

**LOCALIZAÇÃO DE DISPOSITIVOS MÓVEIS  
USANDO ROTEADORES COM ANTENAS  
DIRECIONAIS E CLASSIFICAÇÃO DE DADOS**

ATACÍLIO COSTA CUNHA

**LOCALIZAÇÃO DE DISPOSITIVOS MÓVEIS  
USANDO ROTEADORES COM ANTENAS  
DIRECIONAIS E CLASSIFICAÇÃO DE DADOS**

Dissertação apresentada ao Programa de Pós-Graduação em Informática do Instituto de Ciências Exatas da Universidade Federal do Amazonas como requisito parcial para a obtenção do grau de Mestre em Informática.

**ORIENTADOR: HORÁCIO A. B. FERNANDES DE OLIVEIRA**

Manaus

Maio de 2014

© 2014, Atacílio Costa Cunha.  
Todos os direitos reservados.

Cunha, Atacílio Costa  
A524p Localização de Dispositivos Móveis usando  
Roteadores com Antenas Direcionais e Classificação de  
Dados / Atacílio Costa Cunha. — Manaus, 2014  
xiii, 75 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal do  
Amazonas

Orientador: Horácio A. B. Fernandes de Oliveira

Localização Wifi, kNN, RSSI

CDU 1234

# [Folha de Aprovação]

Quando a secretaria do Curso fornecer esta folha, ela deve ser digitalizada e armazenada no disco em formato gráfico.

Se você estiver usando o `pdflatex`, armazene o arquivo preferencialmente em formato PNG (o formato JPEG é pior neste caso).

Se você estiver usando o `latex` (não o `pdflatex`), terá que converter o arquivo gráfico para o formato EPS.

Em seguida, acrescente a opção `approval={nome do arquivo}` ao comando `\ppgccufmg`.

Se a imagem da folha de aprovação precisar ser ajustada, use:  
`approval=[ajuste] [escala] {nome do arquivo}`  
onde *ajuste* é uma distância para deslocar a imagem para baixo e *escala* é um fator de escala para a imagem. Por exemplo:  
`approval=[-2cm] [0.9] {nome do arquivo}`  
desloca a imagem 2cm para cima e a escala em 90%.

*A Deus, a minha noiva e a minha família...*

# Agradecimentos

Agradeço primeiramente a Deus por conceder todos os dias de vida para chegar até aqui, pela força e perseverança concedida nos momentos mais difíceis, por jamais ter me desamparado e por mais esta realização na minha vida.

À minha noiva Aline por ter me incentivado para que eu não deixasse de fazer o mestrado, por ter orado por mim e também por ter me apoiado a cada momento.

Ao meu orientador Dr. Horácio, pelos conselhos, pelo apoio durante todos os momentos da minha pesquisa, por ter me aceitado como orientando e por ter me guiado durante todos esses dias focando no meu melhor desempenho e rendimento.

À minha família e amigos por terem sido tão compreensivos a cada momento, pelos conselhos diários, pelas orações e por terem me apoiado até aqui.

Ao Luis, Demétrius e Brandão por terem me ajudado bastante na parte técnica deste trabalho.

Ao Juan, pelas indicações e conselhos dados para o melhor desenvolvimento da minha pesquisa.

À Cris, por ter me ajudado nas revisões dos textos em inglês que escrevi durante minha pesquisa.

Ao Carlos e aos amigos de trabalho do SIDIA, que me apoiaram, me aconselharam e me ajudaram durante os momentos finais desta dissertação.

# Resumo

Os sistemas de Localização WiFi estão cada vez mais presentes no nosso dia-a-dia devido à grande facilidade que os dispositivos utilizados nesta tecnologia podem proporcionar para os usuários, como mobilidade, independência de fatores que limitam e causam desconforto na usabilidade como cabos, por exemplo, e também por ser mais conveniente para o usuário utilizar algo de fácil manuseio. Além destes pontos, sabe-se que os dispositivos WiFi estão cada vez mais desenvolvidos, proporcionando mais possibilidades de uso no cotidiano dos usuários.

Sistemas de Localização WiFi realizam a localização de dispositivos ou pessoas em um determinado ambiente com base em medições de potência de sinal, obtidas através da troca de pacotes entre os dispositivos que compõem tais sistemas.

Neste trabalho é proposto um método de Localização WiFi utilizando múltiplos roteadores com antenas direcionais no ambiente de localização. Os roteadores com antenas direcionais são utilizados neste trabalho devido a sua capacidade de limitar a cobertura sinal para uma determinada área, diferente dos sistemas com antenas omnidirecionais. O impacto que um pacote enviado por um dispositivo com posição desconhecida gera nas antenas direcionais do sistema é capaz de proporcionar valores de RSSI diferentes.

Através da fase *offline* do sistema onde é realizado a parte de treinamento, um banco de dados é criado com todos os valores de RSSI obtidos a partir de cada pacote recebido pelos roteadores direcionais e as posições de onde os pacotes foram enviados. Essa base de dados é utilizada como conhecimento para que o algoritmo de classificação possa estimar de onde um pacote foi enviado, somente com base nos valores de RSSI obtidos em cada roteador direcional. Foram comparados dois algoritmos de classificação nesta proposta como forma de analisar as características de comportamento de cada um deles nos ambientes de localização. São eles os algoritmos *kNN* e Naive Bayes, ambos bastante utilizados na literatura.

Durante a experimentação do sistema de localização composto pelos roteadores direcionais, foi analisado também como organizar os roteadores para ser obter resulta-

dos com maior precisão de localização. Desta forma, foi possível analisar nos ambientes de localização a organização dos roteadores unidos apontando para direções diferentes, e separados focados para o centro do ambiente de localização.

Outro ponto importante desta proposta é a aplicação do sistema de localização com base em dois tipos de localização: Localização Precisa e Localização Contextual. A Localização Precisa permite estimar a posição exata de um dispositivo nos eixos  $x$  e  $y$  de um determinado ambiente onde ele está instalado. Já a Localização Contextual, permite estimar posição de um determinado dispositivo com base na análise do contexto do ambiente onde ele está inserido.

Todos os dados obtidos durante os experimentos foram analisados e testados, e demonstram que o sistema proposto neste trabalho pode ser utilizado como sistema de localização WiFi para ambientes internos como alternativa do sistema comumente de localização encontrado na literatura, composto por roteadores omnidirecionais.

Conforme a estrutura do ambiente de localização onde o sistema estava inserido e a organização dos roteadores direcionais foi possível obter resultados bem semelhantes ao do sistema de localização convencional composto por roteadores omnidirecionais. Isto pode ser visto nos resultados obtidos em um ambiente pequeno de localização composto pelos laboratórios do IComp onde, com roteadores direcionais e classificação de dados a partir do algoritmo kNN, foi possível obter 84,26% de precisão na localização comparada com 83,93% conseguida com roteadores direcionais juntos utilizando o mesmo algoritmo de classificação. Outro ponto interessante visto nos resultados deste trabalho é a utilização do algoritmo Naive Bayes tende a ser melhor para o sistema quando o mesmo utiliza roteadores direcionais ao invés de omnidirecionais, já que este algoritmo consegue analisar como o direcionamento de cada antena impacta nas características de recepção de sinal em um determinado ambiente. Isto pode ser visto por exemplo nos resultados obtidos em um ambiente maior de localização composto pelas salas de professores do IComp onde o sistema de localização com antenas direcionais espalhadas obteve 80,16% de precisão na estimativa de localização comparado com o sistema composto por antenas omnidirecionais que obteve 71,68%.

Com isto, pretende-se mostrar que os sistemas de localização podem ser compostos por roteadores direcionais como alternativa para o sistema convencional de localização WiFi.

# Abstract

WiFi Location systems are increasingly present in our days because of the great ease that the devices can provide to users, such as mobility, independence of factors that limit and cause discomfort usability as cables, for example, and also be more convenient for the user to use something easy to handle. Apart from these points, it is known that the WiFi devices are increasingly developed, providing more opportunities for use in the routine of users.

WiFi Location Systems perform device or people discovery in a given environment based on power signal measurements obtained by exchanging packets between devices that compose such systems.

This paper proposes a Location WiFi method using multiple routers with directional antennas in the location environment. Routers with directional antennas are used in this study due to its ability to limit signal coverage for a given area, unlike systems with omnidirectional antennas. The impact that a packet sent by a device with unknown position generates in directional antennas installed at the location system can provide different RSSI values.

Through the system *offline* stage where the training part is done, a database is created with all RSSI values obtained from each packet received by directional routers and positions from where the packages were sent. This database is used as knowledge for the classification algorithm so that it can estimate from where a package was sent, only based on RSSI values obtained at each directional router. Two classification algorithms were compared in this proposal to analyze the performance characteristics of each one in any location environment. They are the algorithms *kNN* and Naive Bayes, both widely used in the literature.

During the the training part of the localization system composed by directional routers, it was also analyzed how to organize the routers to get results with greater location accuracy. Thus, it was possible to analyze the organization of the routers in the location environments, being together and pointing in different directions, and separated and focused to the center of the localization environment.

Another important point of this proposal is the application of the localization system based on two types of location: Precise Location and Contextual Location. The precise location allows to estimate the exact position of a device in the axes  $x$  and  $y$  of a particular environment where it is installed. And the contextual location allows to estimate the position of a particular device based on the analysis of the context of the environment where it is inserted.

All data collected during the experiments were analyzed and tested, and show that the system proposed in this paper can be used as a WiFi positioning system for indoor environments as an alternative to the system commonly found in the location literature, composed of omnidirectional routers.

According to the structure of the location where the system was inserted and the organization of directional routers, it was possible to obtain very similar results compared to the location system composed by omnidirectional routers. This can be seen from the results obtained in a small location environment composed by the ICOMP laboratories where the system with directional routers and data classification from the kNN algorithm, has achieved 84.26% of accuracy in positioning compared with 83.93% achieved with directional routers together using the same classification algorithm. Another interesting point seen in the results of this work is that the use of Naive Bayes algorithm tends to be better when the system uses directional routers rather than omni-directional one. This happens because this algorithm can analyze as the direction of each antenna impacts the characteristics of signal reception in a given environment. This can be seen for example in the results obtained on a larger localization environment composed of ICOMP teachers rooms where the positioning system with scattered directional antennas got 80.16% accuracy in positioning estimation compared with the system composed by omnidirectional antennas that got 71.68%.

Thus, we shall show that the tracking systems can be composed of directional routers as an alternative to the WiFi location conventional system.

# Lista de Figuras

2.1	Classificação dos algoritmos de localização em Redes sem Fio. Fonte: Adaptado de Robles et al. (2010) . . . . .	7
2.2	Exemplo de classificação pelo algoritmo kNN com k=3. . . . .	15
2.3	Padrão de Radiação das Antenas Omnidirecionais. . . . .	18
2.4	Padrão de Radiação das Antenas Direcionais. . . . .	18
4.1	Exemplo do ambiente de treinamento com utilização de roteadores com antenas direcionais espalhados. . . . .	38
4.2	Exemplo do ambiente de treinamento com utilização de roteadores com antenas direcionais unidos. . . . .	39
4.3	Exemplo do ambiente de treinamento com utilização de roteadores com antenas omnidirecionais espalhados. . . . .	39
5.1	Ambiente 1 dos experimentos com roteadores direcionais espalhados. . . . .	44
5.2	Ambiente 1 dos experimentos com roteadores direcionais juntos. . . . .	44
5.3	Ambiente 2 dos experimentos com roteadores direcionais espalhados. . . . .	45
5.4	Ambiente 2 dos experimentos com roteadores direcionais juntos. . . . .	45
5.5	Linksys WRT610N. Fonte: <a href="http://www.ubergizmo.com">www.ubergizmo.com</a> . . . . .	46
5.6	NanoStation Loco M2. Fonte: <a href="http://www.comrede.com.br">www.comrede.com.br</a> . . . . .	46
5.7	Dell Vostro 3560. Fonte: <a href="http://www.dell.com">www.dell.com</a> . . . . .	46
5.8	Organização dos Roteadores Direcionais Espalhados e Omnidirecionais nos Laboratórios. . . . .	47
5.9	Organização dos Roteadores Direcionais Juntos e Omnidirecionais nos Laboratórios. . . . .	47
5.10	Organização dos Roteadores Direcionais Espalhados e Omnidirecionais nas Salas dos Professores. . . . .	48
5.11	Organização dos Roteadores Direcionais Juntos nas Salas dos Professores. . . . .	48
5.12	Função de Distribuição Acumulativa de Erro por Distância de Classificação. . . . .	54

5.13	Distribuição de Acertos por Posição usando Roteadores Direcionais Unidos e algoritmo kNN - Laboratórios. . . . .	55
5.14	Distribuição de Acertos por Posição usando Roteadores Omnidirecionais com algoritmo kNN - Laboratórios. . . . .	55
5.15	Organização dos Roteadores no Ambiente de Localização Composto pelos Laboratórios. . . . .	56
5.16	Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Omnidirecionais nos Laboratórios. . . . .	57
5.17	Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Direcionais Juntos nos Laboratórios. . . . .	57
5.18	Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Direcionais Separados nos Laboratórios. . . . .	58
5.19	Gráfico de Função de Distribuição Acumulativa de Erro para Estimativa de Sala. . . . .	61
5.20	Distribuição de Acertos por Sala usando Roteadores Direcionais Espalhados e algoritmo kNN - Salas dos Professores. . . . .	61
5.21	Distribuição de Acertos por Sala usando Roteadores Omnidirecionais e algoritmo kNN - Sala dos Professores. . . . .	62
5.22	Organização dos Roteadores no Ambiente de Localização Composto pelas Salas dos Professores. . . . .	63
5.23	Variação do Sinal por Sala (Bloco 1) - Roteadores Omnidirecionais nas Salas dos Professores. . . . .	64
5.24	Variação do Sinal por Sala (Bloco 1) - Roteadores Direcionais Juntos nas Salas dos Professores. . . . .	64
5.25	Variação do Sinal por Metro (Bloco 1) - Roteadores Direcionais Separados nas Salas dos Professores. . . . .	65
5.26	Variação do Sinal por Sala (Bloco 2) - Roteadores Omnidirecionais nas Salas dos Professores. . . . .	66
5.27	Variação do Sinal por Sala (Bloco 2) - Roteadores Direcionais Juntos nas Salas dos Professores. . . . .	66
5.28	Variação do Sinal por Metro (Bloco 2) - Roteadores Direcionais Separados nas Salas dos Professores. . . . .	67

# Lista de Tabelas

3.1	Comparação dos trabalhos relacionados a Localização WiFi . . . . .	23
5.1	Porcentagem de instâncias classificadas corretamente nos experimentos realizados nos laboratórios do IComp. . . . .	52
5.2	Porcentagem de instâncias classificadas corretamente nos experimentos realizados nas salas dos professores do IComp. . . . .	59

# Sumário

Agradecimentos	vi
Resumo	vii
Abstract	ix
Lista de Figuras	xi
Lista de Tabelas	xii
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos . . . . .	3
1.1.1 Objetivo Geral . . . . .	4
1.1.2 Objetivos Específicos . . . . .	4
1.2 Organização da Dissertação . . . . .	5
<b>2 Fundamentação Teórica</b>	<b>6</b>
2.1 Localização em Redes Sem Fio . . . . .	7
2.1.1 Estimativa de Distância . . . . .	10
2.1.2 Cálculo de Posicionamento . . . . .	11
2.2 Classificação de Dados . . . . .	12
2.2.1 Algoritmo <i>k-Nearest Neighbor</i> - <i>kNN</i> . . . . .	14
2.2.2 Algoritmo Naive Bayes . . . . .	15
2.3 Antenas . . . . .	16
2.4 Características das Antenas Omnidirecionais e Direcionais . . . . .	17
2.5 Conclusão . . . . .	20
<b>3 Trabalhos Relacionados</b>	<b>22</b>
3.1 Localização Baseada em RSSI . . . . .	23
3.2 Localização Baseada em Antenas Direcionais . . . . .	25

3.3	Localização Baseada em Treinamento . . . . .	29
3.4	Conclusão . . . . .	34
<b>4</b>	<b>Localização de Dispositivos Móveis usando Roteadores com Antenas Direcionais e Classificação de Dados</b>	<b>35</b>
4.1	Método Proposto . . . . .	35
4.2	Abordagens de Localização . . . . .	37
4.2.1	Abordagem com Utilização de Roteadores com Antenas Direcio- nais Espalhados . . . . .	37
4.2.2	Abordagem com Utilização de Roteadores com Antenas Direcio- nais Unidos . . . . .	38
4.2.3	Abordagem com Utilização de Roteadores com Antenas Omni- direcionais Espalhados . . . . .	38
4.3	Conclusão . . . . .	40
<b>5</b>	<b>Análise e Avaliação dos Resultados</b>	<b>42</b>
5.1	Experimentos . . . . .	43
5.1.1	Local dos Experimentos . . . . .	43
5.1.2	Equipamentos Utilizados . . . . .	46
5.1.3	<i>Softwares</i> Utilizados . . . . .	48
5.2	Localização Precisa . . . . .	52
5.3	Localização Contextual . . . . .	59
5.4	Conclusão . . . . .	67
<b>6</b>	<b>Conclusões</b>	<b>69</b>
6.1	Trabalhos Futuros . . . . .	71
	<b>Referências Bibliográficas</b>	<b>72</b>

# Capítulo 1

## Introdução

Redes Sem Fio possuem um número cada vez maior de aplicações no nosso dia-a-dia devido aos benefícios agregados à sua utilização como independência de fios, mobilidade, flexibilidade de cobertura, rápida instalação, baixo custo equipamentos e possibilidade de atender a vários usuários conectados ao mesmo tempo sem necessidade de equipamentos extras. Uma das áreas em Redes Sem Fio que cada vez mais tem recebido destaque é a Localização WiFi (Boukerche et al., 2007; Bahl & Padmanabhan, 2000; Fang et al., 2010). Sistemas de Localização WiFi são geralmente utilizados em ambientes internos, devido às altas taxas de erros na utilização de GPS (*Global Position System*) nestes sistemas. Tal solução é mais utilizada em sistemas de localização para ambientes externos (Boukerche et al., 2007), pois em ambientes internos ocorre a obstrução dos sinais dos satélites através das características do ambiente onde o sistema está implantado.

A Localização WiFi pode ser utilizada para:

- rastreamento: observação contínua de pessoas e objetos a partir da utilização de um modelo que descreva ao sistema os movimentos dos componentes;
- monitoramento: coleta e armazena dados dos nós localizados no ambiente de atuação do sistema;
- posicionamento: gera o posicionamento de nós estacionários ou em movimento num determinado ambiente.

As aplicações de localização têm crescido em várias categorias de mercado como em ambientes organizacionais, transporte público e redes sociais. Tal fator está relacionado ao advento dos *smartphones*, com a grande variedade de aplicações de localização que têm ganhado espaço entre os consumidores. Além desta visão comercial, pode-se

encontrar também aplicações e trabalhos com localização WiFi para as áreas de saúde como, por exemplo, para endoscopia utilizando cápsula sem fio, como visto no trabalho de Khan et al. (2011) e no de Swar et al. (2012). Dessa forma, podemos ver a grande importância da localização WiFi para diversos campos de aplicações e necessidades nos dias atuais.

De maneira geral, os sistemas desenvolvidos para a área de Localização Wifi utilizam nós espalhados em um ambiente, dos quais alguns têm suas posições conhecidas, e por isso recebem o nome de nós âncoras, servindo de referencial para outros nós que não conhecem o seu posicionamento, estes chamados de nós alvos. Os referenciais dos nós alvos a partir dos nós âncoras são obtidos a partir de métodos de estimativa de distância ou de ângulo, e, quando um ou mais destes métodos são utilizados para gerar informação de referencial de localização, é possível calcular o posicionamento do nó a partir de técnicas de multilateração (Barsocchi et al., 2009) ou de treinamento (método *fingerprint*) (Meng et al., 2011).

Sabe-se que a maioria dos sistemas de localização encontrados na literatura utiliza roteadores ou dispositivos com antenas omnidirecionais. Este tipo de antena recebe ou irradia sinais em todas as direções, respectivamente, e, dessa forma, acaba tendo sua potência de sinal reduzida em um determinado ambiente, o que pode levar a muitas vezes não conseguir alcançar um determinado nó para receber ou enviar informações. As antenas direcionais não possuem tal desvantagem pois concentram todo o seu poder de sinal para uma determinada região, não desperdiçando energia em áreas que não são de interesse para a localização de sistemas, melhorando as condições de localização no ambiente em análise, e proporcionando por fim melhores condições para as características de RSSI que serão utilizadas como conhecimento para os algoritmos de localização. Até a escrita desta dissertação não foi encontrado nenhum trabalho que utilizasse roteadores direcionais para estimativa de posicionamento de nós com base na análise do impacto de recebimento de sinal a partir do direcionamento dos roteadores.

Por isso, neste trabalho é implementado e avaliado uma solução de localização WiFi, utilizando 3 ou mais roteadores com antenas direcionais. Tais roteadores serão responsáveis por localizar todos os nós que estiverem na área de alcance do sinal do sistema. Neste sistema, a força do sinal enviado por um determinado dispositivo, chegando ao mesmo tempo nos diversos roteadores com antenas direcionais a ângulos diferentes, gera informações suficientes para que seja possível estimar com precisão a sua posição.

Sem a necessidade de dispositivos dispendiosos para a localização do nó, o nosso sistema pode facilmente ser implementado em qualquer ambiente, sem a necessidade de antenas ou qualquer hardware extra especialmente desenvolvido para localização.

O sistema proposto neste trabalho é baseado em treinamento de dados e por isto possui as seguintes fases:

- fase *offline*: vários pacotes são enviados de posições conhecidas em um ambiente interno, e são registrados tanto os valores de RSSI obtidos em cada roteador direcional como a identificação da posição de onde cada pacote foi enviado para criação da base de dados que é utilizada como conhecimento sobre os aspectos do sinal no ambiente de localização, para utilização na próxima fase;
- fase *online*: dispositivos com posições desconhecidas enviam pacotes para o sistema e os valores de RSSI obtidos pelos roteadores direcionais são utilizados como parâmetros para estimação da posição de envio a partir da atuação do algoritmo de classificação sobre a base de dados utilizada como conhecimento.

Este trabalho utiliza dois algoritmos de classificação: o primeiro é o algoritmo *k-Nearest Neighbor (kNN)*, que prediz a classe de uma determinada instância com base nos  $k$  exemplos de treino mais próximos no espaço de características, e o segundo é o algoritmo *Naive Bayes* que considera que cada atributo não possui relação com nenhum outro atributo e analisa como um atributo pode contribuir para a probabilidade de definição de uma classe.

O algoritmo kNN foi escolhido por, além de ser um referencial como algoritmo de classificação em muitos trabalhos de localização (Lee & Lampe, 2011; Bahl & Padmanabhan, 2000), seu funcionamento se adequar bastante às condições da localização utilizada nesta proposta, já que devido à condição instável do RSSI, diversas instâncias no espaço de características podem ser geradas próximas à localização desejada. Já o algoritmo Naive Bayes foi escolhido por, além de ser um algoritmo interessante de ser explorado na área de localização, seu funcionamento tende a atribuir maior importância às características que melhor definem a classe de um problema, neste caso, a posição de envio do sinal. Com isto, o algoritmo Naive Bayes pode identificar qual sistema funciona de maneira mais estável e confiável para a estimativa de posicionamento: o sistema com antenas omnidirecionais ou com direcionais.

Desta forma, o sistema especificado neste trabalho apresenta um método que contribui para a área de localização de nós móveis em redes WiFi através da implementação e avaliação de uma nova técnica de localização que utiliza os valores de RSSI obtidos a partir de roteadores com antenas direcionais.

## 1.1 Objetivos

A partir da contextualização da localização WiFi nos dias de hoje, podemos definir o objetivo geral e os objetivos específicos deste trabalho.

