

Luís Filipe Pinho Correia

Recolha e Detecção de Padrões Geoespaciais



Departamento de Ciência de Computadores
Faculdade de Ciências da Universidade do Porto
Agosto de 2012

Luís Filipe Pinho Correia

Recolha e Detecção de Padrões Geoespaciais

*Tese submetida à Faculdade de Ciências da Universidade do Porto
para obtenção do grau de Mestre em Ciência de Computadores*

Orientador: Vitor Santos Costa

Co-orientador: Roberto Ugo Di Cera Colazingari

Departamento de Ciência de Computadores
Faculdade de Ciências da Universidade do Porto

Agosto / 2012

Resumo

Dadas as variadas funções disponibilizadas hoje em dia pelos terminais moveis e a ubiquidade destes é de estranhar a ausência de um sistema de localização *indoor*. Se considerarmos a localização *outdoor*, o Global Positioning System (GPS) surge como uma solução viável e eficiente, com mecanismos secundários de e alternativas como o Assisted GPS (A-GPS)¹ entram em utilização aquando a estimativa da nossa localização não é suficientemente boa para o efeito pretendido.

No entanto se um terminal móvel necessitar de se localizar no interior de um edifício ou no subsolo, a receção do sinal GPS não será certamente suficiente para obter a informação essencial para o efeito, pois estes requerem linha de vista. A alternativa, o A-GPS poderá entrar em funcionamento nos terminais compatíveis² mas a margem de erro será ditada pela cobertura da antena que serve o terminal móvel no momento. A menos que o utilizador comum esteja disposto a carregar consigo mais um dispositivo para se localizar dentro de um edifício (efetivamente carregar um dispositivo que sirva de “etiqueta” para um sistema de localização)

Existe no entanto solução para localização *indoor* cujos requisitos já se encontram cumpridos e por isso não requerem alteração ou renovação do *hardware* existente nos dispositivos móveis. Será este o assunto do nosso trabalho: estudar metodologias, aproveitando o *hardware* já em utilização, que permitam localizar dispositivos que comuniquem por ondas rádio. Para além disso, será também estudado a melhor forma de recolher, armazenar e processar esses dados de localização de modo a gerar informação de um modo eficiente, consoante as necessidades específicas de uma aplicação *location-aware* baseada em serviços de localização geográfica.

¹A-GPS entra em uso quando a receção do sinal GPS não é suficientemente bom para o tipo de localização pretendida e consiste na utilização da dados provenientes de outros sinais rádio, tais como estação mais próxima do provedor de serviços.

²a compatibilidade com A-GPS consiste em *hardware* dotado de uma maior sensibilidade no *hardware* de receção de sinais rádio de modo a medir a potência destes com mais precisão. Porém este tipo de comportamento pode ser emulado por *software*.

Deixo os meus agradecimentos à minha família. Deixo também às pessoas que me apoiaram na elaboração deste trabalho. Obrigado.

Conteúdo

Resumo	3
Lista de Tabelas	8
Lista de Figuras	10
1 Introdução	14
1.1 Motivação	15
1.2 Objetivos	15
2 Background	17
2.1 Serviços Baseados em Localização	17
2.2 Tecnologias de Serviços Baseados em Localização <i>indoor</i>	18
2.2.1 Bluetooth	20
2.2.2 Comunicação Entre Dispositivos	23
2.2.3 Wi-Fi	25
2.2.4 GSM	30
2.2.5 Outras Tecnologias para Localização <i>Indoor</i>	36
2.2.5.1 Infravermelhos	36
2.2.5.2 Ultrassons	37
2.2.5.3 <i>Radio Frequency Identifications</i> (RFIDs)	37

2.3	Detecção e Localização de Sinais de Rádio	38
2.3.1	Propagação de Sinal	40
2.3.2	<i>Software Defined Radio</i>	42
2.3.3	Sistemas Ativos e Sistemas Passivos	44
2.3.3.1	Sistemas Ativos	44
2.3.4	Utilizando a força de sinal observada	46
2.3.5	Algoritmos matemáticos	46
2.3.5.1	Trilateração	46
2.3.5.2	Triangulação	51
2.3.5.3	Trilateração com Recurso a Extrapolação	53
2.4	Sistemas de Bases de Dados <i>NoSQL</i>	55
3	Sistema	58
3.1	No local	60
3.1.1	Servidor Local	60
3.1.2	Master e Slave	61
3.1.3	Antenas Captura (BT/WIFI/GSM)	71
3.1.4	Kinect	73
3.1.5	Comunicação entre Componentes	77
3.1.6	Trilateração	78
3.2	O Sistema Distribuído	80
3.2.1	Servidor Global	80
3.2.2	Cassandra	82
3.2.3	PostgreSQL/PostGIS	84
3.2.4	NodeJS	85

4	Resultados	86
4.1	Local de estudo	87
4.2	Resultados obtidos com Bluetooth	89
4.3	Resultados Obtidos com Wi-Fi	90
4.4	Resultados Obtidos com <i>GSM</i>	91
5	Discussão dos Resultados e Trabalho Futuro	94
5.1	Discussão dos Resultados	94
5.1.1	Sobre as Capturas	96
5.2	Conclusões	97
5.3	Trabalho Futuro	98
A	Acronyms	100
	Referências	101

Lista de Tabelas

2.1	Protocolo Bluetooth	21
2.2	Versões da especificação IEEE 802.11(a/b/g/n)	25
2.3	Resumo dos campos de endereçamento de rede da camada <i>MAC</i>	29
2.4	Resumo de Sistemas de Base de Dados <i>NoSQL</i>	57
4.1	Resultados Bluetooth	89
4.2	Valores de <i>Received Signal Strength Indicator</i> (RSSI) da fase <i>offline</i> para os resultados Bluetooth (BT)	89
4.3	Resultados Wi-Fi	91
5.1	Resultados Wi-Fi com diferentes antenas e tempos de recolha	96

Lista de Figuras

2.1	Tipologias de rede Bluetooth	23
2.2	Tipologias de rede Wi-Fi	26
2.3	Tipologia de uma rede <i>Global System for Mobile Communications</i> ou <i>Groupe Spécial Mobile</i>	32
2.4	Sensores <i>cricket</i>	38
2.5	Propagação de Sinais Rádio <i>versus</i> Infravermelhos ou Ultrassons dentro de um edifício	39
2.6	Alternativas para a localização de dispositivos <i>indoor</i>	39
2.7	Ilustração do envio de dados por sinais rádio	40
2.8	Ilustração de uma antena omni-direcional e uma antena direcional (em cima e em baixo, respetivamente)	42
2.9	USRP1 da <i>Ettus Research</i> , com um rádio para envio e outro rádio para recepção	43
2.10	Exemplo de Utilização do Sistema Para Iluminação	46
2.11	Demonstração do processo de trilateração	47
2.12	Exemplo da intersecção de três círculos consoante a precisão das medições obtidas para a distância	49
2.13	Determinação do centróide da intersecção entre dois círculos	49
2.14	Triangulação usando os ângulos entre o ponto a ser calculado e as referências	52
2.15	Exemplo de possível propagação de sinais rádio	54

