, AND TRUSTWORTHY COMPUTING -VOL 1 (SUTC'06). Taichung, China: IEEE.

Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1636200>. Acesso em: 20 jul. 2013, 2006.

INTANAGONWIWAT, C. GOVINDAN, R. e ESTRIN, D. Directed diffusion. In: PROCEEDINGS OF THE 6TH ANNUAL INTERNATIONAL CONFERENCE ON MOBILE COMPUTING AND NETWORKING - MOBICOM '00. New York, New York, USA: ACM Press. Disponível em: <<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Directed+Diffusion:+A+Scalable+and+Robust+Communication+Paradigm+for+Sensor+Networks#8>>. Acesso em: 20 jul. 2013, 2000.

ISHIZUKA, M. e AIDA, M. Performance study of node placement in sensor networks. In: 24TH INTERNATIONAL CONFERENCE ON DISTRIBUTED COMPUTING SYSTEMS WORKSHOPS, 2004. PROCEEDINGS. Tokyo, Japan: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1284093>. Acesso em: 20 jul. 2013, 2004.

JABOUR, E. C. M. G. **Um protocolo de Transporte Colaborativo para Redes de Sensores Sem Fio**. UFRJ - Rio de Janeiro, RJ, Brasil. 2009.

JONES, C. H. **IEEE 1451.4 smart transducers template description language**. Disponível em: <<http://standards.ieee.org/develop/regauth/tut/tdl.pdf>>. Acesso em: 20 jul. 2013.

KIM, D. et al. Minimum Data-Latency-Bound k-Sink Placement Problem in Wireless Sensor Networks. **IEEE/ACM Transactions on Networking**, v. 19, n. 5, p. 1344–1353, doi:10.1109/TNET.2011.2109394, 2011.

KOSTIN, S. e AMORIM, C. L. Controle de Potência de Transmissão para Redes de Sensores em Cenários com Obstáculos. In: XXV SIMPÓSIO BRASILEIRO DE ENGENHARIA DE SISTEMAS COMPUTACIONAIS. Belém, PA, Brasil: LARC - SBC. Disponível em: <<http://brenojac.com/sbrc/ST07 - 03.pdf>>. Acesso em: 19 jul. 2013, 2007.

KOULALI, M. et al. QDGRP : A hybrid QoS Distributed Genetic routing protocol for Wireless Sensor Networks. In: 2012 INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS. Tangier, Marrocos: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6320271>>. Acesso em: 4 jun. 2013, 2012.

LI, Q. LI, C. e LI, J. Data Aggregation Algorithm Based on Grid and Adaptive Genetic Algorithm for Wireless Sensor Networks with a Mobile Sink. In: 2010 2ND INTERNATIONAL WORKSHOP ON INTELLIGENT SYSTEMS AND APPLICATIONS. Wuhan, China: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5473378>>. Acesso em: 20 jul. 2013, 2010.

LIN, J. e LIAO, M. A clustering patch hierarchical routing protocol for wireless sensor networks. In: 2010 5TH INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE & EDUCATION. Hefei, China: IEEE. Disponível em:

<<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5593457>>. Acesso em: 20 jul. 2013, 2010.

LINDEN, R. **Algoritmos genéticos**. 1. ed. Rio de Janeiro, RJ, Brasil: Brasport Livros e Multimídia Ltda, 2006. p. 348

LOBO, E. L. M. **Uma solução do problema de horário escolar via algoritmo genético paralelo**. Dissertação (Mestrado) - Curso de Mestrado em Modelagem Matemática, CEFET-MG - Belo Horizonte, MG, Brazil. 2005.

LOUREIRO, A. A. F. et al. Redes de sensores sem fio. In: XXI SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES. Natal, RN, Brasil: SBRC. Disponível em: <http://www.sensornet.dcc.ufmg.br/publica/pdf/179_Loureiro_Nogueira_Ruiz_Mini_Nakamura_Figueiredo.pdf>. Acesso em: 11 nov. 2012, 2003.

LUGO, P. L. M. e TEIXEIRA JR., R. F. Algoritmo Genético com Busca Local para a Programação da Produção em Sistemas Flow-Shop Híbridos. In: CONGRESSO LATINO-IBEROAMERICANO DE INVESTIGAACIÓN OPERATIVA E SIMPOSIO BRASILEIRO DE PESQUISA OPERACIONAL. Rio de Janeiro, RJ, Brasil: ALIO, 2012.

MACEDO, D. F. **Um protocolo de roteamento para redes de sensores sem fio adaptável por regras de aplicação**. UFMG - Belo Horizonte, MG, Brazil. 2006.

MACEDO, P. J. da S. F. A. De. **Desenvolvimento de Modelos de Simulação de Redes de Sensores sem Fios**. UMinho - Braga, Portugal. 2010.

MATOS, T. B. BRAYNER, Â. e MAIA, J. E. B. Agregação e predição de dados em rede com precisão ajustável no processamento de consultas em redes de sensores sem fio. In: 23RD BRAZILIAN SYMPOSIUM ON DATABASES. Porto Alegre, RS, Brasil: Sociedade Brasileira de Computação. Disponível em: <<http://dl.acm.org/citation.cfm?id=1498947>>. Acesso em: 22 jul. 2013, 2008.

MELO, J. C. e. **Uma Solução de Controle de Densidade Centrado em Dados para Redes de Sensores Sem Fio**. UFMG - Belo Horizonte, MG, Brazil. 2009.

MICHALEWICZ, Z. e SCHOENAUER, M. Evolutionary algorithms for constrained parameter optimization problems. **Evolutionary computation**, 1996.

MINKU, F. L. LUDERMIR, T. B. e ARAÚJO, A. F. R. Computação evolucionária para otimização dinâmica de parâmetros de EFuNNs. In: V ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL. São Leopoldo, RS, Brazil: SBC. Disponível em: <<http://www.lbd.dcc.ufmg.br/bdbcomp/servlet/Trabalho?id=8423>>. Acesso em: 22 ago. 2013, 2005.

MIRANDA, F. A. M. **Proposta de alocação de baterias em redes de sensores sem fio orientada à maximização do tempo de vida**. UNICAMP - Campinas, SP, Brasil. 2011.

NAKAMURA, F. G. **Planejamento dinâmico para controle de cobertura e conectividade em redes de sensores sem fio planas**. Belo Horizonte, MG, Brazil: UFMG, 2003.

NAMBA, M. de O. **Modelagem e Especificação de um Middleware para Redes de Sensores Sem Fio Aplicado à Saúde**. Dissertação (Mestrado) - Pós-Graduação do Instituto de Informática, UFG - Goiânia, GO, Brasil. 2011.

NETO, J. A. T. C. **Análise de Protocolos de Roteamento Unicast em Redes Ad Hoc Móveis Baseada em um Modelo Realístico de Bateria**. UFCG - Campina Grande, PB, Brasil. 2011.

OLIVEIRA, C. E. R. De. **Utilização de nó sorvedouro móvel para balancear a carga de energia em redes de sensores sem fio**. UFV - Viçosa, MG, Brasil. 2011.

PINTO, A. R. DANTAS, M. A. R. e BITENCORT, B. R. Fusão de Dados Tempo Real em Redes de Sensores Sem Fio Multimídia. In: SIMPÓSIO BRASILEIRO DE SISTEMAS MULTIMÍDIA E WEB. Gramado, RS, Brasil: SBC. Disponível em: <<http://www.das.ufsc.br/~montez/publications/2007 Webmedia Alex.pdf>>. Acesso em: 22 jul. 2013, 2007.

QUINTÃO, F. P. MATEUS, G. R. e NAKAMURA, F. G. Uma abordagem evolutiva para o problema de cobertura em Redes de Sensores sem fio. In: XXIV CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. Salvador, BA, Brasil: SBC. Disponível em: <<http://www.sensornet.dcc.ufmg.br/pdf/fred04abordagem.pdf>>. Acesso em: 22 jul. 2013, 2004.

ROCHA, M. L. e SILVA, W. G. Algoritmo Genético Aplicado à Configuração de uma Rede de Sensores Sem Fio para Cobertura de uma Região. In: I ENCONTRO REGIONAL DE PESQUISA OPERACIONAL DO NORTE. Manaus, AM, Brasil: SOBRAPO, 2011.

RUYAN, Z. et al. Genetic clustering route algorithm in WSN. **Sixth International Conference on Natural Computation (ICNC 2010)**, n. Icnc, p. 4023–4026, doi:10.1109/ICNC.2010.5584826, 2010.

SCHONFELD, J. e LOUIS, S. J. The effect of selection on the development of mutational robustness. In: 2007 IEEE CONGRESS ON EVOLUTIONARY COMPUTATION. Singapore: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4425090>>. Acesso em: 20 jul. 2013, 2007.

SENGUPTA, S. et al. An Evolutionary Multiobjective Sleep-Scheduling Scheme for Differentiated Coverage in Wireless Sensor Networks. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 42, n. 6, p. 1093–1102, doi:10.1109/TSMCC.2012.2196996, 2012.

SHAH, R. C. e RABAEY, J. M. Energy aware routing for low energy ad hoc sensor networks. In: WIRELESS COMMUNICATIONS AND NETWORKING

CONFERENCE, 2002. WCNC2002. 2002 IEEE. Orlando, FL, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=993520>. Acesso em: 22 jul. 2013, 2002.

SHAPIRO, S. S. e FRANCA, R. S. An Approximate Analysis of Variance Test for Normality. **Journal of the American Statistical Association**, v. 67, n. 337, p. 215–216, doi:10.1080/01621459.1972.10481232, 1972.

SILVA, A. P. e GOULART, C. C. Multi-Geo: um protocolo de roteamento hierárquico para redes de sensores sem fio. In: XXXV CONFERÊNCIA LATINO-AMERICANA DE INFORMÁTICA. Pelotas, RS, Brasil: CLEI. Disponível em: <<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multi-Geo:+Um+Protocolo+de+Roteamento+Hierarquico+para+Redes+de+Sensores+Sem+Fio#0>>. Acesso em: 22 jul. 2013, 2009.

SILVA, C. V. Da FRERY, A. C. e VIANA, P. Avaliação do Consumo de Energia e Tempo de Vida de Redes de Sensores Sem Fio Comerciais. In: XXIX MECÂNICA COMPUTACIONAL. Buenos Aires, Argentina: Asociación Argentina de Mecánica Computacional, 2010.

SONG, X. WANG, C. e PEI, J. 2ASenNet: A multiple QoS metrics hierarchical routing protocol based on swarm intelligence optimization for WSN. In: 2012 IEEE INTERNATIONAL CONFERENCE ON INFORMATION SCIENCE AND TECHNOLOGY. Wuhan, China: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6221703>>. Acesso em: 20 jul. 2013, 2012.

SWIECH, M. C. S. **Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial**. UTFPR - Curitiba, PR, Brasil. 2005.

TANENBAUM, A. S. **Redes de computadores**. 4th. ed. Rio de Janeiro, RJ, Brasil: Editora Campus, 2003. p. 945

TUNA, G. et al. Mobile Robot Aided Self-Deploying Wireless Sensor Networks for Radiation Leak Detection. In: 2012 5TH INTERNATIONAL CONFERENCE ON NEW TECHNOLOGIES, MOBILITY AND SECURITY (NTMS). Istanbul, Turquia: IEEE. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6208745>>. Acesso em: 4 jun. 2013, 2012.

TÜXEN, M. RÜNGELER, I. e RATHGEB, E. Interface connecting the INET simulation framework with the real world. In: PROCEEDINGS OF THE 1ST INTERNATIONAL CONFERENCE ON SIMULATION TOOLS AND TECHNIQUES FOR COMMUNICATIONS, NETWORKS AND SYSTEMS & WORKSHOPS. Marseille, France: ICST. Disponível em: <<http://dl.acm.org/citation.cfm?id=1416267>>. Acesso em: 22 jul. 2013, 2008.

VARGA, A. e HORNIG, R. AN OVERVIEW OF THE OMNeT++ SIMULATION ENVIRONMENT. In: PROCEEDINGS OF THE FIRST INTERNATIONAL ICST CONFERENCE ON SIMULATION TOOLS AND TECHNIQUES FOR

COMMUNICATIONS NETWORKS AND SYSTEMS. Marseille, France: ICST. Disponível em: <<http://eudl.eu/doi/10.4108/ICST.SIMUTOOLS2008.3027>>. Acesso em: 22 jul. 2013, 2008.

VERONA, A. B. **Simulação e análise de redes de sensores sem fio aplicadas à viticultura**. UEM - Maringá, PR, Brasil. 2010.

VIANA, G. Padrão IEEE 754 para Aritmética Binária de Ponto Flutuante. **Revista CT**, p. 29–43, 1999.

YE, W. HEIDEMANN, J. e ESTRIN, D. An energy-efficient MAC protocol for wireless sensor networks. In: INFOCOM 2002. TWENTY-FIRST ANNUAL JOINT CONFERENCE OF THE IEEE COMPUTER AND COMMUNICATIONS SOCIETIES. PROCEEDINGS. IEEE. New York, NY, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1019408>. Acesso em: 22 jul. 2013, 2002.

YOUNIS, M. e AKKAYA, K. Strategies and techniques for node placement in wireless sensor networks: A survey. **Ad Hoc Networks**, v. 6, n. 4, p. 621–655, doi:10.1016/j.adhoc.2007.05.003, 2008.

ZILIC, Z. RADECKA, K. e CHENARD, J.-S. Architectures of increased availability wireless sensor network nodes. In: 2004 INTERNATIONAL CONFERENCE ON TEST. San Jose, CA, USA: IEEE. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1387396>. Acesso em: 19 jul. 2013, 2004.

ZINI, É. de O. C. **Algoritmo Genético Especializado na Resolução de Problemas com Variáveis Contínuas e Altamente Restritos**. UNIVERSIDADE ESTADUAL PAULISTA - [S.I.]. 2009.

APÊNDICE A Análise Estatística

A.1 Parâmetro Tamanho da População

Inicialmente, foi necessário realizar uma análise descritiva dos dados coletados nos testes do algoritmo. Para esta análise, foi utilizado o *software* de estatística R. A análise descritiva caracterizou os dados partir de medidas que resumiram a informação em: média, mediana, *quartis*, valores máximo e mínimo, desvio padrão, variância, coeficiente de variação e coeficiente de correlação do ganho do Algoritmo em relação ao tamanho da população. Na Tabela A.1 e na Tabela A.2 são explicitados os resultados obtidos pelo software R.

Tabela A.1 Medidas da análise descritivas dos testes para o tamanho da população

Parâmetros	Ganho do Algoritmo
Mínimo	0,029
1º Quartil	0,083
Mediana	0,105
Média	0,104
3º Quartil	0,126
Máximo	0,201
Desvio Padrão	0,035
Variância	0,001
Coeficiente de Variação	33,379

Tabela A.2 Coeficiente de correlação desempenho do algoritmo e tamanho da população

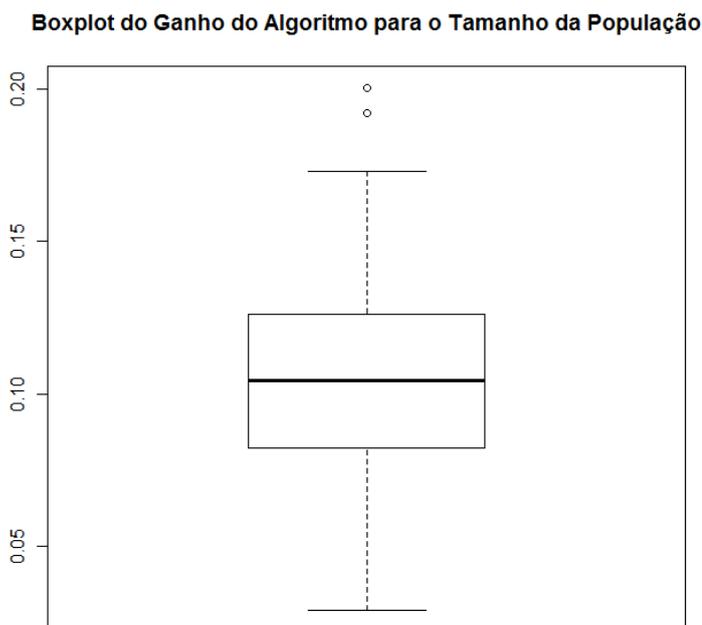
	Ganho do Algoritmo	Tamanho da População
Ganho do Algoritmo	1,000	-0,164
Tamanho da População	-0,164	1,000

Analisando-se os resultados dos testes, em especial o coeficiente de variação que ficou abaixo de 50%, pôde-se concluir que as médias são representativas. Outra informação importante, observada nos resultados da Tabela A.1 e da Tabela A.2 é o coeficiente de correlação, que ficou em (-0,164), bem

próximo de (0,000). Logo, fica demonstrado, que não existe uma relação linear entre as duas variáveis em estudo.

