INSTITUTO FEDERAL DE SANTA CATARINA

DANIEL TREVISAN TATSCH

Localização indoor utilizando Bluetooth Low Energy e aprendizado de máquina

São José - SC

novembro/2018

Localização indoor utilizando Bluetooth Low Energy e aprendizado de máquina

Trabalho de conclusão de curso apresentado à Coordenadoria do Curso de Engenharia de Telecomunicações do campus São José do Instituto Federal de Santa Catarina para a obtenção do diploma de Engenheiro de Telecomunicações.

Orientador: Roberto de Matos

Coorientador: Mario de Noronha Neto

São José - SC

novembro/2018

RESUMO

A necessidade de localização de objetos ou pessoas não se limita somente à aplicações em ambientes externos, tendo a localização *indoor* grande importância nos cenários de *Internet of Things* (IoT). Tecnologias como o Bluetooth Low Energy (BLE), foco deste trabalho, em conjunto com técnicas de aprendizado de máquina usam apenas dados provenientes da comunicação entre os componentes da rede para aferir a localização em ambientes internos. O objetivo deste trabalho é apresentar testes e análises de diferentes técnicas de aprendizado de máquina associadas à localização *indoor* utilizando o BLE. Este documento apresenta a fundamentação teórica necessária para a realização do objetivo proposto, bem como o estabelecimento de parâmetros e cenários em que serão realizados os testes.

Palavras-chave: Localização *indoor*. Bluetooth Low Energy. Aprendizado de máquina.

LISTA DE ILUSTRAÇÕES

Figura 1 – Região de incerteza utilizando o *Received Signal Strength Indication* (RSSI) 13
Figura 2 – *Uniform Linear Array* (ULA) 14
Figura 3 – Direção da origem do sinal recebido com o *Angle of Arrival* (AoA) . . . 15
Figura 4 – Região de incerteza utilizando o *Time Difference of Arrival* (TDoA) . . 16
Figura 5 – Região de incerteza utilizando a trilateração 17
Figura 6 – Estimativa da localização utilizando a triangulação 18
Figura 7 – Região de incerteza utilizando a multilateração 18
Figura 8 – Cenário de localização utilizando o *fingerprint* 19
Figura 9 – Pilha protocolar do BLE 21
Figura 10 – Canais BLE e sobreposição dos canais do Wi-Fi 22
Figura 11 – Formato do quadro BLE 23
Figura 12 – Descrição do cenário de testes 30
Figura 13 – Plataforma de desenvolvimento ESP32 30
Figura 14 – Plataforma Arduino Nano e Módulo BLE AT-09 31

LISTA DE TABELAS

Tabela 1 – Exemplo da fase de testes do algoritmo Naive Bayes 27
Tabela 2 – Cronograma das atividades previstas 32

LISTA DE ABREVIATURAS E SIGLAS

IoT *Internet of Things* 2

BLE Bluetooth Low Energy 2

M2M *Machine-to-Machine* 11

LPWAN *Low Power Wide Area Network* 11

bps bits por segundo 11

NFC *Near Field Communication* 12

RFID *Radio Frequency Identification* 9

GPS *Global Positioning System* 9

RSSI *Received Signal Strength Indication* 3

AoA *Angle of Arrival* 3

ToA *Time of Arrival* 12

ToF *Time of Flight* 15

TDoA *Time Difference of Arrival* 3

ULA *Uniform Linear Array* 3

SIG *Bluetooth Special Interest Group* 21

GFSK *Gaussian Frequnecy Shift Keying* 22

AFH *Adaptative Frequency Hopping* 22

TDMA *Time Division Multiple Access* 22

CRC *Cyclic Redundancy Check* 23

HCI *Host Controller Interface* 21

ATT *Attribute Protocol* 23

GATT *Generic Attribute Profile* 23

GAP *Generic Access Profile* 24

SMP *Security Manager Protocol* 24

L2CAP *Logical Link Control and Adaption Protocol* 23

IA Inteligência Artificial 25

SLAs *Supervised Learning Algorithms* 25

K-NN *K-Nearest Neighbor* 26

AWS Amazon Web Services 29

API Application Programming Interface 30

SUMÁRIO

1 INTRODUÇÃO . . . 9
1.1 Objetivo geral . . . 10
1.2 Objetivos específicos . . . 10

2 FUNDAMENTAÇÃO TEÓRICA . . . 11
2.1 Internet das Coisas . . . 11
2.2 Parâmetros da comunicação . . . 12
2.2.1 Received Signal Strength Indication (RSSI) . . . 12
2.2.2 Angle of Arrival (AoA) . . . 13
2.2.3 Time of Arrival (ToA) . . . 15
2.2.4 Time Difference of Arrival (TDoA) . . . 15
2.3 Técnicas de localização . . . 16
2.3.1 Trilateração . . . 16
2.3.2 Triangulação . . . 17
2.3.3 Multilateração . . . 18
2.3.4 *Fingerprinting* . . . 19
2.4 Tecnologias . . . 20
2.5 Bluetooth Low Energy . . . 21
2.5.1 Camada Física . . . 21
2.5.2 Camada de Enlace . . . 22
2.5.3 Logical Link Control and Adaption Protocol (L2CAP) . . . 23
2.5.4 Attribute Protocol (ATT) e Generic Attribute Profile (GATT) . . . 23
2.5.5 Generic Access Profile (GAP) . . . 24
2.5.6 Security Manager Protocol (SMP) . . . 24
2.6 Aprendizado de Máquina . . . 25
2.6.1 K-Nearest Neighbor (K-NN) . . . 26
2.6.2 Naive Bayes . . . 26
2.6.3 Árvore de decisão . . . 27

3 PROPOSTA . . . 29
3.1 Estudo das técnicas de localização *indoor* . . . 29
3.2 Implementação de uma rede de dispositivos BLE . . . 29
3.3 Análise e testes das técnicas de aprendizado de máquina . . . 31
3.4 Cronograma de atividades previstas . . . 32

REFERÊNCIAS **33**

1 INTRODUÇÃO

Atualmente, a busca por sistemas de controle e monitoramento situados no contexto de IoT tem priorizado equipamentos que possuem baixo custo de implementação e consumo energético ínfimo. Isso permite englobar aplicações altamente escaláveis e com valores aceitáveis a diferentes contextos, como transporte, saúde, automação industrial, ou até o meio ambiente, onde poderia ser feito por exemplo, o monitoramento de lugares propícios a ocorrência de desastres naturais (AL-FUQAHA et al., 2015). Com o intuito de suprir essas demandas, tecnologias de comunicação como o Bluetooth Low Energy (BLE), *Radio Frequency Identification* (RFID) e LoRaWAN possuem como principais características a eficiência energética e custo de implementação reduzido.

À medida que soluções envolvendo IoT se difundem, surge a necessidade de se obter resultados mais aprimorados com tecnologias que antes não eram utilizadas para tais finalidades. No caso do BLE por exemplo, havendo uma rede de dispositivos que se comunicam utilizando essa tecnologia, a mesma pode ter como objetivo determinar a localização desses equipamentos em determinados ambientes.

Estimar o posicionamento de pessoas ou objetos em ambientes abertos, denominados *outdoor*, se tornou algo trivial graças ao *Global Positioning System* (GPS). No entanto, segundo Sadowski e Spachos (2018), sua precisão fica comprometida quando aplicado em ambientes internos, chamados de *indoor*, como prédios ou casas. Isso se dá por conta da dependência de linha de visada dos dispositivos com os satélites de GPS. Alguns fatores que desfavorecem a utilização do GPS para localização *indoor* estão ligados também ao consumo de energia e ao custo de aquisição do sensor.

Para Rezende e Ynoguti (2015), o maior desafio de se implementar a localização *indoor* está nas características do próprio ambiente em si, como a existência de móveis, paredes, estruturas da construção ou equipamentos que utilizam de outras tecnologias de comunicação sem fio. Todos esses fatores podem influenciar na comunicação dos nós que compõem a rede. Para indicar com fidelidade a posição de um usuário ou dispositivo em um ambiente interno, a precisão pode chegar a apenas alguns centímetros, visto que com uma diferença mínima outro ambiente poderia ser registrado.

Diferente do GPS, a localização *indoor* pode apresentar, dependendo da técnica utilizada, além de coordenadas geográficas, resultados na forma simbólica, como setores ou ambientes em que os dispositivos podem estar posicionados. De acordo com Rezende e Ynoguti (2015), a forma em que um sistema de localização *indoor* é utilizado varia de acordo com cada cenário em que é aplicado. Em uma loja por exemplo, pode haver o monitoramento da seção em que cada cliente está, sugerindo ofertas de determinados

produtos que estão próximos a sua posição. Já em um hospital, saber a localização de médicos e pacientes pode ser crucial em casos de emergência.

