

# **Ferramenta de análise de soluções supervisionadas para localização indoor**

**Alexandre Rodrigues Pelagaggi Nunes**

*Dissertação apresentada à Escola Superior de Tecnologia e Gestão para obtenção do Grau de Mestre em Engenharia Eletrotécnica e Industrial*

Trabalho realizado sob a orientação de:

Rui Pedro Sanches de Castro Lopes

Maurício Henrique Costa Dias

**Bragança**  
Junho de 2019



## **Dedicatória**

Dedico esse trabalho a todos que contribuíram para minha evolução profissional e pessoal.



## **Agradecimentos**

Agradeço aos meus pais pela contribuição que realizaram em minha carreira e formação de caráter, Alexandre Pelagaggi Nunes e Luciana Rodrigues Pelagaggi Nunes. A minha Esposa pelo suporte emocional durante todo meu caminho acadêmico e profissional, Maria Luíza de Souza Oliveira Ottoni ao meu sogro por todas as orientações acadêmicas e tempo despendido para os mais diversos fins que vieram a contribuir para minha formação, Dr. Adacto Benedicto Ottoni. Agradeço aos meus Irmãos Fabio Rodrigues Pelagaggi Nunes e Bianca Rodrigues Pelagaggi Nunes e minha cunhada Marianna de Souza oliveira Ottoni pelo fomento do meu sucesso. Aos meus Orientadores Dr. Rui Pedro Lopes e Dr. Maurício Henrique Costa Dias pelo suporte e orientação que possibilitaram minha formação durante o programa de dupla diplomação. Ao Dr. Ronney Mancebo e a Dra. Angela Norte pela condução absolutamente exemplar, inovadora e carinhosa dos programas de cooperação internacionais disponibilizados no Centro Federal de educação tecnológica do Rio de Janeiro. Agradeço à Dra. Ana Lucia Ferreira de Barros e ao Dr. Mauro Sandro dos Reis pelas orientações fornecidas durante minha participação em programas de pesquisa no CEFET. Agradeço aos meus colegas de curso Igor Andrade Araujo, Felipe Augusto Correa Pegado, Vanessa Pereira Cavalcanti e Jamila Café pela amizade e companheirismo durante o curso de Engenharia Eletrônica. Agradeço a Alunas Hannah Alves e pelo suporte intelectual extraordinário e assustadoramente excepcional prestado a todos os alunos de engenharia eletrônica.



## Resumo

Tendo em vista a gama de aplicações capazes de utilizar dados de localização afim de agregar funcionalidades à vidas das pessoas, seja viabilizando navegabilidade, provendo controle de conteúdo entre outras, os avanços sociais e comerciais trazidos por sistemas de navegação via satélite, é natural que se busque formas de viabilizar serviços fundamentados em localização, que compreendam ambientes cuja os métodos disponíveis pequem em precisão ou tempo de resposta que se enquadrem em cenários que possuem grande variabilidade.

Visando viabilizar o estudo dessas tecnologias, bem como garantir a qualidade da implementação prática das mesmas, é fornecido nesse trabalho um contributo capaz de combinar técnicas e ferramentas existentes, em modelo alinhado com tecnologias estáveis e de grande demanda no mercado de software, em um instrumento capaz de analisar, sob diversas métricas, a performance de soluções de localização indoor. A ferramenta foi pensada para atender diversas tecnologias e técnicas e foi implementada para técnicas de correlação de dados comumente denotadas fingerprinting, frequentemente divididas em duas fases, nomeadamente de levantamento de dados de comparação e de utilização.

Foi realizado o levantamento de dados no campus de Bragança do Instituto Politécnico de Bragança (IPB), na Escola Superior de Tecnologia e Gestão (ESTIG), com o propósito de demonstrar a aplicabilidade da ferramenta em ambientes universitários.

O objetivo do trabalho descrito por meio desse, é proporcionar praticidade de configuração de um levantamento de dados, flexibilidade na configuração do uso dos dados, compatibilização com algoritmos e bibliotecas de aprendizagem de máquina de alto nível e a análise dos resultados sob diversas métricas.

Palavras Chave: Indoor-Positioning, RSSI, Wifi, fingerprinting, magnetômetro



# **Abstract**

Given the crescent location-based applications and services that add value to the everyday life of most of the current population, providing navigation or localization capability or even administrative resources, the commercial potential brought by global navigation satellite systems, it becomes natural that researches arise in order to provide these capabilities where the current implementation lacks sufficiency.

Aiming to improve the studies of these new technologies, as well as assure their quality, this work provides a mean of combining existing technologies and tools generating metrics about the implemented solutions. This tool is thought to comprehend indoor locations solutions in general and it is applied to fingerprinting applications, which are usually categorized into two phases, data collection, and data correlation.

In order to demonstrate the proposed tool in its full capability, a data site survey was performed in the campus of the Instituto Politécnico de Bragança (IPB) in School of Technology and Management (ESTIG).

The Goal of this work is to provide a tool that automates the data visualization and metrification of indoor locations solutions, providing data usage flexibility with a high customization level in the current commercial recurrent tools.

**Key Words:** Indoor-Positioning, RSSI, Wifi, fingerprinting, magnetometer

# Índice Geral

Dedicatória.....	iii
Agradecimentos .....	v
Resumo .....	vii
Abstract.....	ix
Índice Geral .....	x
Lista de Siglas.....	xii
Índice de Figuras .....	xiii
Índice de Tabelas .....	xiv
Capítulo 1 Introdução a localização indoor.....	15
1.1. Contextualização em sistemas de posicionamento .....	15
1.2. Contrapontos do uso de GNSS em contexto indoor .....	16
1.3. Justificativa do Trabalho.....	17
1.4. Organização do trabalho .....	18
1.5. Sumário .....	20
Capítulo 2 Estado da Arte de Técnicas de Localização indoor.....	21
2.1. Tecnologias e metodologias utilizadas para localização indoor.....	21
2.2. Alinhamento com escopo do trabalho .....	27
2.3. Implementações em localização Indoor.....	28
2.4. Sumário .....	30
Capítulo 3 Tecnologias adotadas.....	31
3.1. Contextualização com a arquitetura implementada .....	31
3.2. Ferramentas utilizadas .....	32
3.3. Sumário .....	36
Capítulo 4 Trabalho efetuado .....	37
4.1. Escopo do Trabalho .....	37
4.2. Arquitetura do projeto.....	38
4.3. Funcionalidades dos elementos .....	39
4.4. Sumário .....	52
Capítulo 5 Análise e Discussão de Resultados.....	53
5.1. interfaces com utilizador.....	55
5.2. Seleção de resultados .....	56
5.3. Casos de uso.....	59
5.4. Análise entre algoritmos e sinais .....	65
5.5. Sumário .....	66

Capítulo 6	Conclusão e trabalhos futuros.....	69
6.1.	Trabalhos futuros .....	71
Bibliografia.....		73

## Lista de Siglas

POIs	Pontos de Interesse.
SDK	Software Development kit
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
DOP	Dilution of precision
RTK	Real Time Kinematic
LPWANs	Low Power Wide Area Networks
VANTs	Veículos aéreos não tripulados
ESTIG	Escola Superior de Tecnologia e Gestão de Bragança
IPB	Instituto Politécnico de Bragança
IoT	Internet das Coisas
USA	United States of America
TDOA	Time difference of Arrival
TOA	Time of arrival
TOF	Time of Flight
AOA	Angle of Arrival
BLE	Bluetooth Low Energy
PDR	Pedestrian Dead Reckoning
CMX	Cisco Connected Mobile Experiences
SVG	Scalable Vector Graphics
MQTT	MQ Telemetry Transport

# Índice de Figuras

Figura 1: Multilateralidade usando medições TDOA e TOA com hipérbole e círculos, respectivamente, como possíveis emissoras locais [6].....	22
Figura 2: Posicionamento via ângulo de chegada .....	23
Figura 3: efeito Doppler (Jose Wudka 1998) .....	24
Figura 4: possíveis casos de uso de triangulação .....	25
Figura 5: Linearização temporal de múltiplos TOAs .....	26
Figura 6: Localização via Trilateração .....	26
Figura 7: Arquitetura das ferramentas de computação .....	32
Figura 8: Arquitetura do projeto .....	38
Figura 9: Gerenciamento de Sítios de estudo .....	40
Figura 10: Legenda dos marcadores e respectivas cores na página de configuração .....	40
Figura 11: Caixas de diálogo para edição.....	40
Figura 12: Página de serialização de dados.....	41
Figura 13: Processo de edição de Algoritmos .....	42
Figura 14: Criação e utilização de aplicação de teste .....	43
Figura 15: Seleção de ordem de treinamento .....	44
Figura 16: Distribuição de erro percentual.....	44
Figura 17: Espalhamento bidimensional sob a planta baixa.....	45
Figura 18: Espalhamento bidimensional simplificado .....	45
Figura 19: Gráfico de espalhamento 3d com filtragem por faixa de erro.....	46
Figura 20: Estatísticas acerca do resultado selecionado e sua distribuição de erro.....	46
Figura 21: Utilização da aplicação Android para coleta de dados .....	47
Figura 22: Implementação CRUD da API.....	48
Figura 23: Modelo relacional do banco de dados.....	49
Figura 24: Pontuação cubicamente proporcional ao erro .....	52
Figura 25: Andar Térreo da ESTIG .....	54
Figura 26: Andar dos laboratórios da ESTIG .....	54
Figura 27: Resultado da funcionalidade de serialização de dados .....	55
Figura 28: Resultado da organização e disposição dos resultados .....	56
Figura 29: Distribuições de erro para as métricas de “W_T_Tree” .....	59
Figura 30: Conjunto de treino (a) e teste (b) da aplicação “W_T_TREE” .....	60
Figura 31: Gráficos disponibilizado pela ferramenta para aplicação “W_T_TREE” avaliada sob a métrica avaliada pela métrica Intervalo. ....	61
Figura 32: Conjunto de treino (a) e teste (b) da aplicação “W_KNN_LABS” .....	62
Figura 33: Gráficos disponibilizado pela ferramenta para aplicação “W_KNN_LABS” avaliada sob a métrica avaliada pela métrica Cúbico 7.5 .....	63
Figura 34: Conjunto de treino (a) e teste (b) da aplicação “I_W_KNN” .....	64
Figura 35: Gráficos disponibilizado pela ferramenta para aplicação “I_W_KNN” avaliada sob a métrica avaliada pela métrica Cúbico 7.5 .....	64
Figura 36: Distribuição para sinais magnéticos.....	65
Figura 37: Distribuição para sinais wifi + magnético .....	65
Figura 38: Distribuição para sinais wifi .....	66

## Índice de Tabelas

Tabela 1: Algoritmos e sinais implementados no servidor de localização.....	49
Tabela 2: Relação de pontos de interesse coletados em cada experiência .....	53
Tabela 3: Resultados dos treinamentos ESTIG andar térreo .....	56
Tabela 4: Resultados dos treinamentos ESTIG andar térreo para aplicação “W_T_TREE” .....	60
Tabela 5: Estatísticas da distribuição obtida pela aplicação “W_T_TREE” avaliada sob a métrica avaliada pela métrica Intervalo .....	61
Tabela 6: Resultados dos treinamentos ESTIG andar dos laboratórios “W_KNN_LABS” .....	62
Tabela 7: Estatísticas da distribuição obtida pela aplicação “W_KNN_LABS” avaliada pela métrica Cúbico 7.5 .....	63
Tabela 8: Métricas da distribuição obtida pela aplicação “I_W_KNN” avaliada pela métrica Cúbico 7.5.....	65

# Capítulo 1      **Introdução a localização indoor**

A noção de localização e a sua demanda é notada há muitos séculos pela humanidade, tida como essencial desde tempos em que a geolocalização era obtida com instrumentos primitivos como astrolábio e quadrante, que aferiam altura do sol e a altura da estrela polar para obter latitude e longitude. Também se destaca na história a importância da bússola, principalmente desde a era dos descobrimentos. Os avanços na capacidade tecnológica levaram a humanidade a colocar emissores em órbita do planeta afim de utilizar uma rede de satélites para determinar a localização de um objeto, capaz de receber sinais e utilizar os mesmos e suas diferenças de tempo de recepção para calcular sua localização. A importância da localização é expressa em diversos níveis, que variam desde aplicações militares e estratégicas ou até mesmo em contexto doméstico.

## **1.1. Contextualização em sistemas de posicionamento**

Atualmente, a forma mais utilizada para geoposicionamento é o sistema de GNSS, comumente expresso pela sigla de Global Positioning System (GPS). Ele é, entretanto, apenas um dos sistemas de GNSS, GPS é a implementação dos Estados Unidos da América (USA) e coexiste atualmente com o GALILEO, da União Europeia, GLONASS da Rússia e o BeiDou da China. Esses sistemas possuem importâncias estratégicas que impactam a capacidade de industrialização em setores como automobilística autônoma e assistida, veículos aéreos não tripulados (VANTs), popularmente conhecidos pelo acrônimo da língua inglesa (DRONES), desenvolvimento urbano de cidades inteligentes entre outros serviços como navegabilidade para população.

Os sistemas GNSS possuem tal importância que grandes potências mundiais como Japão e USA vem prospectando a implementação de uma força armada dedicada a proteção de propriedades espaciais [1].

Dados de posicionamento além de atrair projetos de pesquisas por grandes empresas como a Uber e sua cidade de teste de veículos autônomos [2], projetos em parcerias público privadas como o projeto dinamarquês TAPAS que por meio de uma rede de estações amplificadoras, Real Time Kinematic (RTK) obteve precisão de 1cm<sup>3</sup> em movimento[3].

## **1.2. Contrapontos do uso de GNSS em contexto indoor**

Fatores como dependência da linha de visão com satélites e alteração de caminho do sinal por reflexão em ambientes de densa urbanização vertical ou atenuação do sinal pelas estruturas dos edifícios, fazem com que o uso de sistemas GNSS seja precário ou não aplicável. Em ambientes indoor, a atenuação do sinal e múltipla reflexão do mesmo, faz com que a diluição da precisão (DOP) aumente de forma a impossibilitar diversas aplicações.

Apesar dos grandes avanços na tecnologia de geolocalização provocarem uma certa estabilidade em aplicações orientadas a localização via GNSS, devido a insatisfatória ou inexistente contribuição para soluções de localização indoor torna-se necessária a busca por novas tecnologias capazes de atuar em cenários onde esse não é capaz.

Entretanto, não existe até hoje um modelo utilizado por padrão que possua baixo custo de implementação, manutenção e alta replicabilidade, com capacidade de atuar em zonas que o GNSS não é capaz.

A localização indoor surge da necessidade de complementar ou substituir as tecnologias de localização e navegação convencionais, visto que esses sistemas possuem ineficácia. Essa necessidade também pode ser percebida ao se observar que a maioria da população atualmente segue modelo social baseado em uma atuação profissional e vida social em ambientes fechados.

A ferramenta proposta no escopo desse trabalho visa viabilizar aplicações nesses ambientes onde geralmente a humanidade passa a maior parte do tempo, além de armazenar bens da indústria e comércio. A fim de atender demandas como rastreabilidade de ativos por parte das empresas, é necessário enfrentar o desafio de obter a localização indoor, aplicações nesse sentido têm tido demanda crescente com a crescente digitalização da indústria (Indústria 4.0) e, portanto, objeto de pesquisa.

### **1.3. Justificativa do Trabalho**

Tendo em vista a crescente demanda por soluções em localização indoor, é natural que seja questionada a metodologia de avaliação das mesmas. Frequentemente se utiliza como métrica da qualidade da proposta a “precisão”, métrica que se baseia em uma propriedade estatística acerca de uma distribuição de erros obtidas de um espaço amostral da solução em questão, por exemplo média aritmética ou mínimo erro estimado (MEE).

Devido à variedade de métricas que podem ser obtidas a partir de uma distribuição de erro e tendo em vista a necessidade de validar a constância de métricas de performance em metodologias abordadas para determinar a localização indoor, é questionado, por este trabalho, a variabilidade da adequação de uma determinada métrica a aplicações variadas. Questiona-se se, sob determinado aspecto, a aplicação que prioriza uma métrica pode ser considerada mais adequada ainda que venha a possuir maior erro máximo ou a faixa de erro mais ampla. O questionamento surge com a possível existência de aspectos de impacto qualitativo e inerentes ao contexto em que a solução venha a se enquadrar. Caberia, então, ao consumidor final da solução de localização a escolha da métrica que melhor se enquadra em sua necessidade. Justifica-se, assim, uma ferramenta capaz de automatizar essa análise em contexto de soluções indoor.

Esse questionamento surge como uma crítica à forma atual utilizada para se determinar a performance, que determina uma precisão como uma métrica que parte de uma distribuição de erros. Essa distribuição normalmente é gerada devido à dificuldade associada na geração de um modelo de propagação fiável a um ambiente genérico. É justificável a dedução a partir de um espaço amostral, o comportamento esperado de um modelo de aferição do posicionamento. Entretanto, devido à diversidade de métricas disponíveis e variabilidade de cenários aplicáveis aos modelos, surge a demanda por uma

forma de determinar se o modelo atua qualitativamente de forma superior avaliado prioritariamente com uma determinada métrica quando em comparação com outra métrica.

Portanto, este trabalho demonstra como pode auxiliar no desenvolvimento, utilização e análise de soluções para localização indoor, viabilizando a comparação entre diferentes abordagens sob diferentes métricas além da comparação entre um mesmo modelo sob diferentes métricas e viabilizando a adequação de um modelo a utilização de parâmetros que melhor se enquadram com a prioridade do consumidor final.

## **1.4. Organização do trabalho**

Este documento encontra-se organizado em 6 capítulos:

### **Capítulo 1 Introdução a localização indoor**

É apresentada uma introdução ao conceito e importância da localização. Os desafios que a localização indoor busca vencer. O impacto do trabalho proposto no tema e a disposição organizacional do texto.

### **Capítulo 2 Estado da Arte de Técnicas de Localização indoor**

Devido à variedade de tecnologias e técnicas que podem ser utilizadas para aferir ou estimar a localização em ambientes indoor, foi realizado um estudo comparativo, apresentado no Capítulo 2. Este estudo tem foco na viabilidade, escalabilidade e aplicabilidade de soluções em âmbito acadêmico, comercial e industrial. O estudo foi utilizado para embasar a decisão de aplicação do trabalho em uma das possíveis abordagens de localização indoor.

### **Capítulo 3 Tecnologias utilizadas**

Será apresentado também no escopo desse contributo o estudo de como tecnologias de software recentes como computação em nuvem podem vir a auxiliar na obtenção de maior velocidade e menor custo de implementação bem como trazer maior escalabilidade nesse contexto. É expresso nesse capítulo a influência na escolha da arquitetura e das

ferramentas pela motivação de obtenção de um produto expansível e integrável em ambientes comerciais, buscando fugir de um resultado estritamente acadêmico.

Será abordada a metodologia de desenvolvimento do projeto quanto às escolhas de ferramentas e limitações que essas viriam a impor na concepção da arquitetura. É feita uma análise voltada da capacidade de interação entre outras possíveis expansões da ferramenta.

