



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

LOCALIZAÇÃO *Indoor* BASEADA EM TECNOLOGIA BLUETOOTH DE BAIXA ENERGIA

WAGNER CIPRIANO

Orientador: Paulo Eduardo Maciel de Almeida
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientador: Flávio Vinícius Cruzeiro Martins
Centro Federal de Educação Tecnológica de Minas Gerais

BELO HORIZONTE
AGOSTO DE 2018

WAGNER CIPRIANO

**LOCALIZAÇÃO *Indoor* BASEADA EM TECNOLOGIA
BLUETOOTH DE BAIXA ENERGIA**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Sistemas Inteligentes

Orientador: Paulo Eduardo Maciel de Almeida
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientador: Flávio Vinícius Cruzeiro Martins
Centro Federal de Educação Tecnológica de Minas Gerais

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
AGOSTO DE 2018

Cipriano, Wagner
C5771 Localização Indoor baseada em tecnologia Bluetooth de baixa energia / Wagner Cipriano. – 2018.
80 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.

Orientador: Paulo Eduardo Maciel de Almeida.

Coorientador: Flávio Vinícius Cruzeiro Martins.

Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Sistemas de posicionamento de interiores (localização sem fio) – Teses. 2. Inteligência computacional – Teses. 3. Processamento de sinais – Técnicas digitais – Teses. 4. Triangulação – Teses. I. Almeida, Paulo Eduardo Maciel de. II. Martins, Flávio Vinícius Cruzeiro. III. Centro Federal de Educação Tecnológica de Minas Gerais. IV. Título.

CDD 621.38223

Agradecimentos

Foi difícil a decisão de inserir esta página devido ao medo de ser injusto e esquecer de pessoas importantes nesta caminhada, mas eu resolvi arriscar e desde já peço desculpas por não conseguir inserir todos, em apenas uma página. Mas eu tenho certeza que cada um sabe o quanto foi importante nesta caminhada, independente de ter ou não o nome mencionado.

Primeiramente agradeço a Deus por ter permitido que eu chegasse até aqui com saúde, apesar de todas as tribulações, para finalizar este trabalho.

Aos meus pais, Maria Augusta e José Cipriano, que não tiveram a oportunidade de estudar, nem de oferecer boas condições para que seus filhos o pudessem fazer, mas eu tenho certeza que participaram deste trabalho como anjos da guarda intercedendo por mim.

Aos meus orientadores que confiaram no meu trabalho e me deram a oportunidade de desenvolver esta pesquisa. Ao CEFET-MG por ter me acolhido e oferecido a infra estrutura necessária para os desenvolvimento dos estudos. Aos responsáveis pela biblioteca do campus II, que foi disponibilizada para execução dos testes e medições. Aos professores que lecionam com profissionalismo e conseguem proporcionar aulas de qualidade. Aos brasileiros que pagam seus impostos e tornam possíveis estudos como este, que por intermédio da CAPES, recebeu auxílio financeiro.

Ao meu amor, Tatiana Santos, que teve que suportar a minha ausência, nervosismo e ansiedade na busca incansável pelos melhores resultados possíveis. À minha família, que mesmo chateados com a minha ausência, entenderam a importância do projeto.

Ao amigo Túlio Vieira, que foi fundamental e ofereceu todo apoio para que este projeto fosse concluído. Aos demais amigos de laboratório, Ramon, Rondy, Samara, Lucas, Renan e Rafaela pela colaboração e companheirismo. Ao Alexandre Frias pela sua vocação docente e toda a sua disponibilidade em ajudar, sempre. Aos amigos de estudos André, Thiago, Otaviano, Cristiano, Fernando e todos os demais que fizemos disciplinas, trabalhos ou estudamos juntos nestes últimos de CEFET-MG.

Resumo

Com o desenvolvimento da Internet das Coisas e o crescente interesse comercial por serviços baseados em localização *indoor*, tornou-se essencial o aperfeiçoamento de técnicas para suportar as aplicações. As pessoas passam de 80% a 90% do tempo em ambientes fechados e algumas das aplicações móveis que utilizam fazem sentido apenas quando a localização do usuário é conhecida. Contexto no qual os serviços dessa natureza se tornam muito importantes. O GPS é o sistema mais utilizado para localização e orientação, consolidado para ambientes externos, porém apresenta limitações para funcionamento em locais fechados. A partir da versão 4.0, de 2010, o protocolo Bluetooth passou a incluir o padrão *Low Energy* (BLE), que possibilitou o lançamento dos *beacons* BLE, em 2013/14, e impulsionou o desenvolvimento das aplicações de localização em ambientes fechados. Neste contexto, o presente trabalho avalia técnicas de localização *indoor* para suportar aplicações móveis. A partir da planta baixa do estabelecimento, é construído um modelo gráfico de representação da área de interesse, onde serão executados os experimentos. O posicionamento dos *beacons* é definido e os experimentos são conduzidos no ambiente. São desenvolvidas técnicas de localização utilizando *Fingerprinting* e Triangulação que são testadas com os dados gerados nos experimentos. Os resultados do trabalho mostram uma precisão de 90% a 100% para as técnicas de *Fingerprinting*, dependendo da granularidade utilizada. Para a Triangulação, a mediana do erro ficou entre 0 e 2 metros para todos os experimentos com foco nessa abordagem e o intervalo interquartil ficou situado entre 0 e 2.25 metros em 3 dos 4 experimentos. A análise dos resultados mostra que a utilização dos *Beacons* BLE é adequada para ambientes divididos com granularidade entre 1 e 3 metros, para um intervalo de confiança entre 90% e 99%. Foi validado, nos experimentos efetuados, que é possível alcançar uma exatidão de 100% em configurações que trabalham com identificação por zona.

Palavras-chave: Localização indoor. Inteligência Computacional. Fingerprinting. Triangulação.

Abstract

With the development of the Internet of Things and the growing business demand for indoor location based services, it has become essential to improve techniques to support the applications. People spend 80% to 90% of their indoors, and some of the mobile applications they use make sense only when the provider knows where the customer is. Reason why services like this become very important. The GPS is the main system to define location and orientation over the earth, that is useful for outdoor location, however it has limitations for indoor environments. Since the version 4.0, from 2010, the Bluetooth protocol included the Low Energy (BLE) standard, which enabled the launch of BLE beacons in 2013/14, and instigated the development of indoor location applications. Context in which this work evaluates indoor location techniques to support mobile applications. From the floor plan of the establishment, a graphic model of the target area is build, where the experiments will be executed. The positioning of the beacons is defined and the experiments are conducted in the environment. Localization techniques are developed using Fingerprinting and Triangulation that are tested with the actual data, generated in the experiments. The results show an accuracy of 90% to 100% for Fingerprinting techniques, depending on the granularity used. For the Triangulation approach, the median error was in the range 0-2 meters for all experiments focused on this approach and the interquartile range was located in-between 0 and 2.25 meters in 3 of the 4 experiments. The result analysis shows that the use of BLE beacons is suitable for granularity of areas between 1 and 3 meters, for a confidence interval between 90% and 99%. It was also corroborated, in the experiments, that it is possible to achieve an accuracy of 100% when working with regions.

Keywords: Indoor navigation. Computational Intelligence. Fingerprinting. Triangulation.

Sumário

1 – Introdução	1
1.1 Motivação	3
1.2 Definição do Problema	4
1.3 Objetivos	5
1.4 Organização do Trabalho	5
2 – Referencial Teórico	7
2.1 Internet das Coisas	7
2.2 Bluetooth	9
2.2.1 Bluetooth Low Energy	10
2.3 Beacons BLE	12
2.3.1 Principais Aplicações	13
2.3.2 Interação com o Mundo Físico	14
2.3.3 Tipos de <i>Frames</i>	14
2.3.4 Distância e Localização	15
2.4 Localização <i>Indoor</i>	16
2.4.1 Localização vs Proximidade	16
2.4.2 Serviços Baseados em Localização	17
2.4.3 Tecnologias de Medição Utilizadas	17
2.4.4 Medição de Distância Baseada em RSSI	18
2.4.5 Localização por <i>Fingerprinting</i>	19
2.4.6 Localização por Triangulação	19
2.5 Redes Neurais Artificiais	21
2.5.1 Perceptron Multicamadas	22
2.6 K Vizinhos Mais Próximos	22
3 – Trabalhos Relacionados	24
3.1 Comportamento do Consumidor	24
3.2 Tecnologia a Serviço do Varejo	25
3.3 Abordagens de Localização <i>Indoor</i>	25
3.4 Discussões Finais	28
4 – Metodologia	29
4.1 Estudo de Caso	30
4.2 Hipóteses	30
4.3 Planejamento de Experimentos	31

4.3.1	Características e Mapa do Ambiente	31
4.3.2	Identificação dos <i>Beacons</i>	32
4.3.3	Experimento A - Análise de Dados	32
4.3.4	Experimento B - Tempo de Leitura	32
4.3.5	Experimento C - Calibração	33
4.3.6	Experimento D - Localização	33
4.4	Avaliação dos Experimentos	34
5	Análise Estatística dos Dados	35
5.1	Coleta dos Dados	35
5.2	Latência	35
5.3	RSSI	36
5.3.1	Normalidade dos dados	37
5.4	Discussões Sobre a Avaliação	38
6	Desenvolvimento do Modelo	39
6.1	Representação do Estabelecimento	39
6.1.1	Área de Cobertura	39
6.1.2	Posicionamento dos <i>Beacons</i>	40
6.1.3	Mapa	41
6.2	Dispositivos	42
6.2.1	Configurações Utilizadas	43
6.3	Calibração dos <i>Beacons</i>	43
6.4	Execução dos Experimentos	45
6.4.1	Aplicativo <i>Android</i>	45
6.4.2	Medições dos Pontos de Coleta	46
6.5	Algoritmos	47
6.5.1	Tratamento dos Dados	47
6.5.2	<i>Fingerprinting</i>	47
6.5.2.1	K Vizinhos Mais Próximos	48
6.5.2.2	Rede Neural <i>Multilayer Perceptron</i>	49
6.5.3	Triangulação	49
6.5.3.1	Pontos Fora da Área de Interesse	50
6.5.3.2	Limite	50
7	Resultados, Análise e Discussão	53
7.1	<i>Fingerprinting</i> KNN	54
7.1.1	Representação Cromatográfica	54
7.1.2	Resultados dos Experimentos	54
7.1.3	Experimento D4	56

7.2	<i>Fingerprinting MLP</i>	58
7.3	Triangulação	61
7.3.1	Baixa Dispersão do <i>Beacons</i>	62
7.3.2	Análise de Exatidão por Ponto	63
7.3.3	Comparação dos Resultados	64
8	– Conclusão	66
8.1	Contribuições	66
8.2	Trabalhos Futuros	67
8.3	Considerações Finais	68
	Referências	70
	 Apêndices	 75
	APÊNDICE A – Mapas dos Experimentos	76
A.1	Experimento D1	76
A.2	Experimento D2	77
A.3	Experimento D3	77
A.4	Experimento D4	78
A.5	Experimento D5	78
A.6	Experimento D6	79
A.7	Experimento D7	79
A.8	Experimento D8	80

Lista de Figuras

Figura 1 – Utilização do tempo das pessoas quanto ao ambiente	2
Figura 2 – Visão geral da IoT - Mercados verticais e integração horizontal	4
Figura 3 – Arquitetura Internet das Coisas	8
Figura 4 – Pilha de protocolos <i>Bluetooth</i>	10
Figura 5 – Canais e Frequências do Espectro BLE	11
Figura 6 – Cálculo da posição em 2d utilizando a Trilateração	20
Figura 7 – Interseção entre duas circunferências. Adaptado de Bourke (1997)	21
Figura 8 – Etapas para desenvolvimento do modelo	29
Figura 9 – Mapa do experimento D1	31
Figura 10 – Intervalo entre as coletas de dados	36
Figura 11 – <i>Boxplot</i> RSSI por <i>beacon</i>	37
Figura 12 – Histograma RSSI	37
Figura 13 – Planta do ambiente	40
Figura 14 – Mapa do experimento D2	42
Figura 15 – Dev <i>Beacon</i>	43
Figura 16 – Curva de referencia para intensidade do sinal até 20m	44
Figura 17 – Tela de configurações implementadas no aplicativo <i>Android</i>	46
Figura 18 – Tempo de coleta de dados	47
Figura 19 – Pontos de interseção dentro da área de interesse	51
Figura 20 – Parâmetro <i>Threshold</i> com distância dentro da área de abrangência	51
Figura 21 – Parâmetro <i>Threshold</i> - experimento sem nenhum ponto de interseção	52
Figura 22 – Divisão da área de interesse por grupos	53
Figura 23 – Mapa de posicionamento dos beacons (esquerda) e mapa de calor do experimento D4	55
Figura 24 – Matriz de confusão do experimento D4 por ponto - KNN	56
Figura 25 – Matriz de confusão do experimento D4 por corredor - KNN	57
Figura 26 – Matriz de confusão do experimento D4 - corredor específico - KNN	58
Figura 27 – Matriz de confusão do experimento D4 por ponto - MLP	59
Figura 28 – Matriz de confusão do experimento D4 por corredor - MLP	60
Figura 29 – Matriz de confusão do experimento D4 - Corredor específico - MLP	60
Figura 30 – Algoritmo de Triangulação	61
Figura 31 – Resultados dos experimentos de Triangulação	62
Figura 32 – Triangulação com baixa dispersão dos <i>beacons</i>	63
Figura 33 – Análise de posicionamento dos beacons	64

Lista de Abreviaturas e Siglas

ANATEL	Agencia Nacional de Telecomunicações
BLE	<i>Bluetooth Low Energy</i>
dBm	Decibel-milliwatts
DECOM	Departamento de Computação
GPS	<i>Global Positioning System</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
m	Metros
ms	Milisegundos
mW	Miliwatt
RSSI	<i>Received Signal Strength Indication</i>
TOA	<i>Time Of Arrival</i>

1 Introdução

A noção de localização sempre foi muito relevante para as pessoas. E nos últimos anos tornou-se muito importante também para os objetos devido às aplicações móveis e ao desenvolvimento da Internet das Coisas. O crescente interesse por esses serviços baseados em localização, aplicados a ambientes fechados, tem impulsionado o crescimento da área. Inúmeros serviços inteligentes fazem sentido apenas quando a localização do usuário é conhecida, sendo importante uma aproximação eficiente, principalmente em ambientes fechados. Desse modo, tecnologias cientes de localização e suas aplicações tornaram-se pré-requisitos para a implementação da computação ubíqua (JIN; LU; PARK, 2006).

Desde os séculos mais antigos, o ser humano sempre teve a necessidade de saber sobre sua localização e orientação geográfica. A principal descoberta que proporcionou um grande avanço para que as pessoas pudessem se orientar sobre sua localização no planeta Terra foi a descoberta do magnetismo. Há pelo menos dois mil anos, os chineses já sabiam que um pedaço de metal magnetizado era atraído para os polos do planeta. Tal descoberta tornou possível a criação da bússola, que revolucionou a história e proporcionou um grande avanço na exploração e colonização de novos continentes. Durante séculos, a bússola foi o instrumento mais utilizado para navegação e orientação.

Com o avanço da era espacial no século XX, novas possibilidades se abriram. Em 1957, o Sputnik 1 foi lançado pela União Soviética, sendo o primeiro satélite artificial da Terra. Posteriormente, novos satélites foram lançados e o desenvolvimento de tecnologias de comunicação espacial possibilitaram a criação dos sistemas de posicionamento global (GOSWAMI, 2012). Inicialmente projetados para uso militar e atualmente disponíveis para uso de toda a população de modo gratuito, é o sistema mais utilizado para localização e orientação de pessoas ou objetos na superfície da terra. O GPS (*Global Positioning System*) é muito útil para definição da localização geográfica, navegação, orientação e direção, para encontrar objetos, lugares ou pessoas próximas, entre outras aplicações e funcionam muito bem em ambientes externos, porém devido à atenuação ou ausência de sinal, são ineficazes em infraestruturas fechadas, conforme Liu et al. (2007) e Jin, Lu e Park (2006).

O crescimento da população traz consigo grandes desafios para a sociedade e um deles é a crescente demanda por espaços de conveniência, entretenimento, socialização e também compras, como *shopping centers*, lojas e supermercados com serviços e infraestrutura adequada para os clientes. Acompanhando essa demanda, verifica-se um crescimento nas dimensões e quantidade de itens nos estabelecimentos de comércio, especialmente supermercados, o que fatalmente provoca um aumento no tempo necessário para que um

cliente consiga encontrar todos os itens que deseja e concluir a sua compra.

Em um ritmo de vida conturbada em que vivem os cidadãos das grandes cidades, economizar tempo nas compras de supermercado com ferramentas adequadas seria com certeza um parâmetro relevante para escolha do fornecedor. Os serviços que tentam entregar esse tipo de valor aos clientes são os mais variados e, muitas vezes, quando não é feita uma análise criteriosa do que é relevante, se verifica serviços subutilizados e recursos mal empregados.

Os principais serviços utilizados pela população estão disponíveis em ambientes fechados como *shopping centers*, cinemas, supermercados e lojas de varejo. Um reflexo deste comportamento pode ser comprovado na pesquisa sobre o uso de tempo das pessoas em ambientes abertos e fechados, cujos resultados são reproduzidos na [Figura 1](#). De acordo com o estudo, as pessoas gastam de 80% a 90% do tempo em ambiente fechados e outros números importantes revelam ainda que 70% das chamadas de celular e 80% das conexões de dados se originam nesses ambientes.



Figura 1 – Utilização do tempo das pessoas quanto ao ambiente

Fonte: Strategy Analytics

No contexto mencionado, as ferramentas de localização e orientação *indoor* se inserem como componentes relevantes para as aplicações móveis, uma vez que possibilitam a entrega de uma experiência mais customizada aos usuários. Acompanhando a evolução desse comportamento das pessoas, a implementação dessas ferramentas no contexto do varejo se faz relevante. Elas podem agregar serviços ao consumidor como forma de expandir o conhecimento sobre o seu comportamento no ambiente de compras. A partir daí, é possível oferecer uma experiência de compras mais personalizada. Assim, o lojista pode aumentar as taxas de conversão, transformando indicações oferecidas aos clientes em vendas.

1.1 Motivação

A motivação para este trabalho reside em questões importantes, ainda em aberto, com relação às demandas de localização no nível de ambientes fechados. Até que ponto conseguimos replicar a precisão dos sistemas de GPS para esses ambientes? Com qual grau de exatidão e precisão podemos alcançar com uma arquitetura de baixo custo e baixo consumo de energia?

O problema de localização e orientação *indoor* é atual e possui grande relevância no meio acadêmico e comercial. Utilizando a ferramenta de registro de patentes ([GOOGLE, 2017](#)) pode ser constatado que, de Janeiro de 2010 a Junho de 2017 existem mais de 120 registros que correspondem a busca pela expressão *indoor location*.

Fundado por 4 grandes empresas em Agosto de 2012, atualmente contando com 5 membros colaboradores, e mais de 20 membros associados, o consórcio de empresas *In Location Alliance* ([ALLIANCE, 2017](#)) reúne grandes empresas do mercado de tecnologia. O consórcio é uma aliança estratégica para acelerar a adoção de soluções de posicionamento em ambientes fechados, com o objetivo de disponibilizar um novo padrão de experiência aos usuários, abrindo novas possibilidades para consumidores e novas oportunidades para lojistas.

Os carrinhos de compras inteligentes já são uma realidade atualmente no mundo e ainda estão pouco explorados no Brasil. As aplicações em evidência neste arcabouço são: leitor de código de barras, conteúdo (receitas, informações nutricionais), integração com dispositivos do cliente (*smartphones, tablets*) e sistemas de navegação. Esse último de grande relevância para clientes que desejam ganhar tempo nas suas compras, uma vez que auxiliam na localização de produtos desejados por meio de rotas de compra.

A importância da implementação de ferramentas baseadas em localização *indoor* se fundamenta, além da experiência de compras personalizada para auxiliar o cliente de varejo, na construção de bases de dados robustas sobre o comportamento dos clientes, de modo similar ao que já existe nos ambientes de comércio virtual, mas ainda pouco explorado no comércio tradicional. O setor de varejo, especialmente o setor de supermercados, é um dos domínios importantes no qual se pode aplicar relevantes soluções de Internet das Coisas ([AL-FUQAHA et al., 2015](#)), conforme visão geral da IoT (*Internet Of Things*), apresentada na [Figura 2](#). Nesta visão, os autores propõem uma abordagem com a separação entre os mercados verticais, com aplicações de domínio específico, e uma integração horizontal, com os serviços independentes do domínio. O fluxo de informações ocorre tanto no contexto intra-domínio (verticais) quanto no contexto amplo, por meio dos serviços independentes de domínios.



Figura 2 – Visão geral da IoT - Mercados verticais e integração horizontal

Fonte: Adaptado de [Al-Fuqaha et al. \(2015\)](#)

1.2 Definição do Problema

O problema tratado neste trabalho pode ser encontrado na literatura como localização *indoor*, porém também referenciado como posicionamento *indoor*, ou posicionamento em ambientes fechados, entre outras variações. A questão consiste em definir a localização, no caso deste trabalho em duas dimensões, de um dispositivo em relação ao ambiente em que ele se encontra.

A partir da planta baixa do estabelecimento, contendo as medidas do ambiente, e das informações sobre a posição e dimensões do mobiliário, um modelo gráfico de representação da área de interesse deve ser gerado. Esse modelo deverá ser utilizado para exibir a posição relativa do objeto alvo. Essa tecnologia de localização *indoor* pode ser aplicada para rastreamento de ativos, micro-navegação assistida, orientação de rotas, localização de pontos de interesse, orientação de deficientes visuais.

Essa funcionalidade deve estar disponível em todo o ambiente local, não apenas nas proximidades dos *beacons*, instrumentos utilizados para estimar a distância de um dispositivo compatível. Além disso, deve ser capaz de definir a localização de todos os dispositivos compatíveis que estiverem na área de interesse, utilizando tecnologias sem fio.

1.3 Objetivos

Esse trabalho propõe o desenvolvimento de tecnologias de localização *indoor* para suportar aplicações que podem beneficiar os principais elementos da cadeia varejista:

- a. Consumidor: propiciando ferramentas que facilitam suas compras e sua experiência nos ambientes de comércio varejista;
- b. Lojista: fornecendo um conjunto de dados valiosos sobre o comportamento dos consumidores que frequentam seu estabelecimento para que se possa, utilizando, por exemplo, técnicas como a mineração de dados, traçar o perfil comportamental do seu cliente e utilizar esta definição para entregar campanhas de *marketing* personalizadas e assertivas para cada consumidor, de modo similar ao que já se verifica atualmente no *e-commerce*;
- c. Fabricante: levantando mapas sobre a interação dos consumidores de cada estabelecimento com os seus produtos, de modo que possam direcionar melhor suas campanhas de incentivo e ações de *marketing* e capacidade de medir as taxas de conversão para suas campanhas junto ao varejo.

O objetivo geral deste trabalho é avaliar técnicas de localização de pessoas ou objetos em ambientes internos, assim como a comparação entre essas técnicas e avaliação dos cenários onde melhor se aplicam, especialmente no âmbito do comércio varejista nacional. Para alcançar o objetivo geral foram traçados os seguintes objetivos específicos:

1. Construir um modelo de representação computacional a partir de um ambiente físico real;
2. Coletar os dados de sensoriamento do ambiente utilizando *beacons* de baixo custo e um dispositivo leitor que já está disponível no ambiente;
3. Aplicar e avaliar técnicas para resolução do problema de localização *indoor* utilizando abordagem *Fingerprinting* e Triangulação;
4. Comparar objetivamente as técnicas e os cenários de aplicação.

1.4 Organização do Trabalho

A partir deste capítulo de introdução, os próximos capítulos deste trabalho estão organizados como se segue: No [Capítulo 2](#) um estudo da literatura é retratado com os conceitos essenciais para o bom desenvolvimento e entendimento desta pesquisa. No [Capítulo 3](#), alguns dos relevantes trabalhos relacionados a esta pesquisa são apresentados ao leitor. No [Capítulo 4](#), a metodologia com o detalhamento de como o trabalho foi conduzido para alcançar os objetivos propostos. Em seguida, no [Capítulo 5](#), é realizada uma análise para validação dos dados coletados dos *beacons* no primeiro experimento do trabalho e o desen-

volvimento do modelo é apresentado no [Capítulo 6](#). Os resultados são exibidos e discutidos no [Capítulo 7](#) e, finalmente, o [Capítulo 8](#) trata das considerações finais, contribuições e trabalhos futuros.

2 Referencial Teórico

Neste capítulo, um estudo da literatura de referência será apresentado com os temas essenciais para o bom desenvolvimento desta pesquisa.

2.1 Internet das Coisas

As pessoas já podem controlar alguns subsistemas domésticos a partir do *smartphone*, acender e apagar luzes, acessar câmeras de segurança, controlar o ar condicionado, entre outras funções. Essas obrigações podem ser realizadas por meio de sistemas automatizados, sem a necessidade das pessoas se preocuparem com essas tarefas. São os equipamentos reagindo à presença das pessoas, abordagem importante da Computação Ubíqua. Frequentemente tratada como sinônimo de Computação Pervasiva, Internet das Coisas, entre outros, é utilizada para revelar a onipresença da computação no cotidiano das pessoas. É quando a computação sai dos computadores pessoais e vai parar nos objetos e locais do dia a dia, como eletrodomésticos, sistemas de refrigeração e iluminação, meios de transporte, entre outros, como forma de prover serviços mais transparentes para os usuários (LOUREIRO et al., 2009).

O tema pode parecer muito recente, mas suas primícias nos remetem à década de 90 quando o cientista da computação e visionário Sr. Mark Weiser, então cientista chefe do XEROX PARC, profetizou com um brilhantismo incrível a nova grande onda da computação:

A computação ubíqua é a terceira onda da computação, que está apenas começando. Primeiro tivemos os mainframes compartilhados por várias pessoas. Estamos na era da computação pessoal com pessoas e máquinas estranhando umas às outras. A seguir vem a computação ubíqua a era da tecnologia 'calma', quando a tecnologia recua para o pano de fundo de nossas vidas. (WEISER, 1991, p. 2).

Este visionário cientista, que é considerado o pai da computação ubíqua, mostrou o caminho a ser seguido pelas tecnologias do novo século em outra célebre afirmação: Weiser (1991, p. 1): "As tecnologias mais profundas são aquelas que desaparecem."

Quando foi criado esse conceito, não existiam tecnologias suficientes para torna-lo realidade na vida das pessoas. Hoje temos muito consolidada a Internet das Coisas (do inglês: IoT - *Internet Of Things*) que desempenha um papel importante nesse contexto. Também chamada de Internet de todas as coisas (IoE: *Internet of Everything*), reúne um conjunto de dispositivos, softwares e pessoas para disponibilizar soluções inovadoras e tecnologias disruptivas, apresentado na Figura 3.

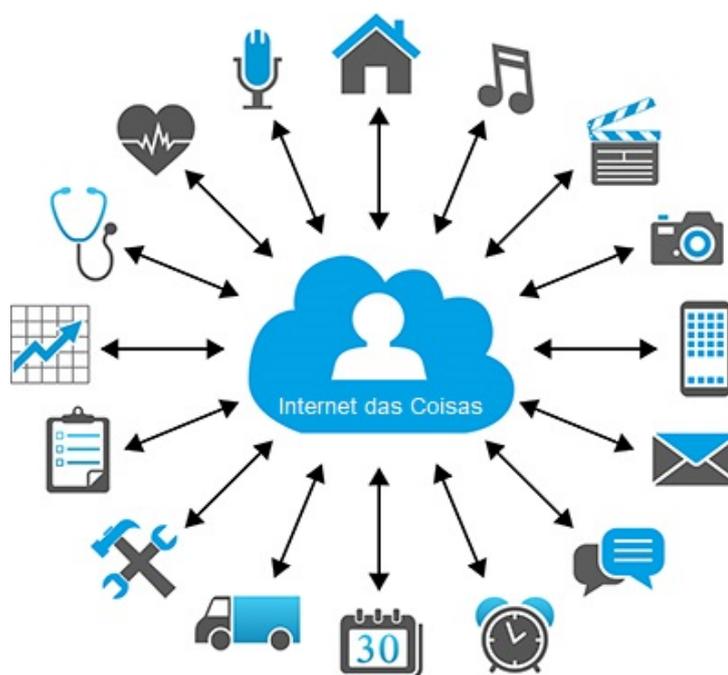


Figura 3 – Arquitetura Internet das Coisas

Fonte: Adaptado de [Gaikwad, Gabhane e Golait \(2015\)](#)

A base dessa tecnologia é que dispositivos estarão conectados à Internet e a outros dispositivos e poderão trocar dados entre si, de modo ubíquo, e desta troca vem a possibilidade de se cruzar dados e extrair informações importantes das pessoas, do ambiente e da sociedade. Conforme pesquisa da Cisco ([COMPTON; BUTANEY, 2015](#)), já alcançamos a expressiva marca de mais dispositivos conectados do que pessoas. A mesma pesquisa estima ainda que teremos 50 Bilhões de dispositivos conectados em 2020, com uma população de 7.6 Bilhões de pessoas.