Na Figura A.1 apresenta-se com mais detalhes o *boxplot* da variável Ganho do Algoritmo para o Tamanho da População (GATP).

Figura A.1 Boxplot da variável ganho do algoritmo para o tamanho da população



Da Figura A.1 e da Tabela A.1 concluiu-se que a variável GATP possuía distribuição simétrica, pois a linha mediana estava equidistante dos extremos da caixa do *boxplot*. A maior concentração dos dados estava entre (0,083) e (0,126). A distância interquartil indica que os dados têm tendência moderada a serem homogêneos e existem duas medidas como possíveis *outlier*.

Nesta avaliação existiam duas variáveis, sendo a variável “Ganho do Algoritmo” é do tipo quantitativo contínuo, pois pôde-se utilizar qualquer valor real entre 0 (zero) e 1 (um) na sua representação. A variável GATP é do tipo qualitativa ordinal, pois o teste realizado para determinar o tamanho da população, que proporcione o melhor desempenho para o algoritmo foi configurado para cinco valores distintos 20, 30, 40, 60, 80 e 100. Logo, têm-se as seguintes variáveis:

- W : variável de trabalho do GATP (em porcentagem);
- Y : variável explicativa do tamanho da população (em unidades).

Para a aplicação dos testes paramétricos com objetivo de verificar se existiam diferenças entre as médias do GATP e o “Tamanho da População” foi preciso verificar se os dados possuíam uma distribuição normal. Os dados para serem considerados normais precisam ter no mínimo 30 amostras, caso o número de amostras seja menor, é preciso aplicar o teste de *Shapiro-Wilk*. Outra condição necessária para a utilização dos testes paramétricos é que a distribuição também tenha homogeneidade de variâncias, que pôde ser verificada pelo teste de *Bartlett*. Na Tabela A.3 são apresentados os resultados do teste de normalidade da distribuição e na Tabela A.4 o teste de homogeneidade da variância, aplicados aos dados do “Ganho do Algoritmo”.

Tabela A.3 Resultado do teste de normalidade da distribuição para tamanho da população

Shapiro-Wilk	Ganho do Algoritmo
W_{calc}	0,993
W_{α}	0,927
<i>p-valor</i>	0,701
α	0,050

Tabela A.4 Resultado do teste de homogeneidade da variância para o tamanho da população

Bartlett	Ganho do Algoritmo
<i>K-squared</i>	2,633
<i>df</i>	4,000
<i>p-valor</i>	0,621

O resultado apresentado na Tabela A.3 confirmou que os dados extraídos do experimento têm uma distribuição normal, valor de $W_{calc} > W_{\alpha}$ e $p\text{-valor} > \alpha$. O resultado na Tabela A.4 confirmou a homogeneidade das variâncias. Portanto, foi possível aplicar os testes paramétricos para verificar se existiam diferenças entre as médias de ganho do algoritmo em relação ao tamanho da população de indivíduos de cada geração.

Neste experimento buscou-se comparar as médias de ganho do algoritmo dos 5 (cinco) grupos, que estavam separados pelo tamanho da população utilizada nos testes de cada grupo. Como se tem mais de 2 (dois) grupos a serem comparados, foi escolhido o teste *ANOVA* de uma via (*Oneway ANOVA*). O resultado do teste *Oneway ANOVA* é exposto na Tabela A.5.

Tabela A.5 Resultado do teste para verificar se existem diferenças entre as médias dos grupos

Oneway ANOVA	Valores
<i>F</i>	1,908
<i>Num DF</i>	4,000
<i>Denom DF</i>	72,317
<i>p-valor</i>	0,118
α	0,050

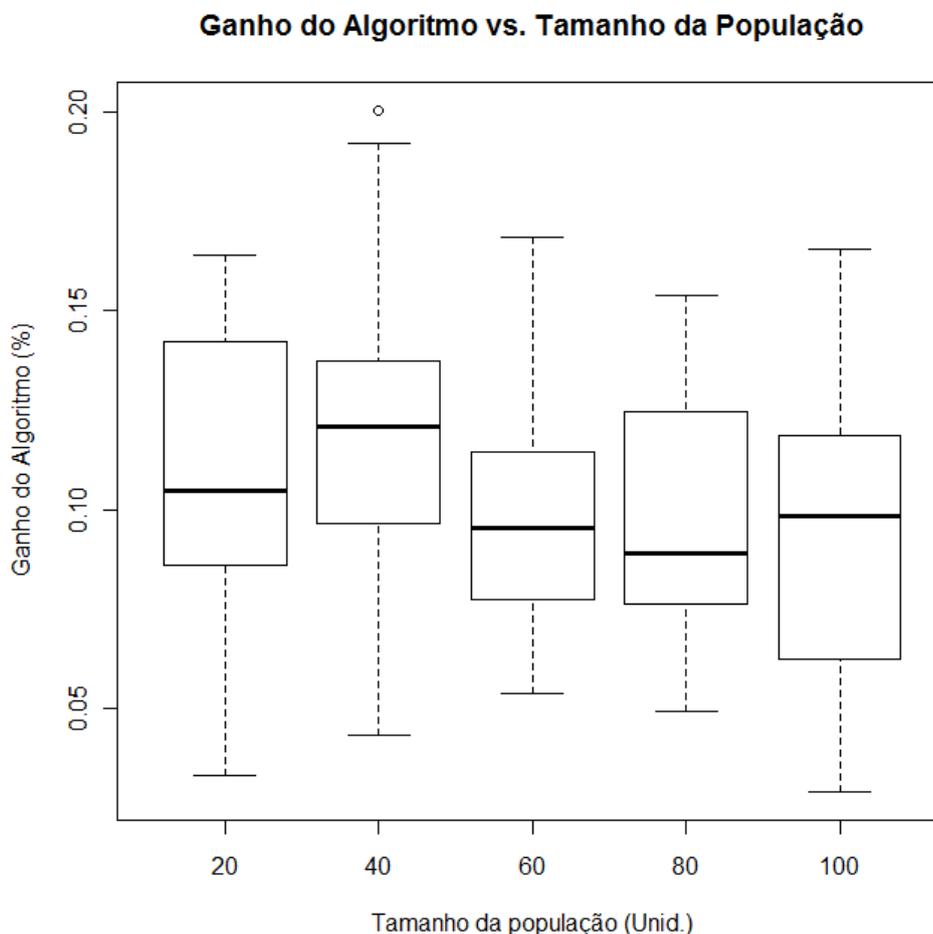
O resultado do *p-valor* do teste ANOVA de uma via, como pode ser verificado na Tabela A.5, obteve um valor maior que o α . Portanto, pode-se afirmar com um nível de confiança de 95,0% que não há diferença significativa entre as médias do ganho do algoritmo em relação aos 5 (cinco) grupos de tamanho de população diferentes. No entanto, ainda precisava-se escolher qual tamanho da população a ser utilizado no algoritmo de busca. Esta escolha foi feita baseado no *boxplot* do ganho do algoritmo em relação aos 5 (cinco) grupos, que pode ser observado na Figura A.2 e também nos resultados dos testes *t* de *Student* apresentados na Tabela A.6.

Tabela A.6 Ganho do algoritmo em função do tamanho da população

Tamanho da População	Média	IC	
20	0,106	0,092	0,120
40	0,118	0,105	0,132
60	0,100	0,088	0,111
80	0,098	0,087	0,109
100	0,096	0,082	0,110

Do *boxplot* da Figura A.2 percebe que a mediana do grupo de tamanho da população de 40 indivíduos foi a que apresentou melhor desempenho em relação aos demais grupos. O mesmo grupo de 40 indivíduos também apresentou o melhor resultado nos testes *t* de *Student*, como pode ser verificado na Figura A.2. Portanto, baseando-se nesta análise gráfica, o valor para o parâmetro tamanho da população a ser utilizado nos demais teste foi de 40 indivíduos.

Figura A.2 Boxplot do Ganho do Algoritmo em relação aos grupos do tamanho da população



A.2 Parâmetro Método de Seleção

Para o estudo estatístico do método de seleção foi realizada a mesma análise descritiva utilizada no estudo do tamanho da população. Na Tabela A.7 e na Tabela A.8 são explicitados os resultados obtidos pelo *software* R.

Analisando-se o resultado do teste do coeficiente de variação, que ficou abaixo de 50% para a variável Ganho do Algoritmo do Método de Seleção (GAMS), pôde-se concluir que as médias foram representativas. Outra informação importante que pôde ser retirada dos resultados na Tabela A.8 foi o coeficiente de correlação. Nos testes, o coeficiente de correlação foi (0,143), demonstrando-se estatisticamente a não existência de uma relação linear entre as duas variáveis em estudo.

Tabela A.7 Medidas da análise descritivas dos testes para o método de seleção

Parâmetros	Ganho do Algoritmo
Mínimo	0,060
1º Quartil	0,106
Mediana	0,137
Média	0,138
3º Quartil	0,167
Máximo	0,267
Desvio Padrão	0,042
Variância	0,002
Coefficiente de Variação	30,285

Tabela A.8 Coeficiente de correlação ganho do algoritmo e o método de seleção

	Ganho do Algoritmo	Método de Seleção
Ganho do Algoritmo	1,000	0,143
Método de Seleção	0,143	1,000

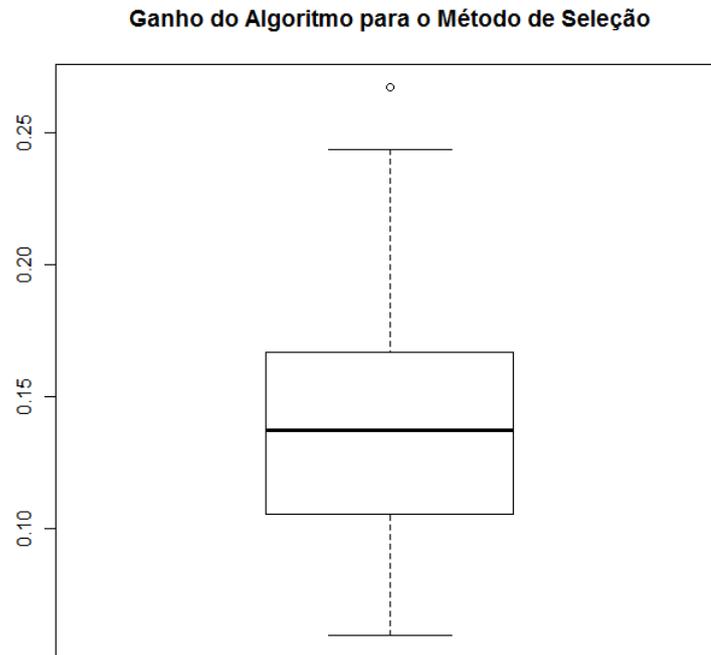
Na Figura A.3, é apresentado em mais detalhes o *boxplot* da variável GAMS.

Observando-se a Figura A.3 e a Tabela A.7, concluiu-se que a variável Ganho do Algoritmo para o método de seleção possuía distribuição levemente assimétrica à direita, pois a linha mediana estava um pouco deslocada para o extremo superior da caixa do *boxplot*. A maior concentração dos dados está entre (0,106) e (0,167), a distância interquartil indica que os dados têm tendência moderada a serem homogêneos e uma das medidas é um possível *outlier*.

Nesta avaliação existiam duas variáveis, sendo a variável Ganho do Algoritmo do tipo quantitativo contínuo, pois pode-se utilizar qualquer valor real entre 0 (zero) e 1 (um) na sua representação. A variável “Método de Seleção” é do tipo qualitativo nominal formada por 5 (cinco) grupos, em que 1 (um) dos grupos utiliza o método da Roleta Viciada e os outros 4 (quatro) grupos utilizam o método por Torneio (Roleta, *rank2*, *rank4*, *rank6* e *rank8*). Logo, têm-se as seguintes variáveis:

- W: variável de trabalho do GAMS (em porcentagem);
- Y: variável explicativa do Método de Seleção.

Figura A.3 Boxplot da variável ganho do algoritmo para o método de seleção



Os mesmo testes estatísticos utilizado no estudo sobre o tamanho da população, foram aplicados aqui. Na Tabela A.9, são apresentados os resultados do teste de normalidade da distribuição e na Tabela A.10 o teste de homogeneidade da variância, aplicados aos dados do GAMS.

Tabela A.9 Resultado do teste de normalidade da distribuição para método de seleção

Shapiro-Wilk	Ganho do Algoritmo
W_{calc}	0,987
W_{α}	0,927
$p\text{-valor}$	0,154
α	0,050

O resultado apresentado na Tabela A.9 confirma que os dados extraídos do experimento possuíam uma distribuição normal, valor de $W_{calc} > W_{\alpha}$ e $p\text{-value} > \alpha$. O resultado da Tabela A.10 confirmou a homogeneidade das variâncias. Portanto, foi possível aplicar os testes paramétricos para verificar se existiam diferenças entre as médias de ganho do algoritmo em relação ao método de seleção dos pais para a geração dos indivíduos da próxima geração.

Tabela A.10 Resultado do teste de homogeneidade da variância para o método de seleção

Bartlett	Ganho do Algoritmo
<i>K-squared</i>	4,497
<i>df</i>	4,000
<i>p-valor</i>	0,343

Neste experimento buscou-se comparar as médias de GAMS dos 5 (cinco) grupos, que são diferenciados pelo método de seleção adotado nos testes de cada grupo. Como se tem mais de dois grupos a serem comparados foi escolhido o teste ANOVA de uma via (*Oneway ANOVA*). O resultado do teste *Oneway ANOVA* é exposto na Tabela A.11.

Verifica-se na Tabela A.11 que o resultado do *p-valor* do teste ANOVA de uma via proporcionou um valor menor que o α . Portanto, pôde-se afirmar, com um nível de confiança de 95,0%, que há diferença significativa entre as médias do GAMS em relação aos cinco grupos do método de seleção. Logo, foi realizado o teste *t* de *Student* para verificar qual dos grupos proporcionaria o melhor ganho do algoritmo e também foi feita uma análise gráfica a partir do boxplot. Na Tabela A.12 e na Figura A.4 são apresentados os resultados dos testes realizados.