O BLE, bem como outras tecnologias de comunicação sem fio de baixo consumo energético e custo de implementação reduzido, busca suprir com essa necessidade de localização em ambientes internos. De acordo com Sadowski e Spachos (2018), essas técnicas podem se basear em diversos parâmetros atrelados à comunicação entre os dispositivos presentes em uma rede de sensores. No entanto, técnicas que utilizam o RSSI oferecem uma demanda maior, visto que não é necessário que se tenha um *hardware* específico e complexo para obtenção desse parâmetro, tornando o sistema mais viável economicamente.

Mesmo que medidas baseadas na potência do sinal recebido sofram grandes variações devido ao ruído presente no ambiente e as atenuações causadas pelo multi percurso, Liu et al. (2007) afirma que modelos que utilizam o RSSI podem obter resultados satisfatórios através da aplicação de técnicas de aprendizado de máquina.

Este trabalho está dividido em 2 partes. Primeiramente, no Capítulo 2 são apresentados os conceitos de *Internet of Things* (IoT), seguido dos parâmetros, técnicas e tecnologias utilizadas na localização *indoor*. Por ser a tecnologia escolhida no desenvolvimento do trabalho, a seção 2.5 aborda as características do Bluetooth Low Energy (BLE). Por fim, a seção 2.6 define o tema aprendizado de máquina, destacando os algoritmos presentes em cada um dos seus métodos. No Tabela 3.4 é apresentada a proposta do projeto que será realizado no decorrer da segunda parte do trabalho de conclusão de curso.

1.1 Objetivo geral

Realizar testes e análises de diferentes técnicas de aprendizado de máquina associadas à localização *indoor*, utilizando como tecnologia de comunicação o BLE.

1.2 Objetivos específicos

- Estudar os parâmetros de comunicação utilizando BLE para uso na localização *indoor*;
- Desenvolver protótipos de *gateways* e *beacons* BLE;
- Coletar informações dos *beacons* para alimentar uma base de dados de acordo com suas posições;
- Compreender, utilizar e comparar algoritmos de aprendizado de máquina para estimar a localização dos *beacons*;

2 FUNDAMENTAÇÃO TEÓRICA

Os tópicos apresentados neste capítulo abordam assuntos considerados relevantes para o desenvolvimento do trabalho. Na seção 2.1, o tema IoT é definido, havendo também uma breve discussão sobre as tecnologias que compõem os seus sistemas. Nas seções 2.2, 2.3 e 2.4 são definidos respectivamente os parâmetros, técnicas e tecnologias de comunicação utilizados para a realização da localização *indoor*. Na seção 2.5 a tecnologia BLE, por ser o foco deste trabalho, é apresentada mais detalhadamente. Por fim, a seção 2.6 se encarrega de definir o conceito de aprendizado de máquina, bem como apresentar suas diversas técnicas que podem ser aplicadas em sistemas de localização *indoor*.

2.1 Internet das Coisas

O conceito de IoT está ligado a capacidade de diversos objetos, dispositivos ou equipamentos estabelecerem comunicação entre si. Essa forma de interação constitui um tipo de comunicação denominada *Machine-to-Machine* (M2M), que possibilita que diversos sensores se comuniquem tipicamente pelo meio sem fio e troquem informações de forma autônoma (XIAO et al., 2016). Além disso, essa conexão viabiliza a criação de aplicações voltadas a diferentes contextos, como transporte, saúde, automação industrial ou até o mesmo ao meio ambiente, onde poderia ser feito por exemplo, o monitoramento de lugares propícios a ocorrência de desastres naturais (AL-FUQAHA et al., 2015).

Com a possibilidade do desenvolvimento de uma rede de IoT específica para cada um dos cenários apresentados, surge então, a necessidade de analisar e delimitar os requisitos que serão impostos sobre a rede, bem como suas restrições. De acordo com Loureiro et al. (2003), os dispositivos que compõem uma rede de sensores sem fio possuem limitações quanto ao processamento, memória, taxa de transmissão de dados, alcance e consumo de energia. Sendo assim, é evidente que uma única tecnologia de comunicação não seja capaz de suprir com todas essas necessidades.

Segundo Samie, Bauer e Henkel (2016), para aplicações em que se faz necessário ter uma vasta área de cobertura aliada ao baixíssimo consumo energético, tecnologias pertencentes as redes *Low Power Wide Area Network* (LPWAN), como SigFox e LoRaWAN tendem a ser opções viáveis. Por mais que possuam baixas taxas de transmissão (até poucos kbits por segundo (bps) dependendo da distância entre os componentes da rede), essas tecnologias levam certa vantagem em relação a rede celular, se comparados os custos de implementação e consumo de energia.

Existem também, tecnologias que se encarregam de cobrir áreas menores, como o

RFID e o *Near Field Communication* (NFC), tendo cobertura de apenas alguns centímetros e baixas taxas de transmissão. A vantagem dessas tecnologias é prover de elementos passivos, chamados de *tags*, para comunicação. No RFID por exemplo, existe um elemento ativo que envia constantemente um sinal de controle. Quando uma *tag* é aproximada, esse sinal é identificado e rotulado por ela com o seu registro único, sendo então, enviado de volta para o transmissor (AL-FUQAHA et al., 2015).

Para os cenários de IoT que não têm como requisito o longo alcance, pode ser necessário no entanto, obter taxas de transmissão mais elevadas com um *hardware* de custo acessível. Tecnologias como o Bluetooth e BLE possuem como atrativos essas características, tendo o BLE vantagens no que diz respeito ao consumo energético, tempo de configuração e suporte às topologias de rede em estrela com número ilimitado de nós (SAMIE; BAUER; HENKEL, 2016). Nas próximas seções são explicitados os parâmetros de comunicação, bem como técnicas que utilizam desses parâmetros para realizar a localização *indoor*. São apresentadas também, tecnologias de IoT que podem ser utilizadas para essas aplicações.

2.2 Parâmetros da comunicação

Com o objetivo de realizar a implementação de um sistema de localização *indoor*, antes da escolha da técnica a ser utilizada, deve-se analisar primeiramente quais características da comunicação entre os dispositivos da rede serão consideradas. Essa escolha impacta diretamente na complexidade, no custo e no consumo de energia dos equipamentos do sistema. Segundo Liu et al. (2007), os modelos de localização mais comuns são baseados em parâmetros como a potência do sinal recebido, chamada de RSSI; o ângulo de incidência do sinal na antena do receptor, conhecido como AoA; o tempo de chegada do sinal, *Time of Arrival* (ToA), ou o TDoA, que especifica a diferença do tempo de chegada dos sinais na antena do receptor.

2.2.1 Received Signal Strength Indication (RSSI)

De acordo com Sadowski e Spachos (2018), o *Received Signal Strength Indication* (RSSI) é o parâmetro mais simples e popular utilizado em técnicas de localização *indoor*. Isso se da pelo fato de que esse dado pode ser obtido sem nenhuma adição ou maior processamento de *hardware*. No entanto, sua precisão fica comprometida por conta da interferência causada pela presença do ruído no sinal propagado. Além disso, o multi percurso elevado nos ambientes internos causado por pessoas e móveis também ocasiona erros na estimativa do posicionamento.

Analisando o fato de que o ruído presente no sinal varia conforme o ambiente em que este sinal é transmitido, sua propagação segue um modelo de perda de caminho conhecido como sombreamento log-normal (RAPPAPORT, 2001). Nele ocorre um decai-

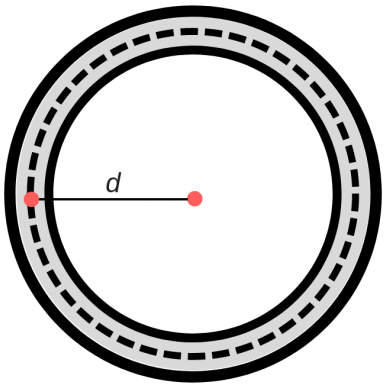
mento logaritmo do sinal em relação a distância, variando de forma aleatória segundo uma distribuição normal. Sendo assim, pode-se representar a perda de caminho $P(d)$ em função da distância e da variável aleatória X de média nula e desvio padrão σ através da Equação 2.1.

$$P(d) = P(d_0) - 10 \cdot log_{10}(\frac{d}{d_0}) + X \tag{2.1}$$

Onde n é o expoente de perda de caminho que indica a velocidade com a qual o sinal é atenuado, d_0 corresponde a uma distância de referência e $P(d_0)$ indica a potência do sinal em relação a essa distância.

Sendo o RSSI conhecido pelo receptor, a distância d entre ele e o dispositivo transmissor pode ser estimada através de um círculo de incerteza, conforme mostra a Figura 1. A área destacada explicita as imprecisões nas medições do RSSI, que podem causar um deslocamento do valor ideal da distância encontrada.