Essa grande população de dispositivos conectados, gerando dados por segundo numa proporção nunca antes vista, gera uma grande oportunidade para uma nova profissão do século 21, o cientista de dados. Entre as abordagens mais importantes para se extrair conhecimento desta grande massa de dados, destacam-se as Redes Neurais Artificiais, *Deep Learning*, *Hadoop*, *Streaming*, entre outros. Alguns autores já citam o *Big Data* como uma nova fronteira da inovação e produtividade ([MANYIKA et al., 2011](#)).

Este tipo de computação está profundamente relacionado às características do mundo físico ([LOUREIRO et al., 2009](#)), uma vez que para prover serviços automatizados de maneira precisa, as aplicações necessitam acesso a informações de contexto como localização, temperatura, quantidade de pessoas em um ambiente, entre outras. Os sensores podem captar, de modo autônomo, dados importantes do meio, tais como: temperatura, luminosidade, umidade, nível de ruído, velocidade, níveis de poluição, presença ou quantidade de pessoas ou objetos, além de proporcionar informações de contexto para possibilitar definir

localização do dispositivo.

2.2 Bluetooth

O protocolo *Bluetooth*, especificado no IEEE 802.15.1 (WPAN/*Bluetooth*), é um padrão de rede sem fio por meio de ondas de rádio, de baixo consumo energético e de curto alcance utilizado em uma ampla gama de dispositivos eletrônicos (SIG, 2018). O nome *Bluetooth* provém do rei da Dinamarca chamado *Harald Bluetooth* que foi responsável pela unificação das tribos dinamarquesas no século X. O protocolo *Bluetooth* foi inicialmente proposto pela empresa suíça de telecomunicação Ericsson com o intuito de, seguindo os feitos do rei Harald, promover a integração de vários dispositivos sem a necessidade de comunicação cabeada. O raio de operação desse protocolo pode chegar a até 100m, quando utilizando a classe 1 (100mW - 20 dBm).

Em 1998, com o foco voltado para difundir e ampliar a utilização do padrão, foi criado o *Bluetooth Special Interest Group* (SIG), formado por gigantes da tecnologia como Nokia, Toshiba, IBM e Intel. Posteriormente, em 1999, outras organizações ingressaram neste grupo (Lucent, Microsoft e Motorola) e a cada ano o número de membros foi aumentando e hoje já conta com mais de 33.000 membros (SIG, 2018).

O *Bluetooth* opera em uma faixa de frequências não licenciada industrial, científica e médica, em inglês ISM (*Industrial, Scientific, and Medical*), na frequência de 2.4GHz (SIG, 2018). Como existem muitos dispositivos que operam na faixa dos 2.4GHz, o *Bluetooth* adota o sistema chamado AFH, do inglês *Adaptive Frequency Hopping*, que consiste em realizar saltos regulares de frequências em torno da faixa de trabalho.

Esses saltos regulares ocorrem entre os diversos canais distribuídos na faixa de frequência. Assim o sistema garante uma reduzida interferência na comunicação por meio deste protocolo. No caso da versão *Classic* (ER/EDR) são 79 canais distribuídos no espectro entre os 2.402 e 2.480GHz com espaçamento de 1 MHz (SIG, 2018).

A partir da versão 4.0, o protocolo *Bluetooth* passou a incluir o padrão *Low Energy* (*Bluetooth Smart*), em uma das atualizações mais importantes. A versão atual 5.0 foi lançada em 2016, com o dobro da velocidade da versão anterior e com melhorias que elevam a capacidade de tráfego de dados em até 800%. O *Bluetooth* se tornou uma solução de conectividade de nível industrial que poderá ser o padrão de comunicação sem fio para IoT nas próximas décadas (SIG, 2018).

2.2.1 Bluetooth Low Energy

A variação de baixo consumo de energia e baixa taxa de transmissão do *Bluetooth*, conhecida como *Smart* ou BLE, foi especificada a partir da versão 4.0, disponibilizada no ano de 2010 (SIG, 2018). O primeiro *smartphone* que implementou esse padrão foi o iPhone 4S (PACHAL, 2011). A inclusão desse padrão viabilizou a criação de pequenos sensores alimentados por pequenas baterias do tipo botão (*coin-cell*) com ciclo de vida de meses ou até anos, como é o caso dos *beacons* (RAZA et al., 2017) que serão abordados a seguir.

A versão 4.0 também adicionou o modo de transmissão *broadcast*, além do já existente modo de emparelhamento orientado por conexão (SIG, 2018). Esse modo foi o principal facilitador para operação dos *beacons*. A principal melhoria que habilitou os recursos da IoT no *Bluetooth* foi a introdução do suporte ao protocolo de Internet IPSP (*Internet Protocol Support Profile*), que disponibiliza conectividade IPv6 no BLE (RAZA et al., 2017). Assim, os dispositivos *Bluetooth Smart* passaram a oferecer suporte ao 6LoWPAN (*IPv6 over Low power Wireless Personal Area Networks*). O 6LoWPAN é um padrão especificado com o objetivo de utilizar o protocolo IPv6 em plataformas de baixo consumo energético.

Com a inclusão deste novo padrão, o *Bluetooth* passou a contar com as famílias *Classic* (BR/EDR) e *Smart* (*Single Mode ou BLE*). Como são fundamentalmente diferentes, não existe possibilidade de comunicação direta entre os dois padrões. Por isso foi criado um terceiro modo para prover a interoperabilidade entre os padrões existentes, o *dual Smart Ready* (BR/EDR/LE ou *dual mode*). Dessa forma os dispositivos mais novos, que saíram de fábrica com o padrão *Smart Ready*, podem estabelecer conexão com todos os dispositivos compatíveis com um dos 3 padrões. Na Figura 4 podem ser visualizadas as configurações entre as versões e os tipos de dispositivos.

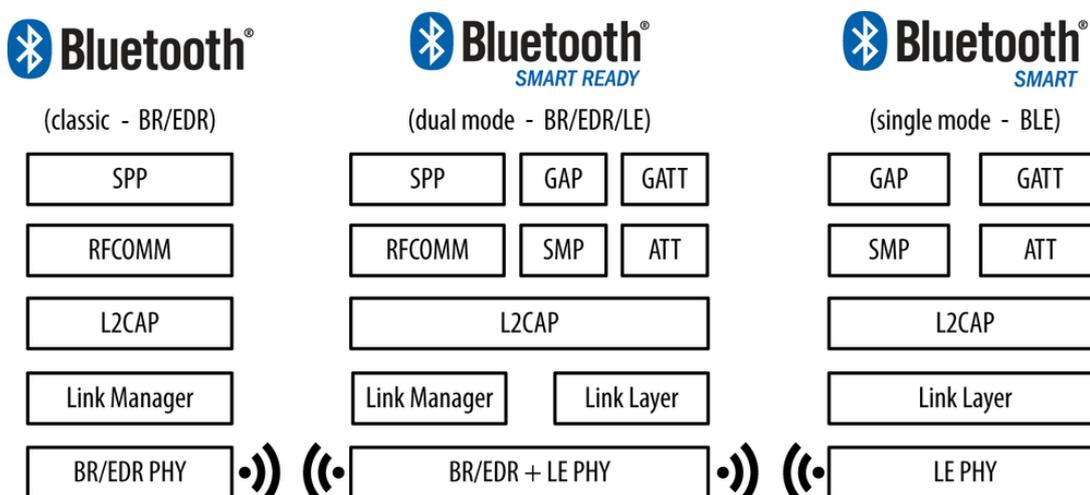


Figura 4 – Pilha de protocolos *Bluetooth*

Fonte: Adaptado de SIG (2018)

O BLE habilita a transmissão de dados em dois modos, o primeiro ocorre com os dispositivos conectados e o segundo é o modo *advertising (broadcast)*. O modo conectado utiliza a camada *Generic Attribute (GATT)*, uma transferência de dados up-para-um. Um importante exemplo de aplicação deste modo ocorre nos aplicativos de configuração dos beacons, onde o usuário pode selecionar os parâmetros mais relevantes para a sua aplicação. Já o modo *advertising* utiliza a camada *Generic Access Profile (GAP)*, transmitindo dados de um-para-muitos. Este é o modo padrão de operação dos beacons para transmissão de dados.

A versão *Smart* é otimizada para transmissões em curto período de tempo e pode utilizar as topologias de rede *Mesh*, *Broadcast* e *Point-to-Point* (incluindo piconets) (SIG, 2018). Ela é dividida em 40 canais de 2MHz, distribuídos entre 2.402 a 2.480GHz, numerados de 0 a 39, sendo que os canais 37, 38, e 39 são utilizados apenas para enviar pacotes de *advertisement*. Durante o envio desse tipo de pacote, o periférico BLE transmite nos 3 canais, um após outro. E um dispositivo cliente que está escaneando os sinais vai ouvir estes canais específicos para descobrir dispositivos próximos.

Tais canais BLE são estrategicamente posicionados, em relação aos canais mais populares e utilizados do *Wi-Fi*, conforme segue: antes do primeiro (37), entre o canal 1 e 6 do *Wi-Fi* (38) e no final, após o canal 11 (39) do espectro, conforme destacado de vermelho na Figura 5. Esse amplo espaçamento entre os canais de *advertisement* ajuda o BLE a gerenciar melhor a interferência do *Wi-Fi*, que compartilha a mesma faixa de frequência (ARGENOX, 2018).

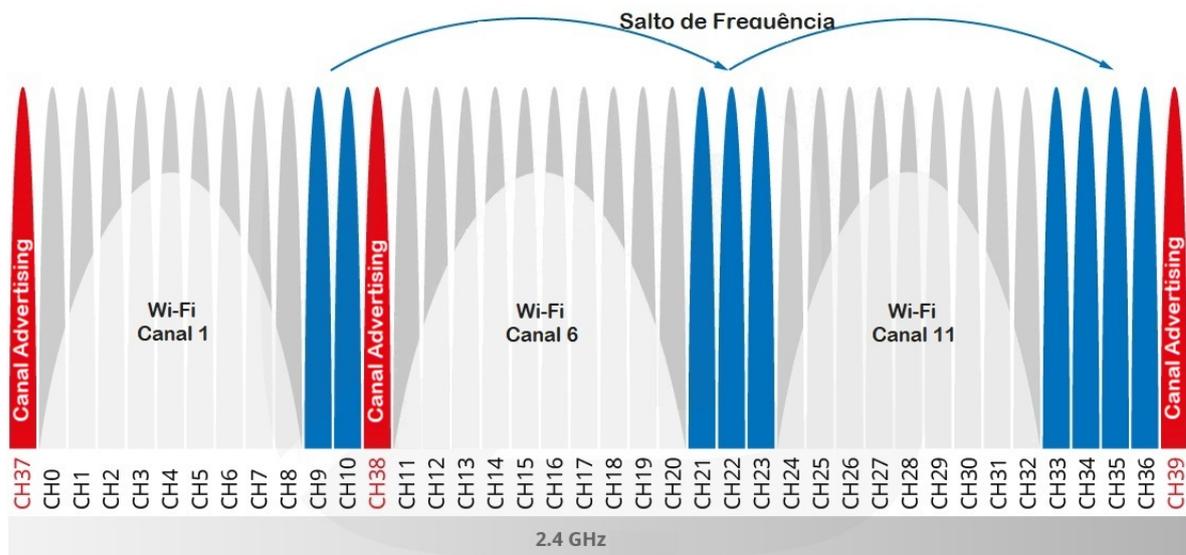


Figura 5 – Canais e Frequências do Espectro BLE

Fonte: Adaptado de Argenox (2018)

A grande vantagem na utilização dessa tecnologia para localização *indoor* consiste no seu suporte a dispositivos móveis contemporâneos, em especial os *tablets* e *smartphones*. Atualmente a grande maioria dos dispositivos já possuem suporte ao BLE, visto que a versão 4.0 foi lançada em 2010, assim como o primeiro *smartphone* com suporte à nova tecnologia. O advento dos *beacons* a partir de 2013, ajudou a impulsionar o lançamento de novos dispositivos com suporte ao *Bluetooth* 4.0, de acordo com o relatório *Bluetooth Market Update* de 2018 (SIG, 2018).

2.3 Beacons BLE

A partir do lançamento da versão 4.0 do *Bluetooth*, foi incorporado o padrão *Low Energy*. Uma variação otimizada para transmissões em curto período de tempo, com baixo consumo de energia e baixa taxa de transmissão, que possibilita um ciclo de vida de meses, ou até anos, para os dispositivos. Essa abordagem viabilizou a criação de sensores alimentados por pequenas baterias como os *beacons* (RAZA et al., 2017).

O primeiro *beacon* BLE foi lançado em Setembro de 2013 pela empresa Estimote (NEWMAN, 2014). Esse *hardware* representa uma nova geração de dispositivos de baixo custo que estão tornando possível aplicações relevantes de localização *indoor*, permitindo que varejistas possam rastrear cada passo dos seus clientes por meio de aplicativos móveis instalados nos *smartphones* de cada consumidor (NEWMAN, 2014).

Os *beacons* são pequenos dispositivos, quase imperceptíveis, que se comunicam via ondas de rádio, utilizando a tecnologia *Bluetooth Low Energy*, e podem trocar informações com dispositivos compatíveis com essa tecnologia a uma distância de até 100 metros (*Bluetooth* classe 1). Esses dispositivos podem ter especificações distintas, dependendo da necessidade de cada projeto ou orçamento disponível, desde os mais simples somente com *Bluetooth* e outras opções com funções extras como sensores de temperatura e acelerômetro.

Eles servem basicamente para transmitir dados em *broadcast*, incluindo identificadores únicos, que podem ser coletados por dispositivos móveis na sua área de abrangência. As aplicações que coletam estes dados utilizam como informação de contexto para definição de proximidade do dispositivo em relação ao ponto onde está afixado o *beacon*, ou para disparar ações específicas dado um evento de entrada (ou saída) do dispositivo na área de cobertura.

Tais dispositivos não coletam nenhum dado do usuário pois nem sequer possuem esta capacidade. Porém, os aplicativos instalados nos dispositivos dos usuários utilizam o sinal enviado pelos *beacons* para estimar a localização do usuário. Conseqüentemente, as questões referentes à privacidade do usuário dependem do tipo de aplicativo e dos seus

termos específicos de aceitação, além do consentimento do usuário com relação às regras e aos dados coletados (NEWMAN, 2014).

Os principais fabricantes mundiais são *Estimote, Kontakt, BlueCats, BlueSense, Gelo, Gimbal, Glimworm*, entre outros (GROOT, 2014), além de *Radius Networks, EM Microelectronic, Minew, BlueUp, Sensoro e Accent Systems*, de acordo com Brachmann (2018). As opções nacionais ainda são limitadas, sendo *DevBeacon* (utilizado neste trabalho) e *Taggen* as principais.

Atualmente alguns autores, ainda que em poucos casos interessados diretamente na promoção da tecnologia, acreditam que o futuro dos *beacons* se funde ao futuro da Internet das Coisas (AUGUR, 2017) e visualizam os *beacons* como os blocos construtores da IoT (NOWODZINSKI; ŁUKASIK; PUTO, 2016).

2.3.1 Principais Aplicações

Os *beacons* são frequentemente utilizados para as empresas iniciarem no mundo da Internet das Coisas, sendo empregados nas aplicações de sistemas de localização em tempo real (RTLS, do inglês, Real Time Location System), navegação, automação industrial, *marketing* contextual, análise comportamental de clientes em compras, entre outras.

O uso dos *beacons* nos estabelecimentos varejistas é uma realidade muito explorada para melhorar a experiência de compras dos clientes em vários países desenvolvidos, enquanto que, no Brasil, a tecnologia está presente em poucos projetos e ainda de forma inexpressiva, conseqüentemente existe um grande potencial a ser explorado. As principais aplicações, não se limitando a esta lista, podem ser:

- a. Autenticação e automação;
- b. *Marketing* baseado em localização;
- c. Acompanhamento e movimentação de objetos e pessoas em ambientes em tempo real (rotas de compras, entre outros);
- d. Exibição de conteúdo personalizado (entreterimento, turismo, museus, entre outros);
- e. Localização assertiva de pessoas e objetos em ambientes (instruções para encontrar produtos em lojas, entre outros);
- f. Em redes de sensores sem fios (WSN - *Wireless Sensor Networks*) para resolução do problema da localização dos nós sensores.

Os *beacons* trabalham em conjunto com dispositivos de comunicação com padrão *Bluetooth* utilizando *push notifications*. Para a aplicação mais clássica, publicidade, por exemplo, é necessária a utilização de pelo menos um *beacon*, e o aparelho do cliente (tablet, *smarthphone*, entre outros) precisa ter capacidade de comunicação via *Bluetooth*, que precisa estar ativado

e o cliente deve ter o aplicativo da loja ou de um parceiro instalado.

No *e-commerce* existem várias ferramentas (*cookies*, integração com redes sociais, histórico de navegação e compras, entre outros) que possibilitam rastrear o comportamento do cliente para traçar seu perfil e conseguir oferecer sugestões e promoções mais assertivas para alcançar um incremento nas vendas. Tais ferramentas não estão disponíveis em um comércio tradicional (nas lojas). Nesse contexto, existe um problema relevante que os *beacons* podem resolver com eficácia, que possibilita trazer para o ambiente do comércio físico, de modo automatizado, informações relevantes que proporcionam estimar o comportamento do consumidor.

A partir de um aplicativo instalado no *smartphone* do cliente e da sua comunicação com os *beacons* é possível habilitar a utilização de métricas importantes para que o lojista proporcione uma experiência totalmente personalizada para o seu cliente, destacando-se:

1. Nome do cliente que está entrando na loja;
2. Quais os últimos itens ele comprou e quando foi a última compra;
3. Quais departamentos e lojas os clientes passam mais tempo;
4. Quais os produtos ele possui maior interesse;

2.3.2 Interação com o Mundo Físico

Os *beacons* também são utilizados para habilitar a *web* física, que pode ser vista como uma extensão da *web* existente, com objetivo de viabilizar a interação com objetos físicos. Segundo [Sneps-Sneppe e Namiot \(2016\)](#), trata-se de um termo genérico para descrever a interconexão entre objetos e a *web*.

A *web* física é uma abordagem aberta que habilita interações rápidas e contínuas com objetos e locais físicos ([GOOGLE, 2018b](#)), que pode ser aplicada com diferentes objetivos. Qualquer objeto pode transmitir conteúdo, como um endereço *web* (*frame Eddystone-URL*), por exemplo. Isso pode gerar uma notificação e direcionar automaticamente os usuários, que estão na área de abrangência, para um conteúdo relacionado ao contexto, e o usuário opta ou não por entrar e interagir.

2.3.3 Tipos de *Frames*

Os tipos de *frames* são formatos de dados definidos para utilização da estrutura do pacote *advertising*, que é enviado nos três canais (37, 38 e 39) do BLE, conforme [Figura 5](#). A estrutura do pacote BLE não é alterada, apenas o formato e os dados enviados pelos dispositivos que podem variar. Assim, cada fabricante pode incorporar seus próprios formatos e dados ou utilizar um padrão já existente. Este formato é uma especificação da camada *Generic*

Attributes (GATT) do BLE que define uma estrutura de dados hierárquica com o objetivo de manter a interoperabilidade total com vários dispositivos *Bluetooth* (SIG, 2018).

O *iBeacon* foi o primeiro tipo de *frame* disponibilizado no mercado, lançado em 2013. É um padrão fechado, desenvolvido pela Apple. As principais informações transmitidas nesse formato se referem à três níveis de identificação contextual, que são detalhadas na Tabela 1 (APPLE, 2014).

Tabela 1 – Principais campos do *frame iBeacon*

Campo	Descrição
UUID	Identificador universal (<i>Universally Unique Identifier</i>). 16 bytes. Hexadecimal.
Major	Identificador de 2 bytes que trabalha como um sub identificador do UUID.
Minor	Similar ao Major, utilizado para mais um nível de detalhamento após o Major.

Apesar de o *frame iBeacon* funcionar tanto com dispositivos Apple quanto dispositivos Android, ele apresenta instabilidade quando utilizado no Android e não existe possibilidade de melhorias pela comunidade por ser um padrão fechado. Como forma de vencer essas limitações, em julho de 2015, a Google lançou um novo protocolo chamado *Eddystone* que pode ser utilizado em qualquer uma das plataformas (COSTA, 2015). O protocolo é de código aberto e está disponível na plataforma github.

Não existe apenas um, mas vários *frames Eddystone* (UID, EID, URL, TLM) que suportam diferentes formatos de dados. O UID é similar ao *iBeacon*, sendo utilizado para transmitir um identificador estático de 10 bytes e um código de instância que serve como um nível a mais de detalhamento. O tipo URL é utilizado para transmitir uma url que pode ser utilizada diretamente pelo APP cliente para disponibilizar um conteúdo *web*, por exemplo. O TLM (*Telemetry*) pode ser utilizado para enviar dados sobre *status* do *beacon*, como carga da bateria e tempo de atividade, que pode ser muito útil para o controle de manutenção dos dispositivos. O *frame EID* (*Eddystone Ephemeral Identifier*) é similar ao UID, com a diferença que os identificadores são criptografados e mudam a cada pequeno intervalo de tempo, de modo que apenas o aplicativo autorizado pode utilizá-lo (GOOGLE, 2018a).

2.3.4 Distância e Localização

Os *beacons* foram inicialmente concebidos para comunicação e publicidade, demais finalidades como localização *indoor* são oportunistas (FARAGHER; HARLE, 2015). A documentação da Apple sobre *iBeacons*, (APPLE, 2014), enfatiza uma limitação em relação a utilização dos *beacons* BLE para determinar distância. Ela limita a resposta em 4 zonas de proximidade, conforme Tabela 2. Esses conceitos se aplicam de forma análoga ao funcionamento dos *beacons* com *frames Eddystone*.

Tabela 2 – Classificação dos estados de proximidade dos *Beacons*

Status	Proximidade
Muito Próximo	Dispositivo muito próximo do <i>beacon</i>
Perto	Existe uma linha de sinal transparente entre os dispositivos. 1 a 3 metros.
Longe	<i>Beacon</i> pode ser detectado, porém com baixa confiança na exatidão para definir se está em um dos níveis mais precisos (perto, muito próximo)
Desconhecido	A proximidade não pode ser determinada. Não possui medições suficientes ou não entrou na área de cobertura

Entretanto muitas pesquisas são desenvolvidas, conforme [Seção 3.3](#), no sentido de alcançar maior acurácia para viabilizar a utilização desses dispositivos para definição da localização aproximada em relação a um ambiente. Essas pesquisas buscam mostrar que os *beacons* podem funcionar também como GPSs para ambientes fechados, podendo alcançar precisão de poucos metros ou até centímetros em alguns casos. Trabalhando em conjunto com sistemas de navegação *indoor*, disponibilizam a localização de objetos, distâncias ou criação de perímetros virtuais.

Os *beacons* são estrategicamente posicionados em locais de interesse e baseado na proximidade do usuário com cada um dos *beacons* é feita a triangulação e assim definida a localização do usuário. A partir dessa localização, quando o usuário está dentro do geofence definido, são disparadas ações programadas que podem ser simples envios de notificações ou interações mais elaboradas.

2.4 Localização *Indoor*

O principal objetivo de localização *indoor* é conhecer a posição de pessoas ou objetos em um ambiente fechado para que esta informação de contexto possa ser utilizada em aplicações LBS (*Location-Based Services*). O termo posição neste sentido expressa a capacidade de localizar a posição física em um espaço conhecido previamente ([HERMERSDORF, 2007](#)).

2.4.1 Localização vs Proximidade

É válido destacar que os conceitos de localização e proximidade, apesar de diferentes, são muitas vezes confundidos. A localização está relacionada com uma posição exata, que possui um vetor de atributos, incluindo a direção ([GAST, 2014](#)). É um parâmetro absoluto em um contexto, geralmente definido por um sistema de coordenadas, caso de um sistema cartesiano, onde a definição da localização de um ponto p é dada por $p = (x, y)$.

Já a utilização do termo proximidade, de modo contrastante, está relacionada a uma medida de distância em relação a algum ponto base, por exemplo, qual distância uma pessoa está de um objeto que possui posição $p = (x, y)$. Nesse caso não é possível definir sua posição absoluta, apenas a distância relativa, sem conhecimento da direção.

2.4.2 Serviços Baseados em Localização

Os serviços baseados em localização (LBS) são serviços que se baseiam na localização do usuário como pré-requisito mais importante para disponibilizar suas principais funcionalidades. Possuem vasta possibilidade de aplicação, como turismo, entretenimento, *smart cities*, *marketing*, entre outros. No ambiente de varejo, as novas tecnologias de localização *indoor* possibilitam um novo paradigma na interação entre os clientes e o estabelecimento e por conseguinte, uma nova oportunidade para redefinir a estratégia e atualizar o modelo de negócios (NOWODZINSKI; ŁUKASIK; PUTO, 2016).

Uma vasta gama de aplicações de *marketing* personalizado e serviços móveis e de auto-atendimento podem oferecer suporte para uma nova era no comércio. É uma oportunidade para que as lojas possam se reinventar. Os serviços vão desde uma simples informação, como envio de indicação de direção para os produtos que o cliente deseja encontrar, até serviços baseados em inteligência artificial para indicações de produtos de acordo com a posição dentro da loja, o histórico de compras e lista de compra automática baseada no perfil de consumo do cliente.

2.4.3 Tecnologias de Medição Utilizadas

O GPS é uma tecnologia consolidada e amplamente utilizada para navegação e localização *outdoor* em toda a superfície terrestre. Entretanto no âmbito dos ambientes internos, com a área em desenvolvimento, ainda não existe uma tecnologia consolidada. Ao contrário, existe uma grande variedade delas e muitos cenários de aplicação que podem impactar diretamente na definição da melhor opção a ser implementada. A seguir, serão apresentadas essas tecnologias, agrupadas por tipo de comunicação.

O primeiro grupo de tecnologias utilizadas é o GPS. O sinal GPS não está disponível na maioria dos ambientes fechados, entretanto, é possível a implementação de soluções de localização *indoor* utilizando essa tecnologia, com a utilização de estratégias que possam garantir a disponibilidade do sinal na área de interesse. Nirjon et al. (2014) implementam um sistema de localização *indoor* utilizando tecnologia GPS. Nesse trabalho é utilizada uma antena direcional de alto ganho como o *front-end* do dispositivo receptor, e, segundo os autores, alcançou resultados com mediana do erro inferior a 10m.

Radio-free é o segundo agrupamento de técnicas relevantes. Neste grupo de propostas

não se utiliza ondas de rádio para comunicação. Uma importante técnica desse grupo é a abordagem que utiliza o campo magnético da Terra. Ela tem seu funcionamento baseado nas flutuações do campo magnético no interior de edifícios, que ocorrem tanto de fontes naturais quanto de origem humana, como estruturas de aço, concreto, sistemas de energia, aparelhos eletro-eletrônicos entre outros. De acordo com [Haverinen e Kempainen \(2009\)](#), assumindo que as anomalias nesses ambientes são quase estáticas, e têm variabilidade local suficiente, é possível obter uma impressão digital magnética para implementação de uma abordagem *Fingerprinting* (veja [Subseção 2.4.5](#)). Outra opção possível é a técnica de Medições Inerciais que utiliza sensores para a medição de passos da pessoa.

O terceiro grupo se trata da utilização de som e luz. O ultrassom é uma tecnologia de propagação de som a uma frequência superior àquela que o ouvido do ser humano pode perceber. Existem sensores ultrassônicos que utilizam esta tecnologia. [Li et al. \(2016\)](#) calculam a posição do dispositivo receptor por meio da técnica de tempo de chegada (ToA, *Time of Arrival*) utilizando ultrassom. A transmissão de sinais infravermelho modulado por largura de pulso também pode ser utilizada como técnica de localização *indoor*. [Want et al. \(1992\)](#) propõem um serviço de localização centralizada no qual os funcionários de uma empresa utilizam um crachá que transmite sinais em infravermelho com informações sobre a sua localização, por meio de uma rede de sensores.