2.16	Solução para a Trilateração com erro através de extrapolação	55
3.1	Visão Global dos Componentes do Sistema	59
3.2	Visão global da implementação dos sistemas Master e Slave	63
3.3	Lógica do processo de filtragem	64
3.4	Visão global da implementação dos sistemas Master e Slave	68
3.5	Diagrama de sequência das rotinas dos componentes Master e Slave do Sistema de Localização	70
3.6	Antenas de 2.4 GHz utilizadas no sistema	72
3.7	Padrão de propagação das antenas 2.4GHz utilizadas	73
3.8	Antenas de GSM-900/DCS-1800 utilizadas no sistema	73
3.9	Esquema UML de sequência das mensagens trocadas desde uma antena Master até ao Servidor Global	78
3.10	Mapeamento do local de trabalho situado em um andar que se encontra dividido entre duas empresas	81
3.11	Organização dos dados para o <i>Cassandra</i>	84
3.12	Modelo Relacional para dados geo-espaciais	84
4.1	<i>Screenshot</i> da aplicação que implementa o algoritmo de localização a um conjunto de dados	86
4.2	Planta do local de estudo	87
4.3	Localização dos pontos (<i>P1</i> a <i>P9</i>) a estudar no local de teste	88
4.4	<i>Screenshot</i> da interface criada para estudar visualmente o comporta- mento do sistema	90
4.5	<i>Screenshot</i> dos dados <i>Global System for Mobile Communications</i> ou <i>Groupe Spécial Mobile</i> (GSM) recolhidos para estimação da localização do terminal móvel	93

GPS Global Positioning System	3
A-GPS Assisted GPS	3
WiFi Wi-Fi	15
BT Bluetooth	8
GSM Global System for Mobile Communications	10
SBL Serviços Baseados em Localização	17
RFID <i>Radio Frequency Identification</i>	5
UWB <i>Ultra-Wide Band</i>	18
TDMA <i>Time Division Multiple Access</i>	22
RSSI <i>Received Signal Strength Indicator</i>	8
SSID <i>Service Set Identifier</i>	27
GSM <i>Global System for Mobile Communications</i> ou <i>Groupe Spécial Mobile</i>	10
BSC <i>Base Station Controller</i>	33
BTS <i>Base Transceiver Station</i>	32
BSS <i>Base Station Subsystem</i>	33

	12
MS <i>Mobile Subscriber</i>	30
SDR <i>Software Defined Radio</i>	43
MSC <i>Mobile Switching Center</i>	33
HLR <i>Home Location Register</i>	
VLR <i>Visitor Location Register</i>	
AuC <i>Authentication Center</i>	
EIR <i>Equipment Identity Register</i>	
IMEI <i>International Mobile Equipment Identity</i>	31
OMC <i>Operation and Maintenance Center</i>	
SIM <i>Subscriber Identity Module</i>	31
PIN <i>Personal Identification Number</i>	35
IMSI <i>International Mobile Subscriber Identity</i>	31
TMSI <i>Temporary Mobile Subscriber Identity</i>	67
SSL <i>Secure Sockets Layer</i>	59
ARM <i>Advanced RISC Machine</i>	71

ARFCN <i>Absolute Radio-Frequency Channel Number</i>	31
LUR <i>Location Updating Request</i>	67
DCCH <i>Dedicated Control Channel</i>	67
SDCCH <i>Standalone Dedicated Control Channel</i>	67
AGCH <i>Access Grant Channel</i>	67
CCCH <i>Common Control Channels</i>	67
NoSQL <i>Not Only SQL</i>	56

Capítulo 1

Introdução

Dadas as variadas funções disponibilizadas hoje em dia pelos terminais moveis e a ubiquidade destes é de estranhar a falta de um sistema de localização *indoor* mais eficiente e sobretudo mais prático do que os atualmente existentes. Se falarmos de localização *outdoor*, o GPS surge como solução viável e eficiente, e alternativas como o entram em utilização aquando a estimativa da nossa localização não é suficientemente boa para o efeito pretendido[EM99].

No entanto se um terminal móvel necessitar de se localizar no interior de um edifício ou no subsolo, a receção do sinal GPS não será certamente suficiente para obter a informação essencial para o efeito, pois estes requerem linha de vista. Como alternativa, o A-GPS poderá entrar em funcionamento nos terminais compatíveis mas a margem de erro será ditada pela cobertura da antena que serve o terminal móvel no momento. A menos que o utilizador comum esteja disposto a carregar consigo mais um dispositivo para se localizar dentro de um edifício (efetivamente carregar um dispositivo que sirva de “etiqueta” para um sistema de localização)

Iremos explorar uma solução para localização *indoor* cujos requisitos já se encontram cumpridos e por isso não requerem alteração ou renovação do *hardware* existente nos dispositivos móveis. Será este o assunto do nosso trabalho: estudar metodologias que permitam localizar dispositivos que comuniquem por ondas rádio, aproveitando o *hardware* já em utilização. Para além disso, será também estudado a melhor forma de recolher, armazenar e processar esses dados de localização de modo a gerar informação de um modo eficiente, consoante as necessidades específicas de uma aplicação *location-aware* baseada em serviços de localização geográfica.

1.1 Motivação

Atualmente o telemóvel é um dispositivo indispensável e a sua utilização encontra-se perfeitamente enraizada no dia-a-dia nos países industrializados. Uma das características facilmente observável nestes países é a forte concorrência entre competidores dentro da mesma área de operação e a resultante evolução desta. É então deste modo que se explica a história da evolução dos telemóveis, principalmente das funcionalidades que hoje oferecem: funcionam autonomamente, sem fios, e oferecem uma abundância de serviços que sejam do interesse do seu utilizador e visam facilitar a sua vida quotidiana. Isto porque permitem que os seus portadores estejam sempre “ ligados ao mundo”, em contacto com quem querem e com o que querem.