Figura 1 – Região de incerteza utilizando o RSSI



Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

Utilizar o RSSI como parâmetro para determinar a localização *indoor* se torna viável graças a facilidade na obtenção de resultados e o baixo custo do *hardware* envolvido. Mesmo que sofra com o ruído e com a atenuação devido ao multi percurso, modelos que utilizam esse parâmetro podem obter melhores resultados realizando múltiplas medições com vários pontos conhecidos e aplicando de técnicas de aprendizado de máquina (LIU et al., 2007).

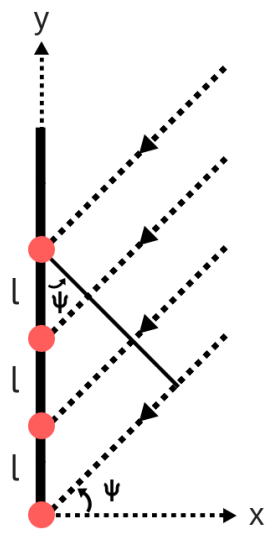
2.2.2 Angle of Arrival (AoA)

Para os modelos baseadas no *Angle of Arrival* (AoA), é necessário que se tenha um arranjo de antenas para determinar o ângulo no qual o sinal é propagado. Consequentemente,

esses modelos possuem um custo maior, pois requerem um *hardware* complexo e calibragens para que se obtenha uma precisão satisfatória (SADOWSKI; SPACHOS, 2018).

Segundo Sahinoglu, Gezici e Güvenc (2008), um dos arranjos de antenas mais comuns é chamado de ULA e consiste no alinhamento e espaçamento uniforme entre todas as antenas, conforme mostra a Figura 2, onde os pontos no eixo y representam as antenas receptoras, espaçadas por uma distância l. Com isso, o ângulo de chegada ψ de um sinal pode ser obtido de acordo com os atrasos subsequentes em cada antena. Por outro lado, para sinais de banda estreita, podem ser utilizadas as diferenças de fase entre os sinais que chegam nas antenas do arranjo, podendo ser combinadas afim de estimar a direção de origem do sinal recebido.

Figura 2 – *Uniform Linear Array* (ULA)



Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

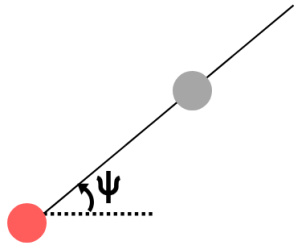
O cálculo do AoA pode ser realizado conhecendo o atraso de propagação τ na antena subsequente, como mostra a Equação 2.2

$$\psi = \arcsin\left(\tau \cdot l \cdot c\right) \tag{2.2}$$

Sendo ψ o valor do AoA, l o espaçamento entre as antenas e c a velocidade da luz.

Diferente do que ocorre com o RSSI, onde pode ser obtida a distância entre dois pontos, a Figura 3 mostra que medida do AoA determina uma direção da origem do sinal recebido de acordo com o ângulo ψ formado entre os pontos de referência (vermelho) e de destino (cinza) (SAHINOGLU; GEZICI; GÜVENC, 2008).

Figura 3 – Direção da origem do sinal recebido com o AoA



Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

2.2.3 Time of Arrival (ToA)

Também conhecido como *Time of Flight* (ToF), o parâmetro *Time of Arrival* (ToA) é utilizado em sistemas que baseiam-se na sincronização dos tempos de envio e chegada de um sinal entre transmissor e receptor para estipular a distância entre os dois pontos (FARID; NORDIN; ISMAIL, 2013). O resultado encontrado é análogo ao apresentado na Figura 1, e pode ser estimado através do produto entre tempo de propagação do sinal até o receptor e a velocidade da luz.

Para Gu, Lo e Niemegeers (2009), técnicas que utilizam desse parâmetro para realizar a localização *indoor* de dispositivos são consideradas as mais precisas, podendo filtrar os efeitos de multi percurso presentes nos ambientes internos. Porém, segundo Farid, Nordin e Ismail (2013), apesar da precisão, os métodos que utilizam ToA necessitam de equipamentos específicos para estabelecer com fidelidade o sincronismo entre todos os dispositivos da rede, aumentando assim o custo do sistema.

2.2.4 Time Difference of Arrival (TDoA)

Diferente do ToA que requer sincronização entre os dispositivos de referência e os que se deseja estimar a localização, as técnicas que utilizam o *Time Difference of Arrival* (TDoA) necessitam apenas da sincronização entre os pontos de referência (FARID; NORDIN; ISMAIL, 2013).

De acordo com Sahinoglu, Gezici e Güvenc (2008), uma forma de obter o TDoA é estimar o ToA em um receptor para dois dos pontos de referência, realizando então a diferença entre esses valores, conforme apresentado na Equação 2.3

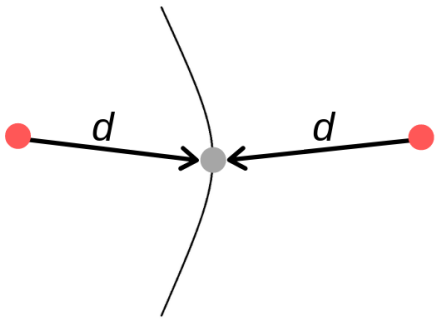
$$\tau = \tau_1 - \tau_2 \tag{2.3}$$

Onde τ_1 e τ_2 representam a estimativa do ToA do primeiro e segundo pontos de referência, respectivamente.

Com essa diferença de tempo calculada, o dispositivo alvo (cinza) fica posicionado

em uma região de incerteza formada por uma hipérbole, conforme mostra a Figura 4, onde os pontos conhecidos estão em vermelho.

Figura 4 – Região de incerteza utilizando o TDoA



Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

2.3 Técnicas de localização

Nesta subseção são apresentadas técnicas de localização *indoor* que atuam sobre um determinado dispositivo através do seu posicionamento em relação aos demais componentes da rede. Para tal, são utilizados os parâmetros apresentados na subseção 2.2.

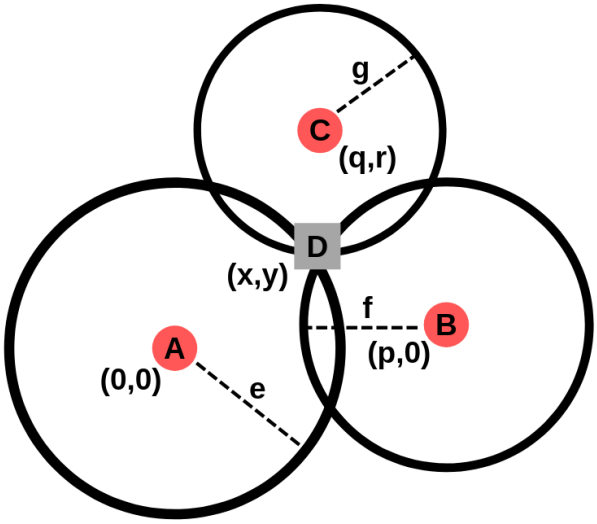
Primeiramente são apresentadas técnicas que utilizam abordagens geométricas para determinar a localização de um dispositivo, através da interseção das posições de um conjunto de pontos de referência. Por fim, é apresentada a técnica utilizada neste trabalho, chamada de *Fingerprinting*, que parte de modelos estatísticos para determinar a localização *indoor*.

2.3.1 Trilateração

O método da trilateração consiste na obtenção da localização em duas dimensões de um dispositivo através da sua distância em relação a três pontos de referência, onde suas posições são conhecidas (SADOWSKI; SPACHOS, 2018).

Parâmetros como o RSSI ou o ToA podem ser utilizados para estabelecer a distância entre cada um dos pontos e o dispositivo alvo. Com isso, uma região de incerteza é formada de acordo com cada uma das distâncias estimadas. A Figura 5 apresenta a região de incerteza D com coordenadas $(x; y)$ formadas pela interseção dos pontos A, B e C, que possuem coordenadas $(0; 0)$, $(p; 0)$ e $(q; r)$ e distâncias e, f e g, respectivamente.

Figura 5 – Região de incerteza utilizando a trilateração



Fonte: Baseado em Sadowski e Spachos (2018).