### 1.1.1 Objetivo Geral

O objetivo geral deste trabalho é implementar e avaliar um sistema de localização baseado em múltiplos roteadores com antenas direcionais que mapeiam a localização de dispositivos móveis dispostos em um ambiente interno com base em treinamento realizado durante uma etapa de pré-processamento.

### 1.1.2 Objetivos Específicos

Para se alcançar o objetivo geral, é necessário alcançar outros objetivos específicos:

1. **Criar um sistema de recepção de pacotes capaz de detectar que o mesmo pacote foi recebido em múltiplos roteadores:** é preciso identificar que o mesmo pacote foi recebido nos diversos roteadores, para que seja possível a extração da potência do sinal e do tempo de chegada dos pacotes iguais;
2. **Criar metodologia para extração de informações do pacote e composição da base de conhecimento:** a partir dos pacotes que chegam no sistema composto por roteadores, é necessário extrair as informações do mesmo pacote que chega nos roteadores e utilizá-las para criação da base de conhecimento. Busca-se nesta metodologia uma forma rápida e precisa para criação da base de dados, de maneira que a mesma não possua instâncias que possam afetar o processo de classificação.
3. **Analisar de maneira comparativa as abordagens propostas neste trabalho com a abordagem convencional:** as abordagens propostas neste trabalho se baseiam na utilização do roteadores com antenas direcionais de maneiras diferentes: a primeira abordagem foca na utilização destes roteadores espalhados pelo ambiente de localização; já a segunda abordagem foca na utilização destes roteadores a partir da junção dos mesmos, onde cada um estará apontando para direções diferentes. A abordagem convencional com utilização de roteadores omnidirecionais espalhados pelo ambiente de localização terá apenas caráter comparativo para a análise dos aspectos de cada uma das abordagens propostas neste trabalho.

4. **Analisar de maneira comparativa as duas vertentes de localização: localização precisa e localização contextual:** com base nas características de cada ambiente de localização utilizado nos experimentos deste trabalho, será possível analisar as características de atuação e desempenho do sistema de localização proposto a partir de duas vertentes de localização específicas, onde a primeira é a Localização Precisa, utilizada para indicar posição exata de um determinado dispositivo em um ambiente, e a Localização Contextual, onde é possível estimar a área de contexto do ambiente de localização onde um determinado dispositivo está inserido.

## 1.2 Organização da Dissertação

Esta proposta está dividida em 5 capítulos. No capítulo 2, é apresentada a fundamentação teórica deste trabalho, além da definição das características de Localização WiFi e seus componentes. Outros pontos importantes deste capítulo são a contextualização de técnica de classificação  $kNN$  e Naive Bayes, a classificação dos sistemas de localização e as características das antenas omnidirecionais e direcionais.

No capítulo 3, são apresentados trabalhos relacionados ao sistema de localização apresentado nesta proposta, assim também como tecnologias e características do objetivo deste sistema.

No capítulo 4, são apresentados todos os pontos concernentes a metodologia do trabalho como características do sistema proposto neste trabalho e as abordagens utilizadas.

No capítulo 5 são apresentadas as especificações dos experimentos, dos ambientes, equipamentos e softwares utilizados, e também os resultados obtidos neste trabalho juntamente com a análise dos mesmos.

Por fim, no capítulo 6 é apresentada a conclusão deste trabalho e a definição de trabalhos futuros.

## Capítulo 2

# Fundamentação Teórica

A Localização WiFi possui cada vez mais destaque entre os sistemas de Redes Sem Fio, principalmente por hoje em dia a utilização de equipamentos sem fio ser comum em diversos tipos de ambiente, como escolas, empresas e áreas públicas. Os sistemas utilizados para Localização WiFi são geralmente compostos por roteadores omnidirecionais, e geralmente se baseiam em medidas de potência de sinal e caracterização do ambiente para estimativa de posicionamento.

Os roteadores omnidirecionais possuem uma área de cobertura de 360 graus e trabalham com tráfego de dados numa determinada frequência do espectro de sinal, parte esta definida como canal. Devido a tais fatores, os roteadores omnidirecionais podem sofrer e causar impactos em outros roteadores dispostos num determinado ambiente. Uma alternativa para solucionar este problema pode ser a utilização de roteadores direcionais que podem se dedicar à transmissão e recebimento de sinal em uma determinada direção, já que tais roteadores possuem uma cobertura entre 90 a 120 graus, e conseqüentemente, possuem maior alcance sobre o sinal e menos chance de afetar outros roteadores do ambiente.

Por isto, a proposta deste trabalho tem como objetivo a criação de um sistema de localização composto por roteadores direcionais e classificação de dados para estimativa de posicionamento, a partir de uma fase de pré-configuração, e também procura mostrar que este tipo de roteador pode ser utilizado como alternativa para localização em ambientes internos.

Nas seções seguintes serão apresentados os conceitos necessários para o melhor entendimento deste trabalho.

## 2.1 Localização em Redes Sem Fio

Os sistemas de localização em Redes sem Fio podem ser classificados em diferentes formas, conforme se pode ver na figura 2.1. As definições das classificações são feitas a seguir:



**Figura 2.1.** Classificação dos algoritmos de localização em Redes sem Fio.  
Fonte: Adaptado de Robles et al. (2010)

- Quanto à mobilidade
  - *estático*: quando o sistema é implantado em um ambiente sem mobilidade, como por exemplo, as Redes de Sensores Sem Fio;
  - *dinâmico*: quando o sistema possui como característica a mobilidade de seus nós, como por exemplo, as Redes WiFi;
- Quanto ao número de âncoras
  - *baseado em âncora*: quando há a necessidade de cálculo de posicionamento a partir de um ou mais nós âncoras;
  - *livre de âncora*: quando o sistema consegue operar sem a necessidade de nós âncoras, proporcionando então para seus nós alvos apenas um posicionamento relativo;

- Quanto ao posicionamento
  - *preciso*: permite estimar a posição exata de um dispositivo nos eixos  $x$  e  $y$  de um determinado ambiente;
  - *contextual*: permite estimar posição de um determinado dispositivo com base na análise do contexto do ambiente;
- Quanto ao cenário
  - *interno*: quando o sistema é desenvolvido para atuação em ambientes internos e sofre em alguns aspectos específicos como *shadowing* ou *multipath*;
  - *externo*: quando o sistema é desenvolvido para atuação em ambientes externos e também sofre em aspectos específicos como condição ambiental e interferência onde há prédios pela proximidade;
- Quanto à utilização de treinamento:
  - *baseado em treinamento*: depende de treinamento de dados para estimativa de localização de nós com posição desconhecida. Este método depende da criação de uma base dados com informações obtidas do ambiente de localização para que, em seguida, consiga estimar a posição de um nó a partir da fase de classificação de dados;
  - *livre de treinamento*: não depende de nenhum conhecimento prévio baseado em treinamento para estimativa de localização.
- Quanto à forma de cálculo
  - *distribuído*: quando os nós são capazes de calcular o posicionamento, seja seu ou de outro nó;
  - *centralizado*: quando todo cálculo de posicionamento dos nós do sistema são realizados por um nó central;
  - *baseado em cluster*: quando no sistema há a divisão de nós em *clusters*, onde em cada cluster há uma instância responsável pelo cálculo de posicionamento dos nós dispostos dentro do seu próprio cluster;
- Quanto à coordenada
  - *relativa*: quando não há necessidade de identificação do posicionamento específico, apenas a importância de saber a referência a partir de algum ponto específico;

- *absoluta*: quando algum sistema proporciona informação de posicionamento, como por exemplo, o GPS;
- Quanto ao tipo de cálculo para estimativa posicionamento:
  - *baseado em probabilidades*: utiliza cálculos de probabilidades para posicionamento;
  - *baseado em estatística*: utiliza cálculos estatísticos para posicionamento;
  - *baseado em geometria*: utiliza cálculos geométricos para posicionamento.

O problema da localização pode ser descrito da seguinte maneira: em um determinado ambiente, tem-se um conjunto de nós localizados em posições diferentes mas respeitando os limites e características de posicionamento no ambiente. Se dois nós conseguem enviar e receber informações um para o outro significa que ambos estão na área de alcance de sinal de cada um, podendo eles estabelecer comunicação entre si de maneira livre.

No contexto da localização de Redes sem Fio, a teoria define que se há comunicação do nó  $i$  com o nó  $j$  há também comunicação do nó  $j$  com o nó  $i$ . Contudo, na vida real tal condição muitas vezes não é respeitada já que muitos fatores podem influenciar no comportamento da comunicação entre os nós que compõem o sistema, sendo esses os fatores:

- *configurações de hardware diferentes entre nós*: devido à grande variedade de marcas de componentes para redes WiFi, muitas vezes a comunicação entre nós se torna incapaz já que assim como existem nós de qualidade superior e com capacidade de comunicação com qualquer outro nós, também existem nós de qualidade inferior e com impossibilidade de comunicação com alguns tipos de nós.
- *fatores gerados pelo ambiente e seus participantes*: tais fatores podem muitas vezes afetar a comunicação entre os nós dispostos em um ambiente. Estes fatores são: objetos estáticos no ambiente, movimentação de objetos e pessoas, condições climáticas (temperatura, umidade) e outros.
- *fatores relacionados aos aspectos das redes WiFi*: tais fatores são concernentes aos aspectos de sinal relacionados a comunicação entre os nós em redes WiFi como os seguintes aspectos definidos no trabalho de Heurtefeux & Valois (2012):
  - path loss*: perda da potência eletromagnética do sinal conforme sua propagação pelo espaço.

*fading*: desvio de atenuação sofrido pelo sinal, muitas vezes responsável pela interferência, amplificação e atenuação da potência do sinal.

*shadowing*: perda do sinal causada por obstáculos ou por qualquer outro fator que possa afetar o sinal transmitido para comunicação de dois nós.

Os pontos citados acima definem os vários aspectos existentes nos sistemas de redes sem fio. Tais pontos devem ser tratados nos componentes de um sistema de localização WiFi, de maneira que seja possível conseguir um bom desempenho no sistema independente das condições adversas do ambiente onde o mesmo está instalado.

Os sistemas de localização WiFi são compostos por alguns componentes necessários para se realizar a localização de um determinado nó cuja informação é desconhecida. Tais componentes são:

- *Estimativa de distâncias*: responsável pela estimativa de distância de um nó a partir das informações obtidas pelas suas características de sinal. Alguns dos métodos que realizam esta função são: *RSSI* (Received Signal Strength Indicator), *ToA* (Time of Arrival) e *AoA* (Angle of Arrival).
- *Cálculo de posicionamento*: responsável por calcular a posição de nós, cuja informação é desconhecida, através de dados disponíveis a partir do contexto do ambiente de redes WiFi. Alguns dos métodos que realizam esta função são: multilateração ou método de treinamento, este último mais conhecido como *fingerprint*.

A partir de tais componentes, qualquer sistema de localização WiFi é capaz de definir a posição de nós que não possuem tal informação. Tais nós são conhecidos como nós alvos, já que os mesmos são o principal interesse deste trabalho devido a problemática de identificação de posicionamento. Os nós alvos só serão localizados a partir das informações de posicionamento já conhecidas dos nós dispostos no ambiente em análise. Estes nós são conhecidos como nós âncoras. Em muitos sistemas, os nós alvos já localizados podem ser utilizados como parâmetros para localização de outros nós, passando a ser conhecidos como nós estabelecidos (Boukerche et al., 2007).

### 2.1.1 Estimativa de Distância

A estimativa de posicionamento possui uma grande importância para os sistemas de localização WiFi, pois tais informações obtidas entre dois nós compõem posteriormente o cálculo da posição do nó desejado.

Existem muitos métodos que são utilizados para obter estas informações para o sistema de localização, diferenciando entre eles características como custo, hardware e precisão. Alguns desses métodos são:

- *Received Signal Strength Indicator (RSSI)*: utilizado para estimativa de distância entre dois nós a partir da potência dos sinais trocados entre eles, que oferece características de análise de dados devido ao seu desvanecimento conforme a propagação pelo espaço. Este método tem sido cada vez mais utilizado em trabalhos relacionados a localização WiFi (Fang et al., 2010; Barsocchi et al., 2009; Ding et al., 2011) devido ao seu baixo custo agregado de utilização, já que tal método pode ser utilizado com qualquer equipamento capaz de receber ou enviar sinais. Uma das desvantagem deste método é a facilidade com que o sinal pode ser afetado com obstruções ou qualquer outro fator externo ao sistema de localização, afetando conseqüentemente os seus valores.
- *Time of Arrival (ToA)*: utilizado para estimativa de distância entre dois nós a partir de medidas de tempo geralmente baseadas no intervalo para a chegada do sinal em um determinado nó. Para que seja possível a utilização deste método, é necessário que todo o sistema de localização esteja bem sincronizado no que diz respeito a medidas de tempo. Uma das desvantagens deste método é o alto custo agregado aos equipamentos utilizados para a obtenção de informação de AoA.
- *Angle of Arrival (AoA)*: utilizado para estimativa de ângulo entre nós. Geralmente este método utiliza três ou mais antenas direcionais dispostas em posições e ângulos diferentes, de maneira que seja possível especificar o impacto causado por um sinal obtido pelo conjunto de antenas. Apesar do custo agregado às antenas de qualidade capazes de gerar bons resultados a localização WiFi, o método AoA pode gerar ótimos resultados na estimativa de ângulo, e trabalhos como o de Hui et al. (2007) demonstram a importância e benefícios encontrados na utilização de tal método.

### 2.1.2 Cálculo de Posicionamento

O cálculo de posicionamento é realizado a partir das informações obtidas na etapa de estimativa de posicionamento com um ou mais métodos dos especificados na subseção acima. A única diferença entre os métodos de cálculo de posicionamento é a maneira como eles utilizam as estimativas para compor as informações de posicionamento. Tais métodos serão definidos nas seções a seguir.

### 2.1.2.1 Multilateração

Este método utiliza as estimativas de distância de um nó geradas para três ou mais outros nós próximos a ele, considerando-as como raios para composição de círculos. A partir da interseção destes círculos é possível calcular a posição do nó com posicionamento desconhecido. A multilateração é um método que proporciona rapidez ao processo de cálculo de posicionamento, mas muitas vezes não retrata todas as condições necessárias para a avaliação do ambiente em que se deseja realizar localização de nós, como por exemplo obstáculos ou movimentação de pessoas e seus impactos gerados sobre o sinal.

### 2.1.2.2 Treinamento

Também conhecido como *análise de cena*, este método é comumente utilizado em ambientes internos para obter, armazenar e analisar características de sinal no ambiente em que os nós do sistema estão dispostos. Ele é formado por duas fases bem distintas:

1. a fase *offline*, responsável por coletar informações e criar uma base de dados com as estimativas de posicionamento a partir de medidas, formando assim um mapa de estimativas do ambiente em que os nós estão dispostos. Tais características podem ser valores de RSSI medidos pelo sistema ou qualquer outra informação medida pelos nós alvos localizados em posições conhecidas.
2. a fase *online*, onde as informações obtidas no ambiente e armazenadas na base de dados são utilizadas para estimar a posição dos nós alvos. Desta forma, pode-se analisar que esta técnica tem como objetivo identificar um conjunto de valores que se equiparam as medidas obtidas na fase online e utilizar as posições associadas a estas características para realizar a estimação do posicionamento dos nós. Tal técnica considera que as características do ambiente na fase online são as mesmas na fase offline, o que nem sempre é válido, precisando sempre gerar uma atualização na base de dados, etapa esta conhecida como *calibração*. As definições de posicionamento para este método podem estar relacionadas com a caracterização de classes para os dados de estimativas, etapa esta conhecida como *Classificação de Dados*, que será melhor especificada na seção seguinte.

## 2.2 Classificação de Dados

A Classificação de Dados é o processo de designação de uma classe para um conjunto de atributos que representam uma entrada ou uma instância do problema em questão.

Este processo depende inteiramente de um conhecimento prévio sobre o problema. Este conhecimento pode ser descrito como uma base de dados com várias instâncias que definem em diversas formas as características do problema. Cada instância do problema é composta por atributos que descrevem o problema seguidos de uma classe que define a composição de tais atributos.

A fase de avaliação de um sistema geralmente depende de informações para serem utilizadas como conhecimento e para serem exemplos de testes. O conjunto de instâncias para teste é importante para a avaliação pois ele é utilizado como forma de avaliar a precisão do algoritmo de classificação em sua configuração. Vale ressaltar que tanto a base de dados para conhecimento como para teste devem descrever o problema de maneira mais próxima da realidade possível, pois, caso isso não aconteça, o sistema pode ser levado a *underfitting* ou *overfitting*, condições estas onde a base de dados descreve de maneira pobre ou de maneira muito específica o problema desejado, respectivamente.

No caso de sistemas de localização baseados em treinamento, vários pacotes são enviados de posições específicas do ambiente para o sistema com roteadores com antenas direcionais. Os valores de RSSI obtidos pelos roteadores são utilizados como atributos para a criação da base e a identificação da posição de onde os pacotes foram enviados define a classe dos atributos.

A utilização de classificação de dados em trabalhos de localização tem se tornado cada vez mais frequente. Podemos encontrar no trabalho de Lee & Lampe (2011), por exemplo, um sistema para localização de nós utilizando classificação sobre valores de RSSI. O autor também apresenta uma comparação entre os resultados dos experimentos de classificação a partir dos algoritmos *Support Vector Machine* (SVM), *Simple Gaussian Classifier* (SGC) e *k-Nearest Neighbor* (kNN), algoritmo este também utilizado em (Bahl & Padmanabhan, 2000). No trabalho de Xu et al. (2012) também podemos encontrar um método de treinamento que utiliza abordagens de classificação probabilística baseadas em análise discriminante.

Este trabalho utiliza dois algoritmos de classificação:

- algoritmo *k-Nearest Neighbor* - *kNN*: bastante utilizado na literatura (Lee & Lampe, 2011; Bahl & Padmanabhan, 2000) devido a suas condições de funcionamento. Este algoritmo foi escolhido devido as características da classificação de dados utilizada neste trabalho. Este algoritmo analisa a classe que mais ocorre entre os vizinhos mais próximos a uma instância para realizar a sua classificação. Seu comportamento favorece ao problema apresentado neste trabalho já que um nó com posição desconhecida terá várias exemplos da mesma classe ao seu redor,

umentando a precisão da classificação;

- algoritmo *Naive Bayes*: algoritmo probabilístico que consegue identificar quais atributos de um problema podem melhor estimar uma instância. A partir deste algoritmo, é possível analisar o impacto da organização dos roteadores no ambiente de localização para a classificação da posição de um determinado nó e, conseqüentemente, definir como organizar os roteadores para obter mais precisão na classificação de posição de nós.

Estes dois algoritmos são baseados em aprendizagem supervisionada, onde eles possuem conhecimento do ambiente a partir de exemplos rotulados, e assim tais algoritmos conseguem estabelecer um mapeamento entre entrada e saída.

### 2.2.1 Algoritmo *k-Nearest Neighbor* - *kNN*

É comum considerar que proximidade entre as características das instâncias pertencentes a uma classe específica de um problema variem pouco. Desta forma, é possível classificar novas instâncias de um problema com base no conhecimento das características que melhor definem uma classe e assim analisar a proximidade existente entre as características da nova instância e das classes existentes no problema. Este comportamento define o funcionamento do algoritmo *kNN*.

Seja uma base de dados  $X = \{x_1, x_2, \dots, x_{(m-1)}, x_m\}$  formada por várias instâncias  $x_i$ , onde  $x_i = \{a_1, a_2, \dots, a_{(n-1)}, a_n, c_i\}$ . Cada  $a_i$  representa uma característica da instância e  $c_i$  a sua classe. Para classificar uma nova instância  $y$ , o algoritmo *kNN* analisa todas as instâncias dispersas no espaço de características das instâncias de treino que estão próximas de  $y$ , e a rotula com a classe que mais aparece entre os seus  $k$  vizinhos mais próximos.

O algoritmo *kNN* possui uma fase *offline* que se dedica em obter informações para a construção da base de dados, e uma fase *online* que identifica a classe que melhor se adequa a instância em análise a partir dos dados obtidos na fase *offline*. Na figura 2.2 é possível analisar um exemplo de classificação pelo método *kNN*.

O conhecimento para classificação nesta técnica depende diretamente das instâncias existentes na base de dados, que são utilizadas para identificação dos vizinhos mais próximos. O armazenamento das características e das classes consiste na fase de *offline* do algoritmo *kNN*. Sua fase *online* é dependente do valor de  $k$  e das características da instância cuja classe se deseja estimar.

A distância métrica entre as instâncias de um espaço de características pode ser definida conforme o tipo das características que as definem. A distância *Euclidiana*, por

exemplo, é utilizada para instâncias com características contínuas. Considerando que  $p = (p_1, p_2, p_3, \dots, p_{n-1}, p_n)$  e  $q = (q_1, q_2, q_3, \dots, q_{n-1}, q_n)$  sejam dois pontos, a distância euclidiana  $d_{euclidiana}$  entre eles pode ser definida por:

$$d_{euclidiana} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (2.1)$$

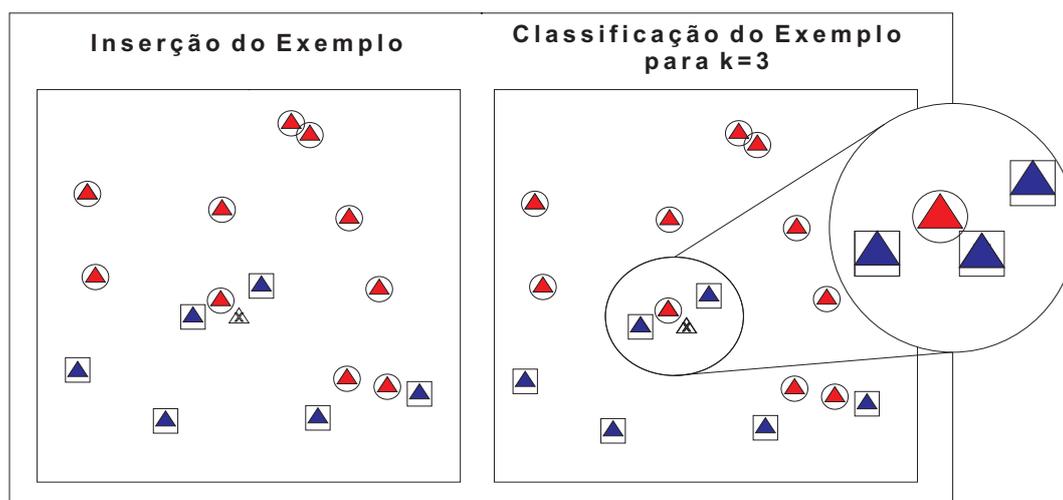
As características das instâncias de sinal utilizadas neste trabalho são discretas e por isto foi utilizado outro tipo de distância conhecida como distância *Manhattan*, que se baseia na soma das diferenças absolutas entre as suas coordenadas cartesianas. Considerando que  $p = (p_1, p_2, p_3, \dots, p_{n-1}, p_n)$  e  $q = (q_1, q_2, q_3, \dots, q_{n-1}, q_n)$  sejam dois pontos, a distância *Manhattan*  $d_{manhattan}$  entre eles pode ser definida por:

$$d_{manhattan} = \sum_{i=1}^n |p_i - q_i| \quad (2.2)$$

## 2.2.2 Algoritmo Naive Bayes

O algoritmo de classificação *Naive Bayes* mede o impacto que um atributo, vinculado ou não ao restante dos outros atributos do ambiente em análise, gera na estimativa de uma classe do problema. Assim, é possível analisar as características das instâncias de um problema desconsiderando o relacionamento que pode haver entre elas.

Um cenário que pode exemplificar a utilização deste algoritmo é a identificação de tipos de veículos a partir de instâncias que descrevem suas características. Consideremos que a análise de uma instância deve dizer se a classe do veículo é carro ou



**Figura 2.2.** Exemplo de classificação pelo algoritmo kNN com  $k=3$ .

caminhão, e cada instância seja formada pelos seguintes atributos: peso, número de rodas, motor e número de portas. O algoritmo Naive Bayes analisa como cada atributo deste problema contribui para a classificação de uma instância como carro ou como caminhão.

