



UNIVERSIDADE FEDERAL DO AMAZONAS
Faculdade de Tecnologia
Programa de Pós-Graduação em Engenharia Elétrica

Alípio de Sales Carvalho

**Localização em Ambientes Internos utilizando
redes IEEE 802.11 e algoritmo WKNN**

Manaus - AM

2021

Alípio de Sales Carvalho

Localização em Ambientes Internos utilizando redes IEEE 802.11 e algoritmo WKNN

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

Programa de Pós-Graduação em Engenharia Elétrica

Universidade Federal do Amazonas

Orientador: Prof. Dr. Celso Barbosa Carvalho

Coorientador: Prof. Dr. José Ferreira de Rezende

Manaus - AM

2021

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

C331l Carvalho, Alípio de Sales
Localização em Ambientes Internos utilizando redes IEEE 802.11 e algoritmo WKNN / Alípio de Sales Carvalho, Celso Carvalho. 2021
58 f.: il. color; 31 cm.

Orientador: Celso Barbosa Carvalho
Coorientador: José Ferreira de Rezende
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. IIs. 2. Rss. 3. Wi-fi. 4. Wknn. I. Carvalho, Celso. II. Universidade Federal do Amazonas III. Título

ALÍPIO DE SALES CARVALHO

**LOCALIZAÇÃO EM AMBIENTES INTERNOS UTILIZANDO
REDES IEEE 802 11 E ALGORITMO WKNN**

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

Aprovado em 01 de outubro de 2021.

BANCA EXAMINADORA

DocuSigned by:

Celso Barbosa Carvalho

Prof. Dr. Celso Barbosa Carvalho, Presidente

Universidade Federal do Amazonas

DocuSigned by:

André Luiz Duarte Cavalcante

Prof. Dr. André Luiz Duarte Cavalcante, Membro

Universidade Federal do Amazonas

DocuSigned by:

Marcel William Rocha da Silva

Prof. Dr. Marcel William Rocha da Silva, Membro

Universidade Federal Rural do Rio de Janeiro

Dedico este trabalho aos meus pais e a Deus.

Agradecimentos

A Deus por permitir esta grande oportunidade, pela força, saúde e fé na conquista dos meus objetivos.

Ao meu orientador, professor Celso Carvalho pela oportunidade, orientações e paciência na consolidação deste trabalho.

Ao meu co-orientador, professor José Ferreira de Rezende pelo suporte nas orientações.

Ao professor Waldir Sabino da Silva Júnior pelo grande incentivo para a realização deste curso de Mestrado.

A minha esposa Elinelma, pelo apoio emocional nos momentos difíceis.

Ao meu saudoso pai Paulo Carvalho, minha querida e amada mãe Ivanilda Sales, meu amado filho Jean Lukas e meus irmãos Erika, Jeane e Paulo Alexandre que sempre me apoiaram e torceram por mim nesta jornada.

Aos meus amigos pelo apoio emocional e companheirismo, em especial ao Jacqson Ribeiro pelo suporte no desenvolvimento de algumas aplicações.

A coordenação do Programa de pós-Graduação do curso de Engenharia Elétrica – PPGEE e Universidade Federal do Amazonas – UFAM por todo apoio neste curso e por disponibilizar os laboratórios para realização dos experimentos.

As instituições que apoiaram esta pesquisa, especificamente pelo financiamento da Samsung Eletrônica da Amazônia Ltda, conforme previsto no Art. 48 do decreto no 6.008/2006, nos termos da Lei Federal no 8.387/1991, através de convênio No 004, firmado com o CETELI/UFAM; a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); pela Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM/Programa PPP), CT-INFRA, MCT/CNPQ, SECTI-AM e Governo do Estado do Amazonas

"A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original."

(Albert Einstein)

Resumo

Nos últimos anos, com a evolução das tecnologias de informação e comunicação, surgiram novas técnicas de localização de objetos ou usuários em ambientes internos. Apesar disso, os sistemas de localização interna (ILSs) continuam sendo um desafio, pois os trabalhos da literatura dependem de sensores com tecnologia específica e alto custo para obter a acurácia adequada para o sistema. Além disso, muitos trabalhos da literatura relacionados a ILS apresentam limitações em sua interface homem máquina (*Human Machine Interface - HMI*), dificultando a operação e a usabilidade. Nesta dissertação de mestrado apresenta-se proposta de ILS de alta capacidade computacional, fácil operação e usabilidade para determinar a localização de um usuário, a partir de seu smartphone, utilizando assinaturas dos sinais de RF (*Radio Frequency*) transmitidos em redes Wi-Fi pré-existent no ambiente, uma base de dados criada a partir das assinaturas de RSSI coletadas e algoritmo de aprendizagem supervisionada WKNN (*Weighted K-Nearest Neighbor*). O ILS proposto apresentou erro de localização RMSE (*Root Mean Square Error*) de 2,87 m, superando trabalhos recentes da literatura, onde o menor erro de localização RMSE obtido foi de 4,31 m, além disso, nesta dissertação propõe-se fornecer facilidades de configuração de HMI.

Palavras-chave: ILS; RSS; Redes Wi-Fi; Algoritmo WKNN

Abstract

In recent years, with the evolution of information and communication technologies, new techniques for locating objects or users in indoor environments have emerged. Despite this, indoor location systems (ILSs) continue to be a challenge since some proposals depend on sensors with specific technology and high cost to obtain adequate system accuracy. In addition, many existing ILS proposals have limitations in their Human-Machine Interface (HMI), making operation and usability difficult. In this work we present a high computational capacity, easy operation and simple usability system. The proposed ILS determines the location of an user, from his smartphone, using signatures of pre-existing Wi-Fi access points in the environment, a database created in an *off-line* phase and WKNN (Weighted K-Nearest Neighbor) supervised learning algorithm. The proposed ILS showed a RMSE (Root Mean Square Error) location error of 2,87 m, surpassing recent work in the literature, where the lowest RMSE location error obtained was 4,31 m. In addition, our proposal provides HMI configuration facilities.

Keywords:ILS; RSS; Wi-Fi; WKNN

Lista de ilustrações

Figura 1 – Método de trilateração.	20
Figura 2 – Método <i>fingerprinting</i>	21
Figura 3 – Conjunto M de APs, $m = 1, \dots, M $	22
Figura 4 – Coleta de 1 amostra de RSS de cada AP do ambiente, em 1 RP do ambiente.	23
Figura 5 – Conjunto R de RPs ($RP_1, \dots, RP_{ R }$), com $r = 1, \dots, R $	23
Figura 6 – Conjunto de FPs, formado por elementos $FP_{i,r}$ onde $i = 1$ instância de <i>fingerprinting</i> por RP_r	24
Figura 7 – Estimativa de localização de TP na fase <i>on-line</i>	25
Figura 8 – Exemplo de método <i>fingerprinting</i> utilizando algoritmo KNN.	27
Figura 9 – Mapa de rádio de exemplo usando método <i>fingerprinting</i>	28
Figura 10 – Determinação de coordenadas usando método <i>fingerprinting</i> e KNN.	29
Figura 11 – Determinação de coordenadas usando método <i>fingerprinting</i> e WKNN.	31
Figura 12 – Arquitetura geral.	36
Figura 13 – Fluxograma da aplicação do smartphone e do servidor.	37
Figura 14 – Validação Cruzada κ -fold do modelo proposto.	40
Figura 15 – Telas da aplicação EZNAV.	43
Figura 16 – Ambiente de experimentação no bloco CETELI.	44
Figura 17 – Erro médio de Localização MAE versus valor \hat{K}	47
Figura 18 – Erro médio MAE versus número amostras RSS.	48
Figura 19 – Erro RMSE x valor do parâmetro \hat{K}	49
Figura 20 – CDF versus erro médio MAE.	50
Figura 21 – Quantidade de APs visíveis em cada TP.	51
Figura 22 – Número Aps visíveis versus erro médio MAE.	52

Lista de tabelas

Tabela 1 – Padrões Wi-Fi.	20
Tabela 2 – Exemplo de cálculo de distância Euclidiana com KNN e $\hat{K}=1$	29
Tabela 3 – Cálculo de distância Euclidiana com WKNN e $\hat{K}=2$	31
Tabela 4 – Características dos trabalhos relacionados.	35
Tabela 5 – Erro Médio MAE x Valor \hat{K} – Coordenada X'	46
Tabela 6 – Erro Médio MAE x Valor \hat{K} – Coordenada Y'	46
Tabela 7 – Erro Médio MAE x Valor \hat{K} – MAE Resultante.	47
Tabela 8 – Erro Médio MAE x N° de amostras RSS.	48
Tabela 9 – Erro RMSE x Valor \hat{K}	49
Tabela 10 – Quantidade de APs visíveis em cada TP.	50
Tabela 11 – Número de APs visíveis x Erro médio MAE.	51

Lista de abreviaturas e siglas

AP	<i>Access Point</i>
APC	<i>Affinity propagation clustering</i>
AGNES	<i>Agglomerative Nesting</i>
BSSID	<i>Basic Service Set Identification</i>
BWP	<i>Between-Within Proportion</i>
BLE	<i>Bluetooth Low Energy</i>
FP	<i>Fingerprinting</i>
GPS	<i>Global Positioning System</i>
HMI	<i>Human Machine Interface</i>
ILS	<i>Indoor Location System</i>
IFCM	<i>Improved Fuzzy C-means</i>
ISM	<i>Industrial Scientific and Medical</i>
IoT	<i>Internet of things</i>
LBS	<i>Location Based Services</i>
KNN	<i>K-Nearest Neighbor</i>
MAC	<i>Media Access Control</i>
NFC	<i>Near Field Communication</i>
OFD	<i>Omnidirectional Fingerprint Database</i>
PDR	<i>Pedestrian Dead Reckoning</i>
PCA	<i>Principle Component Analysis</i>
RF-ID	<i>Radio Frequency Identification</i>
RP	<i>Reference Point</i>
RSS	<i>Received Signal Strength</i>

RSSI	<i>Received Signal Strength Indicator</i>
RMSE	<i>Root Mean Squared Error</i>
TP	<i>Test Point</i>
UWB	<i>Ultra Wide Band</i>
WKNN	<i>Weighted K-Nearest Neighbor</i>
Wi-Fi	<i>Wireless Fidelity</i>
WLAN	<i>Wireless Local Area Network</i>

Sumário

1	INTRODUÇÃO	14
1.1	Motivação	16
1.2	Descrição do problema	17
1.3	Objetivos Gerais	17
1.4	Objetivos Específicos	17
1.5	Organização do Trabalho	17
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	O IEEE 802.11	19
2.2	Métodos de localização em ambientes internos	20
2.2.1	Método da trilateração	20
2.2.2	Método <i>fingerprinting</i>	21
2.3	Algoritmos baseados no método <i>fingerprinting</i>	26
2.3.1	Algoritmo <i>K-Nearest Neighbor</i> (KNN)	26
2.3.1.1	Distância <i>Manhattan</i>	26
2.3.1.2	Distância Euclidiana	27
2.3.2	Algoritmo <i>Weighted K-Nearest Neighbor</i> (WKNN)	30
3	TRABALHOS RELACIONADOS	32
4	MATERIAIS E METODOLOGIA	36
4.1	Arquitetura	36
4.1.1	Fase de treinamento (<i>off-line</i>)	37
4.1.2	Fase de validação e Fase de testes (<i>on-line</i>)	39
4.2	Equipamentos utilizados e setup	42
5	RESULTADOS	44
5.1	Erro médio MAE x Valor \hat{K}	46
5.2	Erro médio MAE x Número de amostras RSS	47
5.3	Erro RMSE versus valor \hat{K}	48
5.4	CDF versus erro médio MAE	49
5.5	Quantidade de APs visíveis em cada TP	50
5.6	Número de APs visíveis x erro médio MAE	51
6	CONCLUSÃO	53

REFERÊNCIAS 54

1 Introdução

Nos últimos anos houve uma crescente demanda por propostas de serviços baseados em localização (*Location Based Services – LBS*) (INSIGHT, 2012). Os LBSs podem disponibilizar localização em ambientes externos ou internos, estes últimos sendo chamados de *Indoor Location Systems - ILSs*. Em ambientes externos existem soluções bem difundidas e consolidadas que utilizam *Global Positioning System (GPS)* para disponibilizar localização. No entanto, os sistemas GPS não alcançam os ambientes internos eficientemente devido a obstáculos tais como paredes, tetos e objetos. Estes obstáculos causam atenuação dos sinais dos satélites, além do efeito de multipercurso, dificultando a recepção destes sinais por receptores posicionados em ambientes internos (LI et al., 2016b) (NIRJON et al., 2014) (MAKKI et al., 2015).

Um fato que motiva a pesquisa e desenvolvimento dos ILSs é que a grande maioria das pessoas gasta a maior parte do seu tempo em ambientes internos como: escolas, escritórios, shoppings, supermercados, hospitais, indústria e outros (BELLONI, 2010). Nestes locais, a existência dos ILSs é importante para promover serviços como a localização de pessoas (ex. escolas e escritórios), otimização de rotas para aquisição de produtos ou serviços (ex. shoppings e supermercados), localização de pacientes, médicos e equipamentos para o caso de hospitais ou o gerenciamento de inventário, navegação de robôs e a otimização do fluxo de produção para as aplicações na indústria. Como a tecnologia da Internet das Coisas (*Internet of things - IoT*) está se tornando cada vez mais difundida, aplicações para IoT necessitarão cada vez mais do desenvolvimento dos ILSs (MACAGNANO; DESTINO; ABREU, 2014)(CARVALHO et al., 2019).

Contudo, para a proposta de ILSs que sejam amplamente utilizados no dia a dia das pessoas, existem definições e questões em aberto que precisam de direcionamento. Algumas destas questões são: é possível o desenvolvimento de ILSs que aproveitem a infra-estrutura de rede sem fio existente nos ambientes, que disponibilizem interface homem máquina (*Human Machine Interface - HMI*) utilizando tecnologias que permitam a fácil interação do usuário com o sistema, e que ao mesmo tempo possuam acurácia de localização adequada?

Em pesquisas existentes na literatura relacionadas com infra-estrutura/tecnologia de rede sem fio utilizada, existem propostas que endereçam soluções para ILSs baseados em tecnologias de *Radio Frequency Identification (RF-ID)* (XI; JIANG; LIAO, 2014), *InfraRed (IR)* (WU; ZHU; ZHANG, 2017), *Ultra Wide Band (UWB)* (YANG; ZHAO; CHEN, 2018) (ALNAFESSAH et al., 2016) e *Bluetooth* (DOUSH et al., 2017) ou *Wi-Fi* (LI et al., 2016a)(JUN et al., 2017)(FERREIRA, 2019)(FERREIRA; SOUZA; CARVALHO, 2020). Comenta-se que apesar da literatura apontar que *Bluetooth* e *UWB* são exemplos

de tecnologias com as quais pode-se desenvolver ILSs de maior acurácia de localização (ZHANG et al., 2006), cita-se que a exceção do Wi-Fi, todas as tecnologias citadas anteriormente não são amplamente encontradas nos ambientes internos do dia a dia das pessoas, não aproveitando assim, a infra-estrutura de rede sem fio existente.

No que tange a facilidades de HMI, os smartphones têm sido amplamente utilizados em ILSs pela facilidade na programação de interfaces de usuários, e por disponibilizarem tecnologias Wi-Fi e *Bluetooth* que facilitam a comunicação entre o smartphone e dispositivos existentes no ambiente (ex. pontos de acesso Wi-Fi) (XIAO et al., 2016)(??)(LANE et al., 2010)(ZAFARI; GKELIAS; LEUNG, 2019). Contudo, a maioria das soluções na literatura que implementam ILSs utilizando smartphones, empregam técnicas que demandam baixo custo computacional (exemplo: técnicas para redução da dimensão de amostras de RSS (*Received Signal Strength*) coletadas) devido restrições de energia e memória do dispositivo, reduzindo a acurácia dos ILSs propostos (SUBBU et al., 2014)(ARYA; GODLEWSKI; MELLÉ, 2010)(FANG; LIN; LIN, 2008)(JIA et al., 2017). De forma a mitigar os problemas de acurácia causados por restrições de energia e memória, disponibilizando IHM (Interação Homem Computador) de fácil interação, desenvolvemos um sistema de localização com alta capacidade computacional, implementando um modelo cliente-servidor para realização de tarefas mais simples no dispositivo smartphone (cliente) e tarefas mais complexas realizadas no computador servidor. Como exemplo, citamos que o dispositivo smartphone coleta e faz pré-processamento de amostras de RSS dos *beacons* transmitidos periodicamente pelos pontos de acesso (*Access Points* - APs), enquanto um computador servidor executa os algoritmos de localização e armazena dados em BD (Banco de Dados) com informações do mapa de rádio do ambiente (ARYA; GODLEWSKI; MELLÉ, 2010) (YIU et al., 2016)(WANG; CONG, 2018)(XUE et al., 2018)(LI et al., 2019)(WANG et al., 2020)(BI et al., 2018).

A respeito da acurácia de localização, constatamos que utilizar a média aritmética para processar, nos dispositivos smartphones, os valores das amostras de RSS, pode reduzir erros de estimativa de localização causados por variações dos valores das amostras de RSS devido a fenômenos como multi-percursos e obstáculos móveis (KHALAJMEHRABADI; GATSIIS; AKOPIAN, 2016). Embora trabalhos recentes da literatura (WU et al., 2016)(XUE et al., 2019)(HUANG; MANH, 2016) têm tentado minimizar os efeitos desses fenômenos para melhorar a acurácia de localização implementando técnicas de filtragem, normalização, transformação lineares ou subdimensionamento no conjunto de amostras de RSS, eles não têm-se preocupado com a quantidade de amostras de RSS coletadas dos APs e densidade de RPs/ m^2 na fase *off-line* que podem influenciar na acurácia de localização na fase *on-line* (YIU et al., 2016).