A utilização da Radiofrequência é outro importante grupo, onde estão a maioria das tecnologias utilizadas. As tecnologias incluem, sem se limitar a: Identificação por radiofrequência ([JIN; LU; PARK, 2006](#); [KHOLOD et al., 2010](#)), Redes *Wi-Fi* IEEE 802.11 ([HE; CHAN, 2016](#); [KRIZ; MALY; KOZEL, 2016](#)), *Bluetooth Low Energy* ([GAST, 2014](#); [LARSSON, 2015](#); [KRIZ; MALY; KOZEL, 2016](#)), *Ultra Wide Band* ([ALARIFI et al., 2016](#)). Na maioria das tecnologias de radio-frequência, as técnicas utilizadas são intensidade do sinal (*RSSI, Received Signal Strength Indication*), Ângulo de Recepção (*AoA, Angle of Arrival*), Tempo de Chegada (ToA, *Time of Arrival*).

2.4.4 Medição de Distância Baseada em RSSI

Um sinal de rádio transmitido a partir de uma antena é propagado através de um espaço ocorrendo perdas na intensidade até que o alvo seja alcançado ([CHO et al., 2015](#)). A distância de um *beacon* é calculada estimando a distância que um sinal deve percorrer para que ocorra uma determinada perda na intensidade que chega ao dispositivo receptor. Essa estimativa é feita utilizando algumas constantes de calibração, que variam de acordo com especificações de hardware, e características físicas do meio ([GAST, 2014](#)).

De acordo com [Young \(2013-2018\)](#), a função exponencial que modela a relação entre o RSSI e a distância é dada pela [Equação 1](#). Nessa função, a distância (d) é obtida a partir da intensidade do sinal medido (RSSI), um valor de referência do RSSI a uma distância de

1 metro ($RSSI_{Ref1m}$) e 3 coeficientes (a, b, c) que são dependentes de características do hardware utilizado e do meio de propagação do sinal.

$$d = a \times \left(\frac{RSSI}{RSSI_{Ref1m}} \right)^b + c. \quad (1)$$

Um dos grandes desafios da abordagem baseada em RSSI é a sensibilidade aos ruídos do canal utilizado para transmissão, como a interferência de sinal, atenuação e reflexão. Todos esses eventos geram impacto significativo na amplitude do sinal. As técnicas baseadas em RSSI também estão muito suscetíveis à variabilidade do transmissor, receptor e antena. Essas características de *hardware* impactam diretamente na precisão e exatidão dos cálculos. Por fim, características físicas do ambiente, como temperatura e umidade geralmente causam variabilidade na velocidade de propagação do sinal (YANG; LIU, 2010).

2.4.5 Localização por *Fingerprinting*

A técnica de localização *Fingerprinting* utiliza a base de dados, que é obtida por meio de medições sucessivas na fase *off-line*, para definir a posição do dispositivo na sua fase *on-line*. Esse *dataset* de aprendizado está diretamente relacionado com as características do meio físico e do hardware utilizado nos experimentos.

Na fase *off-line* são feitas medições para extrair e armazenar uma espécie de impressão digital de cada um dos pontos de interesse. Essa base de dados é armazenada e indexada para facilitar futuras buscas de padrões. A partir de uma nova entrada com as medições de potência de sinal, o sistema efetua uma busca na base armazenada e retorna a instância que mais se assemelha ao registro pesquisado. Essa instância possui o *label* do ponto onde foi feita a medição, finalizando assim a classificação da amostra.

Existem vários algoritmos de localização baseados em *Fingerprinting*, incluindo os tipos determinísticos e estocásticos. Nesse trabalho foram implementados os algoritmos KNN (do tipo determinístico) e MLP (do tipo estocástico) por serem os mais utilizados na literatura e que tem apresentado os melhores resultados.

2.4.6 Localização por Triangulação

A Triangulação é uma das principais técnicas de determinação da localização, e utiliza propriedades geométricas dos triângulos para calcular a posição de um objeto que se deseja localizar (alvo). Ela é dividida em 2 grupos: Angulação e Lateralização. Na primeira abordagem utiliza-se a medida dos ângulos e na segunda, a medida das distâncias entre os pontos âncora (pontos de referência cuja posição é conhecida) e o objeto alvo (YANG; LIU, 2010). Muitos dos trabalhos que aplicam essas técnicas estão nas áreas de Redes

de Sensores Sem Fio (RSSF), para encontrar as coordenadas dos nós de uma rede, ou localização *indoor/outdoor*, para definir a posição de objetos e pessoas num espaço físico. Neste trabalho, a abordagem utilizada será a Lateralização e por isso o foco desta seção será neste tema.

O método de multilateração estima a posição de um alvo utilizando a informação da distância entre o alvo e os n pontos âncora. O caso particular da multilateração mais conhecido é quando $n = 3$, a Trilateração. Na [Figura 6](#) é demonstrada um caso de cálculo da posição em duas dimensões, utilizando a Lateralização com 3 âncoras. A partir dos 3 pontos de referência A_1 , A_2 e A_3 , cada um com as distâncias obtidas até o alvo dadas por r_1 , r_2 e r_3 respectivamente, são obtidas as circunferências. O ponto P_1 no caso **a** ou a área de interseção dessas esferas dada pelos pontos P_1 , P_2 e P_3 , no caso **b**, são, respectivamente, o ponto e a área estimada para localização do alvo.

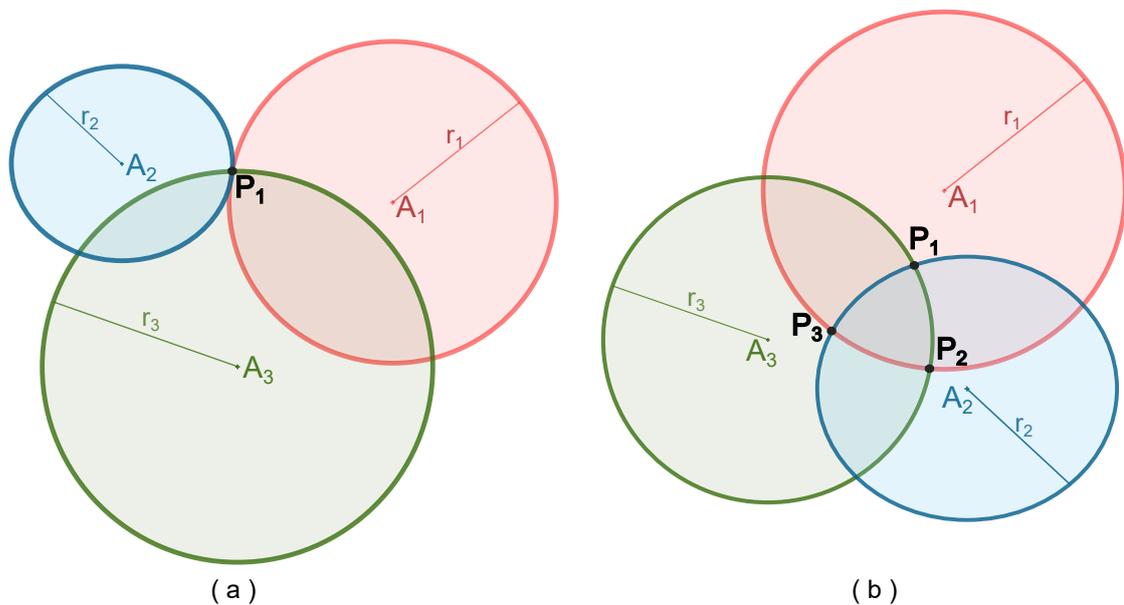


Figura 6 – Cálculo da posição em 2d utilizando a Trilateração

É possível obter a localização do alvo em um ambiente n -dimensional por intermédio de $n + 1$ pontos de referência. Para medições exatas de intensidade do sinal, sem interferências, seria possível obter a distância exata entre o dispositivo a ser localizado e os âncoras, caso em que é gerado um único ponto de interseção entre as 3 esferas, conforme item **a** da [Figura 6](#). Porém, na prática, as aplicações lidam com muitos ruídos nas leituras dos dados de RSSI. E, nesse caso, a resposta do algoritmo pode ser uma área, conforme item **b** da [Figura 6](#). A partir daí, alguma técnica pode ser utilizada para definir o ponto, geralmente o centro geométrico da área estimada. A localização calculada seria única nos dois casos, considerando que os pontos de referência não sejam lineares ([YANG; LIU, 2010](#)).

Cada um dos âncoras são os centros de circunferências com raios iguais à distância

estimada até o alvo. Utilizando técnicas simples de álgebra linear e trigonometria, é possível obter os pontos de interseção gerados pelos pares de circunferências. Na [Figura 7](#) está representado um exemplo com dois pontos de interseção. Existem outros dois casos possíveis, quando apenas um ponto é gerado, e o caso sem nenhum ponto de interseção entre as duas circunferências. A quantidade máxima de pontos de interseção (melhor caso) é dada por um arranjo simples $A_{n,p} = \frac{n!}{(n-p)!}$, onde n é a quantidade de pontos âncora utilizados e $p = 2$. A quantidade mínima de pontos é sempre zero (pior caso).

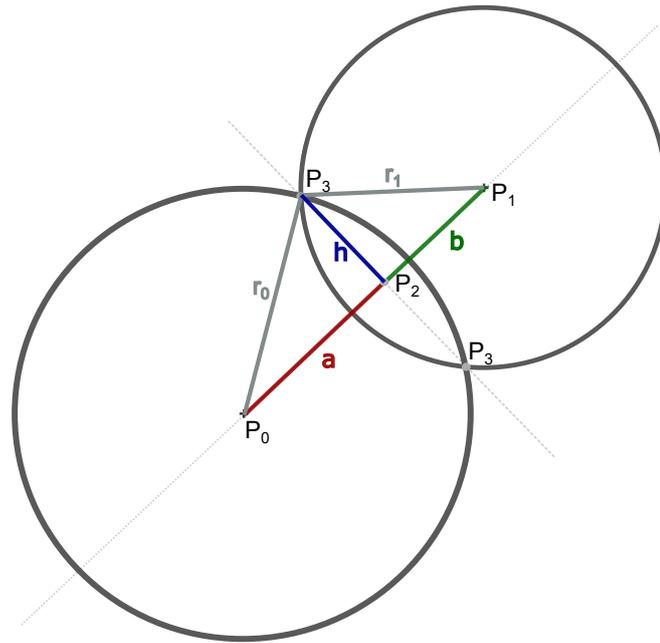


Figura 7 – Interseção entre duas circunferências. Adaptado de [Bourke \(1997\)](#)

2.5 Redes Neurais Artificiais

O estudo das Redes Neurais Artificiais (RNA) é motivado, desde o seu início, pelo reconhecimento do modo como funciona o cérebro humano que é diferente do computador digital convencional. O grande potencial do cérebro serve de fomento para vários pesquisadores que concentram esforços na tentativa de trazer a habilidade de aprendizagem, reconhecimento de padrões e resolução de problemas para os sistemas computacionais ([HAYKIN, 2009](#)).

Uma RNA é um processador distribuído e paralelo, composto de unidades de processamento simples, chamadas neurônios, projetada para armazenar conhecimento baseado em experiência e observação. Esse conhecimento é adquirido a partir de processos de aprendizagem, tal qual o cérebro biológico, e armazenado em estruturas conhecidas como pesos sinápticos, para consulta e resolução de problemas relacionados ([HAYKIN, 2009](#)).

As redes são treinadas utilizando dados reais do problema. Os projetistas buscam sempre pelas melhores técnicas e conjuntos de dados sobre o problema, de modo que a rede neural tenha boa capacidade de generalização para resolução de problemas relacionados. Tais práticas contribuem para que a rede gere uma saída próxima da esperada mesmo quando recebem entrada de dados diferente daquelas utilizadas no processo de aprendizagem.

As RNA oferecem propriedades e capacidades importantes. Entre elas podemos destacar: não linearidade, mapeamento entrada-saída, adaptabilidade, respostas comprobatórias, informação contextual, tolerância a erro, uniformidade de análise e projeto, analogia neurobiológica (HAYKIN, 2009). Essas capacidades de processamento tornam as RNA poderosas ferramentas para resolução de problemas complexos, de larga escala, que são intratáveis para outros tipos de sistemas.

2.5.1 Perceptron Multicamadas

Neste trabalho utilizamos uma RNA do tipo Perceptron Multicamadas, ou MLP (*Multilayer Perceptron*), para implementação de um classificador de padrões não linear. Esse classificador utiliza a base de registros de intensidade de sinal com os dados previamente normalizados para o algoritmo de aprendizagem.

As redes MLP utilizam n camadas de neurônios do tipo perceptron. O perceptron é o tipo mais simples de rede neural, proposto por Frank Rosenblatt em 1958, foi o primeiro modelo de aprendizagem supervisionada. Este modelo possuía apenas uma camada de neurônios com pesos sinápticos e vias ajustáveis. Seu objetivo é classificar, de modo binário, padrões que podem ser separados linearmente (HAYKIN, 2009).

A MLP é uma rede neural composta por várias camadas de neurônios. Possui uma camada de entrada, uma ou mais camadas escondidas, onde é executada a maior parte do processamento da rede e uma camada de saída, onde são concluídas as respostas da rede. Essas redes podem resolver problemas mais complexos, não lineares, eliminando a limitação que existia nas redes perceptron unitárias.

2.6 K Vizinhos Mais Próximos

O algoritmo K Vizinhos mais próximos (do inglês KNN: *K Nearest Neighbors*) é um método não paramétrico utilizado para classificação e regressão. Consiste em atribuir a um ponto não classificado a classe mais presente entre seus k vizinhos mais próximos, assumindo que instâncias similares possuem classificações iguais. Essa característica importante da proximidade faz com que o algoritmo dependa fortemente da métrica utilizada como distância (COVER; HART, 1967).

O KNN utiliza o conceito de pares em um espaço métrico onde a distância entre cada um dos pares está bem definida. E essa distância é utilizada para encontrar o conjunto dos k vizinhos mais próximos de uma instância. A classificação é definida com a classe que tem mais ocorrências nesse conjunto (COVER; HART, 1967).

3 Trabalhos Relacionados

Neste capítulo serão apresentados alguns dos trabalhos mais relevantes relacionados a esta pesquisa, divididos em três seções. A primeira apresenta trabalhos que buscam capturar informações sobre o comportamento do consumidor nos ambientes de compras. A segunda cita alguns trabalhos que apresentam soluções de tecnologia para apoiar os consumidores durante as compras, nos ambiente de varejo. A terceira, e mais importante, apresenta os trabalhos de localização *Indoor*. Finalmente, na última seção, são apresentadas as discussões finais sobre o estado da arte.

3.1 Comportamento do Consumidor

Nesta seção são apresentados os trabalhos que abordam a análise de comportamento do consumidor para identificação de relações entre estes hábitos e o perfil de compras em ambientes de comércio varejista.

Em [Kholod et al. \(2010\)](#), os autores desenvolvem uma abordagem de análise de comportamento de consumidores, que utiliza RFI (*Radio Frequency Identification*) para coletar dados sobre a influência do tamanho da rota no comportamento de compra dos clientes dentro de um supermercado do Japão no ano de 2009. O trabalho investigou como se dá a relação entre a distância percorrida na rota e o valor total da compra encontrando um coeficiente de correlação de Pearson (*ρ de Pearson*) igual a 0.8457. Esse resultado significa que essas duas grandezas possuem uma dependência linear forte, visto que quanto mais próximo de 1, maior a correlação. Outros fatos abordados foram o valor médio por pedido de 3525 yen (aproximadamente R\$100.00 em valores atualizados), e a quantidade média de 20 itens por pedido. Outros trabalhos atuais também abordam esta interessante relação entre as características da rota percorrida e o comportamento do cliente em lojas de comércio varejista ([KANEKO; YADA, 2016](#); [SYAEKHONI; LEE; KWON, 2016](#)).

Análises do comportamento humano em um ambiente virtual são utilizadas por [Tastan e Sukthankar \(2011\)](#) para propor um modelo de movimento para estimar caminhos em ambientes fechados. Um dos objetivos dos autores é diminuir os efeitos causados pelas restrições impostas pelos atores humanos para os problemas de reconhecimento de padrões. Este objetivo é trabalhado colocando um ser humano para interagir com o mesmo ambiente onde o algoritmo vai ser executado para que ele possa utilizar o conhecimento adquirido antecipadamente sobre percepção e desvio de obstáculos. Assim estas informações são utilizadas como entrada para os algoritmos de treinamento do modelo para que se tenha a capacidade de estimar caminhos e rotas para aplicações de computação ubíqua. O modelo

de direção de humanos se propõe a identificar o caminho e o destino final desejado com simulações em ambientes virtuais controlados com inserção de obstáculos nos caminhos e levando em consideração a velocidade, aceleração angular e destino final. Esta abordagem apresenta como principal desvantagem a necessidade de um ser humano realizar a experiência em um ambiente virtual controlado para que possa produzir os resultados esperados.

3.2 Tecnologia a Serviço do Varejo

Nesta seção são apresentados os trabalhos que dissertam sobre ferramentas de tecnologia que podem melhorar a experiência de compras dos consumidores nos ambientes de varejo.

De acordo com [Orel e Kara \(2013\)](#), o mercado mundial de supermercados está assistindo um aumento da utilização de tecnologias de auto-atendimento durante o processo de compras. Segundo os autores, a qualidade do serviço está intimamente relacionada com a satisfação do cliente e o trabalho examina este nível de serviço, principalmente tecnologias de *self-checkout systems* (SCS), e o impacto na satisfação e fidelização dos clientes de supermercados e mercearias em um mercado emergente (Turquia).

[Chen, Li e Hu \(2015\)](#) propõem um algoritmo genético para resolução do problema de planejamento da rota de compras em um supermercado que explora o espaço de busca procurando minimizar a função objetivo para encontrar as melhores soluções. O algoritmo proposto utiliza representação por permutações e mede a qualidade das soluções com uma função que mede a distância percorrida em cada rota. O algoritmo utiliza como entrada uma lista de produtos contendo suas coordenadas (x, y) no plano cartesiano. Esse trabalho alcança o objetivo principal de encontrar a melhor rota, entretanto, o modelo utiliza a distância euclidiana como métrica, não considerando os obstáculos, como gôndolas e freezers, dentro do supermercado.

3.3 Abordagens de Localização *Indoor*

Nesta seção são analisados os trabalhos que utilizam as técnicas de *Fingerprinting* e Triangulação, seja baseados em *Bluetooth* BLE, *Wi-Fi*, ou modelos híbridos.

No trabalho de [Cai \(2014\)](#), o autor propõe um modelo de localização indoor para ambientes de lojas de varejo. São instalados 6 pontos de acesso *Wi-Fi* em uma área interna de $900m^2$ para aferição da localização dos dispositivos, que é feita por meio de Triangulação com medidas de RSSI. Na arquitetura do sistema proposto, os clientes utilizam uma aplicação móvel e os funcionários utilizam uma aplicação *web* para obter as informações gerenciais. O aplicativo móvel coleta dados dos dispositivos como *timestamp*, *mac address*, além das

informações de força do sinal de cada um dos pontos de acesso visíveis, que escaneiam o espectro de 2.4GHz (802.11 b/g/n) e 5GHz (802.11 ac), para definir as coordenadas x e y daquele cliente. Para atingir a maior precisão possível na definição das coordenadas de um cliente, o sistema utiliza técnicas como modelagem bayesiana, agrupamento e regressão com RNA. Este modelo consegue uma melhoria de 25% após aplicação da modelagem bayesiana, porém apresenta a desvantagem da dependência de uma infraestrutura de rede de qualidade instalada que nem sempre está disponível.

Abreu (2014) desenvolve um trabalho de localização *indoor* utilizando como cenário de testes uma casa com área de 145.44 metros quadrados. O autor utiliza 4 roteadores *Wi-Fi* de diferentes fabricantes, configurados em frequências distintas, e dispositivos móveis para recebimento do sinal dos 4 pontos de acesso e armazenamento dos valores de RSSI. Nesse trabalho, o autor alcança uma precisão de poucos metros, porém com alto consumo de energia devido a tecnologia utilizada, se comparada com o BLE. Sua técnica mais assertiva, *Weighted Centroid*, apresentou resultados com o melhor caso em 0.08m de erro, pior caso 7.53m e erro médio de 1.80m. Como os trabalhos lidam com dados não normais, a aplicação de parâmetros e métodos não paramétricos seria mais indicada, porém o autor não cita dados como mediana e desvio padrão, apenas erro mínimo, médio e máximo.

No trabalho de Cho et al. (2015), os autores propõem um método de medição de distância entre dois dispositivos (um *beacon* BLE da *Estimote* e um *smartphone*). A proposta é uma abordagem de correção de erro do ruído, presente no sinal BLE, por meio da inclusão de um dispositivo adicional (nomeado de *beacon* de auto-correção) para utilização na calibração da distância de referência. Porém esse dispositivo adicional é um *smartphone*, ficando o sistema constituído por 2 *smartphones* e um *beacon*. O *beacon* fica sempre na mesma posição e a uma distância constante de 1m do dispositivo adicional, enquanto que o aparelho de medição varia sua distância entre 0.4 e 1.4 metros. O dispositivo adicional recebe o sinal do *beacon* e transmite o valor do RSSI recebido para o dispositivo de medição, que pode utilizar este valor, além do valor da intensidade do sinal que ele recebe diretamente do *beacon*. Como o dispositivo adicional fica sempre a uma mesma distância de 1m do *beacon*, o dispositivo de medição pode utilizar o valor do RSSI de correção para mitigar o ruído. O autor reporta uma melhoria no erro médio da distância em 10% com um intervalo de 1.5m. Essa abordagem é uma boa estratégia para o caso de medição de uma única distância isoladamente, porém inapropriada para ambientes reais e complexos onde são calculadas as distâncias para *n beacons* e utilizados como entrada para o método de Triangulação. Além do aumento da complexidade de implantação e calibração, o principal fator pode ser o custo de se manter um dispositivo de auto-correção a uma curta distância de cada um dos *beacons* utilizados como âncoras.

Um estudo de viabilidade da tecnologia *Bluetooth Low Energy* para sistemas de posiciona-

mento *indoor* é proposto no trabalho de Larsson (2015) com a utilização de 6 *Beacons* e um receptor. Nesse trabalho, o ambiente de testes foi uma pequena área aberta de um edifício comercial, com aproximadamente $35m^2$. Nesse espaço foram demarcados 33 pontos de medição com distância de 1m entre eles. A técnica utilizada foi o *Fingerprinting* e o trânsito de pessoas no local de testes foi impedido durante a coleta de dados. Os *beacons* foram configurados com um intervalo de transmissão de 500ms. A partir dos dados coletados, o autor validou a distribuição de dados RSSI, descartando a hipótese de normalidade dos dados e mostrando 3 separações relevantes nos valores de intensidade do sinal, atribuindo este comportamento ao uso dos 3 canais de *advertisement* (ver Figura 5). Esse resultado foi utilizado como base para a decisão de utilizar apenas um dos canais nos experimentos seguintes. O autor pondera que, como os 3 canais diferem na frequência, suas propriedades de atenuação de sinal também diferem entre si e essa variação gera impacto no sistema, dividindo os valores do RSSI nas 3 faixas distintas. O resultados obtidos mostram um erro médio de 1.9m na abordagem por pontos (*Points*). Na abordagem por rota (deslocamento em linha reta, em uma mesma direção) os resultados foram 1.3m e 1.4m, com paradas a cada 1m e 0.5m, respectivamente para os experimentos *Line 1* e *Line 2*. É importante destacar que neste trabalho foi utilizado apenas o método *Fingerprinting* e que a quantidade de amostras foi reduzida, com 13, 14 e 16 unidades de teste, respectivamente para os experimentos *Line 1*, *Line 2* e *Points*.

Kriz, Maly e Kozel (2016) propõem a inclusão dos *beacons* BLE com o objetivo de melhorar a precisão e exatidão em um sistema de localização *indoor* baseado em *Fingerprinting* utilizando pontos de acesso *Wi-Fi*. Os autores justificam a proposta apresentando resultados para os experimentos antes e depois da inclusão dos *beacons*. Nas configurações utilizadas nos experimentos houve uma melhora nos resultados, de 23% na mediana e redução na variância, quando utilizando a abordagem combinada (BLE + *Wi-Fi*) em comparação com a utilização somente dos pontos de acesso *Wi-Fi*. O resultado alcançado no trabalho utilizando apenas os *beacons* foi um erro de 2m na mediana e com o intervalo interquartil entre 1m e 4m.

Outra característica importante é que Kriz, Maly e Kozel (2016) desenvolveram o sistema de forma a permitir que voluntários incluam novos registros de medições para atualizar continuamente a base de dados. A estimativa de posição é feita utilizando o algoritmo *Weighted K-Nearest Neighbors* (wKNN), uma variação do KNN que promove a inserção de uma estrutura de pesos para os vizinhos em uma relação inversa com a distância para o ponto de interesse. O ambiente dos testes foi uma área de 52m x 43m, utilizando 8 pontos de acesso *Wi-Fi* (em cada marcador do mapa existe 1 ponto de acesso de 2.4GHz e outro de 5GHz), e 17 *beacons* BLE fixados no teto, da fabricante Estimote, pioneira no mercado dos *beacons*. O intervalo de transmissão utilizado foi de 100ms. Este é um dos quesitos que podem variar dependendo do hardware utilizado. Para efeito de comparação, a configuração

mínima dos *beacons* da DEV, utilizados no presente trabalho, é de 200ms (por *frame*). O que significa que, teoricamente, utilizando o dispositivo da DEV é possível alcançar apenas a metade do dados coletados pelo *beacon* da Estimote, no mesmo intervalo de tempo.

3.4 Discussões Finais

A partir da pesquisa bibliográfica, é possível verificar que existem várias tecnologias que são utilizadas para resolução de problemas de localização *indoor* atualmente. Sendo que as tecnologias de Radiofrequência prevalecem entre as mais utilizadas e que alcançam os melhores resultados. Dentre essas tecnologias se destacam a utilização dos pontos de acesso *Wi-Fi* e o *Bluetooth Low Energy*. A utilização do *Ultra Wide Band* (UWB) tende a crescer a partir deste ano devido à previsão de lançamento de novos *beacons* UWB no mercado, o que pode ajudar a reduzir os custos e aumentar a disponibilidade destes equipamentos. O *Wi-Fi* apresenta a desvantagem de alto consumo de energia se comparado com o BLE. Os *beacons* BLE se destacam pela versatilidade, facilidade de instalação, e baixo consumo de energia, mas ainda estão na mesma faixa exatidão da tecnologia *Wi-Fi* devido a fatores importantes como a interferência de outras tecnologias de transmissão sem fio que ocupam as mesmas faixas de frequência do espectro eletromagnético, como o próprio *Wi-Fi*.

As técnicas mais utilizadas atualmente são o *Fingerprinting* e a Triangulação. A primeira possui um custo muito maior na fase de calibração, devido à necessidade de construção da base de dados de aprendizagem para o modelo, onde cada ponto do mapa deve ser medido. Outra desvantagem do *Fingerprinting* é a forte dependência das condições físicas do local, tais como a lotação. Se essas condições variam muito entre o momento das medições da fase *off-line* e o momento da utilização na fase *on-line*, o resultado pode ser comprometido. A Triangulação possui menor dependência da fase de calibração e por isso é a técnica mais escalável. Modelos híbridos também podem gerar bons resultados, como a união do *Wi-Fi* com o *Bluetooth Low Energy*, ou a combinação das técnicas *Fingerprinting* e Triangulação.

É importante salientar que, para comparação entre os resultados dos trabalhos, devem ser levados em consideração fatores importantes como (a) densidade dos *beacons*; (b) aspectos físicos do ambiente de testes incluindo a incidência ou não de obstáculos móveis como pessoas; (c) os equipamentos utilizados como transmissores e receptores dos sinais de radiofrequência; (d) a configuração utilizada nos transmissores, principalmente no que diz respeito à relação potência/consumo de energia; (e) as técnicas de localização utilizadas; e (f) a fórmula de cálculo do erro.

4 Metodologia

Nesta seção será apresentada a forma como o trabalho deverá ser conduzido para o desenvolvimento da proposta, a execução dos experimentos e a validação dos resultados. Foi realizada, conforme apresentado nos capítulos 2 e 3 uma revisão da literatura para as áreas que envolvem o presente trabalho. Essa revisão inclui tecnologias utilizadas como arcabouço para os serviços de localização *indoor*, técnicas e algoritmos como Triangulação, *Fingerprinting*, classificadores, e finalmente as estratégias mais utilizadas para desenvolvimento de sistemas de navegação *indoor*.