#### **Capítulo 4 Implementação do trabalho**

Nesse capítulo será apresentada a arquitetura e a motivação por trás da arquitetura adotada. O capítulo será dividido entre as partes integrantes da solução e cada uma delas será explicada individualmente quanto a sua importância no contexto de geração e disposição de métricas acerca de soluções de localização indoor. Sendo os principais componentes:

- Ferramenta administrativa
- Aplicação Android
- API + Banco de dados
- Algoritmos de Aprendizagem de máquina.

#### **Capítulo 5 Análise e Discussão de Resultados**

Nesse capítulo será apresentado uma demonstração de resultados da ferramenta posta em prática em dois cenários, nas instalações da Escola Superior de Tecnologia e Gestão de Bragança (ESTIG), e em um ambiente residencial. Será discutido como a ferramenta viabilizou a comparação de resultados entre as diferentes métricas e algoritmos bem como a comparação entre diferentes escalas e ambientes.

#### **Capítulo 6 Conclusões e trabalho futuro**

Nesse capítulo serão apresentadas as considerações finais acerca do trabalho desenvolvido e como o mesmo se portou quanto as aplicações futuras e possíveis expansões do trabalho.

## **1.5. Sumário**

O trabalho realizado pretende o fornecimento de uma aplicação capaz de se integrar em soluções de localização indoor, objetivando fornecer maiores detalhes acerca do comportamento esperado da mesma a partir de um espaço amostral obtido. O contributo proposto pelo autor busca atuar como facilitador no processo de análise e escolha de soluções em localização indoor.

O objetivo do trabalho é a produção de uma ferramenta que disponibilize, via interfaces gráficas, a funcionalidade de configuração de locais para aplicação de algoritmos de localização indoor. Será fornecido também uma secção de análise de resultados do método de localização, a qual irá gerar métricas estatísticas da distribuição de erro obtida a partir dos resultados gerados. A solução tem foco na utilização de dados obtidos utilizando a metodologia de localização fingerprinting.

# **Capítulo 2 Estado da Arte de Técnicas de Localização indoor**

A fim de auxiliar na tomada de decisão, esse capítulo apresenta um estudo de soluções comerciais, projetos de pesquisa e desenvolvimento e soluções de localização indoor, tendo como objetivo refletir o estado atual da tecnologia, este estudo será dividido em duas partes, compostas respectivamente pelas técnicas de localização e pelas soluções que utilizam essas técnicas. Devido à diversidade de soluções existentes, a fim de evitar fuga ao objetivo do trabalho, as soluções demonstradas foram selecionadas com base na análise das técnicas utilizadas.

As métricas que esse estudo utiliza assumem elevada demanda por aplicações em Internet das Coisas (IoT) e cidades inteligentes, que se deve à convergência em aplicações práticas e viabilidade de um produto nesse contexto. Será, portanto, priorizado o levantamento de possíveis aplicações com alta escalabilidade e replicabilidade. Tem-se como foco o suporte a aplicações integráveis em ambientes residenciais, urbanos e industriais.

Este capítulo analisa as técnicas existentes, definindo uma diretriz que demonstra a proposta do trabalho, apresenta soluções que tenham sinergia com foco escolhido e qualifica e delimita o escopo do trabalho perante o conteúdo apresentado.

## **2.1. Tecnologias e metodologias utilizadas para localização indoor**

Existem diversas abordagens e soluções existentes capazes de atuar a fim de obter a localização de um elemento em ambientes indoor. Serão apresentadas as abordagens mais frequentes, as propriedades de propagação de sinais e métodos convencionalmente exploradas [4]. Será apresentado, a seguir, um estudo das propriedades de sinais comumente exploradas e as principais metodologias para determinar a localização.

## Propriedades de sinais

As propriedades inerentes a propagação de sinais RF normalmente exploradas são.

### *TOA, Time of Arrival*

Time of arrival (TOA) é uma forma primitiva e eficiente de determinar a distância percorrida por um sinal a partir da diferença de horários de partida e chegada do mesmo, portanto também denominada Time of Flight (TOF).

$$d = v * T$$

$d$ =distância percorrida pelo sinal,  $v$  = velocidade de propagação do sinal e  $T$  = tempo de viagem do sinal. Nota-se que a sincronização é necessária entre dispositivos afim de obter o tempo de emissão e recepção, permitindo o correto calculo da distância, o deesafio de sincronização tema recorrente e diferentes autores buscam formas de contrapo-lo [5].

### *TDOA, Time difference of Arrival*

Time difference of Arrival é uma técnica que utiliza a diferença de tempo de chegada entre sinais emitidos de forma sincronizada, como forma de calcular distância dos emissores que já possuem localização conhecida. A técnica possui como vantagem dispensar da necessidade de sincronismo com o sensor móvel, sendo necessário somente sincronismo entre os emissores que normalmente dispõem de alimentação contínua e menor impacto financeiro na escalabilidade [6].

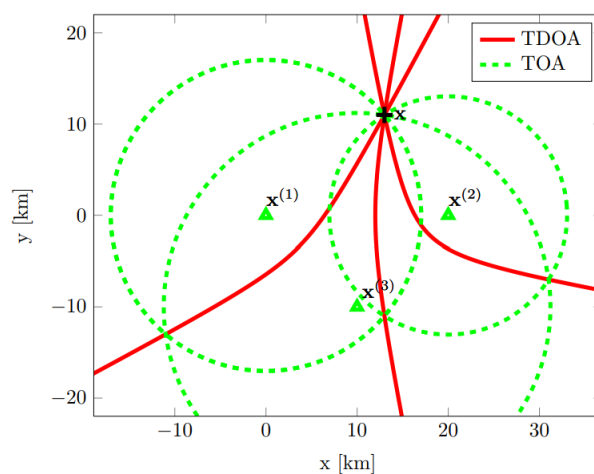


Figura 1: Multilateralidade usando medições TDOA e TOA com hipérbole e círculos, respectivamente, como possíveis emissores locais [6].

Para a utilização dessas propriedades tanto em um cenário ativo, TOA, quanto em passivo, TDOA, por parte do objeto a ser localizado, é necessário um controle síncrono de emissores ou receptores. Entretanto uma flexibilidade adquirida ao se explorar essas propriedades é que o emissor pode ser móvel e a localização calculada por receptores fixos, ou os emissores podem ser fixos e o ativo ser capaz de determinar sua própria localização.

#### *AOA, Angle of Arrival*

Essa propriedade é explorada quando uma estação receptora possui uma disposição de antenas direcionais e utilizando a propriedade TOA de um sinal recebido é possível determinar o ângulo de incidência do mesmo. Quando aplicado a múltiplos sinais torna-se possível estimar a posição a partir dos ângulos estabelecidos com os pontos fixos que podem ser emissores ou receptores. Utilizando dois emissores é possível apenas determinar a posição em duas dimensões e ao acrescentar emissores torna-se possível a posição tri dimensional com maior precisão.[7]

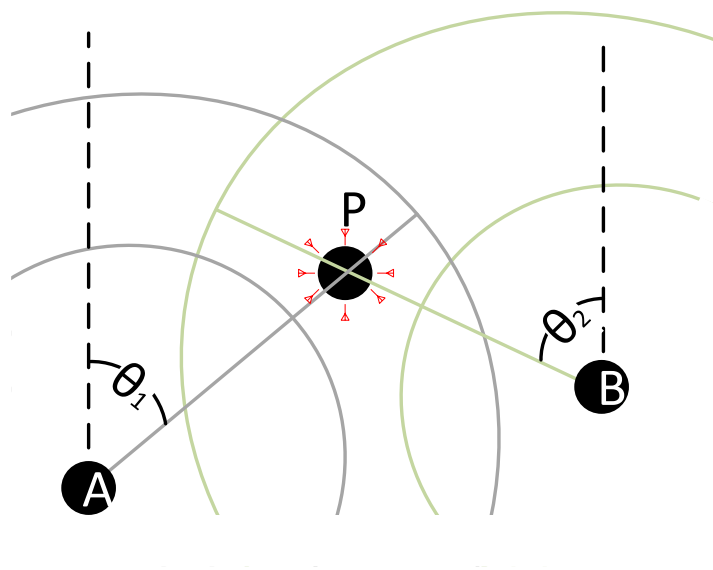


Figura 2: Posicionamento via ângulo de chegada

#### *FOA, Frquency of Arrival*

A análise de frequência de chegada pode ser utilizada para realizar rastreamento de objetos em movimento utilizando o efeito Doppler. Esse efeito consiste na mudança de frequência de um sinal emitido devido ao movimento relativo entre emissor e receptor,

causando um aumento de frequência ao se aproximar (blueshift) e decrescimento ao se afastar (redshift).

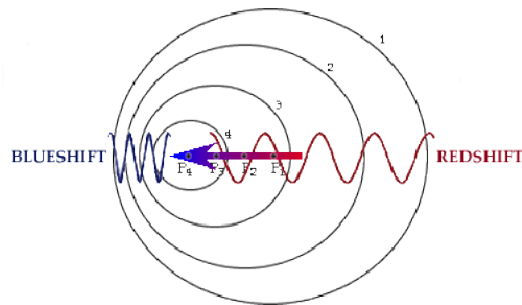


Figura 3: efeito Doppler (Jose Wudka 1998)

O efeito possui aplicações para geolocalização que foram exploradas por Engel para determinar fontes de ruído em sinais de satélites, geolocalização em escala global. Concluiu-se apesar da ineficácia, pode ser utilizado como complemento à TOA afim de aprimorar a localização [8].

### *RSSI, Received Signal Strength Indication*

Devido a imprevisibilidade de percurso de sinal em ambientes fechados, bem como a variabilidade da atenuação do mesmo em diferentes meios, a utilização da potência de sinal não corresponde a uma relação direta com a distância entre emissor e receptor. O RSSI é um indicador de qualidade de sinal que segundo o padrão estabelecido pelo IEE 802.11 pode variar em uma escala de 0 a 255 que varia de acordo com fabricante. O indicador RSSI pode ser encontrado em diversos protocolos de comunicação, tais como WIFI, BLE, LoRaWAN, ZigBee, Bluetooth... Apesar de não possuir relação direta com distância quando utilizado com diversas fontes simultâneas de sinais, é possível a obtenção de um vetor de RSSIs que possui característica relativa a posição em que ele foi obtido. Essa propriedade é comumente explorada em algoritmos de posicionamento por correlação, denominados Fingerprinting, o qual será explicado no escopo desse documento [9].

### **Algoritmos/Metodologias**

#### *Triangulação*

A triangulação utiliza múltiplas vezes a propriedade AOA a fim de determinar por meio da lei dos cossenos a posição de um objeto móvel.

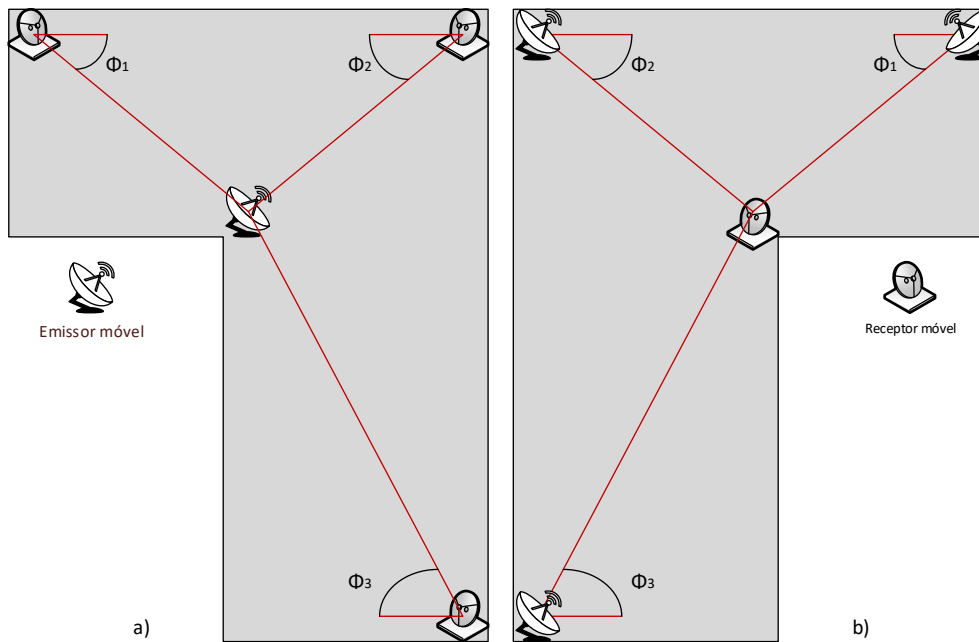


Figura 4: possíveis casos de uso de triangulação

Conforme Figura 4, independentemente da componente móvel ser emissora ou receptora, após a aferição dos ângulos, é necessário que seja realizado o cálculo de triangulação, para o qual é preciso conhecimento prévio da localização dos componentes fixos. Nota-se que na configuração a) o emissor móvel tem seu sinal detectado por receptores fixos que centralizam a informação de ângulo de chegada afim de determinar a posição do sensor, nesse caso para que seja possível a navegabilidade é necessária a transmissão da posição do servidor ao dispositivo móvel. Essa abordagem tem como vantagem a centralização de informações de rastreamento, simplificando gerenciamento de múltiplos ativos em ambientes como estoques e centros de distribuição. Na configuração b) o receptor móvel utiliza os ângulos dos sinais que intercepta e é capaz de determinar sua própria localização, simplificando a necessidade computacional centralizada o que vem a diminuir a complexidade da solução [10].

### *Trilateração*

A Trilateração é o processo que permite obter a posição de um objeto a partir da observância dos tempos de recepção de emissores diferentes. Utilizando TOA ou TDOA determina-se a distância de cada emissor e sabendo-se a localização dos mesmos é possível detectar a posição do ativo móvel [10].



Figura 5: Linearização temporal de múltiplos TOAs

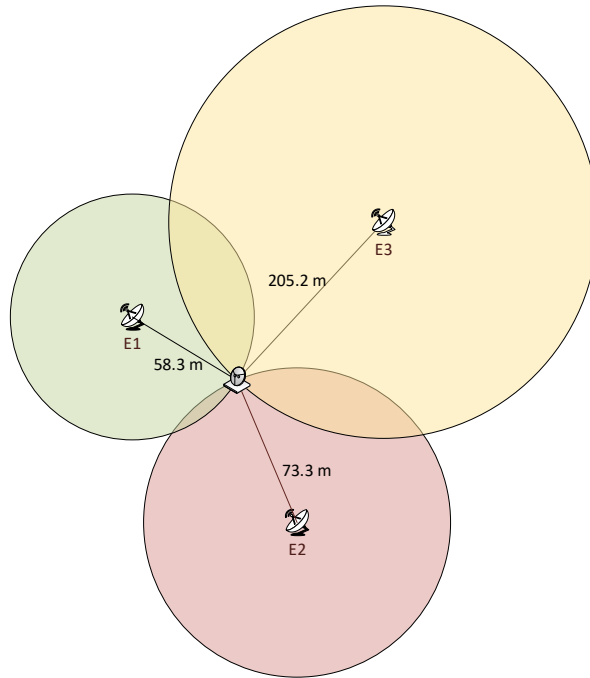


Figura 6: Localização via Trilateração

A Fonte 5, ilustra que após um  $t_0$  de sincronismo entre receptor e emissor a observância de recepção dos sinais em  $t_1$ ,  $t_2$ ,  $t_3$  como TOA dos respectivos sinais dos emissores 1, 2 e 3 e viabiliza a localização conforme Figura 6. Para realizar a trilateração pode-se utilizar sinais RF ou sinais acústicos [4].

### *Proximidade*

Existem sistemas de posicionamento que se baseiam na detecção da proximidade, o componente rastreado normalmente utilizam etiquetas RFID, NFC ou sistemas visuais como QR code. Nessa abordagem a posição obtida via  $N$  pontos fixos de leitura que são responsáveis por observar a movimentação de um ativo. É possível a implementação inversa onde o ativo é responsável por identificar sua movimentação por pontos definidos. nesses casos a precisão do sistema está definida pela disposição dos pontos de leitura e pelo alcance da tecnologia. É válido ressaltar que no caso clássico utilizando por exemplo

RFID passivos como dispositivos móveis, o serviço de localização é normalmente centralizado, o que facilita o gerenciamento dos rastreados [4].

### *Fingerprinting ou correlação*

A técnica de fingerprinting vem da ideia de trabalhar com algum grau de aleatoriedade dos sinais que estão disponíveis no meio. Essa metodologia costuma ser dividida em duas fases, uma de treinamento e outra de rastreamento. Na fase de treinamento é realizada uma exaustiva coleta de dados em diversos pontos de forma sistemática objetivando coletar o maior número possível de perfis e seus pontos associados. Na fase de rastreamento um sensor realiza uma coleta simplificada de dados do ponto em que se encontra utilizando os mesmos meios que o treinamento em suas N coletas por ponto, e então utiliza-se o resultado do processamento dos dados prévios para determinar sua posição. A partir do mapeamento de diversos dados em diversos pontos é possível com uma coleta simples estimar a posição utilizando como parâmetro os dados previamente adquiridos. Para tanto existem algumas abordagens dessa metodologia, uma delas visa utilizar filtro de Kalman em uma série temporal de dados afim de melhorar o posicionamento medido, [6]. Existem trabalhos utilizando sinais magnéticos visando aplicabilidade em smartphones convencionais [11], entretanto, destaca-se a recorrência em que se encontra trabalhos utilizando sinais do espectro não licenciado RF de 2.4Ghz Bluetooth e Wifi além de empresas utilizando LPWANs sub 1Ghz [12].

## **2.2. Alinhamento com escopo do trabalho**

Tendo em vista as técnicas e metodologias apresentadas em 2.1, em conjunto com a viabilidade de implementação no âmbito do trabalho proposto, torna-se necessário escolher uma forma de implementação do trabalho proposto no tema que ele se encontra. Dada a necessidade de infraestrutura que viriam a limitar a capacidade de implementação no tempo útil do trabalho, em conjunto com a maior capacidade de disposição e geração de métricas sob os resultados estatísticos na metodologia de fingerprint, tomou-se a decisão de aplicar a proposta do trabalho nessa metodologia. É importante ressaltar que o trabalho desenvolvido não se limita à mesma, uma vez que possui compatibilidade com

outras abordagens que venha a utilizar sua estrutura de forma a viabilizar análise de métricas diversas acerca da solução.

### **2.3. Implementações em localização Indoor**

Uma vez obtido o foco da implementação do projeto, aqui serão dispostos os casos de destaque no âmbito prioritariamente comercial que utilizem a abordagem estatística e posteriormente será discutido como a proposta desse trabalho se porta diante dos casos apresentados.

#### **CISCO CMX Hyperlocation**

A solução comercial disponível pela multinacional CISCO promete entregar uma precisão de um a três metros utilizando a infraestrutura de rede 2.4Ghz WIFI combinada com Bluetooth Low Energy (BLE). Os requisitos para utilização do hyperlocation são:

- gateway Cisco Aironet® 4800 Access Point
- Cisco Connected Mobile Experiences (CMX) Location Engine

A controladora de rede integrada ao sistema CMX analisa os RSSIs e AOA para cada endereço de rede, MAC adress, dos clientes wifi a fim de determinar via triangulação, nota-se que o sistema opera em máxima precisão quando ambos WIFI e BLE são disponibilizados pelos rastreados. O CMX permite gerar estatísticas, métricas e alarmes dos clientes. A solução é limitada ao rastreamento de 4000 dispositivos móveis simultaneamente e conta com vasta estrutura de rede WIFI além de acesso a controladora de Rede.