Tabela A.11 Resultado do teste para verificar se existem diferenças entre as médias dos grupos para o método de seleção

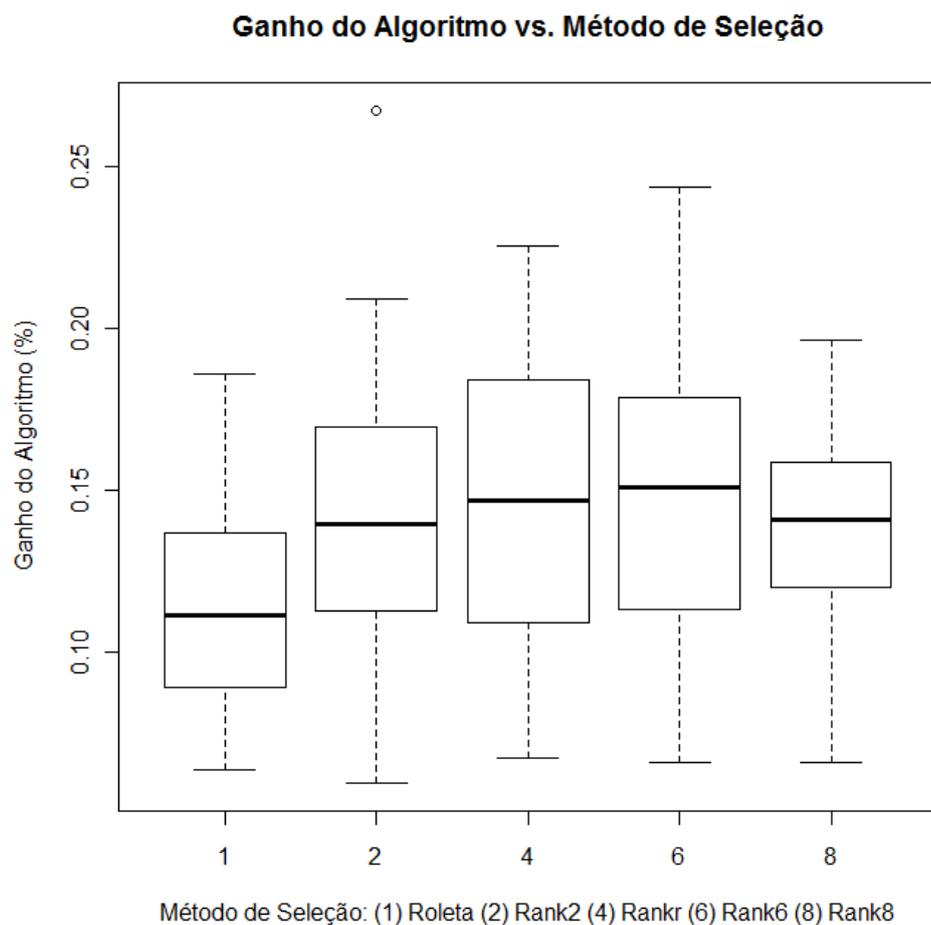
Oneway ANOVA	Valores
<i>F</i>	3,936
<i>Num df</i>	4,000
<i>Denom df</i>	72,191
<i>p-valor</i>	0,006
α	0,050

No *boxplot* da Figura A.4 percebe-se que a mediana do grupo *Rank6* foi a que apresentou melhor desempenho em relação aos demais grupos. O mesmo grupo também demonstrou o melhor resultado no *t-test* como pode ser verificado na Tabela A.12. Logo, baseando-se na análise gráfica e nos resultados dos testes na Tabela A.12 o método de seleção que utilizado no algoritmo foi o baseado em Torneiro com 6 (seis) indivíduos.

Tabela A.12 Ganho do algoritmo em função do método de seleção

Método de Seleção	Médi a	IC	IC
Roleta	0,116	0,103	0,128
Rank2	0,144	0,127	0,160
Rank4	0,146	0,129	0,163
Rank6	0,150	0,134	0,166
Rank8	0,137	0,124	0,150

Figura A.4 Boxplot do ganho do algoritmo em função dos grupos do método de seleção



A.3 Parâmetro Método de Cruzamento

Para o estudo estatístico do método de cruzamento foi realizada a mesma análise descritiva utilizada no estudo do tamanho da população. Na Tabela A.13 e na Tabela A.14 são explicitados os resultados obtidos pelo *software* R.

Tabela A.13 Análise descritiva do ganho do algoritmo para o método de cruzamento

Parâmetros	Ganho do Algoritmo
Mínimo	0,012
1º Quartil	0,062
Mediana	0,106
Média	0,103
3º Quartil	0,141
Máximo	0,242
Desvio Padrão	0,051
Variância	0,003
Coeficiente de Variação	49,275

Analisando-se o resultado do teste do coeficiente de variação, que ficou abaixo de 50,0% para a variável Ganho do Algoritmo para Configuração de Cruzamento (GACC), pôde-se concluir que as médias são representativas. Outra informação importante que pôde ser retirada dos resultados expostos na Tabela A.14 é o coeficiente de correlação. Nos testes, o coeficiente de correlação ficou em (0,086), demonstrando-se estatisticamente a não existência de uma relação linear entre as 2 (duas) variáveis em estudo.

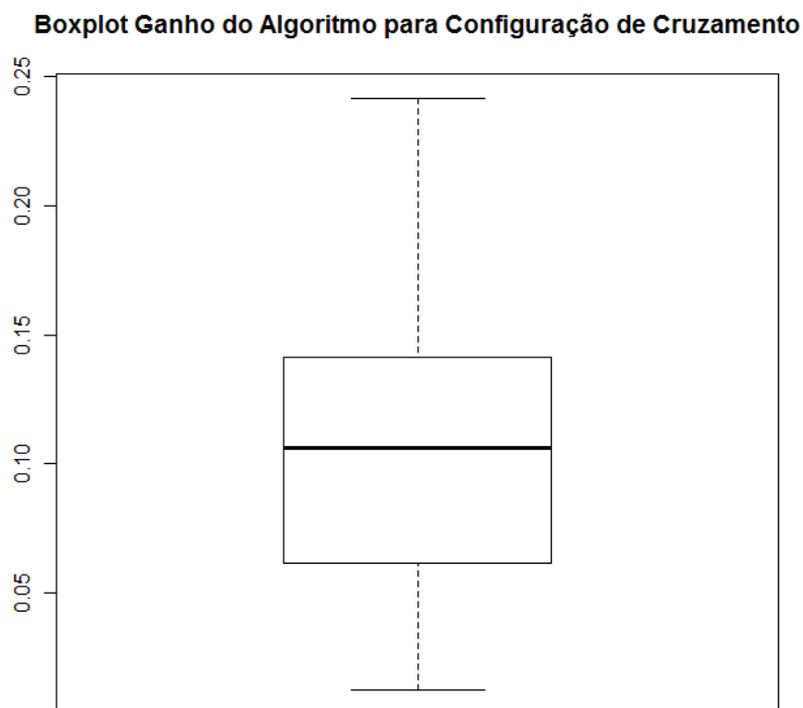
Na Figura A.5, apresenta-se com mais detalhes o *boxplot* da variável GACC para a configuração de cruzamento.

Tabela A.14 Coeficiente de correlação do Ganho do Algoritmo e a configuração de cruzamento

	Ganho do Algoritmo	Configuração de Cruzamento
Ganho do Algoritmo	1,000	-0,086
Configuração de Cruzamento	-0,086	1,000

Utilizando-se da Figura A.5 e da Tabela A.14 concluiu-se que a variável GACC possuía distribuição levemente assimétrica à direita, pois a linha mediana estava um pouco deslocada para o extremo superior da caixa do *boxplot*. A maior concentração dos dados estava entre (0,106) e (0,141) e a distância interquartil indica que os dados possuem tendência moderada a serem homogêneos e não foi encontrado *outlier*.

Figura A.5 Boxplot da variável ganho do algoritmo para a configuração de cruzamento



Nesta avaliação existiam 2 (duas) variáveis, sendo a variável GACC do tipo quantitativo contínuo, pois pode-se utilizar qualquer valor real entre 0 (zero) e 1 (um) na sua representação. A variável “Configuração de Cruzamento” é do tipo qualitativo nominal, formada pela combinação dos métodos de cruzamento (Cruzamento de 1 ponto, Cruzamento de 2 Pontos e Cruzamento Uniforme) e pela probabilidade de cruzamento (0,600; 0,800 e 1,000), gerando-se 9 (nove) grupos distintos (1 Ponto/0,600, 2 Ponto/0,600, Uniforme/0,600 1 Ponto/0,800, 2 Ponto/0,800, Uniforme/0,800 1 Ponto/1,000, 2 Ponto/1,000, Uniforme/1,000). Logo, têm-se as seguintes variáveis:

- W: variável de trabalho do GACC (em porcentagem);
- Y: variável explicativa da Configuração de Cruzamento.

Os mesmos testes estatísticos utilizado no estudo sobre o tamanho da população foram novamente utilizados. Na Tabela A.15 são apresentados os resultados do teste de normalidade da distribuição e na Tabela A.16 o teste de homogeneidade da variância, aplicados aos dados do GACC.

O resultado apresentado na Tabela A.15 evidenciou que a distribuição da variável GACC não se assemelha a uma distribuição normal, pois o teste *Shapiro-*

Wilk rejeitou a hipótese de normalidade com um nível de confiança de 95,0%. O resultado na Tabela A.16 do teste de *Bartlett*, para verificar a homogeneidade da variância das médias de ganho do algoritmo de cada grupo, revelou que as variâncias não são homogêneas. Logo, não foi possível aplicar os testes paramétricos, devido aos dados violarem as duas pré-condições de uso. Portanto, utilizou-se o teste não paramétrico de *Kruskal-Wallis*, para verificar se existiam diferenças entre as médias de ganho do algoritmo em relação à configuração de cruzamento na geração dos indivíduos da população.

Tabela A.15 Resultado do teste de normalidade da distribuição para configuração de cruzamento

Shapiro-Wilk	Ganho do Algoritmo
W_{calc}	0,974
W_{α}	0,927
<i>p-valor</i>	$3,076 \times 10^{-4}$
α	0,050

Tabela A.16 Resultado do teste de homogeneidade da variância para a configuração de cruzamento

Bartlett	Ganho do Algoritmo
<i>K-squared</i>	34,161
<i>df</i>	8,000
<i>p-valor</i>	$3,798 \times 10^{-5}$

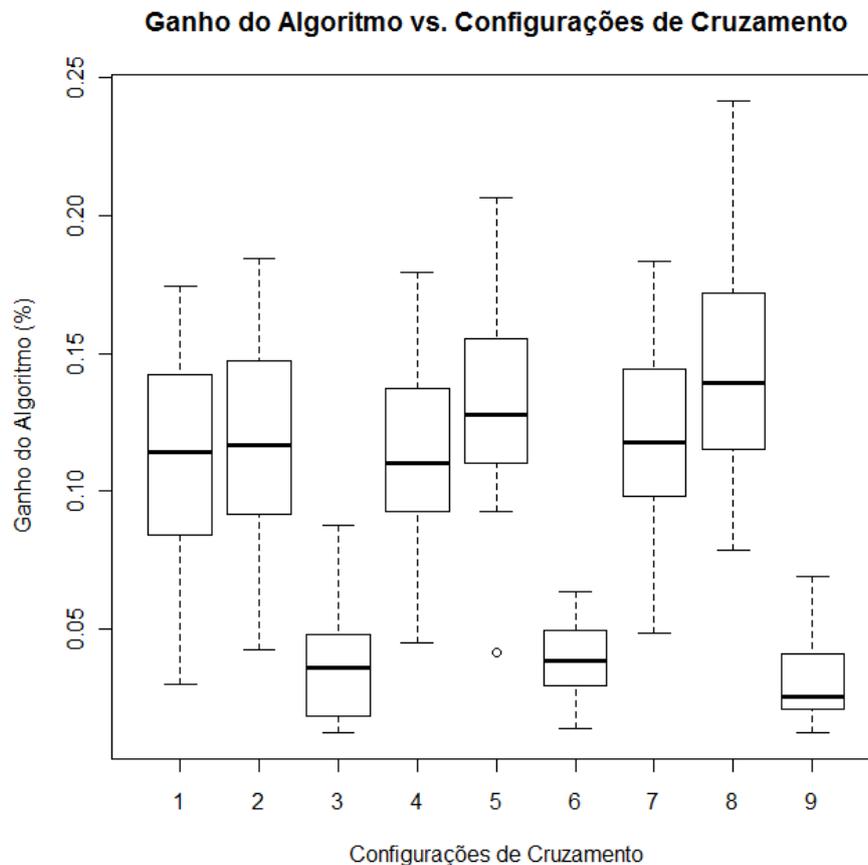
Neste experimento, buscou-se comparar as médias de ganho do algoritmo de cada um dos 9 (nove) grupos, que foram diferenciados pela configuração de cruzamento adotada nos testes de cada grupo. Como havia mais de 2 (dois) grupos a serem comparados, foi escolhido o teste *Kruskal-Wallis*. Na Tabela A.17 apresenta-se o resultado do teste *Kruskal-Wallis*.

Tabela A.17 Resultado do teste para verificar se existem diferenças entre as médias dos grupos para a configuração de cruzamento

Kruskal-Wallis	Valores
<i>chi-square</i>	122,099
<i>DF</i>	8,000
<i>p-valor</i>	$2,200 \times 10^{-16}$
α	0,050

Verifica-se na Tabela A.17 que o resultado do p -valor do teste *Kruskal-Wallis* proporcionou um valor menor que o α . Portanto, pode-se afirmar, com um nível de confiança de 95,0%, que há diferença significativa entre as médias do ganho do algoritmo em relação às configurações de cruzamento. Na Figura A.6 apresenta-se o gráfico do boxplot do ganho do algoritmo em função das configurações de cruzamento.

Figura A.6 Boxplot do ganho do algoritmo em função das configurações de cruzamento



Do *boxplot* da Figura A.6, percebe-se que a mediana do grupo 8 (oito), cruzamento de 2 (dois) pontos, com probabilidade de (1,000) é o que apresenta melhor desempenho em relação aos demais grupos. Logo, baseando-se na análise gráfica, escolheu-se o método de cruzamento de 2 (dois) pontos e a probabilidade de cruzamento de (1,000).

A.4 Parâmetro Probabilidade de Mutação

Para o estudo estatístico da probabilidade de mutação foi realizada a mesma análise descritiva utilizada no estudo do tamanho da população. Na Tabela A.18 e na Tabela A.19 são apresentados os resultados obtidos pelo *software R*.

Tabela A.18 Análise descritiva do ganho do algoritmo para a probabilidade de mutação

Parâmetros	Ganho do Algoritmo
Mínimo	0,057
1º Quartil	0,113
Mediana	0,138
Média	0,140
3º Quartil	0,164
Máximo	0,249
Desvio Padrão	0,040
Variância	0,002
Coefficiente de Variação	28,777

Tabela A.19 Coeficiente de correlação do ganho do algoritmo e a probabilidade de mutação

	Ganho do Algoritmo	Probabilidade de Mutação
Ganho do Algoritmo	1,000	-0,058
Probabilidade de Mutação	-0,058	1,000

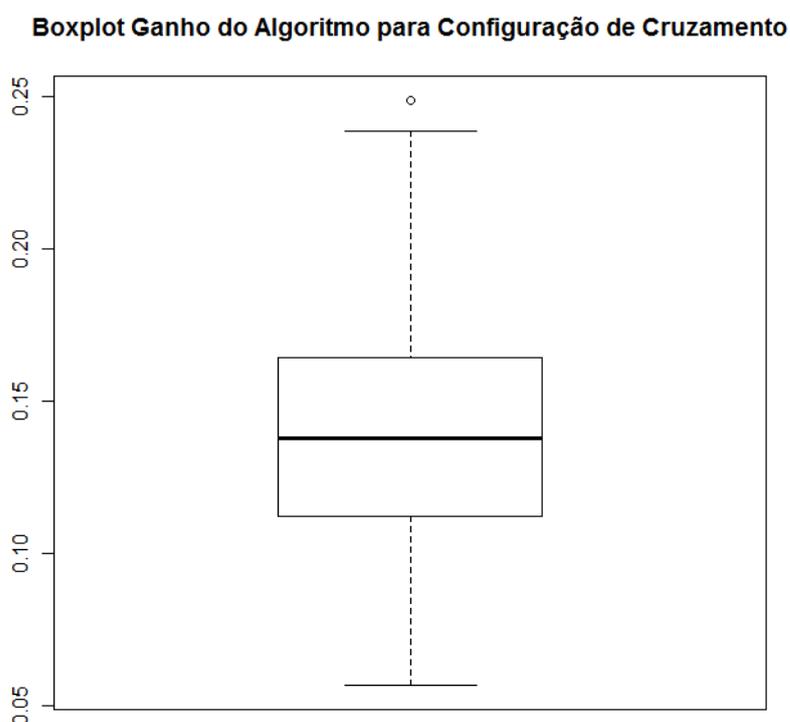
Analisando-se o resultado do teste do coeficiente de variação, que ficou abaixo de 50,0% para a variável Ganho do Algoritmo para a Probabilidade de Mutação (GAPM), pôde-se concluir que as médias são representativas. Outra informação importante que se pôde retirar dos resultados da Tabela A.19 é o coeficiente de correlação. Nos testes, o coeficiente de correlação ficou em (-0,058) demonstrando-se estatisticamente a não existência de uma relação linear entre as 2 (duas) variáveis em estudo.