Sendo conhecidas as distâncias que correspondem aos raios das circunferências na Figura 5, as coordenadas do dispositivo alvo podem ser determinadas através das equações 2.4 e 2.5 apontadas por Sadowski e Spachos (2018).

$$x = \frac{e^2 - f^2 + p^2}{2 \cdot p} \tag{2.4}$$

$$y = \frac{e^2 - g^2 + q^2 + r^2}{2 \cdot r} - x \cdot \frac{q}{r} \tag{2.5}$$

2.3.2 Triangulação

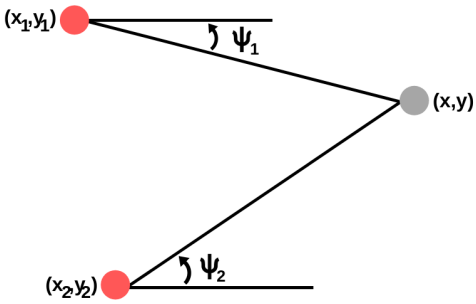
Segundo Sahinoglu, Gezici e Güvenc (2008), a triangulação tem como objetivo estimar a localização *indoor* através do AoA entre apenas dois dispositivos de referência e o que se deseja obter a posição, desde que este não esteja alinhado com os pontos conhecidos . A Figura 6 ilustra os dois pontos conhecidos (em vermelho) com as coordenadas $(x1; y1)$ e $(x2; y2)$, respectivamente.

Seus ângulos $\psi 1$ e $\psi 2$ em relação ao ponto de interesse (cinza) são utilizados para estimar a posição do mesmo através das equações 2.6 e 2.7 propostas por Sahinoglu, Gezici e Güvenc (2008).

$$x = \frac{x_2 \cdot \tan \psi_2 - x_1 \cdot \tan \psi_1 + y_1 - y_2}{\tan \psi_2 - \tan \psi_1} \tag{2.6}$$

$$y = \frac{(x_2 - x_1) \cdot \tan \psi_1 \cdot \tan \psi_2 + y_1 \cdot \tan \psi_2 - y_2 \cdot \tan \psi_1}{\tan \psi_2 - \tan \psi_1} \tag{2.7}$$

Figura 6 – Estimativa da localização utilizando a triangulação

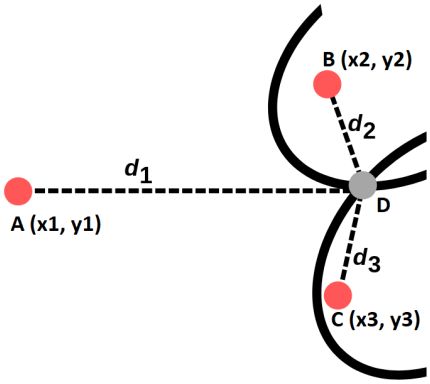


Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

2.3.3 Multilateração

A técnica da multilateração utiliza como parâmetro o TDoA e consiste em agregar no mínimo três dispositivos cuja suas posições são conhecidas (SAHINOGLU; GEZICI; GÜVENC, 2008). A Figura 7 explicita as regiões de incerteza geradas através dos pontos A $(x1; y1)$, B $(x2; y2)$ e C $(x3; y3)$ e o ponto de interseção D (cinza).

Figura 7 – Região de incerteza utilizando a multilateração



Fonte: Baseado em Sahinoglu, Gezici e Güvenc (2008).

A estimativa da localização pode ser feita então, calculando as regiões das duas hipérboles em relação à distância $d1$ através da Equação 2.8 (SAHINOGLU; GEZICI; GÜVENC, 2008).

$$d_{i1} = d_i - d_1 = \sqrt{(x - x_i)^2 + (y - y_i)^2} - \sqrt{(x - x_1)^2 + (y - y_1)^2} \tag{2.8}$$

Onde $i = 2, 3$.

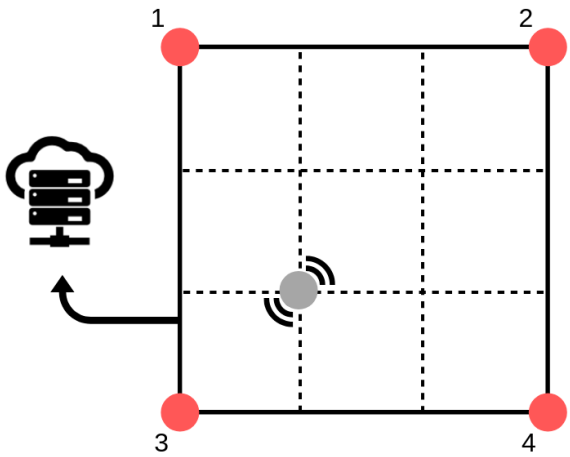
2.3.4 *Fingerprinting*

Apesar das imprecisões associadas à medição do RSSI destacadas na subseção 2.2.1, a técnica *fingerprinting* é uma solução viável para se modelar ambientes que utilizam desse parâmetro para realizar a localização *indoor* (INDOOR..., 2018). Além disso, por estimar a posição dos dispositivos utilizando apenas RSSI, não é necessário utilizar *hardwares* com sincronização de tempo ou arranjos de antenas, diminuindo significativamente o custo e complexidade do sistema.

Segundo Farid, Nordin e Ismail (2013), a localização utilizando *fingerprinting* consiste em duas etapas: calibração ou treino e fase de localização ou testes. A ideia é separar um ambiente em diversos setores, onde na primeira fase é realizado um mapeamento dessas regiões através da medida do RSSI de diferentes dispositivos conhecidos em relação ao que se deseja analisar. Após a obtenção desses valores em todos os setores, os resultados são armazenados em uma base de dados. Com isso, a segunda etapa consiste em determinar a localização de um dispositivo através da comparação entre os dados atuais e os que foram previamente obtidos na primeira fase. Para manter a precisão desse método, a base de dados gerada na fase de treino deve ser periodicamente atualizada.

A Figura 8 apresenta o processo de localização *indoor* utilizando a técnica *fingerprinting*. Os pontos 1, 2, 3 e 4 correspondem aos dispositivos conhecidos que esperam os dados que são enviados constantemente em *broadcast* pelo ponto desconhecido. Eles então, alimentam a base de dados na primeira etapa ou à utilizam juntamente com as informações capturadas para estimar o setor em que o transmissor se encontra na segunda etapa.

Figura 8 – Cenário de localização utilizando o *fingerprint*



Fonte: Desenvolvida pelo autor.

É importante salientar que não há a necessidade de especificar os pontos de referência através de coordenadas $(x; y)$ como na abordagem geométrica, pois essa técnica retorna como resultado o setor mais provável em que o dispositivo alvo se encontra.

Pelo fato do *fingerprinting* realizar análises estatísticas para determinar a localização *indoor*, na seção seção 2.6 são apresentados diferentes algoritmos de aprendizado de máquina que podem ser utilizados para se obter resultados mais precisos.

2.4 Tecnologias

- **Wi-Fi**: Sistemas de localização *indoor* baseados na tecnologia Wi-Fi (IEEE 802.11) possuem vantagens se comparados a outros modelos, por utilizarem uma tecnologia muito difundida e compatível com praticamente todos os dispositivos que a possuem. Com isso e por utilizar o RSSI, os custos com o *hardware* e a manipulação dos dados são inferiores se comparados a outras tecnologias (FARID; NORDIN; ISMAIL, 2013). Porém, os dispositivos que utilizam o Wi-Fi possuem alto consumo energético, além de sofrerem possíveis interferências por conta de outros equipamentos que operam na mesma faixa de frequência dos 2.4GHz.

- **RFID**: A tecnologia *Radio Frequency Identification* (RFID) é usada na localização *indoor* de ambientes complexos e de difícil modelagem, como escritórios e hospitais (GU; LO; NIEMEGEERS, 2009). Utilizando nos receptores (dispositivos que serão localizados) *tags* passivas, ou seja, que não necessitam de uma fonte de alimentação própria, o sistema pode se tornar mais barato. No entanto, o alcance delas é muito pequeno se comparadas às *tags* ativas, que se encarregam de transmitir suas informações e cobrir uma região maior.

  Apesar de possuir dispositivos muito pequenos utilizados na localização de pessoas, Gu, Lo e Niemegeers (2009) afirma que para se utilizar técnicas de localização precisas são necessários muitos dispositivos RFID, fazendo com que o custo de um projeto utilizando essa tecnologia seja elevado.

- **LoRaWAN**: Desenvolvida pela LoRa Alliance, a tecnologia LoRaWAN possui como principal característica o longo alcance aliado a um consumo energético ínfimo. Segundo Sadowski e Spachos (2018), uma vantagem de se utilizar essa tecnologia é o fato dela operar na frequência de 915MHz, não sofrendo com interferências causadas por outras tecnologias de outros dispositivos como no caso do Wi-Fi. Por outro lado, um ponto negativo no uso da LoRaWAN é a necessidade de um *hardware* específico, patenteado pela empresa Semtech Corporation. Com isso, o custo de implantação de um sistema de localização *indoor* com essa tecnologia se torna elevado.