Uma das funcionalidades cada vez mais comuns nos dispositivos móveis é a existência de mecanismos que por uso direto ou indireto dos mesmos providenciam a localização atual do dispositivo. Entre eles, o mais comum e preciso é o GPS, em que o dispositivo recetor recorre a uma rede global de satélites específicos para o efeito e obtém a sua localização atual através da posição de cada satélite e do tempo a que essa posição foi transmitida pelo satélite. A este tipo de sistema, chamaremos sistema direto de localização, pois o seu funcionamento é específico para o efeito e dependente da vontade explícita de quem o usa. Para além do GPS, existem outras funcionalidades nos dispositivos moveis que permitem obter a localização geográfica do seu utilizador. Estes, embora não tenham sido idealizados para tal fim, podem até um certo grau de precisão localizar um dispositivo móvel, devido ao seu modo de funcionamento. A este tipo de localização chamaremos indireto pois o facto de ser possível extrair dados que localizam a origem de transmissão, nenhum tipo de coordenada está explícita ou propositadamente contida na informação transmitida para esse efeito. É ainda parte da nossa caracterização de indireta o facto de ser possível extrapolar a posição do dispositivo móvel pois o utilizador pode não saber, e/ou querer que tal informação seja conhecida. Mas por estar a usar funcionalidades como o Wi-Fi (WiFi), o BT ou ainda o comum GSM do telemóvel (por exemplo) a sua posição - ainda que relativa às antenas da mesma tecnologia em seu redor - pode ser conhecida.

1.2 Objetivos

Neste trabalho usaremos comunicação sem fios das tecnologias BT, WiFi e GSM para localização de terminais móveis sem recorrer a equipamento adicional. Acreditamos

que as particularidades de cada tecnologia, tanto na informação que oferece durante o seu funcionamento, como nos requisitos operacionais necessários para a troca de dados, poderá permitir a implementação de um sistema de localização sem alterar a infraestrutura de rede. Iremos, por isso, idealizar a recolha e o armazenamento desde a fonte para que, dentro do possível, os dados de localização possam servir aplicações *location-aware*, se possível em tempo-real, tal como os sinais GPS. O modo de recolha será o menos intrusivo possível, de modo a requerer uma interação nula ou única por parte do utilizador.

Para tal, restringiremos apenas a recolha de dados de localização: consoante o movimento a detetar, dependendo unicamente da velocidade do alvo¹. Por isso, o comportamento do movimento do objeto a detetar ditará o tempo de recolha necessário antes de determinar uma estimativa precisa para o local de deteção. Só então uma posição será “detetada” pelo sistema. Isto servirá para “atomizar” os dados de uma deteção.

De seguida os dados serão centralizados num servidor aglomerador de deteções para que sejam distribuídos num sistema de base de dados. A finalidade deste sistema é tornar os dados globalmente disponíveis para consulta. Dado que não passam de deteções “em bruto”, uma aplicação *location-aware* terá de consultar os dados consoante as suas necessidades para o momento e processa-los de modo a gerar a informação de acordo com a finalidade pretendida.

¹Um ser humano tem em princípio uma velocidade média bastante inferior que um veículo, e como consequência um detetor apanhará uma menor quantidade de *broadcasts* de *beacons* de localização

Capítulo 2

Background

2.1 Serviços Baseados em Localização

O rápido desenvolvimento a nível tecnológico na área dos dispositivos móveis tem vindo a proporcionar um poder de processamento cujas possibilidades dependem, cada vez mais, apenas da criatividade humana em inventar novos conceitos ou soluções e no talento para tornar esses conceitos possíveis. Também o impulso comercial impele a nossa imaginação a encontrar novas ideias de modo a prevalecer sobre os restantes concorrentes, ministrando para isso os meios necessários para que surjam novas oportunidades de proporcionar mais e melhores funcionalidades. No que toca a Serviços Baseados em Localização (SBL) tem-se verificado ultimamente uma crescente adoção principalmente espalhada pelos meios de comunicação e partilha social, que por esse motivo são cada vez mais comuns[Sri05].

Começaremos por definir precisamente Serviços Baseados em Localização[VMG⁺01]:

“Location Based Services are information services accessible with mobile devices through the mobile network and utilizing the ability to make use of the location of the mobile device.”

A crescente adoção por parte dos consumidores a terminais móveis que suportem especificamente este tipo de serviços, e o correspondente suporte a nível de infraestruturas em grandes ambientes fechados - especialmente comerciais, mas também museus, estações de comboios e aeroportos - tornam possível a implementação e consequente utilização de um serviço de geo-referenciação.

2.2 Tecnologias de Serviços Baseados em Localização *indoor*

Existem variadas alternativas tecnológicas[GLN09] que possibilitam a localização de dispositivos moveis *indoor*. Embora variem entre si em diversos aspetos, principalmente no meio de comunicar com o sistema, todas tendem a recorrer a um mecanismo físico (*hardware*) ou de *software* para localizar o utilizador. Essencialmente qualquer sistema será composto por 2 componentes básicos: o que localiza e o que é localizado. Enquanto que o primeiro é essencial para que o segundo receba as coordenadas corretas que o localizam, sem o localizador nada haveria para localizar. A principal característica dos localizadores será o esforço necessário para a captação dos sinais do dispositivo a localizar. Quanto ao segundo componente, o que é localizado, a principal característica será o modo como transmite ou recebe dados de localização. Tal sistema é então, fundamentalmente, um sistema composto por emissores e recetores. Podemos por isso perceber que quanto maior for a área de serviço abrangida e menor a densidade de localizadores para essa mesma área, mais eficiente será o sistema. Por outro lado, quanto maior for a autonomia e o alcance dos terminais a localizar em relação ao seu peso e tamanho, mais prático será a utilização do sistema. Caso seja necessário não só localizar um terminal móvel, mas também providenciar um modo de o distinguir, é imprescindível a existência de um sistema de identificação. Tal distinção obriga a que haja um mecanismo que garanta a perceção individual dos dispositivos que estão presentes no sistema. Para isso será necessário etiquetar as mensagens de modo a dar a conhecer ao sistema a identidade de cada terminal móvel (ou a origem das mensagens).

As tecnologias mais complexas e dependentes de equipamento específico para o efeito - tanto da parte do emissor como do recetor - tendem a ser as mais precisas. Falamos de tecnologias como os infravermelhos[KB97], os RFID, ultra-sons ou os baseados em *Ultra-Wide Band* (UWB). Neste tipo de opções será necessário que o dispositivo a localizar suporte a receção e/ou emissão num destes espectros passíveis de transmitir dados - a localização por ultra-sons, por exemplo, pode funcionar em conjunto com sinais rádio para sincronizar e coordenar os dispositivos a localizar com os emissores. Apesar de válidas, estas opções não têm sido adotadas em larga escala não só devido à relação custo/benefício demasiada alto que acarretam mas também à falta de comodidade. Tome-se como exemplo uma pessoa que se encontra à entrada de uma zona de grande retalho e pretende, dando a posição atual, primeiro encontrar o andar, seguido da zona, da loja e depois a prateleira onde se situa um determinado objeto. Os SBL que se baseiem nas alternativas mencionadas neste parágrafo - como por exemplo o

Active Badge[WHFG92] ou o *Cricket System*[PCB00], necessitariam que o indivíduo carregasse consigo um dispositivo próprio para o efeito, que funcione numa tecnologia que - ou por fraca adoção, ou por requisitos do sistema - é caro ou funcione apenas com linha de vista com o dispositivo emissor, respetivamente. Este dispositivo extra, teria ainda para além de suportar tal tecnologia, ser capaz de processar a sua localização atual e, para que fosse prático, informar o utilizador das próximas mudanças de direção. Ou seja, seria preciso investir a nível de infraestrutura e alterar os dispositivos próprios para interagir com o utilizador.