O modelo de probabilidade utilizado neste algoritmo, responsável por classificar os exemplos utilizados pelo algoritmo Naive Bayes, pode ser descrito a partir da seguinte modelo condicional:

$$p(w_i|x_j) = \frac{p(x_j|w_i)P(w_i)}{p(x_j)}, \quad (2.3)$$

onde  $p(w_i|x_j)$  é a probabilidade *a posteriori*,  $p(x_j|w_i)$  a verossimilhança,  $P(w_i)$  a probabilidade *a priori*,  $p(x_j)$  a evidência,  $w_i$  um rótulo de classe e  $x_j$  uma característica do problema.

Devido às características do seu funcionamento, o algoritmo *Naive Bayes* não precisa de uma grande quantidade de dados para montar seu conhecimento de classificação. Isto é possível pelo fato do algoritmo analisar apenas o impacto de cada característica para uma determinada classe, e não toda a matriz de covariância.

## 2.3 Antenas

A antena é um dispositivo que transforma energia elétrica em ondas de rádio e também faz o inverso, e para tal função, depende de um rádio transmissor ou um rádio receptor. Tal dispositivo é responsável por muitos serviços utilizados pelas pessoas nos seus dias-a-dias como transmissão de sinal de TV, de rádio, comunicação celular e em redes WiFi.

Tais dispositivos podem ser classificados de acordo com suas características de comportamento e que devem ser levadas em consideração no momento em que sua utilização é necessária. São elas:

- *ganho*: parâmetro utilizado para medir quão direto é padrão de radiação de uma antena. Dessa forma, quanto maior for o ganho de uma antena, maior será sua radiação para uma direção específica. Este fenômeno não faz da antena um agente ativo sobre pois ela não adiciona nenhuma potência ao sinal, e sim apenas redistribui o sinal para ter maior potência em uma direção específica. É importante ressaltar que apesar de antenas de alto ganho terem grandes vantagens para a qualidade do sinal, ela também se limita a uma direção específica, enquanto

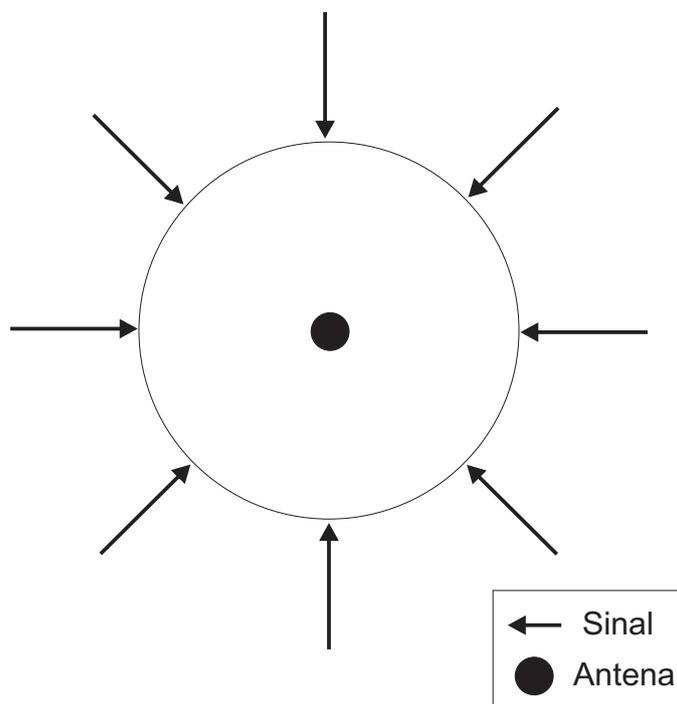
que as antenas que possuem ganho mais baixo podem ter um alcance menor mas consegue atender a mais direções.

- *abertura*: parâmetro que define a porção de energia eletromagnética descrita em área equivalente que a antena oferece para seus receptores.
- *padrão de radiação*: representa a intensidade do campo relativo das ondas de rádio emitidas por uma antena em diferentes. Geralmente tal informação é representada através de um gráfico que apresenta as características do campo eletromagnético em volta da antena.
- *eficiência*: define a relação existente entre a potência irradiada e potência recebida pelos receptores de sinal.
- *polarização*: é a referência da orientação do campo eletromagnético das ondas de rádio geradas por uma antena com respeito a superfície da terra. Tal fator é determinado pela estrutura física da antena e sua orientação.

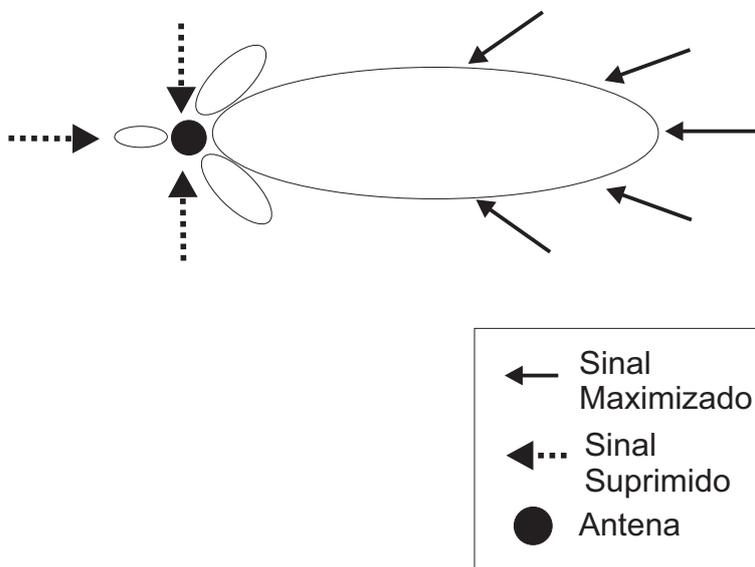
A partir de tais características das antenas é que o sinal eletromagnético é gerado e recebido pelas antenas. O sinal eletromagnético nada mais é do que uma oscilação elétrica e magnética. A partir do sinal é que as informações são enviadas e recebidas entre os dispositivos que possuem antenas.

As antenas podem transmitir e receber ondas de rádio em todas as direções de maneira igualitária, sendo este tipo de antena chamada de omnidirecionais. A figura 2.3 apresenta o padrão de radiação das antenas omnidirecionais, com o seu funcionamento exemplificado a partir do recebimento de alguns sinais. Os sinais recebidos e enviados neste tipo de antena são igualitários, independente das suas direções.

As antenas também podem transmitir e receber ondas de rádio de uma direção específica (Umehira & Ohtomo, 2010), sendo este tipo de antena chamada de direcionais. A figura 2.4 apresenta o padrão de radiação das antenas direcionais, com o seu funcionamento exemplificado a partir do recebimento de sinais incidentes e fora do direcionamento da antena. Esta figura também apresenta o lóbulo principal indicado pela direção da antena, com alguns lóbulos laterais que são perdas de energia bastante comuns nestes tipos de antena. Os sinais que são recebidos e enviados na direção do lóbulo principal são maximizados, enquanto os sinais recebidos e enviados na direção dos lóbulos laterais são suprimidos.



**Figura 2.3.** Padrão de Radiação das Antenas Omnidirecionais.



**Figura 2.4.** Padrão de Radiação das Antenas Direcionais.

## 2.4 Características das Antenas Omnidirecionais e Direcionais

Todo sistema sem fio que utiliza antenas possui 3 propriedades específicas que devem ser levadas em consideração para a sua execução:

- *ganho do sinal*: medida que define a melhoria de potência para transmissão e recepção de informações pelas antenas;
- *direção da antena*: formato do campo de transmissão do sinal em uma determinada antena;
- *polarização*: especificação da orientação de oscilação das ondas para uma determinada antena.

Cada uma dessas propriedades devem ser bem ajustadas de maneira que cada uma delas agregue a antena a melhor utilização e desempenho.

A antena é responsável em transformar energia eletromagnética guiada em irradiada. Ela é utilizada em sistemas para direcionar a energia recebida pelo transmissor de sinal, e de maneira nenhuma ela pode adicionar maior potência ao sinal por ser apenas um elemento passivo. A energia aplicada sobre um sinal está relacionado com o seu alcance e potência. Dessa forma, quanto maior for a direção que uma antena deve atender, menor é a energia que será aplicada sobre o sinal. Desta forma, as antenas que focam em apenas uma direção conseguem aplicar uma energia maior sobre o sinal, fazendo com que o mesmo tenha maior alcance, enquanto as antenas que focam em todas as direções, aplicam menor energia sobre o sinal, fazendo com que o mesmo tenha menor alcance.

As antenas podem ser classificadas em omnidirecionais e direcionais. As antenas omnidirecionais são verticalmente polarizadas e provêm sinal com um padrão de radiação de 360 graus horizontais, sendo então utilizadas para sistemas onde é necessário atender o maior número de nós existentes em qualquer direção de determinada área de atuação. Já as antenas direcionais concentram toda a sua energia para uma direção específica, e assim, seu ganho aumenta juntamente com o seu alcance de cobertura devido a sua diminuição no ângulo de cobertura.

Cada tipo de antena possui suas vantagens e desvantagens:

- Antenas Omnidirecionais
  - Vantagens: fácil instalação; podem cobrir uma grande área em um ambiente interno;

- Desvantagens: possui um sério problema de cobertura de sinal abaixo dos roteadores;
- Antenas Direcionais
  - Vantagens: consegue atender uma cobertura maior de sinal; pode ser utilizada para cobertura baseada em *Line-of-Sight* (LOS), onde a transmissão é feita em linha reta;
  - Desvantagens: não pode ser utilizadas para cobrir grandes áreas.

## 2.5 Conclusão

Neste capítulo foi apresentado a Localização WiFi como um dos problemas mais conhecidos dentro da área de Sistemas de Redes Sem Fio. Foi feita também a classificação dos sistemas de Localização WiFi no que diz respeito a:

- Mobilidade: móvel ou estático;
- Âncora: baseado ou livre;
- Posicionamento: contextual ou preciso;
- Cenário: interno ou externo;
- Treinamento: baseado ou livre;
- Cálculo: distribuído, centralizado ou baseado em cluster;
- Coordenadas: relativas ou absolutas;
- Cálculo para Estimativa de Posicionamento: baseado em probabilidade, estatística ou geometria.

O sistema apresentado neste trabalho pode ser caracterizado sobre os seguintes aspectos de classificação citados acima:

- pode ser utilizado para trabalhar com elementos móveis ou estáticos;
- não depende de âncora para estimativa de posicionamento;
- utilizado em ambiente contextual e preciso;
- seu funcionamento é próprio para ambientes internos;

- depende de treinamento para estimativa de posicionamento;
- seu cálculo de posicionamento é feito de maneira centralizada;
- seu mapeamento de posicionamento para treinamento e estimativa se dá por coordenadas absolutas;
- seu cálculo para estimativa é baseado em probabilidade.

Os sistemas de localização são compostos pelos componentes de estimativa de distância e de cálculo de posicionamento. Para estimativa de distância é possível encontrar inúmeros métodos na literatura como RSSI (indicador de potência de sinal recebido), ToA (tempo de chegada) e AoA (ângulo de chegada). Já com relação ao cálculo de posicionamento, pode-se a multilateração e o método baseado em treinamento para estimativa de posicionamento, que se baseia em classificação de dados como forma de se definir uma classe para uma instância com base em suas características.

Foram utilizados dois algoritmos de classificação de dados separadamente como núcleo de cálculo de posicionamento do sistema proposto neste trabalho: *kNN* (k-Nearest Neighbor) e Naive Bayes.

Este capítulo também apresentou a funcionalidade das antenas e suas características. A caracterização mais importante para este trabalho no que diz respeito às antenas é que elas podem ser omnidirecionais, quando recebem e enviam sinais em todas as direções, ou direcionais, quando recebem e enviam sinais em uma direção específica.

# Capítulo 3

## Trabalhos Relacionados

O problema da localização em Redes Sem Fio tem ganhado cada vez mais atenção. E esta dissertação apresenta uma solução baseada em três frentes diferentes:

- *Localização Baseada em RSSI*: abrange todos os sistemas de localização capazes de estimar a posição de dispositivos localizados em um ambiente através da potência de sinal (RSSI) extraída como informação a partir de pacotes enviados pelos dispositivos e recebidos pelos sistemas. Os *surveys* Boukerche et al. (2007); Honkavirta et al. (2009) apresentam diversos aspectos relacionados aos sistemas que trabalham com este tipo de localização como características de sinal, fatores que causam interferência e detalhes sobre abordagens para estimativa de posicionamento;
- *Localização Baseada em Antenas Direcionais*: abrange todos os sistemas de localização que fazem uso das características do sinal nas antenas direcionais para estimar posição de dispositivos localizados em um determinado ambiente. O trabalho Lu et al. (2010) apresenta as características dos sistemas de localização que utilizam antenas omnidirecionais e os que utilizam antenas direcionais. Além disso, o trabalho citado discute como as antenas direcionais podem melhorar o desempenho dos sistemas de localização;
- *Localização Baseada em Treinamento*: abrange todos os sistemas de localização que se baseiam em classificação de dados a partir de treinamento para estimar posição de dispositivos localizados em um ambiente. O trabalho Gunawan et al. (2012) estuda diferentes meios para classificação de dados para localização baseada em RSSI.

A tabela 3.1 apresenta uma comparação entre os trabalhos relacionados à proposta desta dissertação no que diz respeito aos sistemas de localização que utilizam RSSI, treinamento e antenas direcionais.

**Tabela 3.1.** Comparação dos trabalhos relacionados a Localização WiFi

	<i>Utiliza RSSI</i>	<i>Utiliza Treinamento</i>	<i>Utiliza Antenas Direcionais</i>
<i>Heurtefeux &amp; Valois (2012)</i>	X		
<i>Bahl &amp; Padmanabhan (2000)</i>	X	X	
<i>Ding et al. (2011)</i>	X	X	
<i>Sugano (2006)</i>	X		
<i>Kawauchi et al. (2009)</i>	X		X
<i>Ou (2011)</i>			X
<i>Lu et al. (2010)</i>	X		X
<i>Chaurasiya et al. (2009)</i>	X		X
<i>Yang et al. (2005)</i>	X		X
<i>Robles et al. (2010)</i>	X	X	
<i>Le et al. (2011)</i>	X	X	
<i>Meng et al. (2011)</i>	X	X	
<i>Gogolak et al. (2011)</i>	X	X	
<i>Menendez et al. (2011)</i>	X	X	
<i>Balaguer et al. (2012)</i>	X	X	

### 3.1 Localização Baseada em RSSI

Apesar da variação e da interferência que o sinal é passível de sofrer, a utilização do RSSI tem se tornado cada vez mais frequente devido à sua disponibilidade em todos os dispositivos móveis existentes no mercado, sem necessidade de hardware extra, e também devido à grande riqueza de informações que podem ser obtidas e avaliadas a partir do seu aspecto e propriedades. O trabalho desenvolvido por Heurtefeux & Valois (2012) apresenta várias informações relacionadas ao RSSI, avaliadas de acordo com as propriedades de topologia e com o ambiente criado pela rádio-transmissão. Além disso, outros aspectos também importantes para a sua utilização são apresentados como a correlação entre RSSI e distância, assimetria, isotropia, dinâmica do sinal, *path-loss*, *fading*, *shadowing*. O trabalho apresentado nesta dissertação utiliza o RSSI de maneira

semelhante, com esta informação chegando nos roteadores com antenas direcionais utilizados pelo sistema. A partir disto, a posição do nó que enviou o sinal para o sistema é calculada.

O trabalho de Bahl & Padmanabhan (2000) apresenta um sistema de localização baseado em rádio frequência chamado RADAR. Este sistema tem como objetivo a localização e o rastreamento de pessoas em um determinado ambiente. Para isso, tal sistema utiliza os valores de RSSI obtidos em múltiplas estações bases posicionadas em lugares estratégicos para proporcionar uma melhor cobertura do ambiente em análise. O autor propôs dois tipos de abordagens para determinação da localização dos usuários. A primeira depende completamente das medidas de potência de sinal realizadas pelos pontos de acessos dispostos no ambiente durante a fase *offline* do sistema. Os experimentos realizados sobre tal abordagem demonstraram que muitos aspectos podem afetar a precisão da localização dos usuários durante a fase *online*, sendo eles: orientação do usuário durante o envio de sinal, número de informações obtidas na fase *offline* que compõem a base de dados, números de vizinhos utilizados para cálculo de média de coordenadas e números de pontos de análise em ambiente para composição da base de dados. Outra abordagem apresentada pelo autor neste mesmo trabalho é a modelagem de propagação do sinal considerando aspectos do ambiente como atenuação de paredes e de pisos.

Os experimentos do trabalho citado foram realizados em um piso de uma construção de três andares, onde neste andar haviam três estações bases. Sua base de dados criada em experimento consiste de informações de sinal em função da posição de um determinado usuário equipado com um nó móvel. Tais informações são utilizadas tanto para construção e validação de modelos de propagação de sinal durante a fase *offline* de análise, como para a localização em tempo real do usuário. A partir das informações obtidas durante a fase *offline*, calcula-se as médias, medianas e desvio padrão dos valores de potência de sinal. Para a composição da base de dados, foram obtidas informações de 70 pontos diferentes e, em cada um deles, o sinal era enviado em 4 direções diferentes, totalizando 280 combinações de localização e orientação. Todas estas informações foram necessárias para melhor investigação das características do RSSI com relação aos obstáculos do ambiente e do impacto do usuário sobre o sinal enviado. Nesta dissertação, o impacto da estrutura do ambiente sobre o sistema de localização e a forma de organizar o sistema para obter melhores resultados dentro de tal estrutura também foi avaliado.

Em Ding et al. (2011) podemos encontrar um modelo de erro de localização baseado em um grande número de experimentos. Este modelo possui os seguintes aspectos:

- um amplificador de potência de sinal com o objetivo de evitar a atenuação e a interferência causada pelo RSSI;
- um algoritmo para tratamento de medição de distância;
- melhoria no algoritmo Euclideano com o objetivo de utilizar menos nós âncoras.

O algoritmo para medição de distância utilizado no trabalho citado se divide em algumas fases: a primeira fase, determinada como pré-configuração, armazena valores de RSSI em uma tabela em ordem decrescente para cada distância analisada; a segunda etapa, determinada como a fase de localização, utiliza as tabelas com os valores de RSSI obtidos em cada distância para cálculo de posicionamento onde é analisado qual delas possui menor desvio de valores de RSSI a partir do sinal obtido para comparação. Uma melhoria do algoritmo Euclideano é proposta neste mesmo trabalho para a localização quando se possuem poucos nós âncoras. Tal trabalho conseguiu proporcionar um erro máximo de localização de 2 metros a partir de medições sobre diversos pontos da área de localização em análise. As informações utilizadas neste trabalho para compor a estimativa de posicionamento foi utilizada na pesquisa apresentada nesta dissertação, de maneira que através de uma fase de pré-configuração vários pacotes foram enviados de posições específicas. Assim, o valor do RSSI juntamente com a informação da posição de envio foram utilizados para compor as informações de cada instância da base de dados utilizada em nosso trabalho para estimativa de posicionamento.

O trabalho de Sugano (2006) apresenta um sistema de localização onde cada nó estático pertencente a uma área de observação está sujeito a receber pacotes de dispositivos cujas posições são desconhecidas. Quando três ou mais nós recebem tais informações, eles obtêm o valor de RSSI do pacote recebido, transformam esta informação em distância e a enviam para um nó central, responsável por calcular a posição do dispositivo. Um dos focos deste trabalho é o controle de troca de informações com o intuito de diminuir o número de dados coletados, aumentando o tempo de vida dos nós dispostos na rede. O método utilizado neste trabalho para a estimativa da distância através do RSSI é baseado na estimativa por verossimilhança, chamado de mínimos quadrados. Neste método, é necessário o recebimento de um determinado pacote em pelo menos três nós para geração de uma estimativa mais confiável. Em seus experimentos, foi possível obter um erro máximo de 1,5-2m a partir de uma densidade de 0,27 nós por metros quadrados, mas como os nós alvos nos seus experimentos estavam dispostos nesta mesma faixa de distância, é possível chegar à conclusão de que talvez em alguns momentos o erro de localização pode ser de 100%. O sistema utilizado nesta dissertação também depende do recebimento de um pacote enviado em pelo menos três

nós responsáveis em monitorar os pacotes enviados por cada dispositivo cuja posição é desconhecida, e desta forma, é possível analisar as características de sinal de cada pacote enviado para compor a informação de posicionamento do mesmo.

## 3.2 Localização Baseada em Antenas Direcionais

O trabalho de Ou (2011) apresenta um esquema de localização *range-free* para redes de sensores sem fio usando nós âncoras estáticos equipados com quatro antenas direcionais. Cada nó âncora utilizado nesta abordagem determina sua posição através de informações de GPS e propaga suas coordenadas pelas quatro antenas direcionais conforme sua movimentação pela rede. Os nós alvos espalhados na rede são capazes de detectar tais informações dos nós âncoras através de suas antenas omnidirecionais, e com estes dados é possível calcular suas próprias posições a partir de um esquema de processamento matemático simples. Tal trabalho apresenta alguns dos benefícios da utilização de antenas direcionais para localização:

- concentração do poder de sinal completamente na direção do receptor;
- extensão do alcance de transmissão, ao contrário das antenas omnidirecionais que emitem e recebem a mesma energia em todas as direções;
- não necessita de hardware extra para medição específica;
- evita a necessidade de comunicação entre nós.

Os benefícios listados acima motivaram a utilização de antenas direcionais no sistema de localização apresentado nesta dissertação, principalmente pela maior estabilidade e alcance proporcionado ao sistema devido a concentração do poder de sinal em uma direção específica.

As mensagens enviada pelos nós âncoras no trabalho citado são compostas pelo valor das suas coordenadas  $x$  ou  $y$ , com um campo identificador indicando de qual coordenada se trata tal valor. Dessa forma, cada nó alvo que recebia tais mensagens criava uma lista com as coordenadas enquanto os nós âncoras percorrem o ambiente onde o sistema está implantado. A partir dessas listas, é possível calcular medianas que serão consideradas como a posição do nó alvo.

Durante a definição metodológica e fundamental do problema, Ou (2011) utiliza a definição encontrada em Ramanathan (2001), de que sistemas que utilizam antenas direcionais podem ser classificados em *switched beam*, onde as mensagens são propagadas ou recebidas pelas antenas em direções específicas de acordo com um padrão

determinado, ou *steered beam*, onde a direção e a largura do feixe de sinal podem ser alterados durante a transmissão e recepção pelo sistema. O sistema apresentado nesta dissertação pode ser classificado como *switched beam* pelo fato de o sistema de localização definido utiliza apenas um padrão para envio em recebimento de pacotes.

O trabalho de Lu et al. (2010) é destinado a estimativa da probabilidade de detecção de informações maliciosas a partir do estudo comparativo entre as antenas omnidirecionais e direcionais. Tal trabalho foi desenvolvido a partir do contexto de redes militares ou não suficientemente confiáveis, onde muitas vezes é difícil de se manter a segurança e o anonimato dos nós transmissores. A comunicação existente nestes sistemas de redes sem fio são suscetíveis muitas vezes a análise maliciosa de tráfico de dados devido a sua natureza de propagação de rádio transmissão. Nós hostis dentro da área de atuação destas redes tem como objetivo observá-las para descobrir quando e onde os nós transmitem informações. Dessa forma, este trabalho, apesar não se dedicar à localização, foca na estimativa do impacto da transmissão por antenas direcionais analisando a probabilidade do transmissor ser detectado por um adversário.