Da Figura A.7 e da Tabela A.18 concluiu-se que a variável GAPM possuía distribuição simétrica, pois a linha mediana estava equidistante dos extremos da caixa do boxplot, a maior concentração dos dados estava entre (0,113) e (0,164). A

distância interquartil indica que os dados possuem tendência moderada a serem homogêneos e 1 (uma) das medidas é um possível *outlier*.

Na Figura A.7 apresenta-se em mais detalhes o *boxplot* da variável GAPM.

Figura A.7 Boxplot da variável ganho do algoritmo para a probabilidade de mutação



Nesta avaliação existiam 2 (duas) variáveis, sendo a variável “Ganho do Algoritmo” do tipo quantitativo contínua, pois pode-se utilizar qualquer valor real entre 0 (zero) e 1 (um) na sua representação. A variável “Probabilidade de Mutação” é do tipo qualitativo ordinal, formada por 4 (quatro) grupos criados a partir dos valores de mutação (0,005, 0,010, 0,020 e 0,030). Logo, têm-se as seguintes variáveis:

- W: variável de trabalho do GAPM (em porcentagem);
- Y: variável explicativa da Probabilidade de Mutação.

Os mesmos testes estatísticos utilizado no estudo sobre o tamanho da população foram novamente utilizados. Na Tabela A.20 são apresentados os resultados do teste de normalidade da distribuição e na Tabela A.21 o teste de homogeneidade da variância, aplicados aos dados do “Ganho do Algoritmo”.

Tabela A.20 Resultado do teste de normalidade da distribuição para a probabilidade de mutação

Shapiro-Wilk	Ganho do Algoritmo
W_{calc}	0,991
W_{α}	0,927
<i>p-valor</i>	0,656
α	0,050

Tabela A.21 Resultado do teste de homogeneidade da variância para a probabilidade de mutação

Bartlett	Ganho do Algoritmo
<i>K-squared</i>	1,424
<i>df</i>	3,000
<i>p-valor</i>	0,700

O resultado apresentado na Tabela A.20 confirmou que os dados extraídos do experimento têm uma distribuição normal. O resultado da Tabela A.21 confirmou a homogeneidade das variâncias. Portanto, foi possível aplicar os testes paramétricos para verificar se existem diferenças entre as médias de ganho do algoritmo em relação à probabilidade de mutação.

Neste experimento, buscava-se comparar as médias de ganho do algoritmo de cada um dos 4 (quatro) grupos, que são diferenciados pelo valor da probabilidade de mutação adotado nos teste de cada grupo. Como se tem mais de 2 (dois) grupos a serem comparados, foi escolhido o teste ANOVA de uma via (*Oneway ANOVA*). O resultado do teste *Oneway ANOVA* é exposto na Tabela A.22.

Tabela A.22 Resultado do teste para verificar diferenças entre as médias do ganho do algoritmo de cada valor de probabilidade de mutação

Oneway ANOVA	Valores
<i>F</i>	1,700
<i>Num df</i>	3,000
<i>Denom df</i>	64,250
<i>p-valor</i>	0,176
α	0,050

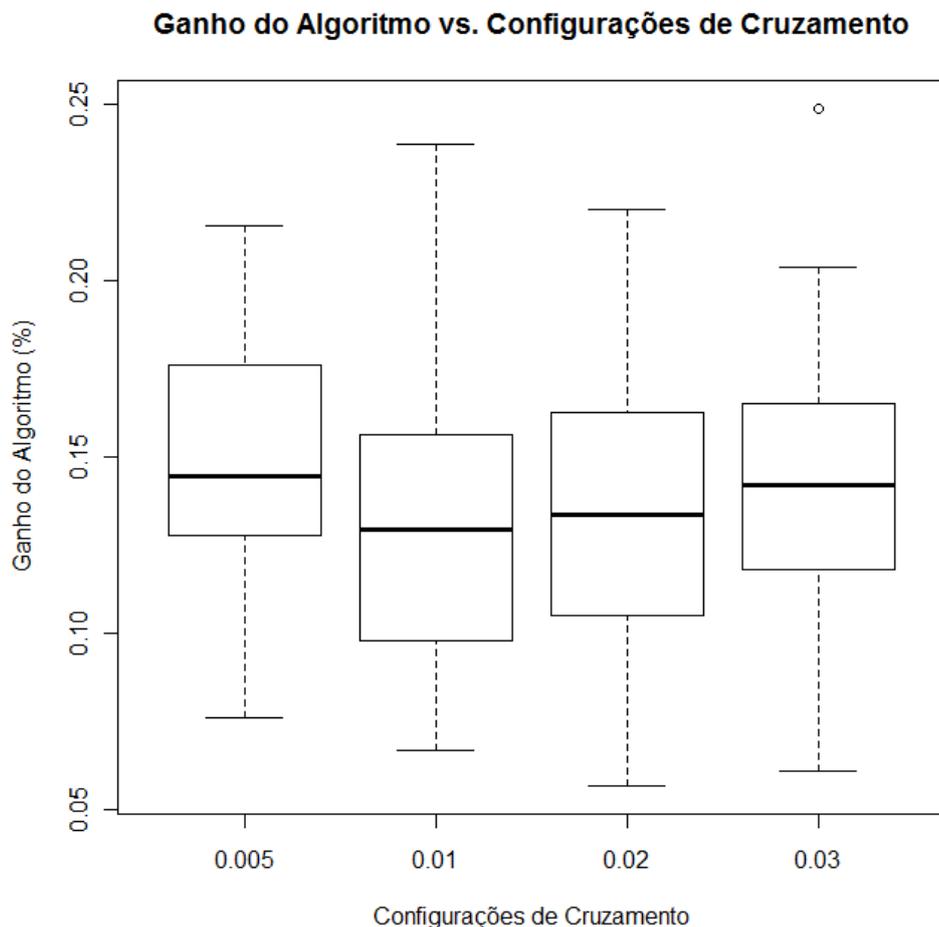
Verificando-se os resultados da Tabela A.22, pôde-se afirmar, com um nível de confiança de 95,0%, que não houve diferença significativa entre as médias do

GAPM utilizado no algoritmo genético. Para determinar o melhor valor para o parâmetro de probabilidade de mutação para este trabalho, foram utilizados os resultados dos testes *t de Student* e a análise gráfica do *boxplot* do ganho do algoritmo para cada um dos valores de probabilidade de mutação. Na Tabela A.23 e na Figura A.8 são apresentados os resultados dos testes realizados.

Tabela A.23 Ganho do algoritmo em função dos valores da probabilidade de mutação

Probabilidade de Mutação	IC	Média
0,005	0,139	0,165
0,010	0,116	0,146
0,020	0,120	0,150
0.030	0,124	0,157

Figura A.8 Boxplot do ganho do algoritmo em função dos valores da probabilidade de mutação



Do *boxplot* na Figura A.8 verificou-se que a mediana da probabilidade de mutação (0,005) foi um pouco melhor que os demais valores de probabilidade de mutação. A mesma probabilidade de mutação também apresentou resultado um pouco melhor nos testes *t de Student*, como pode ser verificado na Tabela A.23. Portanto, baseado na análise gráfica e nos resultados dos testes, escolheu-se a probabilidade de mutação de (0,005) para ser usada neste trabalho.

A.5 Energia Residual

A Hipótese A levantada neste trabalho tem o objetivo de verificar a influência dos obstáculos no consumo de energia da RSSF. Para testar esta hipótese, foi realizado um experimento. Neste experimento, o algoritmo foi configurado para verificar apenas a energia residual, desconsiderando as demais métricas. Foram utilizados os parâmetros disponibilizados no Quadro 6.4 e no Quadro 6.3.

Hipótese A (H_0): A presença de obstáculos nas simulações afeta o consumo de energia da rede, reduzindo o somatório da energia residual dos nós sensores de uma RSSF.

- H_0 : A presença de obstáculos no cenário de simulação não influencia no somatório da energia residual dos nós sensores. Formalmente:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- μ_0 é a média do ganho do algoritmo sem considerar os obstáculos;
- μ_1 é a média do ganho do algoritmo considerando os obstáculos.

O teste estatístico mais apropriado para comparar o somatório da energia residual dos nós sensores da rede é o teste de comparação de 2 (duas) amostras pareadas, teste *t de Student*, mas o seu uso só pode ser aplicado se os dados obedecerem a 2 (duas) condições:

- Os dados precisam ter distribuição normal (teste de *Shapiro-Wilk*);

- Os dados precisam apresentar valores de variância semelhantes (homogeneidade) para os 2 (dois) grupos de dados extraídos do experimento (teste de *Bartlett*).

São apresentados, na Tabela A.24, os resultados do teste de normalidade dos dados. Na Tabela A.25, os resultados do teste de homogeneidade da variância dos dados. Foram realizadas 16 repetições para cada um dos experimentos.

Tabela A.24 Resultado do teste de normalidade da distribuição do somatório da energia residual

Shapiro-Wilk	Resultados
W_{calc}	0,977
W_{α}	0,887
<i>p-valor</i>	0,697
α	0,050

Os resultados apresentados na Tabela A.24 mostram que o valor de W_{calc} é maior que W_{α} e que *p-valor* é maior que α , confirmando que os dados extraídos do experimento têm uma distribuição normal (SHAPIRO e WILK, 1965). O resultado da Tabela A.25 confirma a homogeneidade das variâncias. Portanto, foi possível aplicar os testes paramétricos para verificar se existem diferenças entre as médias da energia residual dos nós, em relação à utilização ou não de obstáculos nos cenários de simulação.

Tabela A.25 Resultado do teste de homogeneidade da variância do somatório da energia residual

Bartlett	Resultados
<i>K-squared</i>	0,445
<i>df</i>	1,000
<i>p-valor</i>	0,505

Na Tabela A.26, são apresentados os resultados do teste *t* de *Student*, utilizado para verificar se existem diferenças significativas nos somatórios da energia residual dos nós sensores em relação à presença ou não de obstáculos no cenário de simulação.

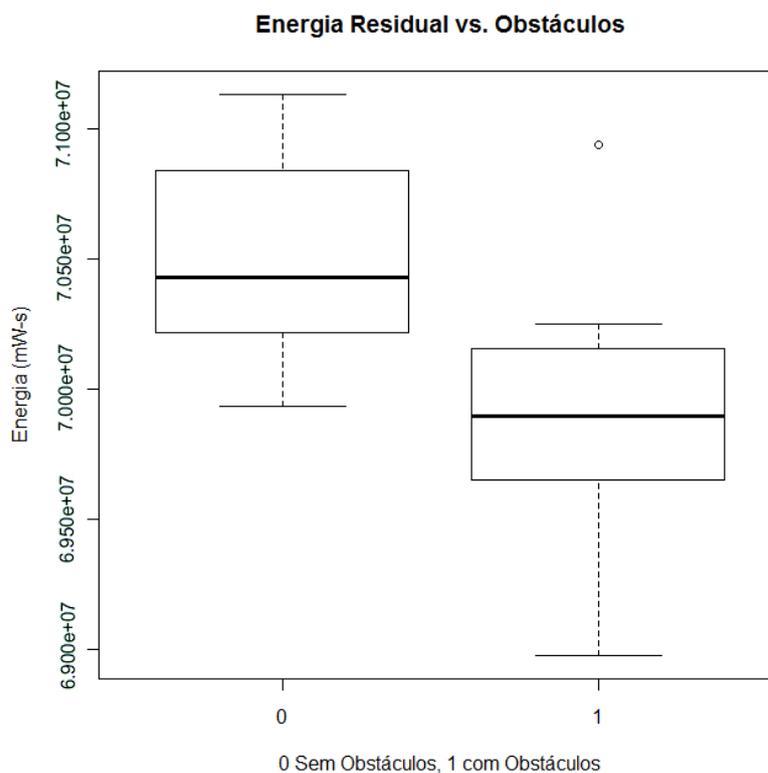
Tabela A.26 Resultado do teste para verificar se existe diferenças entre o somatório da energia residual em relação a presença de obstáculos no cenário

Teste t	Resultados
<i>t</i>	-4,901
<i>df</i>	15,000
<i>p-valor</i>	$1,921 \times 10^{-4}$
α	0,050

O resultado do *p-valor* do teste de *t* de *Student*, como pode ser verificado na Tabela A.26, obteve um valor muito menor que o α . Portanto, é possível afirmar, com um nível de confiança de 95%, que há diferença significativa entre as médias do somatório da energia residual dos nós sensores quando o módulo de obstáculos está ativo, quando comparado com os resultados do módulo desativado.

Outros instrumentos estatísticos utilizados para demonstrar a diferença das médias entre esses 2 (dois) grupos é a utilização do gráfico *boxplot*, apresentado na Figura A.9, e os testes de IC, que têm seus resultados apresentados na Tabela A.27.

Figura A.9 Boxplot do somatório da energia residual em relação a presença de obstáculos no cenário



Ao se observar o *boxplot* da Figura A.9, verifica-se que a mediana do somatório da energia residual sem obstáculos apresenta um nível de energia residual maior ($7,043 \times 10^7$ mW-s) que o do nível com obstáculos ($6,990 \times 10^7$ mW-s). Vale ressaltar que, quanto maior for o nível de energia residual, melhor a qualidade da rede. Para complementar e reforçar a conclusão, baseando-se no *boxplot*, aplicou-se o teste para determinar a média e o IC da energia residual e os resultados são apresentados na Tabela A.27.

Tabela A.27 Média do somatório da energia residual

Obstáculo	Média (mW-s)	IC (mW-s)	
0	$7,053 \times 10^7$	$7,032 \times 10^7$	$7,073 \times 10^7$
1	$6,988 \times 10^7$	$6,964 \times 10^7$	$7,013 \times 10^7$

Nos resultados fornecidos pelos testes estatísticos para a hipótese A, obteve-se o *p-valor* $< \alpha$. Logo, pode-se rejeitar, com um nível de confiança de 95,0%, a hipótese H_0 . Portanto, é possível afirmar que há diferença significativa entre os somatórios da energia residual dos nós em relação à presença ou não de obstáculos no cenário de simulação.

A.6 Tempo de Atraso Fim-a-Fim

A Hipótese B tem como objetivo verificar a influência dos obstáculos sobre os resultados do TAFF. Para verificar a Hipótese B, o algoritmo foi configurado com os valores semelhantes aos utilizados na verificação da energia residual (Hipótese A), com exceção dos valores dos parâmetros β e γ . A configuração completa pode ser consultada no Quadro 6.3 e no Quadro 6.4.

Hipótese B (H_0): a presença de obstáculos nas simulações afeta a média do TAFF dos pacotes entre os nós sensores e o nós sorvedouro.

- H_0 : A média dos TAFF dos pacotes na rede é igual para as simulações com e sem obstáculos no cenário. Formalmente:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- μ_0 é a média do ganho do algoritmo sem considerar os obstáculos;
- μ_1 é a média do ganho do algoritmo considerando os obstáculos.

Foram realizados os mesmos testes aplicados à hipótese A, para verificar a normalidade da distribuição e sua homogeneidade. Na Tabela A.28 e na Tabela A.29, são apresentados os resultados do teste de homogeneidade e da variância dos dados, respectivamente, para os quais foram realizadas 16 repetições do experimento.