As tecnologias mais presentes e difundidas, e por consequente as mais disponíveis, para utilizar são o GPS[EM99] e o A-GPS, o Wi-Fi [GRH⁺06] e o BT [Sal04][YR02]. Tirando o GPS note-se que as restantes opções usam - ou podem usar - comunicação via ondas de rádio na banda dos 2.4GHz. A principal implicação deste modo de comunicação sem fios é a tolerância à falta de linha direta de vista entre emissor e recetor, característica que pode justificar a sua larga adoção principalmente para transferência de conteúdos. Desenvolvendo esta característica podemos dizer que a banda utilizada apresenta um ótimo rácio entre penetração de estruturas em relação à distância útil de comunicação. E isto requerendo uma baixa potência de energia, que poderá ser ajustada consoante a necessidade de maior tempo de vida da bateria ou maior alcance de comunicação[Bal07]. Consideremos para já um dispositivo BT: existem atualmente 3 classes energéticas denominadas[Blu99] *class 1, 2, 3*, cujos alcances são 100, 10 e 1 metros e potência de transmissão 100mW (20dBm) 2.5mW (4dBm) e 1mW (0dBm), respetivamente. Tendo em conta o caso de uso dado no paragrafo anterior, este tipo de tecnologias permitirá que um dispositivo móvel proporcione uma solução de orientação dentro de uma área contida - coberta pela abrangência das antenas - capaz de caber na mão do utilizador, processar os dados de localização e sugerir vários caminhos, com um desgaste de energia suficientemente baixo para que a atual tecnologia de baterias o mantenha ativo por várias horas[Bal07][JP92]. Existe, no entanto, um obstáculo a esta alternativa: dentro de um edifício o GPS e o A-GPS não funcionarão e o Wi-Fi BT não foram idealizados para este tipo de utilizações. Enquanto que não existe, até à data de escrita desta tese, modo prático de solucionar este problema para as duas primeiras tecnologias, as duas restantes não devolvem posições suficientemente precisas caso a observação das ondas rádio propagadas não seja devidamente armazenado e processado.

Ao abordarmos a fraca precisão de estimativas para a localização de dispositivos que usem ativamente ou BT ou WiFi, devemos dizer que, por vezes, isto se deve à sensibilidade que as ondas de rádio na frequência dos 2.4GHz apresentam perante[MT05][FMKK09]:

i) obstáculos constituído primariamente por água; ii) de comprimento maior que a largura da onda (período); iii) estruturas metálicas; iv) dispositivos eletrónicos emissores de micro-ondas;

Nas próximas sub-secções iremos falar mais pormenorizadamente sobre cada uma das tecnologias que iram ser utilizadas pelo nosso sistema de localização.

2.2.1 Bluetooth

O Bluetooth é um protocolo para comunicação sem fios idealizado para pequenas redes locais ou pessoais[Gar03]. Por oferecer baixo consumo energético tem o alcance de comunicação reduzido[Bal07]. É um protocolo aberto¹ cujo uso atingiu grande parte dos dispositivos móveis em todo o mundo. Encontra-se difundido em telemóveis, auriculares, computadores pessoais e periféricos, veículos, relógios, etc[Haa03].

Este protocolo permite estabelecer vários tipos de perfis para a comunicação. Cada perfil implementa um conjunto de regras de transmissão de dados que facilitam a implementação de casos de utilização. Na tabela ?? apresentamos um apanhado dos perfis disponíveis:

¹Sendo um protocolo aberto significa que a propriedade intelectual associada está disponível publicamente e a sua implementação é livre de custos

Perfil	Descrição
Generic Access Profile (GAP)	Define os requisitos genéricos para detectar e estabelecer uma conexão com um dispositivo Bluetooth. Todos os perfis acentam no GAP. Pertence ao núcleo do protocolo e como tal está disponível por defeito
Generic Object Exchange Profile (GOEP)	Define os requisitos para os modelos de uso do OBEX ² . Pertence ao núcleo do protocolo e como tal está disponível por defeito
Serial Port Profile (SPP)	Define os requisitos necessários para portas de série implementadas sobre a camada RFCOMM. Cada porta de série apenas liga dois dispositivos BT entre si. Pertence ao núcleo do protocolo e como tal está disponível por defeito.
Human Interface Device (HID)	Define a comunicação sem fios entre dispositivos HID ³ sobre uma conexão BT.
Headset Profile (HP)	Define os procedimentos que suportam a interoperabilidade entre um auricular e um dispositivo móvel. Transporta comandos AT e depende do SPP.
Object Push Profile (OPP)	Define as trocas de pequenos objetos de dados, tais como <i>vCards</i> , entre dispositivos BT utilizando o perfil OBEX.
Hands-Free Profile (HFP)	Define como utilizar um dispositivo móvel em conjunto com um dispositivo áudio (<i>input</i> e <i>output</i>). Depende do SPP.
File Transfer Profile (FTP)	Define como transferir e manipular objetos de dados no próprio ou com outro dispositivo BT. Depende do GOEP.

Tabela 2.1: Protocolo Bluetooth

Cada perfil pretende facilitar a adoção do protocolo BT consoante a informação a trocar entre os dispositivos. De seguida apresentamos uma descrição de cada perfil.

GAP Define os procedimentos genéricos utilizados para descobrir os dispositivos BT. Define também mecanismos de gestão dos canais de comunicação aquando de uma nova ligação.

GOEP Define os requisitos necessários para a troca de objetos. Descreve por exemplo, modelos de uso para sincronização, transferência de ficheiros ou *push* de objetos.

SPP Define os procedimentos para emular uma porta de serie para que dois dispositivos BT comuniquem entre si através desta.

HID Implementa protocolos, procedimentos e funcionalidades que serão utilizadas por periféricos BT que comuniquem através deste perfil. Na sua essência este perfil emula o funcionamento de portas HID presentes originalmente em periféricos USB.

HP Implementa métodos específicos para serem utilizados com auriculares BT.

OPP Define os requisitos para os protocolos e para os procedimentos que são utilizados pelas aplicações que suportam *push* de objetos.

HFP Define os protocolos e os procedimentos que podem ser utilizados durante o uso do modelo “Operação de um telefone via Dispositivo *In-Car*”.