O objetivo desta pesquisa de dissertação foi conceber um ILS utilizando assinaturas de sinais transmitidos pelos APs Wi-Fi pré-existentes no ambiente, devido a popularidade

da tecnologia, a fim de ir ao encontro da ubiquidade para o sistema de localização. Na proposta de ILS desta dissertação foi desenvolvido um aplicativo para dispositivo smartphone com sistema operacional Android que é um dos sistemas operacionais mais difundidos no mundo (LIMA, 2017) e com facilidade de operação homem-máquina (CORREA et al., 2017). A fim de melhorar a acurácia, o ILS realiza transformações nos valores das amostras de RSS, de forma a suavizar as variações dos sinais de rádio, permitindo reduzir erros de localização. Além disto, o ILS proposto emprega algoritmo de aprendizagem de máquina (*Weighted K-Nearest Neighbor* - WKNN), instalado em um computador servidor, com a capacidade de processamento adequada.

O ILS proposto alcançou erros de localização RMSE (*Root Mean Squared Error*) mínimos com valor de até 2,87 m, que é inferior, em valor numérico absoluto as pesquisas de (YIU et al., 2016)(WANG; CONG, 2018)(XUE et al., 2018) e comparável com a pesquisa de (CHIU, 2017). Sendo que (CHIU, 2017), além de utilizar APs Wi-Fi existentes no ambiente, também necessita posicionar no ambiente, *beacons* bluetooth, utilizados como âncora. Comenta-se que no ILS desenvolvido nesta dissertação foram utilizados somente APs Wi-Fi pré-existentes no ambiente para servirem de transmissores de sinais de referência para localizar um dispositivo smartphone.

O ILS proposto por (SALAMAH et al., 2016) possui erro de localização inferior a presente proposta (ex. 2,29 m), contudo, o ILS proposto em (SALAMAH et al., 2016) utiliza maior quantidade de RPs no ambiente (0,84 RPs/m² quando comparado com 0,1 RPs/m² da proposta desta dissertação).

O ILS desenvolvido e apresentado nesta dissertação utiliza um computador servidor equipado com uma aplicação e algoritmo com a função de estimar a localização do smartphone. Comenta-se que sistemas com computador servidor podem reduzir o tempo para promover a localização, indo em direção a proposta capaz de localizar uma pessoa ou objeto em tempo real.

Por fim, resume-se que esta dissertação contribui com avanços práticos através de um sistema de localização de pessoas ou objetos desenvolvidos com características de ubiquidade, facilidade de instalação e operação e que apresenta maior acurácia de localização do que alguns trabalhos da literatura.

1.1 Motivação

Nos últimos anos houve crescente demanda por propostas de ILSs trazendo desafios a serem vencidos. Um fato que motiva a pesquisa e desenvolvimento dos ILSs é que grande maioria das pessoas gasta a maior parte do seu tempo em ambientes internos como: escolas, escritórios, shoppings, supermercados, hospitais, indústria e outros.

1.2 Descrição do problema

O desenvolvimento de ILSs continua sendo um desafio devido a questões em aberto, que ainda precisam direcionamento relacionados às limitações da infra-estrutura, dificuldades de operação, usabilidade da IHM e aumento da acurácia. Nesta dissertação, como problema de pesquisa, questiona-se como é possível desenvolver ILSs que aproveitem a infra-estrutura de rede sem fio existente nos ambientes, disponibilize interface homem máquina utilizando tecnologias que permitam a fácil interação do usuário com o sistema e que ao mesmo tempo possuam acurácia de localização adequada?

1.3 Objetivos Gerais

Esta pesquisa de mestrado tem por objetivo conceber, desenvolver e avaliar um ILS que determina a localização de um usuário, a partir de seu smartphone, utilizando assinaturas de sinais dos APs Wi-Fi pré-existentes no ambiente, uma base de dados criada na fase *off-line* (mapa de rádio) e algoritmo de aprendizagem supervisionada WKNN para determinar a localização. O ILS proposto apresenta erros de localização similares aos trabalhos (CHIU, 2017)(LI et al., 2019) e comparável a outros trabalhos com menores erros (SALAMAH et al., 2016)(WANG et al., 2020)(BI et al., 2018), porém que utilizam uma quantidade maior de RPs/ m^2 .

1.4 Objetivos Específicos

- Elaborar levantamento bibliográfico com trabalhos do estado da arte da literatura;
- Propor ILS com aplicativo para dispositivo smartphone com sistema operacional Android, que utilize assinaturas de sinais dos APs Wi-Fi pré-existentes no ambiente;
- Propor uma aplicação localizada em computador servidor, baseada em algoritmo de aprendizagem de máquina WKNN para estimar a localização;
- Realizar experimentos com o sistema ILS desenvolvido em um cenário real avaliando seu desempenho e comparando os resultados com outros trabalhos da literatura.

1.5 Organização do Trabalho

Esta dissertação está organizada da seguinte forma:

O Capítulo 2 apresenta uma revisão teórica sobre métodos de localização em ambiente internos, como o método da trilateração, e o método utilizando as assinaturas de sinais transmitidos pelos APs Wi-Fi. Neste capítulo, apresenta-se também uma revisão

teórica sobre algoritmos de aprendizagem de máquina empregados em ILSs, como o KNN (*K-Nearest Neighbor*) e WKNN;

O Capítulo 3 apresenta os trabalhos da literatura relacionados ao tema da presente dissertação de mestrado;

O Capítulo 4 apresenta a arquitetura, metodologia e ferramentas utilizadas no desenvolvimento do ILS proposto nesta dissertação;

O Capítulo 5 apresenta resultados e avaliações de desempenho do ILS proposto, através de experimentos executados em ambiente real;

O Capítulo 6 apresenta a conclusão e trabalhos futuros.

2 Fundamentação Teórica

Nesta seção foi feita uma revisão dos assuntos relacionados com a pesquisa e que foram essenciais para o desenvolvimento da proposta e metodologia.

2.1 O IEEE 802.11

A tecnologia Wi-Fi é baseada no padrão IEEE 802.11 e utiliza rádio frequência para transmitir dados, operando nas faixas de frequências de 2,4 GHz ou 5 GHz. O Wi-Fi tem se tornado a tecnologia mais utilizada em aplicações de redes residenciais, em áreas públicas e trabalho (STELLA; RUSSO; BEGUŠIĆ, 2012). Vários dispositivos podem se interconectar em uma rede Wi-Fi, através de APs que são dispositivos concentradores que possibilitam a comunicação em uma BSS (*Basic Service Set*). Os APs Wi-Fi possuem um endereço físico, também chamado de BSS-ID (*Basic Service Set Identification*), que os identifica de forma única em uma rede sem fio. Uma das formas de promover a localização de um dispositivo em uma rede sem fio, é utilizar a intensidade de sinal recebido (RSS) pelos dispositivos sem fio presentes na rede. Para isto, cada dispositivo sem fio que se deseja localizar deve coletar amostras de RSS de sinais de RF, em um local do espaço, correlacionando o local com as amostras de RSS provenientes dos vários APs existentes.

Os valores das amostras RSS dependem da distância entre transmissor e o receptor, sendo comumente expresso na unidade dBW ou dBm. No entanto, as amostras de RSS podem sofrer variações devido a existência de obstáculos móveis no meio sem fio, tais como pessoas ou objetos, que atuam mudando a direção de propagação das ondas de rádio. Paredes e outras estruturas também podem causar fenômenos de propagação, tal como o multipercurso, que ocasiona flutuações nos valores de amostras de RSS medidos nos dispositivos sem fio. Outra fonte de variações no RSS observado por dispositivos Wi-Fi, é a interferência de outros dispositivos Wi-Fi ou não Wi-Fi (ex. Bluetooth) que podem estar utilizando, simultaneamente, a mesma banda de comunicação ISM (*Industrial Scientific and Medical*). Na Tabela 1 são mostrados vários padrões IEEE 802.11, mostrando nome correspondente, ano de lançamento, taxa máxima e banda utilizada. Nesta dissertação foram utilizados dispositivos com padrão IEEE 802.11g, que transmitem com taxa de até 54 Mbps e operam na banda de 2,4 GHz.

Tabela 1 – Padrões Wi-Fi.

Padrão	Nome	Lançamento	Máx. Velocidade	Banda
802.1b	<i>Wi-Fi</i> 1 (não oficial)	1999	11 Mbps	2,4 GHz
802.1a	<i>Wi-Fi</i> 2 (não oficial)	1999	54 Mbps	5 GHz
802.1g	<i>Wi-Fi</i> 3 (não oficial)	2003	54 Mbps	2,4 GHz
802.1n	<i>Wi-Fi</i> 4	2009	600 Mbps	2,4 GHz e 5 GHz
802.1ac	<i>Wi-Fi</i> 5	2013	3,46 Gbps	5 GHz
802.1ax	<i>Wi-Fi</i> 6	2018/2019	10,53 Gbps	2,4 GHz e 5 GHz

Fonte(s): Wi-Fi.org

2.2 Métodos de localização em ambientes internos

2.2.1 Método da trilateração

Alguns sistemas de localização em redes sem fio (PENEDA; AZENHA; CARVALHO, 2009) utilizam modelos matemáticos para estimar a distância entre dispositivos transmissores (ex. torre da Figura 1) e receptor (ex. dispositivo móvel). Após determinação da distância, podem ser utilizados métodos, como a trilateração (Figura 1), que empregam sistemas de equações para determinar a posição do dispositivo móvel, dada pela intersecção de no mínimo três circunferências, cujos raios são dadas pela distância estimada entre os dispositivos transmissor e o dispositivo receptor (PENEDA; AZENHA; CARVALHO, 2009).

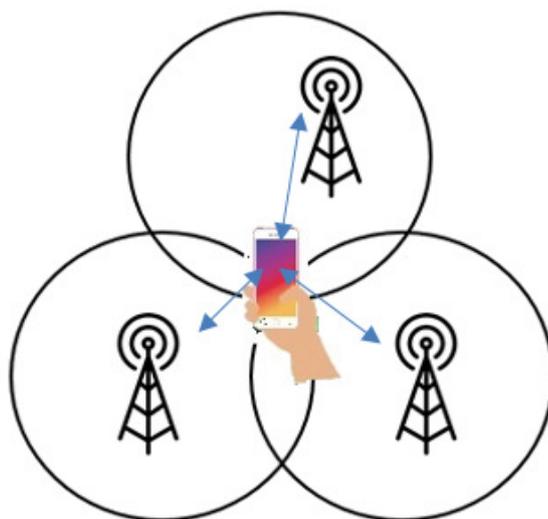


Figura 1 – Método de trilateração.

Outros sistemas de localização utilizam características de propagação das ondas de rádio (BAHL; PADMANABHAN, 2000), que se comportam de maneira distinta para um dado ambiente e local neste ambiente. Estes métodos são chamados de *fingerprinting*, termo em inglês muito comum na literatura que significa impressão digital.

2.2.2 Método *fingerprinting*

O *fingerprinting* (*FP*) representado na Figura 2 é uma das técnicas mais comuns e amplamente usadas (STELLA; RUSSO; BEGUŠIĆ, 2012) na localização em ambientes internos e que envolve duas fases: uma fase *off-line* (3) e uma fase *on-line* (4). Na fase *off-line* (3), um dispositivo móvel, que se deseja localizar, realiza medições de RSS dos *beacons* transmitidos pelos pontos de acessos (*Access Points* - AP) (1) existentes no ambiente. Um conjunto de $|N|$ amostras de RSS (variável $RSS_{n,m}^r$ com $n=1,\dots,|N|$, $m=1,\dots,|M|$ e $r=1,\dots,|R|$) coletadas de $|M|$ APs em $|R|$ pontos de referências (*Reference Points* - RP) (2) em um ambiente, é chamado de mapa de rádio (*Radio Map*) (5) do ambiente. Na Figura 2, mostra-se exemplo onde foi coletado $|N|=1$ amostra, por isso a simplificação da variável RSS (ex. para o RP_1 , $RSS_1^1, \dots, RSS_{|M|}^1$). O mapa de rádio é utilizado para criar uma base de dados, ou *dataset*, que correlaciona as amostras de RSS coletadas em cada RP e provenientes de cada AP, com o local de cada RP no ambiente. Na fase *on-line* (4), um ILS realiza coletas de amostras de RSS em um ponto de teste (*Test Point* - TP) (7) de coordenadas X', Y' na Figura 2, utiliza um algoritmo (6) para buscar em seu dataset o(s) RP(s) com valores de RSS (fase *off-line*), mais próximos dos valores de RSS da fase *on-line*, fazendo uma correspondência entre TP (localização desconhecida) e RP (localização conhecida) (TAHERI; SINGH; EMMANUEL, 2004).

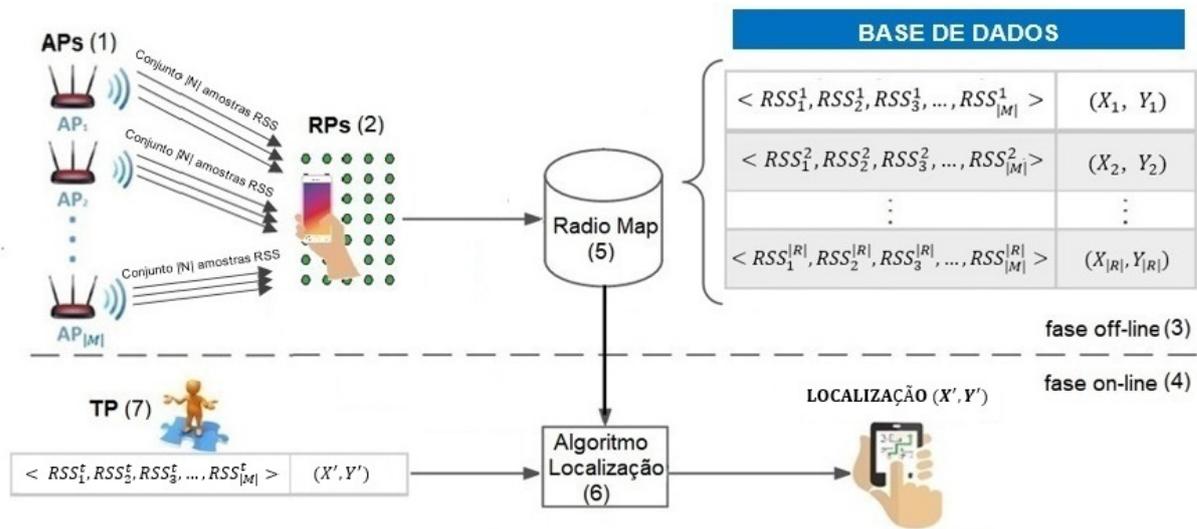


Figura 2 – Método *fingerprinting*.

Alguns termos utilizados nesta dissertação e apresentados na Figura 2 podem ser definidos conforme a seguir:

- **Amostras RSS:** são medições, em um dispositivo smartphone, dos valores de RSS dos quadros de *beacons* transmitidos pelos APs existentes no ambiente de

experimentação. Alguns trabalhos da literatura como Simon Yiu et al. (YIU et al., 2016) usam o termo RSSI (*Received Signal Strength Indicator*) ao invés de RSS. Simon Yiu et al. descreve o valor RSSI como um número de 8 bits de 0 a 255, sendo que 255 representa 0 dBm e 160 representa -95 dBm;

- **RP**s: representam pontos de referências definidos para o ambiente onde desejamos criar o mapa de rádio;
- **TP**: representa um ponto de teste onde desejamos determinar as coordenadas X', Y' de localização;
- **Conjunto de amostras RSS**: corresponde a várias coletas (ex: 2,3,4,...,10) de valores de RSS realizados em um mesmo RP ou TP;
- **Índice N**: representa o conjunto de amostras RSS proveniente de um mesmo AP e, $|N|$ é a quantidade de elementos do conjunto N;
- **Índice |M|**: representa a quantidade de APs existentes no ambiente e visíveis a partir de um dado RP;
- **Índice |R|**: representa a quantidade de RPs definidos para o ambiente onde desejamos criar o mapa de rádio;
- **Instâncias de *fingerprinting* (FP_i)**: formada por um conjunto de $|N| \times |M|$ amostras de RSS coletadas em um RP.

Representando os componentes do modelo de *fingerprinting* na forma de conjuntos, podemos considerar um ambiente contendo um conjunto M de $|M|$ APs conforme Figura 3 e Equação 2.1.

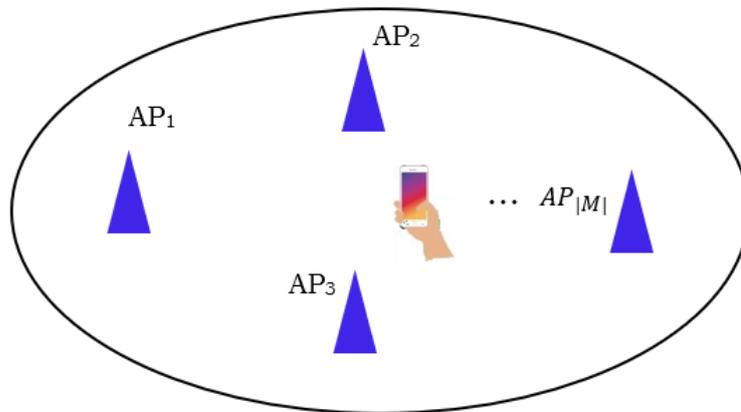


Figura 3 – Conjunto M de APs, $m = 1, \dots, |M|$.

$$m = 1, \dots, |M| \quad (2.1)$$

Cada AP do conjunto M transmite *beacons* cujos valores de amostras de RSS podem ser medidos por um dispositivo sem fio (ex. smartphone) com resultados representados pela variável $RSS_{n,m}^r$ (com $n = 1, \dots, |N|$, $m = 1, \dots, |M|$ e $r = 1, \dots, |R|$), onde na Figura 4 de exemplo, utilizou-se $n = 1$ amostra, $r = 1$ RP, sendo coletada 1 amostra de RSS de cada AP, em 1 RP do ambiente, simplificando a variável que representa as amostras de RSS para RSS_m , $m = 1, \dots, |M|$.