Dentre os casos de utilização publicamente documentados, destaca-se a utilização da ferramenta foi incorporada ao Museu de História Natural de Fernbank, ao qual via um aplicativo para o consumidor final, trouxe melhorias na experiência do visitante. O museu disponibilizou informações sobre exposições e obras de forma interativa e o aplicativo se encontra disponível.

Outro caso destacado pela CISCO foi um Shopping que implementou a solução hyperlocation integrada ao login dos clientes via Facebook , a qual estima que pelo menos

21% dos consumidores teriam configurações necessárias para serem rastreados antes de efetivamente entrarem nas instalações.

A empresa possui também declarações de casos de uso no aeroporto de Copenhague, Stanley Black & Decker, Parque Temático PortAventura e outros os quais não vieram a tornar pública detalhes. Nota-se que em um caso de estudo a empresa alega ter integrado seu sistema à uma rede de mapas de localização indoor de domínio público sendo iniciativa da mazemap [13].

### **Here Indoor position**

Here é uma empresa Norte Americana com foco em desenvolvimento de produtos que forneçam localização e preferencialmente viabilizem navegabilidade. A empresa possui uma solução em localização indoor que pode ser embarcada em um dispositivo com sistema operacional Android e capacidade de coletar dados Wifi e Bluetooth. Ela também oferece o serviço via API web com compatibilidade com sinais 2G, 3G e 4G. Devido a variabilidade de cenários em que a solução se aplica e ao fato de repassarem para o cliente a responsabilidade de compra e instalação de beacons bem como utilização de uma aplicação fornecida para montagem do mapa de correlação, faz sentido que a empresa não se comprometa com uma promessa de precisão tentadora nem apresente um valor que seja muito acima da capacidade atingível do seu produto. Entretanto a solução da Here tem como diferencial a auto suficiência do cliente quanto a implementação o que diminui os custos de escalabilidade e permite maior foco em avanços de software. Um contraponto da solução é a necessidade de se prender como cliente a solução de mapas indoor para consumir a solução de localização porém a empresa fornece atualmente um período experimental gratuito de seu ferramental [14].

### **Localização em tempo real por uso de Filtros de Kalman estendidos fundindo dados WiFi com sensores Inerciais em smartphones**

O caso aqui apresentado se trata de um trabalho de pesquisa no âmbito acadêmico realizado pelo Instituto de Pesquisa em Tecnologias Eletrônicas de Nanjing, em parceria com a Escola de informação e ciência e tecnologia Dalian Maritime University. Nesse trabalho foi desenvolvida uma pesquisa que almejava a fusão de duas distribuições de erro provenientes de dados WIFI e dados inerciais, a técnica utilizada para tanto, consistiu na implementação de filtros de Kalman estendidos, disposto em uma malha fechada de

controle. Ela apresenta resultados demonstrando reduzir os ruídos em soluções fingerprint e evitando o erro acumulado inerente a uso isolado de sensores inerciais. O estudo combina Pedestrian Dead Reckoning, Wifi com filtro de partículas e Filtros Kalman. Esse estudo apresenta resultados utilizando como métricas a média e o desvio padrão da distribuição dos erros nos cenários em que combina ou não os dados e aplica em um filtro de partículas ou Kalman e se trata de uma solução de rastreabilidade contínua onde se assume que o ponto inicial do objeto rastreado é conhecido [15].

## **2.4. Sumário**

A partir do conteúdo disposto nesse capítulo, destaca-se que o foco do trabalho a ser desenvolvido será baseado em metodologia de correlação ou fingerprinting. Isso se deve à maior facilidade de implementação no âmbito do trabalho de mestrado, reduzida demanda de equipamentos, baixo custo de implementação e alta capacidade demonstrativa de métricas acerca de uma distribuição de erro em soluções indoor.

A solução proposta no âmbito desse trabalho se porta de forma a validar em casos comerciais as métricas prometidas pelas empresas como CISCO, bem como analisar a variabilidade da precisão em diferentes cenários. A ferramenta proposta permitirá resguardar soluções como as da empresa HERE, sem deixar de fornecer ao utilizador uma métrica esperada, ou seja, apresentar algum comportamento esperado sem o compromisso diante do cenário de alta variabilidade. Já em estudos acadêmicos como o apresentado pela Universidade marítima de Dalie o trabalho pretende complementar as métricas apresentadas com a automação de estatísticas como intervalo interquartilico, média harmônica entre outras métricas a serem apresentadas a seguir. Ressalta-se trabalho proposto não almeja substituir as soluções citadas e sim obter uma ferramenta capaz de automatizar análise de soluções indoor em diversos âmbitos.

## **Capítulo 3      Tecnologias adotadas**

Nesse capítulo é feita uma breve introdução ao contexto das ferramentas utilizadas na implementação da solução proposta com ocasionais comparações a alternativas disponíveis. O texto aqui apresentado também objetiva dispensar maiores explicações acerca das ferramentas no Capítulo 4, que passa a ter foco majoritário no objetivo almejado pela proposta na apresentação do conteúdo desenvolvido. A leitura deste capítulo passa a complementar ao contexto de engenharia industrial quanto a ferramentas de engenharia de software e IoT servindo como base para o desenho da arquitetura apresentada no capítulo a seguir. Devido a infinidade de linguagens e protocolos existentes no contexto de engenharia de software não é cabível no escopo desse trabalho a realização de um estudo comparativo de todas as tecnologias existentes, sendo o objetivo desse capítulo a disposição do conteúdo utilizado, a justificação da escolha do mesmo e eventuais comparativos com alternativas de destaque.

### **3.1. Contextualização com a arquitetura implementada**

Visando relacionar as ferramentas aqui dispostas com a solução implementada, buscando evitar uma listagem de ferramentas sem contextualização é apresentado aqui uma breve descrição da apresentada na Figura 7.

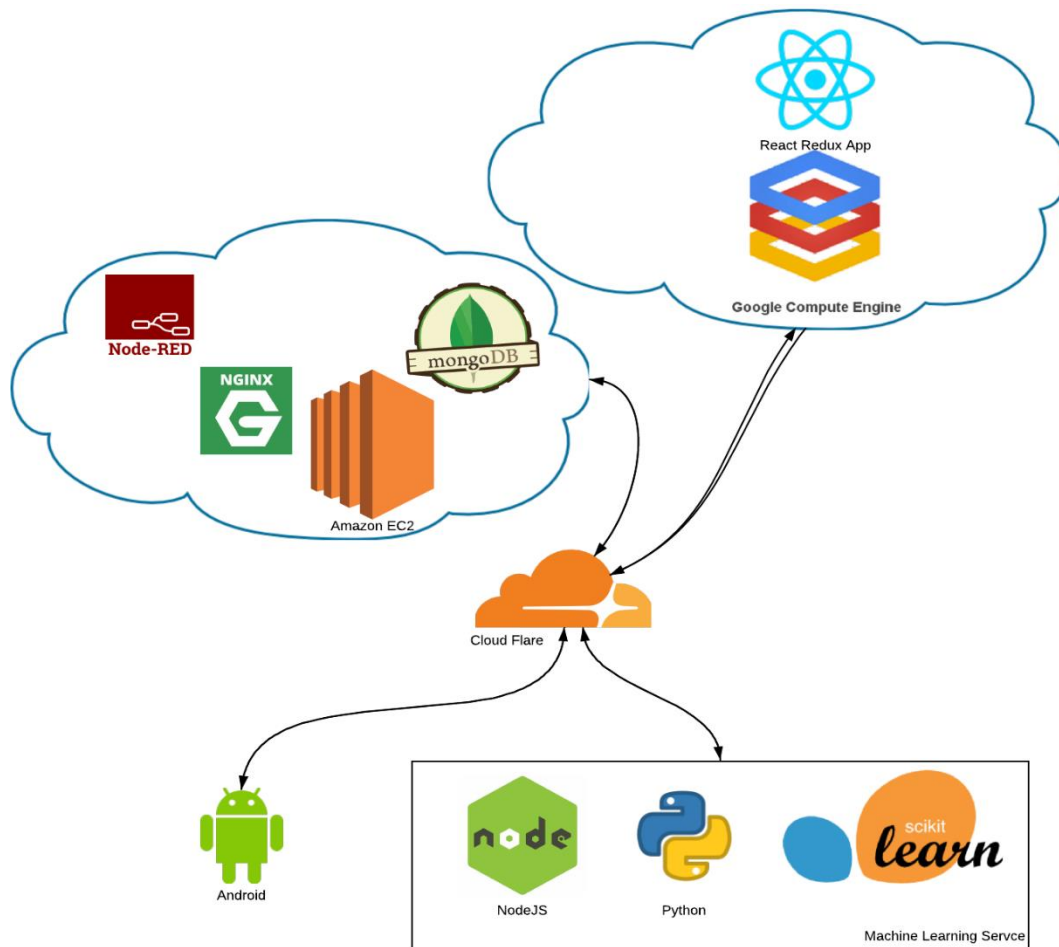


Figura 7: Arquitetura das ferramentas de computação

A arquitetura apresentada é composta por 5 elementos principais:

- Aplicação Android para levantamento dos dados
- Ferramenta administrativa, aplicação Web
- Provedor de localização baseado em aprendizagem de máquina
- API RESTful
- Banco de dados
- Servidor de aplicação

## 3.2. Ferramentas utilizadas

### Banco de Dados

O banco dado escolhido foi o MongoDB, o qual se trata de um banco de dados baseados em documentos [16]. Sua escolha se baseou em sua capacidade de escalabilidade

horizontal que é mais compatível com um eventual crescimento futuro quanto a número de clientes e de elevada compatibilidade com tecnologias web baseadas em Javascript ou typescript, em crescente demanda em ambientes comerciais [17].

## **API RESTfull**

A API foi implementada majoritariamente de forma a performar no banco de dados operações CRUD (Create, Read, Update, Delete) que, em português, representa a capacidade de Criação, Leitura, Atualização e Exclusão dos dados armazenados no banco de dados. A API é classificada pelo acrônimo de Representational State Transfer (REST), ou Transferência de Estado Representacional, e atua via conexões HTTP. Essa implementação é realizada pela framework de programação visual node-red, a qual permite uma programação single thread do nodeJS de facilitada interpretação. A API é hospedada em um servidor do tipo T1 micro da AWS e é acessado via redirecionamento pelo serviço NGINX [16] que é uma ferramenta de controle de carga e redirecionamentos e faz um proxy implementando conexões criptografadas via TLS. O serviço HTTPS foi viabilizado pela organização open-source Let's Encrypt e verificada e disponibilizada por sua distribuição via script python certbot [17].

## **Ferramenta administrativa Aplicação Web**

A Ferramenta administrativa que é responsável por analisar e gerenciar os dados via uma aplicação web. A aplicação web foi hospedada em uma conta gratuita da plataforma de computação em nuvem da Google. A aplicação web foi implementada em ReactJS [18] que é uma framework open-source typescript, linguagem baseada em JavaScript ES6. Essa framework foi desenvolvida pela empresa norte americana Facebook e concorre entre outras com Angular [19] e VueJS [20]. Se trata de um servidor Frontend, ou seja, responsável por compilar e entregar ao cliente o conteúdo da aplicação web, o qual acessa o mesmo por um navegador [21], [22].

No design da ferramenta foi feito uso de bibliotecas Bootstrap e Google material UI que permitem implementar de forma fácil uma interface compreensível com o usuário, mitigando o esforço voltado para parte gráfica da ferramenta.

Para a disposição das plantas que podem ser enviadas para o servidor como imagens foi utilizada a biblioteca leaflet.js que permite interação com mapas não geográficos. Para a

apresentação dos resultados foi utilizado a biblioteca de plot da empresa Google sob licença open-source (React Google Charts) [23], também foi utilizada a biblioteca de gráficos E-charts que se encontra sob licença Apache 2.0 e é open-source.

Afim de buscar maior legibilidade de código, agilidade de implementação e expansibilidade da ferramenta de análise foi implementada a metodologia Redux que é responsável por unificar a fonte de dados da aplicação web e sistematizar a atualização de dados por uma estrutura orientada a eventos que tratam de atualizar somente a parte da interface do usuário que obteve alterações nos dados, dessa forma a aplicação fica mais fluida e portanto possui maior potencial para expansão [24].

Destaca-se que a abordagem de utilização desta ferramenta foi baseada num contrapeso entre facilidade de implementação e capacidade de utilização em ambientes de pesquisa e desenvolvimento de produtos por empresas que venham a se interessar no projeto. É do entendimento do autor que aplicações em ferramentas popularmente acadêmicas como Octave ou Matlab poderiam a vir a demonstrar com a mesma qualidade ou de forma superior os resultados obtidos, porém viriam a se distanciar de um produto comercializável.

### **Aplicação Android**

Optou-se pelo desenvolvimento da aplicação em Android via Android Studio [25]. Essa escolha se baseou na capacidade de integração da linguagem bem como extensa documentação proporcionada pela gigante tecnológica que é proprietária do sistema operacional, Google. A escolha baseou-se também na existência de um maior número de exemplos e discussões na comunidade open-source quando comparado com soluções mais recentes, além de maior performance em implementações nativas. Apesar de interessante o desenvolvimento para múltiplos sistemas operacionais móveis que é possibilitado via frameworks como Ionic, Apache Cordova, ou até mesmo react-native, a adoção de uma implementação nativa se resguardou na total compatibilidade com o kit de desenvolvimento de software (SDK) que viabiliza acesso aos dados dos sensores, e melhor performance.

## **Serviço de aprendizagem de máquina**

A implementação dos algoritmos de localização indoor disponibilizados pela ferramenta visou manter agilidade produtiva e concordância de código com as outras ferramentas do sistema. A opção de um programa em nodeJS [26] foi pensada em viabilização de um servidor single thread o que impede a concorrência entre threads por recursos na máquina em que ele executa. A escolha também se baseou na facilidade de se compatibilizar numa mudança de arquitetura, um programa em javascript com as ferramentas da API, do banco de dados e da plataforma web, isso viabiliza a união de todos os elementos em um único release binário. Devido ao aumento do número de bibliotecas estatísticas de aprendizagem de máquina na linguagem Python esse componente da arquitetura ficou responsável por gerenciar processos filhos que consomem bibliotecas em Python, o que trouxe compatibilidade com o mundo da ciência de dados e evitou limitações futuras nesse sentido. Ressalta-se, entretanto, que grande maioria das bibliotecas de machine learning existentes em python hoje são apenas front-end para implementações em c e c++ o que evita qualquer comprometimento de performance pelo interpretador do python. A biblioteca escolhida foi a Scikit-Learn [27] a qual é executada em um processo filho e tem a troca de dados entre os processos Python e nodeJS é feita via stdin, stdout e um barramento MQTT [28] com tópico dedicado ao processo.

## **Cloudflare**

A fim de realizar a implementação da arquitetura baseada em serviço web foi feito uso de um domínio que se encontrava disponível para o projeto pelagaggi.com. O mesmo serviu para facilitar o roteamento das interações web para a API RESTfull bem como serviço WEB React + Redux [24]. Esse domínio tem o DNS gerenciado pela CloudFlare [29] que é uma ferramenta gratuita de gerenciamento de DNS com proteção a ataques de DDOS e fácil gerenciamento de subdomínios. O subdomínio da API ficou sendo node-red.pelagaggi.com e o subdomínio da ferramenta administrativa web indoor-manager.pelagaggi.com.

### **3.3. Sumário**

O trabalho foi implementado em uma arquitetura priorizando elementos de computação em nuvem de alta escalabilidade, com linguagens de alta demanda em âmbito profissional e acadêmico, priorizando aplicabilidade no mercado profissional e em eventuais pesquisas e desenvolvimento de produtos. Apesar da arquitetura impor uma certa complexidade quanto a diversidade de diferentes ferramentas que ela utiliza, a mesma prioriza produtividade e integrabilidade em cenários reais. Demonstra-se por meio desse que a universidade tem grande capacidade de interagir ativamente com demandas da indústria de software e entregar pesquisas prontas para inserção em contexto empresarial.

## Capítulo 4 Trabalho efetuado

Nesse capítulo será apresentada a arquitetura do projeto bem como a motivação da mesma em função dos objetivos do trabalho e de eventuais trabalhos futuros. Será feita uma descrição das funcionalidades implementadas nas diferentes partes do projeto, tendo foco na interconexão da solução como um todo. O conteúdo será orientado no auxílio que a composição da arquitetura fornece para o enfrentamento da problemática, demonstrando a base tecnológica que a ferramenta utiliza a fim de prestar suporte à análise de soluções em localização indoor. Objetivando evitar maiores detalhamentos técnico acerca das tecnologias utilizadas para resolução da problemática no capítulo de implementação, o qual passa a possuir uma descrição de alto nível acerca do projeto quanto ao tema de localização indoor.

### 4.1. Escopo do Trabalho

Para a realização do trabalho se encontravam disponíveis os seguintes recursos.

- 1 Aparelho celular da fabricante Xiaomi modelo mi A2 com sistema operacional Android 8.1 com capacidade de levantamento de dados de redes Wifi 2.4 [Ghz] e Wifi 5 [Ghz], bem como sensor magnetômetro.
- 1 Banco de dados NOSQL MongoDB hospedado pelo serviço ATLAS com 512MB.
- 1 servidor na categoria “T2.micro” com 1vCPU e 1GiB de memória na plataforma de computação em nuvem Amazon Web Services (AWS)
- 1 servidor na categoria “n1-standard-1” com 1 vCPUj e 3,75 GB de memória em na plataforma de computação em nuvem da Google, Google Cloud Platform

- 1 domínio pelagaggi.com

## 4.2. Arquitetura do projeto

A arquitetura do projeto, apresentada na Figura 8, é composta por uma ferramenta administrativa responsável por configurar a implementação de um experimento, o experimento é o objeto de análise e estudo, sob o qual será coletado um espaço amostral de resultados da solução de localização afim de se obter métricas para a avaliar o desempenho da mesma. A componente administrativa também é pretendida a fim de prestar suporte na análise de resultados. A arquitetura possui ainda uma aplicação Android, é responsável por coletar os dados do ambiente. Também compõem a arquitetura um provedor de localização que consiste em um serviço que implementa algoritmos de aprendizagem de máquina responsáveis por receber os dados e gerar os resultados a serem analisados, uma API que coordena o fluxo de dados e um banco de dados estruturado para a solução proposta.