O Bluetooth Low Energy (BLE) tem sido uma opção viável na localização *indoor*, pois apesar de não possuir grande alcance e sofrer com as atenuações recorrentes dos ambientes internos, possui um custo menor se comparado às tecnologias Wi-Fi e LoRaWAN,

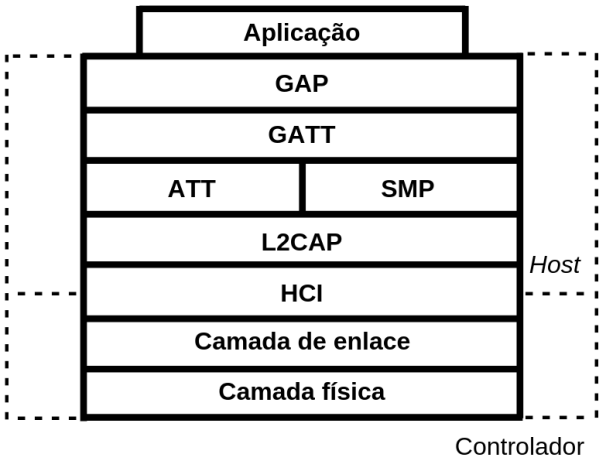
além de um consumo energético ínfimo. Por ser o foco deste trabalho, a seção 2.5 destina-se a apresentar em maiores detalhes as características dessa tecnologia.

2.5 Bluetooth Low Energy

O Bluetooth Low Energy (BLE) é uma tecnologia pertencente ao padrão Bluetooth 4.0, tendo sua origem no ano de 2010 através do *Bluetooth Special Interest Group* (SIG). O BLE tem como principal característica o baixo consumo energético associado à aplicações que necessitam apenas de uma baixa taxa de transmissão de dados, contrapondo os cenários em que a sua versão clássica atua. Além disso, o BLE opera sobre um pequeno número de interações entre os componentes da rede, como o envio de valores de sensores ou pequenas mensagens de controle com uma quantidade mínima de cabeçalho (ČABARKAPA; GRUJIć; PAVLOVIć, 2015).

Segundo Darroudi e Gomez (2017), o BLE atua sob uma pilha de protocolos que operam em duas partes: controlador e *host*. O controlador é responsável pela camada física, apresentada na subseção 2.5.1 e enlace, na subseção 2.5.2. O *host* oferece suporte às camadas superiores para interações entre aplicações, tendo seus protocolos definidos nas subseções 2.5.3 à 2.5.6. A Figura 9 apresenta a estrutura presente na pilha protocolar do BLE, onde a comunicação entre essas duas camadas é padronizada pelo *Host Controller Interface* (HCI).

Figura 9 – Pilha protocolar do BLE



Fonte: Baseado em Chang (2014).

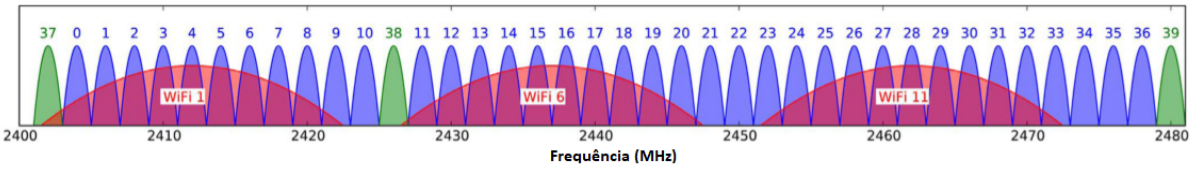
2.5.1 Camada Física

O BLE trabalha na frequência de 2.4GHz, definindo 40 canais de 2MHz cada ao longo do espectro, onde 3 destes são chamados de canais de anúncio e o restante de dados. Na sua versão clássica, o Bluetooth opera com 79 canais, sendo 32 de anúncio e o restante

de dados. Os canais de anúncio têm como finalidade descobrir dispositivos, estabelecer conexões e realizar transmissões em *broadcast*. Já os canais de dados ficam responsáveis por realizar a comunicação bidirecional entre os dispositivos conectados (ČABARKAPA; GRUJIć; PAVLOVIć, 2015).

Por operar na mesma faixa de frequência que outras tecnologias, como o Wi-Fi, o BLE tende a sofrer grandes interferências. Para minimizar esse problema, de acordo com Texas Instruments (2016), os canais 37, 38 e 39 foram escolhidos para serem os de anúncio por não sobreporem os canais mais comuns do Wi-Fi: 1, 6 e 11. A Figura 10 ilustra os respectivos canais do BLE, bem como a faixa de frequência em cada um está localizado. Além disso, destaca o posicionamento dos canais de anúncio (na cor verde) em relação aos canais do Wi-Fi (vermelho).

Figura 10 – Canais BLE e sobreposição dos canais do Wi-Fi



Fonte: Baseado em Indoor... (2018).

Possuindo uma taxa de transmissão de dados de até 1Mbps e utilizando o esquema de modulação *Gaussian Frequnecy Shift Keying* (GFSK), o BLE consegue consumir menos energia por conta do índice de modulação utilizado, alcançando distâncias de até 100 metros (ČABARKAPA; GRUJIć; PAVLOVIć, 2015). Além disso, segundo Gomez, Oller e Paradells (2012), nos canais de dados é utilizada a técnica *Adaptative Frequency Hopping* (AFH), que determina qual dos 37 será utilizado para comunicação durante um tempo pré estabelecido, alternado entre os canais de forma pseudo-aleatória. Essa técnica ajuda a diminuir a possibilidade de interferências e desvanecimento causados pela existência de outras tecnologias que operam na mesma faixa de frequência que o BLE.

2.5.2 Camada de Enlace

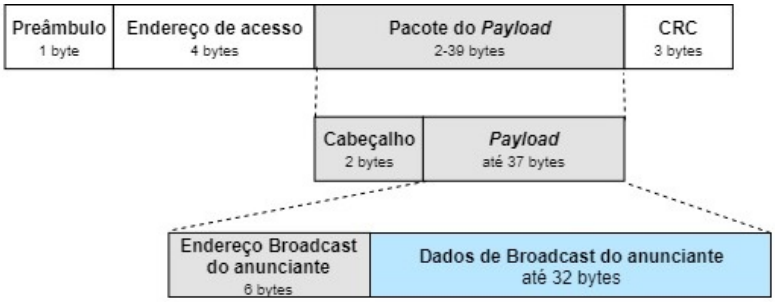
De acordo com Čabarkapa, Grujić e Pavlović (2015), a comunicação entre dois ou mais dispositivos BLE pode ser feita seguindo dois tipos de padrões estabelecidos pela camada de enlace. No primeiro, os dispositivos se comportam como anunciante e *scanner*. O anunciante, como o nome sugere, se aproveita dos canais de anúncio para transmitir seus dados em *broadcast*, enquanto o *scanner* fica basicamente esperando alguma informação.

O segundo padrão a ser utilizado especifica que haja uma conexão entre o anunciante e o *scanner*, estabelecendo características de mestre e escravo, respectivamente. Através da técnica de acesso ao meio *Time Division Multiple Access* (TDMA), o mestre pode estabelecer conexão com diversos escravos, sendo que o escravo pode se conectar com

apenas um mestre por vez. Essa arquitetura de comunicação constitui a chamada *piconet*, que segue uma topologia de rede em estrela (GOMEZ; OLLER; PARADELLS, 2012).

A Figura 11 mostra o formato de um quadro BLE, onde o byte de preâmbulo é utilizado para sincronização, seguido por 4 bytes de endereço de acesso, mais 2 à 39 bytes de informação (sendo 2 o tamanho do cabeçalho fixo) e por fim o *Cyclic Redundancy Check* (CRC), responsável por detectar possíveis erros na recepção do pacote.

Figura 11 – Formato do quadro BLE



Fonte: Baseado em Texas Instruments (2016).

Ao se utilizar o primeiro padrão definido anteriormente, os dados que o anunciante envia não podem ser maiores que 32 dos 37 bytes de *payload* real do BLE, sendo que o seu endereço é enviado nos 6 bytes que antecedem os dados de *broadcast*.

Segundo Chang (2014), o fato do quadro BLE possuir poucos bytes, atribui à tecnologia a possibilidade de estabelecer conexões de forma rápida, transmitindo o pacote com a carga máxima de informações no pacote do *payload* em 300 μs.

2.5.3 Logical Link Control and Adaption Protocol (L2CAP)

Sendo uma versão otimizada e simplificada do mesmo protocolo utilizado no Bluetooth clássico, o *Logical Link Control and Adaption Protocol* (L2CAP) tem como objetivo multiplexar os dados dos protocolos superiores e inferiores, permitindo uma comunicação lógica entre os dispositivos (GOMEZ; OLLER; PARADELLS, 2012). Esse processo não faz uso de mecanismos de retransmissão ou controle de fluxo, presentes nas outras versões do Bluetooth.