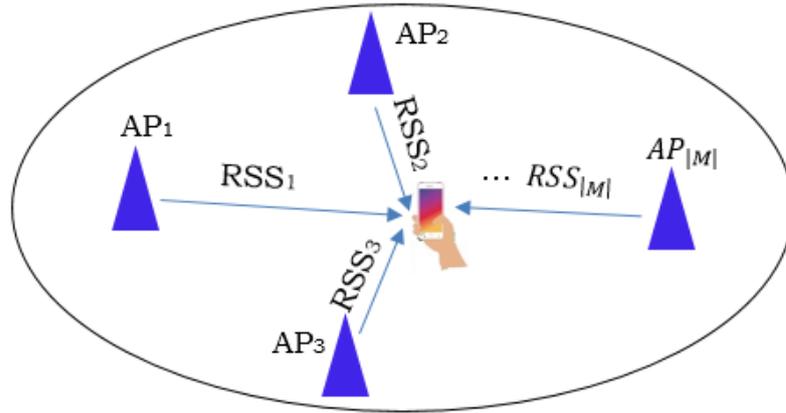


Figura 4 – Coleta de 1 amostra de RSS de cada AP do ambiente, em 1 RP do ambiente.

Para criar um mapa de rádio, inicia-se a fase de treinamento definindo um conjunto R de RPs com índice de $r = 1 \dots |R|$, onde $|R|$ é a quantidade total de RPs, representados pela Figura 5 e a Equação 2.2.

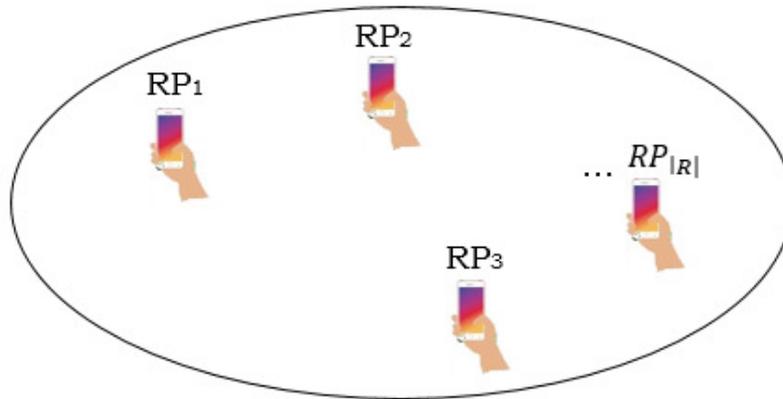


Figura 5 – Conjunto R de RPs ($RP_1, \dots, RP_{|R|}$), com $r = 1, \dots, |R|$.

$$RP_1, \dots, RP_{|R|}, \quad \text{com } r = 1, \dots, |R| \quad (2.2)$$

Para cada um dos $|M|$ APs existentes no ambiente e visíveis a partir de cada RP, o smartphone coleta $|N|$ amostras de RSS (variável $RSS_{n,m}^r$ com $n = 1, \dots, |N|$, $m = 1, \dots, |M|$ e

$r = 1, \dots, |R|$). Após realização das coletas de amostras de RSS para todos os RPs definidos no conjunto R e armazenadas em uma base de dados, teremos então um mapa de rádio formado por um conjunto $FP_{i,r}$ com índice $i = 1, \dots, |FP|$ para cada RP_r , associada à respectiva coordenada espacial (X_r, Y_r) conforme Figura 6. Por simplificação, na Figura 6, é representada a coleta de apenas 1 instância de *fingerprinting* por RP, sendo assim a variável $FP_{i,r}$ foi simplificada como sendo FP_i .

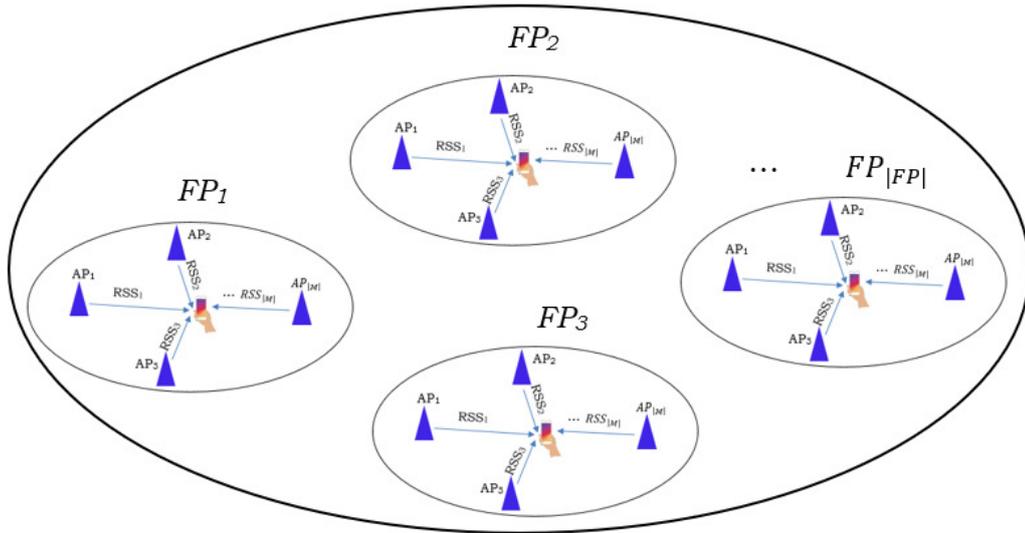


Figura 6 – Conjunto de FPs, formado por elementos $FP_{i,r}$ onde $i = 1$ instância de *fingerprinting* por RP_r .

Com isso, é possível iniciar a fase de teste para estimar a localização de um TP com coordenada desconhecida no ambiente, através de um algoritmo de localização apropriado e utilizando as medições de RSS armazenadas na base de dados conforme a Figura 7.

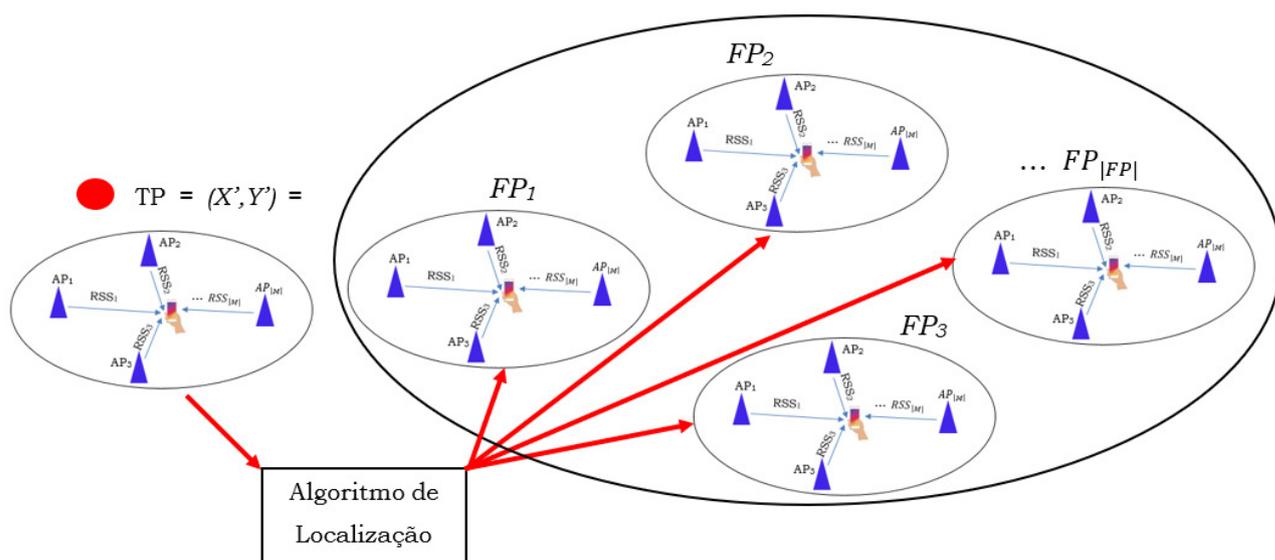


Figura 7 – Estimativa de localização de TP na fase *on-line*.

2.3 Algoritmos baseados no método *fingerprinting*

Nesta seção é feito uma breve revisão sobre os algoritmos de aprendizagem de máquina baseados no método de *fingerprinting* para determinar ou estimar a localização de objetos ou usuários em ambientes internos. Dois algoritmos baseados em *fingerprinting* muito utilizados em problemas de classificação e regressão são, respectivamente, os métodos de aprendizagem de máquina KNN e WKNN (MARZOUQ et al., 2019).

2.3.1 Algoritmo *K-Nearest Neighbor* (KNN)

É um método que consiste em utilizar, para uma instância de teste, a coleta de amostras de RSS dos quadros de *beacons* transmitidos por APs do ambiente para encontrar o(s) \hat{K} vizinho(s) mais próximo(s) do conjunto que foi treinado na fase *off-line*, e assim, determinar a localização aproximada do objeto. Para encontrar o(s) \hat{K} vizinho(s) mais próximo(s) primeiramente é necessário calcular a distância entre os atributos da instância de teste em relação aos atributos da base de dados de treinamento. Dois métodos muito utilizados para o cálculo desta distância são: a distância Euclidiana e a distância *Manhattan*.

2.3.1.1 Distância *Manhattan*

A distância *Manhattan* D_M é calculada pela subtração dos valores das amostras de RSS coletados dos $|M|$ APs, em cada RP, durante a fase de treinamento e, os valores das amostras de RSS coletadas dos $|M|$ APs, no TP com posição desconhecida.

$$D_M = \sum_{m=1}^{|M|} \sum_{n=1}^{|N|} |RSS_{n,m}^r - RSS_{n,m}^t| \quad (2.3)$$

Na Equação 2.3, $RSS_{n,m}^r$ representa o valor das $|N|$ amostra(s) de RSS coletadas dos $|M|$ APs, nos RPs na fase de treinamento sendo comparados um RP_r , $r = 1, \dots, |R|$, por vez com o TP_t . Na mesma equação $RSS_{n,m}^t$, representa o valor das amostras de RSS coletadas dos $|M|$ APs no TP desconhecido da fase de testes. Os valores absolutos calculados das subtrações resultam nas distâncias relativas de valores de coletas de amostras de RSS. O RP com menor distância do TP, no espaço de sinais, determina o RP mais próximo da posição atual do objeto a ser localizado. Para ambientes com áreas grandes e que mudam dinamicamente é preciso dividir o cálculo da distância por $|M|$ conforme Equação 2.4 (GANSEMER; GROSSMANN; HAKOBYAN, 2010).

$$D_M = \frac{1}{|M|} \sum_{m=1}^{|M|} \sum_{n=1}^{|N|} |RSS_{n,m}^r - RSS_{n,m}^t| \quad (2.4)$$

2.3.1.2 Distância Euclidiana

A distância Euclidiana é calculada com a raiz do quadrado da subtração dos valores de coletas de amostras de RSS correspondentes dos $|M|$ APs nos RPs, armazenadas na base de dados após a fase *off-line*, em relação ao valor da coleta de amostras de RSS do TP desconhecido que desejamos determinar. A distância Euclidiana pode ser representada pela Equação 2.5, que foi utilizada nos experimentos desta dissertação.

$$D_E = \sqrt{\sum_{m=1}^{|M|} \sum_{n=1}^{|N|} (RSS_{n,m}^r - RSS_{n,m}^t)^2} \quad (2.5)$$

Na Equação 2.6, $RSS_{n,m}^r$ representa o valor das $|N|$ amostra(s) de RSS coletadas dos $|M|$ APs, nos RPs na fase de treinamento e $RSS_{n,m}^t$ representa o valor das $|N|$ amostra(s) de RSS coletadas dos $|M|$ APs no TP desconhecido da fase de testes. A variável D_E , representa o RP da fase de treinamento com menor distância, no espaço de sinais, do TP da fase de testes. Assim como aplicado no método *Manhatan*, em ambientes com áreas grandes e que mudam dinamicamente é preciso dividir o cálculo da distância por $|M|$ (GANSEMER; GROSSMANN; HAKOBYAN, 2010).

$$D_E = \frac{1}{|M|} \sqrt{\sum_{m=1}^{|M|} \sum_{n=1}^{|N|} (RSS_{n,m}^r - RSS_{n,m}^t)^2} \quad (2.6)$$

Como exemplo de aplicação do método *fingerprinting* combinado com o algoritmo KNN, tomemos o cenário hipotético da Figura 8 composto por $|M|=4$ APs detectados no ambiente, e no qual foram definidos $|R|=8$ RPs, associados as suas respectivas coordenadas X_r, Y_r e, onde desejamos determinar as coordenadas X', Y' do TP.

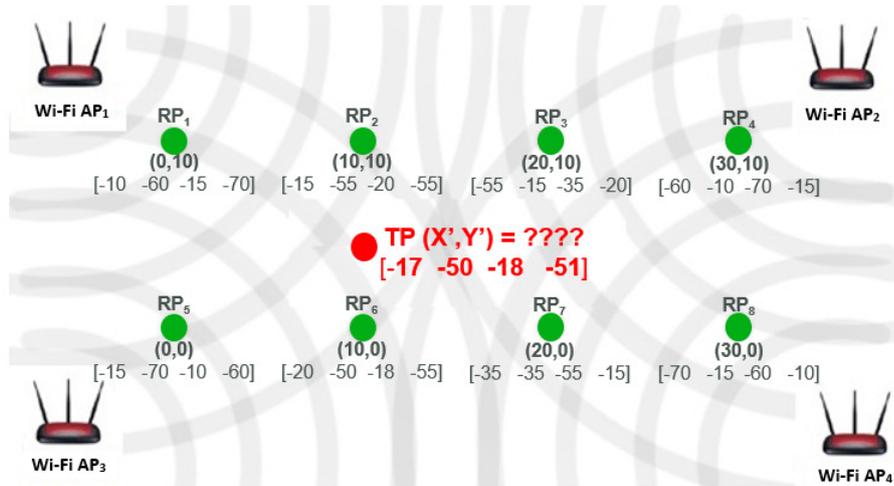


Figura 8 – Exemplo de método *fingerprinting* utilizando algoritmo KNN.

Após a fase de treinamento, tomando $|N|= 1$, os valores coletados de amostras de RSS de todos os RPs formam o mapa de rádio do ambiente conforme apresentado na Figura 9.

```

*****
          [AP1   AP2   AP3   AP4] [Xr   Yr]
*****
RP1 --> [-10   -60   -15   -70] [0    10]
RP2 --> [-15   -55   -20   -55] [10   10]
RP3 --> [-55   -15   -35   -20] [20   10]
RP4 --> [-60   -10   -70   -15] [30   10]
RP5 --> [-15   -70   -10   -60] [0    0]
RP6 --> [-20   -50   -18   -55] [10   0]
RP7 --> [-35   -35   -55   -15] [20   0]
RP8 --> [-70   -15   -60   -10] [30   0]
*****
    
```

Figura 9 – Mapa de rádio de exemplo usando método *fingerprinting*.

Utiliza-se a Equação 2.7 com $|R|= 8$ RPs, $|M|= 4$ APs e $|N|= 1$ amostra por AP, adaptada da Equação 2.6, para calcular a distância euclidiana $D(RSS_{n,m}^r, RSS_{n,m}^t)$ entre valores de amostras de RSS dos atributos dos RPs de treinamento ($RSS_{n,m}^r$) e os valores de amostras de RSS dos atributos correspondentes de teste TP ($RSS_{n,m}^t$). Na Tabela 2 podemos visualizar os resultados da distância Euclidiana calculada nos 8 RPs de treinamento.

$$D_{(RSS_{n,m}^r, RSS_{n,m}^t)} = \sqrt{\sum_{m=1}^{|M|} \sum_{n=1}^{|N|} (RSS_{n,m}^r - RSS_{n,m}^t)^2} \quad (2.7)$$

Tabela 2 – Exemplo de cálculo de distância Euclidiana com KNN e $\hat{K}=1$.

Distância entre instâncias de <i>fingerprinting</i> RSS: $D_{(RSS_{n,m}^r, RSS_{n,m}^t)}$ *	Distância calculada (dB)
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^1 - RSS_{n,m}^t)$	22,78
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^2 - RSS_{n,m}^t)$	7,00
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^3 - RSS_{n,m}^t)$	62,60
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^4 - RSS_{n,m}^t)$	86,31
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^5 - RSS_{n,m}^t)$	23,43
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^6 - RSS_{n,m}^t)$	5,00
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^7 - RSS_{n,m}^t)$	56,69
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^8 - RSS_{n,m}^t)$	86,48

* Na coluna 1 da Tabela 2, tem-se para cada célula, as variáveis, n e m variando de $n = 1, \dots, |N|$ e $m = 1, \dots, |M|$ a fim de determinar as distâncias calculadas na coluna 2 da Tabela 2.

Para $\hat{K}=1$, é necessário encontrar o RP com a menor distância, ou seja, o vizinho mais próximo para determinar a localização da instância de *fingerprinting* do TP. Neste caso, a distância entre valores de amostras de RSS dos atributos ($RSS_{n,m}^6, RSS_{n,m}^t$) possui o menor valor, sendo igual a 5,00 dB (marcação em azul na Tabela 2). Desta forma, a coordenada de localização do TP (X', Y') será igual à coordenada de $RP_6 = (10,0)$ conforme visualizado na Figura 10.