Este trabalho cita uma característica interessante relacionada às antenas omnidirecionais e direcionais: as antenas omnidirecionais radiam e recebem sinais igualmente em todas as direções, e por isto somente uma pequena parte de toda a energia empregada no ambiente por tais antenas é capaz de alcançar o receptor. Já as antenas direcionais não possuem tal desvantagem, pois elas radiam e recebem a potência do sinal em apenas uma direção, amplificando o nível de potência do sinal na direção determinada, formando um feixe direcional para o receptor. Com a utilização das antenas direcionais é possível alcançar uma distância maior para envio do sinal utilizando menos poder de transmissão. As características das antenas citadas acima motivam a utilização de antenas direcionais na nossa pesquisa devido ao seu alcance que possibilita a criação de um ambiente de localização utilizando poucos roteadores e além disse com uma qualidade maior de sinal devido ao melhor nível de sinal.

É óbvio que, como as antenas direcionais focam em apenas uma pequena área de transmissão e envio, um sistema com tais antenas se torna menos provável de interceptação de sinal por adversário, e os experimentos realizados no trabalho de Lu et al. (2010) chegaram ao resultado de 90% de redução na probabilidade de detecção de informações com transmissão a partir de antenas direcionais.

Já em Chaurasiya et al. (2009) é proposto um algoritmo de localização em redes sensores que utiliza antenas direcionais em cada nó pertencente ao sistema. Através da informação de localização do nó, é possível movimentar as antenas para que as mesmas se direcionem para enviar os pacotes desejados para seus destinos específicos. Cada nó disposto neste sistema coleta informações de outros nós e as envia para uma unidade de

processamento centralizada que calcula a localização de todos os nós que não conhecem suas posições. As informações relacionadas aos seus vizinhos que cada nó transmite são a distância e ângulo de envio relativos a cada um deles. Este processo tem início a partir de um nó âncora que envia a informação de sua localização para todos os seus vizinhos, e a partir desta e do ângulo e distância de chegada do sinal os seus nós vizinhos são capazes de identificar suas posições, e conseqüentemente, enviá-las para os seus outros vizinhos.

O diferencial na utilização das antenas direcionais no trabalho citado é que a localização de cada nó âncora é enviada pelas antenas juntamente com o identificador de outra antena, enviado para indicação de direcionamento. No momento que vizinho recebe tal informação ele se torna apto a calcular sua localização aproximada a partir das informações de ângulo e distância de chegada e, posteriormente, poderá propagar tais dados para cálculo do posicionamento de outros nós. Os nós alvos quando recebem a localização de outros nós alvos definem um conjunto de coordenadas que irá compor os pontos extremos da região utilizados para localização dos mesmos. Outro ponto interessante neste trabalho é que, como um nó âncora pode recalculer seu posicionamento não dependendo das coordenadas de GPS, eles também podem calcular um erro estimado de localização associado ao ambiente em que o sistema está disposto e propagá-lo para os nós alvos, proporcionando uma estimativa mais precisa de suas posições.

O trabalho citado ajudou a definir a forma de cálculo de posicionamento dos nós no sistema proposto nesta dissertação, onde tanto os valores de RSSI como os seus identificadores das posições que cada nó enviou os pacotes são passados para uma central responsável por estimar a posição de um nó com base nestas informações e também na base de dados de treinamento do sistema.

Já em Kawauchi et al. (2009) é proposto um método chamado AoE (*Angle of Emission*), que possui as vantagens do AoA e do RSSI, sem a necessidade de utilização de sensores extras ao sistema. A partir deste método, os pontos de acesso podem emitir sinais através da rotação de antenas direcionais com as informações dos ângulos de emissão contidas nestes sinais. Dessa forma, a única coisa que os dispositivos móveis precisam realizar é a extração das informações do ângulo de envio destes sinais, sem a necessidade de utilização de antenas especiais e sincronização de tempo. A estimativa da distância e da direção de um determinado dispositivo para os pontos de acesso são feitas a partir destes sinais recebidos. Um dos pontos defendidos neste trabalho é a da facilidade de utilização do RSSI para sistemas de localização já que:

- a potência do sinal pode ser usada para se estimar a distância entre um dispositivo

móvel e um ponto de acesso, somente utilizando o conhecimento da sua distância para no mínimo três pontos de acesso e das posições dos mesmos, sem nenhum dado a mais de sensoriamento ou hardware específico;

- além disso, a radiação sinalizadora transmitida por estes sistemas não precisa de nenhuma rede para conexão ou para utilização no sistema.

O trabalho especificado nesta dissertação define duas vertentes de posicionamento:

1. *com base em distância*: onde é possível especificar as coordenadas em que um determinado nó se encontra dentro do ambiente de localização;
2. *contextual*: onde ao invés de pontos, as base de dados armazena um conjunto de especificações de sinal para uma determinada área, e assim é possível estimar se um determinado nó pertence a um contexto específico de localização com base nos sinais que o mesmo envia para o sistema.

Outro ponto importante da nossa proposta é a utilização da força do sinal chegando ao mesmo tempo nas múltiplos roteadores com antenas direcionais dispostas em direções diferentes para compor a informação de posicionamento do nó responsável pela transmissão do sinal. Em Yang et al. (2005) pode ser visto um sistema semelhante com o proposto no nosso trabalho. Ele define um sistema de localização com um interruptor que indica a alternância entre o envio ou recepção de informações entre múltiplas antenas direcionais, o que pode afetar os resultados obtidos devido ao estado da antena em um determinado instante.

### 3.3 Localização Baseada em Treinamento

O número de trabalhos de localização baseados em treinamento tem crescido nestes últimos anos (Dawes & Chin (2011); Bahl & Padmanabhan (2000)). Conforme proposto em Meng et al. (2011), os algoritmos baseados em treinamento podem ser divididos em determinísticos, onde os estados das suas variáveis de classificação dependem dos parâmetros do modelo e dos seus estados anteriores; e probabilísticos, onde os estados das suas variáveis são descritos por distribuição de probabilidade. Um dos algoritmos existentes na literatura que é utilizado no trabalho citado se chama *k-nearest neighbor*, que atribui uma classe a um determinado objeto a partir dos seu espaço de vizinhos mais próximos. Esta técnica é utilizada em nossa proposta para classificação do posicionamento do nó responsável pelo sinal enviado e detectado pelos roteadores com antenas direcionais. Outra técnica de classificação utilizada nesta dissertação se chama

*Naive Bayes*, que recebe este nome por ser baseado no teorema de Bayes e por ter forte independência de suposições.

Para realizar treinamento em sistemas de localização WiFi, é necessário criar uma base de dados. As bases de dados para estes sistemas geralmente são formadas por informações obtidas do próprio ambiente. Nesta dissertação, a base de dados é criada a partir dos valores de RSSI dos pacotes que chegam nos roteadores com antenas direcionais. Como cada pacote chega repetidamente nos diversos roteadores, o sistema proposto neste trabalho armazena na base de dados os valores de RSSI obtidos em cada uma das antenas a partir de um pacote enviado de uma determinada posição.

Os trabalhos de Lim et al. (2007) e Chiou et al. (2009) apresentam um ponto fundamental sobre a criação da base de dados em ambientes mais dinâmicos para utilização do método de treinamento: a fase de obtenção de informações do ambiente não deveria ocorrer apenas uma vez, já que a cada momento o ambiente em análise está sujeito a inúmeras mudanças como pessoas caminhando, novos nós podendo ser alocados no ambiente e muitos outros fatores. A precisão da fase de obtenção de dados nestes casos depende de quando e como os dados foram coletados. A desvantagem desta metodologia está no fato de que o erro de localização muitas vezes tende a aumentar devido as suas características, tornando muitas vezes esta técnica ineficiente comparada com a forma padrão de utilização de treinamento, onde a base de dados é criada apenas uma vez e o conhecimento do treinamento é utilizado para todos os momentos durante o funcionamento do sistema. O trabalho definido nesta dissertação utiliza uma fase inicial de pré-configuração para criação da base de dados que é utilizada para treinamento durante toda a fase de utilização do sistema. Esta fase acontece apenas uma vez e caso haja alguma alteração estrutural do ambiente ela deve ser repetida para que o sistema mantenha sua precisão de localização.

O trabalho de Bahl & Padmanabhan (2000) utiliza o kNN como uma de suas técnicas. Em suas análises experimentais para localização de usuários foi possível obter bons resultados com a utilização do algoritmo kNN para cálculo das médias das coordenadas dos vizinhos mais próximos a um ponto específico, considerando as direções do usuário quanto quando o sinal era enviado: 25% do erro de distância é de 1,5m e 50% é de 2,75m, ambos para  $k=5$ . Em outro experimento desconsiderando a direção de envio do sinal e considerando somente os maiores valores de RSSI de cada um dos 70 pontos, foi possível conseguir os seguintes resultados: 25% do erro de distância é de aproximadamente 1m e 50% é de aproximadamente 2,5, ambos para  $k=2$ . Em se tratando do rastreamento de usuários, foi possível alcançar um erro médio de distância de 3,5m. Muitos outros sistemas utilizam valores de RSSI para compor a localização de nós espalhados em um determinado ambiente (Ding et al., 2011; Chengdong et al.,

2011; Scholl et al., 2012; Chang & Liao, 2009; Cherntanomwong & Suroso, 2011), mas o diferencial do RADAR é, além da consideração do piso e parede como fatores para atenuação de sinal, a utilização do algoritmo kNN para a identificação da potência de sinal recebida que melhor se caracteriza com a localização do nó em questão. Sistemas de localização como o descrito no trabalho citado comprovam ainda mais a possibilidade de utilização do algoritmo kNN para se obter bons resultados.

O trabalho de Robles et al. (2010) propõe um algoritmo de localização para redes de sensores distribuído baseado no método de treinamento, onde os nós armazenam informações de medições do ambiente para utilizarem durante o processo de cálculo de posicionamento. Para a realização do processo de localização, este trabalho utiliza um protocolo de comunicação baseado na troca de informações entre o nó âncora e o nó alvo. Durante a fase *online*, este protocolo possui dois estágios:

- no primeiro estágio, o nó alvo propaga em modo broadcast  $n$  requisições de localização, e ao recebê-las os nós âncoras enviam para eles uma resposta contendo a média do valor de RSSI entre eles e a sua localização. O nó alvo então mede o RSSI e, através da propriedade simétrica da conexão, obtém uma média única de RSSI de cada nó âncora que recebeu sua informação. A partir disso, o nó alvo executa o algoritmo chamado *Weighted Centroid Localization*;
- quando há a necessidade de maior precisão na localização do nó alvo, o segundo estágio é executado, e então o nó envia a média de RSSI para o nó âncora do qual obteve a maior medida de RSSI. O nó âncora então calcula a posição do nó alvo a partir da sua base de dados e da utilização do esquema baseado em treinamento com utilização do algoritmo kNN. Durante esta fase, o nó alvo entra em estado de espera e ao seu retorno, envia-se uma requisição ao nó âncora novamente para que ele envie a sua posição já calculada.

O trabalho apresentado nesta dissertação também armazena as informações de RSSI nos roteadores com antenas direcionais durante a fase de pré-configuração. Após esta etapa, todas esses dados são tratados numa central e utilizados para classificação de dados.

Um dos problemas encontrado nos algoritmos de localização é a decisão do ponto de referência que será comparado com os dados de treinamento obtidos durante a fase online. O algoritmo kNN, por exemplo, pode muitas vezes não selecionar o melhor candidato a servir como referência para localização. Para solucionar tal problema, o trabalho de Meng et al. (2011) propõe um método de treinamento baseado em probabilidade de região para reduzir o efeito de fatores causados pelo ambiente que afetam

as características do sinal e melhora a eficácia na localização. O sistema proposto neste trabalho possui as seguintes características:

- na fase de *offline* de treino, as distribuições de probabilidade das potências de sinal recebidas em cada ponto de referência são construídas através do método de histograma probabilístico. Estes histogramas são construídos com valores de RSSI obtidos dos pontos de acesso e partir deles é possível aproximar as funções de densidade. A utilização dos histogramas permite que os valores contínuos se tornem discretos e focam na necessidade de utilização de intervalos não sobrepostos que cobrem todo o alcance da variável unidimensional do seu mínimo até seu máximo. Os parâmetros que afetam a estimação da densidade resultante e que podem ser ajustados são a quantidade e o nível de sinal para esses intervalos.
- durante a fase *online*, foi proposto um algoritmo de localização baseado em três passos de detecção:
  - seu primeiro passo é denominado "detecção e eliminação de impactos sobre o ambiente", onde foi proposto um método não iterativo chamado *Random Sample Consensus* (RANSAC) para detecção e eliminação de parte dos pontos de acesso onde os sinais são obstruídos por efeitos causados pelo ambiente. Um detalhe importante deste passo é que a cada instante somente o sinal de alguns pontos de acesso são distorcidos por efeitos do ambiente. O método proposto neste passo foi adotado a partir do método proposto por Fischler & Bolles (1981), método bastante utilizado na área de processamento de imagem, mas utilizado neste trabalho para armazenar o tempo no processamento do sinal;
  - no segundo passo, é proposto um método de seleção de pontos de referência baseado em região para uma comparação com os resultados obtidos pelos pontos de acesso, melhorando os resultados com relação às mudanças sobre o ambiente. Devido às características do método kNN, este trabalho desenvolveu um método baseado em regiões para seleção dos pontos de referências, funcionando como uma família de probabilidade, ao invés de uma única probabilidade como a proporcionada pelo método kNN. Para isto, utiliza-se  $n$  pontos de referência em um grupo e cada grupo de pontos de referência cobre uma região específica. Para calcular qual melhor ponto se identifica com as medidas do ambiente, é feito a soma de probabilidades no espaço de probabilidade e a região com o maior soma é selecionada;

- no último passo, utiliza-se o método de média ponderada sobre as coordenadas dos vários pontos de referências da região selecionada para localização do dispositivo.

O sistema proposto nesta dissertação também foi analisado nos ambientes de localização para verificar como a disposição dos roteadores com antenas direcionais afetavam o desempenho da localização.

Uma das técnicas mais utilizadas atualmente para classificação de dados se chama multiclassificação, onde ao invés de um classificador, vários classificadores são utilizados para gerar uma determinada saída a partir de uma entrada de dados. O trabalho de Menendez et al. (2011) apresenta um sistema para localização de pessoas em ambientes internos utilizando WiFi e multiclassificação. No trabalho citado, foi possível obter uma taxa de acerto de 86% para estimativa de posicionamento de pessoas em ambientes internos comparado com 82,2% em estimativas utilizando classificador kNN com  $k=10$ . O trabalho apresentado nesta dissertação utiliza o método kNN e *Naive Bayes* para classificação de dados mas estes métodos são utilizados para comparação de resultados.

De forma geral, um problema de localização WiFi pode utilizar qualquer um dos inúmeros algoritmos de classificação de dados encontrados na literatura. Cada classificador obviamente possui suas características de funcionamento e isto deve ser estudado antes de implantá-lo em um sistema de localização, para que tal decisão não afete o desempenho do sistema como um todo. Tais pontos podem ser vistos, por exemplo, nos trabalhos de:

- Balaguer et al. (2012): onde vários classificadores são estudados para utilização no problema de localização de robôs em uma área. São estes os classificadores utilizados no trabalho citado: Modelo Gaussiano, *Support Vector Machine*, kNN, Árvore de Decisão, *Random Forest* e Regressão Logística Multinomial;
- Gogolak et al. (2011): onde podemos encontrar um sistema de localização em redes de sensores com uma proposta do método de treinamento utilizando redes neurais para determinar o posicionamento a partir dos valores de RSSI, algo bastante diferente dos outros trabalhos encontrados na literatura. O processo para determinação de posicionamento dos nós utiliza duas bases de dados para a fase de treinamento: uma com os dados em seus estados naturais, e a outra com dados estatísticos dos valores de RSSI em cada posição. Alguns desses dados são média, mediana e desvio padrão dos valores de RSSI. O treinamento e a simulação deste sistema foram realizados utilizando a ferramenta Matlab. Como resultado deste trabalho, foi possível obter 0,3 metros de precisão na localização.

A proposta do nosso trabalho é realizar a localização de dispositivos móveis utilizando alguns dos conceitos e técnicas citados nos trabalhos anteriores. Tal localização será realizada em um sistema composto por vários roteadores com antenas direcionais que irão receber um determinado pacote enviado por um dispositivo com posição desconhecida. Como as antenas estarão apontadas para direções diferentes, o pacote que será recebido por elas terá valores de RSSI diferentes. Estes valores serão avaliados a partir da base de dados criada durante a fase *offline* do sistema, e através de um algoritmo de classificação (como o kNN ou Naive Bayes que são utilizados neste trabalho) será possível identificar o posicionamento dos dispositivos espalhados no ambiente de localização.

### 3.4 Conclusão

Foram apresentados neste capítulo trabalhos relacionados à proposta do trabalho apresentado nesta dissertação. Eles foram divididos em três frentes de sistemas de localização WiFi:

- Localização Baseada em RSSI;
- Localização Baseada em Antenas Direcionais;
- Localização Baseada em Treinamento.

Cada uma desses frentes são importantes para o estudo do nosso trabalho pois o sistema aqui proposto depende de medidas de RSSI a partir da utilização de antenas direcionais para serem criadas entradas de dados que são utilizadas por algoritmos de classificação para treinamento, capaz de gerar estimativa de posicionamento de nós cujas posições são desconhecidas.

Através de todos os trabalhos aqui citados, é possível ver que a base definida para este trabalho pode ser desenvolvida e aplicada para um sistema eficiente e capaz de fazer uso das três frentes aqui definidas.

## Capítulo 4

# Localização de Dispositivos Móveis usando Roteadores com Antenas Direcionais e Classificação de Dados

Este trabalho tem como base metodológica a localização convencional para ambientes omnidirecionais utilizando roteadores omnidirecionais. A mudança de roteadores omnidirecionais para direcionais utilizada nesta proposta se dá pelas desvantagem encontradas nos sistemas convencionais de localização como:

- necessidade de um grande número de roteadores para ambientes internos maiores;
- não possui foco de localização somente para a área desejada, o que devido a sua característica de sinal, pode ocasionar em erros de localização para o sistema;
- não dedicam a potência de sinal somente para a área desejada, o que prejudica o ambiente de localização como um todo.

A seguir serão apresentados todas os pontos concernentes a metodologia desta proposta de localização utilizando múltiplos roteadores com antenas direcionais e classificação de dados.

### 4.1 Método Proposto

O sistema de localização proposto neste trabalho determina o posicionamento de nós alvos a partir dos valores de RSSI obtidos nos múltiplos roteadores com antenas direcionais. O impacto de um pacote enviado de uma determinada posição, e recebido

repetidamente nos vários roteadores com antenas direcionais, gera valores de RSSI diferentes devido aos fatores como direcionamento e posição das antenas, características do sinal e interferências causadas pelo ambiente.

Antes que o sistema seja utilizado para localização de objetos, pessoas ou dispositivos móveis em um determinado ambiente interno, é necessário haver uma fase de obtenção de informações do ambiente, etapa esta conhecida como fase *offline*. Durante esta fase, os roteadores recebem inúmeros pacotes enviados pelos dispositivos em posições específicas, e cada pacote possui as seguintes informações encapsuladas: endereço MAC do dispositivo, número do pacote e um indicador de posição (um número identificador da posição de onde o pacote foi enviado). O sistema depende do endereço MAC poder focar somente nos pacotes que são enviados pelo dispositivo pertencente ao sistema. Todas as informações dos pacotes são armazenadas em *logs* nos roteadores.

Após enviar pacotes de todos os pontos desejados e armazenar todas as informações necessárias, é necessária uma fase de pré-processamento para extrair o valor de RSSI de um pacote obtido em cada roteador direcional e também a posição de onde cada pacote foi enviado. Todas essas informações são organizadas em uma base de dados para ser utilizada pelo algoritmo de classificação. Vale a pena ressaltar que quanto mais pontos houver numa área de localização, maior será a qualidade de classificação na próxima fase.

A base de dados utilizada como conhecimento para os algoritmos de classificação são formadas por instâncias  $x$  onde  $x_i = \{a_1, a_2, \dots, a_{(n-1)}, a_n, c_i\}$ , cada  $a_i$  é um valor de RSSI de um pacote específico registrado pelo roteador  $i$  instalado no ambiente de localização e  $c_i$  é um número que identifica de que posição o pacote foi enviado.

A próxima etapa de posicionamento é conhecida como fase *online* e, neste momento, o sistema recebe pacotes de dispositivos móveis cujas posições de onde eles foram enviados são importantes de serem descobertas. Esses pacotes são recebidos pelos roteadores e utilizados pelo classificador que vai estimar a posição dos dispositivos móveis baseados na base de dados criada durante a fase *offline*.

Nesta proposta são usados dois algoritmos de classificação separadamente para comparação do sistema de localização utilizando dois tipos diferentes de algoritmos para o cálculo de estimativa de posicionamento. São eles:

- kNN (k-Nearest Neighbor): utiliza o espaço de exemplos rotulados e próximos ao exemplo não rotulado, obtém qual o rótulo que mais aparece entre esses exemplos próximos e aplica este rótulo para o exemplo não rotulado. O valor de  $k$  é uma decisão importante a ser tomada pois isto afeta o comportamento do algoritmo. Depois de estudarmos sobre técnicas de seleção de parâmetros para este algoritmo

Keller et al. (1985), foi possível chegar a conclusão de que o melhor valor para  $k$  precisa ser um número ímpar que vai evitar ao máximo as situações de empate, e mesmo quando houver tal situação, o algoritmo recorre à análise das classes empatadas e escolhe entre elas a com menor distância para o ponto que se deseja estimar. Dessa forma, o valor de escolhido de  $k$  para este trabalho foi 5 porque é importante assegurar que os pontos no espaço de classificação são confiáveis. A técnica usada como cálculo de distância entre pontos foi a *Manhattan*, baseada no caminho em grades entre um ponto e outro.

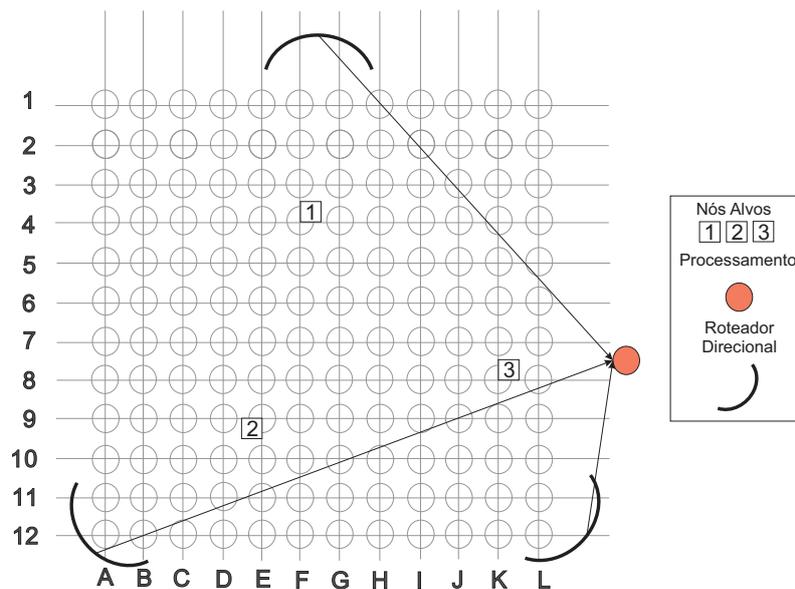
- Naive Bayes: é baseado no teorema de Bayes, considerando que um atributo não possui relação alguma com qualquer outro e analisa como ele pode contribuir para a probabilidade de uma definição específica de classe.

Este trabalho apresenta duas abordagens de localização utilizando roteadores com antenas direcionais: a primeira abordagem utiliza os roteadores espalhados pelo ambiente de localização; já a segunda abordagem utiliza a junção destes roteadores, cada um apontando para direções diferentes. Além destas abordagens, será implementada a abordagem utilizando roteadores com antenas omnidirecionais para avaliação dos resultados e das características de cada um destes sistemas, e também para caráter comparativo.