2.5.4 Attribute Protocol (ATT) e Generic Attribute Profile (GATT)

O *Attribute Protocol* (ATT) é responsável por desempenhar funções de cliente e servidor em uma comunicação entre dois dispositivos. O servidor mantém uma estrutura de dados chamada de atributo, que é encapsulada e gerenciada pelo protocolo superior da pilha, o *Generic Attribute Profile* (GATT). O cliente pode acessar os atributos do servidor (modificando ou obtendo valores) através de solicitações e respostas ou indicações e confirmações (GOMEZ; OLLER; PARADELLS, 2012).

O *Generic Attribute Profile* (GATT) é responsável por definir uma estrutura que, a partir do ATT, realiza a descoberta de serviços e o envio ou requisição de características entre os dispositivos. Uma característica corresponde a um conjunto de dados, incluindo valores e propriedades. Gomez, Oller e Paradells (2012) especifica que tanto as características quanto os serviços são especificados através dos atributos do ATT, como por exemplo: um servidor que executa um serviço "sensor de temperatura"pode possuir uma característica "temperatura", que por sua vez utiliza de atributos para descrever o sensor, armazenar os valores adquiridos e especificar as unidades de medida.

2.5.5 Generic Access Profile (GAP)

O *Generic Access Profile* (GAP) está presente no topo da pilha protocolar do BLE, sendo assim, seu papel é especificar os padrões e procedimentos que as aplicações necessitam para realizar a descoberta de dispositivos, o gerenciamento da conexão e a segurança.

São definidos nesse protocolo quatro padrões com requisitos específicos: *broadcaster*, *observer*, periférico e central. Os dois primeiros padrões correspondem ao envio e recepção de dados sem que aconteça o estabelecimento de conexão entre os dispositivos. Já no padrão estabelecido na central, o dispositivo tem como característica iniciar e gerenciar diversas conexões, enquanto o periférico é projetado para dispositivos que apenas necessitam se conectar a uma central. Esse comportamento entre a central e o periférico explicita a relação mestre e escravo definida na subseção 2.5.2.

2.5.6 Security Manager Protocol (SMP)

Segundo Chang (2014), o *Security Manager Protocol* (SMP) é responsável pelo pareamento dos dispositivos e distribuição de chaves de segurança para garantir a identidade e criptografia de dados. Além disso, ele se encarrega de diminuir a carga de trabalho dos dispositivos escravos, por não possuírem um poder de processamento equivalente ao do mestre.

Para cumprir com os objetivos desse trabalho, são necessários apenas os parâmetros da comunicação entre os dispositivos, sem precisar adquirir dados de sensores específicos. Sendo assim, optou-se por realizar o padrão anunciante e *scanner* especificado na subseção 2.5.2. Com isso, os anunciantes apenas enviam suas identificações, que por sua vez são capturadas pelos *scanners* juntamente com o RSSI, sem a necessidade de se estabelecer conexão entre cada um dos dispositivos.

2.6 Aprendizado de Máquina

Aprendizado de máquina é um segmento da Inteligência Artificial (IA) que tem como objetivo construir algoritmos capazes de adquirir conhecimento de forma automática. Com isso, geram modelos de previsão e análise de dados que se baseiam em resultados conclusivos obtidos através de experiências anteriores (MONARD; BARANAUSKAS, 2003).

Segundo Qiu et al. (2016), diversas áreas, como medicina, astronomia e biologia, que geram uma quantidade complexa e massiva de dados, têm utilizado com frequência técnicas de aprendizado de máquina. Isto se da pelo fato de que estas técnicas podem fornecer possíveis soluções para a obtenção de informações escondidas entre uma enorme quantidade de dados.

Além disso, o aprendizado de máquina tem sido considerado a principal tecnologia para gerenciamento autônomo de redes inteligentes. Cui et al. (2018) afirma que em sistemas de IoT por exemplo, que têm se tornado cada vez mais dinâmicos, heterogêneos e complexos, o uso de aprendizado de máquina permite o desenvolvimento de aplicações mais eficientes. Com isso, é possível prover serviços melhores a fim de se obter mais usuários.

Apesar de ser uma ferramenta poderosa, não existe um método de aprendizado de máquina capaz de representar todo e qualquer sistema. Logo, de acordo com Monard e Baranauskas (2003), é necessário analisar as características de cada uma das técnicas disponíveis, verificando qual algoritmo melhor se encaixa na solução de um determinado problema. O aprendizado de máquina pode ser dividido em três categorias: supervisionado, não supervisionado e aprendizado de reforço, sendo as duas primeiras mais comumente utilizadas.

De acordo com Lovon-Melgarejo et al. (2018), os algoritmos pertencentes ao aprendizado de máquina supervisionado, *Supervised Learning Algorithms* (SLAs), consistem basicamente em duas etapas: treino e classificação. Na primeira fase é gerado um modelo de categorização a partir das amostras dos dados coletados. Esses dados são utilizados posteriormente para inferir a classificação das novas informações coletadas no sistema durante a segunda fase. De modo geral, quando novos dados são coletados, suas informações podem ser estimadas através de um histórico de comportamento do sistema previamente registrado. Como o resultado é conhecido, é possível estimar o desempenho do classificador.

O aprendizado de máquina não supervisionado é aplicado em modelos não possuem valores de saída, chamados de descritivos. Os padrões ou comportamentos do sistema são estimados através de um conjunto de dados previamente disponibilizado (ATHMAJA; HANUMANTHAPPA; KAVITHA, 2018).

Como neste trabalho a localização *indoor* é baseada no método *fingerprinting*, fica claro que através das informações apresentadas na subseção 2.3.4, os modelos de predição

utilizados correspondem ao aprendizado de máquina supervisionado. Sendo assim, as seções 2.6.1, 2.6.2 e 2.6.3 apresentam alguns dos SLAs mais utilizados, segundo Lantz (2015).

2.6.1 K-Nearest Neighbor (K-NN)

De forma geral, os classificadores do tipo *nearest neighbors*, ou vizinhos próximos, têm como objetivo classificar dados desconhecidos (não rotulados) de acordo com suas similaridades com classes previamente definidas. O *K-Nearest Neighbor* (K-NN) é um dos algoritmos que mais utilizam essa abordagem, além de serem um dos mais simples.

O K-NN recebe este nome pelo fato de utilizar informações de K vizinhos próximos para classificar os dados não rotulados. Após a fase de treino, onde vários itens já foram classificados em diversas categorias, o algoritmo identifica K registros nos dados de treinamento que são mais próximos através da similaridade. Com isso, na fase de testes, o dado desconhecido é atribuído à classe da maioria dos K vizinhos próximos.

Lantz (2015) afirma que a escolha de K impacta diretamente no resultado da classificação. Escolhendo um número muito grade é possível reduzir o impacto ou variância causada por dados ruidosos, no entanto, padrões pequenos, mas importantes podem ser desconsiderados. Se um único vizinho for considerado ($K = 1$), pode ocorrer a classificação errônea por conta de dados discrepantes registrados na fase de treino. A escolha do número de vizinhos próximos é realizada então, testando diversos valores para K em vários conjuntos de dados de treinamento e verificando qual oferece melhor desempenho na classificação.

Para calcular a proximidade entre um dado não rotulado e seus vizinhos próximos, o K-NN utiliza a distância Euclidiana, ou seja, o menor caminho entre os dois pontos. Seu cálculo pode ser realizado através do somatório apresentado na Equação 2.9, onde D é a distância entre p e q, que são os respectivos pontos, com n características. (LANTZ, 2015).

$$D = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{N}(p_i - q_i)^2} \tag{2.9}$$

2.6.2 Naive Bayes

A base para entendimento do algoritmo Naive Bayes pode ser analisada através do trabalho realizado no século XVIII pelo matemático Thomas Bayes, cujo desenvolveu os princípios fundamentais para descrever a probabilidade orientada a eventos, ou seja, como uma probabilidade de que ocorra um determinado evendo deve ser revisada à luz de uma nova informação adicional (LANTZ, 2015).

Essas relações entre os eventos dependentes podem ser descritas utilizando o teorema de Bayes, apresentado na Equação 2.10, onde a probabilidade de que ocorra o

evento A, dado que o evento B ocorreu é representada por $P(A|B)$. A probabilidade de ocorrer o evento B dentro de um universo em que já tenha ocorrido A é representada por $P(B|A)$ e $P(A)$ e $P(B)$ são as respectivas probabilidades de ocorrência de A e B.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2.10}$$

O treinamento do método Naive Bayes é realizado através da construção de uma tabela de probabilidades de acordo com os dados fornecidos *a priori*. Lantz (2015) traz como exemplo, a probabilidade de um *e-mail* que contem a palavra dinheiro ser considerado como *spam*, de acordo com a relação entre uma base de *spams* que possuíam tal palavra. De acordo com a Tabela 1, de 20 mensagens consideradas *spam*, 10 possuem a palavra dinheiro, ao passo que de 80 mensagens comuns, apenas 14 continham essa palavra.