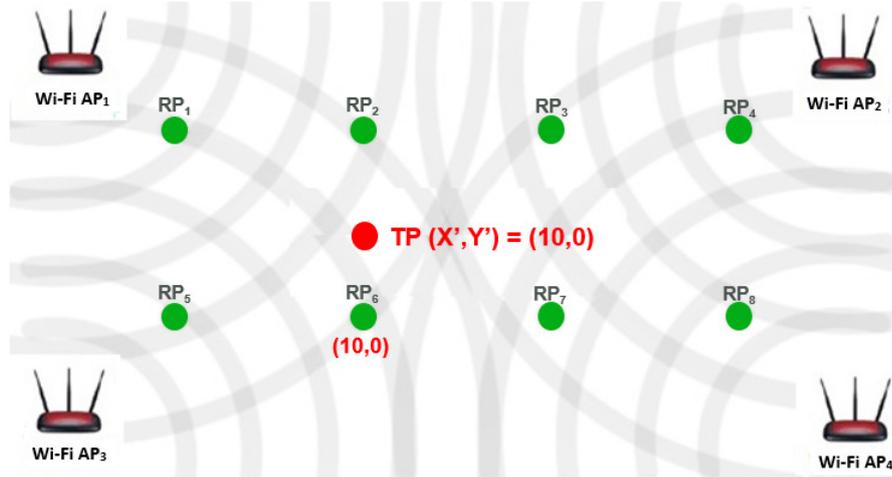


Figura 10 – Determinação de coordenadas usando método *fingerprinting* e KNN.

2.3.2 Algoritmo Weighted K -Nearest Neighbor (WKNN)

Neste algoritmo, a localização desconhecida de ponto TP é estimada através da atribuição de pesos proporcionais à distância entre os atributos da instância de teste e os atributos do(s) \hat{K} RPs mais próximo(s), obtendo assim coordenadas de posição de forma mais acurada (SHIN et al., 2012)(CARVALHO; REZENDE; CARVALHO, 2020). Este algoritmo é representado pelas Equações 2.8 e 2.9.

$$X' = \frac{\sum_{k=1}^{\hat{K}} \left(\frac{x_k}{dist_k} \right)}{\sum_{k=1}^{\hat{K}} \left(\frac{1}{dist_k} \right)} \quad (2.8)$$

$$Y' = \frac{\sum_{k=1}^{\hat{K}} \left(\frac{y_k}{dist_k} \right)}{\sum_{k=1}^{\hat{K}} \left(\frac{1}{dist_k} \right)} \quad (2.9)$$

Onde $dist_k$ representa a distância, em espaços de sinais, do(s) \hat{K} RPs mais próximos, com relação as instâncias i , de amostras de RSS dos RPs e do TP. Enquanto x_k e y_k correspondem as coordenadas de localização X_r, Y_r dos \hat{K} RPs mais próximos. O \hat{K} representa o número de vizinhos mais próximos (TEUBER; EISSFELLER, 2006).

Usando o mesmo cenário da Figura 5. apresenta-se o funcionamento do algoritmo WKNN para determinar as coordenadas de localização de um TP para um $\hat{K}=2$. Inicialmente, calcula-se a distância Euclidiana dos atributos (valores de RSS) entre cada um dos RPs da fase de treinamento e o TP da fase de teste. Em seguida, realiza-se a seleção dos vizinhos mais próximos. Neste caso a distância, em espaço de sinais, entre valores de amostras de RSS dos atributos ($RSS_{n,m}^2, RSS_{n,m}^t$) com valor = 7,00 dB e valores de amostras de RSS dos atributos ($RSS_{n,m}^6, RSS_{n,m}^t$) com valor = 5,00 dB apresentaram os menores valores conforme Tabela 3 (marcação em azul na Tabela 3).

Com os valores das duas menores distâncias obtidos é possível determinar as coordenadas de localização do TP conforme especificado nas equações 2.8 e 2.9.

Substituindo os valores de \hat{K} , $dist_k$ e x_k na Equação 2.8 (com $\hat{K} = 2$, $dist_1 = 7$, $dist_2 = 5$, $x_1 = 10$, $x_2 = 10$), obtêm-se a coordenada X' do TP com valor igual a :

$$X' = \frac{\sum_{k=1}^2 \left(\frac{x_1}{dist_1} \right) + \left(\frac{x_2}{dist_2} \right)}{\sum_{k=1}^2 \left(\frac{1}{dist_1} \right) + \left(\frac{1}{dist_2} \right)} = \frac{\left(\frac{10}{7} \right) + \left(\frac{10}{5} \right)}{\left(\frac{1}{7} \right) + \left(\frac{1}{5} \right)} = \frac{(50) + (70)}{(5) + (7)} = 10$$

Substituindo os valores de \hat{K} , $dist_k$ e y_k na Equação 2.9 (com $\hat{K} = 2$, $dist_1 = 7$, $dist_2 = 5$, $y_1 = 10$, $y_2 = 0$), obtêm-se a coordenada Y' do TP com valor igual a :

$$Y' = \frac{\sum_{k=1}^2 \left(\frac{y_1}{dist_1} \right) + \left(\frac{y_2}{dist_2} \right)}{\sum_{k=1}^2 \left(\frac{1}{dist_1} \right) + \left(\frac{1}{dist_2} \right)} = \frac{\left(\frac{10}{7} \right) + \left(\frac{0}{5} \right)}{\left(\frac{1}{7} \right) + \left(\frac{1}{5} \right)} = \frac{(50)}{(5) + (7)} = 4,16$$

Tabela 3 – Cálculo de distância Euclidiana com WKNN e $\hat{K}=2$.

Distância entre instâncias de <i>fingerprinting</i> RSS: $D_{(RSS_{n,m}^r, RSS_{n,m}^t)}$ *	Distância calculada (dB)
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^1 - RSS_{n,m}^t)$	22,78
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^2 - RSS_{n,m}^t)$	7,00
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^3 - RSS_{n,m}^t)$	62,60
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^4 - RSS_{n,m}^t)$	86,31
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^5 - RSS_{n,m}^t)$	23,43
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^6 - RSS_{n,m}^t)$	5,00
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^7 - RSS_{n,m}^t)$	56,69
$\sum_{m=1}^{ M } \sum_{n=1}^{ N } (RSS_{n,m}^8 - RSS_{n,m}^t)$	86,48

* Na coluna 1 da Tabela 3, tem-se para cada célula, as variáveis, n e m variando de $n = 1, \dots, |N|$ e $m = 1, \dots, |M|$ a fim de determinar as distâncias calculadas na coluna 2 da Tabela 3.

Portanto, as coordenadas de localização do TP(X', Y') = (10 , 4.16) são apresentadas na Figura 11.

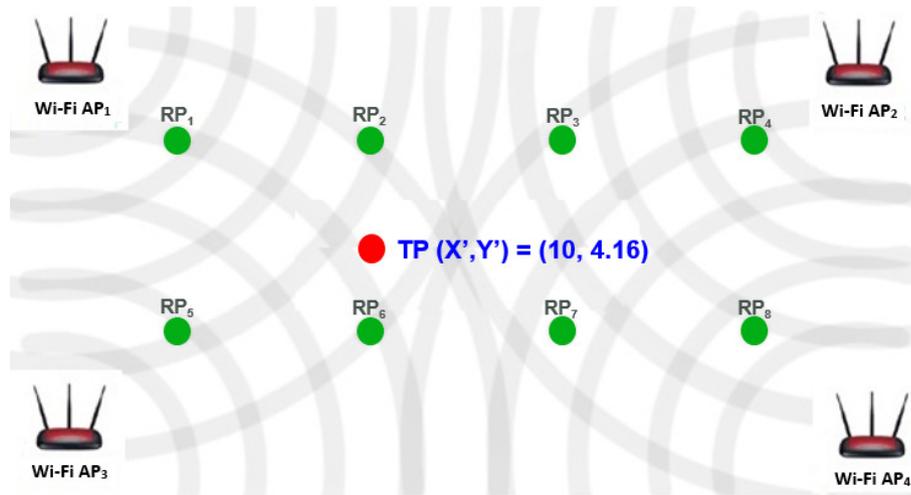


Figura 11 – Determinação de coordenadas usando método *fingerprinting* e WKNN.

3 Trabalhos Relacionados

Muitos trabalhos de pesquisa têm proposto soluções ILSs com o propósito de auxiliar na localização de um usuário ou objeto.

O trabalho de Salamah et al. (SALAMAH et al., 2016) propõe um ILS que utiliza dispositivos com tecnologia Wi-Fi, a técnica de análise de componentes principais (*Principle Component Analysis* - PCA) e algoritmo de aprendizagem de máquina KNN (*K-Nearest Neighbor*) para melhorar a acurácia de localização e reduzir a complexidade computacional. O ILS de (SALAMAH et al., 2016) utilizou smartphone para coletar amostras de RSS e um servidor para armazenar os valores das amostras de RSS coletadas. Para avaliar o desempenho do ILS, Salamah et al. (SALAMAH et al., 2016) conduziu testes com usuário parado (modo estático) e com usuário caminhando (modo dinâmico). Em modo estático, o ILS apresentou erro de localização RMSE de 2,29 m, sendo menor do que a proposta ILS desta dissertação que alcançou erro RMSE de localização de 2,87 m. Contudo o trabalho de (SALAMAH et al., 2016) utilizou maior quantidade de RPs no ambiente, sendo 0,84 RPs/m² quando comparado com 0,1 RPs/m² da proposta desta dissertação.

O trabalho desenvolvido por Chiu (CHIU, 2017) apresenta um ILS, com uma aplicação instalada em dispositivo smartphone, utilizando tecnologias Wi-Fi e *beacons* BLE (*Bluetooth Low Energy*) combinadas, a fim de determinar a localização com maior acurácia. O sistema emprega o método KNN e um filtro médio espacial para tratar as variações dos valores das amostras de RSS transmitidas pelos *beacons* BLE e, recebidas pelo dispositivo smartphone. O autor avaliou a acurácia da localização ao coletar amostras de RSS ao longo de diferentes intervalos de tempo, tendo concluído a coleta de amostras de RSS durante períodos de tempo de 4, 5 e 6s, obtendo valores de acurácia similares para todos os tempos de coleta. O ILS de (CHIU, 2017) alcançou erro RMSE de localização aproximado de 2,8 m, além de também utilizar os APs Wi-Fi disponíveis no ambiente. No entanto Chiu (CHIU, 2017) necessitou instalar no ambiente os sensores BLE para serem utilizados como âncoras.

Na pesquisa de Yiu et al. (YIU et al., 2016) foi realizado um estudo experimental em ILSs considerando métodos puramente baseados em coletas de amostras RSS (métodos tradicionais), e métodos usando modelo de regressão paramétricos (utilizam modelo de propagação de sinais para representar os valores das amostras RSS) e utilizando métodos não paramétricos (utiliza-se um processo Gaussiano para representar os valores das amostras RSS) (HÄHNEL; FOX, 2006). Utilizando estes métodos, foram feitas avaliações de acurácia de localização, apresentando erros RMSE de localização ao variar a densidade de RPs da fase de treinamento, a quantidade de APs, a quantidade de amostras RSS,

entre outros. Nos resultados foi verificado que a acurácia de localização melhora à medida que mais amostras de RSS são coletadas e, caso ocorra redução na densidade dos APs, pode resultar em redução na acurácia de localização. O ILS proposto por Yiu et al. (YIU et al., 2016) utilizou APs disponíveis no ambiente, um tablet com aplicação para coleta de amostras RSS e algoritmo de localização utilizando método KNN. O ILS do trabalho de Yiu et al. apresentou erro RMSE de localização com 4-KNN de 6,73 m, sendo este resultado superior ao obtido nesta dissertação.

O trabalho de Wang (WANG; CONG, 2018) propôs algoritmo que utiliza filtro de Kalman para reduzir o ruído dos valores das amostras de RSS, coletadas na fase *off-line*. Um grande número de experimentos foi realizado para verificar o menor valor de ruído obtido pelo sistema. Depois de filtrar os dados, foi gerado um banco de dados com os valores das amostras de RSS, que teve seus dados aplicados ao algoritmo WKNN para estimar a localização do objeto ou usuário. Apesar de usar APs do ambiente, um smartphone para realizar a coleta de amostras de RSS e filtro de Kalman, o sistema proposto mostrou-se com desempenho inferior em relação ao ILS desenvolvido nesta dissertação, apresentando erro RMSE de localização de 11,07 m.

O trabalho de Xue et al. (XUE et al., 2018) propôs um algoritmo combinado com KNN que melhora o processo de seleção dos K vizinhos mais próximos, com base na análise da proximidade geométrica entre RPs e o TP a ser localizado na fase *on-line*. Esta técnica evita erros de seleção de vizinhos mais próximos quando dois RPs estão a mesma distância do TP. A proposta utiliza os APs do próprio ambiente e um smartphone equipado com aplicativo para coletar amostras de RSS, além do algoritmo de seleção de RP por proximidade geométrica. Os resultados demonstraram que o algoritmo proposto alcançou melhor desempenho que os tradicionais KNN e WKNN, com erros RMSE de localização de 5,75 m e 4,31 m respectivamente, mas superior ao proposto nesta dissertação.

O trabalho de Li (LI et al., 2019) apresenta um método de localização em ambientes internos baseado na combinação dos algoritmos IFCM (*Improved Fuzzy C-means*) e o AGNES (*Agglomerative Nesting*) para realizar divisão por região, com o objetivo de otimizar a quantidade de APs, ao criar um subconjunto dos APs disponíveis. O IFCM apresenta um algoritmo de agrupamento baseado no método *K-means* e o índice BWP (*Between-Within Proportion*) para selecionar o centro de agrupamento inicial ideal e o número de *clusters*. Com este método foi possível reduzir a quantidade de cálculos realizada pelo algoritmo de localização, ao coletar amostras de RSS de testes, durante o processo de localização em tempo real. Os resultados experimentais desta solução mostraram que o método apresentou erro RMSE de localização de 2,8 m, alcançando valor comparável ao proposto nesta dissertação.

O trabalho de Boyuan Wang et al. (WANG et al., 2020) propôs um novo algoritmo WKNN, baseado na obtenção da distância euclidiana ponderada (*Weighted Euclidian*

distance – WED), com análise da diferença entre os valores de RSS dos RPs e TP a ser localizado e da distância calculada pelo modelo de propagação do sinal. Com a combinação da distância euclidiana ponderada e das informações das posições conhecidas dos RPs, é calculada a distância de posição aproximada (*Approximate Position Distance* - APD). O APD é combinado com o algoritmo WKNN para realizar a localização do objeto ou usuário. Os resultados experimentais mostraram que o menor erro RMSE de localização alcançado foi de 2,78 m. No entanto, o trabalho de Boyuan Wang et al. utiliza maior quantidade de RPs no ambiente (1,201 RPs/m²) quando comparado com 0,1 RPs/m² do trabalho proposto nesta dissertação.

No trabalho de Jingxue Bi et al. (BI et al., 2018) foi proposto um método de localização que reduz a influência do corpo humano na propagação de sinais Wi-Fi, adaptando a técnica WKNN com base de dados omnidirecional (*Omnidirectional Fingerprint Database* – OFD). A técnica OFD considera a posição e orientação do smartphone e sequência das medições de RSS realizadas em cada RP. Adicionalmente é introduzido um algoritmo de agrupamento de propagação de afinidade (*Affinity propagation clustering* – APC) na fase *off-line*, que consiste em um agrupamento com base na fusão da distância euclidiana de RSS e da distância da posição dos RPs. Os resultados experimentais mostraram que o menor erro de localização RMSE alcançado foi de 1,51 m apresentando erro inferior ao proposto nesta dissertação, que alcançou erro RMSE de localização de 2,87 m. Contudo, o trabalho de (Bi, 2018) utiliza maior quantidade de RPs no ambiente, 0,473 RPs/m² quando comparado com 0,1 RPs/m² do trabalho proposto nesta dissertação.

Na Tabela 4 são apresentadas as características principais (arquitetura, técnica de pré-processamento, método de localização, erro localização RMSE e outros) das propostas dos trabalhos da literatura e a proposta desenvolvida nesta dissertação. Todos os trabalhos utilizaram APs disponíveis no ambiente, além de smartphone/Tablet com aplicação para coleta de amostras de RSS. Com relação a arquitetura, Salamah et al. e Chiu optaram por uma arquitetura com servidor, enquanto os outros autores da Tabela 4 definiram arquitetura com armazenamento de banco de dados com valores de RSS e execução do algoritmo de localização no smartphone.

Esta dissertação de mestrado apresenta um ILS com alta capacidade computacional para determinar a localização de um smartphone utilizando assinaturas dos sinais recebidos de APs Wi-Fi existentes no ambiente, uma base de dados criada na fase *off-line* e um algoritmo de localização WKNN. O ILS proposto nesta dissertação alcançou erros RMSE de localização mínimos com valor de até 2,87 m, que é inferior, em valor numérico absoluto as pesquisas de (YIU et al., 2016)(WANG; CONG, 2018)(XUE et al., 2018) e comparável com a pesquisa de (CHIU, 2017). Sendo que (CHIU, 2017), além de utilizar APs Wi-Fi existentes no ambiente, também necessita posicionar no ambiente, *beacons* bluetooth, utilizados como âncora.

Tabela 4 – Características dos trabalhos relacionados.