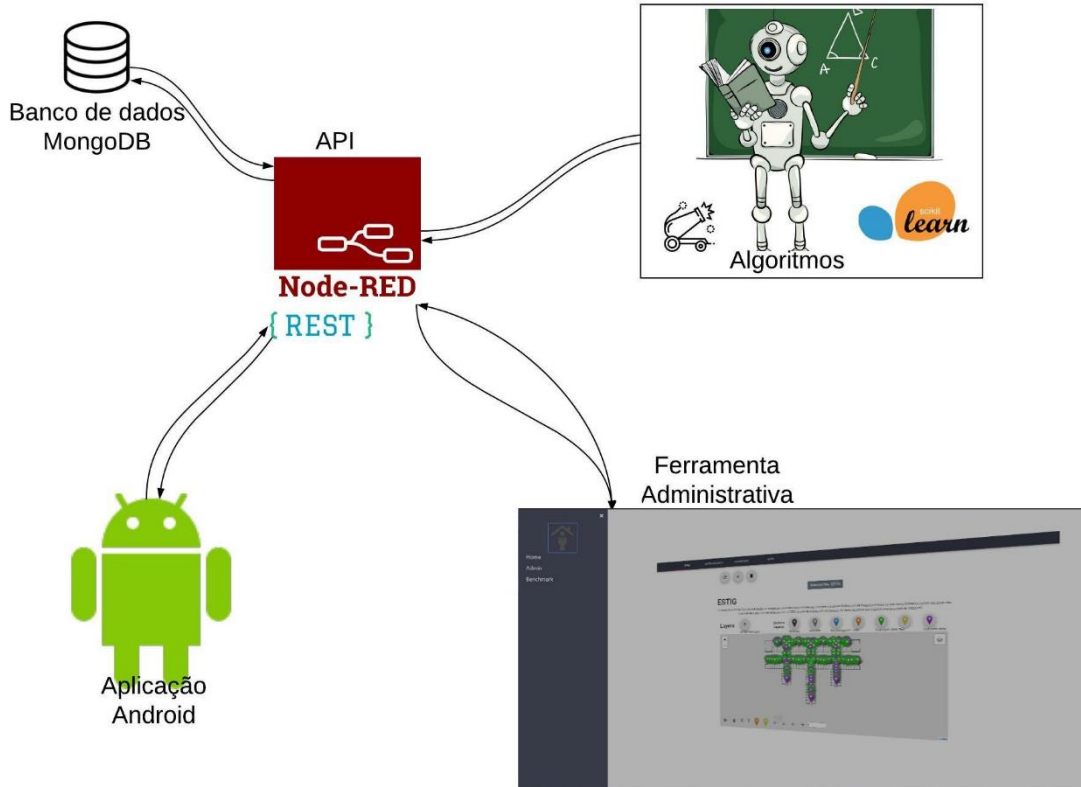


Figura 8: Arquitetura do projeto

Visando a viabilização desse ecossistema foi implementado como componente central uma API web responsável pela recepção de solicitações de escrita, leitura, atualização e exclusão dos dados armazenados em um banco de dados. Além de operações simples a API também interpreta o estado do banco de dados de forma a entregar dados trabalhados, como, por exemplo, solicitação dos dados a serem utilizados pelo serviço responsável por prover a localização.

O banco de dados escolhido permitiu maior facilidade de interação com a API devido a compatibilidade entre NodeJS e MongoDB. É importante ressaltar que o banco de dados possui concordância com outras ferramentas do projeto e como ele foi aplicado e permitiu a implementação dos requisitos no tempo necessário, sem comprometer a expansibilidade da arquitetura.

### **4.3. Funcionalidades dos elementos**

Esta seção aborda cada um dos elementos da arquitetura implementada afim de permitir perceber o fluxo de dados na ferramenta como um todo.

#### **Ferramenta Administrativa**

A ferramenta administrativa foi implementada para viabilizar a criação de sítios de estudos, upload e configuração de plantas baixas (layers), criação e configuração de pontos de interesse. A página de sítios (Figura 9), permite a criação de um sítio de estudo que inclui a disposição de um texto para descrição de seu propósito, além da viabilização da exclusão dos sítios caso venha a ser necessário. O programa define cada planta baixa como uma camada ou um layer e ele permite o upload de múltiplas plantas arquitetônicas, em formatos de imagem vetorizadas, às quais os dados de localização serão atrelados. Essa página permite a adição e configuração de pontos de interesse (POIs) além de identificar quais os pontos de interesse possuem dados ou não. Foram implementadas funcionalidades facilitadoras do processo de configuração, as quais visam melhor atender ambientes de larga escala, essas funcionalidades são:

- Cópia da coordenada X ou Y de um ponto para o outro
- Réplica de um ponto em offset definido para uma direção ortogonal
- Configuração de treino ou teste facilitada com um clique

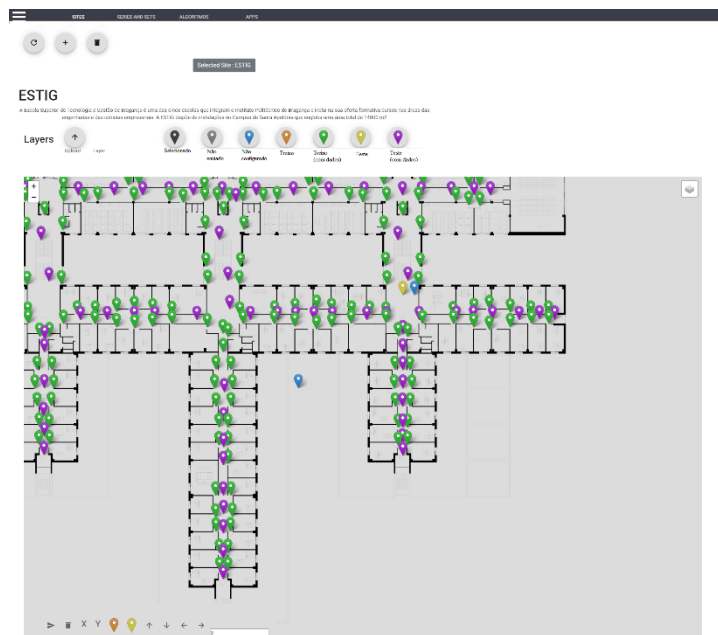


Figura 9: Gerenciamento de Sítios de estudo

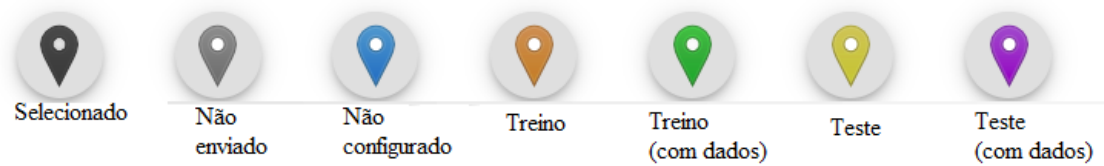


Figura 10: Legenda dos marcadores e respectivas cores na página de configuração

A edição detalhada de cada componente é feita por uma caixa de diálogo, estas são ilustradas pela Figura 11.

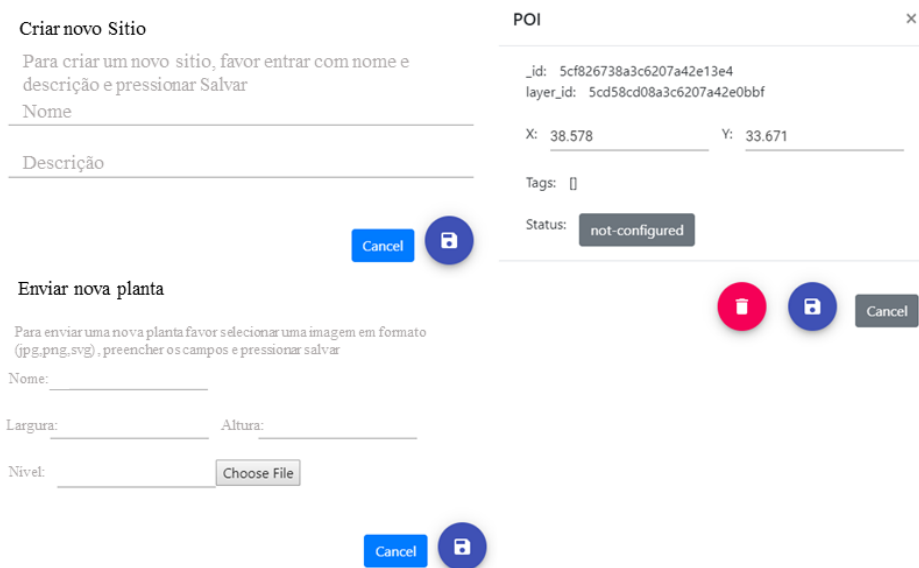


Figura 11: Caixas de diálogo para edição

Destaca-se que as plantas podem ser enviadas como imagem, mas a fim de se evitar distorção ou problemas de resolução recomenda-se o envio de uma fonte vetorizada, como por exemplo formato Scalable Vector Graphics (SVG).

### Serialização dos dados

Uma das principais funcionalidades pretendidas pela ferramenta é a facilitação do uso e reuso de dados em múltiplas combinações. A serialização dos dados é o processo no qual os dados são agregados em diferentes série e posteriormente em um conjunto de series (Figura 12).

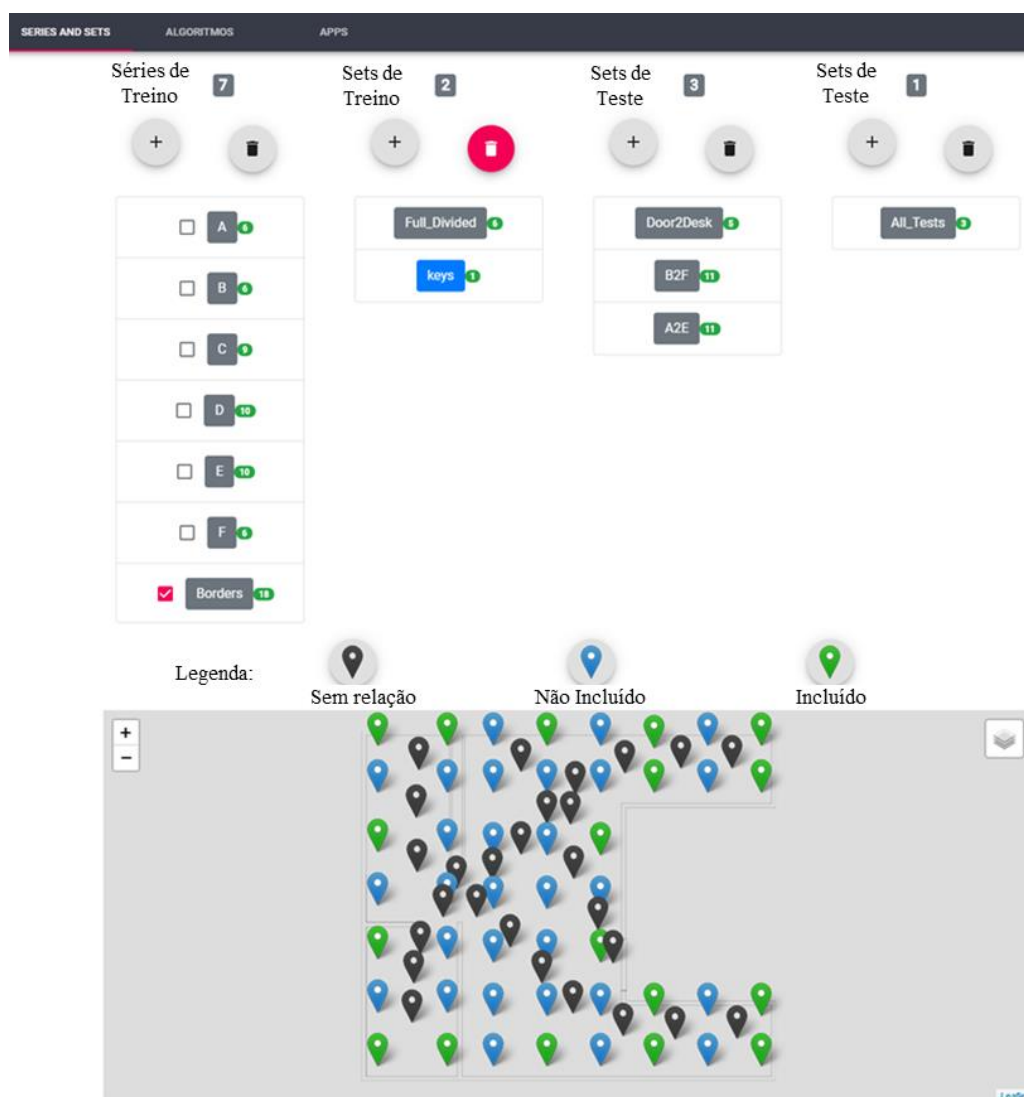


Figura 12: Página de serialização de dados.

A imposição dessa etapa viabiliza a utilização de um conjunto de dados de múltiplas formas em múltiplas ferramentas.

### Registro de algoritmos

Os algoritmos, são formas de controlar qual configuração vai ser utilizada quanto ao provedor de localização. Os algoritmos permitem, em caso de serviços com múltiplas opções de algoritmos, a seleção de qual deve ser utilizada pelo teste a ser realizado. O processo de edição dos Algoritmos é ilustrado na Figura 13.

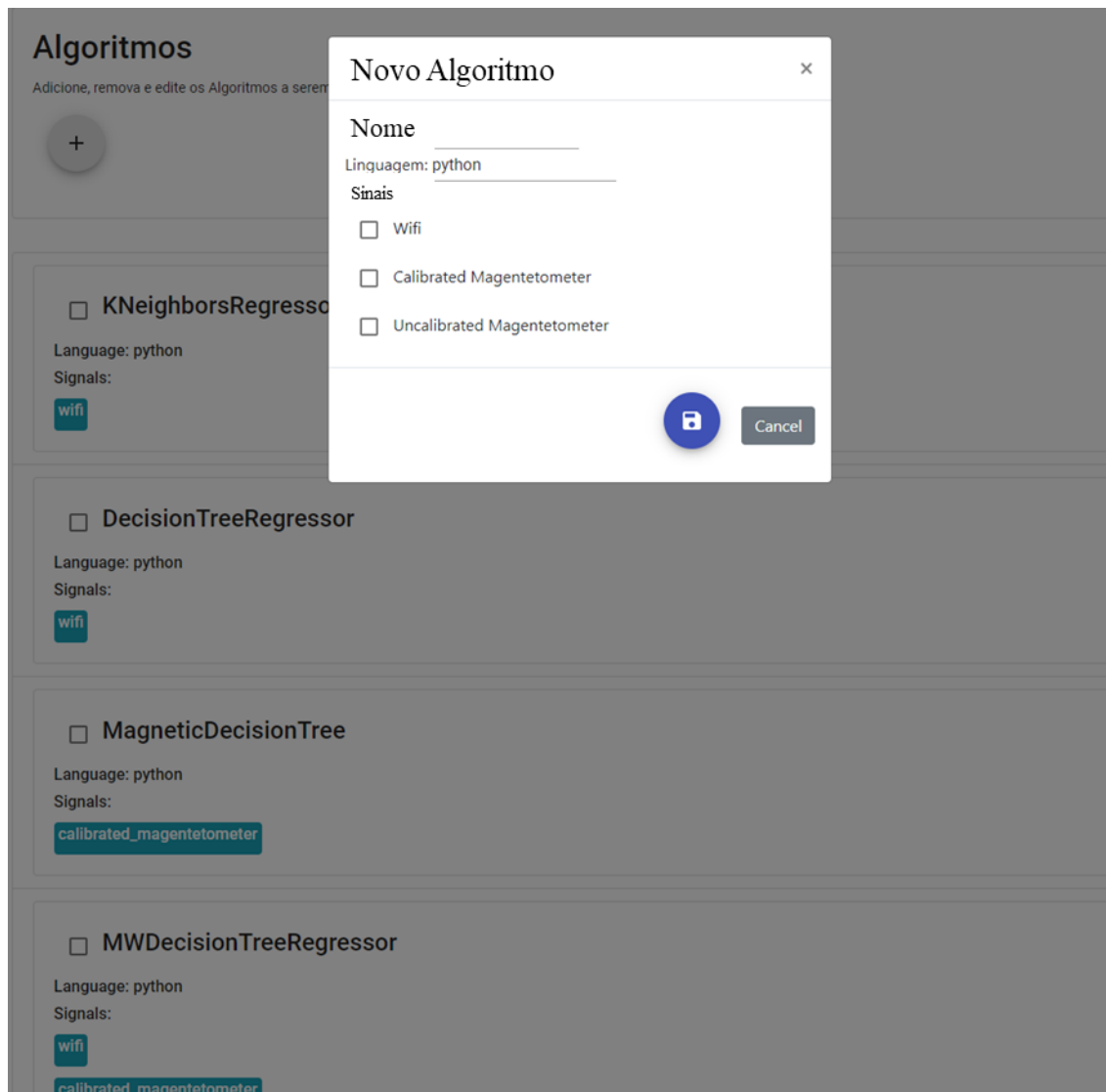


Figura 13: Processo de edição de Algoritmos

Esta etapa viabiliza ao utilizador especificar qual a forma que o servidor de localização deve utilizar para gerar resultados e apesar de não ser de grande mais valia em provedores de único algoritmo facilita a integração a ferramentas de aprendizagem supervisionada automatizada como Google Auto ML ou AWS Sagemaker. Para o demonstrar as funcionalidades desse projeto foi desenvolvido um serviço provedor de localização próprio que é abordado nesse capítulo.

### Aplicações e ordens de execução

As aplicações são as fontes finais de experimentos ou testes e são compostas por um Nome, um conjunto de dados de teste, um conjunto de dados de treino e um algoritmo, ou servidor de localização. Após a criação de uma aplicação ela possui como funcionalidade o envio de uma ordem de execução. Cada ordem de execução será disponibilizada ou enviada ao provedor de localização, que gerará os resultados para análise posterior. Essa divisão foi feita afim de viabilizar que uma mesma configuração da ferramenta de geração de métricas seja usada para fins comparativos entre múltiplas possíveis configurações do provedor de localização, permitindo análise de resultados com diferentes parâmetros (Figura 14).