## 4.2 Abordagens de Localização

### 4.2.1 Abordagem com Utilização de Roteadores com Antenas Direcionais Espalhados

Na abordagem apresentada na figura 4.1, os roteadores com antenas direcionais estão dispostos em diferentes lugares do ambiente de localização, de maneira que todo o ambiente seja coberto pelo alcance das antenas. Com isto, um sinal enviado por um dispositivo para o sistema é capaz de gerar impactos diferentes sobre as antenas direcionais, de maneira que cada uma delas receba valores de RSSI diferentes para um mesmo pacote. Através deste impacto, é possível gerar a localização do dispositivo que enviou o sinal.



**Figura 4.1.** Exemplo do ambiente de treinamento com utilização de roteadores com antenas direcionais espalhados.

#### 4.2.2 Abordagem com Utilização de Roteadores com Antenas Direcionais Unidos

Na abordagem apresentada na figura 4.2, os roteadores com antenas direcionais estão unidos e cada um deles é posicionado para direções diferentes, de maneira que todo o ambiente onde o sistema de localização foi implantado, esteja dentro do alcance das antenas. Em tal caso, ainda há impactos diferentes sobre as antenas direcionais quando um pacote é enviado por um dispositivo cuja a posição é desconhecida, e a partir destes impactos é possível realizar a sua localização.

#### 4.2.3 Abordagem com Utilização de Roteadores com Antenas Omnidirecionais Espalhados

Para fins de comparação, foi implementada a forma convencional de localização utilizando os roteadores com antenas omnidirecionais espalhados pelo ambiente onde se deseja realizar a localização dos nós pertencentes ao sistema. O impacto de sobreposição de campo magnético causado por cada um dos roteadores sobre os seus vizinhos é menor neste caso já que, mesmo em um ambiente interno, eles geralmente não são instalados próximos um dos outros. Com isto, é possível avaliar o impacto de recepção de pacotes sobre o sistema com esta configuração. A figura 4.3 apresenta a disposição do sistema citado.

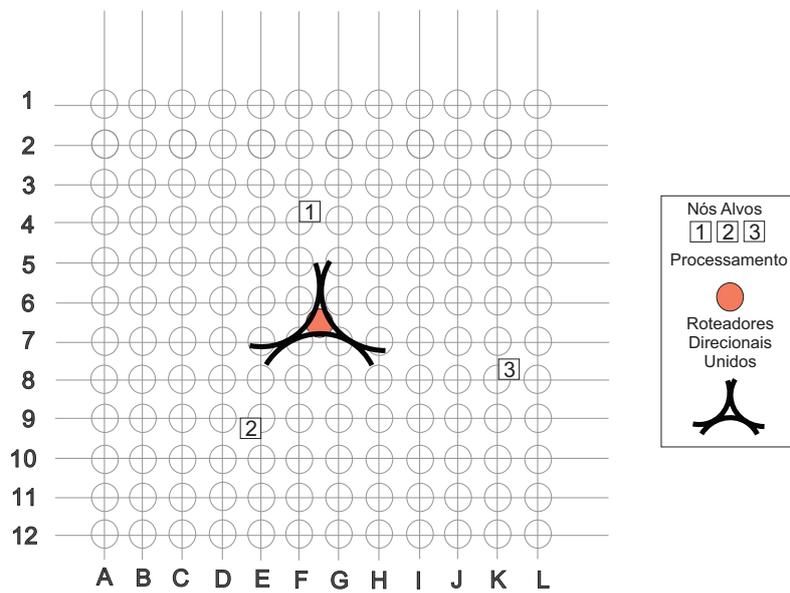


Figura 4.2. Exemplo do ambiente de treinamento com utilização de roteadores com antenas direcionais unidos.

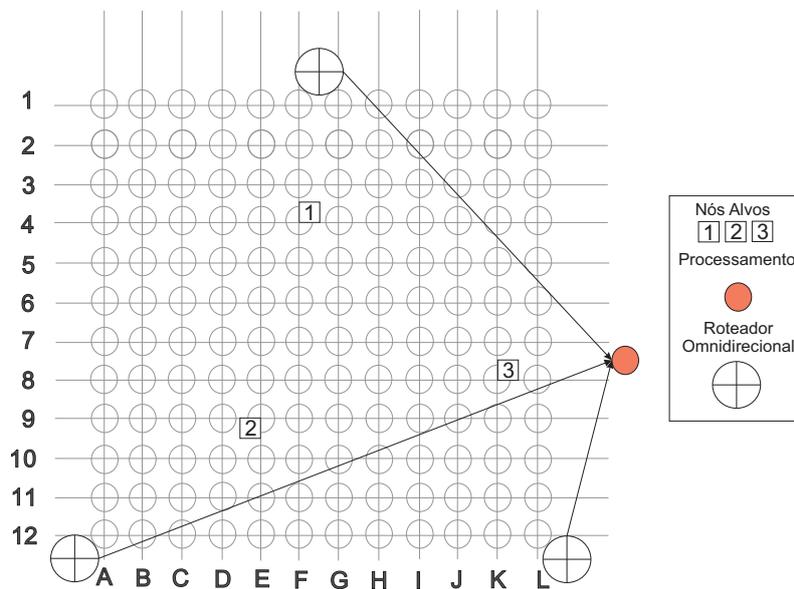


Figura 4.3. Exemplo do ambiente de treinamento com utilização de roteadores com antenas omnidirecionais espalhados.

As figuras 4.1 e 4.2 exemplificam melhor a proposta deste trabalho, e a figura e 4.3 apresentam a abordagem utilizada para análise comparativa. Em todas as figuras, é possível dizer que cada tupla (letra,número) indica uma determinada posição de onde foram enviados pacotes para que os roteadores obtivessem os seus valores de RSSI. Tais valores formarão a base de dados utilizada como conhecimento dos aspectos de sinal obtidos por cada roteador no ambiente durante a fase *offline*. Os nós alvos 1, 2 e 3 estão localizados nas proximidades dos pontos de medição e a partir do sinal transmitido por eles sendo recebido pelos roteadores com antenas direcionais, o sistema será capaz de localizar qual a posição mais se identifica com suas medições. Dessa forma, com uma determinada margem de erro, o nó 2 será identificado como localizado na posição (E,9), o 1 em (F,4) e o 3 em (K,8).

### 4.3 Conclusão

Este capítulo apresentou a motivação deste trabalho de utilizar roteadores direcionais em um sistema de localização ao invés de roteadores direcionais. Esta decisão se dá pelo fato de os roteadores direcionais focarem apenas na área específica de localização, onde o sistema de localização pode tirar proveito do direcionamento dos roteadores como forma de obter uma melhor qualidade de sinal no que diz respeito ao alcance, potência de sinal, e obviamente o direcionamento, que permite o sistema não desperdiçar abrangência de sinal para áreas que não interessam.

Foi apresentado também neste capítulo os detalhes de funcionamento da proposta. Espera-se com este sistema que, a partir dos valores de RSSI obtidos extraídos pelo sistema a partir dos sinais gerados por um nó com posição desconhecida, seja possível estimar o seu posicionamento dentro do ambiente de localização. Para isto, o sistema depende primeiramente de uma fase *offline*, onde são extraído informações do ambiente de localização a partir do mapeamento do mesmo onde são enviados vários pacotes para o sistema a partir de cada posição mapeada no ambiente. Com a extração das informações de RSSI de cada pacote enviado pelos nós em cada posição do ambiente são criadas as bases de dados utilizadas na próxima fase do sistema.

A próxima fase, conhecida como fase *online*, é responsável por estimar o posicionamento de um nó que envia sinal para o sistema de localização. De tal sinal são extraídas as informações de RSSI e, com a execução do algoritmo de classificação sobre a base de dados criada durante a fase *offline*, é possível estimar o posicionamento do nó. Foram utilizados dois algoritmos separadamente como meio de analisar comparativamente o funcionamento do sistema de localização proposto neste trabalho com cada

um destes algoritmos. São eles os algoritmos  $kNN$  e Naive Bayes.

Este trabalho utilizou duas abordagens com roteadores direcionais e as comparou seus funcionamentos com a abordagem convencional onde se utiliza roteadores omnidirecionais. A primeira abordagem com roteadores direcionais é feita através da organização dos mesmos de maneira espalhada no ambiente de localização, de maneira que cada um destes roteadores tenha como direção de foco o centro do ambiente de localização. Já a segunda abordagem é feita através da organização dos roteadores unidos, com direção de foco para as extremidades dos ambiente de localização.

## Capítulo 5

# Análise e Avaliação dos Resultados

Esse capítulo apresenta os detalhes dos experimentos e também os detalhes da análise e da avaliação dos resultados obtidos nos experimentos realizados em dois ambientes do Instituto de Computação da UFAM (IComp): o primeiro composto de 3 laboratórios e o segundo composto por todas as salas dos professores. No primeiro ambiente composto pelos laboratórios, foi trabalhada a *Localização Precisa*, onde é possível estimar a posição exata de um dispositivo nos eixos  $x$  e  $y$  de um determinado ambiente onde ele está instalado. Já no segundo ambiente composto pelas salas dos professores, foi trabalhada a *Localização Contextual*, onde é possível estimar posição de um determinado dispositivo com base na análise do contexto do ambiente onde ele está inserido.

Vale ressaltar que a Localização Precisa permite atribuir uma tupla  $(x,y)$  para indicar a posição exata de um dispositivo em um determinado ambiente. Por outro lado, a Localização Contextual atribui um rótulo como estimativa de localização para um determinado positivo, com base nas informações contextuais do ambiente onde o mesmo está inserido. O rótulo, desta forma, indica uma determinada área do ambiente de localização onde qualquer outro dispositivo inserido nela se comporta da mesma forma.

Através destes métodos de localização, foi possível identificar importantes características do sistema de localização proposto neste trabalho e também analisar como é o comportamento do sistema em cada tipo de ambiente, verificando as vantagens e desvantagens nos três tipos de abordagens utilizadas para estudo e avaliação nesta pesquisa.

As seções a seguir apresentam os detalhes dos experimentos realizados neste trabalho e, em seguida, a análise e a avaliação dos resultados obtidos com base nas duas vertentes de localização citadas acima: Localização Precisa na seção 5.2 e Localização Contextual na seção 5.3.

## 5.1 Experimentos

As subseções a seguir apresentam especificações do local, equipamentos e softwares utilizados para a realização da proposta de localização definida neste trabalho.

### 5.1.1 Local dos Experimentos

A primeira etapa do sistema de localização proposto neste trabalho, conhecida como fase *offline*, é necessária para criar a base de dados contendo os valores de RSSI obtidos em cada roteador através de um pacote enviado e sua posição de envio. Esta base de dados é utilizada pelo sistema como conhecimento das características do sinal no ambiente de localização para ser utilizada durante a classificação de dados. Os experimentos realizados nesta proposta foram feitos em dois ambientes: o primeiro é composto de 3 laboratórios do Instituto de Computação <sup>1</sup> (figuras 5.1 e 5.2) com uma área de  $7.5 \times 14.7 \text{ m}^2$ , e o segundo ambiente é composto por todas as salas dos professores do mesmo instituto (figuras 5.3 e 5.4) com uma área de  $10.73 \times 58.08 \text{ m}^2$ .

Foram considerados dois tipos de localização neste trabalho:

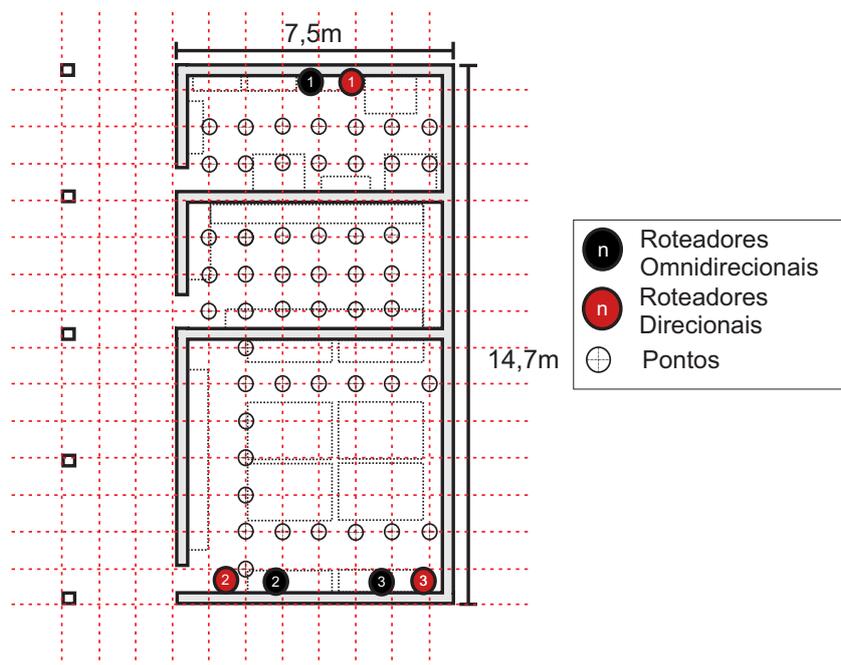
- *Localização Precisa*: utilizada no ambiente composto pelos laboratórios onde foi possível estimar de que ponto mapeado no ambiente de localização cada pacote foi enviado ao sistema. O ambiente utilizado para este tipo de localização pode ser visto nas figuras 5.1 e 5.2, onde os roteadores com antenas direcionais se encontram espalhados e juntos, respectivamente;
- *Localização Contextual*: utilizada no ambiente composto pelas salas dos professores onde foi possível estimar de que sala cada pacote foi enviado ao sistema. O ambiente utilizado para este tipo de localização pode ser visto nas figuras 5.3 e 5.4, onde os roteadores com antenas direcionais se encontram espalhados e juntos, respectivamente.

Alguns elementos podem ser encontrados nas figuras 5.1, 5.2, 5.3 e 5.4, e cada um deles possui um significado específico:

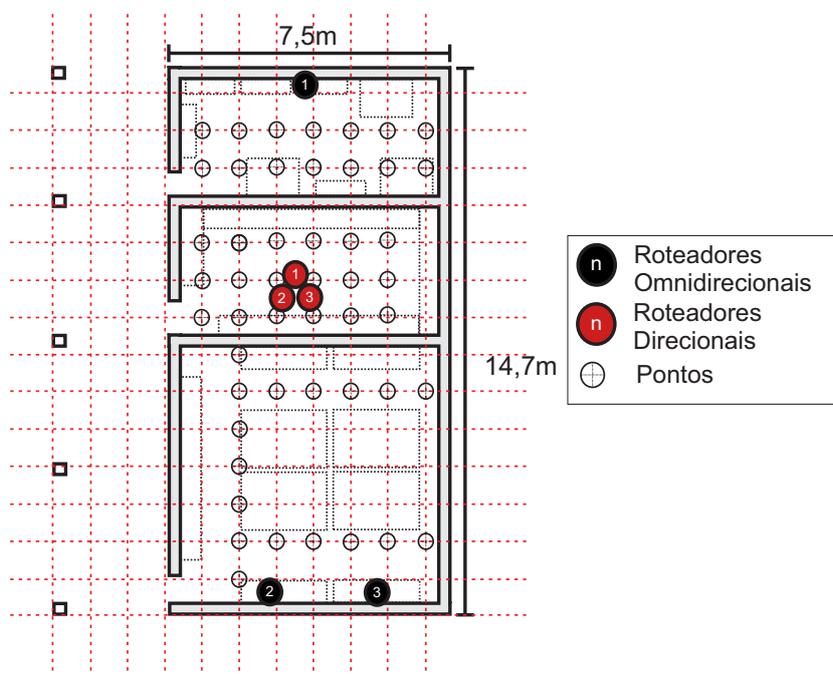
- pontos: posições de onde os pacotes foram enviados;
- linhas: paredes;
- linhas pontilhadas: mesas ou armários;

---

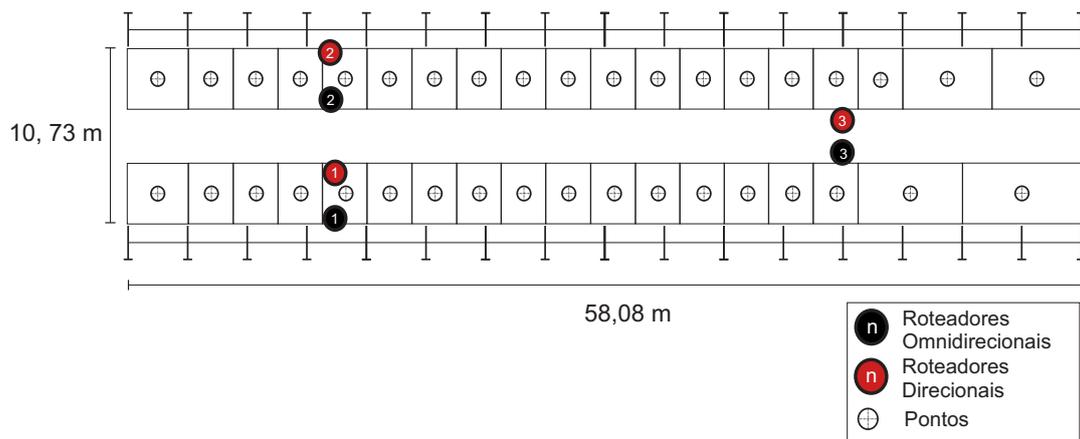
<sup>1</sup>Os laboratórios utilizados foram: Computação Móvel e Ubíqua, Computação Inteligente e Autônoma e Tecnologias Emergentes e Segurança de Sistemas



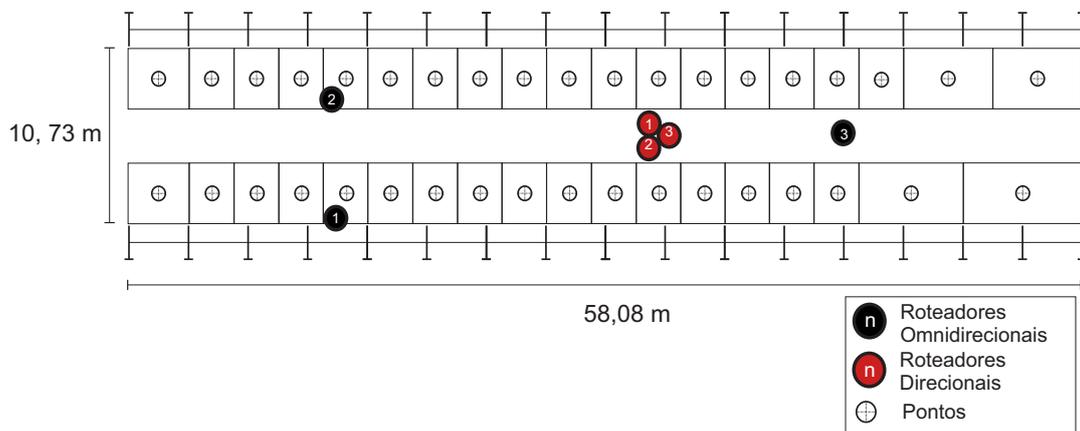
**Figura 5.1.** Ambiente 1 dos experimentos com roteadores direcionais espalhados.



**Figura 5.2.** Ambiente 1 dos experimentos com roteadores direcionais juntos.



**Figura 5.3.** Ambiente 2 dos experimentos com roteadores direcionais espalhados.



**Figura 5.4.** Ambiente 2 dos experimentos com roteadores direcionais juntos.

- bolas pretas: roteadores com antenas omnidirecionais;
- bolas vermelhas: roteadores com antenas direcionais;

De cada posição mapeada no ambiente de localização, foram enviados 200 pacotes para o sistema composto pelos roteadores e o número de posições variou conforme o ambiente de localização. No ambiente onde foi aplicada a localização precisa, foram mapeadas 49 posições, e no ambiente de localização contextual foram mapeados 37 posições.

### 5.1.2 Equipamentos Utilizados

Os experimentos realizados neste trabalho foram todos feitos em ambientes internos reais. Foram utilizados durante estes experimentos os seguintes equipamentos:



**Figura 5.5.** Linksys WRT610N. Fonte: [www.ubergizmo.com](http://www.ubergizmo.com)



**Figura 5.6.** NanoStation Loco M2. Fonte: [www.comrede.com.br](http://www.comrede.com.br)



**Figura 5.7.** Dell Vostro 3560. Fonte: [www.dell.com](http://www.dell.com)

- Roteadores Wireless Linksys WRT610N com o *firmware* DD-Wrt (figura 5.5), como os roteadores omnidirecionais do sistema;
- Roteadores Wireless NanoStation Loco M2 2,4ghz com uma versão modificada do *firmware* OpenWrt (figura 5.6), como os roteadores direcionais com ângulo de cobertura de 120 graus;

- Notebook Dell Vostro 3560 (figura 5.7 com o sistema operacional Ubuntu 12.04 LTS na versão 64 bits. Ele responsável por enviar pacotes em cada ponto do ambiente de localização;

As imagens 5.8, 5.9, 5.10 e 5.11 mostram a organização dos roteadores nos ambientes onde foram realizados os experimentos.



**Figura 5.8.** Organização dos Roteadores Direcionais Espalhados e Omnidirecionais nos Laboratórios.



**Figura 5.9.** Organização dos Roteadores Direcionais Juntos e Omnidirecionais nos Laboratórios.



**Figura 5.10.** Organização dos Roteadores Direcionais Espalhados e Omnidirecionais nas Salas dos Professores.



**Figura 5.11.** Organização dos Roteadores Direcionais Juntos nas Salas dos Professores.

### 5.1.3 Softwares Utilizados

O envio e a recepção de pacotes entre roteadores e dispositivos móveis depende de *softwares* utilitários capazes de gerar os pacotes com as configurações necessárias do projeto e recebê-los para registro dos mesmos em *logs*. Este trabalho utilizou os seguintes *softwares*:

- RssiCapSend: este *software* implementado em C foi desenvolvido pelo laboratório de Computação Móvel e Ubíqua para enviar pacotes com informações encapsuladas conforme a configuração desejada. No caso deste projeto, as seguintes informações eram encapsuladas nos pacotes: endereço MAC do dispositivo, número

do pacote e um indicador de posição. Este *software* foi utilizado no notebook utilizado neste projeto. A sua execução era realizada pelo terminal do Ubuntu com o seguinte formato:

```
./rssiCapSend [interface de rede para envio de pacotes] [delay em segundos entre
envio de cada pacote] [número de pacotes] [indicador de posição] 0 0,
```

onde os dois últimos zeros do comando poderiam ser usados para especificar o indicador dos eixos  $x$  e  $y$ , respectivamente mas para os experimentos realizados neste trabalho sua utilização não foi necessária;

- *RssiCapMultiAir*: este *software* implementado em C foi desenvolvido pelo laboratório de Computação Móvel e Ubíqua para receber pacotes com informações encapsuladas, extraí-las e medir o valor do RSSI de cada um deles. Este *software* extrai as seguintes informações de cada pacote: endereço MAC (utilizado pelo programa como filtro para receber apenas pacotes de um dispositivo específico), número do pacote e indicador de posição. Este programa era deixado executando em segundo plano nos roteadores omnidirecionais através do acesso por terminal a cada um deles, utilizando *Secure Shell* (SSH). A sua execução possui o seguinte formato:

```
./rssiCapMultiAir [interface de rede para monitoramento dos pacotes] 0 0 0,
```

onde os 3 últimos zeros do comando poderiam ser usados para especificar o indicador dos eixos  $x$ ,  $y$  e  $z$ , respectivamente mas para os experimentos realizados neste trabalho sua utilização não foi necessária;

- *tcpdump*: este *software* é um analisador de pacotes para Linux. Através de sua utilização, foi possível medir os valores de RSSI de cada pacote e extrair as informações enviadas pelo *software* *RssiCapSend*: endereço MAC (utilizado pelo programa como filtro para receber apenas pacotes de um dispositivo específico), número do pacote e indicador de posição. Este programa era deixado executando em segundo plano nos roteadores direcionais através do acesso por terminal a cada um deles, utilizando *Secure Shell* (SSH). A sua execução possui o seguinte formato:

```
tcpdump -e -i 'ether [endereço MAC para filtrar os pacotes recebidos]' -c
[quantidade de pacotes a registrar],
```

onde -e imprime o cabeçalho de nível de ligação, -i especifica a interface de rede que se deseja ouvir e -c especifica a quantidade de pacotes que o programa deve receber antes de terminar sua execução.