Solução	Arquit	Tecnol	PPRO	MLOC	Área (m ²)*	Qtde APs	Qtde RPs	Dens RP/Área	Quant amost RSS	Erro RMSE (m)**
Salamah, 2016	AP + Smart phone + Servidor	FP Wi-Fi	PCA	KNN	53	6	45	0,840	Não informado	2,29 m
Chiu, 2017	AP + Smart phone + Servidor	FP(Wi-Fi + BLE)	-	KNN	441	11	22	0,050	24.200	2,8 m
Simon Yiu, 2016	AP + Tablet	FP Wi-Fi	-	KNN	2.500	174	235	0,094	204.450	6,73 m
Wang, 2018	AP + Smart phone	FP Wi-Fi	KF	WKNN	108.703	520	933	0,008	19.937	11,07 m
Xue, 2018	AP + Smart phone	FP Wi-Fi	GPA	KNN	2.756	6	160	0,058	Não informado	4,31 m
Li, 2019	AP + Smart phone	FP Wi-Fi	K-means + IFCM	WKNN	860	65	82	0,095	Não informado	2,8 m
Wang, Boyuan 2020	AP + Smart phone	FP Wi-Fi	APD	WKNN	164	10	197	1,201	Não informado	2,78 m
Jingxue Bi, 2018	AP + Smart phone	FP Wi-Fi	IOFPD	AWKNN	3.200	8	1.516	0,473	Não informado	1,51 m
ILS Proposto, 2021	AP + Smart phone + Servidor	FP Wi-Fi	FRA	WKNN	804	15	85	0,105	13.440	2,87 m

Arquit: Arquitetura implementada no ILS; Tecnol: Tecnologias utilizadas no ILS; PPRO: Técnica de pré-processamento; MLOC: Método de localização; Dens RP/Área: Densidade de RPs por Área; Quant amost RSS: Quantidade de amostras de RSS coletadas; KNN: k-nearest neighbor; WKNN: Weighted k-nearest neighbor; PCA: Principle Component Analysis; AP: Access Points; KF: Kalman filter; FP: Fingerprinting; BLE: Bluetooth Low Energy; IFCM: Improved Fuzzy C-means; APD: Approximate Position Distance; IOFPD: Improved Omnidirectional Fingerprint Database; GPA: Geometric Proximity Analysis; FRA: Funções de Resumo de Amostras RSS; * área total do ambiente fase off-line/on-line; **RMSE: Root Mean Square Error

4 Materiais e Metodologia

Nesta seção são apresentados os materiais e a metodologia usados no desenvolvimento do ILS desta dissertação. Inicialmente, apresentamos cada bloco da arquitetura (Figura 12) do ILS desenvolvido, bem como os procedimentos realizados nas fases de treinamento e validação/testes de localização. Para estas fases, são apresentadas as equações e ferramentas utilizadas.

4.1 Arquitetura

A arquitetura da Figura 12 é composta de APs existentes no ambiente, um dispositivo “smartphone” equipado com uma aplicação Android e, por fim um computador “servidor” do sistema.

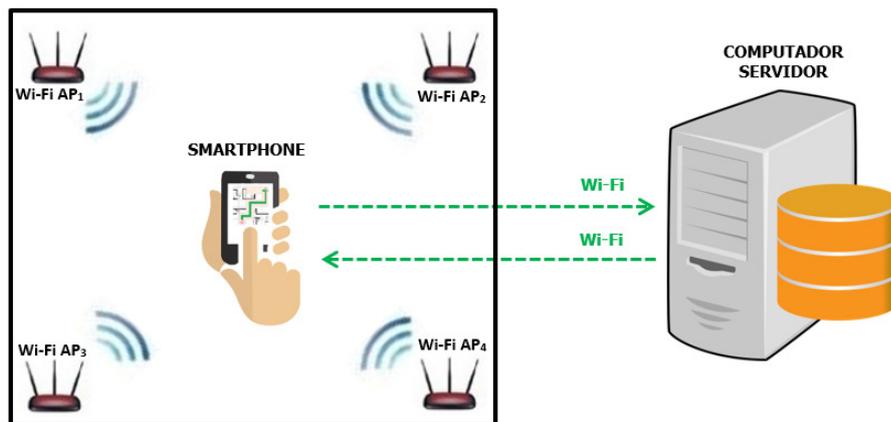


Figura 12 – Arquitetura geral.

Os APs IEEE 802.11, comumente utilizados em ambientes residenciais e comerciais, transmitem periodicamente seus SSIDs (*Service Set Identifiers*), utilizando quadros de *beacon*. Em nossa proposta, medimos, em um dispositivo smartphone, o RSS dos quadros de *beacons* transmitidos por um ou mais APs de referência existentes no ambiente de experimentação. O smartphone mede o RSS dos quadros de *beacon* transmitidos pelos APs próximos através de uma aplicação que desenvolvemos, com o objetivo de montar um mapa de rádio do ambiente, que corresponde ao conjunto de registros dos valores RSS medidos em determinados locais. Um computador “servidor” foi equipado com uma aplicação e algoritmo que desenvolvemos, sendo que o algoritmo tem a função de estimar a localização do smartphone utilizando, para isto, o algoritmo WKNN.

4.1.1 Fase de treinamento (*off-line*)

Na Figura 13, visualiza-se o fluxograma que mostra o funcionamento das aplicações residentes no smartphone (cor azul) e no servidor (cor laranja) do ILS proposto. Conforme figura, o processo de localização envolve duas fases principais: treinamento e testes. Na fase de treinamento, ou *off-line*, são executados os passos conforme fluxograma: (1) No bloco 1, “Cadastro de RP” o usuário se posiciona em cada local de medição, chamado de ponto de referência (RP) e, utiliza o aplicativo Android desenvolvido para cadastrar o RP_r , onde $|R|$ é o total de RPs na Equação 4.1. No cadastro de RPs, utilizou-se mapa, previamente desenhado do ambiente de experimentação, para graficamente posicionar o RP no mapa. No cadastro do RP, o aplicativo identifica os APs que estão no raio de comunicação do smartphone. O aplicativo utiliza os SSIDs e endereços físicos (*MAC- Media Access Control*) ou BSS-ID para identificar cada AP de maneira única. Ao marcar o RP no mapa do ambiente, o aplicativo associa ao RP, os valores das coordenadas (X_r, Y_r) do RP_r sendo cadastrado; (2) No bloco 2, “Coleta amostras de RSS”, os APs são utilizados no ILS como equipamentos transmissores de *beacons*, cujos valores de amostras de RSS são medidos no RP definido no bloco (1).

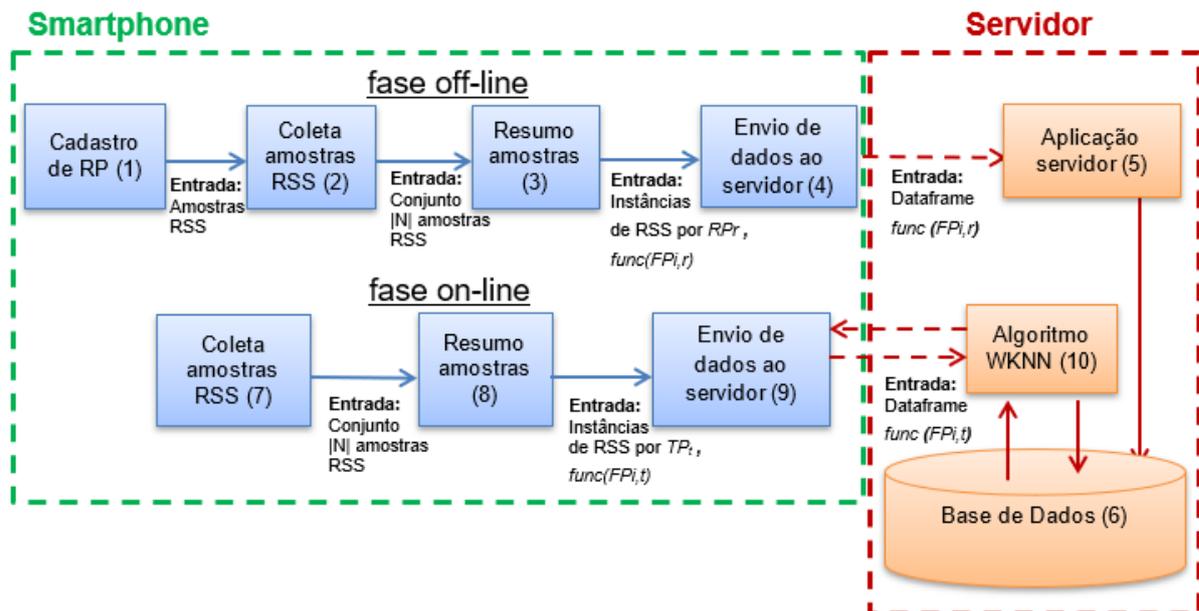


Figura 13 – Fluxograma da aplicação do smartphone e do servidor.

No processo de coleta das amostras de RSS, a aplicação do smartphone coletou $|N|$ amostras de RSS representadas pelo conjunto S_m^r na Equação 4.2, para cada um dos $|M|$ APs existentes no ambiente e visíveis a partir do RP_r . Conforme Equação 4.3, os conjuntos de S_m^r dos $|M|$ APs, coletadas no RP_r , foram agrupadas para criar uma instância i de assinatura $FP_{i,r}$ ou instância de *fingerprinting*, de um dado RP_r , de coordenada (X_r, Y_r) .

Caso os valores das amostras de RSS de um determinado AP tivessem valor menor que -90 dBm, em um instante de coleta da amostra, o valor foi igualado a -200 por convenção, significando fora de alcance de comunicação, e o valor -200 não foi considerado nos cálculos de função média, mediana, mínimo, máximo ou combinada do resumo das instâncias de *fingerprinting*, conforme será apresentado a seguir.

No bloco 3 da Figura 13, chamado de “Resumo de amostras”, executou-se uma função característica ($\text{func}(\cdot)$) sobre a instância de *fingerprinting* $FP_{i,r}$ de RSS de cada RP. A ideia foi criar um vetor de valores que resumisse as características da instância de *fingerprinting* $FP_{i,r}$, sem a necessidade de transmitir todas as amostras de RSS para o servidor. Além disso, em trabalhos da literatura, verificou-se que devido ao multipercurso dos sinais de rádio transmitidos e obstáculos móveis existentes nos ambientes internos, existem cenários em que, por exemplo, dois RPs diferentes, mesmo estando a diferentes distâncias de um certo AP, podem coletar amostras de RSS semelhantes deste AP específico. Isto pode gerar valores de RSS que afetam a acurácia de localização (LIN; LIN, 2005). Com o objetivo de minimizar esses problemas, aplicamos sobre as instâncias de *fingerprinting* $FP_{i,r}$ as funções de resumo de amostras RSS relacionadas abaixo, para resumir as características da instância de *fingerprinting* $FP_{i,r}$:

- **Média:** correspondendo a média das $|N|$ amostras de RSS coletadas para um dado RP_r ;
- **Máximo:** correspondendo ao valor máximo das $|N|$ amostras de RSS coletadas para um dado RP_r ;
- **Mínimo:** correspondendo ao valor mínimo das $|N|$ amostras de RSS coletadas para um dado RP_r ;
- **Mediana:** correspondendo à mediana das $|N|$ amostras de RSS coletadas para um dado RP_r ;
- **Funções combinadas:** correspondendo à média dos valores das funções de resumo de média, máximo, mínimo e mediana das $|N|$ amostras de RSS coletadas para um dado RP_r . Utiliza-se esta função com o objetivo de observar o comportamento das funções de resumo combinadas, em relação a acurácia de localização.

$$RP_r, \quad \text{com } r = 1, \dots, |R| \quad (4.1)$$

$$S_m^r = RSS_{1,m}^r, RSS_{2,m}^r, \dots, RSS_{|N|,m}^r \quad (4.2)$$

$$FP_{i,r} = S_1^r, S_2^r, \dots, S_{|M|}^r \quad (4.3)$$

$$FP_{i,t} = S_1^t, S_2^t, \dots, S_{|M|}^t \quad (4.4)$$

$$func(FP_{i,r}) = func(S_1^r), func(S_2^r), \dots, func(S_{|M|}^r) \quad (4.5)$$

$$func(FP_{i,t}) = func(S_1^t), func(S_2^t), \dots, func(S_{|M|}^t) \quad (4.6)$$

O cálculo da função de resumo aplicado na Equação 4.3 (relacionado a fase de treinamento), resultou na Equação 4.5. Enquanto o cálculo da função de resumo aplicado na Equação 4.4 (relacionado a fase de testes), resultou na Equação 4.6. Nas Equações 4.5 e 4.6, *func* representa uma das funções de resumo de média, máximo, mínimo, mediana e combinada. (4) No bloco 4, “Envio de dados ao servidor” o vetor resultante *fingerprint* $func(FP_{i,r})$ foi transmitido para o servidor de aplicação.

Para transmissão de informações entre o smartphone e o servidor Web utilizou-se os métodos PUT e GET do HTTP (*Hyper Text Transfer Protocol*) que foi aplicado para transmitir objetos no formato JSON (*Java Script Object Notation*). O servidor Web foi configurado para fornecer serviços utilizando a API (*Application Program Interface*) Web REST (*Representational State Transfer*), permitindo comunicação aplicação-servidor. O servidor foi equipado com o algoritmo de aprendizagem de máquina WKNN que utilizamos para propor o ILS desta dissertação. Adotamos a arquitetura cliente-servidor a fim de transferir os cálculos computacionais do algoritmo para o servidor, uma vez que o servidor possuía maior capacidade de processamento e armazenamento. Na Figura 13, bloco 5 “Aplicação servidor” e, ainda na fase de treinamento, foi implementado o serviço que permite adicionar vetores objetos (*dataframes*) de $func(FP_{i,r})$ à “base de dados” localizado no servidor, bloco 6.

4.1.2 Fase de validação e Fase de testes (*on-line*)

Antes da fase de testes (*on-line*), realizou-se a validação do modelo preditivo do algoritmo de aprendizagem de máquina do ILS proposto, utilizando a técnica de validação cruzada (κ -*fold*), por ser uma técnica robusta e muito utilizada na validação de modelos que empregam algoritmos de aprendizagem de máquina (LEI, 2020). No processo de validação, primeiramente separou-se a base de dados de treinamento em um subconjunto de treinamento (80%) e, os restantes (20%) da base para o subconjunto de testes. Dividiu-se o subconjunto de treinamento em $\kappa = 5$ grupos de tamanhos iguais com regiões de treinamentos e testes distribuídos conforme Figura 14. Em seguida, foram

feitas 5 interações, sendo que, em cada interação foi realizado o treinamento dos dados nas regiões de treinamento e avaliação na região de testes, resultando em respectivo valor de acurácia do erro de localização. A média dos valores das acurácias das interações resultaram na acurácia total do ILS do modelo preditivo proposto, alcançando 97,75% na coordenada X' e de 98,32% para a coordenada Y' . Com o objetivo de avaliar a capacidade de generalização do modelo, utilizou-se dados do subconjunto de testes para avaliar a acurácia do ILS do modelo preditivo proposto utilizando dados não processados da base de dados. Os resultados apresentaram 92,12% na coordenada X' e 93,07% para a coordenada Y' .

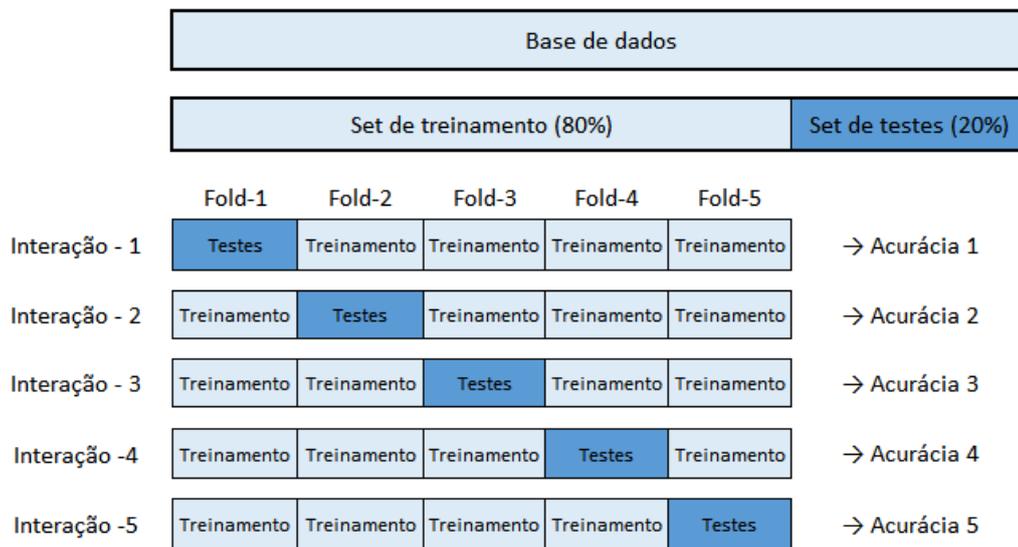


Figura 14 – Validação Cruzada κ -fold do modelo proposto.

No experimento de validação, o parâmetro \hat{K} , que representa o número de vizinhos mais próximos das instâncias de treinamento, assumiu os valores de 1, 2, 3, 4, 5, 6, 7, 8 e 9 com o objetivo de avaliar o ILS proposto em relação a acurácia. Desta forma, com $\hat{K} = 5$, foi obtido melhor desempenho da acurácia na fase de validação.