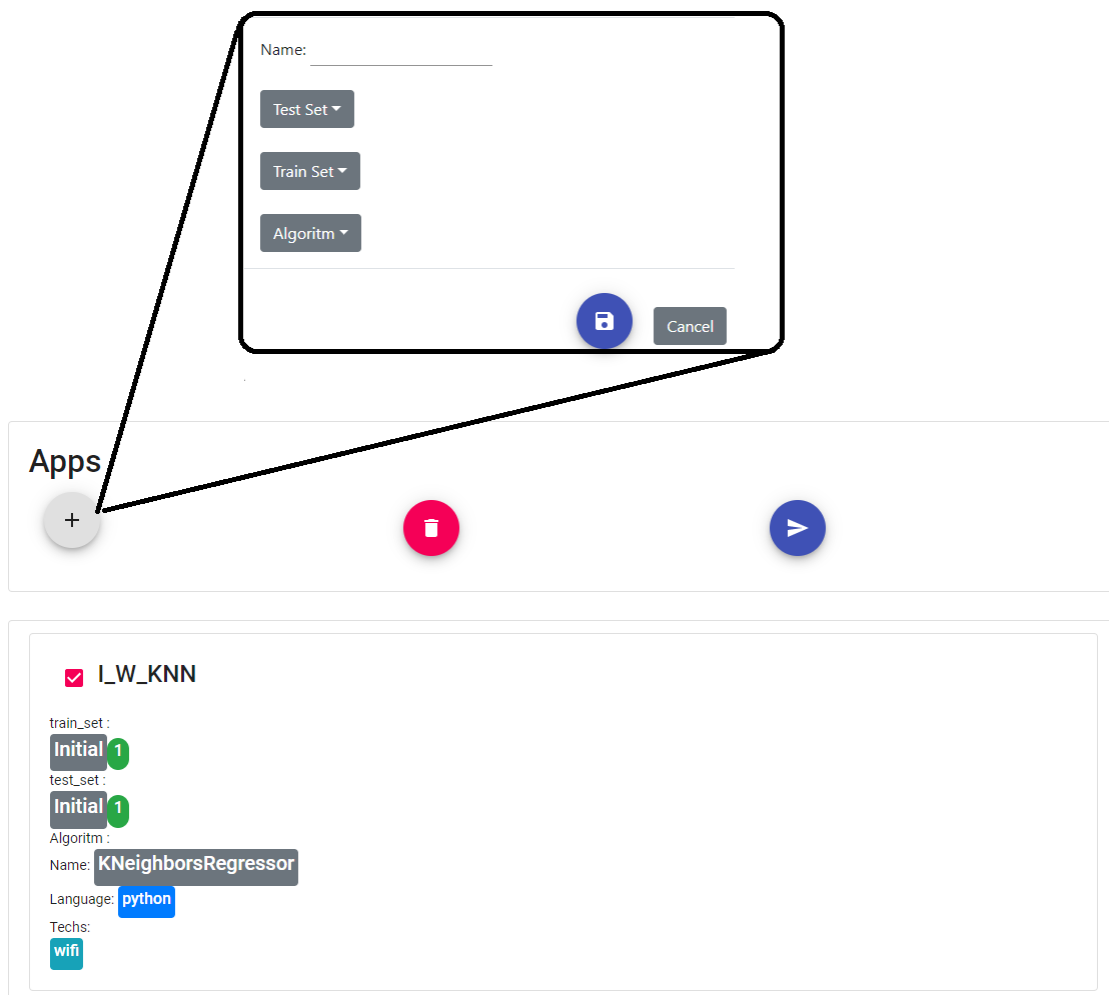


Figura 14: Criação e utilização de aplicação de teste

Quando o provedor possuir uma variedade muito grande de parâmetros, passa a ser inviável a utilização da interface gráfica para esse fim e, portanto, a utilização da API permite a automatização desse processo.

## Resultados

Após a atualização da ordem de execução ser realizada pelo provedor de localização, contendo então os resultados com formato especificado pela API, os mesmos poderão ser analisados quanto as métricas estabelecidas. Destaca-se que a análise dos resultados será feita posteriormente no Capítulo 5, e que é pretendido nesse tópico demonstrar as funcionalidades implementadas na seção de resultados. A utilização da página de resultados parte da seleção de uma ordem de treinamento a qual é disposta em uma tabela contendo todas os resultados e é pode ser selecionada de acordo com a métrica desejada, os resultados podem ser ordenados de acordo com a métrica de forma crescente ou decrescente (Figura 15).

Selecione a Configuração :

Selected Site : ESTIG  
Selected Layer : Térreo

Resultados									
App Name	Evaluation method	Metric Value	Mean Error	Error Range	Variance	Standard deviation	Median		
<input type="checkbox"/>	L_W_KNN	Statistics_Sum	6.084	0.938	1.232	0.179	0.423	1.104	0
<input type="checkbox"/>	L_W_KNN	Harmonic_Mean	0.011	1.641	3.920	2.025	1.423	1.725	2
<input type="checkbox"/>	L_W_KNN	Geometric_mean	0.433	1.453	3.933	2.120	1.456	1.280	2
<input type="checkbox"/>	L_W_KNN	Mean	0.938	0.938	1.232	0.179	0.423	1.104	0
<input type="checkbox"/>	L_W_KNN	Variance	0.131	1.893	1.067	0.131	0.362	1.978	0

Rows per page: 5 1-5 of 45

Figura 15: Seleção de ordem de treinamento

Após a marcação da caixa de seleção do resultado desejado, é apresentado para o usuário um gráfico contendo a distribuição do erro, com eixo vertical o percentual de incidências de erro e eixo horizontal a faixa em metros relacionada a esse percentual (Figura 16).



Figura 16: Distribuição de erro percentual.

Também é disponibilizado para o usuário o espalhamento bidimensional sob a planta fornecida dos pontos de teste, seu gabarito, ou conjunto de resultados esperados, é representado em azul e seus resultados em vermelho representando seus erros (Figura 17).

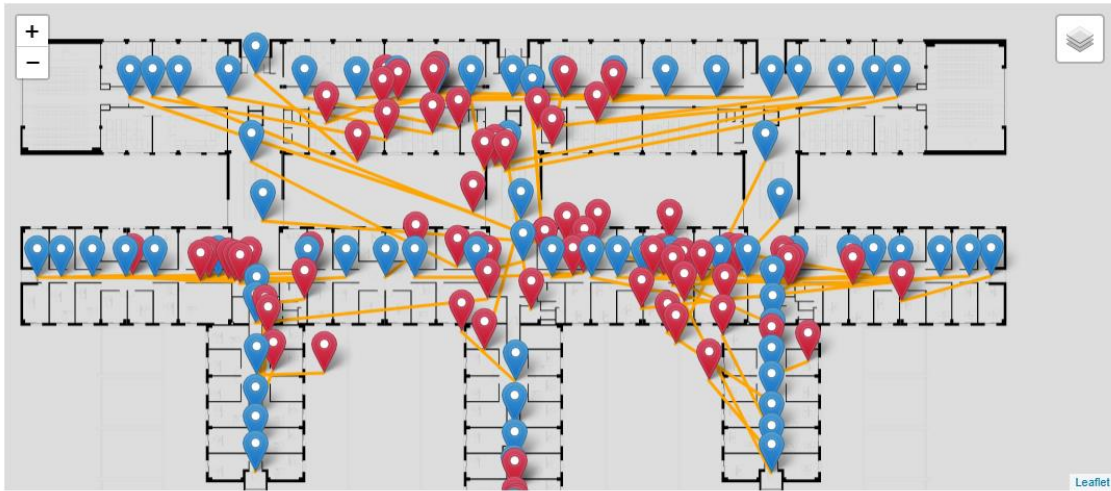


Figura 17: Espalhamento bidimensional sob a planta baixa

Utilizando o espalhamento incidente na planta, é viabilizado ao usuário a navegação dos resultados em um experimento em larga escala, permitindo a melhor visualização dos dados em contexto de localização indoor. Entretanto, visando evitar que excesso de informação venha a atrapalhar a análise do espalhamento, é fornecido um gráfico de espalhamento de erro similar porém sem informações da planta em questão, o qual conforme, Figura 18, mostra um conjunto de vetores partindo ponto em que o erro seria 0 e com um círculo no ponto em que o resultado foi apresentado como destino.

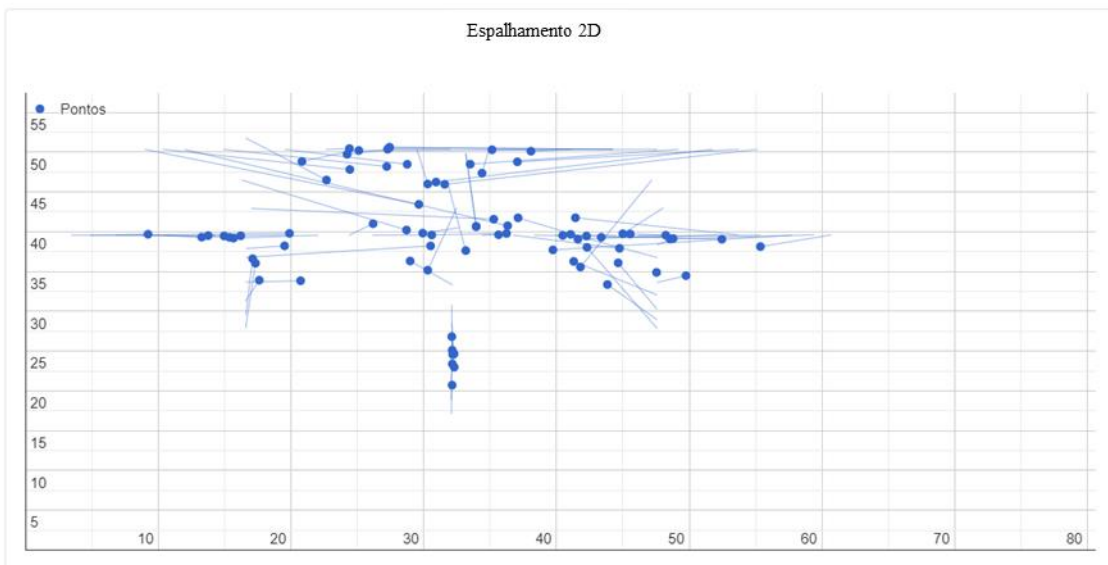


Figura 18: Espalhamento bidimensional simplificado

Também é disponibilizada uma forma de identificar regiões de maior erro por meio de um gráfico um gráfico interativo (Figura 19).

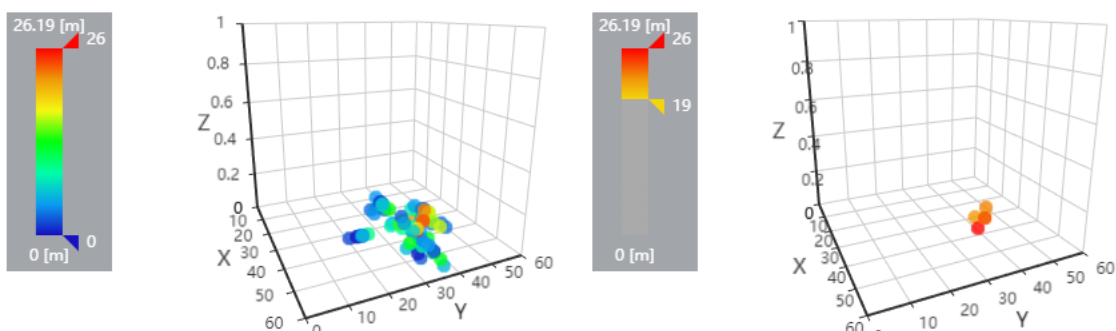


Figura 19: Gráfico de espalhamento 3d com filtragem por faixa de erro

É apresentado também ao usuário uma tabela com informações estatísticas acerca da distribuição de erros do resultado em questão, além do tempo de tratamento da ordem e parâmetros utilizados pela ordem de execução do teste, Figura 20.

Erro médio:	8.05 [m]
Intervalo:	0.208 [m] --- 26.2 [m]
Variância:	38.2 [m <sup>2</sup> ]
Desvio Padrão:	6.18 [m]
Mediana:	6.41 [m]
IQR:	9.07 [m]
Iniciado:	0h :48m :39 s
Finalizado:	2h :21m :10 s
Parâmetros:	{'p':2,'weights':'uniform','algorithm':'auto','n_neighbors':18,'defaultValue':-140}

Figura 20: Estatísticas acerca do resultado selecionado e sua distribuição de erro

### Aplicação Android

A aplicação desenvolvida tem como objetivo realizar a coleta de dados WIFI 2.4 Ghz e 5 Ghz, bem como dados do magnetômetro do aparelho celular. Ela foi pensada de forma a manter a mesma interface gráfica disponibilizada pela ferramenta administrativa e usou a mesma linguagem visual de marcadores de POIs, ou seja mostra quais são os pontos que já possuem dados ou não (Figura 10). Por meio de sua utilização a aplicação permite selecionar um dos sítios de estudo, selecionar a planta em que se deseja coletar os dados, selecionar o ponto de interesse com um duplo clique, afim de evitar toques acidentais,

visualizar o progresso do levantamento de dados e a posterior atualização de seu estado, esse processo é visualizado na Figura 21.

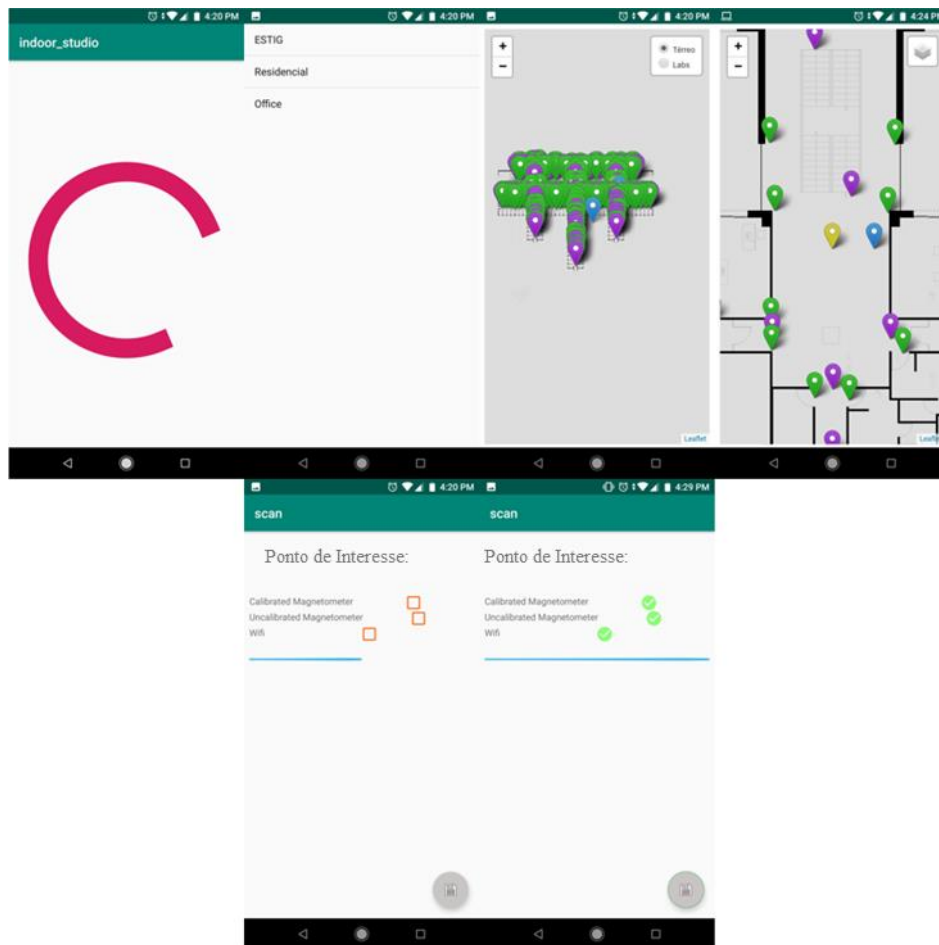


Figura 21: Utilização da aplicação Android para coleta de dados

Pensando em acidentes de utilização durante a operação de levantamento de dados in lócus, a ferramenta disponibiliza a funcionalidade de exclusão dos dados associados a um ponto de interesse em caso de detecção de má utilização. Para exclusão dos dados o usuário pode manter pressionado um dos POIs que possua dados, e após a exclusão dos mesmos a interface é atualizada para continuar o levantamento.

A aplicação foi pensada quanto a seu desenvolvimento na eventual integração com hardware proprietário e nas modificações que viriam a ser necessárias para coleta de dados externos ao SDK Android. Para coleta de outras fontes de dados basta alterar uma das páginas da aplicação, ou como descrita na arquitetura Android uma Activity.

## API e banco de dados

A API é responsável por coordenar o fluxo de dados, bem como garantir a integridade deles. A comunicação com o banco de dados é feita única e exclusivamente por intermédio da API, isso é imposto como medida de prevenção a erros em tempo de desenvolvimento, garantindo a integridade e segurança dos dados coletados, bem como medida de prevenção a ataques cibernéticos, tendo o banco uma reduzida lista de IPs autorizados a conexão. A implementação dos dados inseridos com documentos nas coleções do MongoDB foi feita segundo a metodologia CRUD conforme a framework utilizada ilustra na Figura 22.

As coleções implementadas são:

- Sítios
- Camadas
- POIs
- Series
- Algoritmos
- Apps
- Ordens

A componente relacional entre coleções do banco foi feita de forma orientada a geração e disposição de métricas para soluções indoor Figura 23.

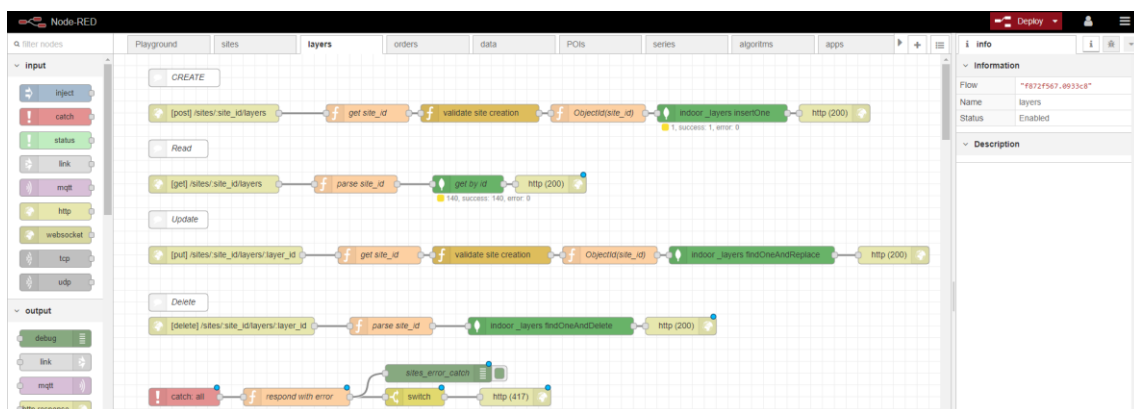


Figura 22: Implementação CRUD da API

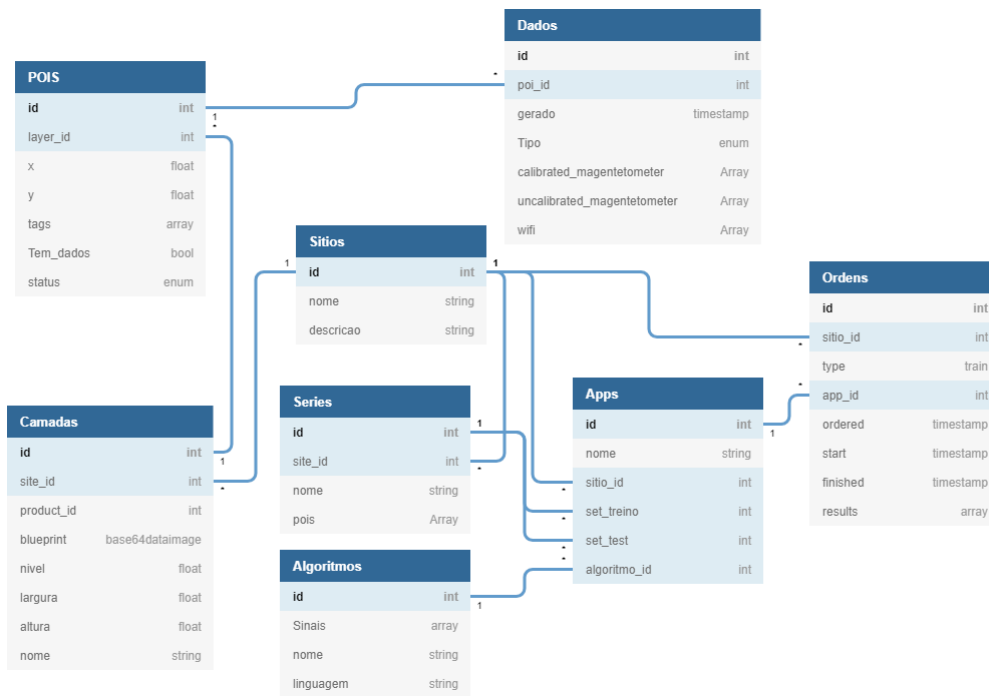


Figura 23: Modelo relacional do banco de dados

## Provedor de localização

Para a geração de resultados é necessária a utilização de um serviço capaz de receber uma ordem de treinamento e gerar os resultados supervisionados, para tanto foi implementado um script em NodeJS que é responsável por consumir a API afim de ordenar os dados de WIFI e campo magnético, combinar os mesmos e envia-los para um subprocesso em Python. O processo de formatação dos dados consiste em definir cada roteador com OUI padronizado pelo IEE como um parâmetro do algoritmo bem como cada um dos registros obtidos pelo magnetômetro. O subprocesso em Python recebe os dados e os aplica na biblioteca Sckiti-learn, que é uma biblioteca recorrente que, assim como a TensorFlow da Google, age como front-end para uma implementação em c ou c++.