Este trabalho será elaborado por etapas, seguindo uma sequência lógica desde a identificação da área de interesse até a etapa de execução dos experimentos e análise dos resultados. A sequência e a descrição dessas 5 etapas a serem executadas no desenvolvimento deste trabalho são apresentadas de forma resumida na [Figura 8](#).



Figura 8 – Etapas para desenvolvimento do modelo

4.1 Estudo de Caso

Este trabalho possui como foco a aplicação no varejo. Aplicações de localização *indoor* são relevantes no ambiente de comércio varejista devido à grande quantidade de itens disponíveis, as extensas dimensões e a relevância de se obter informações sobre o comportamento de compras dos consumidores na busca por oferecer uma experiência mais personalizada. Com as aplicações de micro-navegação assistida, os lojistas passam a ter um mapa comportamental individual dos clientes e podem assim ofertar produtos e serviços com alto grau de personalização. As técnicas que serão avaliadas neste trabalho pretendem suportar essas aplicações nestes ambientes.

Foi utilizado como estudo de caso um supermercado de Belo Horizonte, que forneceu dados relevantes como a planta de uma das lojas e registros sobre a localização de produtos. Neste material foram obtidas informações sobre as dimensões do ambiente, tamanho das gôndolas, corredores e localização dos demais mobiliários. Entretanto, alguns fatores como a necessidade de acessibilidade e o grande volume de testes necessários nas fases iniciais, acabaram tornando inviável a utilização do ambiente do supermercado para execução dos testes. Assim, foi feita uma opção por um ambiente mais similar possível, dentro do CEFET-MG, a biblioteca do campus II. A disposição dos móveis e corredores é similar ao supermercado com suas gôndolas e áreas de circulação de pessoas. Assim, os resultados experimentais tendem a ser mais próximos possível do ambiente comercial.

4.2 Hipóteses

Neste trabalho os principais parâmetros a serem testados são eficiência e desempenho da solução proposta, conseqüentemente as hipóteses que devem ser testadas permeiam estes requisitos. O parâmetro de desempenho está especialmente ligado à factibilidade da técnica para resolver o problema em tempo real.

Com relação à eficiência, serão utilizadas duas métricas para avaliação das soluções, dependendo da técnica em questão. Para a Triangulação será utilizado o erro em metros, que é quantificado por meio da distância euclidiana entre o resultado calculado e o local real do ponto onde foi feita a medição (armazenado na fase de coleta de dados). Já para a técnica *Fingerprinting*, como serão utilizados os classificadores, a avaliação da performance se dá por meio da taxa de acertos.

As hipóteses são exibidas abaixo:

1. O modelo proposto gera soluções melhores ou similares (de acordo com as duas métricas já abordadas) aos resultados disponíveis na literatura (conforme apresentados na [Seção 3.3](#)), apresentando uma solução de baixo custo e utilizando equipamentos

nacionais;

2. O modelo proposto resolve com tempo de latência adequado o problema de localização de um dispositivo leitor.

4.3 Planejamento de Experimentos

Os experimentos foram divididos em quatro grupos, sendo (A) Análise de Dados, (B) Tempo de Leitura, (C) Calibração e (D) Localização. Este último, o principal, para definição da localização *indoor*, que vai contar com variações que serão detalhadas em momento oportuno. As características do ambiente onde serão realizados os experimentos são importantes para definição de componentes dos mapas que serão apresentados no trabalho e serão apresentadas a seguir.

4.3.1 Características e Mapa do Ambiente

A estrutura da biblioteca, no primeiro piso, é composta por duas filas de prateleiras de livros que dividem os corredores verticais e três corredores horizontais, em conformidade com a [Figura 9](#). Os pontos no ambiente são definidos em um plano cartesiano com duas dimensões, sendo:

1. x: parede paralela ao estacionamento frontal e ao Restaurante Estudantil (RE);
2. y: parede perpendicular ao estacionamento do RE, ou seja, da entrada da biblioteca.

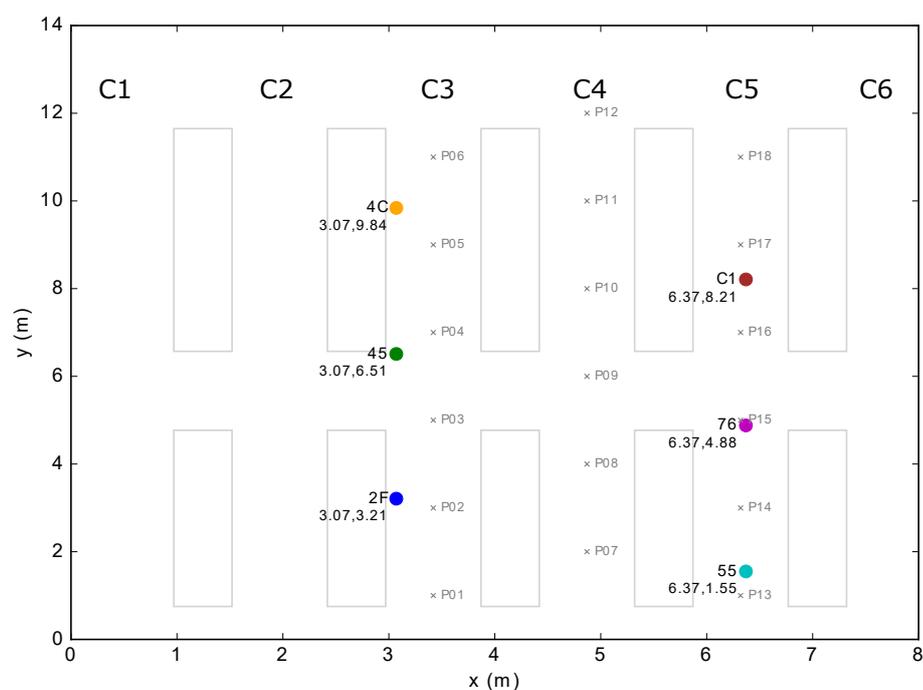


Figura 9 – Mapa do experimento D1

Os corredores são contados a partir do ponto $x,y = 0,0$ como C_1 até C_n . No mapa do experimento D1, [Figura 9](#), a área de cobertura está concentrada nos corredores C_3 , C_4 e C_5 , conforme pontos 1 a 6, 7 a 12 e 13 a 18, respectivamente, de acordo com indicação na própria figura. Para definição da localização dos pontos dentro dos corredores, foi utilizada uma distância de 2m entre eles, iniciando-se o primeiro ponto de cada corredor a 1m ou 2m, alternadamente. O mapa do experimento D1, também exibe a posição relativa dos *beacons*, que na figura estão representados pelos pontos coloridos, com *labels* para o nome do *beacon*, com dois dígitos hexadecimais (ver [Subseção 4.3.2](#)) e a sua posição no plano cartesiano. Os mapas de todos os experimentos podem ser consultados no [Apêndice A](#).

4.3.2 Identificação dos *Beacons*

Para facilitar a identificação, os *beacons* são identificados a partir de seus números de série simplificados, e especificados por Algarismos hexadecimais. Para referenciar os dispositivos neste trabalho, são utilizados os dois últimos dígitos do endereço físico, conforme [Tabela 3](#).

Tabela 3 – Nomenclatura dos *beacons*

Beacon	Mac address	UUID
2F	C1:F5:F3:92:93:2F	E9FF
45	F0:64:27:1A:1E:45	0CB6
4C	D9:7A:00:C6:5C:4C	90BF
55	D5:86:4C:3B:DB:55	4F34
76	D3:6C:52:4F:49:76	53DC
C1	DF:3F:A3:8F:5C:C1	5150

4.3.3 Experimento A - Análise de Dados

Com o intuito de validar os dados gerados pelos *beacons*, serão conduzidos experimentos com cada um dos *beacons*, e com todos concomitantemente, a uma distância fixa do aparelho receptor, e com os dados coletados serão realizadas análises da distribuição gerada, da dispersão dos dados e análise do ruído presente nas medições de RSSI com base no hardware específico utilizado.

4.3.4 Experimento B - Tempo de Leitura

Este experimento será conduzido utilizando diferentes distâncias e diferentes *beacons* e a cada coleta de dados será calculada a média dos dados para a definição do tempo de leitura necessário em cada ponto, por meio da verificação do tempo (em segundos) onde o valor médio do RSSI se estabiliza. Esta definição será utilizada nos demais experimentos, como o tempo mínimo de coleta para que se possa utilizar os parâmetros estatísticos de cada amostra de medições.

4.3.5 Experimento C - Calibração

Em seguida será conduzido o experimento de calibração dos dispositivos, processo que é conduzido com leituras sucessivas a uma mesma distância dos *beacons* e repetindo o procedimento para cada uma das distâncias de interesse, que será entre 0.25 e 20m, nos seguintes intervalos: de 0 a 1m com intervalo de 0.25m; de 1m a 10m com intervalo de 1m; e finalmente de 10 a 20m com intervalo de 2m.

4.3.6 Experimento D - Localização

Serão realizados 8 experimentos distintos, numerados de D1 a D8, para definição da localização, variando a posição dos *beacons*, de acordo com a área de interesse, para análise do impacto nos resultados. Os locais utilizados para fixar os *beacons* serão as vigas de ferro presentes no teto (LV), com o objetivo de obter um maior ângulo de visão possível para os pontos de medição, e as prateleiras (LP), de modo que a altura dos *beacons* seja a mesma do dispositivo de leitura, que será colocado sobre um banco de madeira.

A [Tabela 4](#) exibe a configuração que será utilizada em cada um dos experimentos (Exp). O item posição dos *beacons* exibe as coordenadas em que serão fixados os dispositivos em cada experimento. Os *beacons* são identificados a partir de seus números de série simplificados, e especificados por algarismos hexadecimais (2F, 45, 4C etc.), conforme [Subseção 4.3.2](#). É importante lembrar que os mapas de todos os experimentos podem ser consultados no [Apêndice A](#).

Tabela 4 – Configuração dos experimentos

Exp.	Local	Posição dos <i>beacons</i> (x, y)					
		2F	45	4C	55	76	C1
D1	LV	3.07, 3.21	3.07, 6.51	3.07, 9.84	6.37, 1.55	6.37, 4.88	6.37, 8.21
D2	LV	3.07, 0.87	2.69, 6.51	3.07, 11.71	6.37, 0.87	5.59, 6.51	6.37, 11.71
D3	LP	3.07, 3.21	3.07, 6.51	3.07, 9.84	6.37, 1.55	6.37, 4.88	6.37, 8.21
D4	LP	4.40, 6.66	4.40, 8.66	4.40, 10.66	5.32, 7.66	5.32, 9.66	5.32, 11.66
D5	LP	2.95, 6.66	4.40, 8.66	4.40, 10.66	5.32, 7.66	5.32, 9.66	6.77, 11.66
D6	LP	4.40, 6.60	2.95, 8.27	2.95, 9.97	6.75, 8.27	6.75, 9.97	5.32, 11.63
D7	LP	1.53, 9.12	2.67, 4.77	3.90, 9.97	5.58, 4.77	5.30, 8.27	6.75, 9.12
D8	LP	1.52, 0.75	1.52, 9.12	2.67, 4.77	5.58, 6.57	6.77, 2.76	6.77, 11.66

As medições serão realizadas com o dispositivo de leitura parado em um ponto fixo, dentro do ambiente, com a biblioteca em funcionamento normal (o que inclui trânsito de alunos e servidores). Para cada um dos pontos de leitura, o *tablet* será colocado em cima de um banco de madeira, posicionado com o seu centro em cima do ponto demarcado no chão, conforme numerados de P01 a P18 na [Figura 9](#). Após posicionado, a coleta de dados será iniciada manualmente com o comando *play* do aplicativo, e será finalizada automaticamente de acordo com configuração da quantidade de registros desejados.

4.4 Avaliação dos Experimentos

Cada um dos experimentos será analisado objetivamente a partir de técnicas estatísticas. No experimento A serão avaliadas as questões relativas à dispersão e distribuição de probabilidade, utilizando gráficos *boxplot* e histogramas. Testes de normalidade de dados também serão efetuados utilizando análises Shapiro-Wilk e Anderson-Darling para as variáveis de latência e RSSI, mais importantes no contexto de localização *indoor*.

O experimento B será avaliado utilizando médias móveis da potência do sinal em cada instante de tempo. A investigação do experimento C será finalizada por meio da geração de um gráfico que inclui o *boxplot* do RSSI de cada distância medida e a curva de referência para a função de distância.

Os experimentos referentes à localização serão avaliados de acordo com cada uma das técnicas utilizadas. Por se tratar da utilização de classificadores, as técnicas de *Fingerprinting* serão avaliadas recorrendo-se à matriz de confusão (uma representação padrão para avaliação de modelos estatísticos, em formato específico que permite a visualização e comparação entre os valores reais e aqueles preditos pelo classificador). Finalmente, a performance das técnicas de Triangulação serão medidas em metros, por meio da utilização de gráficos *boxplot*.

5 Análise Estatística dos Dados

A fase de validação dos dados coletados, referente ao Experimento A ([Subseção 4.3.3](#)), tem como objetivo realizar uma análise estatística dos dados gerados pelos *beacons* e coletados pelo aplicativo móvel. Esta análise inclui testes sobre a distribuição de probabilidade dos dados e comparação entre as amostras de dados gerados por diferentes dispositivos. Essa etapa é incluída no trabalho tendo em vista que características de hardware e de ambiente impactam diretamente na transmissão de dados via ondas eletromagnéticas. E para o bom funcionamento do modelo, é importante o conhecimento das especificidades dos equipamentos utilizados e possíveis impactos na geração dos dados.

5.1 Coleta dos Dados

A coleta de dados para experimento de análise estatística foi feito em uma sala fechada de 3.5m por 2.5m. Foi garantida a inexistência de obstáculos entre os *beacons* e o dispositivo de leitura, sem a presença de pessoas ou objetos móveis no ambiente. Os *beacons* foram fixados no teto e o receptor foi colocado sobre um armário resultando em uma distância de 1m entre eles.

A coleta de dados foi feita com os 3 *beacons*, selecionados aleatoriamente, posicionados a 1m de distância (em linha reta e sem obstáculos) do dispositivo de leitura, que neste caso foi um *smartphone Galaxy S7*, modelo SM-G930F. As medições foram feitas durante 2:42h, utilizando uma potência de transmissão (*TxPower*) de $-4dBm$ e um intervalo de transmissão (*Advertisement interval*) de $333ms$. Após o término do experimento foi obtida uma base de dados com 65705 registros, sendo 6.75 registros por segundo e 2.25 registros por segundo por *beacon*.

5.2 Latência

Testes para validação do intervalo de tempo entre as coletas de dados enviados pelos *beacons* foram feitos para identificar possíveis gargalos no recebimento dos dados. Neste experimento, o intervalo entre os registros, agrupados por *beacon*, foi calculado e armazenado para a análise. O *boxplot* com o tempo de leitura de cada *beacon* pode ser visto na [Figura 10](#).

Para os *beacons* 2F e 4C o comportamento foi conforme esperado, com as amostras muito próximas de zero, condizente com o tempo de transmissão de $333ms$. Porém, o

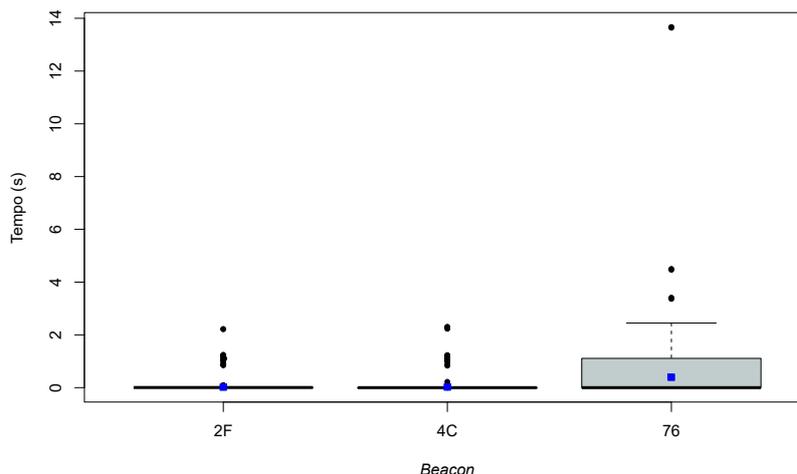


Figura 10 – Intervalo entre as coletas de dados

beacon 76 apresentou uma dispersão acima do previsto, com o terceiro quartil acima de 1 segundo e cerca de 30% dos valores entre 0.5 e 1.5 segundos. No momento dos testes, nas proximidades funcionavam equipamentos de rede, incluindo um roteador *Wireless* (802.11n/g/ac) e o *beacon* mais próximo do roteador era o 76. Acredita-se que esse seja o motivo de tal resultado

Este tipo de comportamento pode prejudicar o desempenho dos algoritmos de localização *indoor*, uma vez que utilizam um tempo de coleta para recebimento das medições e dados estatísticos (média e desvio padrão) da amostra obtida. Se um *beacon* está demorando mais para transmitir, prejudica a qualidade da amostra uma vez que a quantidade de registros obtidos é menor que o esperado e pode ser insuficiente para atingir a estabilização do valor médio do RSSI.

5.3 RSSI

Nesta seção é apresentada a análise dos dados de intensidade do sinal coletados dos *beacons*. De acordo como o *boxplot* referente ao RSSI por *beacon*, [Figura 11](#), o dispositivo 76 mais uma vez obteve um resultado diferente dos demais, a efeito do problema identificado na análise da latência. Ele foi o que apresentou a menor dispersão, justificável devido à menor quantidade de dados coletados. O dispositivo também foi o que apresentou a maior quantidade de *outliers*, (valores atípicos que estão localizados nos extremos da série de dados), o que também pode deteriorar a qualidade da amostra, pois apesar de serem tratados, quanto mais *outliers* menos dados serão utilizados nos algoritmos.

O histograma dos dados de intensidade do sinal recebido por cada *beacon* é apresentado na [Figura 12](#). Ao analisar as distribuições é possível verificar muita assimetria nos conjuntos de dados. E uma função de densidade de probabilidade gaussiana não modela adequadamente

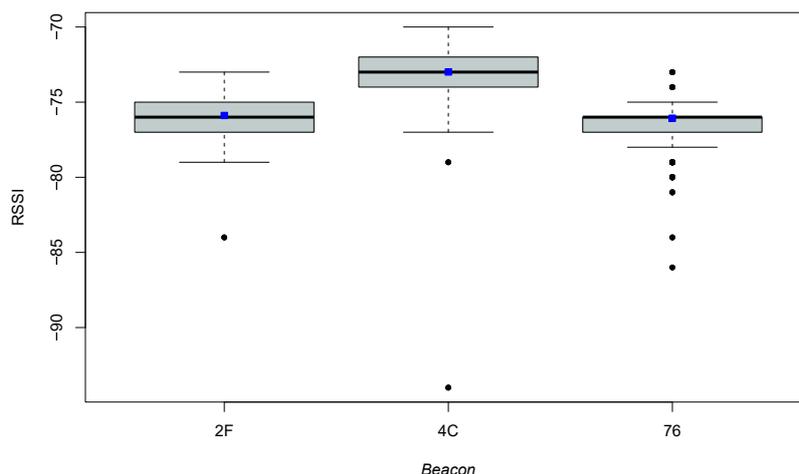


Figura 11 – Boxplot RSSI por beacon

a distribuição da variável RSSI. A seguir, testes de normalidade serão conduzidos para complementar o resultado e realizar a comparação entre os dados gerados por diferentes dispositivos.

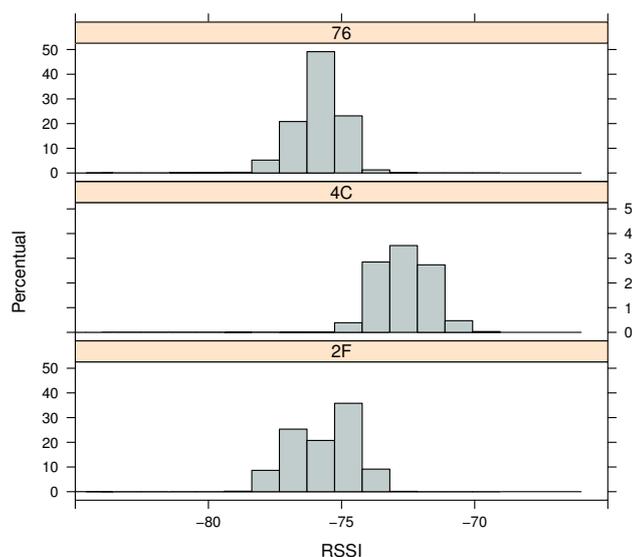


Figura 12 – Histograma RSSI

5.3.1 Normalidade dos dados

Os testes de normalidade utilizados para validação dos dados foram Shapiro-Wilk e Anderson-Darling, por serem os métodos mais utilizados nesse tipo de verificação e pela disponibilidade nos principais softwares ou bibliotecas de análise estatística. Os testes foram feitos com a base de dados completa e também com amostras aleatórias de tamanho 1000, 3000 e 5000 registros, para cada um dos 3 beacons. Em ambos os casos, o resultado obtido foi um p-valor igual ou menor a 2.2^{-16} , e desta forma, para um nível de significância

de 5%, podemos descartar a hipótese nula, que sugere a normalidade dos dados. Esse resultado corrobora com outros achados recentes (LARSSON, 2015), (RAMAKRISHNAN; GAUR; SINGH, 2016) onde também é verificada muita assimetria nos grupos de dados analisados.

Como não ficou comprovada a normalidade dos dados, será necessária a aplicação de métodos não paramétricos. O teste de Kruskal-wallis é uma alternativa não paramétrica para validação da hipótese de que as amostras de dados (geradas por cada um dos *beacons*) são independentes. Com esse teste podemos definir se as distribuições das amostras são idênticas sem assumir que elas sigam uma distribuição normal. De acordo com o p-valor encontrado pelo teste Kruskal-wallis, a hipótese de igualdade dos dados pode ser rejeitada, com 95% de confiança. Sendo assim, é considerado que existe diferença estatística entre os valores obtidos de cada um dos *beacons*.

5.4 Discussões Sobre a Avaliação

A hipótese de normalidade dos dados não foi observada nos experimentos realizados e métodos não paramétricos foram utilizados para comparar as amostras. As medições obtidas de cada *beacon* utilizado para enviar os sinais de identificação variam consideravelmente. Consequentemente, parâmetros de calibração de cada modelo devem ser levados em consideração no momento da análise dos dados recebidos para definição da localização. Com relação ao dispositivo utilizado para fazer a leitura dos dados, também foram feitos os mesmos testes comparativos utilizando diferentes *hardwares* e os resultados foram similares, ou seja, os dados são provenientes de amostras não relacionadas.

6 Desenvolvimento do Modelo

Neste capítulo serão apresentadas as etapas seguidas para a implementação das técnicas propostas, na sequência em que o estudo foi desenvolvido.

6.1 Representação do Estabelecimento

Com a planta da biblioteca, disponibilizada pelo departamento de projetos do CEFET-MG, conforme [Figura 13](#), foram obtidas as medidas principais para criação do modelo de representação do ambiente. Na planta podem ser observadas as dimensões básicas do ambiente, incluindo as dimensões das prateleiras e as vigas em estrutura de ferro fundido (tracejados azuis).

No ambiente físico foram feitas as medições para definição do posicionamento dos mobiliários e demais estruturas presentes no local. Estes dados foram tabulados em formato de arquivo texto estruturado CSV (*comma-separated values*) para atuar como entrada do algoritmo de criação do mapa do ambiente, sendo esse mapa uma parte importante do modelo de localização. Os atributos obtidos neste arquivo podem ser gerados facilmente por meio de uma consulta ao sistema de gestão do varejista, por exemplo, e apresenta os seguintes campos: Obj(tipo de objeto), label (descrição), loc_x e loc_y (localização 2d), dim_x e dim_y (dimensões do objeto). As prateleiras estão representadas por duas fileiras horizontais de retângulos, de acordo com a posição dos objetos no ambiente real.

6.1.1 Área de Cobertura

A partir da planta do estabelecimento, é definida a área de cobertura, que pode variar em cada experimento. As medidas utilizadas no mapa foram aquelas obtidas por meio da planta e de acordo com a área de interesse do experimento. Essas áreas estão definidas por corredores, e nos corredores estão os pontos de medição, numerados de forma crescente a partir da posição (0, 0) do mapa. Demarcada de verde na [Figura 13](#), abrangendo os corredores C3, C4 e C5, está representada a área de interesse dos experimentos D1 a D7. No experimento D8, o corredor C2 é adicionado. A área de cobertura de cada experimento pode ser consultadas também nos mapas, conforme [Apêndice A](#).

A garantia da cobertura de todos os pontos da área por todos os beacons utilizados no experimento é validada por meio de testes na base de dados de treinamento. A acurácia da localização depende diretamente dessa cobertura, que resulta do posicionamento dos beacons.

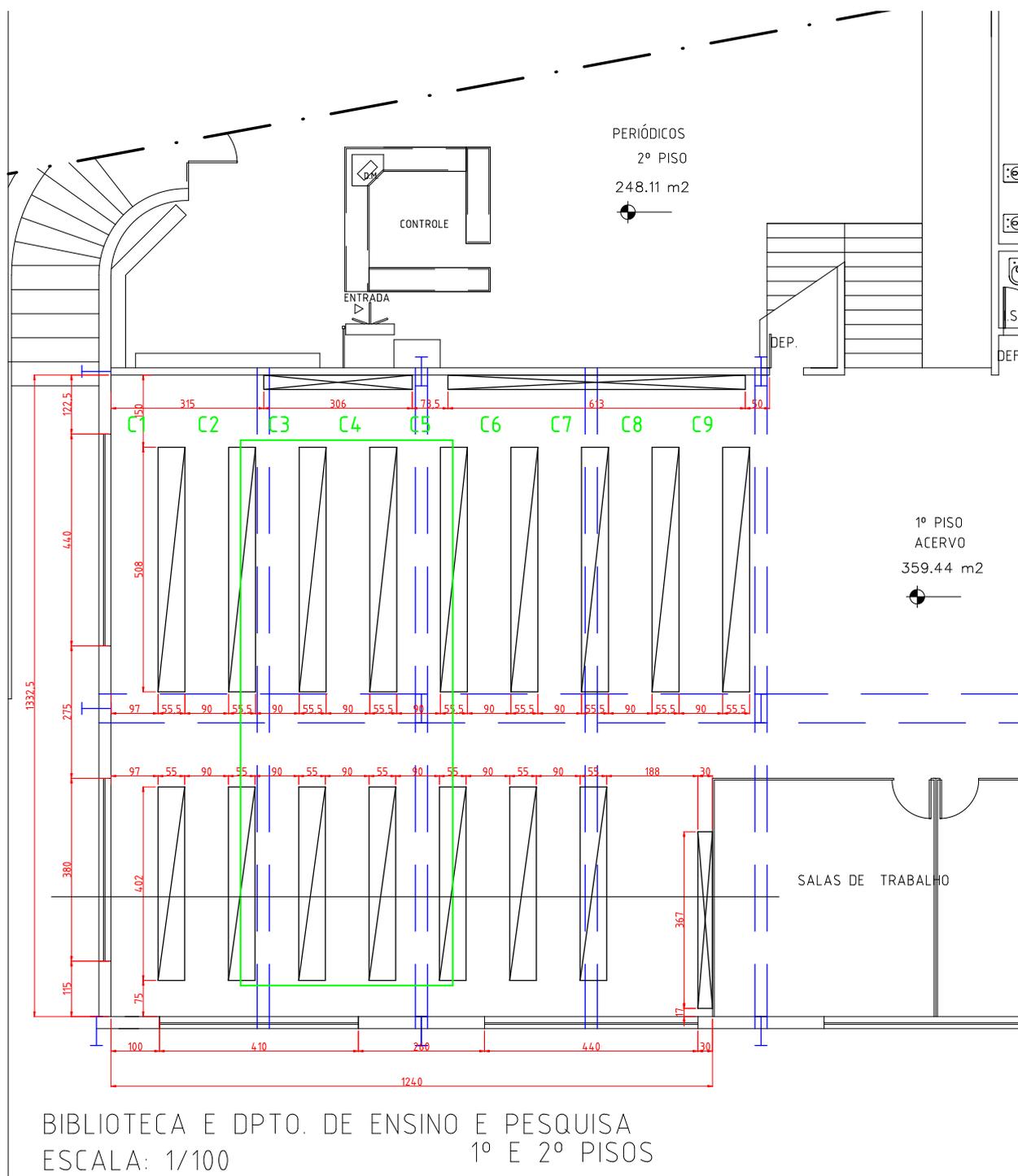


Figura 13 – Planta do ambiente

6.1.2 Posicionamento dos *Beacons*

A definição do posicionamento dos *beacons* foi realizada com o objetivo de obter a cobertura da área de interesse definida no experimento. Os *beacons* estão representados por pontos coloridos e nomeados de acordo com os dois últimos caracteres do *mac address* e tem suas posições variantes no mapa de acordo com o experimento. Conforme descrito na

Subseção 4.3.6, os *beacons* foram fixados nas vigas (teto) em alguns experimentos, e nas prateleiras em outros.