Na fase de testes (*online*), o ILS desenvolvido foi utilizado para permitir a um usuário de smartphone utilizar uma aplicação Android para obter sua localização. A localização foi realizada a partir da coleta de amostras de RSS dos APs em TPs (*Test Points*) cujas coordenadas coincidem com as coordenadas dos RPs. A Figura 13 mostra o fluxograma do processo de localização utilizado na fase de testes. O processo de testes utilizou os blocos 7, 8 e 9 com funções similares aos blocos 2, 3 e 4 utilizados na fase de treinamento, e um novo bloco (bloco 10), cuja função será explicada mais adiante. O bloco 7 da fase de testes possuiu função semelhante ao bloco 2 da fase de treinamento. Contudo, na fase de testes, utilizamos a Equação 4.2 e a Equação 4.4, para formar o conjunto S_m para cada um dos $|M|$ APs existentes no ambiente e visíveis a partir do

TP_t , de índice t , de coordenadas (X', Y') a serem determinadas. O bloco 8 da fase de testes (Bloco “Resumo de amostras”) realizou função idêntica ao do bloco 3 da fase de treinamento, sendo executado uma função $func(.)$ sobre a instância de *fingerprinting* $FP_{i,t}$ de RSS de cada TP de índice t na Equação 4.6, onde $func(.)$ foi a função média, mediana, mínimo, máximo ou combinada, igual a função de sumarização de dados utilizada na fase de treinamento. O bloco 9 da fase de testes (“Envio de dados para servidor”) possuiu mesma função do bloco 4 da fase de treinamento. No bloco 10 (“Algoritmo WKNN”) foi implementado o algoritmo de localização do ILS, utilizando o algoritmo WKNN. O bloco utilizou os valores de amostras de RSS coletados da instância de testes, num instante de tempo e atribuiu pesos conforme a distância entre os atributos da instância de teste e os atributos do(s) \hat{K} vizinho(s) mais próximo(s), obtendo assim coordenadas de posição do TP a ser localizado. Neste bloco, o algoritmo de localização do ILS, inicialmente calculou a distância euclidiana $D_{(r,t)}$, em espaço de sinais, entre os atributos dos resumos das instâncias de *fingerprinting* de um dado TP_t , determinado na fase de testes, e os atributos do RP_r da base de treinamento, conforme representado na Equação 4.7, onde $func(FP_{i,r,m})$ e $func(FP_{i,t,m})$ são os elementos individuais dos conjuntos $func(FP_{i,r})$ e $func(FP_{i,t})$ das Equações 4.3 e 4.4. Utilizamos a variável m para percorrer cada elemento dos conjuntos $func(FP_{i,r})$ e $func(FP_{i,t})$ e, onde $func(FP_{i,r})$ e $func(FP_{i,t})$ são os resumos da instância de *fingerprinting* i das fases de treinamento e testes.

$$D_{(r,t)} = \sqrt{\sum_{m=1}^{|M|} [(func(FP_{i,r,m})) - (func(FP_{i,t,m}))]^2} \quad (4.7)$$

Em seguida foi feita uma seleção das $D_{(r,t)}$ menores distâncias em espaço de sinais com base no parâmetro $K=1, \dots, \hat{K}$, para um dado TP_t da fase de testes. Estes valores foram armazenados em um vetor $Dist$ em ordem crescente de distância de sinais, conforme representado na Equação 4.8.

$$Dist = dist_1, dist_2, \dots, dist_{\hat{K}} \quad (4.8)$$

Os \hat{K} RPs vizinhos mais próximos de um TP_t foram utilizados para determinar as coordenadas (X', Y') do TP_t , conforme a Equação 4.9 e Equação 4.10. Nestas equações, os pesos $dist_K$, aplicados às coordenadas dos \hat{K} vizinhos mais próximos, são proporcionais à distância, em espaço de sinais, de cada um dos \hat{K} vizinhos para o TP_t da fase de testes, enquanto x_K e y_K são as coordenadas dos \hat{K} vizinhos mais próximos.

$$X' = \frac{\sum_{k=1}^{\hat{K}} (\frac{x_k}{dist_k})}{\sum_{k=1}^{\hat{K}} (\frac{1}{dist_k})} \quad (4.9)$$

$$Y' = \frac{\sum_{k=1}^{\hat{K}} \left(\frac{y_k}{dist_k} \right)}{\sum_{k=1}^{\hat{K}} \left(\frac{1}{dist_k} \right)} \quad (4.10)$$

Utilizou-se a biblioteca científica da linguagem de programação *Python scikit-learn*, empregada para desenvolver aplicações de aprendizagem de máquina (PEDREGOSA et al., 2011), a fim de desenvolver o algoritmo de localização do ILS proposto.

4.2 Equipamentos utilizados e setup

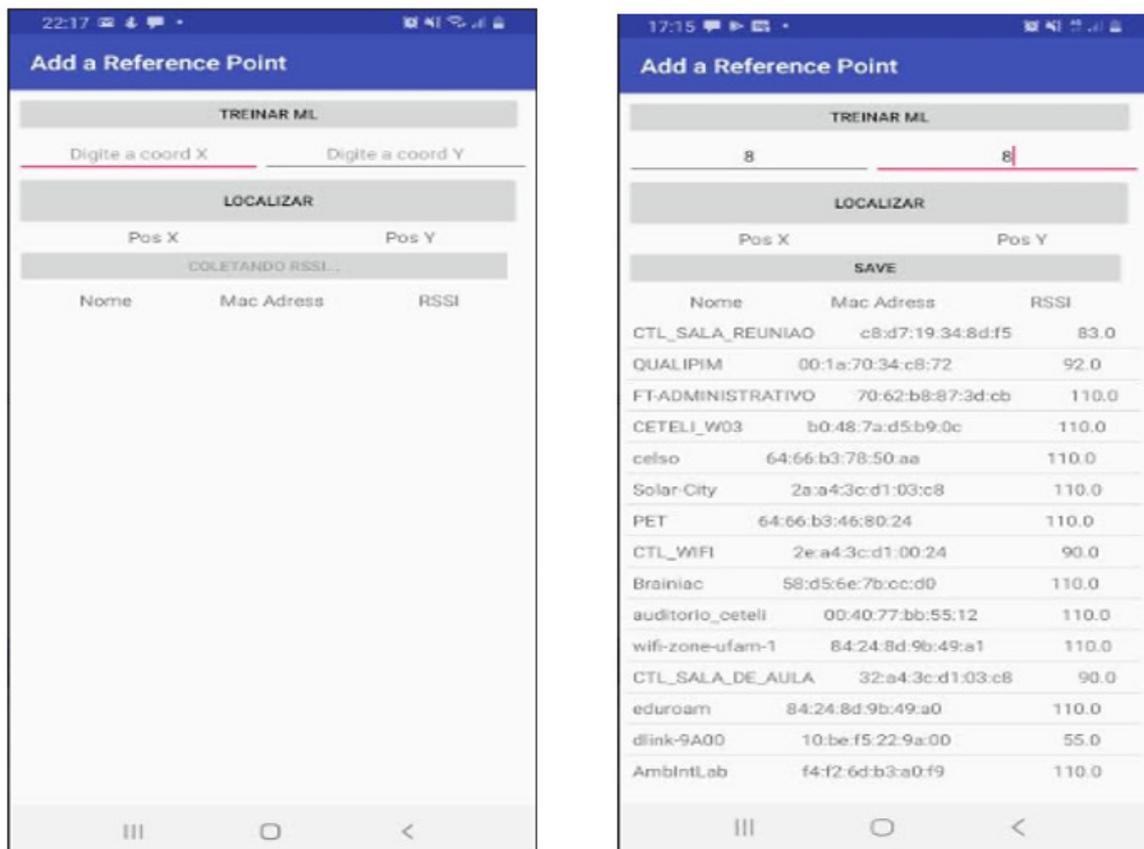
Para realizar os experimentos utilizou-se alguns equipamentos que foram configurados no ambiente de testes permitindo a implementação e avaliação do sistema ILS proposto.

Neste experimento utilizou-se um smartphone da marca Samsung modelo Galaxy A50, equipado com aplicação EZNAV (*Easy Navigation*), que foi desenvolvida empregando a metodologia definida neste trabalho. Esta aplicação EZNAV foi desenvolvida no Android Studio que é uma plataforma de desenvolvimento de aplicações para dispositivos smartphones com sistema operacional Android na versão 3.5 para Windows. Na Figura 15(a) pode-se visualizar a tela de cadastro de RPs, com os campos para inserção das coordenadas X_r, Y_r . Enquanto na Figura 15(b), pode-se visualizar a tela de coleta de amostras RSS, com o botão "Treinar ML" para enviar dados para o Servidor (fase de treinamento) e o botão "Localizar" para determinar coordenadas X', Y' (fase de testes) de um TP.

Utilizou-se um computador Notebook Gateway modelo NE57007B com processador Intel core i5 e 4GB de memória RAM, para abrigar a aplicação do servidor com o algoritmo de localização WKNN e da base de dados, com capacidade de processamento adequada, permitindo robustez e praticidade de operação.

Para desenvolvimento da aplicação do servidor, utilizou-se a ferramenta Notebook Jupyter que é um ambiente computacional web de código aberto, com padrões abertos e serviços para computação interativa em várias linguagens de programação incluindo o *Python*.

Para desenvolver um algoritmo de modelo preditivo de aprendizagem de máquina usando método WKNN no servidor, utilizou-se uma ferramenta de aplicação Web API Rest baseado em *Flask* que é um micro framework que possibilita desenvolver aplicações Web utilizando o *Python*. Na aplicação *Flask* do servidor Web, utilizou-se o objeto *Request* para recebimento e envio de dados via HTTP, uma função *Flask* chamada *Jsonify* para serializar um objeto no formato JSON e a biblioteca *numpy* para manipulação de vetores recebidos da aplicação do smartphone. Para salvar o modelo preditivo de aprendizagem de máquina na "base de dados", após cada adição de novos vetores objetos na fase de treinamento, utilizou-se o objeto *pickle*, que é uma biblioteca capaz de transformar qualquer variável



(a) Tela de cadastro de RPs

(b) Tela de coleta amostras RSS

Figura 15 – Telas da aplicação EZNAV.

em um tipo binário (permitindo salvar como um arquivo). Este objeto *pickle* possibilitou salvar o modelo treinado no computador servidor e, com isso, usa-lo para prever novos dados sem precisar treinar um novo modelo.

Para testes da API, através da submissão de dados para a aplicação no servidor, contendo coletas de amostras de RSS dos APs Wi-Fi, e o retorno de informações do modelo preditivo com as coordenadas X', Y' do TP, utilizou-se a aplicação *Insomnia* (PLAZA, 2021) para executar serviços HTTP, utilizando a API desenvolvida no servidor de aplicação.

5 Resultados

Nesta seção são apresentados os resultados alcançados com os experimentos realizados em um cenário real utilizando o ILS desenvolvido nesta dissertação. Na Figura 16, observa-se o mapa do ambiente de experimentação, representado em um plano bidimensional, com todos os RPs (quadrados na cor verde e azul) utilizados na fase *off-line* e os TPs (quadrados na cor azul) utilizados na fase *on-line*. Os valores apresentados entre parênteses representam as coordenadas (X, Y) em metros. Este cenário representa o andar térreo do bloco (Centro de P&D em Tecnologia Eletrônica e da Informação - CETELI), localizado na UFAM (Universidade Federal do Amazonas).

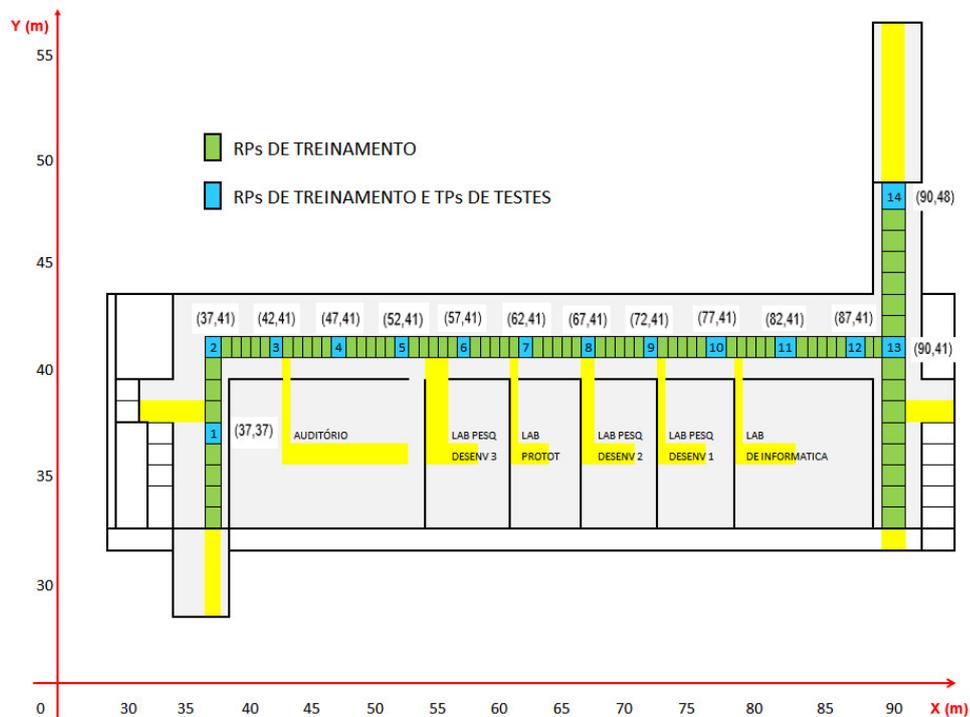


Figura 16 – Ambiente de experimentação no bloco CETELI.

Neste ambiente, para a fase de treinamento, foi definido um total de $R=85$ RPs. Cada RP possuía área de $1\text{ m} \times 1\text{ m}$. Em cada RP foram coletadas $|N|=10$ amostras de RSS de cada AP visível (AP capaz de se comunicar com smartphone localizado no RP), sendo este conjunto de amostras utilizado para criar uma instância de RP_r e assinatura $FP_{i,r}$, sendo utilizado nos experimentos 10 instâncias. Como exemplo, cita-se que se em um RP_r hipotético, fossem percebidos 7 APs pela aplicação Android, teriam sido coletadas um total de 700 amostras de RSS (ex. 7 APs x 10 instâncias i de *fingerprinting* x 10 amostras de RSS/AP/instância) no RP_r .

Comenta-se que neste exemplo hipotético, as 700 amostras seriam inicialmente

armazenadas no aplicativo Android e, em seguida, o aplicativo resumiria as 700 amostras em 70 instâncias de *fingerprinting* utilizando funções de resumo das amostras de RSS. Estas 70 instâncias seriam então transmitidas pelo aplicativo Android para o servidor da base de dados, conforme realização dos experimentos individuais de utilização das funções de resumo de média, mediana, mínimo, máximo e funções combinadas. Após a fase de treinamento no cenário de experimento, prosseguiu-se com a fase de testes definindo um caminho com 14 TPs marcados em azul na Figura 21, e coincidentes com os RPs já mapeados no ambiente na etapa anterior. Na fase de testes, realizou-se a localização do smartphone nos TPs variando o número $|N|$ de amostras de RSS coletadas para o conjunto S_m^r formado por um conjunto de $|N|$ amostras = 2, 4, 6, 8, 10, para cada um dos $|M|$ APs visíveis em cada TP do ambiente e, para cada valor de \hat{K} . Com isso foi possível avaliar o desempenho do ILS calculando os erros de localização dos valores das coordenadas reais (X, Y) em comparação as coordenadas preditas (X', Y') pelo ILS.

Utilizou-se o erro médio absoluto ou erro médio MAE (*Mean Absolut Error*) e a raiz do erro quadrático médio ou erro RMSE (*Root Mean Squared Error*) para calcular os erros de localização. No cálculo do erro médio MAE, cada erro contribui de forma proporcional ao valor absoluto do erro. Enquanto no cálculo do erro RMSE, é considerado a quadratura das diferenças dos erros, de tal forma que algumas diferenças grandes aumentarão o valor do erro RMSE em um grau maior que o MAE (PONTIUS; THONTTEH; CHEN, 2008)(JAMES et al., 2013). As Equações 5.1, 5.2 e 5.3 foram utilizadas para calcular o erro médio MAE de cada coordenada X', Y' e o erro médio MAE resultante de um dado TP na fase de testes. A variável j representa a quantidade de conjunto de dados de coordenadas observadas comparadas ao conjunto de dados de coordenadas preditas, sendo utilizado um $j = 1$ conjunto de dados para o cálculo do erro de localização.

$$MAE_X = \frac{\sum_{v=1}^j |X_v - X'_v|}{j} \quad (5.1)$$

$$MAE_Y = \frac{\sum_{v=1}^j |Y_v - Y'_v|}{j} \quad (5.2)$$

$$MAE = \sqrt{MAE_X + MAE_Y} \quad (5.3)$$

As Equações 5.4, 5.5 e 5.6 foram utilizadas para determinar o erro RMSE de cada coordenada (X', Y') e o erro RMSE resultante de um dado TP na fase de testes.

$$RMSE_X = \sqrt{\frac{1}{j} \sum_{v=1}^j (X_v - X'_v)^2} \quad (5.4)$$

$$RMSE_Y = \sqrt{\frac{1}{j} \sum_{v=1}^j (Y_v - Y'_v)^2} \quad (5.5)$$

$$RMSE = \sqrt{RMSE_X^2 + RMSE_Y^2} \quad (5.6)$$

5.1 Erro médio MAE x Valor \hat{K}

As Tabelas 5, 6 e 7 apresentam os erros médio MAE (em metros), respectivamente, para a coordenada X' , coordenada Y' e o erro médio MAE resultante da localização estimada para o TP da fase de testes, em função do parâmetro \hat{K} para as funções de resumo de amostras RSS da média, máximo, mínimo, mediana e funções combinadas.

Tabela 5 – Erro Médio MAE x Valor \hat{K} – Coordenada X' .