Tabela 1: Algoritmos e sinais implementados no servidor de localização

Algoritmo	Wifi	Magnetômetro	Magnetômetro calibrado
KNeighborsRegressor	x		
KNeighborsRegressor	x	x	x
DecisionTree	x		
DecisionTree		x	

DecisionTree			x
DecisionTree		x	x
DecisionTree	x	x	
DecisionTree	x	x	x
SVR	x		

### Métricas

O serviço de localização é responsável por gerar um ou mais resultados para o algoritmo especificado na ordem de treinamento. Para cada algoritmo implementado diferentes parâmetros podem ser utilizados e com diferentes variações, a especificação detalhada de cada um dos algoritmos utilizados pode ser encontrada no Anexo I. O Serviço gera um vetor com combinações dos parâmetros especificados ao algoritmo em questão. É feita então uma varredura de parâmetros, tendo como princípio a imprevisibilidade dos sinais devida à dinamicidade dos meios pode provocar variações de resultados em função da planta e dos testes. O Provedor de localização então para cada combinação de parâmetros analisa os resultados quanto as métricas acerca da distribuição probabilidade de erros obtida como resultado, sendo as métricas [30]:

- Média harmônica

A média harmônica é uma das três médias de Pitágoras e é frequentemente utilizada para cenários em que ocorrências tem relação com a sua frequência.

$$\text{Média Harmônica} = MH = \frac{n}{\sum_{i=0}^n \frac{1}{e_i}}$$

$n$  = Número de elementos na distribuição de erro

$e_i$  = Elemento de index  $i$  na distribuição

- Média geométrica

A média indica a tendência central ou o valor típico de um conjunto de números usando o produto dos seus valores.

$$\text{Média geométrica} = MG = \left( \sum_{i=0}^n e_i \right)^{1/n}$$

$n$  = Número de elementos na distribuição de erro

$e_i$  = Elemento de index  $i$  na distribuição

- Variância

A variância medida da sua dispersão estatística, indicando "o quão longe" em geral os seus valores se encontram do valor esperado

$$\text{Média Aritmética} = \bar{x} = \frac{\sum_{i=0}^n ei}{n}$$

$$\text{Variância} = s^2 = \frac{\sum_{i=0}^n (ei - \bar{x})^2}{n - 1}$$

$n$  = Número de elementos na distribuição de erro

$ei$  = Elemento de index  $i$  na distribuição

- Intervalo

Diferença entre o maior e o menor erro da distribuição probabilidade de erro, ou último menos primeiro.

$$\text{Intervalo} = I = en - e0$$

$en$  = Maior erro da distribuição ordenada de erros

$e0$  = Menor erro da distribuição ordenada de erros

- Mediana

Mediana é o valor que separa a metade maior e a metade menor de uma amostra, uma população ou uma distribuição de probabilidade.

$$\text{Mediana} = Me = e \frac{n}{2}$$

$e \frac{n}{2}$  = valor central da distribuição ordenada de erros

- Distância entre quartis - IQR

Sendo um quartil a mediana da primeira metade da distribuição de probabilidade de erro, a distância entre quartis é definida como a diferença entre o segundo e o primeiro quartil.

$$\text{IQR} = Q2 - Q1$$

- Soma estatística

Sendo todas as métricas acerca da distribuição probabilidade de erro inversamente proporcionais ao resultado desejado de erro 0 em todas as incidências, a implementação de uma métrica que soma todas as métricas acerca da distribuição é implementada a fim de se buscar um resultado desejável.

$$SE = \text{IQR} + Me + I + s^2 + \bar{x} + MG + MH$$

- Cúbico 7.5

A fim de demonstrar a viabilidade de utilização de diversas métricas, inclusive métricas proprietárias na solução proposta, foi desenvolvida uma métrica que pontua de forma cúbica centrada inversamente em 7.5 metros, os erros na distribuição. A justificativa é que o autor entende que para muitas aplicações industriais um erro de até 7.5 metros é aceitável e, portanto, essa métrica estaria viabilizando aplicações como localizar grandes objetos em uma sala perante a uma planta de grandes dimensões.

$$\text{Cúbico 7.5} = \frac{\sum_{i=0}^n \frac{-e_i + 7.5^3}{7.5^3}}{n}$$

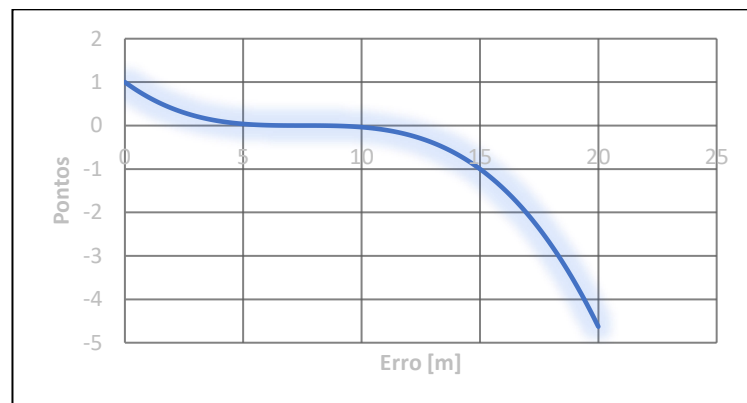


Figura 24: Pontuação cúbicamente proporcional ao erro

#### 4.4. Sumário

Com o intuito de gerar como contributo uma ferramenta de análise de soluções de localização indoor, foi realizado o desenvolvimento de uma arquitetura que é capaz de se integrar a soluções proprietárias, comparar soluções de diferentes provedores de localização e implementou-se o próprio provedor de localização. Em sua abrangência consta a flexibilidade de coleta e tratamento de dados proporcionada pela API e pela aplicação Android. Uma estrutura de armazenamento de dados projetada exclusivamente para o objetivo dessa ferramenta viabilizou a expansibilidade dela sem perda de compatibilidade com as funcionalidades existentes. O projeto foi implementado com compatibilização com ferramentas de software recorrentes na indústria, mas também com tendências em inovações de aprendizagem de máquina supervisionada como o Google Cloud Auto ML. Sendo assim a implementação do projeto cumpre com a proposta e em concordância com a justificativa.

## Capítulo 5      **Análise e Discussão de Resultados**

Nesse capítulo será apresentada uma análise dos resultados atingidos quanto a proposta e justificativa do projeto. Essa análise foi dividida conforme os diferentes componentes do projeto.

A arquitetura proposta pela ferramenta de avaliação de soluções indoor, inclui um serviço provedor de localização, essencial para demonstrar a funcionalidade da plataforma. A análise dos resultados acerca desse serviço será incluída no contexto da solução proposta.

Para os resultados apresentados, será demonstrada a utilização da plataforma no andar Térreo (Figura 25) e andar dos laboratórios (Figura 26), nos quais o volume de dados consta na Tabela 2.

Tabela 2: Relação de pontos de interesse coletados em cada experiência

	Treino	Teste	Total	% Treino	% Teste
ESTIG térreo	164	76	240	68.33%	31.67%
ESTIG laboratórios	60	30	90	66.67%	33.33%

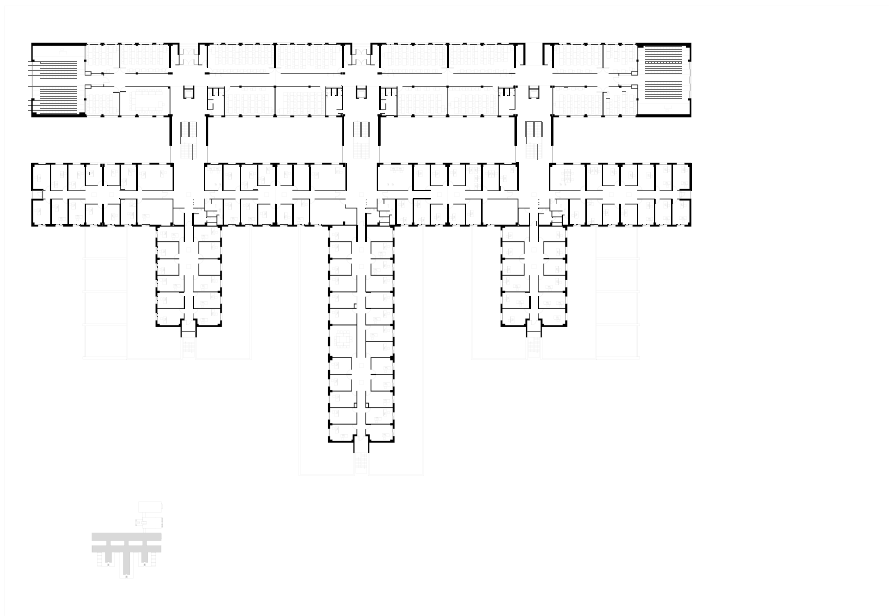


Figura 25: Andar Térreo da ESTIG

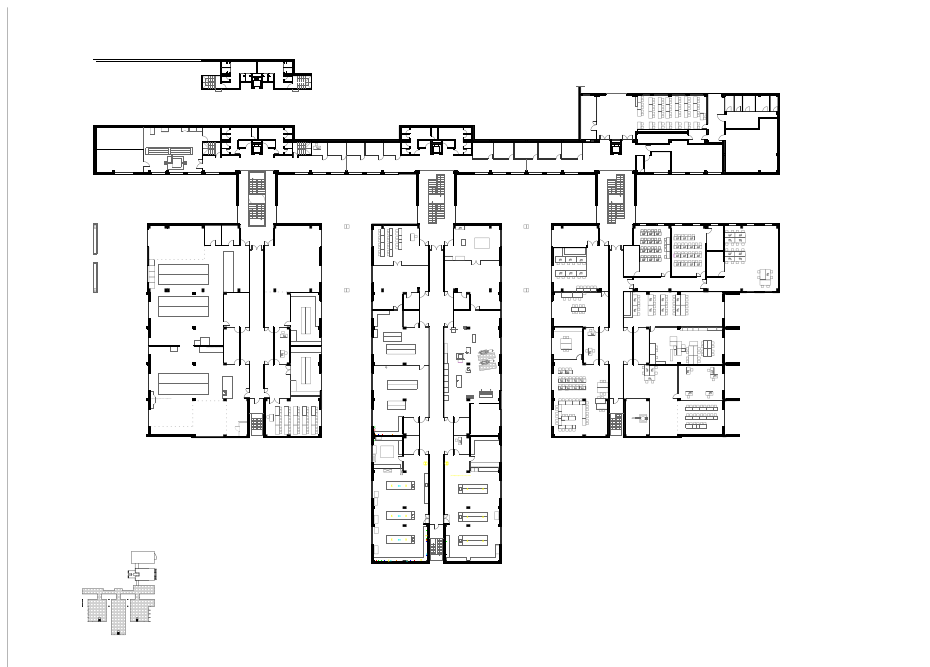


Figura 26: Andar dos laboratórios da ESTIG

## 5.1. interfaces com utilizador

### Ferramenta administrativa

A viabilização gráfica do controle de serialização de dados. A serialização de dados consistem na divisão de dados atrelados a pontos de interesse em grupos denominados series, e em grupos de series denominados sets, isso ocorre tanto para dados de treino quanto de teste, permite de forma intuitiva e passiva de esforço a experimentação do utilizador do conjunto de dados que ele possui quanto ao seu provedor de localização, conforme ilustrado na Figura 27. O provedor de localização é a parte da arquitetura responsável por receber os dados e utilizá-los para determinar a localização associada a eles, esse componente pode ser distribuído ou centralizado.

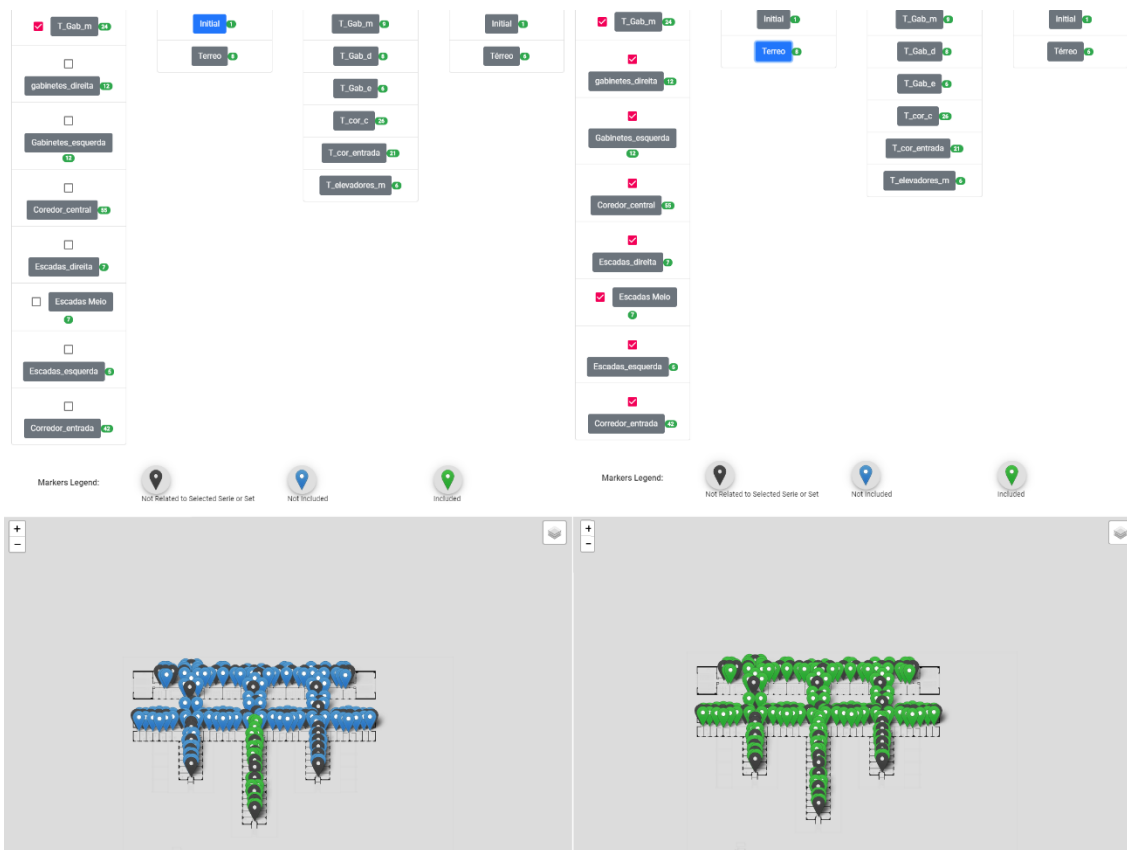


Figura 27: Resultado da funcionalidade de serialização de dados

A fim de cumprir com o objetivo do projeto, é fornecido ao utilizador uma tabela dos resultados que pode ser ordenável perante as métricas estabelecidas. Esse componente permite que seja feita uma análise qualitativa da solução em localização indoor devido capacidade de suporte a diversas métricas estatísticas e gráficos para fornecer maiores detalhes acerca das opções, conforme Figura 28.

App Name	Evaluation method	Metric Value	↑ Mean Error	Error Range	Variance	Standard deviation	Median
<input checked="" type="checkbox"/> L_W_KNN	Statistics_Sum	6.084	0.938	1.232	0.179	0.423	1.104
<input type="checkbox"/> L_W_KNN	Mean	0.938	0.938	1.232	0.179	0.423	1.104
<input type="checkbox"/> L_W_KNN	cubic7.5	0.677	0.938	1.232	0.179	0.423	1.104
<input type="checkbox"/> L_AMW_KNN	Mean	1.066	1.066	2.111	0.532	0.730	0.947
<input type="checkbox"/> L_AMW_KNN	cubic7.5	0.652	1.066	2.111	0.532	0.730	0.947

Figura 28: Resultado da organização e disposição dos resultados

## 5.2. Seleção de resultados

Entre os resultados obtidos na solução implementada destaca-se a sistematização fornecida ao utilizador no processo exploratório dos resultados. Relativo aos 240 pontos de interesse cujo dados foram coletados no andar térreo da ESTIG, foram criadas aplicações que combinam os conjuntos de dados serializados com os 9 Algoritmos desenvolvidos e dos conjuntos da Tabela 1. Os 45 resultados dispostos na Tabela 3 demonstram a necessidade de uma ferramenta para o auxílio na tomada de decisão.

Na Tabela 3: Resultados dos treinamentos ESTIG andar térreo cada linha corresponde a um resultado destacado pelo provedor de localização. Na coluna Aplicação consta o nome da aplicação que gerou esse resultado destacado, o nome da aplicação é um texto livre ao usuário que pode vir a implementar ou não algum padrão de nomenclatura, em algumas das aplicações dispostas foi utilizado o caractere “\_” afim de demarcar abreviaturas entre a Inicial da Aplicação que gerou o resultado, o conjunto de dados utilizado e o algoritmo que gerou o resultado. A segunda coluna, métrica, representa sob qual métrica esse resultado se destacou como melhor para essa aplicação e as colunas a seguir representam as métricas acerca do resultado destacado.