Como visto, em cada tipo de roteador foi utilizado um *software* específico para analisar e extrair informações dos pacotes. Isto se deve pelo fato de cada um dos tipos de roteador possuírem firmwares diferentes, e devido a versão modificada do firmware OpenWrt no roteador direcional não possível utilizar o RssiCapMultiAir para analisar pacotes.

Todos os registros com informações de pacotes armazenados nos roteadores eram recolhidos ao término de cada experimento para a parte de pré-processamento. A realização da parte de pré-processamento é importante para organizar todas as informações para a base de dados utilizada posteriormente como conhecimento pelo algoritmo de classificação dos aspectos do sinal no ambiente onde o sistema está implantado. A extração das informações contidas nos *logs* foi feita através de programas em C e scripts *bash*.

Os *logs* utilizados como entrada para os algoritmos de classificação possuíam o seguinte formato:

$$\begin{aligned}
 & \langle RSSI_1^{\text{roteador1}}, RSSI_1^{\text{roteador2}}, RSSI_1^{\text{roteador3}}, POSIÇÃO_1 \rangle \\
 & \langle RSSI_2^{\text{roteador1}}, RSSI_2^{\text{roteador2}}, RSSI_2^{\text{roteador3}}, POSIÇÃO_1 \rangle \\
 & \langle RSSI_3^{\text{roteador1}}, RSSI_3^{\text{roteador2}}, RSSI_3^{\text{roteador3}}, POSIÇÃO_1 \rangle \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \langle RSSI_{199}^{\text{roteador1}}, RSSI_{199}^{\text{roteador2}}, RSSI_{199}^{\text{roteador3}}, POSIÇÃO_1 \rangle \\
 & \langle RSSI_{200}^{\text{roteador1}}, RSSI_{200}^{\text{roteador2}}, RSSI_{200}^{\text{roteador3}}, POSIÇÃO_1 \rangle \\
 & \langle RSSI_1^{\text{roteador1}}, RSSI_1^{\text{roteador2}}, RSSI_1^{\text{roteador3}}, POSIÇÃO_2 \rangle \\
 & \langle RSSI_2^{\text{roteador1}}, RSSI_2^{\text{roteador2}}, RSSI_2^{\text{roteador3}}, POSIÇÃO_2 \rangle \\
 & \langle RSSI_3^{\text{roteador1}}, RSSI_3^{\text{roteador2}}, RSSI_3^{\text{roteador3}}, POSIÇÃO_2 \rangle \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \langle RSSI_{199}^{\text{roteador1}}, RSSI_{199}^{\text{roteador2}}, RSSI_{199}^{\text{roteador3}}, POSIÇÃO_{MAX} \rangle \\
 & \langle RSSI_{200}^{\text{roteador1}}, RSSI_{200}^{\text{roteador2}}, RSSI_{200}^{\text{roteador3}}, POSIÇÃO_{MAX} \rangle
 \end{aligned}$$

onde *MAX* é o número máximo de posições mapeadas no ambiente de localização de onde os dados de RSSI foram extraídos.

Para a realização da avaliação e análise de resultados do sistema proposto neste trabalho, é necessário ter além da base de dados como conhecimento uma outra base

para testar se o sistema funcionou de maneira adequada. Diante disto, foi utilizada a ferramenta de aprendizagem de máquina *Waikato Environment for Knowledge Analysis* (Weka) <sup>2</sup> que possui os vários algoritmos de classificação supervisionada ou não-supervisionada <sup>3</sup> e também muitas outras ferramentas para análise de dados. Foram utilizados para avaliação separada e comparativa os algoritmos kNN e Naive Bayes já implementados no Weka e a divisão da base de dados em treino e teste foi feita utilizando a configuração de análise de dados chamada Divisão por Porcentagem com valor 66% (66% da base para treino e 34% para teste). Esta técnica de divisão da base de dados se baseia o método chamado *hold-out*, que avalia a generalização do conhecimento gerado no sistema de classificação a partir das informações utilizadas como entrada. Isso se dá justamente pela divisão do conhecimento em dois conjuntos: o de treino e o de teste.

O algoritmo kNN foi utilizado com distância *Manhattan* como configuração de distância entre as instâncias definidas como entrada de dados, já que as informações dos *logs* foram mapeadas como dados discretos. Este algoritmo foi utilizado para analisar como ficam a disposição das medidas de RSSI obtidas nos três roteadores utilizados pelo sistema, para uma determinada posição de envio. Dependendo da proximidade em que os pontos do mapeamento foram definidos, é possível que os pontos de RSSI de um ponto possam se sobrepor a outros. A execução do algoritmo kNN através do *Weka* nos permite analisar este fator, já que além da porcentagem de instâncias de teste classificadas corretamente, a ferramenta também apresenta a quantidade de estimativas erradas de uma posição com relação a posição que o sistema a estimou. Assim, é possível dizer quantas instâncias da posição 2 foram estimadas erroneamente como posição 3, por exemplo. O valor de  $k$  para o algoritmo kNN utilizado neste trabalho foi 5, escolhido com base no estudo de qual valor de  $k$  gera resultados melhores na classificação realizada sobre os dois ambientes de localização, e também com base no fator de que, sendo  $k$  um número ímpar, é possível tender o resultado da classificação para haver menos empates possíveis e mesmo que isto aconteça o algoritmo realiza desempate com base nos critérios de instância mais próxima e classe que mais ocorre entre as instâncias, respectivamente.

Vale ressaltar que a classificação é dada corretamente quando a instância analisada como teste tem sua posição estimada pelo sistema igual a posição de onde foram extraídas suas informações. Ou seja, se uma instância com valores de RSSI extraídos da posição 1 tiver sua posição estimada pelo sistema como posição 1 também, ela foi

---

<sup>2</sup>Site da ferramenta: [www.cs.waikato.ac.nz/ml/weka](http://www.cs.waikato.ac.nz/ml/weka)

<sup>3</sup>Algoritmos capazes de extrair conhecimento de uma base de dados onde cada uma de suas instâncias não possui rotulação.

classificada corretamente.

O algoritmo Naive Bayes foi utilizado neste trabalho pois o mesmo pode identificar quão importante são os componentes do sistema para a estimativa do posicionamento. Neste caso, o algoritmo consegue definir se a composição do sistema com roteadores direcionais é melhor que a composição com roteadores omnidirecionais com base nas entradas de dados utilizadas pelo algoritmo.

## 5.2 Localização Precisa

Através da Localização Precisa, é possível estimar a posição exata em que um determinado dispositivo se encontra dentro do ambiente de localização. A partir do conhecimento prévio formado através da etapa de pré-configuração, o sistema de localização consegue definir uma tupla  $(x,y)$  como estimativa da posição do dispositivo cujo localização é desconhecida. Assim, a posição estimada se baseia no mapeamento do ambiente de localização, através da etapa de pré-configuração onde são registradas as características de sinal em cada ponto do sistema de localização pertencentes aos eixos x e y.

Os experimentos onde foi utilizada a Localização Precisa é composto por três laboratórios do IComp, onde foi possível mapear toda sua área em pontos distantes de 1 metro, tanto no eixo x quanto no eixo y, conforme pode ser visto nas figuras 5.13, 5.14 e 5.15.

A tabela 5.1 apresenta os resultados dos experimentos realizados nos laboratórios do IComp. Estes resultados são baseados na porcentagem de instâncias classificadas corretamente.

**Tabela 5.1.** Porcentagem de instâncias classificadas corretamente nos experimentos realizados nos laboratórios do IComp.

	<i>Roteadores Omnidirecio- nais</i>	<i>Roteadores Direcionais Espalhados</i>	<i>Roteadores Direcionais Juntos</i>
kNN{k=5}	84,26%	75,98%	83,93%
<i>Naive Bayes</i>	74,25%	74,63%	81,86%

É possível ver claramente que o sistema proposto neste trabalho obteve bons resultados para o ambiente composto de laboratórios, mas se percebe que a forma na qual os roteadores direcionais estão organizados numa determinada área afeta diretamente os resultados obtidos pelo sistema. Tal situação pode ser vista na diferença existente

entre os resultados obtidos utilizando roteadores direcionais espalhados e unidos. O ambiente composto por laboratórios é uma área pequena, e com isto, os roteadores direcionais unidos conseguem gerar melhores resultados para localização do que quando eles estão separados.

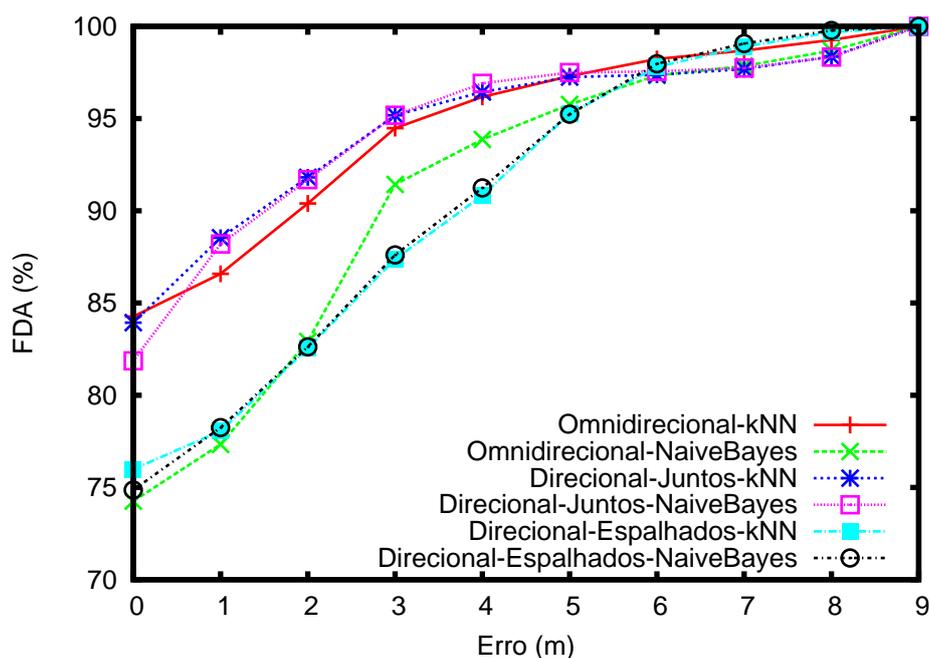
A organização dos roteadores direcionais de maneira unida em uma área pequena também contribui para que o sistema perca o mínimo de pacotes possível no que diz respeito a distância de envio, e também, caracteriza de maneira melhor o comportamento do sinal obtido pelo sistema, já que cada roteador está focado em uma área específica, onde nenhum outro roteador está focado.

Outro fator que afeta os resultados obtidos é a escolha do classificador que melhor atua conforme nas condições do ambiente onde o sistema de localização está inserido. Isto pode ser visto na tabela 5.1 onde os roteadores omnidirecionais utilizando kNN com  $k=5$  obtiveram resultados melhores que os roteadores direcionais, enquanto os roteadores direcionais tiveram melhores resultados com o classificador *Naive Bayes*. Tal comportamento se dá pelas seguintes causas:

- O algoritmo kNN consegue classificar com maior precisão uma classe quando a mesma é bem caracterizada na base de dados utilizada como conhecimento. Desta forma, quanto melhor for descrita e caracterizada uma determinada classe, melhor será o conhecimento que o algoritmo kNN terá e conseqüentemente maior precisão será dada no momento que uma instância for classificada. Como a potência de sinal dedicada ao roteador omnidirecional é menor que a do direcional, os valores de RSSI para os roteadores omnidirecionais variam mais do que nos roteadores direcionais, descrevendo melhor a classe que é representada por pontos nos experimentos realizados nos laboratórios;
- O algoritmo *Naive Bayes* consegue atribuir melhor precisão na classificação conforme a importância que cada atributo possui sobre a classe. Por isto, os experimentos utilizando este algoritmo foram melhores com os roteadores direcionais do que com os roteadores omnidirecionais, já que os direcionais geram melhores informações para o algoritmo devido ao seu foco em apenas uma direção do ambiente. O direcionamento do roteador permite que o algoritmo perceba que o comportamento de sinal gera informações mais confiáveis para classificação.

A figura 5.12 mostra a Função de Distribuição Acumulativa (FDA) em metros com base nos experimentos realizados nos laboratórios. A FDA identifica a porcentagem de precisão (eixo y) com erro menor ou igual que um parâmetro ou condição específica (eixo x). Esta figura mostra que o comportamento do sistema de localização

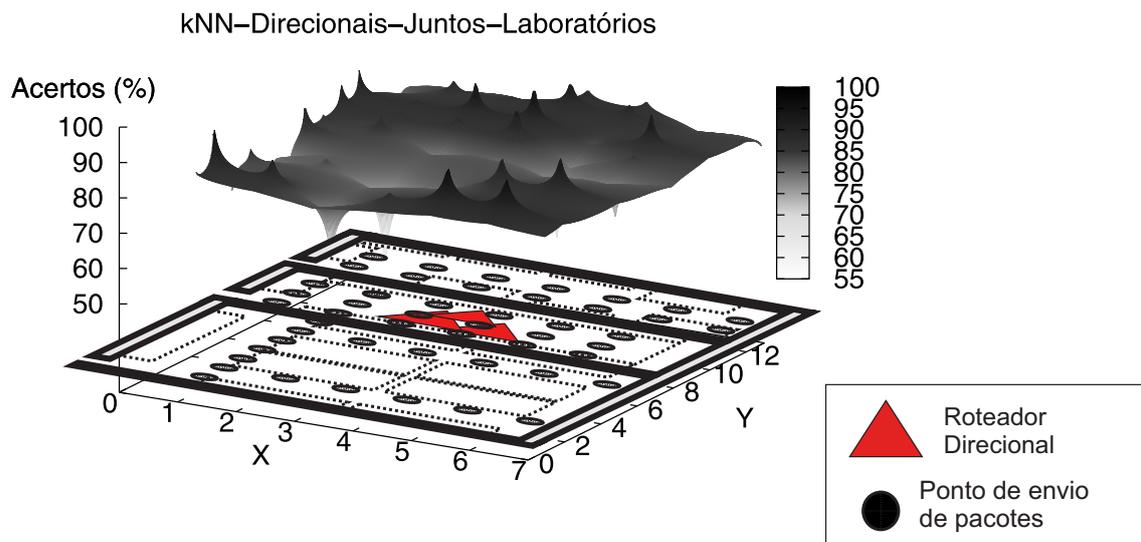
instalado nos laboratórios e utilizando roteadores direcionais, tanto separados quanto unidos, possui o mesmo comportamento que o sistema de localização utilizando antenas omnidirecionais. A figura 5.12 também mostra que os roteadores direcionais juntos conseguiram ter melhores resultados que os roteadores direcionais separados, o que nos faz concluir que a abordagem de união de roteadores direcionais para localização possui um bom desempenho para ambientes pequenos.



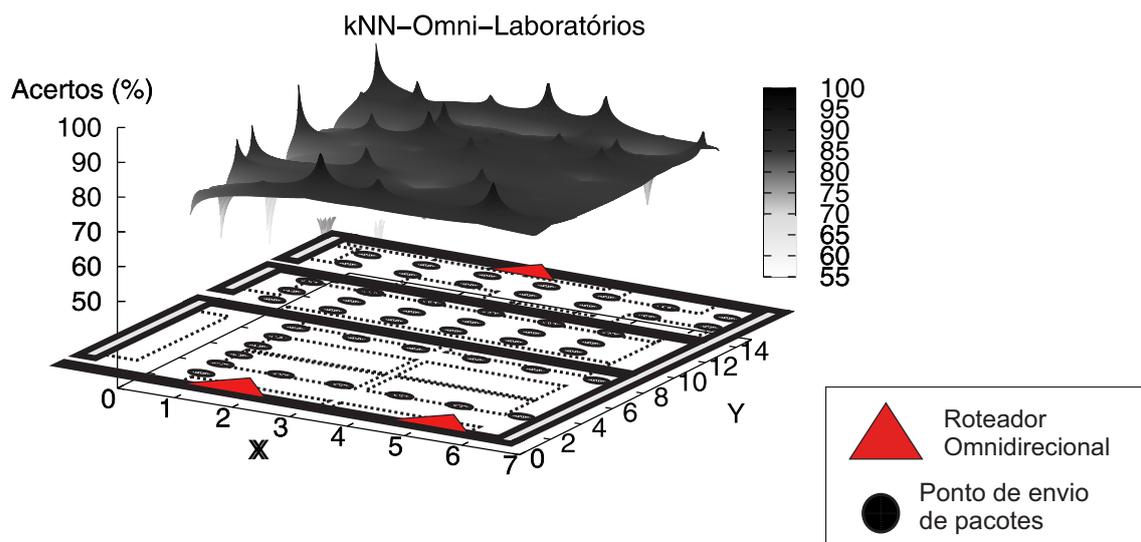
**Figura 5.12.** Função de Distribuição Acumulativa de Erro por Distância de Classificação.

As figuras 5.13 e 5.14 mostram a porcentagem de acertos por pontos dos melhores resultados obtidos pelo sistema no ambiente de localização composto pelos laboratórios: utilizando roteadores direcionais juntos e roteadores omnidirecionais, ambos com algoritmo kNN.

É possível ver na figura 5.13, comparada com 5.14, tanto o impacto de se colocar as antenas direcionais juntas no meio do ambiente em análise, como também a melhoria da classificação para pontos próximos às paredes e objetos da sala. Nestes dois cenários, a precisão de acertos por ponto está entre 85% e 100%. Vale ressaltar que os roteadores direcionais juntos foram melhores do que quando espalhados pelo fato de que neste último caso eles acabavam afetando a qualidade um do outro por estarem na mesma parede e por acabarem sobrepondo muito a área de alcance do sinal um do outro, fato este analisado com a ferramenta Weka através da verificação de ganho de informação de cada atributo da base de dados.



**Figura 5.13.** Distribuição de Acertos por Posição usando Roteadores Direcionais Unidos e algoritmo kNN - Laboratórios.

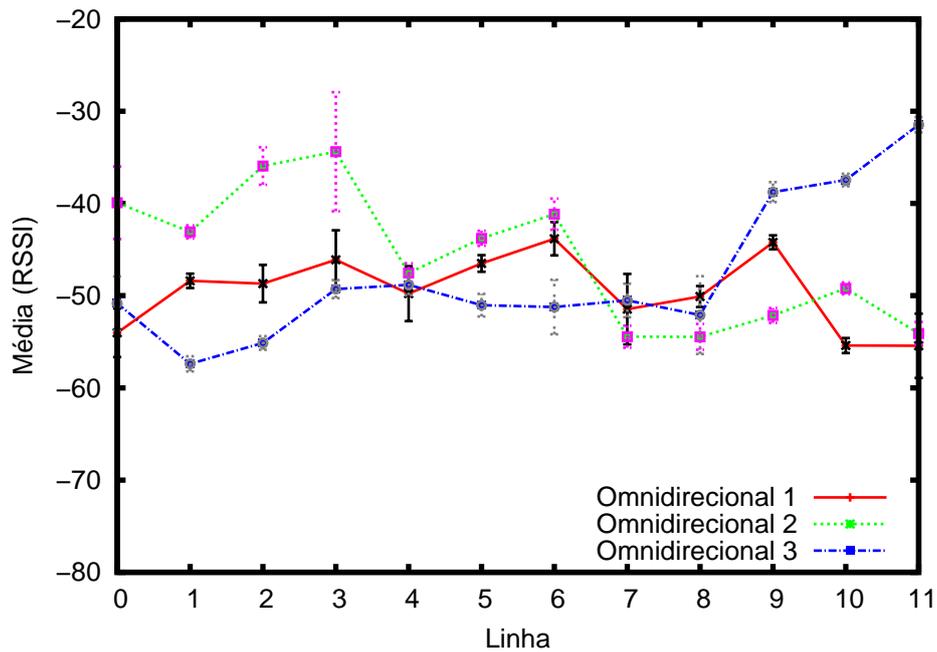


**Figura 5.14.** Distribuição de Acertos por Posição usando Roteadores Omnidirecionais com algoritmo kNN - Laboratórios.

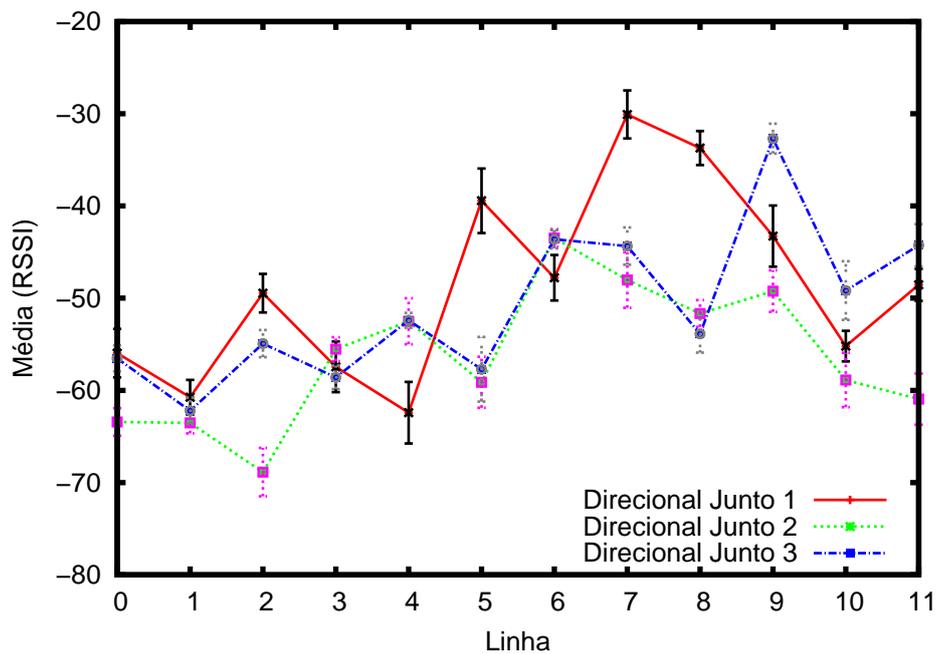
A figura 5.15 apresenta as posições dos roteadores omnidirecionais e direcionais (seja separados ou juntos) no ambiente de localização composto por laboratórios e também um mapeamento do ambiente em linhas e colunas. As figuras 5.16, 5.17 e 5.18 apresentam os aspectos de variação do sinal no ambiente de localização composto pelos laboratórios.

A variação do sinal nas figuras 5.16, 5.17 e 5.18 foram medidas a partir do caminho de pontos percorrido da linha 1 para a linha 10 na coluna 2, onde havia o maior grupo



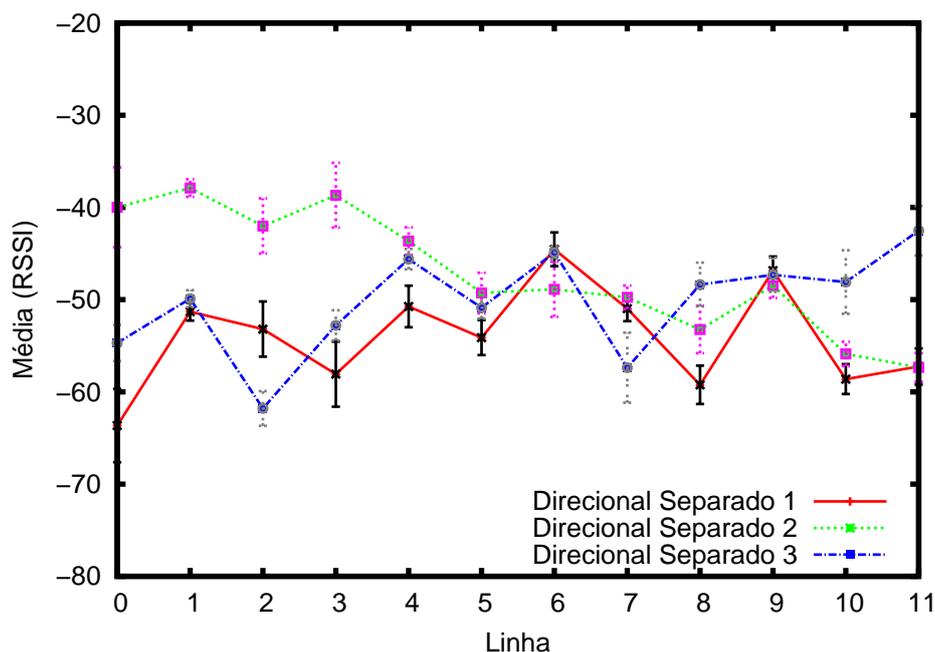


**Figura 5.16.** Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Omnidirecionais nos Laboratórios.



**Figura 5.17.** Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Direcionais Juntos nos Laboratórios.

que o sinal possui um comportamento específico dentro de um determinado ambiente. É possível ver que no ambiente de localização composto pelos laboratórios os



**Figura 5.18.** Variação do Sinal por Posicionamento nas Linhas de 1 a 10 da Coluna 2 - Roteadores Direcionais Separados nos Laboratórios.

valores médios de RSSI para os roteadores direcionais possuem uma variação maior entre posições subsequentes, comparado com a variação dos valores médios de RSSI dos roteadores omnidirecionais para posições subsequentes. Um exemplo disto pode ser visto em 5.17, onde, da distância de 6 metros para 7 metros, há uma variação maior do que a variação existente entre as mesmas distâncias na figura 5.16.