Erro médio MAE versus valor \hat{K} (m) - coordenada X'					
Valor \hat{K}	media	máximo	mínimo	mediana	Funcões combinadas
1	2,07	4,57	3,00	3,93	4,05
2	3,57	3,50	3,64	4,79	4,26
3	2,24	3,05	3,74	4,45	3,61
4	3,18	3,50	3,64	4,79	4,04
5	2,01	2,83	4,06	4,09	3,43
6	3,15	2,89	4,62	3,79	3,54
7	2,71	3,53	4,36	3,61	3,67
8	3,51	3,70	4,47	3,46	3,94
9	2,78	3,37	3,79	3,36	3,70

Tabela 6 – Erro Médio MAE x Valor \hat{K} – Coordenada Y' .

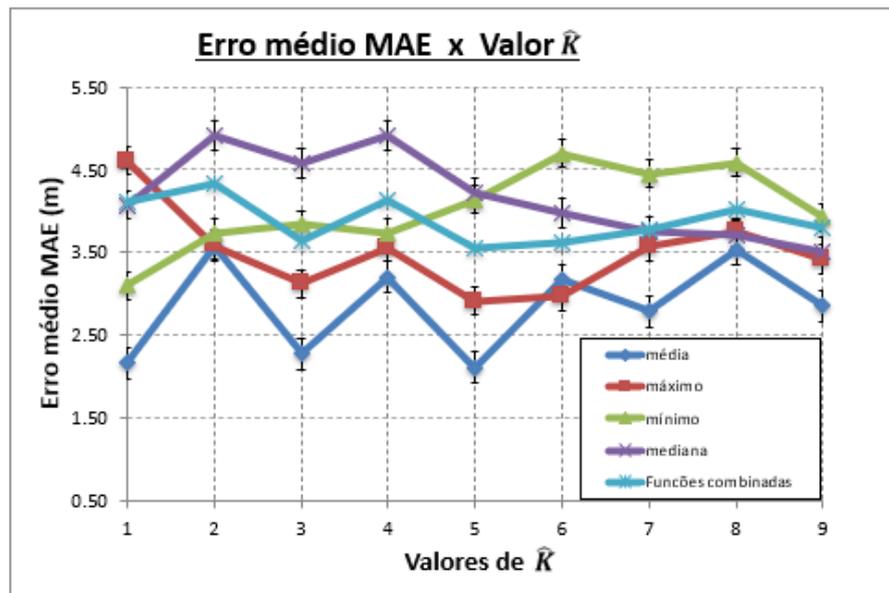
Erro médio MAE versus valor \hat{K} (m) - coordenada Y'					
Valor \hat{K}	média	máximo	mínimo	mediana	Funcões combinadas
1	0,64	0,57	0,79	1,07	0,73
2	0,29	0,79	0,86	1,11	0,78
3	0,43	0,69	0,86	1,05	0,57
4	0,36	0,64	0,86	1,11	0,87
5	0,66	0,70	0,83	1,06	0,94
6	0,31	0,65	0,82	1,20	0,79
7	0,66	0,58	0,90	1,05	0,89
8	0,43	0,60	0,97	1,31	0,84
9	0,69	0,58	1,03	1,04	0,91

Com o objetivo de observar o comportamento do erro médio MAE resultante do ILS em função do parâmetro \hat{K} , a Figura 17 apresenta a variação do valor de \hat{K} nos valores 1,2,3,4,5,6,7,8 e 9 para cada função de resumo de amostras RSS média, máximo, mínimo, mediana e combinada. Nestes resultados foi possível identificar que o ILS proposto alcançou menor erro médio MAE resultante com 2,12m utilizando o parâmetro $\hat{K}=5$ e a função de resumo média (ver Tabela 7). Verificou-se também que o erro médio MAE para a coordenada X' foi maior para todos os testes porque a coordenada Y' é fixa e,

Tabela 7 – Erro Médio MAE x Valor \hat{K} – MAE Resultante.

Erro médio MAE versus valor \hat{K} (m)					
Valor \hat{K}	média	máximo	mínimo	mediana	Funcões combinadas
1	2,17	4,61	3,10	4,07	4,12
2	3,58	3,59	3,74	4,91	4,33
3	2,28	3,12	3,84	4,57	3,65
4	3,20	3,56	3,74	4,91	4,14
5	2,12	2,91	4,14	4,22	3,55
6	3,17	2,97	4,69	3,97	3,63
7	2,79	3,58	4,45	3,76	3,78
8	3,54	3,74	4,58	3,70	4,03
9	2,86	3,42	3,93	3,52	3,81

assim, erra-se menor quantidade de vezes a coordenada Y' que a coordenada X' e, sendo assim, a coordenada Y' estimada é mais próxima da coordenada Y real. Observou-se que as funções combinadas apresentaram menor variação nos erros médios MAE em função de \hat{K} , apesar de alcançar erros médios MAE maiores que a função de resumo da média, máximo e mediana.

Figura 17 – Erro médio de Localização MAE versus valor \hat{K} .

5.2 Erro médio MAE x Número de amostras RSS

A Figura 18 apresenta a avaliação do erro médio MAE do ILS em função do $|N|$ número de amostras RSS coletadas para cada um dos $|M|$ APs visíveis (representadas pelo conjunto S_m), variando $|N|=2, 4, 6, 8, 10$. Neste experimento ajustou-se o ILS para operar em melhor desempenho obtido no experimento da Figura 18, ou seja, utilizando a função

de resumo média e o parâmetro $\hat{K} = 5$. A Tabela 8 apresenta os resultados obtidos neste experimento.

Tabela 8 – Erro Médio MAE x N° de amostras RSS.

Erro médio MAE x N° de amostras RSS ($\hat{K} = 5$)	
N° de amostras RSS	Erro médio MAE (m)
2	5,75
4	6,39
6	3,81
8	6,07
10	2,12

Nos resultados, foi verificado que ao variarmos o número de amostras $|N|$ na fase de testes, o ILS alcançou o menor erro médio MAE com 2,12 m para uma quantidade de 10 amostras de RSS. Por outro lado, com quantidade de amostras RSS menores 2 e 4, o ILS proposto alcançou maiores erros médios MAE com 5,75 m e 6,39 m respectivamente. Portanto, nesta configuração, o ILS proposto apresentou melhor desempenho quando aumentou-se o número de amostras RSS de $|N|=2$ para $|N|=10$.

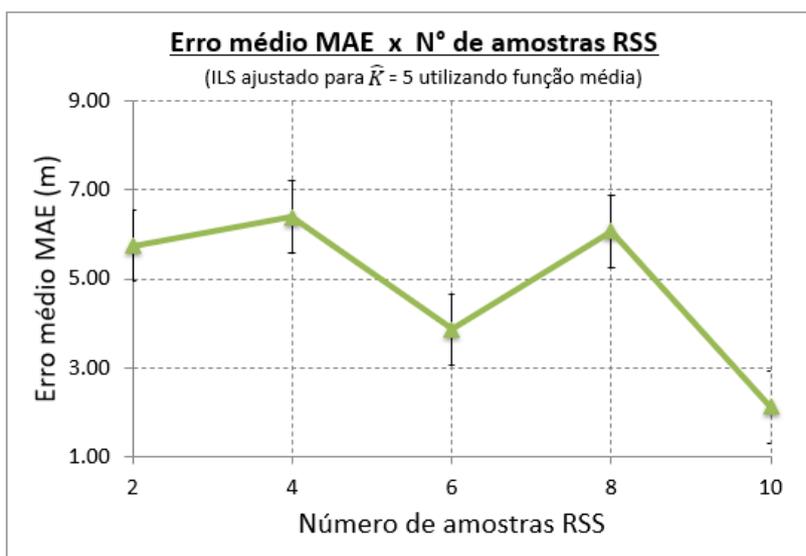


Figura 18 – Erro médio MAE versus número amostras RSS.

5.3 Erro RMSE versus valor \hat{K}

Na Figura 19, é mostrado o comportamento do erro RMSE (em metros) em função do valor \hat{K} para cada função de resumo média, máximo, mínimo, mediana e as funções combinadas. Na Tabela 9, verificou-se que ao utilizar a função de resumo média, foi alcançado o melhor desempenho obtendo-se o menor valor de erro RMSE com valor igual a 2,87 m utilizando $\hat{K} = 1$. Neste experimento, observou-se também que as funções

combinadas apresentaram menor variação nos erros RMSE em função de \hat{K} , apesar de alcançar erros RMSE maiores que a função de resumo da média, máximo e mediana.

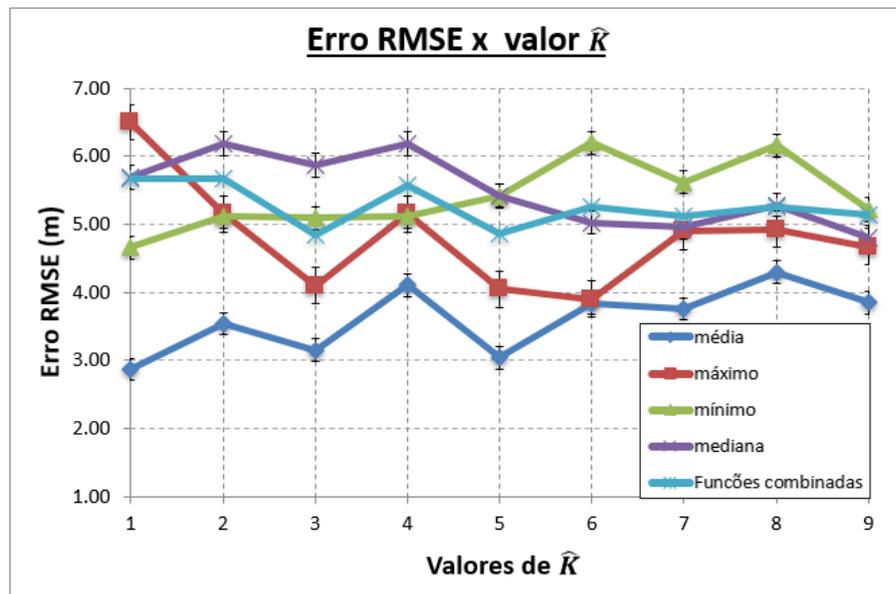


Figura 19 – Erro RMSE x valor do parâmetro \hat{K} .

Tabela 9 – Erro RMSE x Valor \hat{K} .

Erro RMSE (m) x Valor \hat{K}					
Valor \hat{K}	media	máximo	mínimo	mediana	funções combinadas
1	2,87	6,50	4,66	5,69	5,68
2	3,54	5,15	5,11	6,19	5,67
3	3,15	4,10	5,10	5,86	4,85
4	4,11	5,15	5,11	6,19	5,58
5	3,04	4,04	5,42	5,42	4,86
6	3,85	3,90	6,20	5,03	5,26
7	3,76	4,90	5,62	4,96	5,12
8	4,30	4,93	6,16	5,29	5,25
9	3,85	4,67	5,22	4,81	5,13

5.4 CDF versus erro médio MAE

Na Figura 20 visualiza-se o gráfico da Função de Distribuição Acumulada (*Cumulative Distribution Function* – CDF) para cada função de resumo das amostras RSS em relação ao erro médio MAE com parâmetro $\hat{K}=5$. Na Figura 19 verifica-se que em 79,69% das estimativas de localização, a função resumo da média apresentou erro médio MAE inferior a 4 m, superando as demais funções através das quais foi alcançado o mesmo valor de erro médio com probabilidades menores: máximo=64,27%, funções combinadas: 55,21%, mínimo=48,90% e mediana=47,50%.

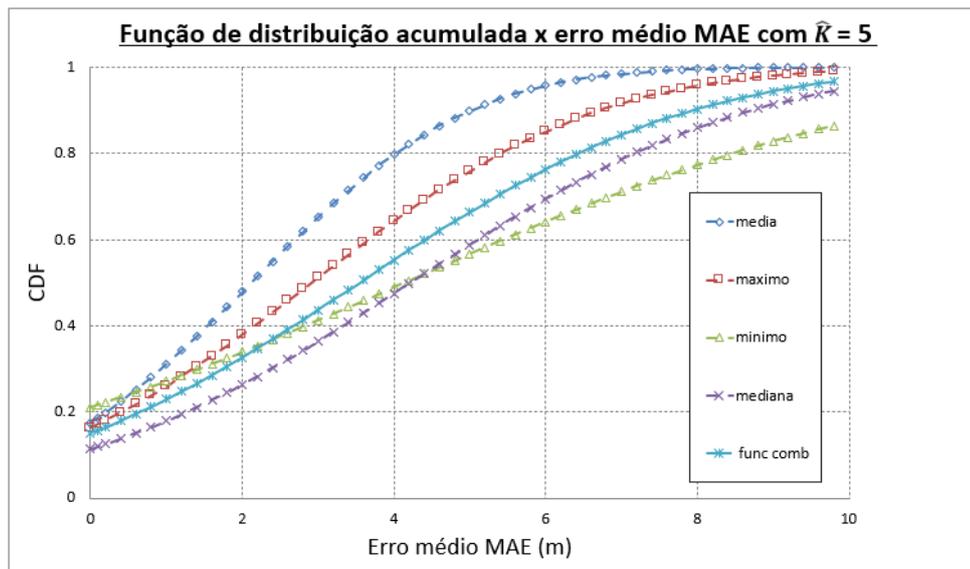


Figura 20 – CDF versus erro médio MAE.

5.5 Quantidade de APs visíveis em cada TP

A Tabela 10 apresenta o número de APs visíveis em cada TP com suas respectivas coordenadas reais X e Y (em metros) no ambiente de experimentação da Figura 16. Com base nos dados verificamos no TP 1 e TP 2, a aplicação desenvolvida conseguiu coletar amostras de RSS de 14 APs, enquanto no TP 11 a aplicação desenvolvida conseguiu coletar somente 3 APs.

Tabela 10 – Quantidade de APs visíveis em cada TP.

N. do TP	Coordenada X (m)	Coordenada Y (m)	Quantidade de APs visíveis
1	37	37	14
2	37	41	14
3	42	41	10
4	47	41	10
5	52	41	8
6	57	41	8
7	62	41	5
8	67	41	6
9	72	41	5
10	77	41	4
11	82	41	3
12	87	41	4
13	90	41	5
14	90	48	5

A Figura 26 apresenta a quantidade de APs visíveis em cada TP.

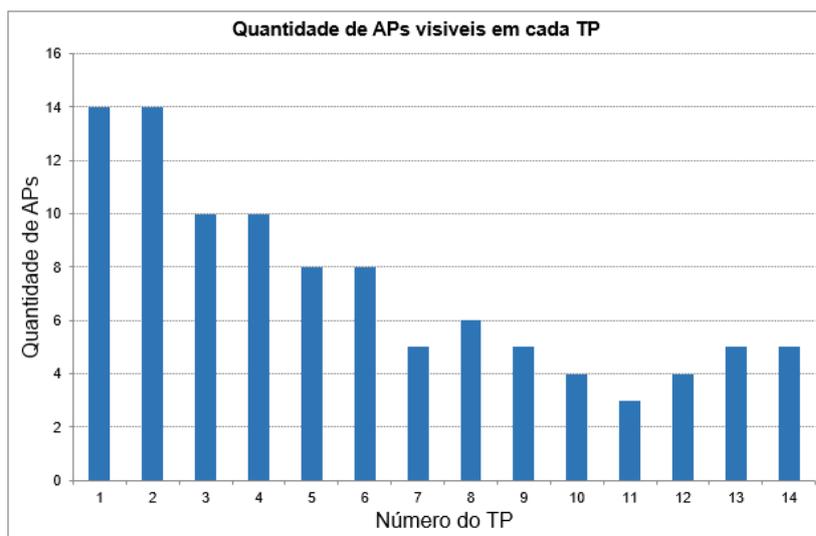


Figura 21 – Quantidade de APs visíveis em cada TP.

5.6 Número de APs visíveis x erro médio MAE

A Tabela 11 apresenta o erro médio MAE em função da quantidade de APs visíveis em cada TP no ambiente de experimentação.

Tabela 11 – Número de APs visíveis x Erro médio MAE.

N. do TP	Coordenada X (m)	Coordenada Y (m)	Nº de APs visíveis	Erro médio MAE (m)
1	37	37	14	1,36
2	37	41	14	3,04
3	42	41	10	5,75
4	47	41	10	7,64
5	52	41	8	3,16
6	57	41	8	4,16
7	62	41	5	2,28
8	67	41	6	6,56
9	72	41	5	3,48
10	77	41	4	3,72
11	82	41	3	4,36
12	87	41	4	6,72
13	90	41	5	8,92
14	90	48	5	11,16

Na Figura 22, o eixo x representa o número do TP que pode variar de 1 a 14. O eixo y e as barras azuis representam a quantidade de APs visíveis a partir dos TPs que variam de 1 a 14 e, a curva laranja representa o erro médio MAE medido em cada um dos 14 TPs. Nestes resultados verificamos que o menor erro médio MAE foi alcançado no TP1 com erro médio de 1,36 m com 14 APs detectados, enquanto que o maior erro médio MAE

foi alcançado no TP14 com erro médio MAE de 11,16 m com 5 APs visíveis. É importante observar outros TPs com o 7 no qual foi alcançado erro médio MAE de 2,28 m com 5 APs detectados e o TP8 no qual foi calculado erro médio MAE de 6,56 m com 6 APs. Nestes pontos e próximos verificou-se pelos resultados, um comportamento diferente em relação aos resultados alcançados nos TPs iniciais e finais do cenário devido a fatores não identificados durante o experimento. Desta forma, não foi possível determinar uma relação linear para todo o cenário de teste, visto que pode haver outros fatores a ser considerados.