Para os *beacons* fixados nas vigas, a altura para o dispositivo de leitura é de 2.39m. Esse valor é utilizado no algoritmo de Triangulação para definir a distância euclidiana 2d. Por meio da utilização do Teorema de Pitágoras, o efeito da coordenada Z (altura) é compensado. Já para o caso dos *beacons* fixados nas prateleiras, a altura utilizada será exatamente a mesma do dispositivo de leitura, sendo desprezado o efeito desta variável.

Os experimentos com os *beacons* fixados no teto são utilizados com o intuito de aumentar o ângulo de visão direta entre os *beacons* e os pontos. No caso dos *beacons* fixados nas prateleiras, o sinal está mais suscetível à atenuação. Essa atenuação ocorre devido ao bloqueio físico gerado pelos obstáculos (prateleiras, livros, etc) existentes no local. Entretanto estes são obstáculos fixos, que não representam um grande problema, devido ao fato de que a atenuação tende a ocorrer de forma mais constante. Já no caso de obstáculos móveis (pessoas, pacotes, carrinhos de compra, entre outros) o problema é complexo para ser modelado. A maneira de enfrentar esse tipo de problema inclui a coleta de várias amostras e a utilização de variáveis estatísticas como medianas e desvios padrão.

Os experimentos D1, D2, D3 e D8 foram elaborados com o foco na abordagem da Triangulação, enquanto que os numerados como D4, D5, D6 e D7 com foco na abordagem *Fingerprinting*. O experimento D4 foi elaborado com foco na identificação de uma área de interesse para o problema binário, onde o objetivo principal é responder se o alvo está ou não nessa área. Entretanto, todos os 8 experimentos foram testados em todas as técnicas.

6.1.3 Mapa

O layout do estabelecimento é um pré-requisito que pode ser recebido por meio de arquivos texto (CSV) extraídos de um banco de dados. A primeira tabela de dados é a Planta, que possui os dados: largura, comprimento. A tabela seguinte será a Gondola, contendo os seguintes dados: código, largura, comprimento, tamanho dos slots, coordenadas (x, y) em relação à origem (que será o ponto mais ao norte/oeste da planta).

Um exemplo do *layout* do arquivo referente ao experimento D2 é exibido na Tabela 5. Nessa tabela é apresentada uma linha para cada tipo de objeto. O primeiro campo (Obj) é o identificador do tipo de objeto, sendo G para as gôndolas (ou prateleiras), B para os *beacons* e P para os pontos. O segundo campo (*label*) é o identificador do objeto. Os campos *loc_x* e *loc_y* são referentes à localização relativa do objeto no ambiente. Os campos *dim_x* e *dim_y* são referentes às dimensões do objeto.

Com estas informações foi possível gerar o mapa do ambiente físico no qual foram realizados

Tabela 5 – Exemplo de arquivo do *layout* do estabelecimento

Tipo	Descrição
Cabeçalho	Obj;label;loc_x;loc_y;dim_x;dim_y
Gôndolas	G;1A;0.97;0.75;0.55;4.02
Beacon	B;2F;3.07;0.87;0.10;0.10
Ponto	P;10;3.42;1.00;0.10;0.10

os experimentos para definição da posição do dispositivo leitor de acordo com cada medição de intensidade de sinal recebida dos *beacons*. O mapa do experimento D2 pode ser visto na [Figura 14](#). Esse experimento projetado com os *beacons* fixados nas vigas do teto, nas extremidades da área de interesse, com foco na técnica de Triangulação. Os pontos indicados com um x nos corredores e com o label iniciando com o caractere P são os pontos de coleta que serão utilizados no experimento.

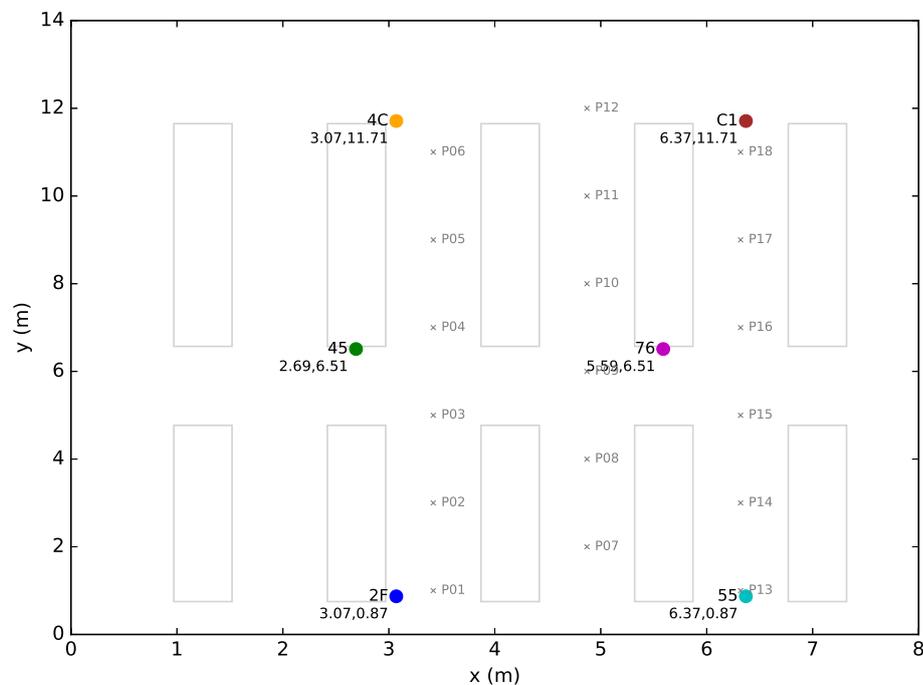


Figura 14 – Mapa do experimento D2

6.2 Dispositivos

Os 6 *beacons* utilizados nos experimentos são da fabricante nacional Dev Tecnologia, modelo *DEV Beacon Bluetooth Dual IP67 (Bluetooth Low Energy proximity-beacon)* e possuem certificação ANATEL. Conforme pode ser visto na [Figura 15](#), é composto por uma placa de circuito integrado, um case e uma bateria CR2477 que é utilizada como fonte de energia. As especificações técnicas mais importantes do dispositivo são replicadas na [Tabela 6](#), de acordo com o manual técnico disponibilizado pelo fabricante ([DEV, 2018](#)).



Figura 15 – Dev Beacon

Tabela 6 – Especificações técnicas

Especificação	Valor	Detalhe
Slots	4 unidades	Frames: UID, URL, TLM, EID, <i>iBeacon</i>
Adv Interval	0.1 a 10 segundos	Programável por slot
Tx Power	-40, -30, -20 a 4dBm (4dB)	Programável por slot
Radio	Bluetooth 4.1	Bluetooth Low Energy (<i>Bluetooth Smart</i>)
Antena	Omnidirecional	Gain 2dB (max.)
Fonte	1 Bateria	CR2477 (3V)
Certificação	Certificado nacional ANATEL	04186-16-03155

6.2.1 Configurações Utilizadas

As configurações utilizadas nos 6 *beacons* em todos os experimentos, exibida na Tabela 7, foi feita utilizando um aplicativo disponibilizado pelo fabricante e que pode ser baixado da loja de aplicativos do *Android*.

Tabela 7 – Configuração dos *beacons*

Atibuto	Valor	Observação
Frames	1 <i>Eddystone</i> UUID	demais frames inativos
Potência	4dBm	Potência de transmissão máxima
Intervalo	200ms	Intervalo de transmissão (<i>advertising interval</i>)

6.3 Calibração dos *Beacons*

O experimento de calibração é conduzido com o objetivo de ajustar os dispositivos correlacionando a distância da medição com o RSSI. São feitas medidas da intensidade do sinal a cada uma das distâncias (em metros): 0.25, 0.50, 1.00 a 10.00 (com intervalo de 1m), 12 a 20m (com intervalo de 2m).

Os dados obtidos neste experimento (distância versus intensidade do sinal) são utilizados

para obter os valores de referência para a distância padrão de 1m. Os mesmos dados servem como entrada para o algoritmo de regressão (modelo não linear) Gauss-Newton, utilizado para encontrar os coeficientes da função de cálculo da distância, conforme [Equação 1](#) da [Subseção 2.4.4](#), que será utilizada no algoritmo de Triangulação.

A partir dessa relação é gerada também a curva de referência da potência do sinal em função da distância da medição. O resultado analítico deste experimento pode ser visualizado na [Figura 16](#). Cada caixa do gráfico é referente a uma amostra de medidas de intensidade de sinal coletadas em uma distância fixa, conforme valores do eixo x (metros), e no eixo y estão os valores do RSSI (dBm).

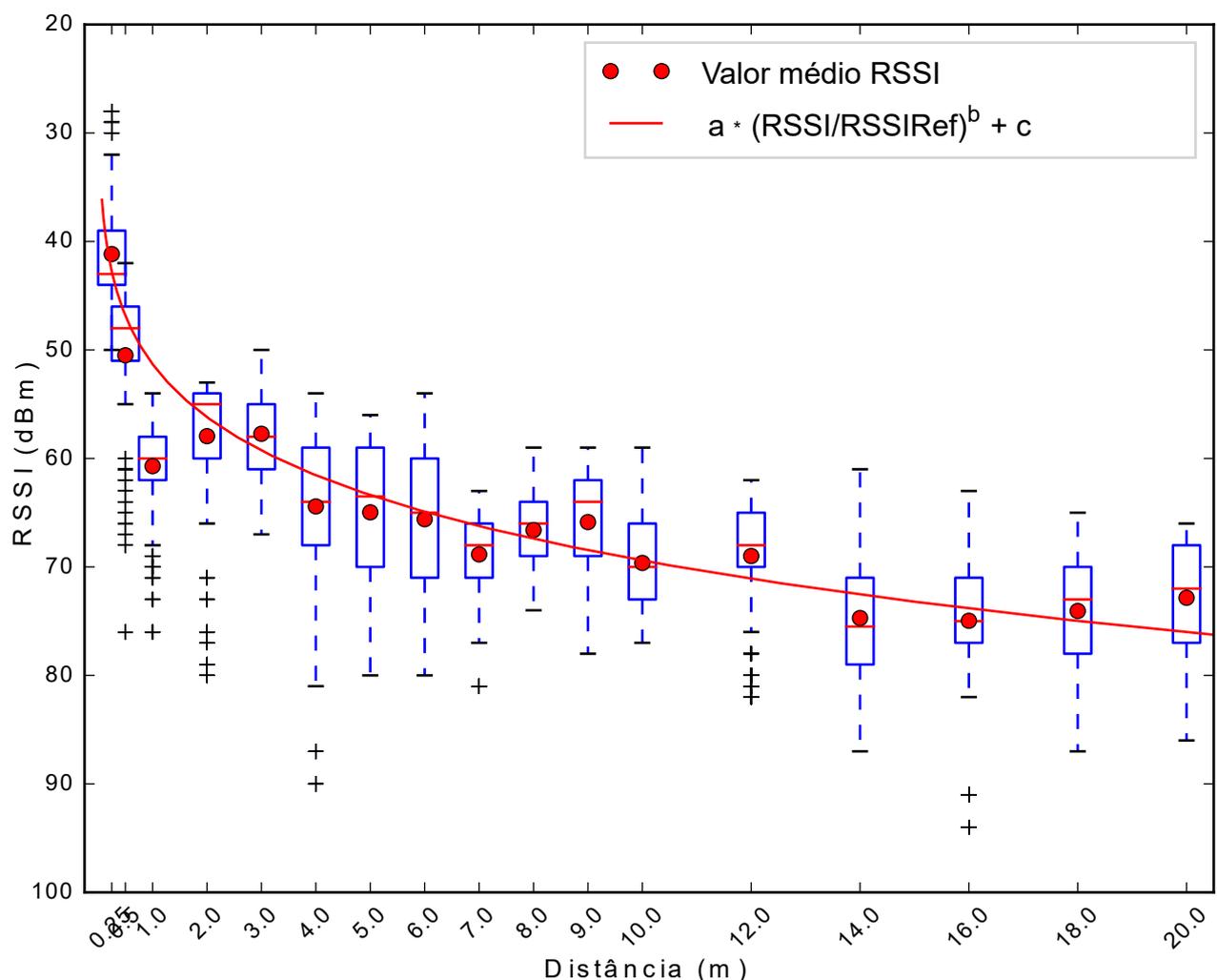


Figura 16 – Curva de referência para intensidade do sinal até 20m

O resultado dos parâmetros estimados pelo método de regressão não linear são apresentados na [Tabela 8](#). Após um estudo sobre a relevância dos coeficientes obtidos com este modelo, verificou-se que, para um intervalo de confiança de 95%, o parâmetro c não interfere no modelo, visto que o t -value obtido foi > 0.05 .

Tabela 8 – Análise de aptidão dos parâmetros

Parâmetro	Valor	Desvio Padrão	Erro	t-value
a	4.980	2.080	2.395	0.03118
b	8.048	2.022	3.981	0.00137
c	-0.896	2.109	-0.425	0.67737

Conseqüentemente, o parâmetro c foi removido do modelo e a nova fórmula ficou conforme [Equação 2](#). A partir deste novo modelo, foi feita uma nova regressão para encontrar os parâmetros a e b e esta fórmula foi adotada para os experimentos nas técnicas de Triangulação.

$$d = a \times \left(\frac{RSSI}{RSSI_{Ref1m}} \right)^b \quad (2)$$

6.4 Execução dos Experimentos

A coleta de dados foi feita na biblioteca, utilizando-se um aplicativo instalado em um *tablet* (Samsung Galax Tab A - modelo SM-P550), rodando *Android* versão 7.1.1 (Nougat). Foram feitas medições de 15 segundos em cada ponto, utilizando o *tablet* colocado sobre um banco de 0.60m de altura, centralizado nos pontos demarcados no chão. Os dados foram salvos em um arquivo texto, no formato CSV, contendo dados do RSSI medido, com indicação do ponto, *mac address* do *beacon*, entre outros dados. Após a coleta das medições de cada ponto *in loco*, as bases de dados foram armazenadas para atuação como entrada para os algoritmos de localização *indoor*.

6.4.1 Aplicativo *Android*

O aplicativo utilizado foi o *Beacon Scanner*, de código aberto, versão 2.1.0, disponível no *Github* ([BRIDOUX, 2018](#)). Foram feitas alterações no aplicativo para adequação às necessidades dos experimentos deste trabalho. Foi inserido um *log* local no dispositivo via arquivo de texto CSV, pois o padrão de *log* disponível funcionava apenas via HTTP/JSON. A tela de configurações do aplicativo que mostra as novas opções implementadas é exibida na [Figura 17](#).

Conforme tela de configurações apresentada, foram inseridas novas opções de *log* para a fase de coleta de dados. O primeiro item é a distância entre o *beacon* e o *tablet*, parâmetro *Distance Ref*, que é utilizado para o experimento de calibração dos *beacons*. O segundo item é o registro do ponto que está sendo feita a coleta (*Measured point*) para a fase *off-line* do *Fingerprinting*. Um terceiro parâmetro para definição do tamanho do *log* (*Log Max Length*)

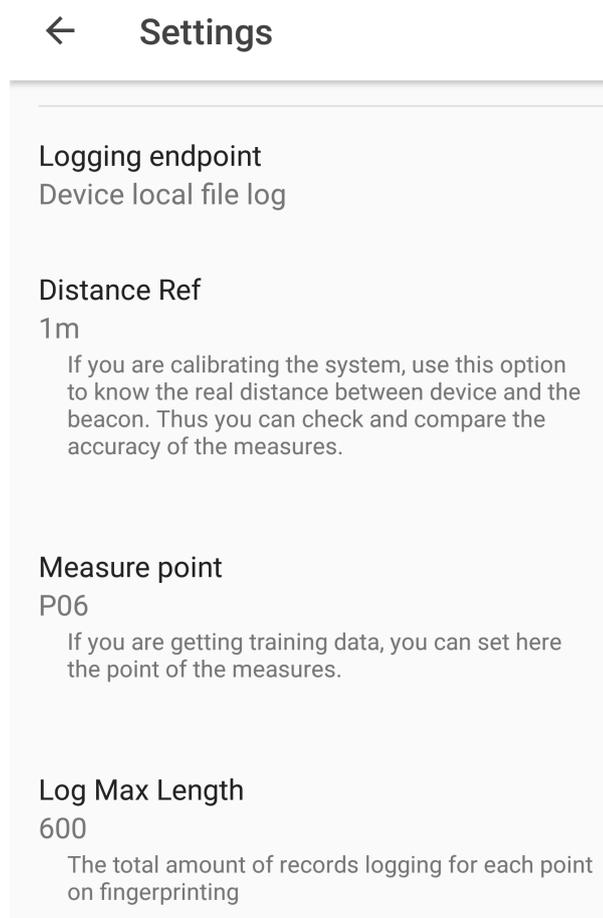


Figura 17 – Tela de configurações implementadas no aplicativo *Android*

também foi inserido, de forma a possibilitar ao aplicativo interromper, automaticamente, a coleta após alcançar a quantidade de registros desejada para uma amostra.

Outra melhoria implementada foi a visualização dos dados do experimento que está sendo executado, como o número do ponto ou a distância de referência, sempre que a coleta de dados está em andamento. Esta melhoria permite um maior controle na coleta de dados para garantir que os experimentos estão sendo conduzidos conforme planejado.

6.4.2 Medições dos Pontos de Coleta

Para definição do tempo de medição necessário, utiliza-se a análise de estabilidade das medições. Como pode ser analisado no gráfico da [Figura 18](#), a partir de 15 segundos de medições, a média já atinge um ponto de equilíbrio, com pequenas variações que não causam impacto significativo na exatidão dos cálculos de localização. O eixo x apresenta o tempo, em segundos, e no eixo y está a potência do sinal (valor médio do RSSI).

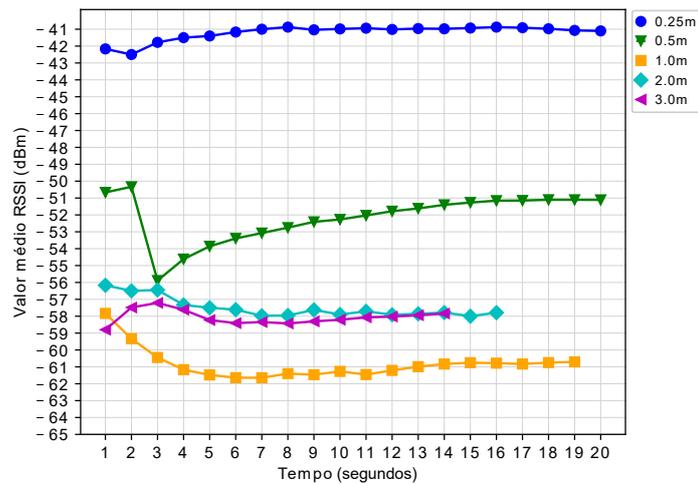


Figura 18 – Tempo de coleta de dados

6.5 Algoritmos

Nesta seção são apresentados os algoritmos mais importantes utilizados no modelo e as principais características.

6.5.1 Tratamento dos Dados

Na fase de coleta de dados, cada registro que foi armazenado no arquivo CSV possuiu o *timestamp*, RSSI, identificador do *beacon* e do ponto. Este arquivo foi utilizado para validação dos dados coletados com relação à integridade e em seguida os dados foram agrupados por ponto e pelo conjunto de 6 *beacons* (quantidade utilizada nos experimentos) gerando assim um novo arquivo. Cada registro desse novo arquivo possui a informação do ponto de coleta e os dados da intensidade do sinal de cada um dos *beacons*.

Em seguida foi realizada uma análise estatística dos dados com o objetivo de identificar e remover os *outliers*, uma vez que estes podem ser provenientes de erros na medição, dada a característica aleatória da distribuição dos dados. Os *outliers* podem enviesar negativamente os resultados do modelo aumentando o erro nos cálculos. Neste trabalho foram rotulados como *outliers* os dados dispersos mais de 2 desvios padrão da mediana.

6.5.2 Fingerprinting

Os algoritmos baseados em *Fingerprinting* que foram implementados utilizam os dados já tratados, agrupados por ponto e pelo conjunto de *beacons*. Esses dados caracterizam uma identificação digital de cada um dos pontos. Existem vários algoritmos de localização baseados em *Fingerprinting*, neste trabalho foram utilizados dois dos classificadores mais

difundidos e que têm apresentado os melhores resultados em recentes estudos, KNN e MLP.

Na implementação dos classificadores, os dados coletados após a realização do experimento são divididos em dois grupos. O primeiro grupo, composto por 80% dos dados, é utilizado como treinamento. O segundo grupo, correspondente aos 20% restantes, é utilizado apenas para validação do classificador.

6.5.2.1 K Vizinhos Mais Próximos

Neste trabalho foi implementado um algoritmo classificador KNN (k vizinhos mais próximos) para atuação sobre os dados coletados com a técnica *Fingerprinting*. O KNN emprega uma métrica de distância para estimar quais os pontos armazenados na base de dados de treinamento são os mais próximos da instância pesquisada (na fase *on-line*). Utilizando o parâmetro $k = 1$, caso particular 1-NN, a classificação é finalizada retornando o primeiro vizinho mais próximo.

Na implementação deste trabalho, a métrica de distância utilizada foi a diferença absoluta de vetores. Assim, a distância é calculada utilizando a diferença entre a intensidade do sinal das leituras armazenadas no conjunto de dados de treinamento e a amostra da fase *on-line*, conforme [Equação 3](#). Onde dav é a soma da distância absoluta entre os vetores, n é a quantidade de *beacons*, b_i é o *beacon* da posição i , md é a função que retorna a mediana do RSSI para o conjunto de dados do treinamento, e, de modo similar, \tilde{md} para a amostra da fase *on-line*.

$$dav = \sum_{i=1}^n (|md(b_i) - \tilde{md}(b_i)|) \quad (3)$$

O algoritmo foi implementado utilizando o valor médio das medições de cada *beacon* em cada ponto coletado na fase *off-line* como valor de referência, de forma que cada ponto possui um vetor de n posições (com $n = 6 =$ quantidade de *beacons* utilizada) correspondente ao valor de referência de cada um dos *beacons*, sempre ordenados em ordem alfabética pelo nome, conforme [Subseção 4.3.2](#).

Foi utilizado um vetor de pesos, que define uma maior prioridade para os *beacons* que estão teoricamente mais próximos do leitor (menor perda na intensidade do sinal), conforme [Equação 4](#). As diferenças encontradas para os 3 primeiros *beacons* foram reduzidas, respectivamente em 7%, 5% e 3%. Ao passo que para os 2 últimos a diferença é aumentada em 3% e 6%. A motivação para o vetor de peso dessa forma se dá pela natureza da relação entre a distância e a exatidão das medições. Quanto mais próximo do *beacon*, menor a

probabilidade de interferências no sinal.

$$weight = [0.93, 0.95, 0.97, 1.00, 1.03, 1.06] \quad (4)$$

6.5.2.2 Rede Neural *Multilayer Perceptron*

Além do KNN, foi implementado também um classificador MLP, utilizando uma Rede Neural *Multilayer Perceptron*, com algoritmo de treinamento supervisionado L-BFGS, um método iterativo para solução de problemas de otimização não linear irrestritos, conforme Fletcher (1987). Esse é um algoritmo da família quase-newton que aproxima o algoritmo BFGS (referente às iniciais dos nomes dos autores: Broyden–Fletcher–Goldfarb–Shanno) usando uma quantidade limitada de recursos computacionais. Isso se deve ao fato de que esses métodos utilizam apenas informações de primeira ordem, buscando uma aproximação iterativa da inversa da matriz hessiana (ANDREW; GAO, 2007). A utilização desse método é muito difundida para estimar parâmetros na área de aprendizagem de máquina.

Foram feitos vários testes para se obter a melhor relação entre performance e exatidão com relação à configuração de neurônios nas camadas escondidas da rede. Nesses testes, o melhor resultado foi alcançado utilizando duas camadas escondidas, com 150 e 25 neurônios, com o parâmetro $alpha = 1^{-5}$ (penalidade), para controle da magnitude dos pesos da rede. A base de dados do treinamento foi a mesma utilizada para o KNN.

Para validação do desempenho da rede, na avaliação da capacidade de generalização do modelo MLP, neste trabalho foi utilizada a técnica de validação cruzada chamada *k-fold cross-validation*. Nessa abordagem, um *dataset* D é randomizado e dividido em k partes, $D_1; D_2; \dots; D_k$, de tamanhos aproximadamente iguais. O modelo é treinado em $D \setminus D_t$ e testado em D_t k vezes, para cada *round* $t \in \{1, 2, \dots, k\}$ (KOHAVI et al., 1995). A exatidão da rede é calculada de forma global, utilizando a soma do total de classificações corretas de cada *round*, em razão do número de instâncias totais existentes no *dataset*. No caso particular deste trabalho, $k = 5$, conseqüentemente, dividindo a base de dados em 5 partes com aproximadamente 20% dos registros em cada.

6.5.3 Triangulação

Neste trabalho aplica-se as abordagens de multilateração, sendo que a distância é obtida a partir da intensidade do sinal recebido de cada um dos *beacons* (âncoras) no dispositivo do usuário (alvo). São utilizados 6 *beacons* e, a cada ponto de medição, todos aqueles que geraram pontos de interseção são utilizados como forma de melhorar a eficiência do modelo. Esses pontos são classificados de acordo com a menor distância total entre ele e os demais pontos existentes. Dos melhores classificados, são selecionados 6 pontos

para formar o melhor *cluster*. A partir do *cluster* selecionado, calcula-se o centroide para definição da localização do alvo.

O algoritmo de Triangulação por Lateralização utilizado neste trabalho foi implementado conforme referencial disponível na [Subseção 2.4.6](#), com introdução de algumas melhorias, que serão apresentadas a seguir.

6.5.3.1 Pontos Fora da Área de Interesse

Foi implementado um parâmetro para desconsiderar os pontos de interseção gerados fora do mapa (área de interesse), quando o número de pontos na área de interesse é igual ou maior a 6. Desse modo, sempre que existir pelo menos esta quantidade de pontos qualificados dentro do mapa, os pontos que estão fora dele são desconsiderados para o processo de eleição do melhor *cluster*.

Na [Figura 19](#), é possível verificar os pontos de interseção entre as circunferências, representados por triângulos pretos e vermelhos. Todos os pontos visíveis na área de interesse (mapa) foram utilizados para eleição do melhor *cluster*, sendo que os pontos vermelhos foram os eleitos. Apesar de não visíveis, existem pontos de interseção fora da área de interesse, como aqueles gerados pela interseção dos beacons 55 e 4C, 55 e C1, 55 e 45, entre outros.

6.5.3.2 Limite

No algoritmo de Triangulação foi verificado que algumas circunferências não geravam pontos de interseção, mesmo quando muito próximas, em alguns casos a poucos centímetros de distância. Consequentemente, um promissor ponto para participar da eleição do melhor *cluster* era perdido. Para melhorar esta busca pelos pontos de interseção e incorporar no modelo a característica de incerteza que existe nas medições (provenientes de distribuições com características aleatórias) foi implementado o parâmetro *Threshold* que define um limite superior na distância entre as circunferências para que seja considerado um ponto de interseção no modelo.

Assim sendo, esses pontos promissores serão aproveitados como um ponto de interseção sempre que estiverem dentro do limite definido. Consequentemente, o objetivo de trabalhar com regiões de grande potencial como candidatas no processo de seleção do melhor *cluster* é alcançado, ao considerar estes pontos onde as circunferências se aproximam muito, porém sem se tocarem.