FTP Define os requisitos para os protocolos e para os procedimentos a utilizar aquando da utilização do modelo de uso “Transferência de Ficheiros”. (FTP)

Quanto à tipologia de rede, esta pode ser ponto-a-ponto ou ponto-a-multiponto tal como na figura 2.1 em a) e b), formando uma *piconet*. Cada *piconet* é constituída por um *Master* e vários *Slaves*. Um dispositivo Bluetooth ao agir como *Master* serve como referência de sincronização com o(s) restante(s) dispositivo(s) para estabelecer um sistema de comunicação *Time Division Multiple Access* (TDMA)⁴, e o relógio interno do *Master* (em conjunto com os últimos 28 bits do seu endereço de rede) servirá como referência para sincronização do relógio interno de cada *Slave*. Deste modo um ou mais dispositivos BT na mesma *piconet* partilham um canal físico para comunicação, falando entre si cada um a seu devido tempo.

⁴*Time Division Multiple Access* é um método de comunicação que partilha a mesma frequência de relógio entre vários dispositivos criando canais de comunicação, divididos entre si pelos vários dispositivos através do tempo, para que possam comunicar nessa frequência.

No máximo sete dispositivos Bluetooth podem comunicar ativamente entre si, dentro da mesma *piconet*. É possível que mais dispositivos pertençam à mesma rede, no entanto, para além do limite acima referido, os restantes deveram permanecer no estado de *parked*. Assim que necessário puderão comunicar com os restantes dispositivos da rede sem ser necessário recorrer ao procedimento de estabelecer uma ligação a uma rede. Cabe ao *Master* gerir tanto os dispositivos ativos como os *parked*. Se necessário será possível estabelecer conexões inter-rede criando uma *scatternet* (2.1 em c)). Para isso um dispositivo (quer *Master*, quer *Slave*) deverá criar uma conexão com um outro numa rede adjacente. A referência de sincronização não poderá ser a mesma entre duas *piconets*.

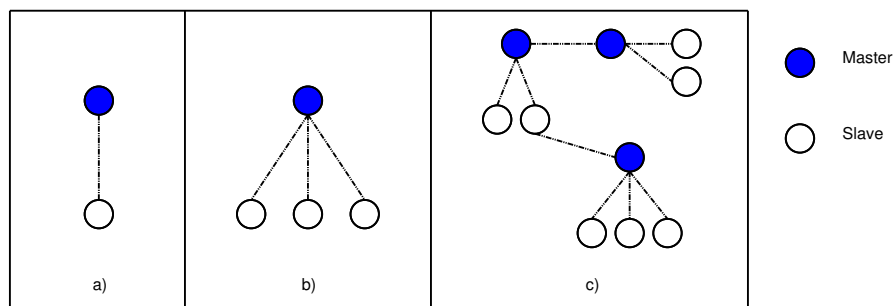


Figura 2.1: Tipologias de rede Bluetooth

2.2.2 Comunicação Entre Dispositivos

Para que dois dispositivos possam interagir entre si será necessário, em primeiro lugar, que se descubram. Partindo do princípio de que ambos se encontram no alcance das suas antenas de rádio, um dos dispositivos deverá dar início ao modo chamado de *inquiry scan*. Neste modo um dispositivo BT inicia a procura por outros dispositivos utilizando códigos de *inquiry* definidos na especificação permitindo que apenas sejam encontrados tipos de dispositivos segundo a sua classe (PDAs, periféricos, auriculares, etc.). Ao iniciar esta pesquisa o dispositivo BT gera uma sequência de *hopping*⁵ derivada do seu próprio relógio interno e do código de *inquiry* escolhido. Serão criados 32 canais de comunicação possíveis (dos 79 disponíveis do espectro Bluetooth). De seguida serão transmitidas mensagens de inquirição em cada um destes canais.

⁵*Frequency Hopping* é um método para partilha de largura de banda de um dado espectro restrito para comunicação que consiste em alternar a frequência do sinal de rádio sequencialmente. O emissor e o recetor sabem de antemão a sequência ou intervalo de alternância. Em cada intervalo existe uma oportunidade de transferência de dados a que se chama canal de comunicação.

Do outro lado, dispositivos BT que estejam visíveis entrarão periodicamente no modo de *inquiry scan* à escuta por sinais de inquirição. O intervalo de escuta é definido pela sua própria classe em conjunto com o seu relógio local. Aquando de um encontro de mensagens de *inquiry* com *inquiry scan* o dispositivo que recebe a mensagem de inquirição responde com uma mensagem de *inquiry response* que inclui o seu endereço de rede e o seu relógio. Segundo a especificação serão necessários 10,24 segundos para detetar todos os dispositivos BT na vizinhança. É no entanto possível terminar mais cedo o processo de procura definindo limites para o número de respostas, para o número de dispositivos a encontrar ou ainda por um intervalo inferior aos 10.24 segundos. Após este processo de pesquisa, o utilizador terá a oportunidade de abrir uma conexão com os dispositivos encontrados. Primeiro entrando em modo de *paging* para acordar numa frequência de *hopping* fixa para comunicar e depois abrindo efetivamente a conexão.

Para o objetivo do nosso trabalho iremos estudar o uso da tecnologia Bluetooth para localizar um dispositivo que esteja visível para outros dispositivos Bluetooth através da força do sinal recebido, ou RSSI⁶. Este valor está presente sempre que se estabelece uma conexão⁷ entre dois dispositivos BT ou, a partir da versão 2.1 da especificação BT, na resposta a uma mensagem de inquirição. Utilizaremos apenas a inquirição por dispositivos durante o nosso trabalho. Este procedimento tem como única limitação (para o nosso trabalho) o facto de ser necessário o utilizador colocar o dispositivo BT como visível, de modo a que este responda ao pedido de pesquisa por parte dos dispositivos vizinhos. Se tivéssemos optado por abrir uma conexão o utilizador teria de introduzir uma palavra segredo de cada vez que: 1) fosse necessário recorrer a localização *indoor*; 2) perdesse a conexão com o *Master*. Isto acarretaria um custo demasiado penoso para a usabilidade do sistema e implicaria ainda uma aplicação que fizesse a ponte entre o sistema de deteção e o dispositivo móvel para que o utilizador tomasse conta do estado da ligação e introduzisse a palavra segredo cada vez que passasse para uma nova estação de deteção. Embora abrir uma conexão proteja o utilizador do ponto de vista de segurança - pois requer o passo de autenticação, ou a referida palavra segredo - ao utilizarmos os dados presentes no procedimento de inquirição cremos que teremos um equilíbrio suficientemente bom entre privacidade e segurança já que o utilizador tem a opção de colocar o seu dispositivo oculto para pesquisas, mantendo as conexões atuais que tiver ativas e apenas o tornar visível

⁶RSSI é uma medida utilizada genericamente em comunicações sem fios para medir a potência de sinal recebido [OP05]

⁷Existe outro valor que poderia ser utilizado, de nome *Link Quality*. No entanto viemos a descobrir que é um campo que é populado consoante critérios subjectivos ao fabricante e por isso não determinístico.

sempre que quiser abrir novas ligações.