Tabela A.28 Resultado do teste de normalidade da distribuição do TAFF

Shapiro-Wilk	Resultados
W_{calc}	0,913
W_α	0,881
$p\text{-valor}$	0,018
α	0,050

Tabela A.29 Resultado do teste de homogeneidade da variância do TAFF

Bartlett	Resultados
$K\text{-squared}$	13,027
df	1,000
$p\text{-valor}$	$3,070 \times 10^{-4}$

A partir dos resultados apresentados nas Tabela A.28, é possível afirmar que os dados do TAFF não seguem uma distribuição normal, pois o valor do $p\text{-valor}$ foi de (0,018), indo de encontro à segunda condição do teste *Shapiro-Wilk* para a normalidade. O segundo problema ocorrido com os dados do TAFF foi a apresentação de características de não homogeneidade da variância dos dados, como pode ser observado na Tabela A.29.

Devido ao fato de os dados de TAFF não apresentarem um distribuição normal, que é a primeira condição para a aplicação dos testes estatísticos paramétricos, e também não possuem homogeneidade das variâncias, foram

realizadas transformações nos dados de TAFF, com o objetivo de normalizá-los. Na Tabela A.30 e na Tabela A.31, são apresentados os resultados dos testes de normalidade (*Shapiro-Wilk*) e homogeneidade de variância (*Bartlett*) para a transformação logarítmica dos dados de TAFF.

Tabela A.30 Resultado do teste de normalidade da distribuição do logaritmo do TAFF

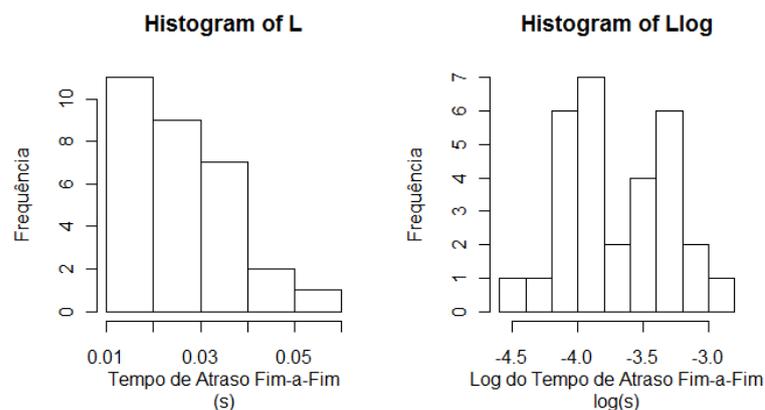
Shapiro-Wilk	log(Atraso)
W_{calc}	0,973
W_{α}	0,881
<i>p-value</i>	0,633
α	0,050

Tabela A.31 Resultado do teste de homogeneidade da variância do logaritmo do TAFF

Bartlett	log(Atraso)
<i>K-squared</i>	1,158
<i>df</i>	1,000
<i>p-value</i>	0,282

Com a transformação aplicada aos dados de TAFF, obteve-se uma melhoria considerável nos testes de normalidade e homogeneidade da variância, conforme pode ser observado nos resultados apresentados na Tabela A.30 e na Tabela A.31, satisfazendo as condições do teste de *Shapiro-Wilk* (SHAPIRO e WILK, 1965). Essa melhoria também pode ser confirmada observando-se os histogramas dos dados originais e os modificados pela transformação logarítmica, apresentados na Figura A.10.

Figura A.10 Histogramas dos dados originais e do logaritmo do TAFF



Assim, foi possível aplicar os testes paramétricos, para verificar a existência de diferenças no TAFF entre os testes realizados com a presença ou não de obstáculos nos cenários de simulação. O teste estatístico *t de Student* é o mais apropriado para verificar a Hipótese B, uma vez que se têm apenas 2 (dois) grupos a serem comparados: o grupo com obstáculos e o grupo sem obstáculos. Os resultados dos testes encontram-se na Tabela A.32.

Tabela A.32 Resultado do teste para verificar se existe diferenças entre as médias do TAFF

Teste t	Resultados
<i>T</i>	8,644
<i>Df</i>	14,000
<i>p-valor</i>	$5,496 \times 10^{-7}$
α	0,050

Observando a Tabela A.32, é possível verificar que o resultado do *p-valor* do teste *t de Student* foi menor que o do α . Logo, é possível afirmar, com um nível de confiança de 95,0%, que há diferença significativa nas médias do TAFF entre os grupos que utilizam obstáculos e os que não utilizam obstáculos nos cenários de simulação. Novamente, foi utilizado o gráfico e os testes de IC, representados pela Figura A.11 e pela Tabela A.33, respectivamente.

Figura A.11 Boxplot da média dos TAFF em relação aos obstáculos no cenário

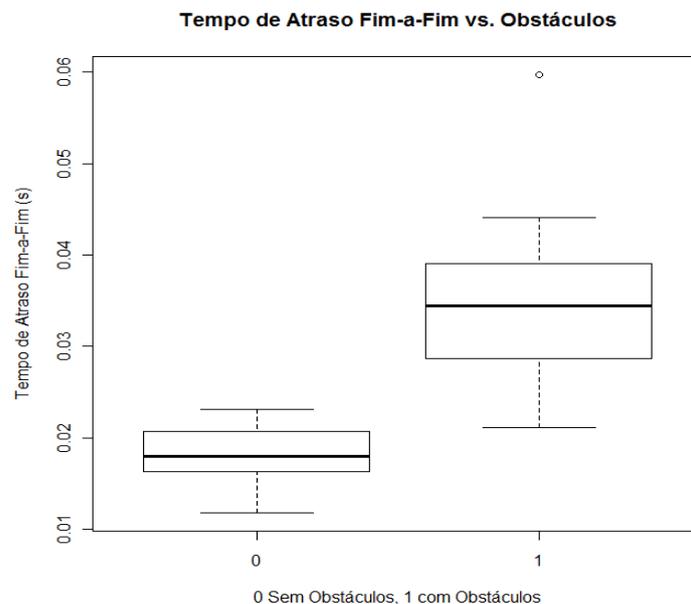


Tabela A.33 Média do TAFF

Obstáculo	Média (s)	IC (log(s))	
0	0,018	0,016	0,020
1	0,035	0,030	0,040

Ao se observar o *boxplot* da Figura A.11, é possível verificar que a mediana do TAFF do grupo sem obstáculos obteve o tempo de (0,018) segundos, menor que o grupo com obstáculos, que foi de (0,034) segundos. Especificamente nesta métrica, o objetivo do algoritmo é encontrar a configuração que apresente o menor TAFF. O segundo estudo realizado foi utilizando os testes para determinar o IC dos 2 (dois) grupos. Neste teste, também se verificou que a média do grupo sem obstáculos, que é de (0,018) segundos, é melhor que a média do grupo com obstáculos, que foi de (0,035) segundos. Os resultados para o IC dos grupos são apresentados na Tabela A.33.

Assim, após a análise dos resultados dos testes estatísticos (Tabela A.32), (Tabela A.33) e do *boxplot* (Figura A.11), pode-se rejeitar, com o nível de significância de 5,0%, a hipótese B de que os dados do TAFF apresentam a mesma média quando aplicados a cenários com ou sem a presença de obstáculos.

A.7 Taxa de Perda de Pacotes

Na verificação da influência dos obstáculos sobre a TPP, foi formulada a Hipótese C e, novamente, utilizou-se um experimento semelhante ao da Hipótese A, modificando-se os parâmetros β e δ para configurar o algoritmo de busca da melhor TPP. Os valores dos parâmetros de configuração estão disponíveis no Quadro 6.4 e no Quadro 6.3.

Hipótese C (H_0): A presença de obstáculos nas simulações afeta a TPP da rede de sensores sem fio.

- H_0 : A presença de obstáculos no cenário de simulação não influencia na média da TPP da rede.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Em que:

- μ_0 é a média do ganho do algoritmo sem considerar os obstáculos;
- μ_1 é a média do ganho do algoritmo considerando os obstáculos.

Foram realizados os testes para verificação da normalidade da distribuição e da sua homogeneidade. Os resultados são apresentados na Tabela A.34 e na Tabela A.35, respectivamente. Neste experimento, foi necessária a realização de 19 repetições.

Tabela A.34 Resultado do teste de normalidade da distribuição da TPP

Shapiro-Wilk	Resultados
W_{calc}	0,852
W_{α}	0,901
$p\text{-valor}$	1,419e-04
α	0,050

A partir dos resultados apresentados na Tabela A.34, é possível afirmar que os dados da TPP não seguem uma distribuição normal, pois as 2 (duas) condições do teste *Shapiro-Wilk*, $W_{calc} > W_{\alpha}$ e $p\text{-valor} > \alpha$ (SHAPIRO e WILK, 1965), foram violadas. Quanto à segunda premissa para o uso dos testes paramétricos, esta também não foi satisfeita, uma vez que o $p\text{-value}$ do teste de homocedasticidade (teste de *Bartlett*) foi maior que α , como pode ser observado na Tabela A.35.

Tabela A.35 Resultado do teste de homogeneidade da variância da TPP

Bartlett	Resultados
$K\text{-squared}$	31,659
df	1,000
$p\text{-valor}$	1,837e-08

Os dados da TPP não apresentaram uma distribuição normal. Não satisfazendo, assim, a primeira condição para a aplicação dos testes estatísticos paramétricos. Além disso, os dados também não apresentam homocedasticidade, o que não satisfaz a segunda condição. Para tentar normalizar os dados, foram aplicadas algumas transformações. Na Tabela A.36 e na Tabela A.37, são apresentados os resultados dos testes de normalidade (*Shapiro-Wilk*) e

homogeneidade de variância (*Bartlett*), para as transformações logarítmicas, raiz quadrada, arco seno e de *Box-Cox* dos dados da TPP.

Tabela A.36 Resultado do teste de normalidade das transformações nos dados da TPP

<i>Shapiro-Wilk</i>	$\log(TPP)$	$\sqrt{(TPP)}$	$\arcsin(TPP)$	$\frac{(TPP)^\lambda - 1}{\lambda}$
W_{calc}	0,867	0,862	0,852	0,869
W_α	0,881	0,881	0,881	0,881
<i>p-value</i>	3,407e-04	2,472e-04	1,426e-04	3,759e-04
α	0,050	0,050	0,050	0,050

Os resultados dos testes de *Shapiro-Wilk* aplicados às transformações dos dados da TPP, apresentados na Tabela A.36, mostraram que nenhuma das transformações aplicadas aos dados foi suficiente para tornar a sua distribuição normal. Os testes de homocedasticidade de *Bartlett*, conforme a Tabela A.37, apresentaram resultados positivos para a transformação logarítmica e *Box-Cox*, e negativas para as transformações raiz quadrada e arco seno.

Tabela A.37 Resultado do teste de homogeneidade da variância do logaritmo da TPP

<i>Bartlett</i>	$\log(TPP)$	$\sqrt{(TPP)}$	$\arcsin(TPP)$	$\frac{(TPP)^\lambda - 1}{\lambda}$
<i>K-squared</i>	2,707	14,501	34,293	0,040
<i>df</i>	1,000	1,000	1,000	1,000
<i>p-value</i>	0,100	1,401e-04	4,741e-04	0,841

Apesar dos resultados do teste de homocedasticidade para as transformações logarítmicas e *Box-Cox* terem sido satisfatórias, não foi possível a aplicação dos testes estatísticos paramétricos. Portanto, para verificar a hipótese C se fez necessária a utilização dos testes não paramétricos, *Wilcoxon*.

Na Tabela A.38, encontram-se os resultados do teste *Wilcoxon* para verificar se existem diferenças significativas nas médias da TPP em relação à presença ou não de obstáculos no cenário de simulação.

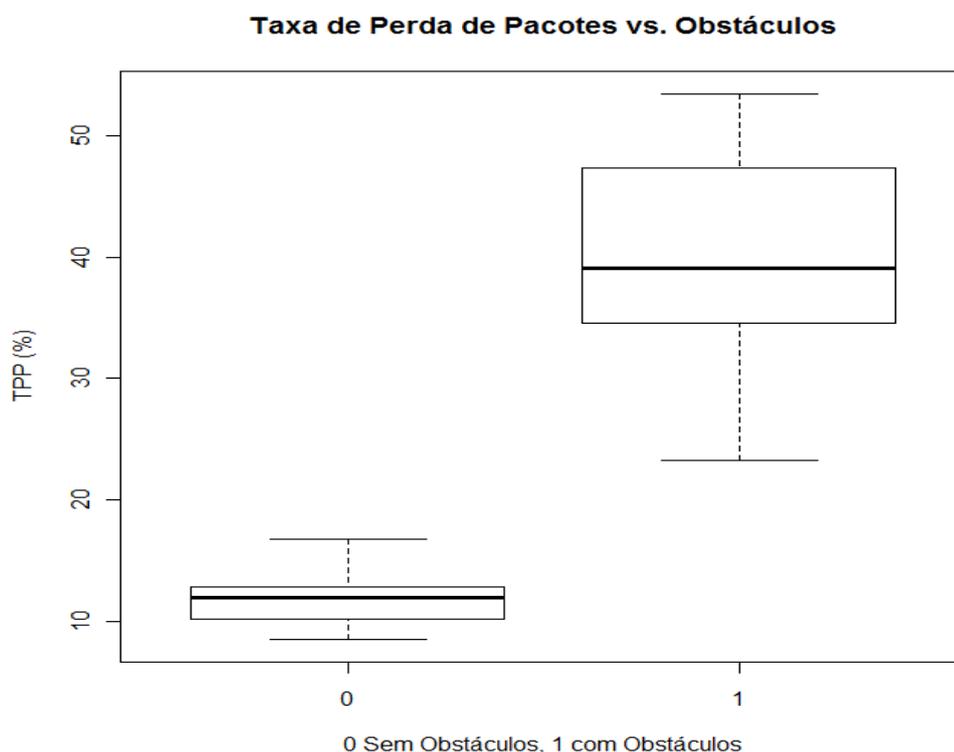
Tabela A.38 Resultado do teste para verificar se existe diferenças entre as médias da TPP

<i>Wilcoxon</i>	Valores
<i>V</i>	0,000
<i>p-valor</i>	3,815e-06
α	0,050

O resultado do *p-valor* do teste de *Wilcoxon*, como pode ser verificado na Tabela A.38, obteve um valor muito menor que o α . Portanto, pode-se afirmar, com um nível de confiança de 95%, que há diferença significativa entre as médias da TPP dos grupos com e sem obstáculos.

Definido que existe diferença entre as médias da TPP, foram realizados mais testes com o objetivo de determinar qual grupo apresenta maior TPP. O teste não paramétrico utilizado foi uma variação do teste *Wilcoxon* e o resultado está exposto na Tabela A.39. Outro instrumento estatístico utilizado para demonstrar a diferença das médias entre esses 2 (dois) grupos foi o gráfico *boxplot*, que está ilustrado pela Figura A.12.

Figura A.12 Boxplot da média da TPP em relação aos obstáculos



Observando-se o *boxplot* da Figura A.12, é possível verificar que a mediana da TPP do grupo sem obstáculos (12,0%), apresenta melhor desempenho em relação ao grupo com obstáculos (39,1%). É importante lembrar que, quanto menor for à TPP, melhor a qualidade da rede. O mesmo ocorre nos resultados apresentados na Tabela A.39 do teste *Wilcoxon*, que apontam que a média da TPP

para o grupo que utiliza obstáculos nos cenários é maior que a média da TPP do grupo sem obstáculos.

Tabela A.39 Teste de Wilcoxon para determinar qual o grupo, com ou sem utilização de obstáculos no cenário, apresenta maior TPP

<i>Wilcoxon (Maior)</i>	Sem Obstáculos > Com Obstáculos	Com Obstáculos > Sem Obstáculos
V	0,000	190,000
<i>p-valor</i>	1,000	1,907e-06
α	0,050	0,050

Pelos resultados obtidos nos testes estatísticos e na análise do *boxplot*, pode-se rejeitar, com um nível de confiança de 95%, a hipótese H_0 de que as médias da TPP são iguais para os testes realizados com e sem presença de obstáculos no cenário.