Tabela 3: Resultados dos treinamentos ESTIG andar térreo

Aplicação	Métrica	Erro médio	Intervalo	Variância	Desvio padrão	Mediana	IQR	Média geométrica	Média harmônica
I_W_KNN	Soma estatística	0.938	1.232	0.179	0.423	1.104	0.589	0.811	0.64

I_W_KNN	Média harmônica	1.641	3.92	2.025	1.423	1.725	2.266	0.528	0.011
I_W_KNN	Média geométrica	1.453	3.933	2.12	1.456	1.28	2.456	0.433	0.071
I_W_KNN	Média	0.938	1.232	0.179	0.423	1.104	0.589	0.811	0.64
I_W_KNN	Variância	1.893	1.067	0.131	0.362	1.978	0.535	1.86	1.827
I_W_KNN	Range	1.625	0.921	0.135	0.367	1.542	0.609	1.591	1.559
I_W_KNN	Mediana	1.79	5.693	4.26	2.064	0.734	2.574	0.936	0.508
I_W_KNN	IQR	1.57	2.125	0.38	0.616	1.458	0.274	1.485	1.413
I_W_KNN	Cúbico 7.5	0.938	1.232	0.179	0.423	1.104	0.589	0.811	0.64
I_AMW_KNN	Soma estatística	1.126	1.552	0.225	0.474	1.15	0.4	1.039	0.953
I_AMW_KNN	Média harmônica	2.671	6.306	6.477	2.545	1.728	4.444	1.06	0.115
I_AMW_KNN	Média geométrica	1.296	2.949	1.163	1.078	1.096	1.763	0.82	0.438
I_AMW_KNN	Média	1.066	2.111	0.532	0.73	0.947	1.078	0.857	0.689
I_AMW_KNN	Variância	1.589	1.239	0.152	0.39	1.559	0.525	1.544	1.497
I_AMW_KNN	Range	1.616	0.932	0.176	0.419	1.816	0.817	1.564	1.51
I_AMW_KNN	Mediana	1.368	3.952	1.792	1.339	0.761	1.286	0.98	0.763
I_AMW_KNN	IQR	1.609	1.352	0.156	0.395	1.618	0.395	1.561	1.506
I_AMW_KNN	Cúbico 7.5	1.066	2.111	0.532	0.73	0.947	1.078	0.857	0.689
Téreo	Soma estatística	3.151	13.222	9.175	3.029	2.042	3.483	2.029	1.158
Téreo	Média harmônica	6.987	50.318	71.774	8.472	3.609	8.326	3.244	0.112
Téreo	Média geométrica	2.952	15.575	9.957	3.155	1.736	2.826	1.7	0.47
Téreo	Média	2.952	15.575	9.957	3.155	1.736	2.826	1.7	0.47
Téreo	Variância	3.425	16.395	9.115	3.019	2.506	3.928	2.249	1.039
Téreo	Range	3.298	13.052	11.173	3.343	2.448	3.059	1.971	1.065
Téreo	Mediana	3.225	19.194	12.373	3.518	1.673	3.258	1.891	1.039
Téreo	IQR	3.288	19.763	12.25	3.5	2.52	2.274	2.035	1.052
Téreo	Cúbico 7.5	2.952	15.575	9.957	3.155	1.736	2.826	1.7	0.47
AM_Téreo	Soma estatística	7.987	25.268	38.813	6.23	6.005	8.546	5.544	2.838
AM_Téreo	Média harmônica	10.981	45.485	132.662	11.518	7.115	14.931	4.582	0.506
AM_Téreo	Média geométrica	7.827	45.452	113.283	10.643	2.836	9.251	2.878	0.788
AM_Téreo	Média	6.773	31.293	57.892	7.609	3.71	8.128	3.349	1.42
AM_Téreo	Variância	8.048	25.981	38.152	6.177	6.414	9.034	5.581	2.783

AM_Térreo	Range	7.991	25.122	39.686	6.3	6.005	8.623	5.512	2.805
AM_Térreo	Mediana	7.331	45.103	88.402	9.402	2.075	12.572	3.102	1.508
AM_Térreo	IQR	7.651	35.234	50.084	7.077	5.771	6.816	4.814	2.627
AM_Térreo	Cúbico 7.5	7.226	26.798	42.846	6.546	5.537	8.965	4.036	1.161
W_T_Tree	Soma estatística	4.177	15.247	15.241	3.904	2.539	4.74	2.648	1.636
W_T_Tree	Média harmônica	7.874	38.595	60.784	7.796	5.222	7.281	4.179	0.111
W_T_Tree	Média geométrica	3.823	20.073	19.38	4.402	2.334	4.156	1.913	0.593
W_T_Tree	Média	3.77	26.767	22.729	4.768	2.156	2.893	2.017	0.867
W_T_Tree	Variância	3.776	19.884	13.789	3.713	2.765	3.501	2.421	0.854
W_T_Tree	Range	4.177	15.247	15.241	3.904	2.539	4.74	2.648	1.636
W_T_Tree	Mediana	4.179	17.037	20.881	4.57	1.79	5.452	2.29	1.29
W_T_Tree	IQR	5.373	36.826	79.816	8.934	2.09	2.51	2.306	1.129
W_T_Tree	Cúbico 7.5	4.179	17.037	20.881	4.57	1.79	5.452	2.29	1.29

Devido ao eventual desafio envolvido na seleção de um modelo de localização indoor, o qual pode vir a possuir diversas possíveis configurações, ou ser selecionado priorizando diferentes métricas, a componente organizacional na interface gráfica, pretende facilitar a seleção dos resultados (Figura 28).

Entre as análises gráficas ofertadas pela plataforma, a primeira oferecida é a distribuição de probabilidade obtida a partir do espaço amostral de teste supervisionado. A ferramenta nesse contexto viabiliza a comparação entre as métricas dispostas das distribuições para uma dada aplicação (Figura 29).

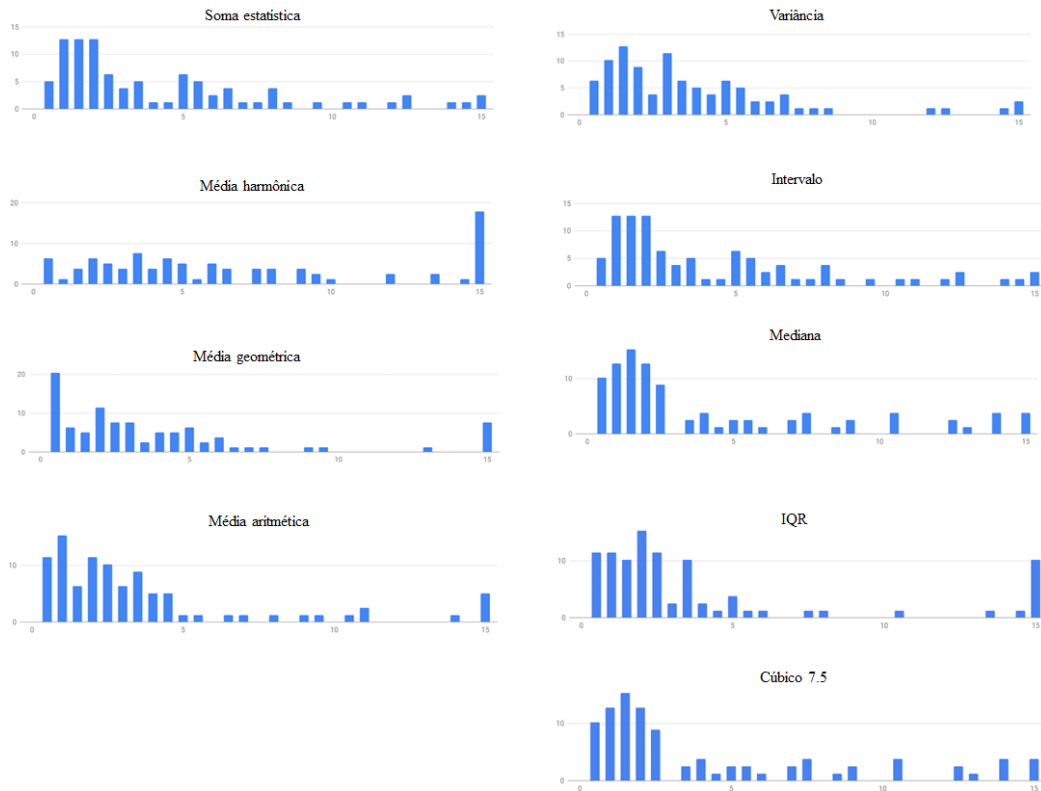


Figura 29: Distribuições de erro para as métricas de “W\_T\_Tree”

### 5.3. Casos de uso

Sob as métricas estabelecidas será avaliada a capacidade da ferramenta prestar suporte na análise de três aplicações utilizando dados das potências WIFI dos roteadores no andar térreo e dos laboratórios da ESTIG.

#### Aplicação no Andar Térreo

Utilizando dados das redes WIFI dos pontos coletados no Andar Térreo (Figura 30), aplicados ao algoritmo de regressão baseado em árvore de decisão, “Decision Tree Regressor”, foi possível obter a aplicação denominada “W\_T\_TREE”, a qual tem seus resultados dispostos na Tabela 4.

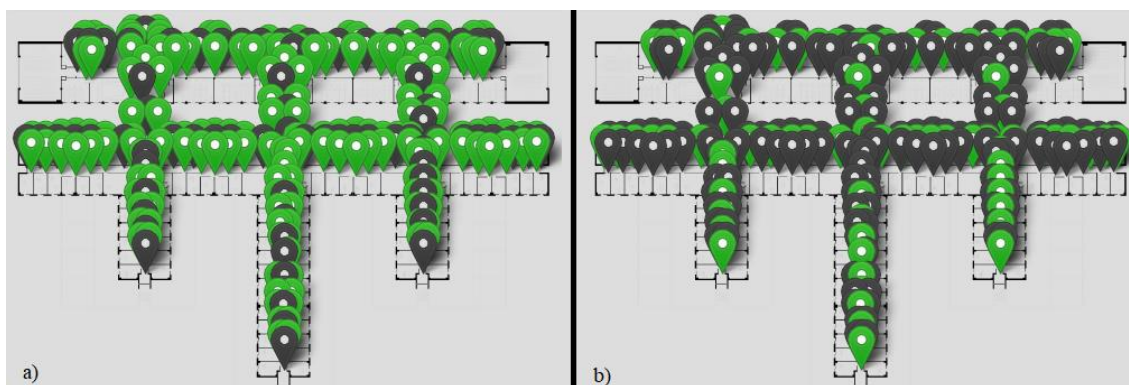


Figura 30: Conjunto de treino (a) e teste (b) da aplicação “W\_T\_TREE”

Tabela 4: Resultados dos treinamentos ESTIG andar térreo para aplicação “W\_T\_TREE”

	Erro médio	Intervalo	Variância	Desvio padrão	Mediana	IQR	Média geométrica	Média harmônica
Soma estatística	4.177	15.247	15.241	3.904	2.539	4.74	2.648	1.636
Média harmônica	7.874	38.595	60.784	7.796	5.222	7.281	4.179	0.111
Média geométrica	3.823	20.073	19.38	4.402	2.334	4.156	1.913	0.593
Média	3.77	26.767	22.729	4.768	2.156	2.893	2.017	0.867
Variância	3.776	19.884	13.789	3.713	2.765	3.501	2.421	0.854
Intervalo	4.177	15.247	15.241	3.904	2.539	4.74	2.648	1.636
Mediana	4.179	17.037	20.881	4.57	1.79	5.452	2.29	1.29
IQR	5.373	36.826	79.816	8.934	2.09	2.51	2.306	1.129
Cúbico 7.5	4.179	17.037	20.881	4.57	1.79	5.452	2.29	1.29

A ferramenta prestou suporte a análise do resultado por meio da disponibilização do gráfico de dispersão do erro, espalhamento 2D em planta e em versão limpa, além de uma forma de filtragem de erro afim possibilitar detecção de zonas de maior concentração de erro (Figura 31).

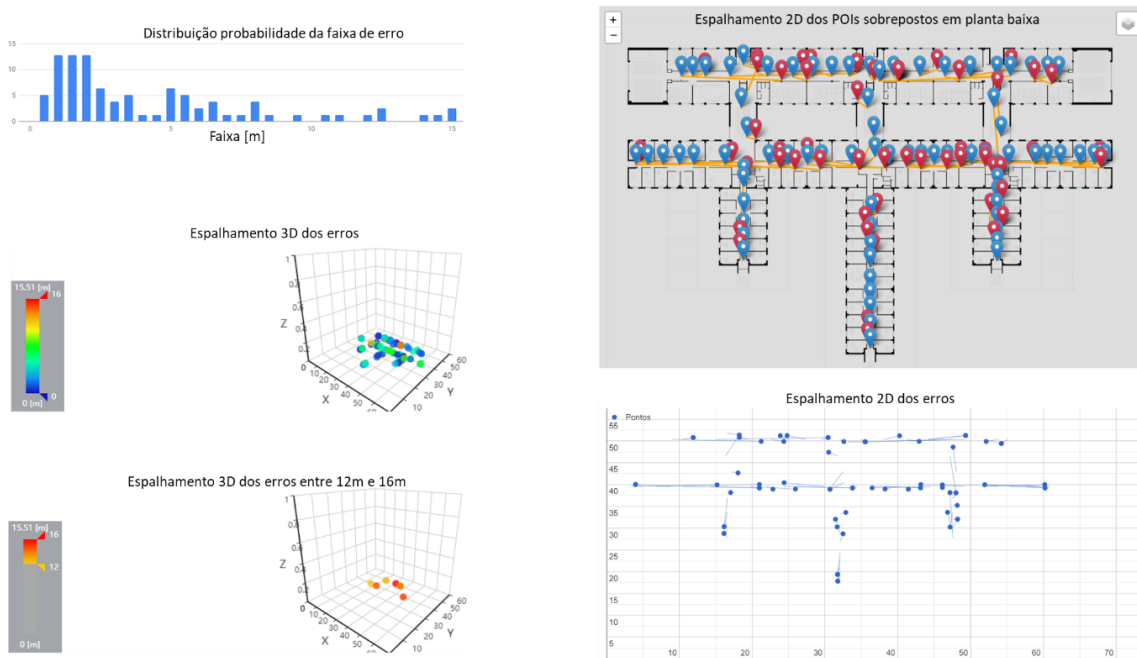


Figura 31: Gráficos disponibilizado pela ferramenta para aplicação “W\_T\_TREE” avaliada sob a métrica avaliada pela métrica Intervalo.

As estatísticas levantadas acerca da distribuição de erro obtida para aplicação “W\_T\_TREE” foram fornecidas de forma automatizada pela aplicação e constam na Tabela 5.

Tabela 5: Estatísticas da distribuição obtida pela aplicação “W\_T\_TREE” avaliada sob a métrica avaliada pela métrica Intervalo

Erro médio	4.18 [m]
Intervalo de erro	3.88 [m]
Variância	0.265 [m] ---- 15.5 [m]
Desvio padrão	3.90 [m]
Mediana	2.54 [m]
IQR	4.74 [m]
Ordem iniciada	2h :29m :20 s
Ordem finalizada	3h :07m :09 s

### Aplicação no Andar dos Laboratórios

O mesmo procedimento realizado no andar térreo para a aplicação “W\_T\_TREE” foi realizado no andar dos laboratórios para o conjunto de dados relacionados aos pontos da Figura 32. Gerando os resultados dispostos na Tabela 6, e ao selecionar a métrica Cúbico 7.5 a ferramenta proporciona os gráficos da Figura 33 e as estatísticas da Tabela 6.



Figura 32: Conjunto de treino (a) e teste (b) da aplicação “W\_KNN\_LABS”

Tabela 6: Resultados dos treinamentos ESTIG andar dos laboratórios “W\_KNN\_LABS”

Métrica	Erro médio	Intervalo	Variância	Desvio padrão	Mediana	IQR	Média geométrica	Média harmônica
Soma estatística	3.219	8.524	5.79	2.406	2.61	3.289	2.213	1.189
Média harmônica	3.804	10.984	10.032	3.167	2.85	3.966	2.208	0.283
Média geométrica	2.918	9.043	6.825	2.613	1.951	3.618	1.613	0.579
Média	2.918	9.043	6.825	2.613	1.951	3.618	1.613	0.579
Variância	3.458	10.106	5.563	2.359	2.951	2.678	2.498	0.969
Intervalo	3.614	8.068	6.619	2.573	3.322	4.221	2.603	1.523
Mediana	3.165	14.325	10.573	3.252	1.819	3.491	2.005	1.296
IQR	3.531	15.325	9.873	3.142	2.73	2.189	2.56	1.646
Cúbico 7.5	2.918	9.043	6.825	2.613	1.951	3.618	1.613	0.579

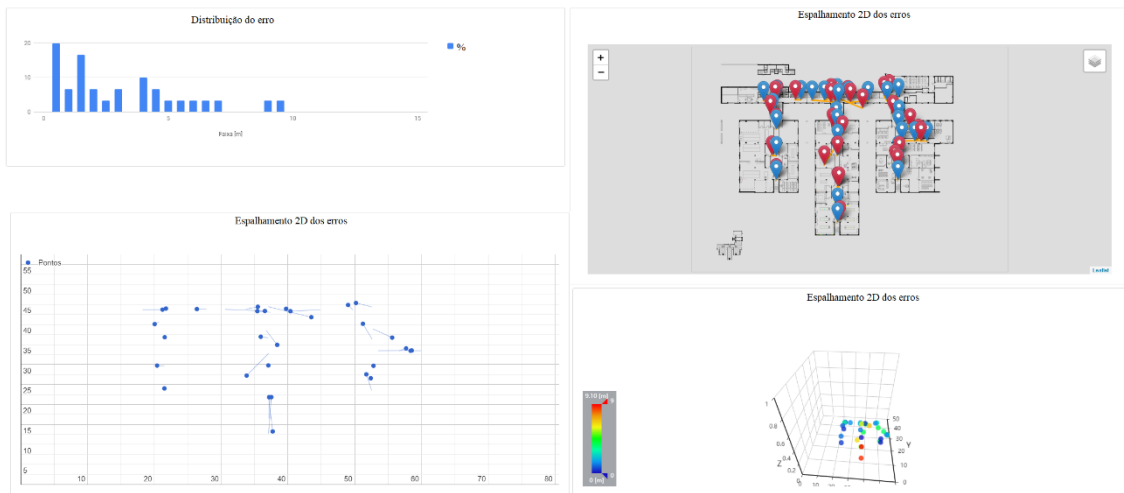


Figura 33: Gráficos disponibilizado pela ferramenta para aplicação “W\_KNN\_LABS” avaliada sob a métrica avaliada pela métrica Cúbico 7.5

Tabela 7: Estatísticas da distribuição obtida pela aplicação “W\_KNN\_LABS” avaliada pela métrica Cúbico 7.5

Erro médio	2.92 [m]
Intervalo de erro	0.0583 [m] ---- 9.10 [m]
Variância	6.83 [m <sup>2</sup> ]
Desvio padrão	2.61 [m]
Mediana	1.95 [m]
IQR	3.62 [m]
Ordem iniciada	0h :08m :25 s
Ordem finalizada	0h :14m :23 s

### Aplicação em escala reduzida

Afim de demonstrar o potencial ofertado pela ferramenta na flexibilização do uso dos dados coletados, foi implementada uma aplicação em escala reduzida contendo dados de potência dos roteadores de rede WIFI aplicados ao algoritmo K Nearest Neighbours Regressor, para essa aplicação os pontos de interesse foram restritos aos gabinetes centrais da planta do térreo da ESTIG, Figura 34.