Tabela 1 – Exemplo da fase de testes do algoritmo Naive Bayes

	Dinheiro		
verossimilhança	**Sim**	**Não**	**Total**
spam	10/20	10/20	20
mensagem comum	14/80	66/80	80
Total	24	76	100

Fonte: Modificado de Lantz (2015).

Através desses dados, a fase de testes conclui que se um novo *e-mail* for recebido e este conter a palavra dinheiro, a chance de ser considerado *spam* é de $10/20 = 0.5$ ou 50%, ao passo que ele também pode ser considerado uma mensagem comum com probabilidade igual a $14/80 = 0.175$ ou 17.5%.

Por fim, vale ressaltar que como a ocorrência de cada evento é considerada como independente, ou seja, que não influencia na ocorrência de outros eventos, a probabilidade de um novo conjunto de dados pertencer a este grupo será dada pela multiplicação de cada uma das razões encontradas na fase de treino, como no exemplo da Tabela 1.

2.6.3 Árvore de decisão

Classificadores que realizam o algoritmo da árvore de decisão utilizam uma estrutura de ramificação, que começa em um único ponto denominado raiz. Quando um novo dado começa a ser analisado, este é passado pelos pontos de decisão, exigindo a escolha de um determinado caminho, ou ramo, de acordo com os seus atributos. Quando uma decisão final é realizada, a árvore encontra seus nós de folha, também chamados de nós terminais, que fornecem o resultado esperado dado a série de eventos ocorrida.

O benefício de se utilizar a árvore de decisão está no fato de que a sua estrutura não se limita a apenas o uso do algoritmo de aprendizado de máquina. Após a criação

do modelo é possível gerar uma estrutura em formato legível por humanos, fornecendo uma visão sobre como e por qual motivo este modelo funciona bem ou mal para uma determinada tarefa.

Através dessas características, Lantz (2015) apresenta alguns exemplos de sistemas que utilizam o algoritmo da árvore de decisão:

- Análise de crédito e pontuações bancárias, nos quais os critérios fazem com que o cliente tenha um empréstimo aprovado ou não;
- Estudo de *marketing* de comportamento do cliente através de pesquisas de satisfação, sendo compartilhadas com agências de publicidades;
- Diagnóstico de condições médicas com base em medições laboratoriais, sintomas ou taxa de progressão da doença.

Outros pontos positivos em se utilizar a árvore de decisão estão no fato dela poder ser utilizada em grandes ou pequenos conjuntos de dados, sendo que é possível interpretar modelos pequenos sem a necessidade de utilizar técnicas matemáticas, tornando essa técnica mais eficiente.

Por outro lado, é comum que se tenha a sobreposição dos dados em um modelo, além de problemas na própria modelagem, devido à dependência de divisões paralelas aos ramos. Árvores muito grandes tendem a ficar mais difíceis de se interpretar, visto que certas decisões podem ser consideradas contraditórias. Por fim, é visto que apenas uma mudança nos dados de treinamento pode ocasionar grandes mudanças na lógica de decisão.

3 PROPOSTA

Para o cumprimento dos objetivo propostos, este trabalho está dividido em três etapas: estudo das técnicas de localização *indoor*, desenvolvimento de uma rede de dispositivos BLE e análise e testes de diferentes técnicas de aprendizado de máquina aplicadas na localização dos componentes presentes na rede implementada.

3.1 Estudo das técnicas de localização *indoor*

Nesta etapa é necessário compreender os diversos métodos utilizados para se obter a localização *indoor*, desde os parâmetros da comunicação que poderão ser utilizados (RSSI, AoA, ToA ou TDoA), até a implementação do algoritmo responsável por determinar a posição dos equipamentos.

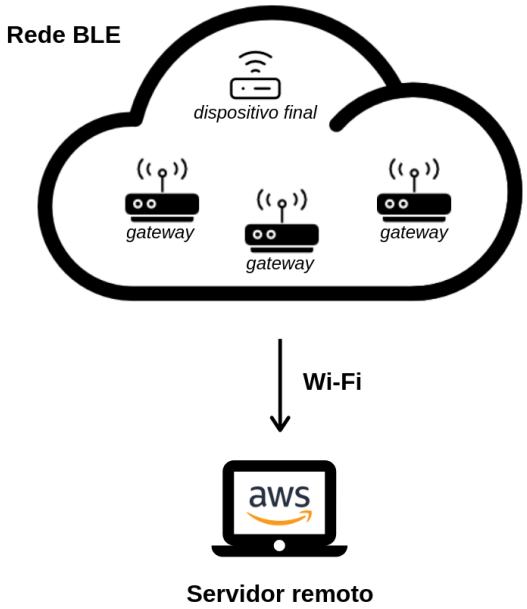
Por conta da praticidade e custo de aquisição dos equipamentos, optou-se por utilizar o parâmetro RSSI com a técnica *fingerprint* para realizar a localização *indoor*. Serão realizados testes variando a quantidade de *gateways* para verificar a precisão nas medições.

3.2 Implementação de uma rede de dispositivos BLE

A partir da definição da técnica de localização utilizada, parte-se então para a implementação de uma rede de dispositivos se comunicando via BLE. A arquitetura dessa rede constitui uma topologia em estrela, onde os dispositivos finais, ou nós, enviam suas informações aos *gateways* ou concentradores. Eles ficam responsáveis por enviar as informações coletadas e a identificação dos nós para um servidor remoto, onde serão aplicados os algoritmos de aprendizado de máquina para determinar a localização deles. A Figura 12 apresenta um possível cenário de testes, onde estão presentes três *gateways* que coletam o RSSI do nó desconhecido, chamado de dispositivo final, através dos seus pacotes de anúncio.

Os componentes da rede que será desenvolvida estão descritos a seguir, sendo que será utilizado o serviço de computação em nuvem da Amazon, Amazon Web Services (AWS) como servidor.

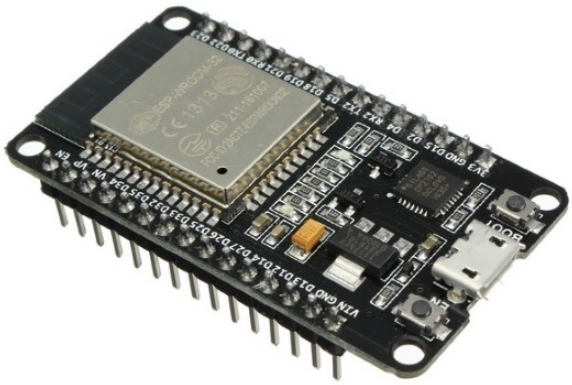
Figura 12 – Descrição do cenário de testes



Fonte: Desenvolvida pelo autor.

- *Gateway*: Para realizar esta função foi escolhida a plataforma de desenvolvimento ESP32[1], mostrada na Figura 13, que possui as interfaces de comunicação sem fio BLE e Wi-Fi.

Figura 13 – Plataforma de desenvolvimento ESP32



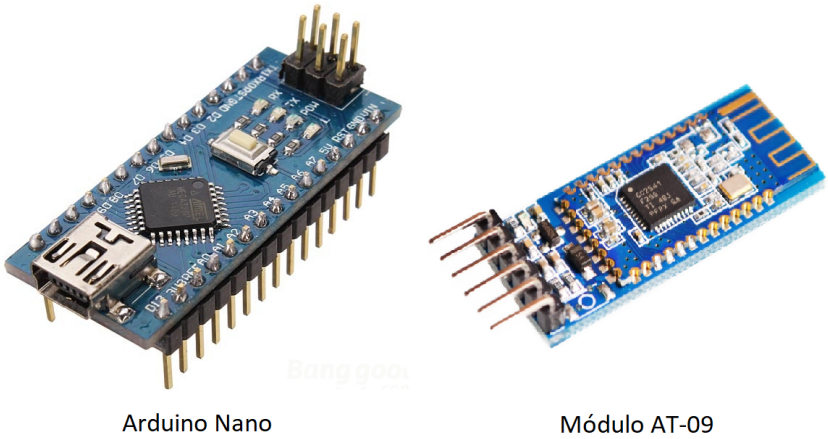
Fonte: Filipeflop.