2.2.3 Wi-Fi

O Wi-Fi é também um protocolo de comunicação sem fios que está idealizado para ter um alcance bastante superior do que o Bluetooth[Dha07]. Como está este protocolo requer mais energia por parte do sistema. Enquanto que o BT tenta minimizar o consumo de energia, o Wi-Fi foi pensado para atingir maiores taxas de transferência e maiores alcances[Sal04]. A sua finalidade é ultimamente complementar a rede *ethernet* substituindo-a onde esta não chega. Para isso o protocolo mantém inalteradas as camadas mais altas da especificação *ethernet* e substitui as camadas mais baixas de transmissão de dados. O Wi-Fi não implementa novas funcionalidades na rede do mesmo sentido que o BT faz com os variados perfis que suporta. Tem, no entanto, vindo a evoluir na velocidade e alcance através de revisões à especificação, mantendo a compatibilidade com as versões anteriores (excepto com a 802.11a), conforme a tabela 4.2:

Especificação	Frequência	Largura de banda
802.11a	5 GHz	54 Mbps
802.11b	2.4 GHz	11 Mbps
802.11g	2.4 GHz	54 Mbps
802.11n	2.4 GHz, 5 GHz, 2.4 ou 5 GHz (seleccionável), ou 2.4 and 5 GHz (concorrente)	450 Mbps

Tabela 2.2: Versões da especificação IEEE 802.11(a/b/g/n)

Uma rede Wi-Fi é composta por vários dispositivos capazes de receber e emitir ondas rádio na frequência dos 2.4Ghz ou dos 5Ghz, traduzir esses sinais em dados por pacotes, e entrega-los ao sistema operativo do dispositivo *host* tal como um dispositivo de rede *ethernet*. O funcionamento de um dispositivo que permita o acesso a uma rede *wireless* é especificado pela norma IEEE 802.11[OP05]. Nesta especificação podemos encontrar dois tipos de equipamentos para criar e usar uma rede Wi-Fi:

- **Estação Wi-Fi** é um dispositivo com a capacidade de aceder a uma rede Wi-Fi
- **Ponto Acesso (PA)** permite que estações Wi-Fi acessem a uma rede com

fios *ethernet*, à qual o AP está diretamente ligado. Um *PA* pode servir várias redes distintas.

O mesmo *standard* define dois modos que permitem funcionamento em rede:

- **Infra-estrutura** no qual as estações se conectam a um *Access Point*
- **Ad-Hoc** no qual todas as estações Wi-Fi se conetam entre si

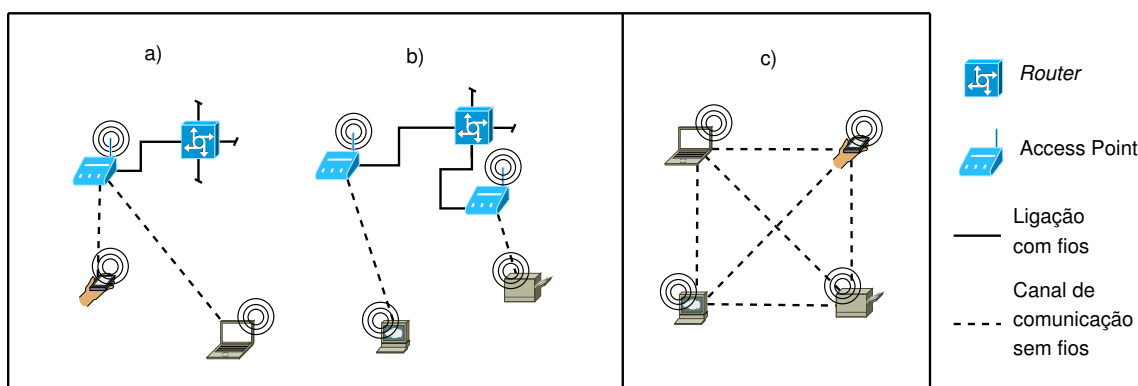


Figura 2.2: Tipologias de rede Wi-Fi

Para que vários dispositivos *wireless* comuniquem é então necessário que um deles seja *Ponto de Acesso* e os restantes em *client mode*, ou então que todos funcionem em modo *Ad-Hoc*. Para que uma rede Wi-Fi esteja ligada ao “exterior”, ou duas redes Wi-Fi ligadas entre si, é necessária a existência de um dispositivo roteador que faça a ponte com a rede exterior. Só por si, uma estação ou um *Ponto de Acesso* não é capaz de interagir com outras redes. As tipologias aqui mencionadas estão visualmente representados na figura 2.2, onde se pode ver uma rede com um único *PA* na alínea a), dois *PAs* na mesma rede ligando vários dispositivos Wi-Fi na alínea b) e uma rede *Ad-Hoc* isolada na alínea c)

Resumindo, cada dispositivo Wi-Fi terá que ter configurados os seguintes parâmetros básicos:

- **Modo de operação** *Ad-Hoc* ou *Infra-estrutura*. Qualquer dispositivo Wi-Fi pode criar ou juntar-se a redes *Ad-Hoc*. Apenas um *Access Point* ou um *router* podem criar uma rede em modo *Infra-estrutura*

- **Canal de Operação** a extensão 802.11 a usar (a, b, g, n. . .) e o país. Dados estes parâmetros os canais a disponíveis serão automaticamente configurados pelo *firmware* do dispositivo Wi-Fi

- **Nome da Rede** é o *Service Set Identifier* (SSID), essencialmente o nome da rede a ser transmitido. O tamanho máximo é de 32 *bytes*. A transmissão poderá ser omitida como medida adicional de segurança

Para que uma estação descubra outras estações na sua proximidade terá que recorrer a um aplicativo de pesquisa, normalmente fornecido pelo fabricante. O procedimento de associação a uma rede é composto por três passos:

- a descoberta de redes disponíveis
- a autenticação na rede escolhida, caso este se encontre protegida por palavra passe
- a associação a essa rede

A descoberta de redes pode ser efetuada em todos os canais Wi-Fi disponíveis, ou apenas no canal pretendido. Analogamente, pode ser feita uma pesquisa por determinado SSID ou por todos os disponíveis. Os *Pontos de Acesso* transmitem pacotes *MAC*⁸ chamados de *beacons* com o propósito de anunciar a rede para as estações Wi-Fi interessadas. Estes *beacons* podem ser dirigidos a estações que estejam a fazer uma listagem activa ou passiva da rede. A diferença entre estes dois modos é a seguinte:

- **Listagem Activa** a estação alterna nos os canais definidos na especificação IEEE 802.11 para o seu país em busca de *beacons* de resposta a listagem ou periódicos. Estes dividem-se ainda em dois modos:

- **específico** é pedida uma listagem por SSIDs específico à qual cada *PA* representando esse SSID responde
- **broadcast** é pedida uma listagem geral (sem SSID) à qual cada *Ponto de Acesso* responde com uma resposta a pedido de listagem contendo todos os SSIDs que representa

⁸pacotes *MAC* ou pacotes da camada de *Medium Access Control* da especificação IEEE 802.3 (*Ethernet*). Esta situa-se logo após a camada física especificada pela IEEE 802.11

- **Listagem passiva** é efetuada uma “escuta” por *beacons* alternando pelos canais definidos na especificação IEEE 802.11 para o seu país na expectativa de receber um *beacon* periódico (voluntário). Por defeito estes são emitidos a cada 100 ms pelos *APs*

Durante uma listagem por *APs* uma estação está impedida de comunicar com uma rede a que esteja já associada. Isto pode prejudicar a ligação Wi-Fi consideravelmente se pensarmos que o dispositivo pode estar em movimento. Para minimizar o impacto desta ação um dispositivo Wi-Fi pode seguir várias estratégias normalmente implementadas pelo fabricante. Daremos como exemplo duas:

i) **Comportamento em *background*** na eventualidade de ser necessário mudar de *PA* para melhor a receção de sinal (efetuar *roaming* para outro *PA*) o dispositivo pode fazer uma listagem em *background* sempre que não houverem transmissões de dados a decorrer, adquirindo assim conhecimento do ambiente *wireless* em seu redor e otimizar o processo de *roaming*;

ii) **Comportamento do processo de *roaming*** cada fabricante pode criar a sua estratégia de *roaming* alterando o tempo de escuta e os canais que serão listados de modo a otimizar o tempo de demora do processo;

A comunicação dá-se efetuando trocas de pacotes entre o *Ponto de Acesso* e as estações Wi-Fi no caso de a rede ser Infra-estrutura, ou diretamente com a estação destino do pacote no caso de se tratar de uma rede *Ad-Hoc*. Cada pacote *wireless* quanto à camada *MAC* é constituído por 4 endereços (só 3 serão usados a maior parte das vezes) que implementam a estratégia de comunicação:

BSSID: identifica o *PA* de uma rede em Infra-estrutura. Caso seja uma rede *Ad-Hoc* é um número gerado aleatoriamente

Endereço de Destino (ED): identifica o recipiente final do pacote de dados

Endereço de Origem (EO): identifica o emissor do pacote de dados

Endereço do Recipiente (ER): identifica o recipiente imediato (o *PA*)

Endereço de Transmissão (ET): identifica o endereço do *PA* que transmitiu o pacote de dados

O número de endereços de rede contidos está relacionado com o tipo de pacote a ser transmitido. Este pode conter dados específicos de um dispositivo para outro, que no caso de ser numa rede em Infra-estrutura terá de passar obrigatoriamente pelo *PA*, ir diretamente de um dispositivo para outro no caso de ser uma rede *Ad-Hoc*, ser um pacote contendo um *beacon* com o SSID de um *PA* ou ser um pacote de rede com dados para todos os dispositivos *wireless* da rede. De seguida apresenta-se uma tabela que resume os endereçamentos consoante o propósito do pacote:

<i>PA</i> destino	<i>PA</i> origem	Endereço 1	Endereço 2	Endereço 3	Endereço 4
×	×	ED	EO	BSSID	n/a
×	✓	ED	BSSID	EO	n/a
✓	×	BSSID	EO	ED	n/a
✓	✓	ER	ET	ED	EO

Tabela 2.3: Resumo dos campos de endereçamento de rede da camada *MAC*

Quanto às colunas de endereço na tabela podemos reparar que

Endereço 1: contém sempre o endereço do recepiente, que será sempre a estação Wi-Fi receptora do pacote

Endereço 2: contém sempre o endereço da estação Wi-Fi que está a transmitir fisicamente o pacote de dados

Endereço 3: será ou a origem do pacote ou o destinatário pretendido pela origem

Endereço 4: é o destino final para um pacote que é transmitido e recebido numa rede

Quanto às linhas podemos dizer que na primeira o pacote viaja entre um *Access Point* e as estações a ele associadas; segunda e terceira linha definem os endereços necessários para pacotes que são passados entre um *Access Point* e um roteador; na quarta linha está o caso especial em que o pacote viaja entre dois roteadores *wireless*, criando por isso a necessidade de mais dois endereços adicionais para o *PA* origem e estação destino.

2.2.4 GSM

O GSM é um standard internacional para comunicação entre dispositivos móveis[MP92]. Foi criado em 1984, em França, para ser utilizado por toda a Europa como sistema de comunicação telefónico sem fios. Apesar da sua adoção a níveis globais ter acontecido em meados dos anos 90 [CHRW08] o GSM é utilizado ainda hoje. Mais concretamente é um protocolo de comunicação digital, sem fios, e que de apesar de na altura do seu aparecimento oferecer uma largura de banda superior em relação à linha telefónica terrestre, hoje em dia já começa a tornar-se obsoleta em comparação com protocolos mais recentes. Apesar de novos protocolos serem mais rápidos, mais eficientes permitirem uma maior oferta de serviços, o GSM ainda perdura devido à adoção em grande escala que teve. Ao longo dos anos sofreu várias atualizações para aumentar a taxa de transferência de dados e criar novos meios de utilização, principalmente o acesso à Internet.