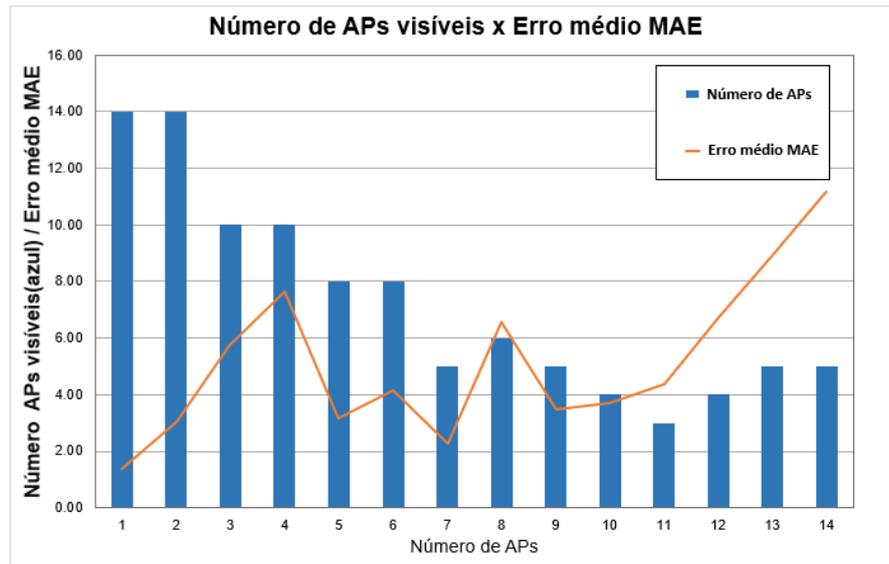


Figura 22 – Número Aps visíveis versus erro médio MAE.

Os APs existentes no ambiente e utilizados no experimento não ficaram visíveis a partir de todos os TPs uma vez que os valores de RSS estavam abaixo do valor de sensibilidade mínima de recepção do adaptador Wi-Fi do smartphone utilizado nos experimentos.

6 Conclusão

Neste trabalho de dissertação, foi apresentada proposta de ILS utilizando dispositivo smartphone equipado com aplicação Android que foi desenvolvido para medir e armazenar em uma base de dados, o RSS dos sinais emitidos por APs existentes no ambiente. Além da aplicação Android, desenvolvemos aplicação, localizada em computador servidor, baseada no algoritmo de aprendizagem de máquina WKNN e que estima a localização de um dispositivo smartphone equipado com a aplicação Android desenvolvida.

Concluimos que ao utilizar a função média sobre as amostras de RSS, obtivemos resultados promissores alcançando erro RMSE de 2,87 m. Constatamos também que o ILS proposto apresentou melhor desempenho quando aumentou-se o número N de amostras de RSS de 2 para 10.

Como trabalhos futuros podemos relacionar:

- Avaliar o tempo de resposta do ILS proposto neste artigo e verificar a operação em tempo real;
- Analisar o comportamento do ILS em períodos de tempo diferentes do dia (manhã, tarde e noite) e da semana (dias da semana e fim de semana);
- Implementar ILS em plano tridimensional (X,Y,Z) de forma a considerar vários andares de um prédio;
- Implementar e avaliar outros tipos de funções de resumo de amostras no ILS desenvolvido.

Referências

- ALNAFESSAH, A.; AL-AMMAR, M.; AL-HADHRAMI, S.; AL-SALMAN, A.; AL-KHALIFA, H. Developing an Ultra Wideband Indoor Navigation System for Visually Impaired People. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 12, n. 7, p. 6152342, 2016. Citado na página 14.
- ARYA, A.; GODLEWSKI, P.; MELLÉ, P. A Hierarchical Clustering Technique for Radio Map Compression in Location Fingerprinting Systems. In: IEEE. *2010 IEEE 71st Vehicular Technology Conference*. [S.l.], 2010. p. 1–5. Citado na página 15.
- BAHL, P.; PADMANABHAN, V. N. RADAR: An In-Building RF-Based User Location and Tracking System. In: IEEE. *Proceedings IEEE INFOCOM 2000. Conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064)*. [S.l.], 2000. v. 2, p. 775–784. Citado na página 20.
- BELLONI, F. Bringing Navigation Indoors. *Nokia Research Center Radio Systems Laboratory, Finland*, 2010. Citado na página 14.
- BI, J.; WANG, Y.; LI, X.; QI, H.; CAO, H.; XU, S. A Novel Adaptive Weighted K-Nearest Neighbor Positioning Method Based on Omnidirectional Fingerprint Database and Twice Affinity Propagation Clustering. In: IEEE. *2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS)*. [S.l.], 2018. p. 1–9. Citado 3 vezes nas páginas 15, 17 e 34.
- CARVALHO, A. de S.; REZENDE, J. F. de; CARVALHO, C. B. Sistema de Localização em Ambientes Internos usando Wi-Fi e algoritmo WKNN. SBrT, 2020. Citado na página 30.
- CARVALHO, C.; MOTA, E.; FERRAZ, E.; SEIXAS, P.; SOUZA, P.; TAVARES, V.; FILHO, W. L.; FERREIRA, D.; MANZONI, P.; CALAFATE, C. Entropy Based Routing for Mobile, Low Power and Lossy Wireless Sensors Networks. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 15, n. 7, p. 1550147719866134, 2019. Citado na página 14.
- CHIU, C.-C. *Indoor Positioning System with BLE and Wi-Fi technology-Data Analysis and Accuracy Improvement*. Dissertação (Mestrado) — České vysoké učení technické v Praze. Vypočetní a informační centrum., 2017. Citado 4 vezes nas páginas 16, 17, 32 e 34.
- CORREA, A.; BARCELO, M.; MORELL, A.; VICARIO, J. L. A Review of Pedestrian Indoor Positioning Systems for Mass Market Applications. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 17, n. 8, p. 1927, 2017. Citado na página 16.
- DOUSH, I. A.; ALSHATNAWI, S.; AL-TAMIMI, A.-K.; ALHASAN, B.; HAMASHA, S. ISAB: Integrated Indoor Navigation System for the Blind. *Interacting with Computers*, Oxford University Press, v. 29, n. 2, p. 181–202, 2017. Citado na página 14.

- FANG, S.-H.; LIN, T.-N.; LIN, P.-C. Location Fingerprinting in a Decorrelated Space. *IEEE Transactions on Knowledge and Data Engineering*, IEEE, v. 20, n. 5, p. 685–691, 2008. Citado na página 15.
- FERREIRA, D.; SOUZA, R.; CARVALHO, C. QA-KNN: Indoor Localization Based on Quartile Analysis and the KNN Classifier for Wireless Networks. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 20, n. 17, p. 4714, 2020. Citado na página 14.
- FERREIRA, D. A. d. O. Localização em Ambientes Internos Utilizando Redes IEEE 802.11. Universidade Federal do Amazonas, 2019. Citado na página 14.
- GANSEMER, S.; GROSSMANN, U.; HAKOBYAN, S. RSSI-Based Euclidean Distance Algorithm for Indoor Positioning Adapted for the Use in Dynamically Changing WLAN Environments and Multi-Level Buildings. In: IEEE. *2010 International conference on indoor positioning and indoor navigation*. [S.l.], 2010. p. 1–6. Citado 2 vezes nas páginas 26 e 27.
- HÄHNEL, B. F. D.; FOX, D. Gaussian Processes for Signal Strength-Based Location Estimation. In: CITESEER. *Proceeding of robotics: science and systems*. [S.l.], 2006. Citado na página 32.
- HUANG, C.-C.; MANH, H.-N. RSS-Based Indoor Positioning Based on Multi-Dimensional Kernel Modeling and Weighted Average Tracking. *IEEE Sensors Journal*, IEEE, v. 16, n. 9, p. 3231–3245, 2016. Citado na página 15.
- INSIGHT, B. *Mobile Location-Based Services*. 2012. Citado na página 14.
- JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. *An Introduction to Statistical Learning*. [S.l.]: Springer, 2013. v. 112. Citado na página 45.
- JIA, B.; HUANG, B.; GAO, H.; LI, W. On The Dimension Reduction of Radio Maps with a Supervised Approach. In: IEEE. *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*. [S.l.], 2017. p. 199–202. Citado na página 15.
- JUN, J.; HE, L.; GU, Y.; JIANG, W.; KUSHWAHA, G.; VIPIN, A.; CHENG, L.; LIU, C.; ZHU, T. Low-Overhead WiFi Fingerprinting. *IEEE Transactions on Mobile Computing*, IEEE, v. 17, n. 3, p. 590–603, 2017. Citado na página 14.
- KHALAJMEHRABADI, A.; GATSIS, N.; AKOPIAN, D. Structured Group Sparsity: A Novel Indoor WLAN Localization, Outlier Detection, and Radio Map Interpolation Scheme. *IEEE Transactions on Vehicular Technology*, IEEE, v. 66, n. 7, p. 6498–6510, 2016. Citado na página 15.
- LANE, N. D.; MILUZZO, E.; LU, H.; PEEBLES, D.; CHOUDHURY, T.; CAMPBELL, A. T. A Survey of Mobile Phone Sensing. *IEEE Communications magazine*, IEEE, v. 48, n. 9, p. 140–150, 2010. Citado na página 15.
- LEI, J. Cross-Validation with Confidence. *Journal of the American Statistical Association*, Taylor & Francis, v. 115, n. 532, p. 1978–1997, 2020. Citado na página 39.
- LI, J.; GAO, X.; HU, Z.; WANG, H.; CAO, T.; YU, L. Indoor Localization Method Based on Regional Division with IFCM. *Electronics*, Multidisciplinary Digital Publishing Institute, v. 8, n. 5, p. 559, 2019. Citado 3 vezes nas páginas 15, 17 e 33.

- LI, W.; WEI, D.; YUAN, H.; OUYANG, G. A Novel Method of WiFi Fingerprint Positioning Using Spatial Multi-Points Matching. In: IEEE. *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. [S.l.], 2016. p. 1–8. Citado na página 14.
- LI, X.; XU, D.; WANG, X.; MUHAMMAD, R. Design and Implementation of Indoor Positioning System Based on iBeacon. In: IEEE. *2016 International Conference on Audio, Language and Image Processing (ICALIP)*. [S.l.], 2016. p. 126–130. Citado na página 14.
- LIMA, W. D. Android e a Influência do Sistema Operacional Linux. *Tecnologias em Projeção*, v. 8, n. 1, p. 100–111, 2017. Citado na página 16.
- LIN, T.-N.; LIN, P.-C. Performance Comparison of Indoor Positioning Techniques Based on Location Fingerprinting in Wireless Networks. In: IEEE. *2005 International Conference on Wireless Networks, Communications and Mobile Computing*. [S.l.], 2005. v. 2, p. 1569–1574. Citado na página 38.
- MACAGNANO, D.; DESTINO, G.; ABREU, G. Indoor Positioning: A Key Enabling Technology for IoT Applications. In: IEEE. *2014 IEEE World Forum on Internet of Things (WF-IoT)*. [S.l.], 2014. p. 117–118. Citado na página 14.
- MAKKI, A.; SIDDIG, A.; SAAD, M.; BLEAKLEY, C. Survey of WiFi Positioning Using Time-Based Techniques. *Computer Networks*, Elsevier, v. 88, p. 218–233, 2015. Citado na página 14.
- MARZOUQ, M.; FADILI, H. E.; LAKHLIAI, Z.; MECHAQRANE, A.; ZENKOUAR, K. New Distance Weighted K Nearest Neighbor Model for Hourly Global Solar Irradiation Estimation. In: IEEE. *2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*. [S.l.], 2019. p. 1–5. Citado na página 26.
- NIRJON, S.; LIU, J.; DEJEAN, G.; PRIYANTHA, B.; JIN, Y.; HART, T. COIN-GPS: Indoor Localization from Direct GPS Receiving. In: *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. [S.l.: s.n.], 2014. p. 301–314. Citado na página 14.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V. et al. Scikit-Learn: Machine Learning in Python. *the Journal of machine Learning research*, JMLR. org, v. 12, p. 2825–2830, 2011. Citado na página 42.
- PENEDA, L.; AZENHA, A.; CARVALHO, A. Trilateration for Indoors Positioning within the Framework of Wireless Communications. In: IEEE. *2009 35th Annual Conference of IEEE Industrial Electronics*. [S.l.], 2009. p. 2732–2737. Citado na página 20.
- PLAZA, N. H. *Analysis of Web and API Security*. Dissertação (B.S. thesis) — Universitat Politècnica de Catalunya, 2021. Citado na página 43.
- PONTIUS, R. G.; THONTTEH, O.; CHEN, H. Components of Information for Multiple Resolution Comparison Between Maps that Share a Real Variable. *Environmental and Ecological Statistics*, Springer, v. 15, n. 2, p. 111–142, 2008. Citado na página 45.

SALAMAH, A. H.; TAMAZIN, M.; SHARKAS, M. A.; KHEDR, M. An Enhanced WiFi Indoor Localization System Based on Machine Learning. In: IEEE. *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. [S.l.], 2016. p. 1–8. Citado 3 vezes nas páginas [16](#), [17](#) e [32](#).

SHIN, B.; LEE, J. H.; LEE, T.; KIM, H. S. Enhanced Weighted K-Nearest Neighbor Algorithm for Indoor Wi-Fi Positioning Systems. In: IEEE. *2012 8th International Conference on Computing Technology and Information Management (NCM and ICNIT)*. [S.l.], 2012. v. 2, p. 574–577. Citado na página [30](#).

STELLA, M.; RUSSO, M.; BEGUŠIĆ, D. RF Localization in Indoor Environment. *Radioengineering*, v. 21, n. 2, 2012. Citado 2 vezes nas páginas [19](#) e [21](#).

SUBBU, K. P.; ZHANG, C.; LUO, J.; VASILAKOS, A. V. Analysis and Status Quo of Smartphone-Based Indoor Localization Systems. *IEEE wireless communications*, IEEE, v. 21, n. 4, p. 106–112, 2014. Citado na página [15](#).

TAHERI, A.; SINGH, A.; EMMANUEL, A. Location Fingerprinting on Infrastructure 802.11 Wireless Local Area Networks (WLANs) using Locus. In: IEEE. *29th Annual IEEE International Conference on Local Computer Networks*. [S.l.], 2004. p. 676–683. Citado na página [21](#).

TEUBER, A.; EISSFELLER, B. WLAN Indoor Positioning Based on Euclidean Distances and Fuzzy Logic. In: *Proceedings of the 3rd Workshop on Positioning, Navigation and Communication*. [S.l.: s.n.], 2006. p. 159–168. Citado na página [30](#).

WANG, B.; GAN, X.; LIU, X.; YU, B.; JIA, R.; HUANG, L.; JIA, H. A Novel Weighted KNN Algorithm Based on RSS Similarity and Position Distance for Wi-Fi Fingerprint Positioning. *IEEE Access*, IEEE, v. 8, p. 30591–30602, 2020. Citado 3 vezes nas páginas [15](#), [17](#) e [33](#).

WANG, X.; CONG, S. An Advanced Algorithm for Fingerprint Localization based on Kalman Filter. In: IEEE. *2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS)*. [S.l.], 2018. p. 1–5. Citado 4 vezes nas páginas [15](#), [16](#), [33](#) e [34](#).

WU, C.; YANG, Z.; ZHOU, Z.; LIU, Y.; LIU, M. Mitigating Large Errors in WiFi-Based Indoor Localization for Smartphones. *IEEE Transactions on Vehicular Technology*, IEEE, v. 66, n. 7, p. 6246–6257, 2016. Citado na página [15](#).

WU, C.; ZHU, M.; ZHANG, Y. The Design and Implementation of an Infrared Indoor Positioning and Lighting System. In: IEEE. *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. [S.l.], 2017. p. 2134–2138. Citado na página [14](#).

XI, L.; JIANG, M. Y.; LIAO, C. Z. Indoor Positioning System Based on RFID. In: TRANS TECH PUBL. *Applied Mechanics and Materials*. [S.l.], 2014. v. 543, p. 2627–2631. Citado na página [14](#).

XIAO, J.; ZHOU, Z.; YI, Y.; NI, L. M. A Survey on Wireless Indoor Localization from the Device Perspective. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 49, n. 2, p. 1–31, 2016. Citado na página [15](#).

- XUE, M.; SUN, W.; YU, H.; TANG, H.; LIN, A.; ZHANG, X.; ZIMMERMANN, R. Locate the Mobile Device by Enhancing the WiFi-Based Indoor Localization Model. *IEEE Internet of Things Journal*, IEEE, v. 6, n. 5, p. 8792–8803, 2019. Citado na página 15.
- XUE, W.; HUA, X.; LI, Q.; YU, K.; QIU, W. Improved Neighboring Reference Points Selection Method for Wi-Fi Based Indoor Localization. *IEEE Sensors Letters*, IEEE, v. 2, n. 2, p. 1–4, 2018. Citado 4 vezes nas páginas 15, 16, 33 e 34.
- YANG, X.; ZHAO, F.; CHEN, T. NLOS Identification for UWB Localization Based on Import Vector Machine. *AEU-International Journal of Electronics and Communications*, Elsevier, v. 87, p. 128–133, 2018. Citado na página 14.
- YIU, S.; DASHTI, M.; CLAUSSEN, H.; PEREZ-CRUZ, F. Locating User Equipments and Access Points using RSSI Fingerprints: A Gaussian process approach. In: IEEE. *2016 IEEE International Conference on Communications (ICC)*. [S.l.], 2016. p. 1–6. Citado 6 vezes nas páginas 15, 16, 22, 32, 33 e 34.
- ZAFARI, F.; GKELIAS, A.; LEUNG, K. K. A Survey of Indoor Localization Systems and Technologies. *IEEE Communications Surveys & Tutorials*, IEEE, v. 21, n. 3, p. 2568–2599, 2019. Citado na página 15.
- ZHANG, Y.; LIU, W.; FANG, Y.; WU, D. Secure Localization and Authentication in Ultra-Wideband Sensor Networks. *IEEE Journal on Selected areas in communications*, IEEE, v. 24, n. 4, p. 829–835, 2006. Citado na página 15.