Um exemplo do impacto, neste caso positivo, pode ser analisado na [Figura 20](#). Na imagem da esquerda, sem o parâmetro implementado, as circunferências com centro 4C e 2F não geram ponto de interseção por diferença de milímetros e o ponto calculado (estrela

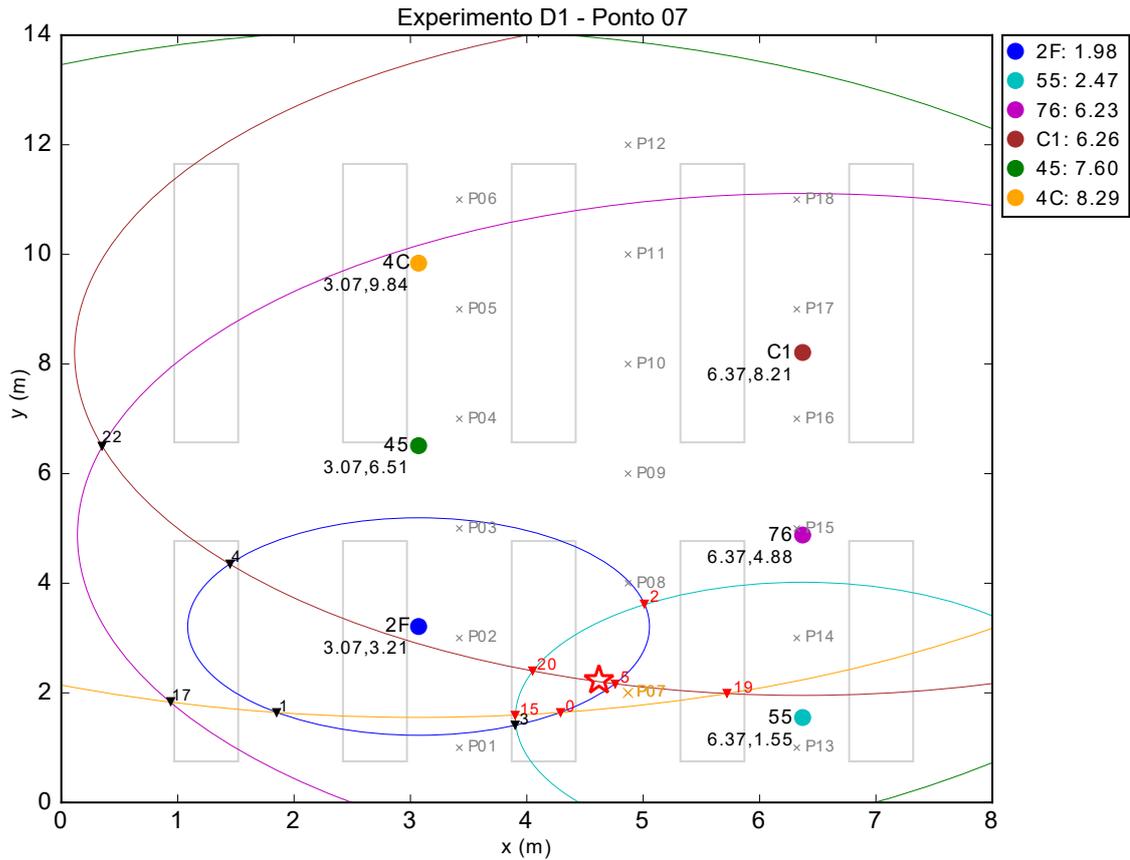


Figura 19 – Pontos de interseção dentro da área de interesse

vermelha) fica próxima ao ponto 14. Enquanto que na figura da direita é gerado o ponto de interseção 0 e o ponto calculado se aproxima do ponto real (P11).

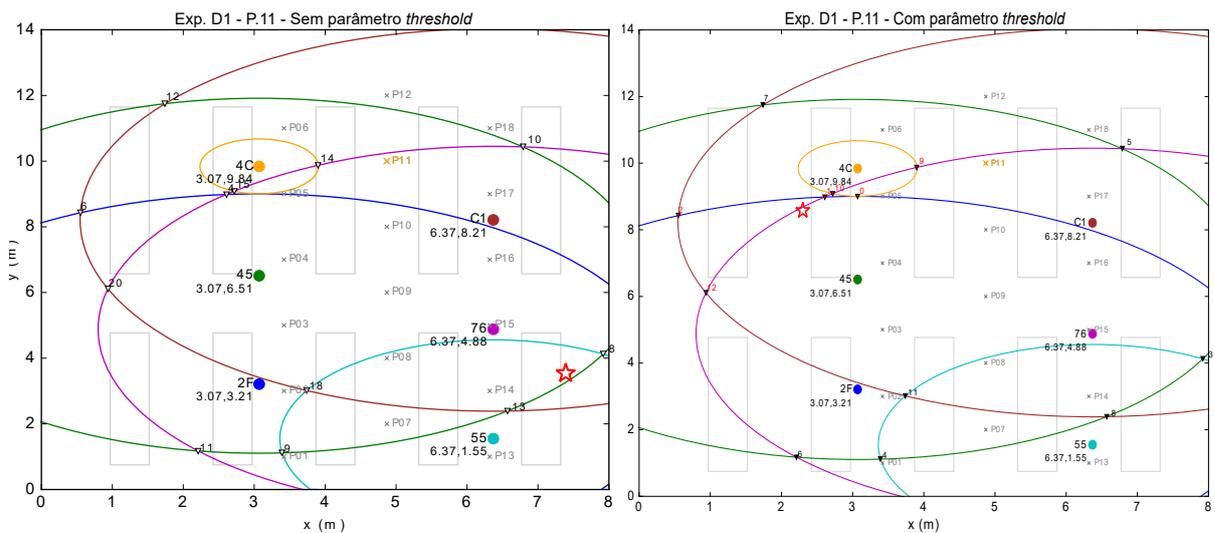


Figura 20 – Parâmetro *Threshold* com distância dentro da área de abrangência

Um outro exemplo desse comportamento pode ser visto em situações nas quais não é detectada nenhuma interseção entre as circunferências, consequentemente, não existe

nenhuma solução válida. Nesse caso, o algoritmo processa uma busca por uma solução adequando o valor do parâmetro *threshold* até que se obtenha pelo menos um ponto de interseção (ou atinja uma condição de parada, um percentual sobre o valor original do parâmetro), conforme Figura 21 do experimento D4, para o ponto 11.

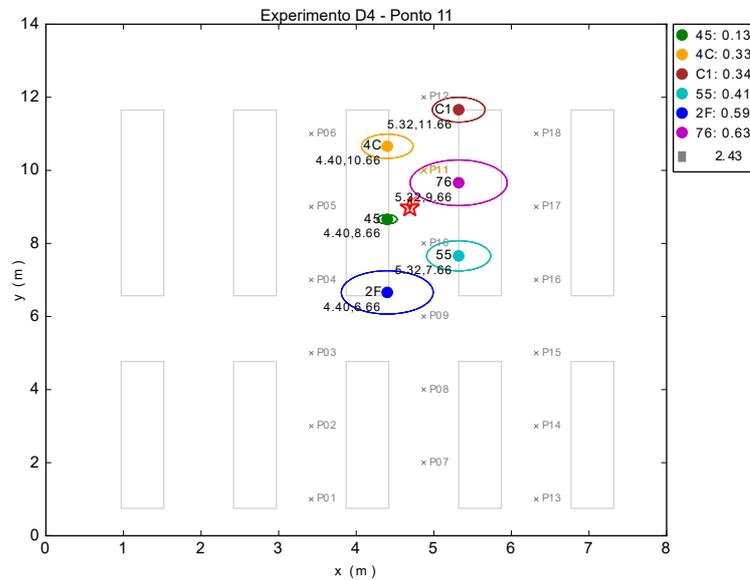


Figura 21 – Parâmetro *Threshold* - experimento sem nenhum ponto de interseção

7 Resultados, Análise e Discussão

Neste capítulo serão apresentados e discutidos os resultados alcançados no trabalho com as três técnicas implementadas. As discussões se alinham no sentido de extrair comportamentos e especificidades de cada experimento em cada uma das técnicas implementadas e por conseguinte, as seções serão divididas desta maneira.

Foram realizados 8 experimentos de localização distintos, numerados de D1 a D8, variando a posição dos *beacons*, de acordo com a área de interesse, conforme detalhado na [Subseção 4.3.6](#). Três agrupamentos para os resultados dos classificadores (KNN e MLP) foram utilizados, de acordo com os níveis de granularidade (ponto, corredor e corredor específico), em concordância com a [Figura 22](#). O objetivo dos classificadores é acertar o ponto exato onde foi feita a medição. Já nos experimentos de Triangulação, o propósito é encontrar o ponto mais próximo possível. O nível mais detalhado é por ponto, conforme pode ser visto nas caixas de cor azul no mapa. Em seguida, o agrupamento por corredor foi realizado acumulando os pontos compreendidos dentro ou próximos a cada corredor. Como foram 18 pontos, e cada grupo de 3 pontos foram agregados em um corredor, totalizando 6 corredores (cor laranja no mapa). Finalmente, no caso do corredor específico, exibido na cor verde, trata-se de um problema binário, para responder se o dispositivo leitor está ou não no corredor compreendido pelos pontos 10, 11 e 12.

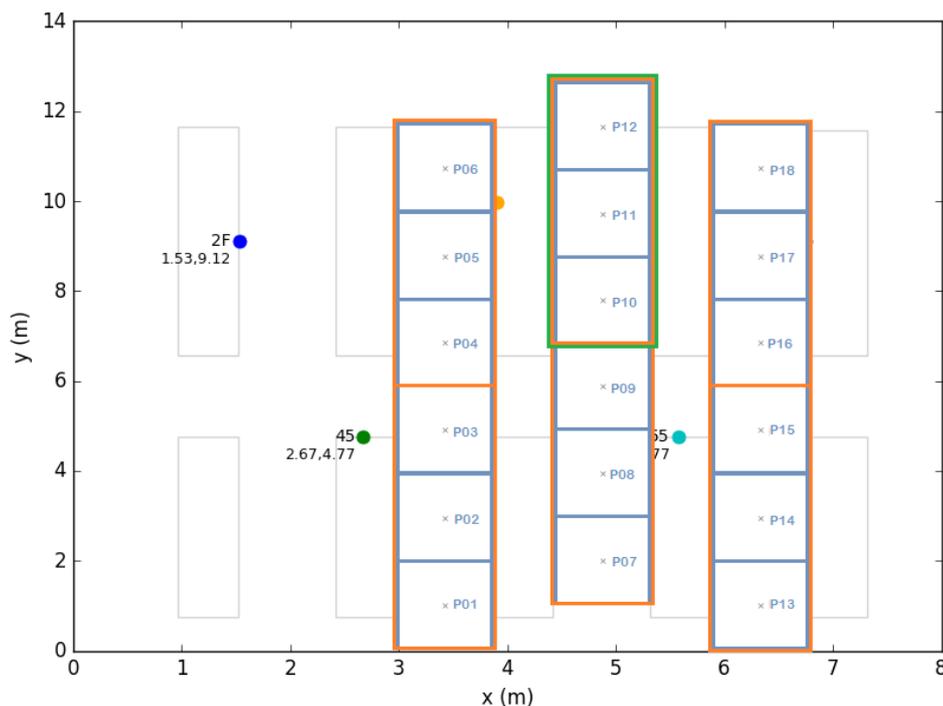


Figura 22 – Divisão da área de interesse por grupos

Como foi gerado um volume extenso de resultados, eles serão exibidos em forma de tabelas, e somente os resultados mais importantes serão destacados, a partir das representações mais relevantes de cada técnica, sendo a matriz de confusão para os classificadores e o *boxplot*, com erro em metros, para a Triangulação. As matrizes de confusão são representações padrão para avaliação de modelos estatísticos. Um formato específico que permite a visualização e comparação entre os valores reais (no eixo x), e aqueles preditos pelo classificador (no eixo y).

O desempenho dos classificadores foi medido utilizando a técnica de validação cruzada *k-fold cross-validation*, com $k = 5$, de acordo com definição contemplada na [Subsubseção 6.5.2.2](#). Os resultados de cada rodada de validação foram armazenados para criação da matriz de confusão.

7.1 *Fingerprinting* KNN

Nesta seção serão apresentados os resultados obtidos com a implementação do algoritmo KNN. Os resultados são apresentados por ponto específico do mapa e também agrupados por corredor e corredor específico.

7.1.1 Representação Cromatográfica

A partir dos dados coletados, uma representação cromatográfica foi obtida utilizando uma escala de cores e atribuindo uma cor para cada um dos pontos de medição do experimento, de acordo com o valor médio do RSSI (6 *beacons*) de cada um dos pontos. Na [Figura 23](#), é apresentado o mapa de localização dos *beacons* e o mapa de calor do experimento D4, no qual os *beacons* estão concentrados em apenas um corredor. Este tipo de representação evidencia os locais onde a potência do sinal chega mais forte ou com mais atenuação. Pelo mapa do experimento D4, fica clara a área de maior intensidade do sinal, no corredor de concentração dos dispositivos. As colunas do mapa de calor representam os três corredores onde estão os pontos 1 a 6, 7 a 12, e 13 a 18, e cada célula representa um dos pontos.

7.1.2 Resultados dos Experimentos

Os resultados gerais do KNN, agrupados por experimento, com detalhamento por ponto, por corredor e corredor específico, podem ser visualizados na [Tabela 9](#). Nesta tabela, os percentuais apresentados nas quatro últimas colunas correspondem a taxa de acerto para as granularidades por ponto, corredor e região (problema binário), e na última coluna, a média dos três resultados. A taxa de acerto é o percentual de vezes em que o algoritmo analisado acertou exatamente o ponto onde foi feita a medição.

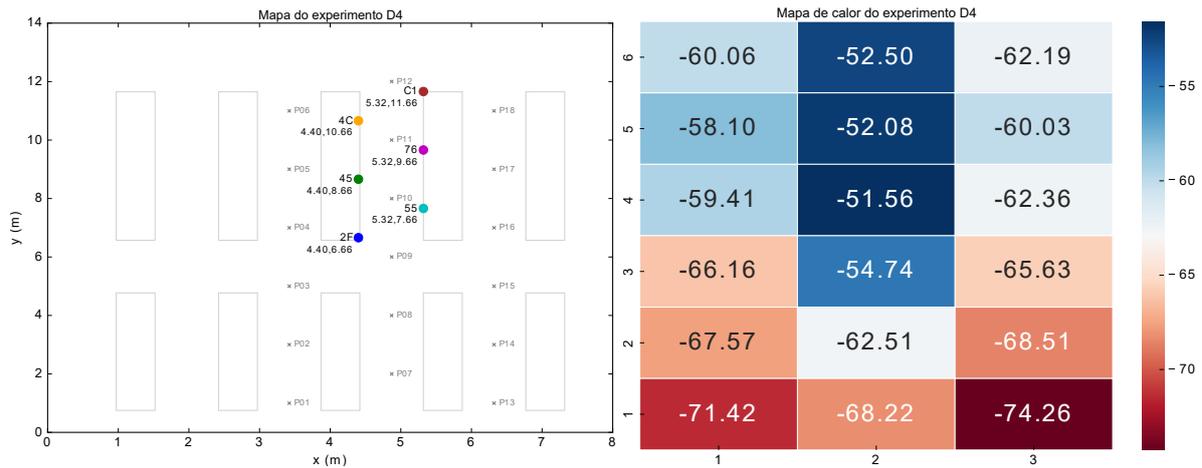


Figura 23 – Mapa de posicionamento dos beacons (esquerda) e mapa de calor do experimento D4

Na avaliação por ponto, a variável de resposta é o *label* do ponto, numerados de 1 a 18. No caso da avaliação por corredor, os pontos são agrupados por corredor, totalizando 6 corredores. E no caso do problema binário, a classificação visa responder apenas se o dispositivo leitor está ou não no corredor alvo. Cada um desses tipos de avaliação é abordado com mais detalhes no decorrer desta seção.

Tabela 9 – Resultados KNN por experimento

Exp.	Amostras	Ponto (%)	Corredor (%)	Binário (%)	Média (%)
D1	17797	81.96	93.41	98.09	91.15
D2	1684	78.92	91.33	100.00	90.08
D3	1677	90.88	93.02	98.21	94.04
D4	1606	87.86	89.10	100.00	92.32
D5	1616	92.64	92.64	100.00	95.09
D6	1511	92.72	94.77	99.34	95.61
D7	1380	97.10	97.10	100.00	98.07
D8	2057	85.48	95.62	98.54	93.21

Os melhores resultados para esta técnica ocorrem nos experimentos com os *beacons* fixados nas prateleiras (D3 ao D8), se considerarmos a média dos 3 níveis de granularidade (última coluna da tabela). Este comportamento se deve ao fato de que as prateleiras atuam como obstáculos, que geram atenuação do sinal para os pontos que não possuem ângulo de visão direta para os *beacons* e, dessa forma, geram maior dissimilaridade entre a intensidade recebida nos pontos que estão em corredores distintos. Quando os *beacons* foram fixados no teto, a amplitude do ângulo de visão direta entre os *beacons* e os pontos ficou maior, minimizando estas diferenças.

7.1.3 Experimento D4

Esse experimento consistiu em concentrar os *beacons* em um espaço pequeno, dentro de um corredor de interesse (conforme Figura 23), a fim de identificar uma configuração de área possível com um percentual de acerto próximo ou igual a 100%. A matriz de confusão resultante do experimento D4 pode ser visualizada na Figura 24. Nesse resultado o nível de detalhamento é por ponto, em cada um dos 18 pontos de coleta do mapa, e o percentual de acertos obtidos foi de aproximadamente 88%.

Valores previstos	1	118 7.35%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	118 100% 0.00%	
	2	4 0.25%	30 1.87%	0	0	0	0	15 0.93%	0	0	0	0	0	0	0	0	0	0	0	49 61.22% 38.78%
	3	7 0.44%	9 0.56%	89 5.54%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	105 84.76% 15.24%
	4	0	0	0	64 3.99%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64 100% 0.00%
	5	0	0	0	0	87 5.42%	0	0	0	0	0	0	0	0	0	0	0	0	0	87 100% 0.00%
	6	0	0	0	0	0	82 5.11%	0	0	0	0	0	0	0	0	0	0	0	0	82 100% 0.00%
	7	0	20 1.25%	0	0	0	0	60 3.74%	0	0	0	0	0	0	22 1.37%	0	0	0	0	102 58.82% 41.18%
	8	0	0	0	0	0	0	0	48 2.99%	0	0	0	0	0	0	0	24 1.49%	0	0	72 65.67% 34.33%
	9	0	0	0	24 1.49%	0	0	0	0	89 5.54%	0	0	0	0	0	0	0	0	0	113 78.76% 21.24%
	10	0	0	0	0	0	0	0	0	85 5.29%	0	0	0	0	0	0	0	0	0	85 100% 0.00%
	11	0	0	0	0	0	0	0	0	0	88 5.48%	0	0	0	0	0	0	0	0	88 100% 0.00%
	12	0	0	0	0	0	0	0	0	0	0	100 6.23%	0	0	0	0	0	0	0	100 100% 0.00%
	13	14 0.87%	0	0	0	0	0	0	0	0	0	0	86 5.35%	0	0	0	0	0	0	100 86.00% 14.00%
	14	7 0.44%	10 0.62%	0	0	0	0	0	0	0	0	0	0	66 4.11%	0	0	0	0	0	83 79.52% 20.48%
	15	7 0.44%	0	0	0	0	0	0	16 1.00%	0	0	0	0	0	0	100 6.23%	0	0	0	123 81.30% 18.70%
	16	0	0	0	0	0	0	0	16 1.00%	0	0	0	0	0	0	0	48 2.99%	0	0	64 75.00% 25.00%
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92 5.73%	0	0	92 100% 0.00%
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79 4.92%	0	79 100% 0.00%
	157 75.16% 24.84%	69 43.48% 56.52%	89 100% 0.00%	88 72.73% 27.27%	87 100% 0.00%	82 100% 0.00%	75 80.00% 20.00%	80 80.00% 20.00%	89 100% 0.00%	85 100% 0.00%	88 100% 0.00%	100 100% 0.00%	86 100% 0.00%	88 75.00% 25.00%	100 100% 0.00%	72 66.67% 33.33%	92 100% 0.00%	79 100% 0.00%	1606 87.85% 12.14%	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
	Valores reais																			

Figura 24 – Matriz de confusão do experimento D4 por ponto - KNN

Ao analisar a correlação dos dados, foi possível verificar que alguns erros ocorrem entre pontos próximos ou estão dentro dos mesmos corredores. São os casos das instâncias (previsto/real) 2/1, com total de 4 erros, 3/1 (7 erros) e 3/2 (9 erros), todos próximos da diagonal principal, totalizando 20 erros dentro do mesmo corredor. Desse modo, mudando a granularidade para o corredor, é possível conseguir esta melhora no resultado final, de 87.86% para 89.10%, referente a esses 20 erros, conforme pode ser visto na [Figura 25](#).

Valores previstos	1-3	257 16.00%	0	15 0.93%	0	0	0	272 94.49%	5.51%					
	4-6	0	233 14.51%	0	0	0	0	233 100%	0.00%					
	7-9	20 1.25%	24 1.49%	197 12.27%	0	22 1.37%	24 1.49%	287 68.64%	31.36%					
	10-12	0	0	0	273 17.00%	0	0	273 100%	0.00%					
	13-15	38 2.37%	0	16 1.00%	0	252 15.69%	0	306 82.35%	17.65%					
	16-18	0	0	16 1.00%	0	0	219 13.64%	235 93.19%	6.81%					
		315 81.59%	257 90.66%	244 80.74%	273 100%	274 91.97%	243 90.12%	1606 89.10%	18.41%	9.34%	19.26%	0.00%	8.03%	9.88%
	1-3	4-6	7-9	10-12	13-15	16-18	Valores reais							

Figura 25 – Matriz de confusão do experimento D4 por corredor - KNN

Ao verificar as correlações desse novo nível de detalhamento, é perceptível que os pontos de 10 a 12, pertencentes ao corredor onde estão concentrados os *beacons* nesse experimento, possuem uma taxa de acerto de 100%, visto que a intensidade do sinal é mais forte nestes pontos que possuem uma visão direta entre todos os *beacons* e os pontos de coleta. Conseqüentemente, todos os erros estão correlacionados no universo de pontos fora do corredor alvo desse experimento. Desse modo, foi feito um novo agrupamento dos resultados para um problema binário, conforme [Figura 26](#).

Nessa nova granularidade foi possível obter uma taxa de acerto de 100% nos experimentos realizados. Essa abordagem é válida quando o objetivo final é apenas saber se o dispositivo leitor está ou não neste corredor (pontos 10 a 12). Esse experimento, apesar de mínimo, sendo resolvido com uma precisão de 100% pode ser aplicável para um estabelecimento



Figura 26 – Matriz de confusão do experimento D4 - corredor específico - KNN

que planeje implantar uma solução por etapas. Essa configuração pode ser aplicada, por exemplo, em um corredor do supermercado onde o lojista queira anunciar a propaganda de um produto daquela seção apenas para os clientes que já estão no local.

7.2 Fingerprinting MLP

A apresentação dos resultados do MLP segue o mesmo padrão da técnica anterior, tabelando os experimentos por agrupamentos, conforme Tabela 10. De modo geral, os resultados do MLP foram melhores que o KNN, e os melhores resultados foram obtidos nos experimentos D7 e D2, sendo D2 o que possui a maior taxa de dispersão (soma das distâncias) entre os *beacons*.

Tabela 10 – Resultados MLP por experimento

Exp.	Amostras	Ponto (%)	Corredor (%)	Binário (%)	Média (%)
D1	17797	88.64	94.47	97.75	93.62
D2	1684	94.77	96.38	100.00	97.05
D3	1677	89.74	93.56	98.63	93.98
D4	1606	88.92	90.66	100.00	93.19
D5	1616	90.16	92.08	98.45	93.56
D6	1511	93.05	96.16	99.21	96.14
D7	1380	100.00	100.00	100.00	100.0
D8	2057	96.21	96.21	97.28	96.57

No detalhamento por ponto, a média foi de 92.69% e no agrupamento por corredor, todos os resultados foram maiores que 90%, com uma média de 94.94% de taxa de acerto, e no caso binário, a média dos resultados foi de 98.91%, sendo o pior resultado acima de 97%. Esse resultado se mostrou importante para identificação de um limite de granularidade no qual se

possa trabalhar para uma precisão próxima de 100% com os *beacons* BLE utilizados nos experimentos. Assim, verifica-se um potencial de sucesso para implantação de soluções em ambientes similares para o problema binário e por zona.

Para o experimento D4 no problema binário, obteve 100% de acertos, assim como o resultado alcançado no classificador KNN. E nos demais agrupamentos, os resultados também foram melhores, 88.92% contra 87.86% (por ponto), e 90.66% contra 89.10% (por corredor). O pior caso aconteceu no experimento D4. Esses resultados estão detalhados abaixo, na matriz de confusão, com agrupamento por ponto na [Figura 27](#), com agrupamento por corredor na [Figura 28](#) e com agrupamento binário na [Figura 29](#).

Valores previstos	1	141 8.78%	10 0.62%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151 83.36% 5.62%	
	2	9 0.56%	40 2.49%	0	0	0	0	15 0.93%	0	0	0	0	0	0	0	0	0	0	0	64 82.50% 87.50%
	3	0	9 0.56%	89 5.54%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98 80.82% 9.18%
	4	0	0	0	76 4.73%	0	0	0	16 1.00%	17 1.06%	0	0	0	0	0	0	0	0	0	109 69.72% 80.28%
	5	0	0	0	0	87 5.42%	0	0	0	0	0	0	0	0	0	0	0	0	0	87 100% 0.00%
	6	0	0	0	0	0	82 5.11%	0	0	0	0	0	0	0	0	0	0	17 1.06%	0	99 82.83% 17.17%
	7	0	0	0	0	0	0	60 3.74%	0	0	0	0	0	0	0	0	0	0	0	60 100% 0.00%
	8	0	0	0	0	0	0	0	16 1.00%	0	0	0	0	0	0	8 0.50%	0	0	0	24 86.67% 13.33%
	9	0	0	0	0	0	0	0	0	72 4.48%	0	0	0	0	0	0	0	0	0	72 100% 0.00%
	10	0	0	0	12 0.75%	0	0	0	0	0	85 5.29%	0	0	0	0	0	0	0	0	97 87.63% 12.37%
	11	0	0	0	0	0	0	0	0	0	0	88 5.48%	0	0	0	0	0	0	0	88 100% 0.00%
	12	0	0	0	0	0	0	0	0	0	0	0	100 6.23%	0	0	0	0	0	0	100 100% 0.00%
	13	7 0.44%	0	0	0	0	0	0	0	0	0	0	0	86 5.35%	0	0	0	0	0	93 82.47% 17.53%
	14	0	10 0.62%	0	0	0	0	0	0	0	0	0	0	0	88 5.48%	0	0	0	0	98 89.80% 10.20%
	15	0	0	0	0	0	0	0	16 1.00%	0	0	0	0	0	0	92 5.73%	0	0	0	108 85.19% 14.81%
	16	0	0	0	0	0	0	0	32 1.99%	0	0	0	0	0	0	0	72 4.48%	0	0	104 88.23% 11.77%
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	75 4.67%	0	75 100% 0.00%
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79 4.92%	79 100% 0.00%
		157 89.81% 10.19%	69 57.97% 42.03%	89 100% 0.00%	88 86.36% 13.64%	87 100% 0.00%	82 100% 0.00%	75 80.00% 20.00%	80 80.00% 20.00%	89 80.90% 19.10%	85 100% 0.00%	88 100% 0.00%	100 100% 0.00%	86 100% 0.00%	88 100% 0.00%	100 92.00% 8.00%	72 100% 0.00%	92 81.52% 18.48%	79 100% 0.00%	1606 88.81% 11.08%
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
		Valores reais																		

Figura 27 – Matriz de confusão do experimento D4 por ponto - MLP

Valores previstos	1-3	298 18.56%	0	15 0.93%	0	0	0	313 95.21% 4.79%
	4-6	0	245 15.26%	33 2.05%	0	0	17 1.06%	295 83.05% 16.95%
	7-9	0	0	148 9.22%	0	8 0.50%	0	156 94.87% 5.13%
	10-12	0	12 0.75%	0	273 17.00%	0	0	285 95.79% 4.21%
	13-15	17 1.06%	0	16 1.00%	0	266 16.56%	0	299 88.96% 11.04%
	16-18	0	0	32 1.99%	0	0	226 14.07%	258 87.60% 12.40%
		315 94.60% 5.40%	257 95.33% 4.67%	244 60.66% 39.34%	273 100% 0.00%	274 97.08% 2.92%	243 93.00% 7.00%	1606 90.66% 9.34%
		1-3	4-6	7-9	10-12	13-15	16-18	
		Valores reais						

Figura 28 – Matriz de confusão do experimento D4 por corredor - MLP

Valores previstos	Não	1333 83.00%	0	1333 100% 0.00%
	Sim	0	273 17.00%	273 100% 0.00%
		1333 100% 0.00%	273 100% 0.00%	1606 100% 0.00%
		Não	Sim	
		Valores reais		

Figura 29 – Matriz de confusão do experimento D4 - Corredor específico - MLP

7.3 Triangulação

Os resultados da técnica de Triangulação são apresentados na forma de diagrama de caixas agrupados por experimento (eixo x) e com erro apresentado em metros (eixo y). Cada ponto em cada experimento gera como resultado a definição de uma localização que é validada com a localização real do ponto, armazenada no momento das medições. Assim, o erro é aferido, em metros, calculando a distância euclidiana entre a localização estimada e a localização real.