Apesar de o poder de sinal ser mais concentrado nos roteadores direcionais do que nos roteadores omnidirecionais, é possível ver que o desvio padrão em cada posição nos gráficos 5.17 e 5.18 é maior do que nas posições no gráfico 5.16. Tal aspecto permitiu que a classificação de dados utilizando o algoritmo *Naive Bayes* fosse melhor para o sistema com roteadores direcionais do que com roteadores omnidirecionais, já que quanto maior for a variação de sinal em cada posição melhor será a caracterização de cada posição durante a fase de pré-configuração. Vale ressaltar que o sinal dos roteadores direcionais podem sofrer tais condições pelo fato de sua utilização não ser destinada para ambientes internos, onde o sinal pode sofrer diversos efeitos facilmente devido às condições estruturais do ambiente. Os roteadores direcionais em si foram criados para a comunicação de dados em larga distância, e geralmente tais roteadores são instalados em posições altas de maneira que o sinal não sofre tantos efeitos quanto no ambiente interno.

### 5.3 Localização Contextual

A Localização Contextual permite estimar a posição contextual de um determinado dispositivo dentro do ambiente de localização. A posição contextual pode ser definida como uma área dentro do ambiente de localização onde as características de sinal dentro desta área possuem um certo padrão. Isto explica o porquê do termo contextual atrelado a localização, já que todas as condições e características de uma determinada área permitem que o sinal sempre tenha um comportamento específico, identificando assim o seu contexto.

Os experimentos com Localização Contextual foram realizados no ambiente composto pelas salas dos professores do IComp. Durante a etapa de pré-configuração do sistema foram enviados vários pacotes das salas dos professores para que o sistema pudesse registrar o comportamento do sinal em cada área contextual de localização, nesta caso, cada sala. As figuras 5.20 e 5.21 apresentam as salas dos professores onde os experimentos utilizando Localização Contextual foram aplicados.

A tabela 5.2 apresenta os resultados dos experimentos realizados nas salas dos professores do IComp. Estes resultados são baseados na porcentagem de instâncias classificadas corretamente.

**Tabela 5.2.** Porcentagem de instâncias classificadas corretamente nos experimentos realizados nas salas dos professores do IComp.

	<i>Roteadores Omnidirecio- nais</i>	<i>Roteadores Direcionais Espalhados</i>	<i>Roteadores Direcionais Juntos</i>
kNN{k=5}	84,12%	80,85%	73,68%
<i>Naive Bayes</i>	71,68%	80,16%	73,96%

O sistema proposto neste trabalho também obteve bons resultados para o ambiente composto pelas salas de professores do IComp. A forma de organização dos roteadores para um ambiente específico pode ser tratado como fator de impacto para os resultados de localização e isto pode ser visto nos resultados obtidos neste ambiente. O ambiente composto pelas salas dos professores sendo uma área maior faz com que os roteadores direcionais separados consigam gerar melhores resultados para localização do que quando eles estão unidos. Isto acontece pelo fato de os roteadores direcionais não sobreporem bruscamente as áreas de cobertura de sinal de cada um, já que o ambiente de localização é maior.

Um ponto importante que pode ser visto tanto nos resultados para localização

precisa quanto para localização contextual é a atuação de cada classificador sobre a base de dados. É possível ver que o algoritmo kNN possui melhores resultados para o sistema convencional com roteadores omnidirecionais do que o sistema com roteadores direcionais. Isto acontece porque o sistema com roteadores omnidirecionais recebe e irradia o seu poder de sinal em todas as direções e isto faz com que o sinal varie bastante pelo fato de este tipo de roteador não dedicar toda a sua potência para uma determinada área, e seu sinal se torna instável.

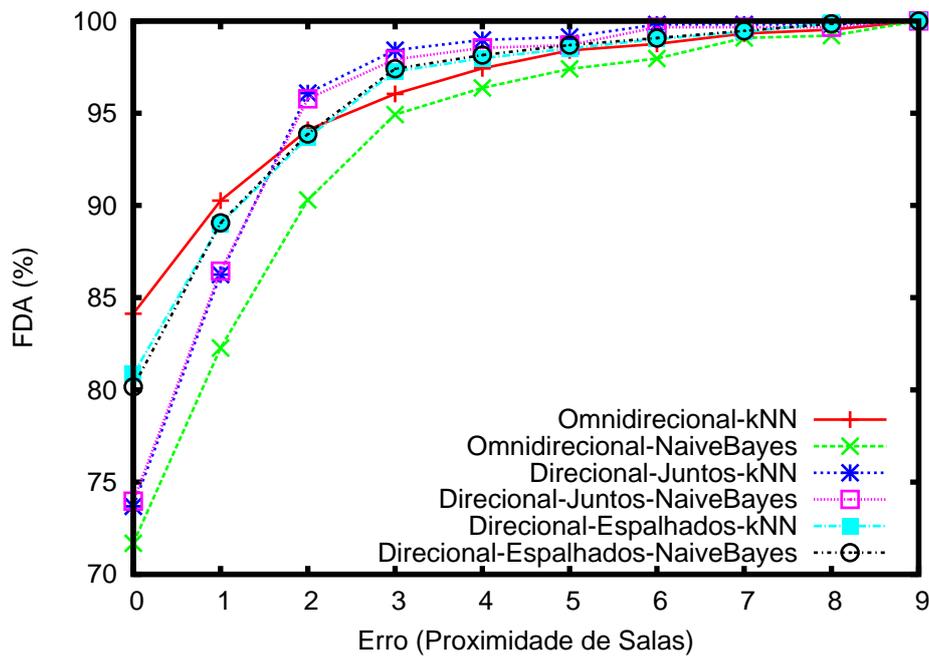
Outro fator que influencia bastante para que o sistema com roteadores omnidirecionais seja melhor que os sistemas com roteadores direcionais quando utilizando o algoritmo kNN é que, como a potência de sinal é reduzida para roteadores omnidirecionais, pode haver muitas instâncias diferentes na base de dados para uma mesma posição, o que gera melhor conhecimento para localização. Já os roteadores direcionais acabam sendo prejudicados por causa do potência de sinal melhorada que é concentrada somente para a área de localização, fazendo com que o RSSI varie menos para diferentes pontos próximos uns dos outros, prejudicando assim o conhecimento para localização.

Por outro lado, o algoritmo *Naive Bayes* obteve melhores resultados com os roteadores direcionais do que com os roteadores omnidirecionais. Isto ocorreu porque os roteadores direcionais recebem e irradiam sinal somente em uma direção específica, e então cada roteador direcional tem mais importância de uma forma probabilística para o classificador *Naive Bayes* pelo fato de seu sinal variar menos do que nos roteadores omnidirecionais. Dessa forma, pode-se dizer que o algoritmo *Naive Bayes* é mais justo diante das circunstâncias de localização do que o algoritmo kNN, que se torna tendencioso as condições do RSSI para uma área específica.

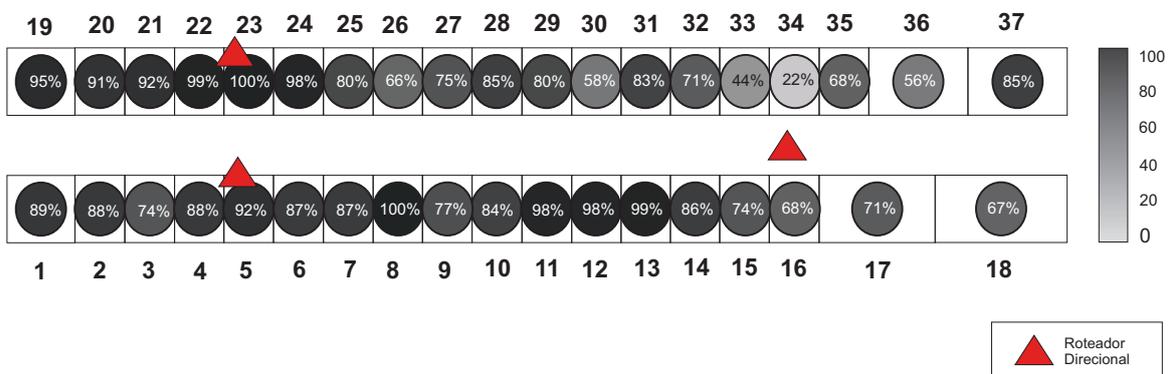
A figura 5.19 mostram a Função de Distribuição Acumulativa (FDA) em proximidade de salas. Em tal figura é possível ver que, por mais que até a precisão com erro de no máximo 1 sala o sistema de localização com roteadores omnidirecionais tenha a melhor curva de precisão, a partir da precisão com erro de no mínimo 2 salas em diante todas as curvas com melhores precisões são dos sistemas de localização com antenas direcionais.

As figuras 5.20, 5.21 mostram a porcentagem de acertos dos melhores resultados obtidos pelo sistema no ambiente de localização composto pelas salas dos professores: roteadores direcionais espalhados e roteadores omnidirecionais, ambos usando algoritmo kNN.

O cenário de experimentos na figura 5.20 comparado com 5.21 mostra o impacto de se utilizar roteadores direcionais espalhados para concentrar a potência do sinal no centro da área de análise. Foi possível obter mais acertos por posição usando



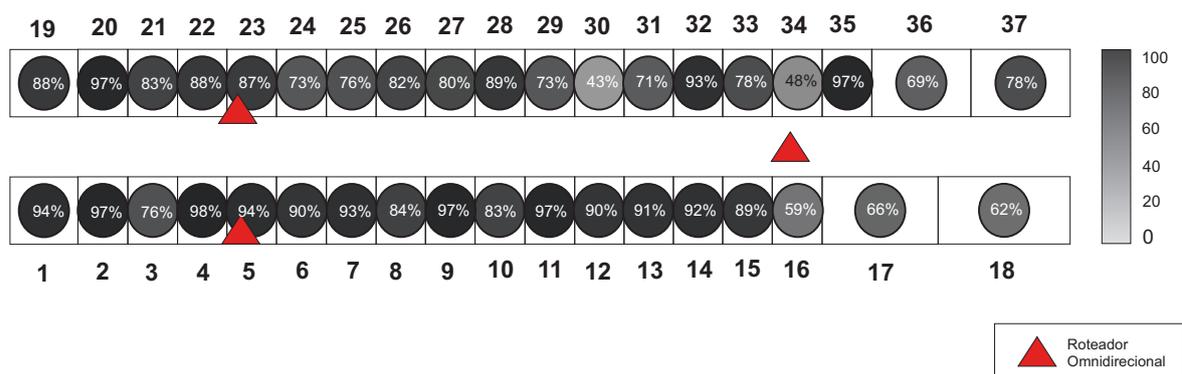
**Figura 5.19.** Gráfico de Função de Distribuição Acumulativa de Erro para Estimativa de Sala.



**Figura 5.20.** Distribuição de Acertos por Sala usando Roteadores Direcionais Espalhados e algoritmo kNN - Salas dos Professores.

roteadores direcionais espalhados para uma área específica do que usando roteadores omnidirecionais. Nos dois casos, a precisão do número de acertos está entre 75% e 100%. Isto foi possível pelo fato de que os roteadores direcionais espalhados focam a potência do sinal para a área que é realmente importante, enquanto os roteadores omnidirecionais perdem poder de sinal para fora da área de localização.

É possível ver na figura 5.20 que a sala com pior resultado de classificação utilizando roteadores direcionais separados é a sala 34, que teve precisão de 22%. Isto



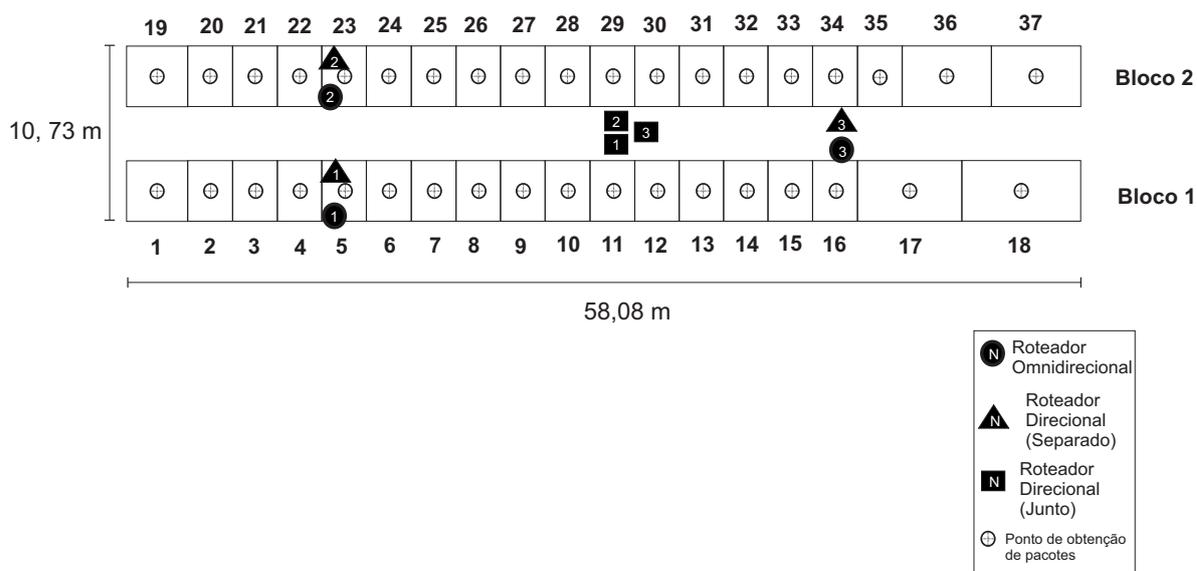
**Figura 5.21.** Distribuição de Acertos por Sala usando Roteadores Omnidirecionais e algoritmo kNN - Sala dos Professores.

ocorreu devido ao fato de que, quando os experimentos foram realizados, esta sala era utilizada como depósito de diversos materiais como vidro, gabinetes, armários e outros, o que afetou bastante a precisão de localização em tal sala. Pode-se ver também que a sala 33, localizada ao lado da 34, sofreu um impacto também devido a tais condições, chegando à precisão de 44%.

A figura 5.21 representando a classificação com roteadores omnidirecionais mostra que a sala 34 já possui um resultado melhor, chegando a precisão de 48%, mas nesta figura a sala com pior precisão é a sala 30, que teve precisão de 43%, por ser uma sala que possuía armários dispostos na parede entre a sala 30 e a 31, o que afetou bastante a classificação por se tratar de uma sala localizada no meio do ambiente de localização. Se a mesma sala estivesse por trás dos roteadores o impacto dos armários não afetaria tanto.

A figura 5.22 apresenta as posições dos roteadores omnidirecionais e direcionais (seja separados ou juntos) no ambiente de localização composto pelas salas dos professores, a identificação dos blocos e também a numeração das salas. As figuras 5.23, 5.24 e 5.25 apresentam os aspectos de variação do sinal no ambiente de localização composto pelas salas dos professores no bloco 1, e as figuras 5.26, 5.27 e 5.28 apresentam os aspectos de variação do sinal no bloco 2.

Como o ambiente composto pelas salas dos professores é maior, é possível ver claramente o comportamento do sinal nos roteadores a partir dos pacotes enviados de cada sala subsequente. O sentido de envio de pacotes em cada sala foi a partir da sala 1, em direção ao roteador 3, até a sala 18. Tal situação pode ser vista na figura 5.23, onde o roteador 3 registra um valor de RSSI cada vez maior conforme os pacotes são enviados mais próximos dele, enquanto os roteadores 1 e 2 vão registrando valores



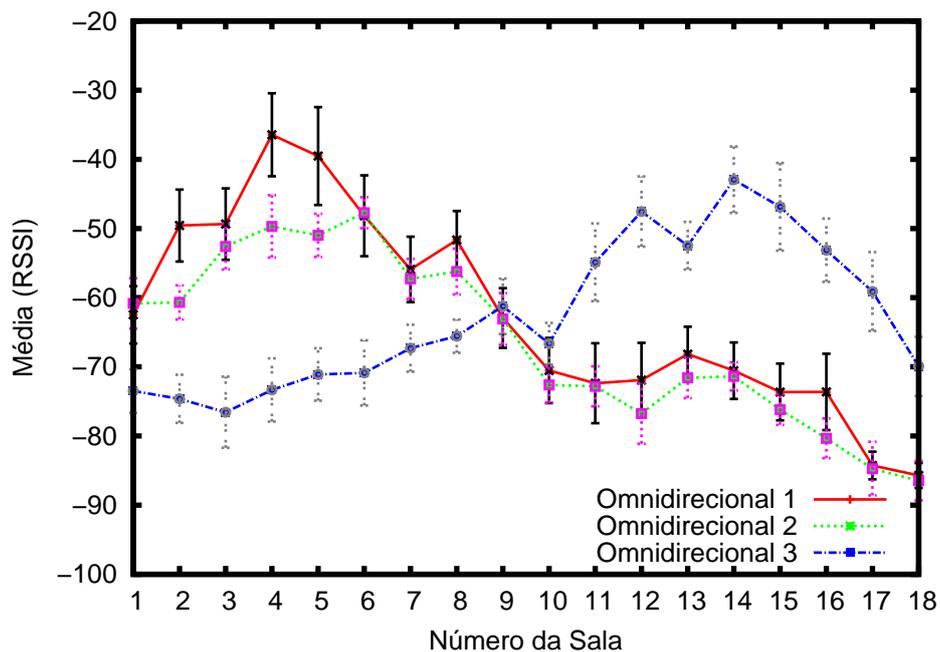
**Figura 5.22.** Organização dos Roteadores no Ambiente de Localização Composto pelas Salas dos Professores.

de RSSI cada vez menores. Pode ser ver também que da sala 1 a sala 4 o RSSI médio cresce para os roteadores 1 e 2 e isto acontece porque a sala 4 antecede a sala onde os mesmos estão instalados. Da sala 5 em diante o RSSI médio dos roteadores 1 e 2 só diminui, enquanto para o roteador 3 o RSSI médio cresce desde a sala 1.

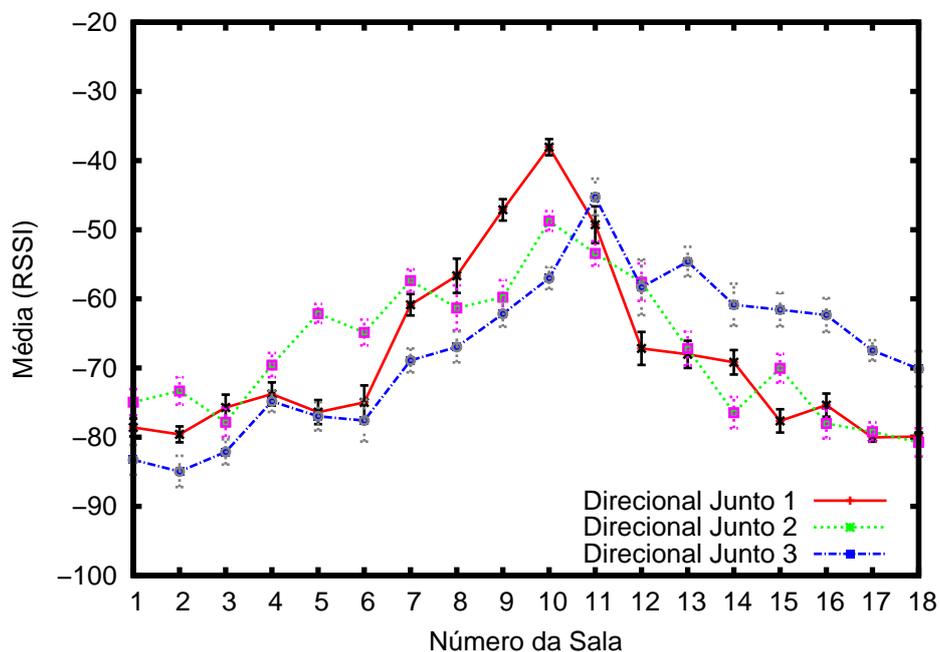
Na figura 5.24 é possível ver um aspecto interessante do sistema com roteadores direcionais juntos. O RSSI médio cresce para os três roteadores até a sala 9, que antecede a sala onde os roteadores foram instalados na frente. Logo após a sala 9, é possível ver as três curvas caindo, conforme os pacotes são enviados de cada vez mais distante dos roteadores direcionais juntos.

Na figura 5.25 é possível ver que a curva do RSSI médio do roteador 1 até a sala 4 cresce mais que a curva do roteador 2, pelo fato de as salas de 1 a 18 serem todas no mesmo bloco de salas onde o roteador 1 está instalado. Desta forma, o roteador 2 que está no outro bloco de salas foca em uma direção diferente. A partir da sala 5, o RSSI médio dos roteadores 1 e 2 já possuem um comportamento mais semelhante, já que agora os pacotes estão sendo enviados nas salas localizadas em frente aos dois roteadores, e enquanto isto o roteador 3 possui um RSSI médio cada vez maior, conforme os pacotes são enviados das salas cada vez mais próximas a ele.

Outro ponto interessante é que agora numa área maior de experimento e com mais salas, o roteador omnidirecional apresentou maior desvio padrão para o RSSI médio obtido pelo envio de pacotes em cada sala, comparado com os desvio padrão dos



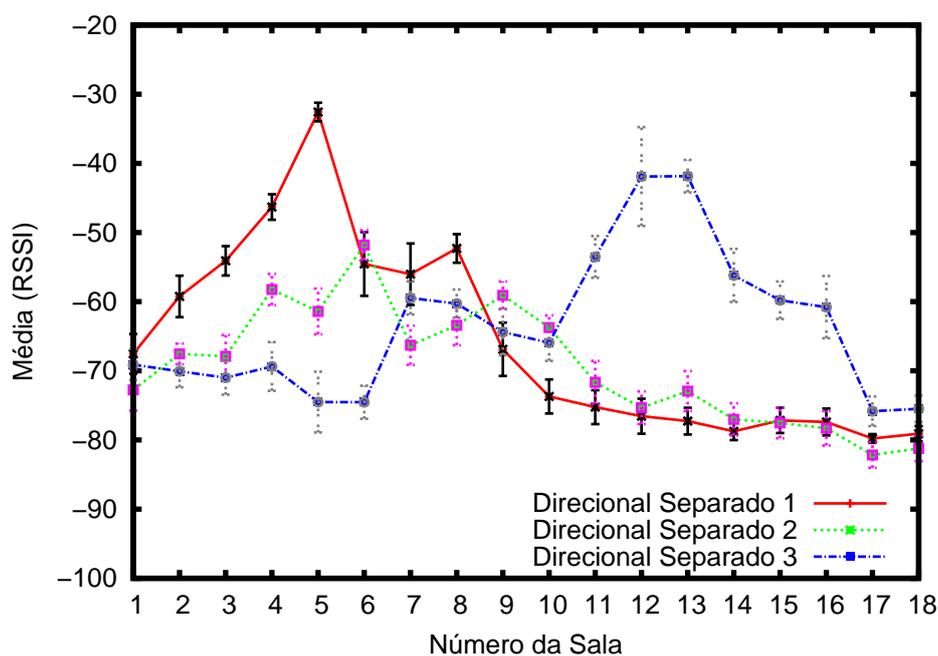
**Figura 5.23.** Variação do Sinal por Sala (Bloco 1) - Roteadores Omnidirecionais nas Salas dos Professores.