A.8 Ganho do Algoritmo sobre a QoS

Neste experimento, o objetivo foi verificar o ganho na QoS após a busca da melhor configuração de implantação realizada pelo modelo proposto. Foram verificados ganhos para as métricas de TAFF e da TPP, com e sem obstáculos, no cenário de simulação. A configuração dos parâmetros do AG utilizada neste experimento foi retirada do Quadro 6.4 e do Quadro 6.3. Foram realizados 10 repetições do experimento para aumentar a representatividade estatística das médias do TAFF e da TPP.

O ganho na métrica do TAFF, proporcionado pelo algoritmo neste experimento, caracteriza-se pelo módulo da diferença entre o melhor TAFF obtido na primeira geração e o melhor TAFF conseguido na última geração. Este ganho foi obtido pelo algoritmo proposto dividido pelo primeiro melhor TAFF. Apresenta-se, na Equação (A.1), o ganho do TAFF.

$$G_{TAF\!F} = \frac{t_{primeira} - t_{ultima}}{t_{primeira}} \quad (A.1)$$

Do mesmo modo, o ganho obtido na métrica da TPP pelo algoritmo caracteriza-se pelo módulo da diferença entre a melhor TPP da primeira geração e a

melhor TPP da última geração, dividido pela primeira TPP. Apresenta-se, na Equação (A.2), o ganho da TPP.

$$G_{TPP} = \frac{t_{primeira} - t_{Ultima}}{t_{primeira}} \quad (A.2)$$

Para que seja possível a aplicação dos testes estatísticos paramétricos, é necessário que a distribuição dos dados satisfaça as condições de normalidade e homocedasticidade. Portanto, os resultados extraídos do experimento foram submetidos ao teste *Shapiro-Wilk*, para verificar a normalidade da distribuição, e ao teste de *Bartlett*, para verificar a homocedasticidade dos dados.

Na Tabela A.40, são apresentados os resultados dos testes de normalidade das 2 (duas) distribuições dos dados das médias de ganho no TAFF. Na primeira distribuição, os dados foram extraídos do experimento realizado com o módulo de obstáculos desabilitado, e na segunda distribuição, com o módulo de obstáculos habilitado. Testes semelhantes foram realizados com os dados do ganho na TPP e os resultados são apresentados na Tabela A.41.

Tabela A.40 Resultado do teste de normalidade da distribuição com e sem obstáculos para o experimento do ganho no TAFF

Shapiro-Wilk	TAFF sem Obstáculos	TAFF com Obstáculos
W_{calc}	0,938	0,956
W_{α}	0,842	0,842
<i>p-valor</i>	0,526	0,737
α	0,050	0,050

Tabela A.41 Resultado do teste de normalidade da distribuição com e sem obstáculos para o experimento do ganho na TPP

Shapiro-Wilk	TPP sem Obstáculos	TPP com Obstáculos
W_{calc}	0,940	0,933
W_{α}	0,842	0,842
<i>p-valor</i>	0,557	0,475
α	0,050	0,050

Na Tabela A.42, é apresentado o resultado do teste de homocedasticidade dos dados para os 2 (dois) experimentos, em relação à utilização ou não de

obstáculos nos cenários de simulação. No teste de *Bartlett*, os dados dos 2 (dois) grupos são verificados simultaneamente.

Tabela A.42 Resultado do teste de homocedasticidade da variância da TPP para o experimento do ganho no TAFF e do ganho na TPP

Bartlett	TAFF	TPP
<i>K-squared</i>	0,145	2,261
<i>df</i>	1,000	1,000
<i>p-valor</i>	0,703	0,133
α	0,050	0,050

Os resultados dos testes de normalidade (*Shapiro-Wilk*) confirmam que o TAFF e a TPP apresentam distribuição normal nos 2 (dois) grupos (com e sem obstáculos). E os testes de homocedasticidade (*Bartlett*), também confirmam a homogeneidade da variância dos dados para os 2 (dois) experimentos. Portanto, foi possível aplicar os testes do tipo paramétrico para encontrar a média dos ganhos de TAFF e TPP.

Utilizou-se o IC para estimar o ganho médio do TAFF e da TPP. Os resultados estão descritos na Tabela A.43 e na Tabela A.44, respectivamente.

Tabela A.43 Média do ganho do TAFF

Obstáculo	Média	IC
0	0,506	0,328
1	0,587	0,430

Tabela A.44 Média do ganho da TPP

Obstáculo	Média	IC
0	0,439	0,340
1	0,507	0,448

Observa-se que o ganho médio proporcionado pelo algoritmo proposto neste trabalho é de, aproximadamente, 50,0%, ou seja, quando o algoritmo é utilizado para determinar a melhor posição de instalação dos nós sensores, o TAFF é reduzido pela metade. O mesmo ganho também ocorreu para a TPP. Nota-se, também, que o

algoritmo consegue melhorar o ganho quando as buscas são realizadas com a influência ativa dos obstáculos nos cenários.

A.9 Análise inferencial do modelo RSSF-AG

A análise inferencial utilizada para verificar o modelo RSSF-AG proposto neste trabalho utilizará um modelo normal linear com parte aleatória, dada pela Equação (A.3), e com a parte sistemática, dada pela Equação (A.5):

$$y_i \sim N(\mu_i; \sigma^2) \quad (\text{A.3})$$

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i \quad (\text{A.4})$$

$$i = 1, 2, \dots, n \quad (\text{A.5})$$

em que:

- $y_i \in \mathbb{R}$ representa a variável resposta, função de avaliação ($f_{Avaliação}$);
- $x_{1i}; x_{2i}; x_{3i}; x_{4i} \in \mathbb{R}$ representam as variáveis explicativas: taxa de cobertura, energia residual, TAFF e TPP, respectivamente;
- $\alpha \in \mathbb{R}$ representa o ajuste de escala, em que a reta corta o eixo das ordenadas. Ou seja, é o valor esperado para a variável de trabalho quando as variáveis explicativas forem iguais a zero;
- $\beta_1 \in \mathbb{R}$ representa o incremento na função de avaliação quando se aumenta a variável taxa de cobertura, mantendo as demais variáveis fixas;
- $\beta_2 \in \mathbb{R}$ representa o incremento na função de avaliação quando se aumenta a variável energia residual, mantendo as demais variáveis fixas;
- $\beta_3 \in \mathbb{R}$ representa o incremento na função de avaliação quando se aumenta a variável TAFF, mantendo as demais variáveis fixas;
- $\beta_4 \in \mathbb{R}$ representa o incremento na função de avaliação quando se aumenta a variável TPP, mantendo as demais variáveis fixas;
- $\varepsilon_i \in \mathbb{R}$ representa os erros independentes do modelo que não serão observados pela função de avaliação. Essa componente tem uma distribuição da por $N(0; \sigma^2)$;

Com isso, construímos o modelo inicial, que considera todas as nossas variáveis, como apresentada na Equação (A.6).

$$f_{Avaliação} = \alpha + Cobertura x_{1i} + Energia x_{2i} + TAFF x_{3i} + TPP x_{4i} + \varepsilon_i \quad (A.6)$$

Para ajustar o modelo, foi utilizado o método AIC. Este método consiste em um algoritmo iterativo, em que, a cada iteração, é eliminada a variável que tenha o menor valor de AIC dentre as que sejam menores que o valor de AIC geral.

O modelo selecionado indicado pelo critério AIC não removeu nenhuma das variáveis explicativas, o que fornece fortes indícios de que o modelo está bem ajustado. Sendo assim, o modelo final está ilustrado na Equação (A.7).

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i \quad (A.7)$$

Na Tabela A.45 são apresentados os valores para os parâmetros estimados do modelo final.

Tabela A.45 Valores dos parâmetro do modelo RSSF-AG

Parâmetro	Valor
α	0,873
β_1	-0,082
β_2	-13,179
β_3	0,068

Assim, na Equação (A.8) é apresentada a reta ajustada para o modelo RSSF-AG.

$$y_i = 0,873 - 0,082x_{1i} - 13,179x_{2i} + 0,068x_{3i} + \varepsilon \quad (A.8)$$

$$i = 1, 2, \dots, 80 \quad (A.9)$$

Na Tabela A.46, são apresentadas as estimativas finais para os parâmetros significativos do modelo inferencial do RSSF-AG

Tabela A.46 Estimativas finais para os parâmetros significativos no modelo

	Estimativa	Erro Padrão	T	P-valor	R²	S
α	0,873	0,018	50,018	< 2e-16		
β_1	-0,082	0,022	-4,033	1,470e-04		
β_2	-13,179	0,385	-34,326	< 2e-16	0,956	3.982
β_3	0,068	0,016	4,347	4,960e-05		

Para auxiliar no entendimento do modelo descrito na Equação (A.8), serão utilizados os erros padrões dos parâmetros estimados α , β_1 , β_2 e β_3 , respectivamente $EP(\alpha)$, $EP(\beta_1)$, $EP(\beta_2)$ e $EP(\beta_3)$. Na Tabela A.47, é possível ver os erros padrão para os parâmetros estimados.

No ajuste do modelo, verificou-se um p-valor de < 2.2e-16, menor que o nível de significância de 0,050. Portanto, pode-se afirmar que o modelo obtido utilizando os critérios de AIC é adequado.

Tabela A.47 Valores dos erros dos parâmetro do modelo RSSF-AG

Parâmetro	Valor
$EP(\alpha)$	0,018
$EP(\beta_1)$	0,022
$EP(\beta_2)$	0,385
$EP(\beta_3)$	0,016

Para verificar a qualidade do ajuste do modelo de regressão normal linear utilizado neste trabalho, foi utilizada a medida coeficiente de determinação - R^2 , que representa a proporção da variabilidade total aplicada à reta ajustada. Usando o aplicativo *R Statistic*, o valor obtido para o coeficiente de determinação foi de 0,956. Em princípio, o melhor modelo tem o valor do coeficiente de determinação próximo a 1 (um). Portanto, conclui-se que o modelo é, de fato, representativo.

APÊNDICE B Gerador do Arquivo de Configuração

B.1 Entrada dos Dados de Configuração

Para executar o aplicativo, que foi desenvolvido utilizando a linguagem de programação Python e o sistema operacional Ubuntu, faz-se necessário abrir o Terminal ilustrado na Figura B.1. Com o terminal aberto, digita-se o comando “python configAG.py” que irá abrir o aplicativo de configuração ilustrado na Figura B.2.

Figura B.1 Tela do aplicativo Terminal do Ubuntu



O algoritmo proposto neste trabalho tem a flexibilidade de ser configurado para realizar a busca da melhor configuração de implantação dos nós sensores utilizando 4 (quatro) métricas: a taxa de cobertura; a energia residual; o TAFF; e o TPP. De acordo com a necessidade do usuário do algoritmo, pode-se configurar o algoritmo para fazer buscas individuais de cada uma das métricas ou uma combinação entre elas, quem irá determinar quais métricas serão utilizada na busca são os valores dados para os parâmetros: Cobertura (0..1); Energia (0..1); Latência

(0..1); e Taxa de Perda de Pacotes (0..1), do aplicativo (Figura B.2). Os possíveis valores para os 4 (quatro) parâmetros podem variar entre 0 (zero), nesse caso a métrica é desconsiderada, e 1 (um) que faz com que o algoritmo só utilize esta métrica. É importante informar que esses 4 (quatro) parâmetros tem a seguinte restrição: a soma total dos valores para os 4 (quatro) parâmetros não pode ser maior que 1 (um).

O segundo valor é o fator de importância da métrica energia residual, que é o somatório de toda energia disponível nos nós sensores após o termino da simulação da RSSF. Também utiliza a mesma variação dos valores da métrica taxa de cobertura.

O terceiro valor é o fator de importância do TAFF, que é o tempo necessário para que um pacote gerado em um nó sensor percorra toda a RSSF e chegue ao nó sorvedouro. Segue as mesmas restrições adotadas para os valores da métrica taxa de cobertura.

O quarto valor é o fator de importância TPP, que corresponde à razão entre, diferença entre o número de pacotes enviados menos o número de pacotes recebidos e o número de pacotes enviados. Com os valores seguindo as restrições adotadas para os valores da métrica de cobertura.

O quinto valor é o tempo máximo de atraso dos pacotes, que determina qual o valor crítico admissível para o atraso na transmissão dos pacotes entre nós sensores e sorvedouro. Esse valor é dado em segundos e deve ser ajustado de acordo com as restrições imposta pela aplicação que irá utilizar a RSSF.

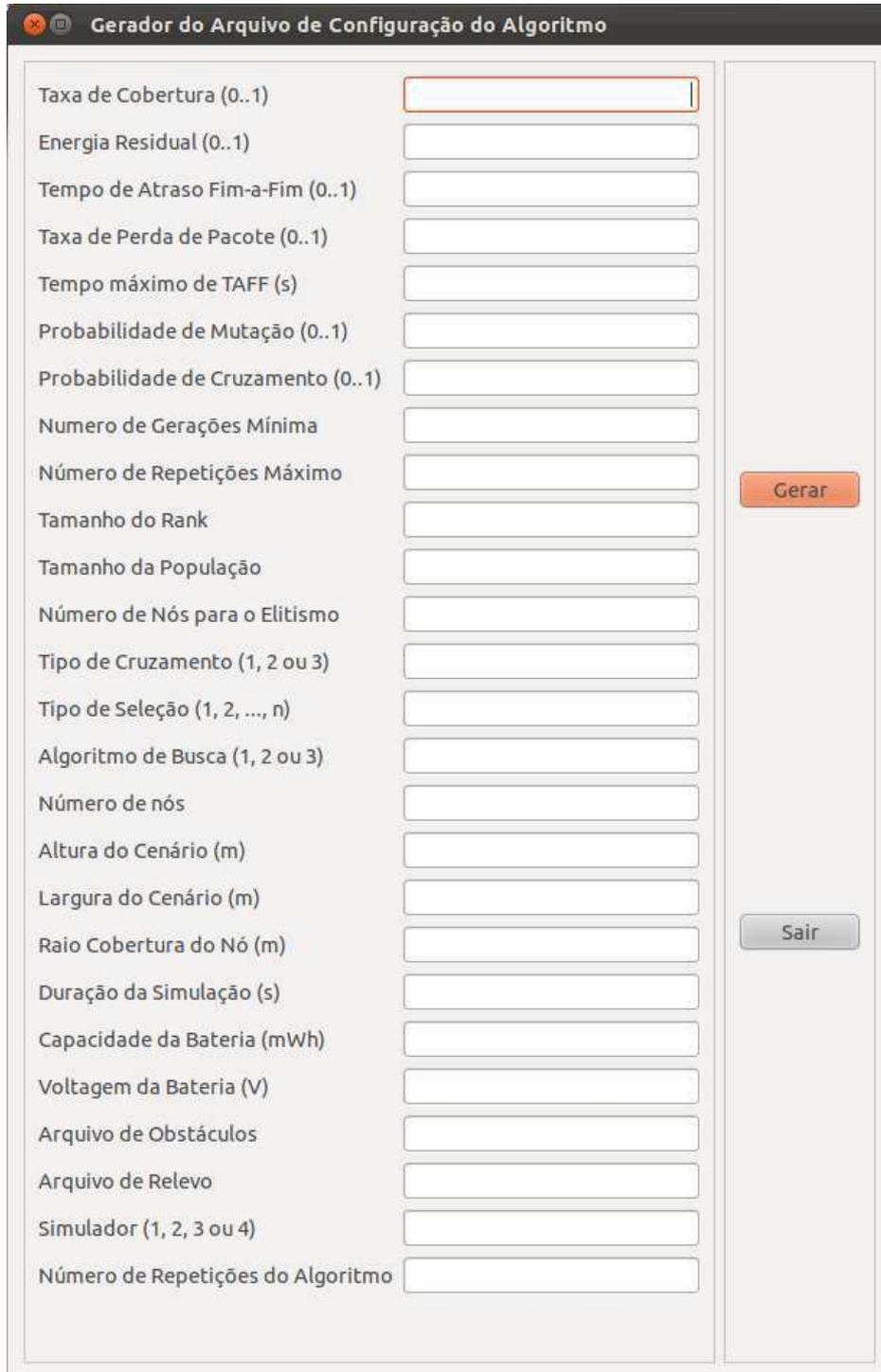
Os próximos valores são utilizados na configuração dos parâmetros do AG proposto para este trabalho. Os parâmetros do AG são: a probabilidade de mutação; a probabilidade de cruzamento; o número de gerações mínimo; o número de repetições máximo; o tamanho do *rank*; o tamanho da população; o número de nós para o elitismo; o tipo de cruzamento; o tipo de seleção; e o algoritmo de busca.