Com essas tecnologias, além da praticidade de implementação por conta de bibliotecas disponíveis para tal, o ESP32 é uma opção de custo viável para se desenvolver o *gateway*. Seu papel é se comportar como *scanner* BLE, identificando os pacotes de anúncio que serão enviados pelos dispositivos finais e encaminhando os dados da comunicação (RSSI e endereço MAC do dispositivo) via Wi-Fi para o servidor remoto, através de uma Application Programming Interface (API) RESTful desenvolvida no mesmo.

[1] https://www.espressif.com/en/products/hardware/esp32/overview

- Dispositivos finais: Irão realizar o papel de *beacons*, enviando em um intervalo de tempo pré determinado suas informações (nome, endereço MAC, identificador de serviço, entre outras características) através dos pacotes de anúncio do BLE. Para desenvolvê-los, será utilizado o módulo BLE AT-09, que utiliza o microcontrolador CC2541[2] da Texas Instruments. Seu papel será enviar constantemente pacotes de anúncio com a sua identificação. Para realizar suas configurações nele, será utilizada a plataforma Arduino Nano, no qual serão enviados comandos AT através da comunicação serial entre as duas plataformas. Após os devidos ajustes, o módulo BLE pode ser apenas alimentado pelo Arduino, ou através de uma fonte de alimentação externa. A Figura 14 apresenta a plataforma Arduino Nano, bem como o módulo BLE AT-09.

Figura 14 – Plataforma Arduino Nano e Módulo BLE AT-09



Fonte: Baseado em Filipeflop.

3.3 Análise e testes das técnicas de aprendizado de máquina

Os testes consistem em aplicar os SLAs estudados na base de dados gerada no servidor. As comparações verificam quais classificadores obtêm maior precisão na estimativa da localização. Estuda-se a possibilidade de utilizar bibliotecas específicas da linguagem *Python* para realizar os testes com os algoritmos de aprendizado de máquina.

[2] http://www.ti.com/lit/ds/symlink/cc2541.pdf

3.4 Cronograma de atividades previstas

A Tabela 2 apresenta o cronograma de atividades que são previstas para o cumprimento dos objetivos propostos no trabalho.

Tabela 2 – Cronograma das atividades previstas

Tarefa	Mês							
	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul
A1	√							
A2		√	√					
A3			√	√				
A4				√	√			
A5					√			
A6						√		
A7						√	√	√
A8							√	√

- A1: Ajustes finais no documento do pré projeto;
- A2: Desenvolver e configurar os protótipos de *gateways* e *beacons* BLE;
- A3: Configurar o servidor AWS;
- A4: Desenvolvimento do banco de dados e escolha da API RESTful;
- A5: Testes de integração dos *gateways* com o servidor;
- A6: Montar cenários de testes com a rede BLE;
- A7: Desenvolver programas de teste com os SLAs selecionados de acordo com os dados coletados no servidor;
- A8: Comparar resultados e elaborar o documento final.

REFERÊNCIAS

AL-FUQAHA, A. et al. Internet of things: A survey on enabling technologies, protocols, and applications. *Communications Surveys & Tutorials, IEEE*, IEEE, USA, v. 17, n. 4, p. 2347–2376, 2015. ISSN 1553-877X. Citado 3 vezes nas páginas 9, 11 e 12.

ATHMAJA, S.; HANUMANTHAPPA, M.; KAVITHA, V. A survey of machine learning algorithms for big data analytics. In: . Coimbatore, Tamil Nadu, India: [s.n.], 2018. v. 2018-January, p. 1 – 4. Disponível em: <http://dx.doi.org/10.1109/ICIIECS.2017.8276028>. Acesso em: 20 nov. 2018. Citado na página 25.

CHANG, K.-H. Bluetooth: A viable solution for iot? *Wireless Communications*, IEEE, USA, p. 6–7, 2014. Citado 3 vezes nas páginas 21, 23 e 24.

CUI, L. et al. A survey on application of machine learning for internet of things. *International Journal of Machine Learning and Cybernetics*, Springer Berlin Heidelberg, Berlin/Heidelberg, v. 9, n. 8, p. 1399–1417, 2018. ISSN 1868-8071. Citado na página 25.

DARROUDI, S. M.; GOMEZ, C. Bluetooth low energy mesh networks: A survey. *Sensors (Switzerland)*, v. 17, n. 7, 2017. ISSN 14248220. Citado na página 21.

FARID, Z.; NORDIN, R.; ISMAIL, M. Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications*, Hindawi Publishing Corporation, v. 2013, 2013. ISSN 2090-7141. Citado 3 vezes nas páginas 15, 19 e 20.

GOMEZ, C.; OLLER, J.; PARADELLS, J. Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors (Basel, Switzerland)*, Molecular Diversity Preservation International (MDPI), v. 12, n. 9, p. 11734–11753, 2012. ISSN 1424-8220. Citado 3 vezes nas páginas 22, 23 e 24.

GU, Y.; LO, A.; NIEMEGEERS, I. A survey of indoor positioning systems for wireless personal networks. *Communications Surveys & Tutorials, IEEE*, IEEE, USA, v. 11, n. 1, p. 13–32, 2009. ISSN 1553-877X. Citado 2 vezes nas páginas 15 e 20.

INDOOR positioning system based on BLE location fingerprinting with classification approach. *Applied Mathematical Modelling*, v. 62, p. 654 – 663, 2018. ISSN 0307-904X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0307904X18302841>. Acesso em: 25 out. 2018. Citado 2 vezes nas páginas 19 e 22.

LANTZ, B. *Machine Learning with R*. 2nd. ed. [S.l.]: Packt Publishing, 2015. ISBN 1784393908, 9781784393908. Citado 3 vezes nas páginas 26, 27 e 28.

LIU, H. H. et al. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, IEEE, USA, v. 37, n. 6, p. 1067–1080, 2007. ISSN 1094-6977. Citado 3 vezes nas páginas 10, 12 e 13.

LOUREIRO, A. A. et al. Redes de sensores sem fio. In: SN. *Simpósio Brasileiro de Redes de Computadores (SBRC)*. [S.l.], 2003. p. 179–226. Citado na página 11.

LOVON-MELGAREJO, J. et al. Supervised learning algorithms for indoor localization fingerprinting using ble4.0 beacons. In: . Arequipa, Peru: [s.n.], 2018. v. 2017-November, p. 1 – 6. Disponível em: <http://dx.doi.org/10.1109/LA-CCI.2017.8285716>. Acesso em: 10 nov. 2018. Citado na página 25.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas Inteligentes-Fundamentos e Aplicações*, v. 1, n. 1, p. 32, 2003. Citado na página 25.

QIU, J. et al. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, v. 2016, n. 1, p. 67, May 2016. ISSN 1687-6180. Disponível em: <https://doi.org/10.1186/s13634-016-0355-x>. Acesso em: 22 nov. 2018. Citado na página 25.

RAPPAPORT, T. *Wireless Communications: Principles and Practice, 2nd Edition*. [S.l.: s.n.], 2001. v. 2. ISBN 0130422320. Citado na página 12.

REZENDE, E. M.; YNOGUTI, C. A. Tecnologias para localização interna (*indoor location*). *II RST - Seminário de Redes e Sistemas de Telecomunicações*, INATEL, BR, 2015. ISSN 2358-1913. Citado na página 9.

SADOWSKI, S.; SPACHOS, P. Rssi-based indoor localization with the internet of things. *IEEE Access*, IEEE, v. 6, p. 30149–30161, 2018. Citado 7 vezes nas páginas 9, 10, 12, 14, 16, 17 e 20.

SAHINOGLU, Z.; GEZICI, S.; GÜVENC, I. *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. [s.n.], 2008. ISBN 9781139472319. Disponível em: <https://books.google.com.br/books?id=ygXzSEzQy4oC>. Acesso em: 20 out. 2018. Citado 6 vezes nas páginas 13, 14, 15, 16, 17 e 18.

SAMIE, F.; BAUER, L.; HENKEL, J. Iot technologies for embedded computing: A survey. In: ACM. *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*. [S.l.], 2016. p. 8. Citado 2 vezes nas páginas 11 e 12.

TEXAS INSTRUMENTS. *Bluetooth low energy Beacons*. [S.l.], 2016. Disponível em: <http://www.ti.com/lit/an/swra475a/swra475a.pdf>. Acesso em: 17 nov. 2018. Citado 2 vezes nas páginas 22 e 23.

XIAO, Y. et al. Full-duplex machine-to-machine communication for wireless-powered internet-of-things. In: . [S.l.]: IEEE, 2016. p. 1–6. ISBN 9781479966646. ISSN 1938-1883. Citado na página 11.

ČABARKAPA, D.; GRUJIć, I.; PAVLOVIć, P. Comparative analysis of the bluetooth low-energy indoor positioning systems. In: *2015 12th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS)*. [S.l.: s.n.], 2015. p. 76–79. Citado 2 vezes nas páginas 21 e 22.