Nos gráficos apresentados nesta seção, os eixos x e y são as medidas do ambiente, em metros, e os demais componentes são detalhados a seguir: os pontos coloridos destacados são os *beacons*, e na legenda cada um é apresentada a distância calculada. A estrela vermelha é a posição calculada pelo algoritmo. Conforme já foi mencionado no [Capítulo 6](#), os pontos indicados com um *x* nos corredores são os pontos de coleta, sendo que o ponto destacado na cor dourada é o ponto onde foi feita a medição atual. Para cada ponto de cada experimento, é gerado um gráfico que é parte do processo de cálculo da posição estimada. Um exemplo deste gráfico é apresentado na [Figura 30](#), com o melhor caso do experimento D2, que ocorreu no ponto 06, com um erro de 0.22m.

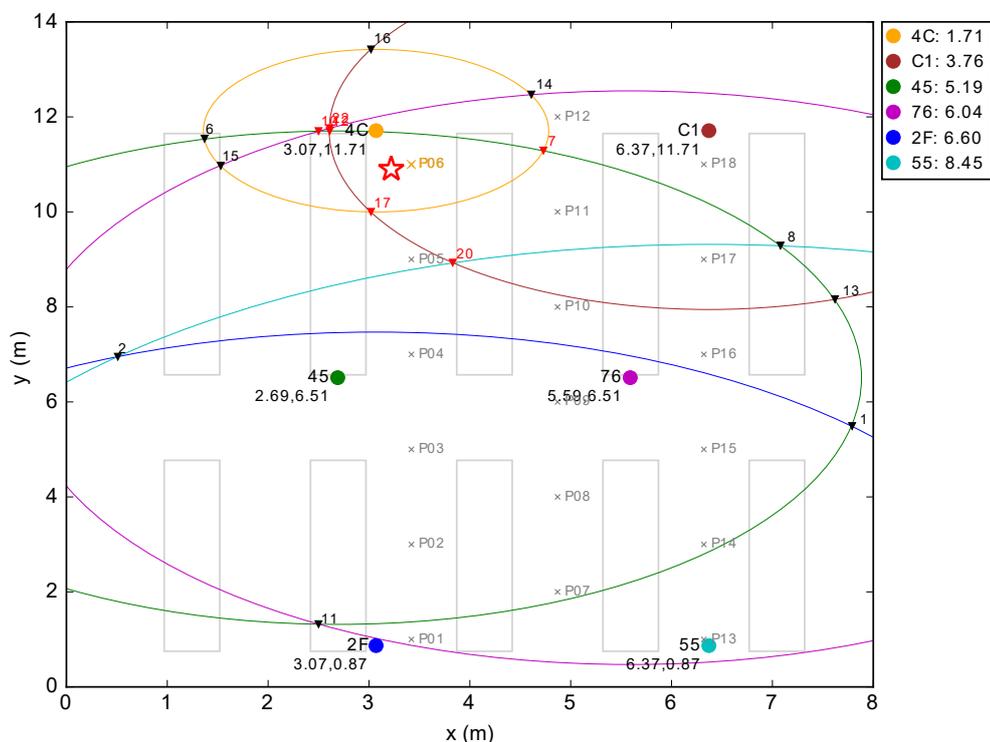


Figura 30 – Algoritmo de Triangulação

Os resultados finais, agrupados por experimento, para a técnica de Triangulação são exibidos na [Figura 31](#). Cada uma das amostras dos resultados da técnica de Triangulação são agrupadas por experimento (eixo x). E para cada experimento é possível avaliar o

centro, a dispersão e a amplitude da distribuição, com o erro dado em metros (eixo y). Em todos os experimentos o intervalo interquartil está compreendido entre 0 e 4m e na maioria dos experimentos com a mediana posicionada abaixo dos 2m.

A menor variabilidade ocorreu nos experimentos 2, 3 e 8, com um intervalo interquartil em torno de 1m. A medida de variabilidade dos dados é importante para estimativa da confiabilidade do modelo.

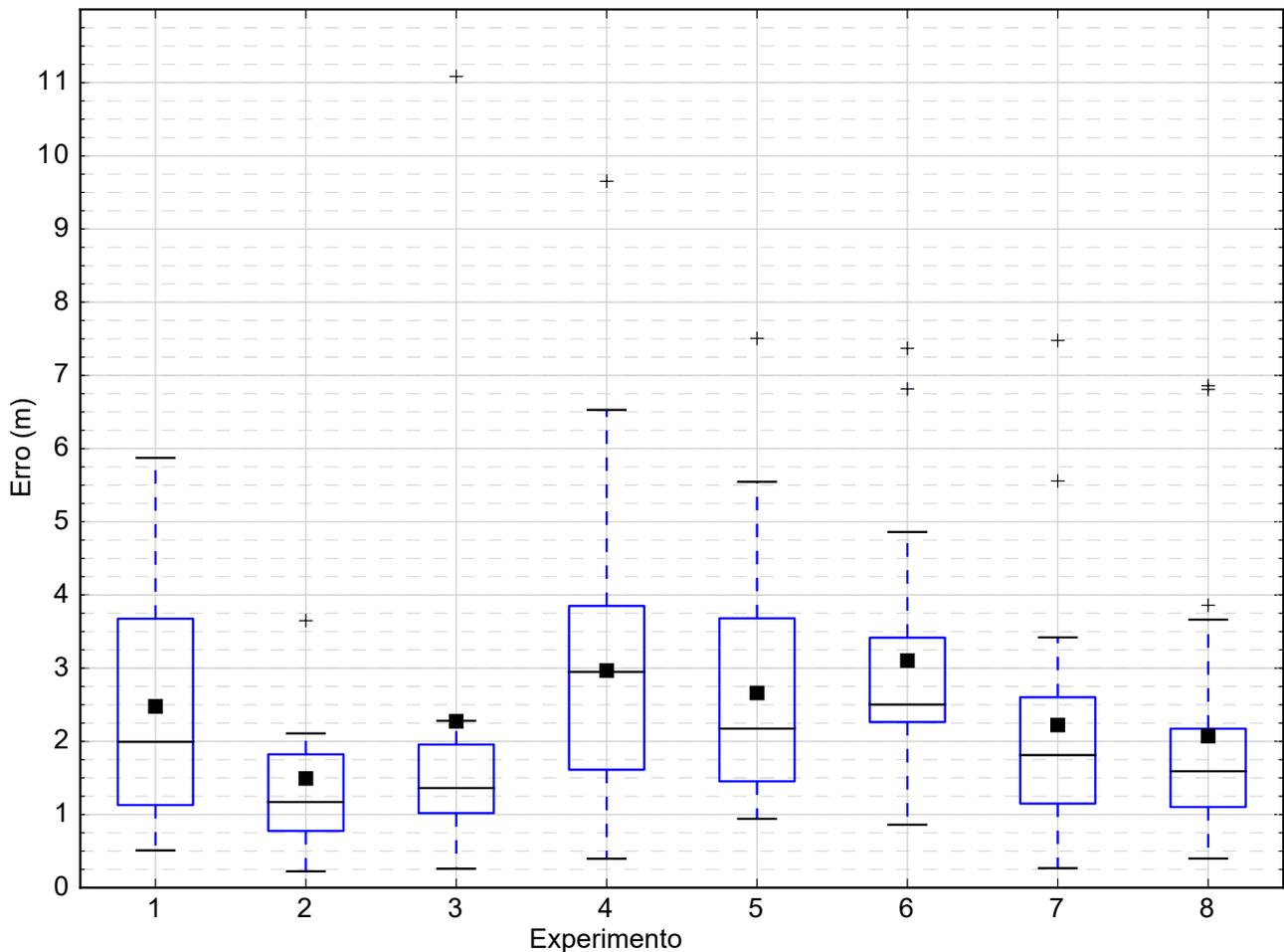


Figura 31 – Resultados dos experimentos de Triangulação

7.3.1 Baixa Dispersão do *Beacons*

Ao concentrar os *beacons* em um espaço reduzido dentro da área de interesse, como no experimento D4, é verificada uma tendência dos círculos ficarem quase coincidentes. Isso ocorre devido aos seus centros serem muito próximos, especialmente em distâncias menores que as faixas de estado de proximidade (Tabela 2). Como consequência existe uma grande quantidade de pontos de qualidade ruim inserindo um peso que pode mover a

interesse, de modo que a maioria dos pontos de interesse estejam localizados dentro da fronteira gerada por esse polígono. Essa teoria é sustentada pelos resultados obtidos no experimento D2, que possui tal configuração e obteve os menores valores de erro para os percentis 25, 50 e 75 de todos os experimentos. Em [Abreu \(2014\)](#) o autor também dá sustentação a essa análise nos seus resultados.

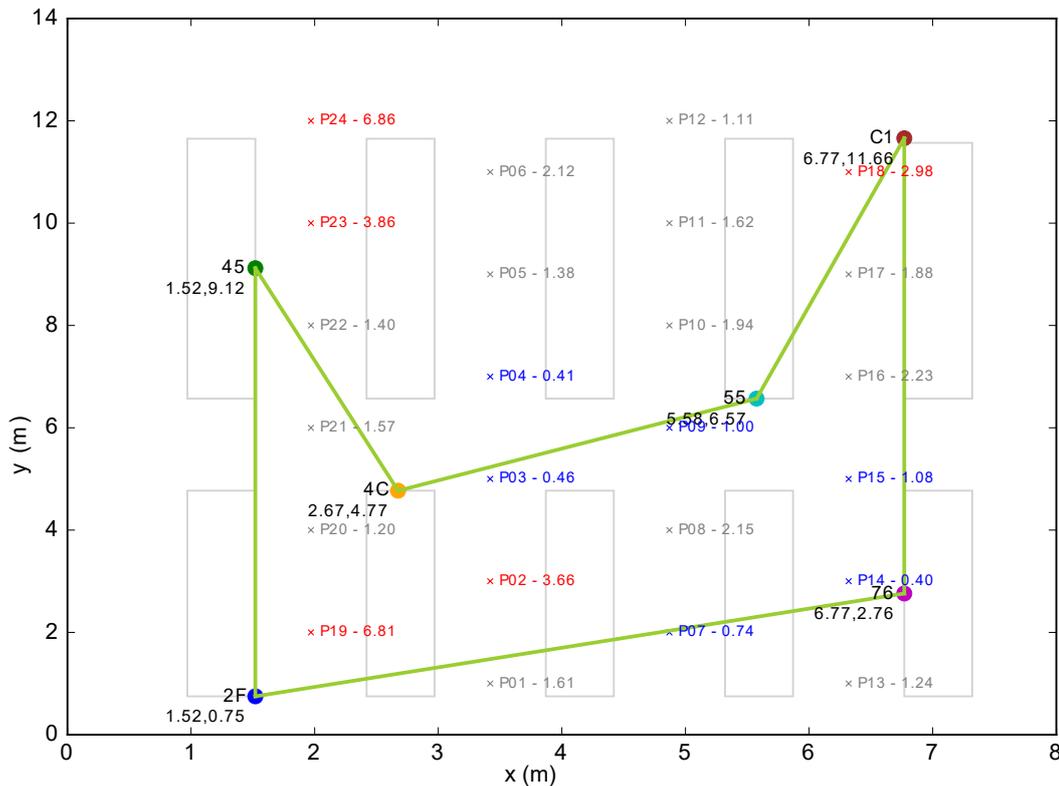


Figura 33 – Análise de posicionamento dos beacons

7.3.3 Comparação dos Resultados

Nesta seção são comparadas as técnicas avaliadas neste trabalho, além de confrontar os resultados da literatura com o presente estudo. A técnica *Fingerprinting* apresenta bons resultados, porém possui um custo maior na fase de calibração, devido à indispensabilidade da base de dados de aprendizagem. O *Fingerprinting* também apresenta a desvantagem de dependência das condições físicas, tais como umidade, temperatura e lotação de pessoas. Se na execução do algoritmo (fase *on-line*), essas condições forem muito diferentes daquelas observadas no momento da carga da base de dados de treinamento (fase *off-line*), os resultados podem divergir consideravelmente.

A Triangulação é uma técnica mais escalável, pois não necessita de uma base de dados prévia com as medições. Na fase de configuração, é necessário apenas o processo de calibração dos beacons para obtenção dos coeficientes da fórmula de cálculo da distância, conforme [Seção 6.3](#). Uma vez concluído este processo, os beacons podem ser posicionados

e utilizados. Sendo assim, esta técnica é mais adequada para implantação de larga escala ou grandes ambientes, enquanto que o *Fingerprinting* é mais adequado para pequenos ambientes.

Conforme já discorrido no capítulo dos Trabalhos Relacionados, na [Seção 3.3](#), as comparações dos resultados dos trabalhos devem ser ponderadas com alguns critérios que os influenciam, não se limitando a:

- a. Hardware utilizado (*beacon* e dispositivo leitor);
- b. Quantidade de *beacons*;
- c. Configurações dos *beacons*, principalmente potência e intervalo de transmissão;
- d. Tamanho da área de cobertura.

Na [Tabela 11](#) são apresentados os resultados das principais referências do estado da arte e do presente estudo. No trabalho atual foram considerados apenas os experimentos que foram propostos com foco na técnica de Triangulação (Exp. D1, D2, D3, D7, D8). Isto porque o erro foi calculado como nos trabalhos comparados, ou seja, com a distância euclidiana, em metros, entre o ponto calculado e o ponto real da medição. O melhor resultado, de 1.20m, se refere ao experimento D2 e o pior, de 2.00m pertence ao experimento D1.

Tabela 11 – Comparação dos Resultados

Autor	Área (m^2)	Tecnol.	Qtde	Técnica	Erro Médio (m)
Abreu (2014)	145	Wi-Fi	4	Triangulação	1.80
Larsson (2015)	35	BLE	6	<i>Fingerprinting</i>	1.90 (por pontos) e 1.30 (por rotas)
Kriz (2016)	2236	BLE	17	<i>Fingerprinting</i>	2.00
Este trabalho	112	BLE	6	Triangulação	Entre 1.20 e 2.00 (Exp. D1, D2, D3, D7, D8)

Nessa tabela, a quantidade (Qtde) se refere ao número de roteadores (*Wi-Fi*) ou de *beacons* (BLE) utilizados. É importante destacar alguns pontos relevantes. No trabalho de [Larsson \(2015\)](#), o intervalo de transmissão utilizado foi de 500ms. Enquanto que em [Kriz, Maly e Kozel \(2016\)](#) foi de 100ms e está sendo considerado o resultado com utilização dos *beacons* apenas e não o modelo híbrido (incluindo *Wi-Fi*).

O resultado alcançado no trabalho ([KRIZ; MALY; KOZEL, 2016](#)) com a utilização apenas dos *beacons* foi um erro de 2m na mediana e intervalo interquartil 1-4m. Como pode ser visto na [Figura 31](#), para a técnica de Triangulação, todos os experimentos do presente trabalho possuem um intervalo menor e os experimentos D2, D3, D7 e D8 ainda possuem a mediana do erro menor.

8 Conclusão

Analisando os resultados do trabalho, verifica-se boas oportunidades para desenvolvimento de um modelo de navegação *indoor* sensível a obstáculos de grande relevância para ambientes de comércio varejista. Os *beacons* são insumos viabilizadores importantes nesse cenário. A evolução constante desses dispositivos, aliada ao desenvolvimento e aperfeiçoamento de aplicações que fazem uso dos seus dados de contexto, é um indicador de que essa tecnologia poderá se consolidar para atendimento das demandas de localização e navegação *indoor*.

Importantes desafios ainda existem para alcançar confiabilidade, robustez e exatidão nos sistemas de localização *indoor*. Os *beacons* com tecnologia Bluetooth Low Energy são muito suscetíveis à interferência de outras tecnologias de transmissão sem fio que utilizam a mesma faixa do espectro de frequências, o que dificulta alcançar ordem de grandeza de centímetros, relevante em alguns tipos de aplicações. Porém, os resultados obtidos no presente trabalho corroboram que é possível encontrar uma configuração específica para a área de interesse onde os resultados possam estar na faixa de 100% de exatidão, conforme resposta do nível de granularidade binário, do classificador KNN, para os experimentos D2, D4, D5 e D7.

8.1 Contribuições

Este trabalho apresentou contribuições importantes para a área de localização *indoor* utilizando *beacons* com a tecnologia *Bluetooth Low Energy* e explorou conceitos importantes para definição da melhor localização dos *beacons* para cobertura da área de interesse. Este estudo também abordou a questão da granularidade para a subdivisão da área de interesse por meio de agrupamento de pontos para investigação da melhoria no grau de exatidão.

As contribuições mais relevantes deste trabalho se estendem nos contextos acadêmico, tecnológico, mercadológico e social, os quais serão detalhados nesta seção. A publicação do artigo [Silva e Almeida \(2017\)](#) sugere a importância das ferramentas de localização *indoor* e a relevância acadêmica dos resultados obtidos nesse trabalho.

Muitos autores compartilham o entendimento de que os *beacons*, dispositivos inovadores e transformadores, estão habilitando a utilização em larga escala da Internet das Coisas no dia a dia das pessoas. [Nowodzinski, Łukasik e Puto \(2016\)](#) classificam os *beacons*, juntamente com os aplicativos móveis, como os blocos construtores da Internet das Coisas no setor de varejo. Nesse cenário de inegável importância tecnológica dos *beacons*, o presente

trabalho contribui, no contexto tecnológico, ao explorar a capacidade de localização *indoor*, uma das principais habilidades dos *beacons*, para demonstrar, de forma empírica, resultados relevantes de algoritmos para oferecer sustentação a um sistema de posicionamento interno.

Os entendimentos de que o conhecimento é a mais valiosa fonte de vantagem competitiva (BOLISANI; SCARSO, 1999), e o ativo organizacional mais importante (MYBURGH, 2000), foi amplamente difundido no últimos anos e reforça a importância de uma cultura arraigada na consolidação de bases de informações consistentes sobre o comportamento do cliente no ambiente de comércio varejista tradicional, de forma similar ao âmbito do *e-commerce*. Nesse contexto de mercado, o trabalho mostra o caminho a ser trilhado para alicerçar tais bases, destacando a importância e mostrando formas eficientes de prover contexto às aplicações para coletar e processar dados relevantes sobre a jornada do cliente no ambiente de comércio.

No contexto da convivência social, os resultados deste trabalho contribuem para o desenvolvimento de novas aplicações e atualização de aplicações existentes para torná-las mais inteligentes com a incorporação de informações de contexto. Com esse novo paradigma de interação com os objetos do mundo real, as aplicações trazem mais produtividade na realização de demandas cotidianas como rota de compras, localização e navegação em ambientes internos de grande circulação de pessoas como *shoppings* e lojas.

8.2 Trabalhos Futuros

Para melhoria das soluções apresentadas neste trabalho, é sugerida a utilização dos *beacons* UWB (*Ultra Wide Band*). Esses dispositivos já são uma realidade atualmente. Os estudos e projetos estão crescendo, ainda que em estágio incipiente. Tais dispositivos podem ser testados para melhoria da exatidão das técnicas de localização *indoor*, uma vez que eles podem alcançar uma precisão de centímetros. Isto porque a tecnologia UWB utiliza baixa potência de transmissão (-41.3 dBm/MHz) evitando interferência de outras tecnologias sem fio que utilizam a mesma faixa do espectro. Os *beacons* UWB também possuem desvantagens, principalmente com relação ao custo e interoperabilidade. Com relação a esse último fator, já existem *beacons* com *dual* radio, incluindo no mesmo dispositivo os dois padrões de comunicação: *Bluetooth Low Energy* (IEEE 802.15.4) e *Ultra Wide Band* (IEEE 802.15.3), o que possibilita contornar essa restrição. Desta forma, é importante estudar a viabilidade desses dispositivos para demandas correlatas àquelas estudadas no atual trabalho, avaliando quesitos positivos como exatidão e precisão, sem esquecer de tratar questões como custo e interoperabilidade.

Este trabalho cria uma importante gama de possibilidades de aplicações que o complementem. A partir da disponibilização de uma ferramenta de navegação *indoor* em um

estabelecimento comercial, sua utilização pelos clientes caracteriza uma proximidade importante e um grande poder de influência no comportamento de compras do consumidor.

Com a implementação deste modelo, informações importantes para estratégias de *marketing* baseadas em localização do cliente dentro da loja serão disponibilizadas. É possível aumentar as taxas de conversão nas sugestões de compras, ao disponibilizar 3 contextos importantes para personalização da indicação: (a) o perfil do cliente que está autenticado no carrinho de compras inteligente, (b) sua localização, (c) a lista de compras atual.

Entretanto esses dispositivos inteligentes não serão utilizados somente para aferir a localização exata dos clientes e os departamentos da loja por ele visitados. A partir dessa localização, quando o usuário está dentro do *geofence* (perímetro virtual) definido, são disparadas ações programadas. Essas ações podem ser simples envios de notificações ou interações mais elaboradas, como *marketing* baseado em localização e perfil, de modo que a aplicação, ao notar a aproximação dos clientes, avalia o seu perfil e oferece sugestões de produtos personalizados.

Os produtos sugeridos podem ser definidos por meio de uma abordagem de inteligência artificial, utilizando *Big Data*:

1. Sugestões baseadas em cupons de compras de outros clientes, correlação entre produtos: "80% dos clientes que compram produto x também compram o produto y";
2. Sugestões baseadas na lista de compras do próprio cliente: "Você tem na sua lista pão, será que não está esquecendo da manteiga?";
3. Sugestões baseadas no histórico de compras do cliente: "Tem muito tempo que você não compra arroz".

Outro importante viés comercial que pode ser explorado pelo estabelecimento de varejo está relacionado com as parcerias comerciais. O lojista pode oferecer produtos baseados em acordos comerciais realizados com os fabricantes dos itens, promoções patrocinadas, entre outras. A integração da aplicação com o *backoffice* do estabelecimento, especialmente o cupom fiscal, permite ainda a apuração das taxas de conversão das indicações feitas.

8.3 Considerações Finais

Os objetivos propostos no início deste projeto incluíam (1) o modelo de representação do ambiente físico, (2) a coleta de dados utilizando o dispositivo previamente disponível, (3) a aplicação e avaliação das técnicas *Fingerprinting* e Triangulação, e (4) a avaliação e comparação das técnicas. Esses objetivos foram alcançados no desenvolvimento do trabalho, conforme segue. O modelo de representação foi elaborado a partir da planta e da localização do mobiliário no ambiente de testes (biblioteca), conforme [Subseção 6.1.3](#).

A partir do planejamento, os experimentos foram executados no ambiente e os dados foram coletados utilizando o dispositivo disponível com um aplicativo que foi adaptado para armazenar os dados no formato compatível, conforme [Seção 6.4](#). Os algoritmos das técnicas de *Fingerprinting* (KNN e MLP) e Triangulação foram implementados conforme [Seção 6.5](#) e testados utilizando os dados coletados. A avaliação dos resultados e a comparação das técnicas implementadas foram discutidas no [Capítulo 7](#).

As duas hipóteses propostas neste trabalho foram (1) alcançar performance melhor ou similar aos resultados disponíveis na literatura e (2) tempo de latência adequado. Conforme [Subseção 7.3.3](#), os resultados alcançados, considerando a técnica de Triangulação, foram na mesma faixa dos trabalhos disponíveis na literatura, ou seja, mediana de erro entre 1 e 2m (distância euclidiana). Para o *Fingerprinting* (KNN e MLP), o modelo de apresentação dos resultados adotado, por se tratar de classificadores, foi a matriz de confusão, porém os resultados dos trabalhos analisados não apresentam uma forma equivalente, o que impossibilitou a comparação. Sobre a latência, o tempo de coleta de dados verificado na [Subseção 6.4.2](#) indica a necessidade de 15 segundos de coleta para que a média do valor do RSSI alcance um ponto de equilíbrio. Este tempo é a restrição apenas para a primeira resposta do algoritmo depois que um dispositivo adentra na área de interesse. Após a primeira resposta, os demais cálculos utilizam os dados históricos disponíveis, sendo suficiente o tempo de processamento do algoritmo, que nos testes não passaram de 2 segundos.

A presente pesquisa, utilizando equipamentos nacionais e recentemente lançados, apresentou resultados similares aos trabalhos divulgados na literatura que utilizam os *beacons* de fabricantes globais e de reconhecida excelência e pioneirismo no mercado. Com experimentos conduzidos em ambiente similar (biblioteca) ao estudo de caso apresentado (supermercado), a possibilidade de replicação dos resultados no ambiente de comércio varejista pode ser vista com otimismo.

Nos últimos anos, principalmente após o lançamento do *iBeacon* pela Apple em 2014 e do *Eddystone* pela Google em 2015, se observa a grande importância que a área de localização *indoor* tem alcançado, seja por meio de número de trabalhos publicados, seja pelo número de patentes registradas, ou pelo desenvolvimento das novas tecnologias que suportarão as principais aplicações. As estimativas para o futuro, de modo equivalente, são muito promissoras, de acordo com previsão da pesquisa [Unacast \(2016\)](#), até 2020 serão 400 milhões de *beacons* implantados em todo o mundo. Essa disponibilidade de sensores inteligentes transformarão a forma como as aplicações interagem com as coisas do mundo real e, conseqüentemente, a forma como as pessoas interagem com estas aplicações.

Referências

ABREU, R. **Localização indoor em ambientes inteligentes**. Dissertação (Mestrado) — Universidade do Minho, 2014. Citado 2 vezes nas páginas 26 e 64.

AL-FUQAHA, A. et al. Internet of things: A survey on enabling technologies, protocols, and applications. **IEEE Communications Surveys & Tutorials**, IEEE, v. 17, n. 4, p. 2347–2376, 2015. IoT Survey. Citado 2 vezes nas páginas 3 e 4.

ALARIFI, A. et al. Ultra wideband indoor positioning technologies: Analysis and recent advances. **Sensors**, v. 16, n. 5, 2016. ISSN 1424-8220. Disponível em: <<http://www.mdpi.com/1424-8220/16/5/707>>. Citado na página 18.

ALLIANCE, I. I. **Bringing Mobile Positioning Indoors**. 2017. Acessado em: 06 Jun 2017. Disponível em: <<http://inlocationalliance.org>>. Citado na página 3.

ANDREW, G.; GAO, J. Scalable training of l_1 -regularized log-linear models. In: ACM. **Proceedings of the 24th international conference on Machine learning**. [S.l.], 2007. p. 33–40. Citado na página 49.

APPLE. **Getting Started with iBeacon**. [S.l.], 2014. Acessado em: 25 Out 2017. Disponível em: <<https://developer.apple.com/ibeacon/Getting-Started-with-iBeacon.pdf>>. Citado na página 15.

ARGENOX, T. **A BLE Advertising Primer**. 2018. Acessado em: 25 Jun 2018. Disponível em: <<http://www.argenox.com/a-ble-advertising-primer/>>. Citado na página 11.

AUGUR, H. **The future of Bluetooth beacons is the future of the IoT**. 2017. Acessado em: 12 Jan 2018. Disponível em: <<https://blog.bluetooth.com/beacons-and-bluetooth-bridging-the-physical-and-digital-worlds>>. Citado na página 13.

BOLISANI, E.; SCARSO, E. Information technology management: a knowledge-based perspective. **Technovation**, Elsevier, v. 19, n. 4, p. 209–217, 1999. Citado na página 67.

BOURKE, P. **Circles and Spheres**. 1997. Acessado em: 18 Set 2017. Disponível em: <<http://paulbourke.net/geometry/circlesphere/>>. Citado 2 vezes nas páginas ix e 21.

BRACHMANN, F. **In 2018 who are the best beacon manufacturers out there?** 2018. Acessado em: 06 Jun 2018. Disponível em: <<http://blog.beeem.co/2018/01/in-2018-who-are-the-best-beacon-manufacturers-out-there/>>. Citado na página 13.