**Figura 5.24.** Variação do Sinal por Sala (Bloco 1) - Roteadores Direcionais Juntos nas Salas dos Professores.

cenários que utilizam roteadores direcionais.

A figura 5.26 já começa a mostrar o comportamento da variação do sinal no

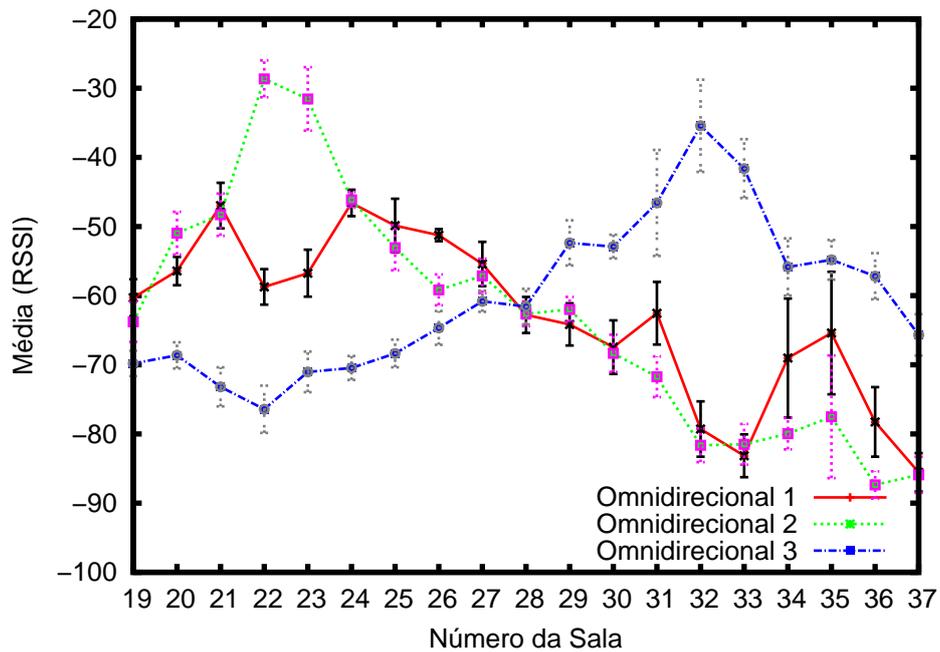


**Figura 5.25.** Variação do Sinal por Metro (Bloco 1) - Roteadores Direcionais Separados nas Salas dos Professores.

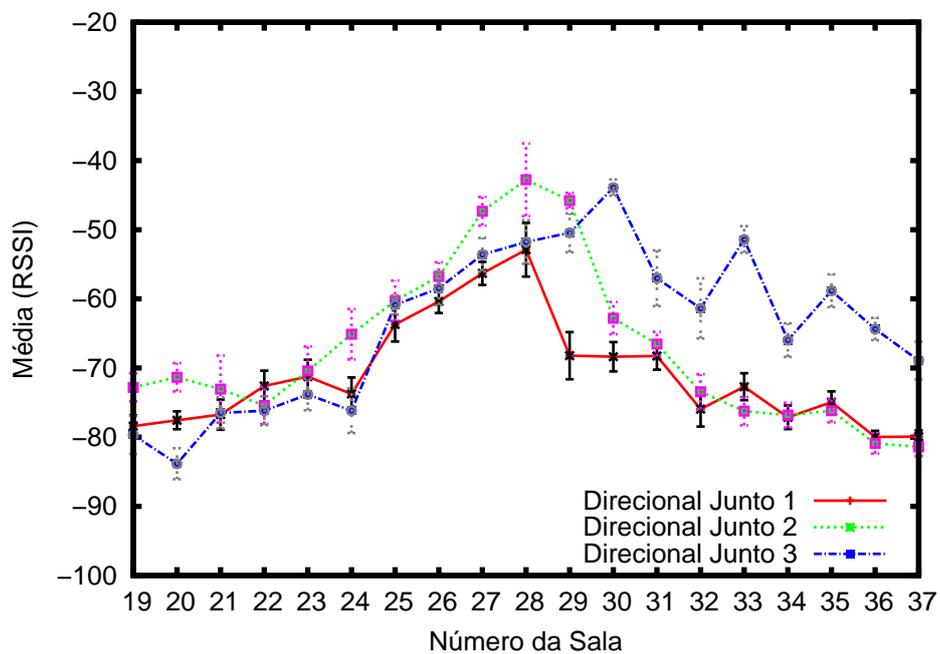
bloco 2. Como citado acima, a sala 34 era utilizada para depósito e por isto não teve resultados tão satisfatórios de precisão, e se pode ver na figura 5.26 que esta sala causa um comportamento anormal nas curvas, fazendo com que a curva do roteador omnidirecional 3 caia bruscamente enquanto as curvas dos roteadores omnidirecionais 1 e 2 subam. Pode-se concluir então que a disposição dos objetos nesta sala fizeram com que os roteadores 1 e 2 fossem favorecidos enquanto o roteador 3 foi prejudicado.

Na figura 5.27 é possível ver mais uma vez o aspecto interessante do sistema com roteadores direcionais juntos. O RSSI médio cresce para os três roteadores até a sala 28, que antecede a sala onde os roteadores foram instalados na frente. Logo após a sala 28, é possível ver as três curvas caindo, conforme os pacotes são enviados de cada vez mais distante dos roteadores direcionais juntos.

Já na figura 5.28, pode ser ver o comportamento do sistema de localização com antenas direcionais separadas no bloco 2. É possível verificar que o roteador direcional 2 agora possui maior média de RSSI pelo fato de as salas estarem dispostas no mesmo bloco que o roteador está instalado. Isto também explica o fato de a curva do roteador direcional 1 possuir um comportamento semelhante a curva do roteador direcional 1, só que com uma média de RSSI mais baixa no início, por ser um roteador que está no bloco 1. Conforme a aproximação da sala para o roteador direcional 3, a sua curva do RSSI médio vai aumentando, enquanto a curva dos roteadores direcionais 1 e 2 vão



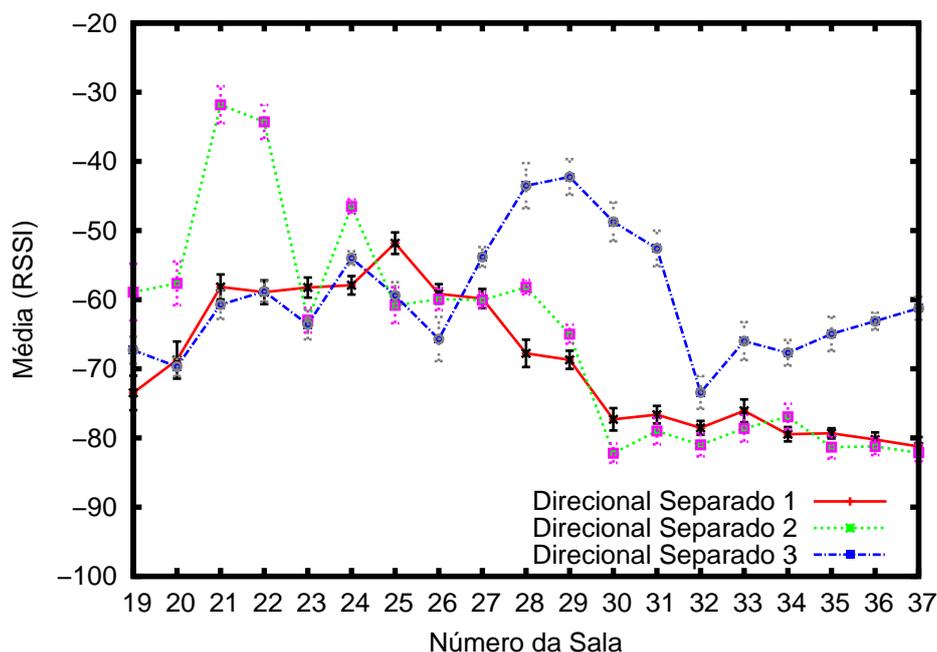
**Figura 5.26.** Variação do Sinal por Sala (Bloco 2) - Roteadores Omnidirecionais nas Salas dos Professores.



**Figura 5.27.** Variação do Sinal por Sala (Bloco 2) - Roteadores Direcionais Juntos nas Salas dos Professores.

caindo.

Por fim, pode-se ver também que os gráficos de variação do sinal da Localização



**Figura 5.28.** Variação do Sinal por Metro (Bloco 2) - Roteadores Direcionais Separados nas Salas dos Professores.

Contextual tem desvio padrão menor para o sistema de localização quando o mesmo utilizava antenas direcionais, enquanto na Localização Precisa a utilização dos roteadores direcionais permitia que o desvio padrão fosse maior. Já nos gráficos de variação do sinal onde o sistema de localização utilizava antenas omnidirecionais, é possível ver que na Localização Contextual houve um desvio padrão maior do que na Localização Precisa. Tal comportamento valida o fato de os roteadores direcionais terem sido criados para trabalhar com grandes distâncias, já que na Localização Precisa o espaço de experimentos era bem menor.

## 5.4 Conclusão

Esta seção avaliou os resultados obtidos a partir das abordagens de localização com roteadores direcionais utilizadas em cada um dos ambientes onde os experimentos deste trabalho foram realizados.

Em cada um dos ambientes utilizados para experimento deste trabalho foi utilizado um tipo específico de localização. No ambiente menor, composto pelos laboratórios do IComp, foi utilizada a Localização Precisa, que permite estimar de que ponto mapeado no ambiente de localização um pacote foi enviado ao sistema. Já o ambiente maior composto pelas salas dos professores do IComp, foi utilizada a Localização Con-

textual, que permite estimar a área de contexto, neste caso as sala, de onde um pacote foi enviado ao sistema.

Para cada ambiente foram analisados a quantidade de instâncias classificadas corretamente, a organização de roteadores juntamente com o algoritmo de classificação de dados que melhor se adequa as configurações do ambiente de localização, características do sistema em cada parte do ambiente de localização e comportamento do sinal.

Apesar de o sistema proposto neste trabalho ter conseguido bons resultados em comparação com o sistema convencional composto por roteadores omnidirecionais, ele não se saiu melhor que o mesmo, mas nos mostrou que ele pode ser utilizado como alternativa para tal proposta.

Além disto, o comportamento do sinal num ambiente real não foi o que se esperava com base na fundamentação teórica sobre o mesmo, já que mesmo em direções contrárias ao direcionamento dos roteadores ainda foi possível obter níveis de sinal consideráveis. De qualquer forma, foi possível ver que direcionamento das antenas em um sistema de localização pode trazer qualidade ao sistema, já que o direcionamento de cada roteador pode agregar melhores informações de localização ao sistema, como vista a partir da utilização do algoritmo Naive Bayes para classificação de dados.

# Capítulo 6

## Conclusões

Este trabalho propôs um novo sistema de posicionamento para ambientes internos usando roteadores com antenas direcionais. Este sistema é baseado em treinamento e classificação de dados, e foi analisado o comportamento do sistema utilizando cada um dos algoritmos *Naive Bayes* e *kNN*, separadamente, como classificadores para análise de dados comparativa. Um dos pontos também abordados neste trabalho foi a organização dos roteadores direcionais de acordo com as características do ambiente.

Os experimentos deste sistema foram realizados em dois ambientes diferentes: laboratórios e salas dos professores do Instituto de Computação da UFAM (IComp). O sistema foi avaliado utilizando duas formas diferentes de organização dos roteadores direcionais e seus resultados foram comparados com os resultados do sistema de posicionamento convencional, composto de roteadores omnidirecionais instalados nas mesmas áreas de localização.

Os resultados obtidos durante os experimentos mostram que os roteadores direcionais podem ser utilizados para localização em ambientes internos porque eles possuem maior potência de sinal concentrada para uma área específica, menor variação de valores de RSSI e podem ser utilizados para analisar uma área maior de posicionamento. Concluiu-se com este trabalho que a forma que os roteadores estão organizados e o algoritmo utilizado para classificação de dados afetam a precisão do sistema de localização, já que tais fatores são afetados pelas condições estruturais do ambiente, e pela utilização das características do sinal no ambiente e no sistema para composição do conhecimento.

Um exemplo de como a organização dos roteadores pode afetar a precisão do sistema pode ser vista a partir da análise comparativa entre as condições dos experimentos realizados neste trabalho. Os roteadores omnidirecionais possuem maior variação de sinal quando utilizados em ambientes de localização maiores, enquanto os roteadores

direcionais possuem maior variação de sinal quando utilizados em ambientes de localização menores.

Apesar de os melhores resultados de precisão na estimativa de localização terem sido obtidos quando o sistema utilizava o algoritmo *kNN* para classificação, os resultados obtidos no sistema utilizando o algoritmo Naive Bayes não ficaram distantes. No que diz respeito a porcentagem das instâncias classificadas corretamente nos experimentos realizados nos laboratórios do IComp, podemos ver que o melhor resultado do sistema utilizando o algoritmo *kNN* foi de 84,26% com a composição dos roteadores formada por roteadores omnidirecionais. A utilização do algoritmo Naive Bayes nas mesmas condições obteve 81,86% com a composição dos roteadores formada por roteadores direcionais juntos no ambiente de localização. Vale ressaltar que com os roteadores direcionais juntos utilizados no sistema de localização juntamente com o algoritmo *kNN* foi possível obter 83,93% de precisão na classificação correta das instâncias.

Já no que diz respeito à porcentagem das instâncias classificadas corretamente nos experimentos realizados nas salas dos professores, é possível ver que o melhor resultado do sistema utilizando o algoritmo *kNN* foi de 84,12% com a composição dos roteadores formada por roteadores omnidirecionais. A utilização do algoritmo Naive Bayes considerando as mesmas características do ambiente obteve 80,16% com a composição dos roteadores formada por roteadores direcionais espalhados pelo ambiente de localização. É importante ressaltar que com os roteadores direcionais espalhados utilizados no sistema de localização juntamente com o algoritmo *kNN* foi possível obter 80,85% de precisão na classificação correta das instâncias.

Pode-se considerar com base em tais pontos que, caso haja a necessidade de localização contextual em um ambiente grande, os roteadores direcionais podem ser organizados separados. Se a necessidade de localização precisa se der em um ambiente pequeno, os roteadores direcionais podem estar juntos. É importante considerar que a utilização do algoritmo Naive Bayes nestes dois casos pode render bons resultados ao sistema de localização, podendo ser utilizada como alternativa ao algoritmo *kNN* bastante utilizado na literatura. Não há dúvidas de que o algoritmo *kNN* conseguiu os melhores resultados em classificação mas o objetivo deste trabalho não é propor a exclusão deste algoritmo para localização, mas sim mostrar que o algoritmo Naive Bayes pode ser utilizado como opção para sistemas de localização.

Este trabalho nos proporcionou também a conclusão de que os roteadores direcionais não funcionaram totalmente como esperávamos já que é possível ver nos resultados que áreas distantes da parte traseira das antenas direcionais ainda conseguem receber sinal com potência suficiente para se ouvir os pacotes enviados por trás das mesmas.

Tal fator foi importante para analisarmos o comportamento ideal e real sobre tal tipo roteador, e também para analisarmos como podemos usar tais informações de comportamento para melhor obter resultados de posicionamento dentro dos ambientes de localização.

## **6.1 Trabalhos Futuros**

Como trabalhos futuros, pretende-se analisar o uso de roteadores omnidirecionais e direcionais em um mesmo sistema de localização para ambientes internos e estudar o impacto dos dois tipos roteadores atuando numa mesma área para um mesmo sistema.

Pretende-se também a realização de estudos sobre a utilização de classificadores em série e em pipeline para a estimativa de localização em ambientes internos. Tal estudo permitirá a análise sobre outros algoritmos destinados a classificação de dados que podem contribuir para o objetivo deste sistema.

# Referências Bibliográficas

- Bahl, P. & Padmanabhan, V. (2000). Radar: An in-building rf-based user location and tracking system. Em *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pp. 775 –784 vol.2.
- Balaguer, B.; Erinc, G. & Carpin, S. (2012). Combining classification and regression for wifi localization of heterogeneous robot teams in unknown environments. Em *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 3496–3503.
- Barsocchi, P.; Lenzi, S.; Chessa, S. & Giunta, G. (2009). A novel approach to indoor rssi localization by automatic calibration of the wireless propagation model. Em *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pp. 1 –5.
- Boukerche, A.; Oliveira, H.; Nakamura, E. & Loureiro, A. (2007). Localization systems for wireless sensor networks. *Wireless Communications, IEEE*, 14(6):6 –12.
- Chang, C.-H. & Liao, W. (2009). Revisiting relative location estimation in wireless sensor networks. Em *Communications, 2009. ICC '09. IEEE International Conference on*, pp. 1 –5.
- Chaurasiya, V.; Lavavanshi, R.; Verma, S.; Nandi, G. & Srivastava, A. (2009). Localization in wireless sensor networks using directional antenna. Em *Advance Computing Conference, 2009. IACC 2009. IEEE International*, pp. 131 –134.
- Chengdong, W.; Shifeng, C.; Yunzhou, Z.; Long, C. & Hao, W. (2011). A rssi-based probabilistic distribution localization algorithm for wireless sensor network. Em *Information Technology and Artificial Intelligence Conference (ITAIC), 2011 6th IEEE Joint International*, volume 1, pp. 333 –337.

- Cherntanomwong, P. & Suroso, D. (2011). Indoor localization system using wireless sensor networks for stationary and moving target. Em *Information, Communications and Signal Processing (ICICS) 2011 8th International Conference on*, pp. 1 –5.
- Chiou, Y.-S.; Wang, C.-L.; Yeh, S.-C. & Su, M.-Y. (2009). Design of an adaptive positioning system based on wifi radio signals. *Comput. Commun.*, 32(7-10):1245--1254.
- Dawes, B. & Chin, K.-W. (2011). A comparison of deterministic and probabilistic methods for indoor localization. *Journal of Systems and Software*, 84(3):442 – 451.
- Ding, X.; Zhao, H.; Zhu, J.; Zhang, K. & Li, D. (2011). A novel localization algorithm based on rssi for wireless sensor networks. Em *Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on*, pp. 1 –4.
- Fang, Z.; Zhao, Z.; Geng, D.; Xuan, Y.; Du, L. & Cui, X. (2010). Rssi variability characterization and calibration method in wireless sensor network. Em *Information and Automation (ICIA), 2010 IEEE International Conference on*, pp. 1532 –1537.
- Fischler, M. A. & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381--395.
- Gogolak, L.; Pletl, S. & Kukolj, D. (2011). Indoor fingerprint localization in wsn environment based on neural network. Em *Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium on*, pp. 293 –296.
- Gunawan, M.; Li, B.; Gallagher, T.; Dempster, A. G. & Retscher, G. (2012). A new method to generate and maintain a wifi fingerprinting database automatically by using rfid. Em *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, pp. 1 –6.
- Heurtefeux, K. & Valois, F. (2012). Is rssi a good choice for localization in wireless sensor network? Em *Advanced Information Networking and Applications (AINA), 2012 IEEE 26th International Conference on*, pp. 732 –739.
- Honkavirta, V.; Perala, T.; Ali-Loytty, S. & Piche, R. (2009). A comparative survey of wlan location fingerprinting methods. Em *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on*, pp. 243 –251.

- Hui, T.; Shuang, W. & Huaiyao, X. (2007). Localization using cooperative aoa approach. Em *Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on*, pp. 2416 –2419.
- Kawauchi, K.; Miyaki, T. & Rekimoto, J. (2009). Directional beaconing: A robust wifi positioning method using angle-of-emission information. Em *Proceedings of the 4th International Symposium on Location and Context Awareness, LoCA '09*, pp. 103--119, Berlin, Heidelberg. Springer-Verlag.
- Keller, J.; Gray, M. & Givens, J. (1985). A fuzzy k-nearest neighbor algorithm. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-15(4):580–585.
- Khan, U.; Pahlavan, K. & Makarov, S. (2011). Comparison of toa and rss based techniques for rf localization inside human tissue. Em *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 5602 –5607.
- Le, T.; Le, H.; Nguyen, N.; Tran, D. & Nguyen, N. (2011). Convert wi-fi signals for fingerprint localization algorithm. Em *Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on*, pp. 1 –5.
- Lee, K.-C. & Lampe, L. (2011). Indoor cell-level localization based on rssi classification. Em *Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on*, pp. 000021 –000026.
- Lim, C.-H.; Wan, Y.; Ng, B.-P. & See, C.-M. (2007). A real-time indoor wifi localization system utilizing smart antennas. *Consumer Electronics, IEEE Transactions on*, 53(2):618 –622.
- Lu, X.; Wicker, F. D.; Towsley, D.; Xiong, Z. & Lio', P. (2010). Detection probability estimation of directional antennas and omni-directional antennas. *Wirel. Pers. Commun.*, 55(1):51--63.
- Menendez, P.; Campomanes, C.; Trawinski, K. & Alonso, J. (2011). Topology-based indoor localization by means of wifi fingerprinting with a computational intelligent classifier. Em *Intelligent Systems Design and Applications (ISDA), 2011 11th International Conference on*, pp. 1020–1025.
- Meng, W.; Xiao, W.; Ni, W. & Xie, L. (2011). Secure and robust wi-fi fingerprinting indoor localization. Em *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, pp. 1 –7.

- Ou, C.-H. (2011). A localization scheme for wireless sensor networks using mobile anchors with directional antennas. *Sensors Journal, IEEE*, 11(7):1607 –1616.
- Ramanathan, R. (2001). On the performance of ad hoc networks with beamforming antennas. Em *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing, MobiHoc '01*, pp. 95--105, New York, NY, USA. ACM.
- Robles, J.; Deicke, M. & Lehnert, R. (2010). 3d fingerprint-based localization for wireless sensor networks. Em *Positioning Navigation and Communication (WPNC), 2010 7th Workshop on*, pp. 77 –85.
- Scholl, P.; Kohlbrecher, S.; Sachidananda, V. & Van Laerhoven, K. (2012). Fast indoor radio-map building for rssi-based localization systems. Em *Networked Sensing Systems (INSS), 2012 Ninth International Conference on*, pp. 1 –2.
- Sugano, M. (2006). Indoor localization system using rssi measurement of wireless sensor network based on zigbee standard. Em *Wireless and Optical Communications*, pp. 1--6. IASTED/ACTA Press.
- Swar, P.; Ye, Y.; Ghaboosi, K. & Pahlavan, K. (2012). On effect of transmit power variance on localization accuracy in wireless capsule endoscopy. Em *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pp. 2119 –2123.
- Umehira, M. & Ohtomo, Y. (2010). Impact of antenna directivity for carrier sensing in high density wlan using adaptive directional antenna. Em *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, pp. 1–5.
- Xu, C.; Firner, B.; Zhang, Y.; Howard, R.; Li, J. & Lin, X. (2012). Improving rf-based device-free passive localization in cluttered indoor environments through probabilistic classification methods. Em *Proceedings of the 11th international conference on Information Processing in Sensor Networks, IPSN '12*, pp. 209--220, New York, NY, USA. ACM.
- Yang, C.-L.; Bagchi, S. & Chappell, W. (2005). Location tracking with directional antennas in wireless sensor networks. Em *Microwave Symposium Digest, 2005 IEEE MTT-S International*, p. 4 pp.