O valor da probabilidade de mutação é utilizado para quantificar as chances de aplicação do operador de mutação do AG durante a busca da melhor configuração de implantação dos nós sensores. Esse valor pode variar entre 0 (zero) e 1 (um), no entanto tipicamente são utilizados valores ate (0,01).

A probabilidade de cruzamento, de modo semelhante à probabilidade de mutação, é responsável por definir as chances de aplicação do operador de

cruzamento e os possíveis valores também podem variar entre 0 (zero) e 1 (um), sendo mais comuns valores próximo ou exatamente igual a 1 (um).

Figura B.2 Tela do aplicativo do gerador do arquivo de configuração do algoritmo proposto



Parâmetro	Valor
Taxa de Cobertura (0..1)	<input type="text"/>
Energia Residual (0..1)	<input type="text"/>
Tempo de Atraso Fim-a-Fim (0..1)	<input type="text"/>
Taxa de Perda de Pacote (0..1)	<input type="text"/>
Tempo máximo de TAFF (s)	<input type="text"/>
Probabilidade de Mutação (0..1)	<input type="text"/>
Probabilidade de Cruzamento (0..1)	<input type="text"/>
Numero de Gerações Mínima	<input type="text"/>
Número de Repetições Máximo	<input type="text"/>
Tamanho do Rank	<input type="text"/>
Tamanho da População	<input type="text"/>
Número de Nós para o Elitismo	<input type="text"/>
Tipo de Cruzamento (1, 2 ou 3)	<input type="text"/>
Tipo de Seleção (1, 2, ..., n)	<input type="text"/>
Algoritmo de Busca (1, 2 ou 3)	<input type="text"/>
Número de nós	<input type="text"/>
Altura do Cenário (m)	<input type="text"/>
Largura do Cenário (m)	<input type="text"/>
Raio Cobertura do Nó (m)	<input type="text"/>
Duração da Simulação (s)	<input type="text"/>
Capacidade da Bateria (mWh)	<input type="text"/>
Voltagem da Bateria (V)	<input type="text"/>
Arquivo de Obstáculos	<input type="text"/>
Arquivo de Relevô	<input type="text"/>
Simulador (1, 2, 3 ou 4)	<input type="text"/>
Número de Repetições do Algoritmo	<input type="text"/>

Botões: Gerar, Sair

O primeiro valor fornecido para o aplicativo de geração do arquivo de configuração é o fator de importância da métrica taxa de cobertura. A taxa de

cobertura é a relação entre a área coberta por algum nó sensor e a área total do cenário. Os possíveis valores para esse parâmetro estão entre 0 (zero) e 1 (um).

O valor número de geração mínima é quem determina o número mínimo de gerações que serão utilizadas para que o algoritmo proposto finalize a busca da melhor configuração. Esse parâmetro pode variar no intervalo de 1 a n geração.

O número de repetições máximo, que também é um parâmetro utilizado pelo algoritmo para determinar se o mesmo deve finalizar a busca, tem a função de garantir a evolução máxima dos indivíduos de uma população. A cada nova geração criada pelo o algoritmo, o valor do melhor indivíduo é comparado com o valor do melhor indivíduo da geração anterior, caso esses valores sejam iguais, o contador de número de repetições é incrementado, caso contrario o contador é zerado. Quando esse contador atinge o valor do número de repetições máximo o algoritmo é finalizado.

O valor do tamanho do *rank* determina quantos indivíduos serão utilizados na competição, caso o método de seleção utilizado no algoritmo seja o de *rank*; para determinar o indivíduo que será utilizado para gerar os indivíduos da próxima geração. Esse valor pode variar entre 2 (dois) e o tamanho da população.

O valor para tamanho da população define a quantidade de indivíduos ou cromossomos de uma geração. O tamanho da população deve ser grande o suficiente que garanta uma boa diversidade de indivíduos e conseqüentemente uma evolução satisfatória do algoritmo. Para esse parâmetro os valores podem variar entre 1 (um) e n.

O número de nós para o elitismo é o valor utilizado pelo algoritmo para preservar os melhores indivíduos de uma geração, pois os demais serão eliminados quando for criada uma nova geração.

O tipo de cruzamento define qual o método de cruzamento que será utilizado pelo operador de cruzamento do algoritmo, quando o valor utilizado nesse parâmetro for: 1 (um) o método selecionado será o cruzamento com um ponto; 2 (dois) cruzamento com dois pontos; e 3 (três) cruzamento uniforme.

O tipo de seleção define qual o método de seleção dos pais que serão utilizados para reprodução, quando o valor for 1 (um) o método utilizado será a roleta viciada e 2 (dois) ou mais para o método *rank*.

O valor do campo algoritmo de busca é utilizado para escolher o método de busca que será utilizado pelo algoritmo proposto. Os possíveis valores para esse parâmetro são: 1 (um) para escolher o AG com descarte de indivíduos; 2 (dois) para o AG sem descarte de indivíduos; e 3 (três) para uma busca aleatória.

O número de nós determina a quantidade de nós que cada cromossomo terá, portanto, determina a quantidade de dígitos utilizada na representação de cada cromossomo. Para cada nó são necessário 24 bits, sendo 8 (oito) para a coordenada X, 8 (oito) para a coordenada Y e 8 (oito) para a coordenada Z.

Os valores seguintes são utilizados para configurar alguns parâmetros do simulador, através do script de configuração do simulador. Esses parâmetros são: a altura e largura do cenário; o raio de cobertura do nó; a duração da simulação; a capacidade e voltagem da bateria; e o arquivo de obstáculos e relevo.

O valor da altura do cenário em conjunto com a largura do cenário define a área do cenário no qual os nós sensores serão posicionados. Esses valores são dados em metros.

O valor do campo raio de cobertura do nó, define o alcance dos dispositivos sensores instalado em cada nó. Esse valor é dado em metros.

O valor do campo duração da simulação define o tempo utilizado na simulação da RSSF, esse valor é dado em segundos.

Os valores dos campos capacidade da bateria e voltagem da bateria define a capacidade de energia armazenada em cada nó sensor da rede e sua voltagem de trabalho. O valor da capacidade é dado em (mWh) e a voltagem em (V).

O arquivo de obstáculos determina a posição e a forma de todos os obstáculos presentes no cenário, esse arquivo é construído através de uma matriz com 256 linhas e 256 colunas, quando uma posição deve ser considerada com obstáculos, o valor dessa célula é setada⁸ e caso não seja um obstáculo ela será resetada⁹.

O arquivo de relevo é utilizado para determinar o valor da variável Z de cada possível ponto de instalação de nós sensores no cenário, a arquivo de relevo

⁸ Setado o valor é colocado em 1 (um)

⁹ Resetada o valor é colocado em 0 (zero)

também é formado por uma matriz de 256 linhas por 256 colunas e os valores de cada célula dessa matriz é definido com o valor da altitude do relevo.

O valor para o parâmetro simulador, defini qual simulador será utilizado pelo algoritmo para testar as configurações geradas, mas, devido a limitações da maioria dos simuladores quanto à disponibilização de módulos de obstáculos, esse valor deve ser fixado em 4 (quatro) que corresponde ao simulador OMNeT utilizando o *framework* INET, pois, dentre os simuladores testado, ele foi o único que tinha o módulo de obstáculo funcional.

O ultimo valor é o número de repetições do algoritmo, utilizado para determinar quantas repetições do mesmo experimento será realizado pelo algoritmo proposto, esse parâmetro é utilizado para que o algoritmo gere vários arquivos de configuração de implantação dos nós da RSSF, para cada repetição é gerado 1 (um) arquivo de configuração.

Com todos os valores para os parâmetros definidos explicitamente no aplicativo, para gerar o arquivo de configuração do algoritmo proposto é necessário um clique no botão “Gerar”, logo após, será criado um arquivo como nome “parameters.cfg” que contém os valores de todos os parâmetros que serão utilizados para configurar o AG e o simulador. A Figura B.3 ilustra um exemplo de um arquivo de configuração.

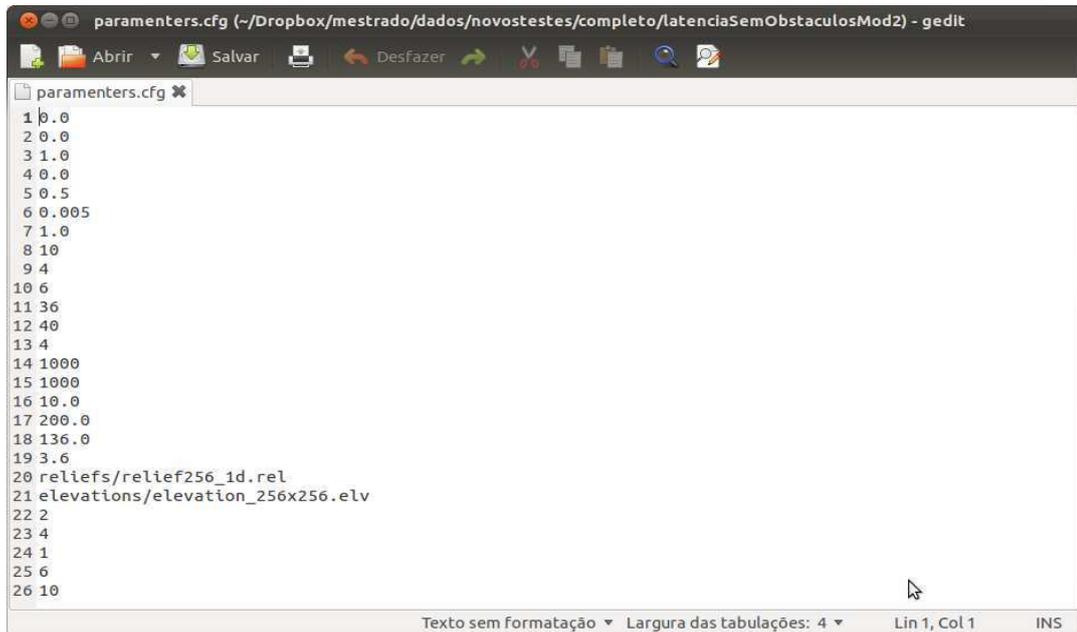
B.2 Execução do Algoritmo

Ao final da fase de definição dos valores para os parâmetros de configuração do modelo proposto o aplicativo gera um arquivo com os valores de configuração. Na Figura B.3, é ilustrada um exemplo de arquivo de configuração.

O primeiro passo dessa fase é abrir o aplicativo Terminal e executar o comando de inicialização do algoritmo de busca. Na Figura B.4, é ilustrada a tela do aplicativo Terminal e o comando de inicialização do algoritmo de busca.

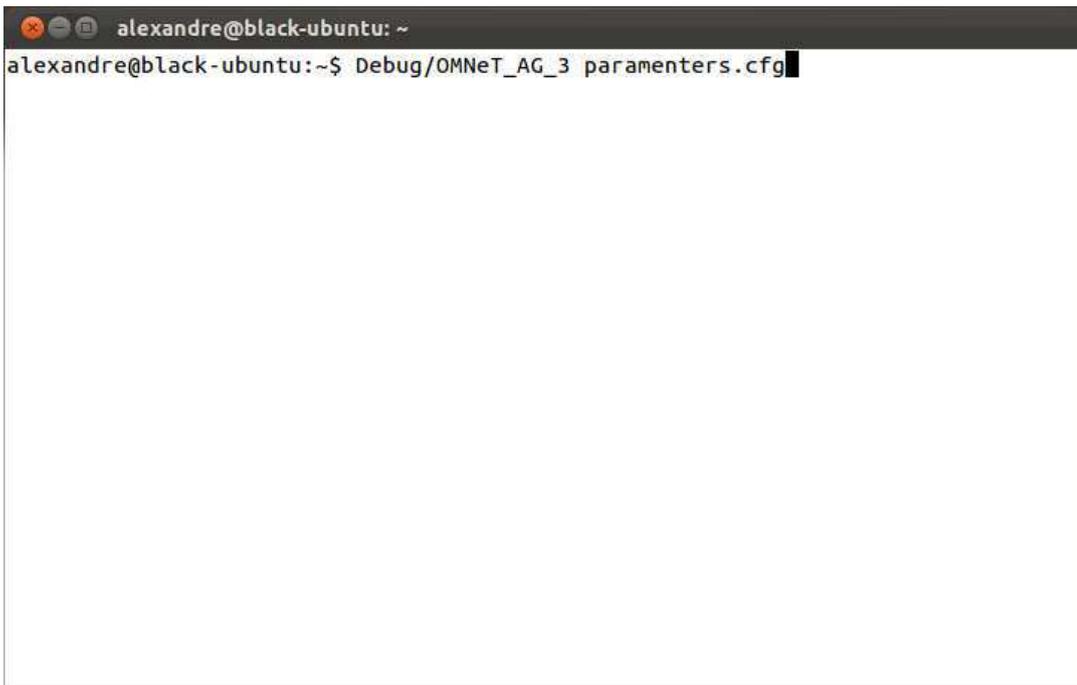
Um dos vários parâmetros configurado na primeira fase é o número de repetições do algoritmo. Por tanto, ao final da execução do algoritmo de busca é gerado um arquivo com a melhor configuração de implantação dos nós sensores para cada repetição. Na Figura B.5, é ilustrado um exemplo do conteúdo do arquivo gerado pelo algoritmo de busca ao final de cada repetição.

Figura B.3 Exemplo do arquivo de configuração do algoritmo proposto



```
paramenters.cfg
1 0.0
2 0.0
3 1.0
4 0.0
5 0.5
6 0.005
7 1.0
8 10
9 4
10 6
11 36
12 40
13 4
14 1000
15 1000
16 10.0
17 200.0
18 136.0
19 3.6
20 reliefs/relief256_1d.rel
21 elevattons/elevation_256x256.elv
22 2
23 4
24 1
25 6
26 10
```

Figura B.4 Tela do terminal com o comando de execução do algoritmo

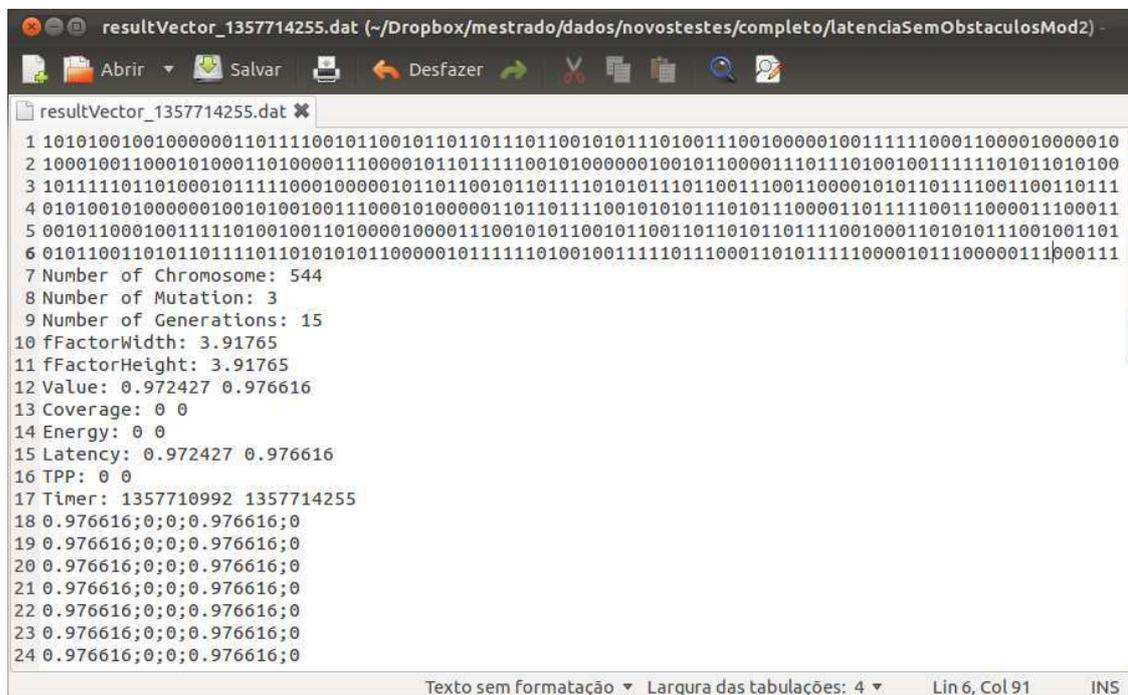


```
alexandre@black-ubuntu: ~
alexandre@black-ubuntu:~$ Debug/OMNeT_AG_3 paramenters.cfg
```

No exemplo ilustrado na Figura B.5, nas 6 (seis) primeiras linhas são apresentadas o vetor, em representação binária, do cromossomo ou individuo que obteve o valor máximo de aptidão calculado utilizando a função de avaliação apresentada na Equação (6.5). São com base neste vetor, que a fase final da

abordagem proposta, criar o arquivo com a configuração de implantação dos nós sensores da RSSF em estudo.