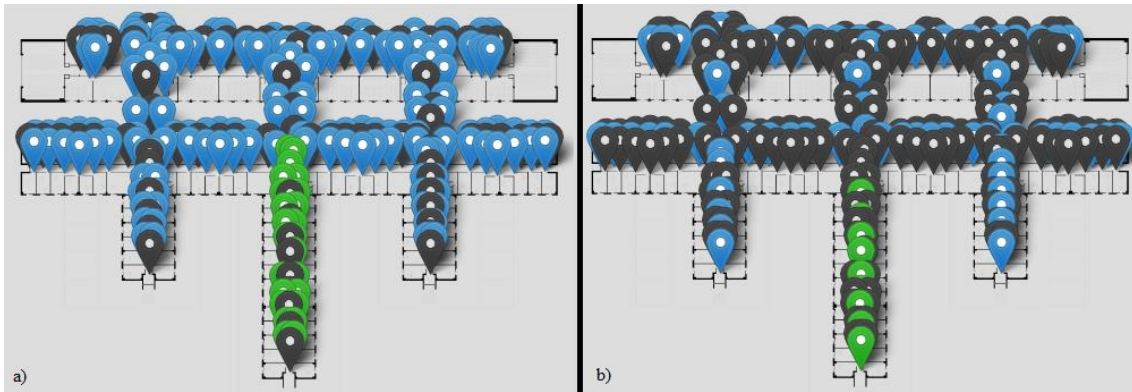


Figura 34: Conjunto de treino (a) e teste (b) da aplicação "I\_W\_KNN"

Para essa aplicação será avaliado o resultado sob a métrica média harmônica a qual a distribuição de erro obtida apresentou os resultados estatísticos relatados na Tabela 8. É apresentado também a forma gráfica da distribuição probabilidade dessa aplicação sob essa métrica, além dos gráficos de espalhamento, Figura 35.

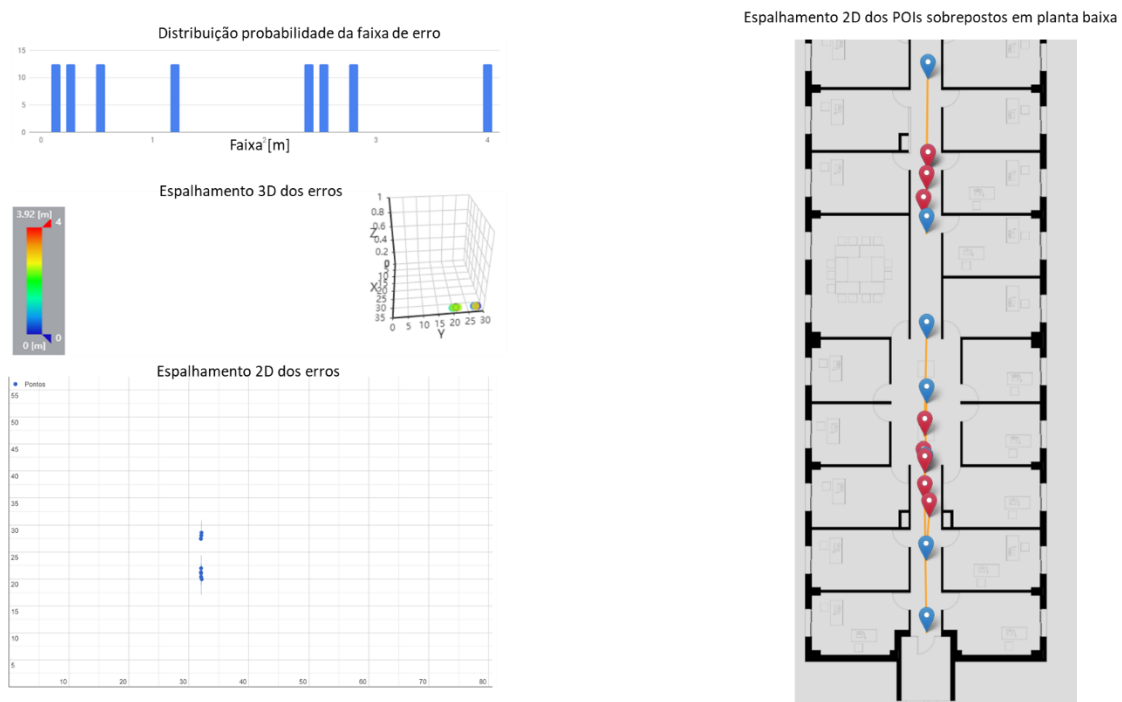


Figura 35: Gráficos disponibilizado pela ferramenta para aplicação "I\_W\_KNN" avaliada sob a métrica avaliada pela métrica Cúbico 7.5

Tabela 8: Métricas da distribuição obtida pela aplicação “I\_W\_KNN” avaliada pela métrica Cúbico 7.5

Erro médio	0.938 [m]
Intervalo de erro	0.207 [m] ---- 1.44 [m]
Variância	0.179 [m <sup>2</sup> ]
Desvio padrão	0.423 [m]
Mediana	1.10 [m]
IQR	0.592 [m]
Ordem iniciada	13h :26m :25 s
Ordem finalizada	13h :30m :01 s

## 5.4. Análise entre algoritmos e sinais

Um dos resultados da ferramenta produzida é a comparação facilitada entre o uso de diferentes algoritmos e de diferentes fontes de sinais. Utilizando o algoritmo KNeighborsRegressor avaliado sob o menor IQR é feita a comparação utilizando como parâmetros de entrada sinais magnéticos (Figura 36), combinando sinais WIFI e magnéticos (Figura 37), e somente utilizando sinais WIFI Figura 38.

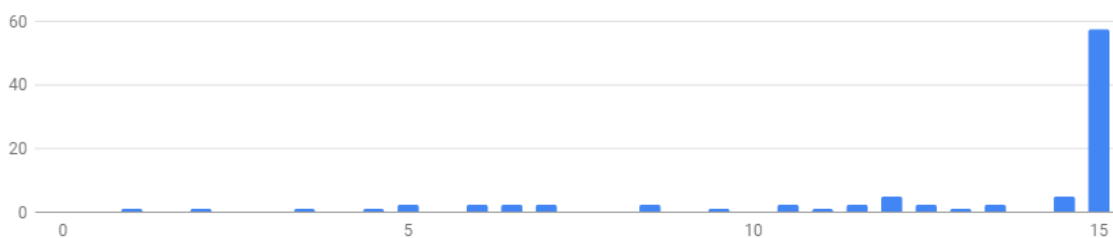


Figura 36: Distribuição para sinais magnéticos



Figura 37: Distribuição para sinais wifi + magnético



Figura 38: Distribuição para sinais wifi

Essa comparação permite destacar uma melhor performance somente ao utilizar como parâmetro o RSSI dos roteadores WIFI. Essa funcionalidade permite ao utilizador traçar uma estratégia quanto a solução a ser aplicada seja ela abandonando o uso de um tipo de parâmetro, nesse caso do campo magnético, ou até mesmo mudando a forma de tratamento de dados pelo algoritmo em questão.

## 5.5. Sumário

Foi possível graças ao aplicativo Android, que o levantamento de dados passasse a ficar dispensado de interação com a ferramenta como um todo, porém mantendo um feedback e controle do processo graças a interface fornecida. Esse tipo de implementação via uma API viabiliza unificação de informações em operações de coleta de dados como um processo em equipe. A aplicação se mostrou resiliente a erros do utilizador graças a duas estratégias, forçar duplo clique para coleta de dados, evitando que acidentalmente seja gerada uma fonte errada de dados, e também fornecendo uma forma simplificada de correção de má interpretação da planta cometidos no momento da coleta, isso graças a funcionalidade que permite a aplicação Android apagar os dados associados a um ponto.

A arquitetura da solução implementada permitiu o desenvolvimento do projeto por meio de sua centralização de funcionalidades em volta de uma API e banco de dados baseada em operações simples como leitura, criação, atualização e exclusão de dados do banco de dados. A arquitetura tem como característica a abertura para integrações futuras como um serviço de execução em tempo de produção dos modelos ou modificações na aplicação Android para levantamento de outros tipos de dados.

Quanto a pertinência dos algoritmos implementados frente aos dados fornecidos considera-se que a ferramenta demonstrou um resultado positivo uma vez que ainda nos piores resultados quanto as métricas implementadas não houve ocorrências fora de contexto como por exemplo todos os erros demasiadamente grandes. Na melhor performance identificada o provedor de localização se mostrou capaz de detectar com erro máximo de 1,5 metros a localização em um setor de uma das experiências.



## Capítulo 6 Conclusão e trabalhos futuros

O contributo desse trabalho foi demonstrado via a implementação de um provedor de localização capaz de trabalhar com sinais WIFI e de campo magnético coletados pela aplicação Android também desenvolvida para o escopo desse projeto. É possível graças ao nível de automação fornecido pela estrutura proposta a análise de resultados sob múltiplas métricas estatísticas acerca da solução, auxiliando no processo de desenvolvimento e escolha de soluções em localização indoor.

A arquitetura implementada permitiu a integração do projeto de forma a possibilitar que os dados fossem coletados, configurados, armazenados, processados e analisados. A implementação em caráter distribuído permitiu maior inserção em contexto de computação e nuvem, o que trouxe convergência com ferramentas de uso recorrente na indústria de software. A arquitetura foi pensada ainda visando integração com sistemas e algoritmos de terceiros uma vez que foi desacoplada a propriedade intelectual associada a localização indoor do resto da solução, isso viabiliza que a ferramenta seja utilizada em contexto de outra solução ou até mesmo de múltiplas soluções concomitantemente.

A centralização de funcionalidades em uma API, viabilizou o armazenamento dos dados em um banco de dados, o qual tem aplicabilidade em ambiente de produção, dispensando os consumidores de maior compreensão acerca da arquitetura estabelecida na componente relacional do banco de dados. O Banco de dados e sua arquitetura relacional implementada permitiu a obtenção de uma plataforma orientada a análise de soluções indoor. A facilidade de transferência e processamento dos dados ofertada pela escolha da framework NodeJS em conjunto com um banco de dados NOSQL baseado em documentos de notação de objetos na linguagem JavaScript, JavaScript Object Notation

(JSON). Entretanto apesar de maior impacto no processo de desenvolvimento da ferramenta, um banco de dados SQL poderia vir a apresentar maior capacidade relacional documentada o que tende a diminuir o número de requisições, aumentar a estabilidade e torna o projeto aplicável em cenários de maior fluxo de dados

A aplicação Android contribuiu como componente indispensável para o funcionamento do projeto, atuando como fonte de dados de sinais de potência dos roteadores WIFI e do sensor magnetômetro. A aplicação demonstrou contribuir também com a automação do processo de ligação dos dados aos pontos de interesse, processo que tende a crescer em complexidade conforme cresce a escala do caso de estudo.

Adicionalmente, visando permitir o reuso da ferramenta como facilitadora no processo de levantamento de dados, a aplicação desenvolvida em Android foi pensada para expansibilidade e compatibilização com terceiros. Na forma de estruturação de um código de aplicação nativa Android a aplicação é desenvolvida em seus componentes gráficos, em XML, e de funcionalidade em Java, onde cada Activity, ou tela, tem sua componente funcional atrelada. Em sua arquitetura a aplicação é independente da fonte de sinal até o momento que entra na tela de aquisição de dados, bastando então a adequação ou implementação de uma parte do projeto Android.

O serviço que atuou como provedor de localização cumpriu o propósito de demonstração da funcionalidade plena do projeto proposto, devido a utilização de algoritmos recorrentes na literatura, implementadas em bibliotecas de linguagens de alto nível, foi possível a obtenção de performances compatíveis com soluções comerciais.

A combinação de diferentes fontes de sinais nos algoritmos de localização serviu para demonstrar a capacidade da ferramenta de auxiliar na detecção de falhas de implementação, visto que os sinais de campo magnético vieram a impactar negativamente o processo de localização. A ferramenta permite ao utilizador uma percepção da atuação do sinal quanto as métricas obtidas, o que pode indicar um uso incorreto do mesmo, seja no emprego de uma fonte de dado indevida ou o mesmo que essa fonte foi utilizada de forma incorreta. Os casos de uso do provedor em conjunto com a flexibilidade de uso dos dados ofertada pela ferramenta administrativa, permitiram a prospecção de novas soluções, fator demonstrado pela maior precisão na aplicação em escala reduzida.

A interface gráfica da ferramenta administrativa demonstrou fornecer uma experiência simplista devido a utilização de bibliotecas de layout que seguem princípios de linguagem visual, tais como a Google Material UI e Bootstrap. Nota-se que a inclusão da planta baixa nas etapas de utilização da ferramenta permite que o utilizador se comunique constantemente com seu objeto de seu objetivo quanto a análise de resultados.

Destaca-se que devido as funcionalidades desenvolvidas na interface gráfica serem consequência da arquitetura implementadas no banco de dados e na API, é possível a interpretação do utilizador da mesma. A interface como reflexo visual direto da arquitetura provoca uma habituação do utilizador facilitando utilizar o projeto para integrações em outros contextos.

A interface proporcionou praticidade na configuração de experimentos e disposição dos resultados e a concordância visual entre a ferramenta administrativa e a aplicação Android contribuiu para o processo de coleta de dados.

Embora implementações de novas funcionalidades, métricas e interpretações visuais venham a ser prospectadas, conclui-se que a ferramenta proporcionou em conjunto com as outras partes integrantes do projeto a capacidade de análise de soluções em localização indoor sob diversas métricas.

## **6.1. Trabalhos futuros**

Entre as possíveis futuras evoluções da ferramenta administrativa destacam-se:

- Página de análise de estabilidade dos sinais ao longo do tempo
- Página de análise da variabilidade dos sinais ao longo dos planos cartesianos
- Suporte a detecção de andar e classificação da localização por setores da planta por implementação de sistema de cercas para viabilizar divisão em modelos específicos a setores da aplicação.
- Análise de percentual de POIs de teste e treino
- Análise de POIs por m<sup>2</sup>
- Fornecimento de uma API WIFI ou bluetooth do aparelho para integração com hardware externos ao aparelho celular
- Logística de armazenamento e escoamento de dados coletados offline

- Configuração de API via interface gráfica
- Integração com interfaces de linhas de comando de ferramentas de computação em nuvem
- Disponibilização nativa de ferramentas de aprendizagem automatizada de modelos supervisionados como Google Cloud Auto ML

## Bibliografia

- [1] Wang Peng, “Japan accelerates space militarization process - China Military.” [Online]. Available: [http://english.chinamil.com.cn/view/2019-02/25/content\\_9433921.htm](http://english.chinamil.com.cn/view/2019-02/25/content_9433921.htm). [Accessed: 14-Jun-2019].
- [2] “Uber Is Testing Its Self-Driving Cars on a Fake City in Pittsburgh | Inc.com.” [Online]. Available: <https://www.inc.com/business-insider/uber-built-fake-city-pittsburgh-almono-test-self-driving-cars.html>. [Accessed: 14-Jun-2019].
- [3] “What is TAPAS - TAPAS.” [Online]. Available: <http://www.tapasweb.dk/english/what-is-tapas>. [Accessed: 14-Jun-2019].
- [4] Wilson Sakpere, Michael Adeyeye-Oshin, and Nhlanhla B.W. Mlitwa, “A state-of-the-art survey of indoor positioning and navigation systems and technologies,” *SACJ* 29, pp. 1–53, Dec. 2017.
- [5] Enyang Xu, Zhi Ding, and Soura Dasgupta, “Source Localization in Wireless Sensor Networks From Signal Time-of-Arrival Measurements,” *IEEE Trans. SIGNAL Process. VOL 59 NO 6 JUNE 2011*, vol. 59, pp. 1–11, Jun. 2011.
- [6] Salim OUTEMZABET and Chahe NERGUIZIAN, “Accuracy Enhancement of an Indoor ANN-based Fingerprinting Location System Using Kalman Filtering,” *IEEE*, 2008.
- [7] A. Lin and Hao Ling, “Location tracking of indoor movers using a two-frequency Doppler and direction-of-arrival (DDOA) radar,” in *2006 IEEE Antennas and Propagation Society International Symposium*, Albuquerque, NM, USA, 2006, pp. 1125–1128.
- [8] U. Engel, “A geolocation method using ToA and FoA measurements,” in *2009 6th Workshop on Positioning, Navigation and Communication*, Hannover, Germany, 2009, pp. 77–82.
- [9] S. Sadowski and P. Spachos, “RSSI-Based Indoor Localization With the Internet of Things,” *IEEE Access*, vol. 6, pp. 30149–30161, 2018.
- [10] Rafael Saraiva Campos, “CONTRIBUTIONS TO RADIO FREQUENCY INDOOR POSITIONING AND THROUGH-THE-WALL MAPPING.” COPPE UFRJ, Jun-2017.
- [11] Hongwei Xie, Tao Gu, Xianping Tao, Haibo Ye, and Jian Lv, “A Practical Magnetic Fingerprinting Approach to Indoor Localization using Smartphones,” presented at the UBICOMP, 2014.
- [12] Tareq Alhmiedat, Ghassan Samara, and Amer O. Abu Salem, “An Indoor Fingerprinting Localization Approach for ZigBee Wireless Sensor Networks,” *Eur. J. Sci. Res.*, vol. 105, no. 2, pp. 190–202, Jul. 2013.
- [13] CISCO, “Cisco Hyperlocation Solution.” CISCO, 05/18.
- [14] T. Here, “Indoor Positioning,” *HERE Developer*. [Online]. Available: <https://developer.here.com/products/indoor-positioning>. [Accessed: 02-Jun-2019].
- [15] Z.-A. Deng, Y. Hu, J. Yu, and Z. Na, “Extended Kalman Filter for Real Time Indoor Localization by Fusing WiFi and Smartphone Inertial Sensors,” *Micromachines*, vol. 6, no. 4, pp. 523–543, Apr. 2015.
- [16] “nginx documentation.” [Online]. Available: <http://nginx.org/en/docs/>. [Accessed: 12-Jun-2019].
- [17] “Node-RED : Documentation.” [Online]. Available: <https://nodered.org/docs/>. [Accessed: 12-Jun-2019].

- [18] “Getting Started – React.” [Online]. Available: <https://reactjs.org/docs/getting-started.html>. [Accessed: 12-Jun-2019].
- [19] “Angular - Introduction to the Angular Docs.” [Online]. Available: <https://angular.io/docs>. [Accessed: 12-Jun-2019].
- [20] “Introduction — Vue.js.” [Online]. Available: <https://vuejs.org/v2/guide/>. [Accessed: 12-Jun-2019].
- [21] “Documentation - Let’s Encrypt - Free SSL/TLS Certificates.” [Online]. Available: <https://letsencrypt.org/docs/>. [Accessed: 12-Jun-2019].
- [22] “HTTPS,” *Wikipedia*. 06-Jun-2019.
- [23] “React Google Charts.” [Online]. Available: <https://react-google-charts.com/>. [Accessed: 12-Jun-2019].
- [24] “Getting Started with Redux · Redux.” [Online]. Available: <https://redux.js.org/>. [Accessed: 12-Jun-2019].
- [25] “Documentation,” *Android Developers*. [Online]. Available: <https://developer.android.com/docs>. [Accessed: 12-Jun-2019].
- [26] N. js Foundation, “Docs,” *Node.js*. [Online]. Available: <https://nodejs.org/en/docs/>. [Accessed: 12-Jun-2019].
- [27] “Documentation scikit-learn: machine learning in Python — scikit-learn 0.21.2 documentation.” [Online]. Available: <https://scikit-learn.org/stable/documentation.html>. [Accessed: 12-Jun-2019].
- [28] “Mosquito man page,” *Eclipse Mosquitto*, 17-Apr-2019. [Online]. Available: <https://mosquitto.org/man/mosquitto-8.html>. [Accessed: 12-Jun-2019].
- [29] “Cloudflare - The Web Performance & Security Company | Cloudflare.” [Online]. Available: <https://www.cloudflare.com/>. [Accessed: 12-Jun-2019].
- [30] K. Krishnamoorthy, *Handbook of statistical distributions with applications*. Boca Raton: Chapman & Hall/CRC, 2006.