BRIDOUX, N. **android-beacon-scanner**. 2018. Acessado em: 05 Jan 2018. Disponível em: <<https://github.com/Bridouille/android-beacon-scanner>>. Citado na página 45.

CAI, D. A retail application based on indoor location with grid estimations. In: IEEE. **Computer, Information and Telecommunication Systems (CITS), 2014 International Conference on**. [S.l.], 2014. p. 1–4. Citado na página 25.

CHEN, X.; LI, Y.; HU, T. Solving the supermarket shopping route planning problem based on genetic algorithm. In: IEEE. **Computer and Information Science (ICIS), 2015 IEEE/ACIS 14th International Conference on**. [S.l.], 2015. p. 529–533. Citado na página 25.

CHO, H. et al. Measuring a distance between things with improved accuracy. **Procedia Computer Science**, Elsevier, v. 52, p. 1083–1088, 2015. Citado 2 vezes nas páginas 18 e 26.

COMPTON, K.; BUTANEY, V. **Enabling the Internet of Everything: Cisco's lot architecture**. [S.l.], 2015. Disponível em: <<http://www.slideshare.net/Cisco/enabling-the-internet-of-everything-ciscos-iot-architecture>>. Acesso em: 02 Nov 2015. Citado na página 8.

COSTA, M. **Utilizando Beacons com Android**. 2015. Acessado em: 22-Maio-2018. Disponível em: <<http://redspark.io/utilizando-beacons-com-android/>>. Citado na página 15.

COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE transactions on information theory**, IEEE, v. 13, n. 1, p. 21–27, 1967. Citado 2 vezes nas páginas 22 e 23.

DEV, T. **Bluetooth Low Energy proximity-beacon Dual IP67: Especificações técnicas**. [S.l.], 2018. Citado na página 42.

FARAGHER, R.; HARLE, R. Location fingerprinting with bluetooth low energy beacons. **IEEE journal on Selected Areas in Communications**, IEEE, v. 33, n. 11, p. 2418–2428, 2015. Citado na página 15.

FLETCHER, R. **Practical Methods of Optimization**. Second. New York, NY, USA: John Wiley & Sons, 1987. Citado na página 49.

GAIKWAD, P. P.; GABHANE, J. P.; GOLAIT, S. S. 3-level secure kerberos authentication for smart home systems using iot. In: IEEE. **Next Generation Computing Technologies (NGCT), 2015 1st International Conference on**. [S.l.], 2015. p. 262–268. Citado na página 8.

GAST, M. S. **Building Applications with iBeacon: Proximity and Location Services with Bluetooth Low Energy**. 1st. ed. [S.l.]: O'Reilly Media, Inc., 2014. ISBN 1491904577, 9781491904572. Citado 2 vezes nas páginas 16 e 18.

GOOGLE. **Advanced Patent Search**. 2017. Acessado em: 08 Jul 2017. Disponível em: <https://www.google.com/advanced_patent_search>. Citado na página 3.

GOOGLE. **Eddystone format**. 2018. Acessado em: 22/05/2018. Disponível em: <<https://developers.google.com/beacons/eddystone>>. Citado na página 15.

GOOGLE. **Physical Web**. 2018. Acessado em: 16 Mai 2017. Disponível em: <<https://google.github.io/physical-web/>>. Citado na página 14.

GOSWAMI, S. **Indoor Location Technologies**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 1.

GROOT, T. **List of the 9 biggest beacon manufacturers**. 2014. Acessado em: 10 Dez 2017. Disponível em: <<https://www.nodesagency.com/list-9-biggest-beacon-manufacturers/>>. Citado na página 13.

HAVERINEN, J.; KEMPPAINEN, A. Global indoor self-localization based on the ambient magnetic field. **Robotics and Autonomous Systems**, Elsevier, v. 57, n. 10, p. 1028–1035, 2009. Citado na página 18.

HAYKIN, S. S. **Neural networks and learning machines**. [S.l.]: Pearson Upper Saddle River, NJ, USA, 2009. v. 3. Citado 2 vezes nas páginas 21 e 22.

HE, S.; CHAN, S. H. G. Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons. **IEEE Communications Surveys Tutorials**, v. 18, n. 1, p. 466–490, Firstquarter 2016. ISSN 1553-877X. Citado na página 18.

HERMERSDORF, M. **Location and Position Measurements**. [S.l.], 2007. v. 2007. Disponível em: <http://metrology.tkk.fi/courses/S-108.4010/2007/position_measurements_MHe_2007.pdf>. Acesso em: 11 Jun 2018. Citado na página 16.

JIN, G. yao; LU, X. yi; PARK, M.-S. An indoor localization mechanism using active rfid tag. In: **IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)**. [S.l.: s.n.], 2006. v. 1, p. 4 pp.–. Citado 2 vezes nas páginas 1 e 18.

KANEKO, Y.; YADA, K. Fractal dimension of shopping path: Influence on purchase behavior in a supermarket. **Procedia Computer Science**, Elsevier, v. 96, p. 1764–1771, 2016. Citado na página 24.

KHOLOD, M. et al. The influence of shopping path length on purchase behavior in grocery store. In: _____. **Knowledge-Based and Intelligent Information and Engineering Systems: 14th International Conference, KES 2010, Cardiff, UK, September 8-10, 2010, Proceedings, Part III**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 273–280. ISBN 978-3-642-15393-8. Disponível em: <https://doi.org/10.1007/978-3-642-15393-8_31>. Citado 2 vezes nas páginas 18 e 24.

KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA. **Ijcai**. [S.l.], 1995. v. 14, n. 2, p. 1137–1145. Citado na página 49.

KRIZ, P.; MALY, F.; KOZEL, T. Improving indoor localization using bluetooth low energy beacons. **Mobile Information Systems**, Hindawi, v. 2016, 2016. Citado 3 vezes nas páginas 18, 27 e 65.

LARSSON, J. **Distance Estimation and Positioning Based on Bluetooth Low Energy Technology**. Dissertação (Mestrado) — Kth Royal Institute of Technology, Stockholm, Sweden, 2015. Citado 4 vezes nas páginas 18, 27, 38 e 65.

LI, J. et al. An indoor ultrasonic positioning system based on toa for internet of things. **Mobile Information Systems**, Hindawi, v. 2016, 2016. Citado na página 18.

LIU, H. et al. Survey of wireless indoor positioning techniques and systems. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 37, n. 6, p. 1067–1080, Nov 2007. ISSN 1094-6977. Citado na página 1.

LOUREIRO, A. A. F. et al. Computação ubíqua ciente de contexto: Desafios e tendências. **27o Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**, p. 99–149, 2009. Citado 2 vezes nas páginas 7 e 8.

MANYIKA, J. et al. **Big data: The next frontier for innovation, competition, and productivity**. 2011. Disponível em: <http://www.mckinsey.com/Insights/MGI/Research/Technology/_and_Innovation/Big_data_The_next_frontier_for_innovation>. Citado na página 8.

MYBURGH, S. The convergence of information technology & information management. **Information Management Journal**, v. 34, n. 2, p. 4–16, 04 2000. Citado na página 67.

NEWMAN, N. Apple ibeacon technology briefing. **Journal of Direct, Data and Digital Marketing Practice**, v. 15, n. 3, p. 222–225, Jan 2014. ISSN 1746-0174. Disponível em: <<https://doi.org/10.1057/dddmp.2014.7>>. Citado 2 vezes nas páginas 12 e 13.

NIRJON, S. et al. Coin-gps: Indoor localization from direct gps receiving. In: **Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services**. New York, NY, USA: ACM, 2014. (MobiSys '14), p. 301–314. ISBN 978-1-4503-2793-0. Disponível em: <<http://doi.acm.org/10.1145/2594368.2594378>>. Citado na página 17.

NOWODZINSKI, P.; ŁUKASIK, K.; PUTO, A. Internet of things (iot) in a retail environment. the new strategy for firm's development. **European Scientific Journal, ESJ**, v. 12, n. 10, 2016. Citado 3 vezes nas páginas 13, 17 e 66.

OREL, F. D.; KARA, A. Supermarket self-checkout service quality, customer satisfaction, and loyalty: Empirical evidence from an emerging market. **Journal of Retailing and Consumer Services**, Elsevier, v. 21, n. 2, p. 118–129, 2013. Citado na página 25.

PACHAL, P. **The iPhone 4S Is the First Bluetooth 4.0 Phone**. 2011. Acessado em: 15 Fev 2018. Disponível em: <<https://www.pcmag.com/article2/0,2817,2395178,00.asp>>. Citado na página 10.

RAMAKRISHNAN, R.; GAUR, L.; SINGH, G. Feasibility and efficacy of ble beacon iot devices in inventory management at the shop floor. **International Journal of Electrical and Computer Engineering (IJECE)**, v. 6, n. 5, p. 2362–2368, 2016. Citado na página 38.

RAZA, S. et al. Building the internet of things with bluetooth smart. **Ad Hoc Networks**, Elsevier BV, v. 57, p. 19–31, mar 2017. Disponível em: <<https://doi.org/10.1016/j.adhoc.2016.08.012>>. Citado 2 vezes nas páginas 10 e 12.

SIG, B. **Bluetooth Special Interest Group**. 2018. Acessado em: 25 Jun 2018. Disponível em: <<https://www.bluetooth.com>>. Citado 5 vezes nas páginas 9, 10, 11, 12 e 15.

SILVA, W. C. da; ALMEIDA, P. E. M. de. An Obstacle-aware Model to Indoor Navigation Applied to Supermarket Shopping Routing. In: **Proceedings of the XXXVIII Iberian Latin American Congress on Computational Methods in Engineering**. ABMEC Brazilian Association of Computational Methods in Engineering, 2017. Disponível em: <<https://doi.org/10.20906/cps/cilamce2017-1091>>. Citado na página 66.

SNEPS-SNEPPE, M.; NAMIOT, D. On physical web models. In: **2016 International Siberian Conference on Control and Communications (SIBCON)**. [S.l.: s.n.], 2016. p. 1–6. Citado na página 14.

SYAEKHONI, M. A.; LEE, C.; KWON, Y. S. Analyzing customer behavior from shopping path data using operation edit distance. **Applied Intelligence**, Springer, p. 1–21, 2016. Citado na página 24.

TASTAN, B.; SUKTHANKAR, G. Leveraging human behavior models to predict paths in indoor environments. **Pervasive and Mobile Computing**, Elsevier, v. 7, n. 3, p. 319–330, 2011. Citado na página 24.

UNACAST. **The Proxbook Report - The State of the Proximity Industry Q4 2015**. [S.l.], 2016. Acessado em: 23 Jun 2018. Disponível em: <https://unacast.s3.amazonaws.com/The_Proxbook_Report_Q4_2015.pdf>. Citado na página 69.

WANT, R. et al. The active badge location system. **ACM Transactions on Information Systems (TOIS)**, ACM, v. 10, n. 1, p. 91–102, 1992. Citado na página 18.

WEISER, M. The computer for the 21st century. **Scientific american**, Nature Publishing Group, v. 265, n. 3, p. 94–104, 1991. Citado na página 7.

YANG, Z.; LIU, Y. Quality of trilateration: Confidence-based iterative localization. **IEEE Transactions on Parallel and Distributed Systems**, v. 21, n. 5, p. 631–640, May 2010. ISSN 1045-9219. Citado 2 vezes nas páginas 19 e 20.

YOUNG, D. G. **android-beacon-library**. 2013–2018. Acessado em: 06 Jun 2018. Disponível em: <<https://github.com/AltBeacon/android-beacon-library>>. Citado na página 18.

Apêndices

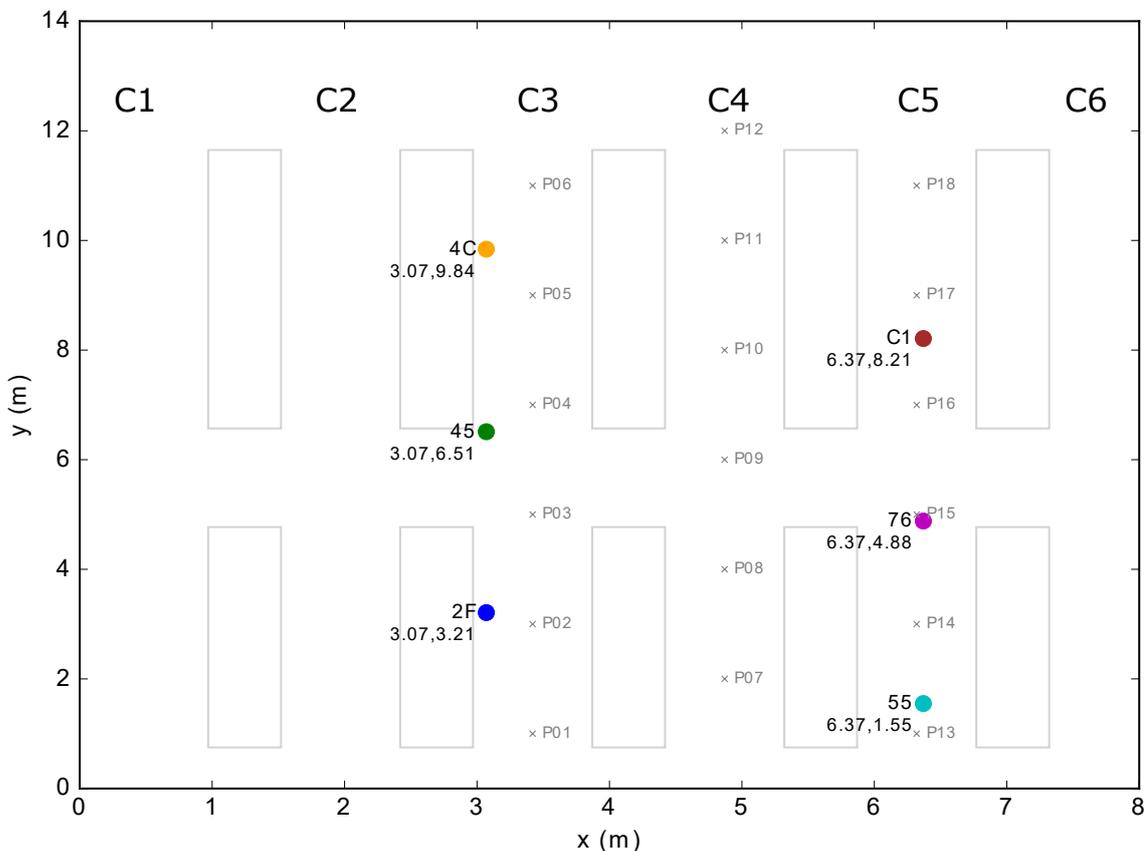
APÊNDICE A – Mapas dos Experimentos

Nesta seção são apresentados os mapas de todos os experimentos de localização (grupo D). Eles refletem exatamente a tabela de configuração dos experimentos (D1 a D8) exibida na Tabela 4 da Subseção 4.3.6. Os mapas são inseridos como anexo para facilitar a visualização de cada um dos experimentos.

A numeração dos corredores é a mesma para todos os experimentos e é exibida apenas no mapa do experimento D1. A escolha dos corredores é feita de acordo com a área de interesse de cada experimento. As pontas dentro dos corredores são numerados a partir do ponto de origem, $(0, 0)$, de forma crescente, de acordo com os valores (x, y) .

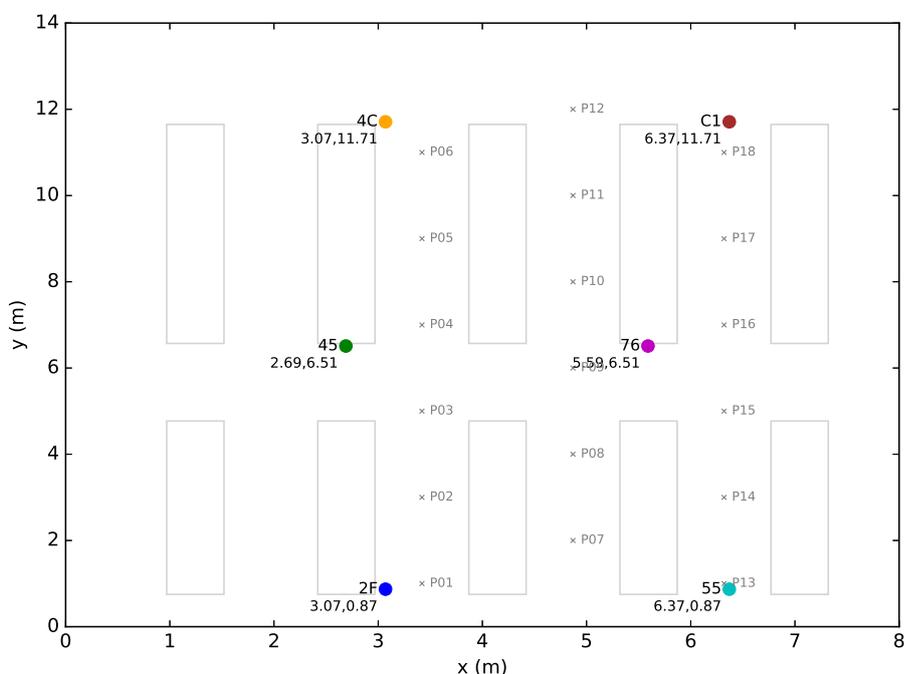
A.1 Experimento D1

Os *beacons* foram fixados nas vigas (teto) a uma altura de $2.39m$ do dispositivo leitor.



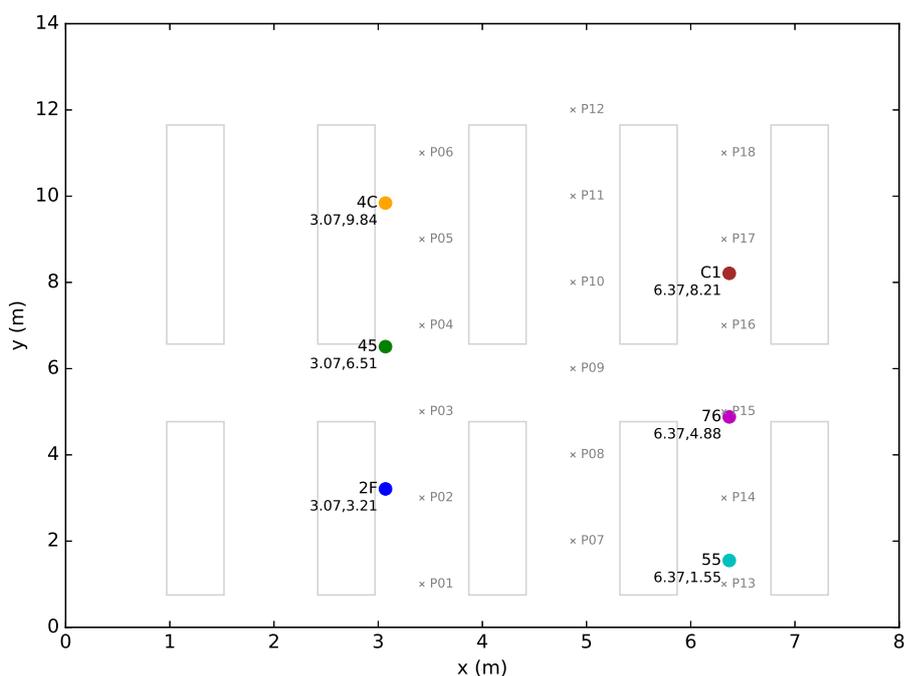
A.2 Experimento D2

Os *beacons* foram fixados nas extremidades nas vigas (teto), a uma altura de $2.39m$ do dispositivo leitor.



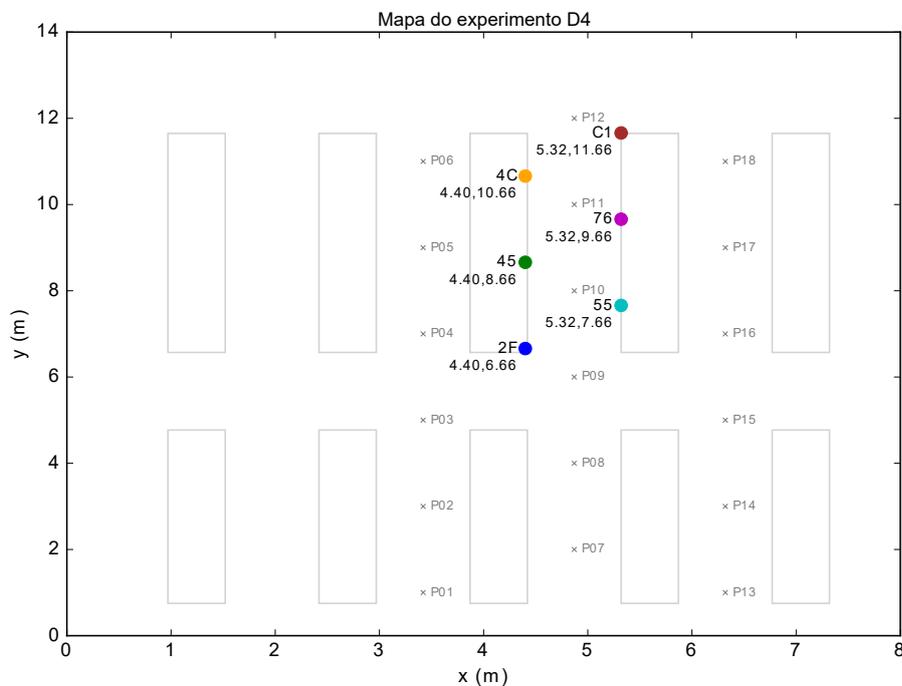
A.3 Experimento D3

Os *beacons* estão nas mesmas posições (x, y) do experimento D1, mudando apenas a posição z ($0.60m$), fixados nas prateleiras, na mesma altura do dispositivo de leitura.



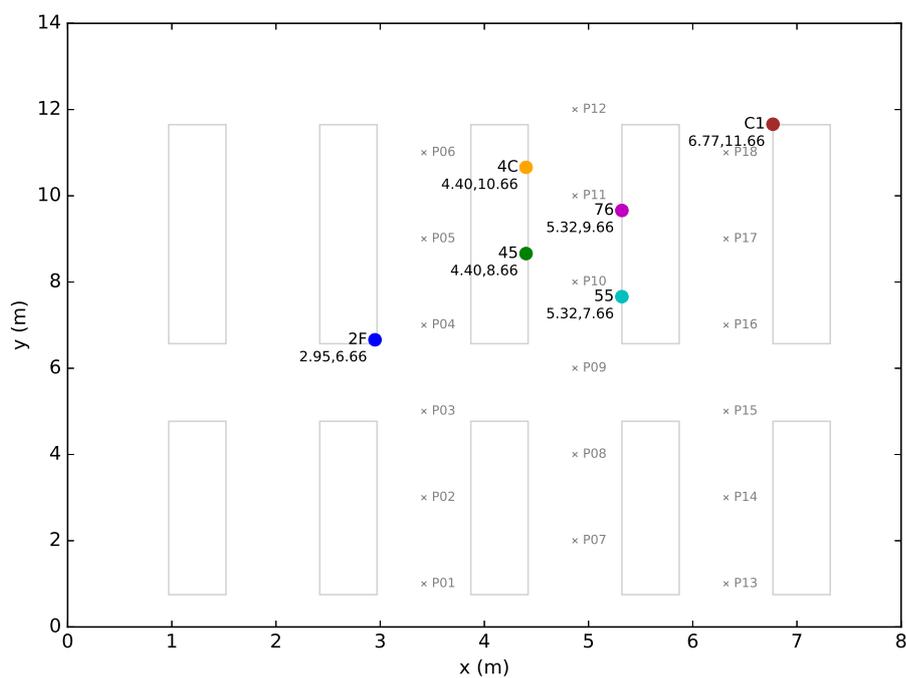
A.4 Experimento D4

Todos os *beacons* foram concentrados em um único corredor, fixados nas prateleiras, na mesma altura (0.60m) do dispositivo de leitura.



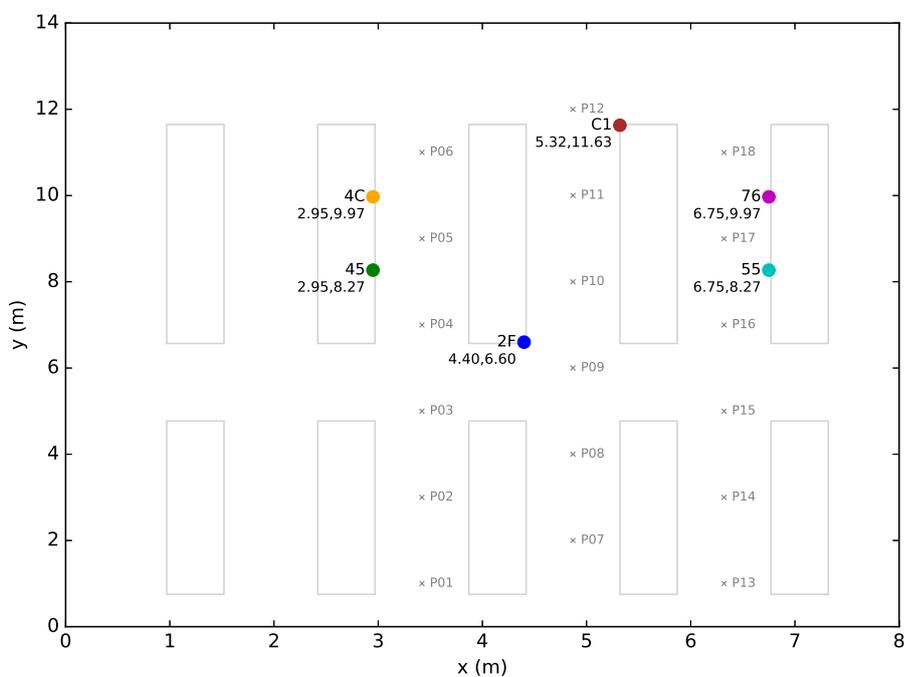
A.5 Experimento D5

Uma variação do experimento D4, mantendo o foco em um corredor, com *beacons* fixados nas prateleiras, na mesma altura (0.60m) do dispositivo de leitura.



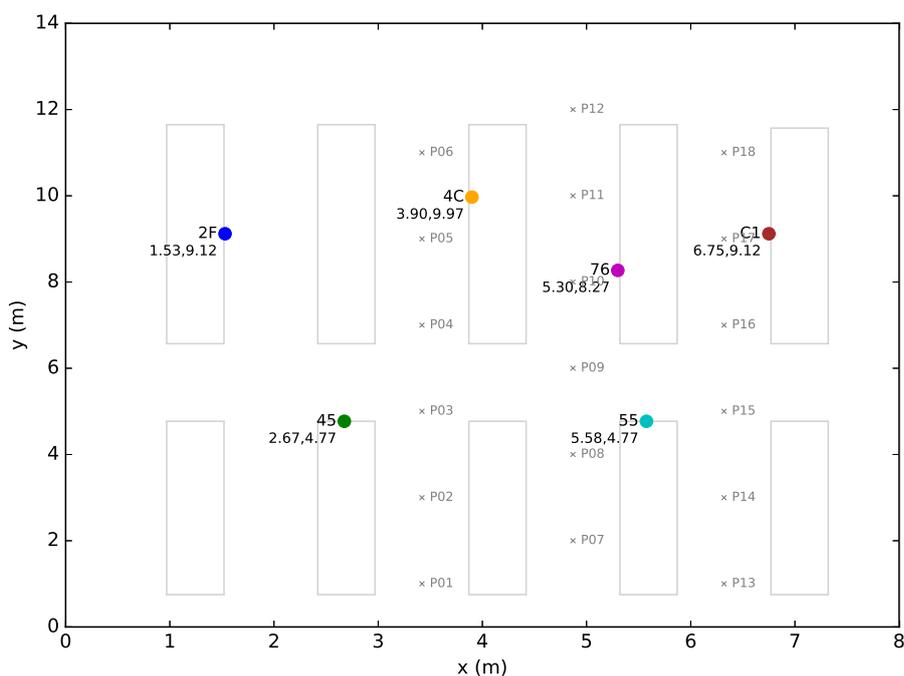
A.6 Experimento D6

Outra variação do experimento D4, com *beacons* mais espalhados, fixados nas prateleiras, na mesma altura (0.60m) do dispositivo de leitura.



A.7 Experimento D7

Foco nos corredores superiores, com *beacons* fixados nas prateleiras, na mesma altura (0.60m) do dispositivo de leitura.



A.8 Experimento D8

Os *beacons* foram fixados nas prateleiras, nas extremidades da área de interesse, na mesma altura (0.60m) do dispositivo de leitura.