Figura B.5 Exemplo do arquivo gerado ao final da execução do algoritmo de busca



```
resultVector_1357714255.dat (~/.Dropbox/mestrado/dados/novostestes/completo/latenciaSemObstaculosMod2)
resultVector_1357714255.dat ✕
1 10101001001000000110111100101100101101101101100101011101001110010000010011111100011000010000010
2 1000100110001010001101000011100001011011110010100000010010110000111011101001001111101011010100
3 10111101101000101111000100000101101100101101111010101110110011100110001010110111001100110111
4 010100101000000100101001001110001010000011011011110010101110101110000110111100111000011100011
5 001011000100111110100100110100001000011100101011001011001101101011011110010001101010111001001101
6 010110011010110111101101010110000010111110100100111110111000110101111100001011100000111000111
7 Number of Chromosome: 544
8 Number of Mutation: 3
9 Number of Generations: 15
10 fFactorWidth: 3.91765
11 fFactorHeight: 3.91765
12 Value: 0.972427 0.976616
13 Coverage: 0 0
14 Energy: 0 0
15 Latency: 0.972427 0.976616
16 TPP: 0 0
17 Timer: 1357710992 1357714255
18 0.976616;0;0;0.976616;0
19 0.976616;0;0;0.976616;0
20 0.976616;0;0;0.976616;0
21 0.976616;0;0;0.976616;0
22 0.976616;0;0;0.976616;0
23 0.976616;0;0;0.976616;0
24 0.976616;0;0;0.976616;0
Texto sem formatação ▾ Largura das tabulações: 4 ▾ Lin 6, Col 91 INS
```

B.3 Saída do Arquivo de Implantação

Para executar aplicativo que gerar o arquivo com a lista de coordenadas de localização para cada nó sensor da RSSF, é necessário novamente utilizar o Terminal e digitar o comando conforme ilustrado na Figura B.6. O aplicativo irá extrair as informações e gerar um arquivo com o mesmo nome, mas com a extensão “cn”. Na Figura B.7 é ilustrado um exemplo do arquivo de saída, que contém as coordenadas de implantação de cada nó sensor da RSSF em estudo.

A primeira linha do arquivo gerado nessa fase contém a informação da duração da simulação utilizada nas buscas. As linhas de 3 (três) a 8 (oito) apresentam informações sobre o tamanho da área do cenário utilizado na simulação. A linha 10 (dez) apresenta a quantidade de nós sensores utilizados na configuração encontrada pelo algoritmo que apresentou o maior valor de aptidão. As demais linhas, a partir da décima segunda, são apresentados os valores das coordenadas X, Y e Z nas respectivas linhas 12 (doze), 13 (treze) e 14 (catorze) para cada um dos nós sensores que compõem a RSSF em estudo.

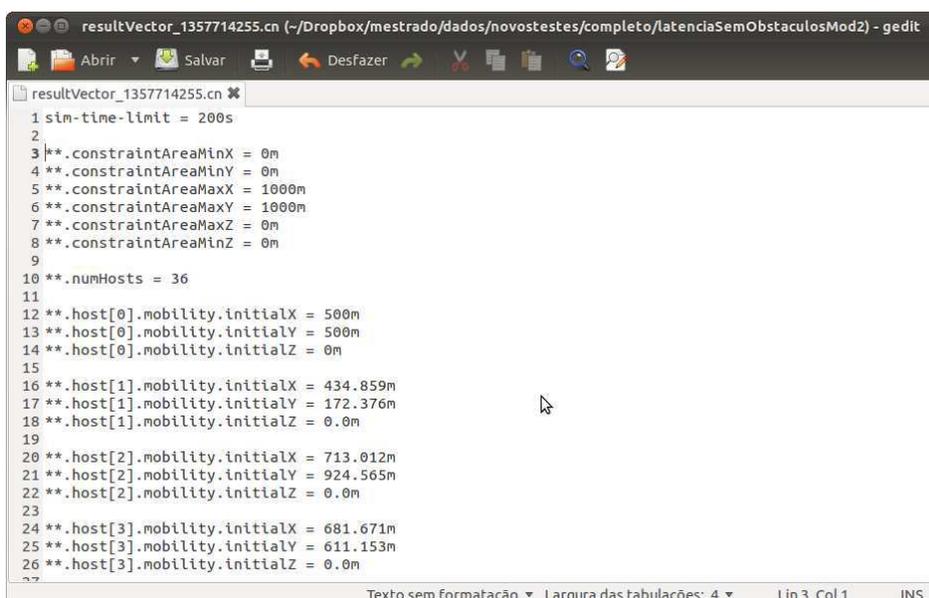
Figura B.6 Tela do terminal com o comando de geração do arquivo de saída da abordagem proposta



```
alexandre@black-ubuntu: ~  
alexandre@black-ubuntu:~$ Debug/OMNeT_Ger_01 resultVector_1362859089
```

Utilizando as informações contidas no arquivo de saída é possível criar várias formas de apresentações dos resultados, um delas faz uso do aplicativo CorelDRAW¹⁰ X5 para ilustrar graficamente o resultado encontrado pela aplicação do algoritmo proposto nesse modelo.

Figura B.7 Exemplo do arquivo de saída gerado pela abordagem proposta



```
resultVector_1357714255.cn (~/.Dropbox/mestrado/dados/novostestes/completo/latenciaSemObstaculosMod2) - gedit  
1 sim-time-limit = 200s  
2  
3 **.constraintAreaMinX = 0m  
4 **.constraintAreaMinY = 0m  
5 **.constraintAreaMaxX = 1000m  
6 **.constraintAreaMaxY = 1000m  
7 **.constraintAreaMaxZ = 0m  
8 **.constraintAreaMinZ = 0m  
9  
10 **.numHosts = 36  
11  
12 **.host[0].mobility.initialX = 500m  
13 **.host[0].mobility.initialY = 500m  
14 **.host[0].mobility.initialZ = 0m  
15  
16 **.host[1].mobility.initialX = 434.859m  
17 **.host[1].mobility.initialY = 172.376m  
18 **.host[1].mobility.initialZ = 0.0m  
19  
20 **.host[2].mobility.initialX = 713.012m  
21 **.host[2].mobility.initialY = 924.565m  
22 **.host[2].mobility.initialZ = 0.0m  
23  
24 **.host[3].mobility.initialX = 681.671m  
25 **.host[3].mobility.initialY = 611.153m  
26 **.host[3].mobility.initialZ = 0.0m  
27
```

¹⁰ CORELDRAW aplicativo desenvolvido pela Corel Corporation.

APÊNDICE C Script de Configuração do INET

Nos teste realizado com o simulador OMNeT++ e o *framework* INET, as configurações utilizadas para simular a rede foram:

- Camada de aplicação – foi utilizado o módulo de aplicação *UDPBasicBurst*, que utiliza o protocolo UDP para enviar rajadas de pacotes com 1024 bytes para o endereço do nó sorvedouro (*host[0]*), em intervalos variando entre 14 e 16 segundos. O primeiro pacote começa a ser enviado 2 (dois) segundos após o início da simulação e termina faltando 1 (um) segundo do tempo de duração da simulação;
- Camada de transporte – foi utilizado o módulo *UDP*, que é um protocolo sem conexão e não garante a entrega dos pacotes (datagrama), além de não garantir a ordem de recebimento das mensagens. No entanto, ele apresenta um pequeno cabeçalho, o que reduz o *overhead* nas transmissões e também torna a transmissão dos pacotes mais rápida em comparação com o protocolo TCP;
- Camada de rede – foi utilizado o módulo *DSRUU* para implementar o protocolo de rede DSR, que é mais indicado para redes estáticas ou de baixa velocidade de deslocamento. Ele tem como característica armazenar todo o caminho (saltos) para chegar ao nó destino, diminuindo o tempo de atraso fim-a-fim, caso a rede não tenha sofrido mudanças;
- Camada de enlace – foi utilizado o módulo *Ieee80211Nic*, que implementa uma rede sem fio baseada na IEEE 802.11 em modo de operação “g”, frequência de operação de 2.4GHz com 2.0mW de potência de transmissão. Nesse módulo, também são implementados os modelos de propagações. Para nosso trabalho, foi utilizado o *RayleighModel*, que é um modelo de desvanecimento variando de acordo com uma distribuição de Rayleigh. Ele é visto como um modelo razoável para propagações de sinais nas zonas urbanas.

A configuração do *framework* INET 2.0.0 utilizada nos experimentos realizados neste trabalho é apresentada no Quadro C.1. Essa configuração tem um impacto direto nos resultados dos ganhos médios obtidos pelo modelo RSSF-AG. No entanto, mesmo que outras configurações das camadas sejam modificadas, o modelo RSSF-AG irá sempre encontrar um caminho que proporcione um ganho na métrica que esteja sendo utilizada na busca. Um detalhe que precisa ser ressaltado é em relação à métrica taxa de cobertura. Devido sua natureza, quando o modelo RSSF-AG é configurado apenas para obter o maior ganho nessa métrica, não existe nenhuma influência da configuração utilizada no simulador de rede, pois essa métrica não faz uso do simulador no cálculo de sua aptidão.

Quadro C.1 Configuração do simulador de RSSF

```
[General]
network = Network

cmdenv-express-mode = true
#parallel-simulation = true

**.udpApp[*].scalar-recording = true
**.udpApp[*].vector-recording = true
**.endToEndDaly.scalar-recording = true
**.manetRouting.scalar-recording = true

**.debug = true
**.coreDebug = false

# nic settings
**.wlan[*].bitrate = 11Mbps

**.wlan[*].typename="Ieee80211Nic"
**.wlan[*].opMode="g"
**.wlan[*].mac.EDCA = false
**.wlan[*].mgmt.frameCapacity = 10
**.wlan[*].mac.address = "auto"
**.wlan[*].mac.maxQueueSize = 14
**.wlan[*].mac.rtsThresholdBytes = 3000B
**.wlan[*].mac.basicBitrate = 6Mbps # 24Mbps
**.wlan[*].mac.retryLimit = 7
**.wlan[*].mac.cwMinData = 31
**.wlan[*].mac.slotTime = 9us #
**.wlan[*].mac.AIFSN = 2 #DIFS

# channel physical parameters
*.channelControl.carrierFrequency = 2.4GHz
*.channelControl.pMax = 2.0mW
**.channelControl.sat = -77.5dBm
*.channelControl.sat = -110dBm
*.channelControl.alpha = 2
*.channelControl.numChannels = 1
*.channelControl.propagationModel = "RayleighModel"
```

```

# radio setup
**.radio.drawCoverage = true
**.wlan[*].radio.transmitterPower = 2.0mW
**.wlan[*].radio.pathLossAlpha = 2
**.wlan[*].radio.snirThreshold = 4dB # in dB
###.wlan[*].radio.sensitivity = -77dBm
**.wlan[*].radio.thermalNoise = -110dBm
**.radio.sensitivity = -85dBm
**.wlan[*].radio.berTableFile = "per_table_80211g_Trivellato.dat"

# Annotation parameters
*.annotations.draw = true

# Obstacle parameters
*.obstacles.debug = true
*.obstacles.obstacles = xmldoc("test.obstacles5.xml")

# Battery
**.batteryType = "InetSimpleBattery"

**.usage_radio_idle = 216mA
**.usage_radio_recv = 260mA
**.usage_radio_send = 375mA
**.usage_radio_sleep = 9mA

**.battery.nominal = 136
**.battery.capacity = 136
**.battery.voltage = 3.6
**.battery.resolution = 1s
**.battery.publishDelta = 0.05
**.battery.publishTime = 1s
**.battery.ConsumedVector = true

# Configure Applications Nodes
**.host[*].numUdpApps = ${1,1,1,1}
**.host[*].udpApp[*].typename = "UDPBasicBurst"
**.host[*].udpType = "UDP"
**.host[*].udpApp[*].destAddresses = "host[0]"
**.host[*].udpApp[*].chooseDestAddrMode = "perBurst"
**.host[*].udpApp[*].destPort = 1234
**.host[*].udpApp[*].messageLength = 1024B
**.host[*].udpApp[*].burstDuration = 0
**.host[*].udpApp[*].sleepDuration = 4s
**.host[*].udpApp[*].sendInterval = 15s + uniform(-1.0s,1.0s)
**.host[*].udpApp[*].localPort = 1234
**.host[*].udpApp[*].startTime = 2s
**.host[*].udpApp[*].stopTime = 195s
**.host[*].udpApp[*].delayLimit = 1s

**.host[*].mobility.initFromDisplayString = false
**.host[*].mobilityType = "StationaryMobility"

**.routingProtocol="DSRUU"
**.PrintDebug = true
**.FlushLinkCache = true
**.PromiscOperation = false
**.UseNetworkLayerAck = false
**.BroadcastJitter = 20 # 20 ms

```

```

**.RouteCacheTimeout = 300 #300 seconds
**.SendBufferTimeout = 300# //30 s
**.SendBufferSize = -1
**.RequestTableSize = -1
**.RequestTableIds = -1
**.MaxRequestRexmt = -1 #// 16,
**.MaxRequestPeriod = 10 #//10 SECONDS
**.RequestPeriod = 500 #//500 MILLISECONDS
**.NonpropRequestTimeout = 30# //30 MILLISECONDS
**.RexmtBufferSize = -1 #//MAINT_BUF_MAX_LEN
**.MaintHoldoffTime = 250# //250 MILLISECONDS
**.MaxMaintRexmt = 2 # //2
**.TryPassiveAcks = true #//1
**.PassiveAckTimeout = 100# //100 MILLISECONDS
**.GratReplyHoldOff = 1 #, //1 SECONDS
**.MAX_SALVAGE_COUNT = 15 # //15
**.LifoSize = 20
**.PathCache = true
**.ETX_Active = false
**.ETXHelloInterval = 1 #, // Second
**.ETXWindowNumHello = 10
**.ETXRetryBeforeFail = -1
**.RREPDestinationOnly = false
**.RREQMaxVisit = 5 # // Max Number that a RREQ can be processes by a node

#include scenary.scn

[Config config0]
description="config0"
include scenary0.scn

[Config config1]
description="config1"
include scenary1.scn

[Config config2]
description="config2"
include scenary2.scn

[Config config3]
description="config3"
include scenary3.scn

```