O GSM é constituído essencialmente por uma rede de células onde os dispositivos móveis, chamados de subscritores móveis (*Mobile Subscriber* (MS)), podem comunicar entre si e com a rede telefónica terrestre. Para além da comunicação de voz o GSM possibilita ainda a troca de mensagens curtas, caixa de correio por voz, *fax*, reencaminhamento de chamadas, identificação do número de origem, entre outros. A banda utilizada para comunicação varia consoante a região do planeta, sendo as mais comuns as frequências 450MHz, 850MHz, 900MHz, 1800MHz e 1900MHz. A troca de dados entre as células e os terminais é efetuada através de canais de comunicação separados, isto é, existe um intervalo de frequências para o envio de dados entre o terminal móvel e a rede, e uma outra para o envio de dados da rede para o terminal. Ao primeiro intervalo chama-se frequências de *uplink* e ao segundo frequências de *downlink*. Para além disso a especificação do protocolo não permite a troca em simultâneo de dados, o que gera um desfasamento temporal entre o envio e a receção. De modo a utilizar o mesmo rádio para enviar e receber, é utilizado um sistema que opera sobre a antena que a sintoniza na frequência de emissão e receção alternadamente; e apesar do desfasamento temporal entre transmissão e receção o GSM comunica em *full duplex*⁹ reservando intervalos de tempo distintos para o terminal móvel na célula onde este opera. Para cada banda de operação estão definidos vários canais de comunicação cujas direções estão separadas por um deslocamento para garantir que não haja sobreposição de frequências entre os dois sentidos de comunicação. A

⁹O termo *full duplex* é usado para descrever um sistema de comunicação ponto-a-ponto que transmite e recebe dados no mesmo instante de tempo.

quantidade de canais existentes depende das características específicas de cada banda. A cada canal é atribuído um *Absolute Radio-Frequency Channel Number* (ARFCN) representando o número do canal dentro da banda em utilização. Para se saber quais as frequências de operação de cada canal podemos seguir a seguinte fórmula:

GSM-900:

$$\text{Uplink} = 890.0 + (\text{ARFCN} \times 0.2)$$

$$\text{Downlink} = \text{Uplink} + 45.0 \text{ (deslocamento)}$$

GSM-1800:

$$\text{Uplink} = 1710.0 + ((\text{ARFCN} - 512) \times 0.2)$$

$$\text{Downlink} = \text{Uplink} + 80.0 \text{ (deslocamento)}$$

Se, por exemplo, quisermos saber quais as frequências dos canais de *uplink* e de *downlink* da banda GSM-900 cujo ARFCN é o 77, utilizando a fórmula correspondente supra mencionada chegamos às seguintes frequências:

$$\text{Uplink} = 890.0 + (77 \times 0.2)$$

$$\text{Uplink} = 890.0 + (15.4)$$

$$\text{Uplink} = 905.4 \text{ MHz}$$

$$\text{Downlink} = \text{Uplink} + \text{deslocamento}$$

$$\text{Downlink} = 905.4 \text{ MHz} + 45.0$$

$$\text{Downlink} = 950.4 \text{ MHz}$$

Posto isto, sabemos que na banda do GSM-900 o canal 77 envia dados do terminal móvel para a rede nos 905.4 MHz e recebe dados da rede nos 950.4 MHz.

Um terminal é composto por dois componentes básicos: o equipamento móvel e um cartão *Subscriber Identity Module* (SIM). Cada um destes componentes pode ser identificado respetivamente, através do *International Mobile Equipment Identity* (IMEI) e do *International Mobile Subscriber Identity* (IMSI). O terminal móvel, só por si, não é capaz de operar numa rede GSM apesar de poder permitir ao utilizador aceder a outras funcionalidades (fora do serviço de rede) que suporte. De igual modo, um cartão SIM é inútil sem um terminal. Em conjunto, estes dois componentes interagem

para que o utilizador possa aceder aos serviços que subscreveu na rede e vice-versa (para que a rede possa aceder ao utilizador através do terminal móvel).

A tipologia de rede é composta por um conjunto de células que deveram estar distribuídas pela área a cobrir e respetiva arquitetura de componentes para controlar o seu funcionamento - entre dispositivos moveis, células, intra-rede e inter-redes móveis e fixas. As células são responsáveis por fornecer cobertura de sinal para os terminais móveis através de sinais de rádio. Uma célula (conjunto antena e *Base Transceiver Station* (BTS)) é o último nó numa rede GSM. Está responsável por lidar com todos os processos de comunicação entre um terminal móvel e a restante rede. Enquanto que a antena transmite e recebe dados através de sinais rádio a BTS está encarregue do tráfego da rede: receber e enviar não só os dados de e para cada utilizador (voz, sms, etc.) como os comandos de controlo necessários para a gestão da área que abrange. Para além disto, está também encarregue de efetuar a codificação e decodificação dos dados e a respetiva encriptação e desencriptação. Na figura 2.3 será feita uma ilustração dos extremos da tipologia de rede, que contém os componentes que serão estudados para o trabalho desta tese.

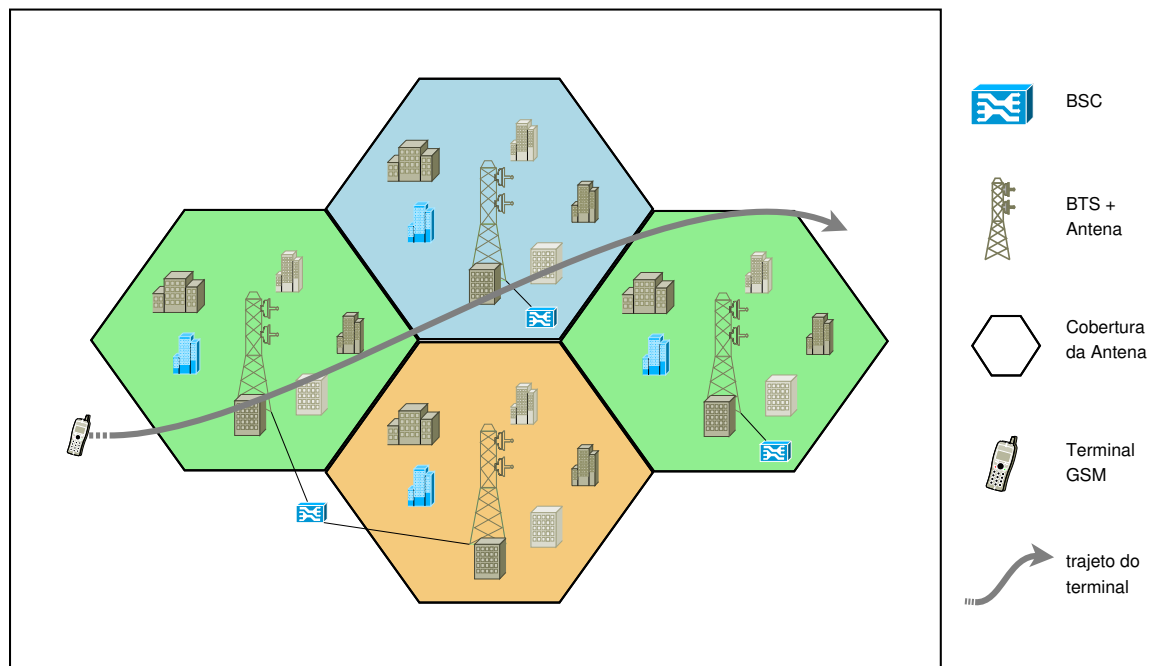


Figura 2.3: Tipologia de uma rede *Global System for Mobile Communications* ou *Groupe Spécial Mobile*

As BSC presentes na ilustração estão encarregues de administrar os recursos de rádio

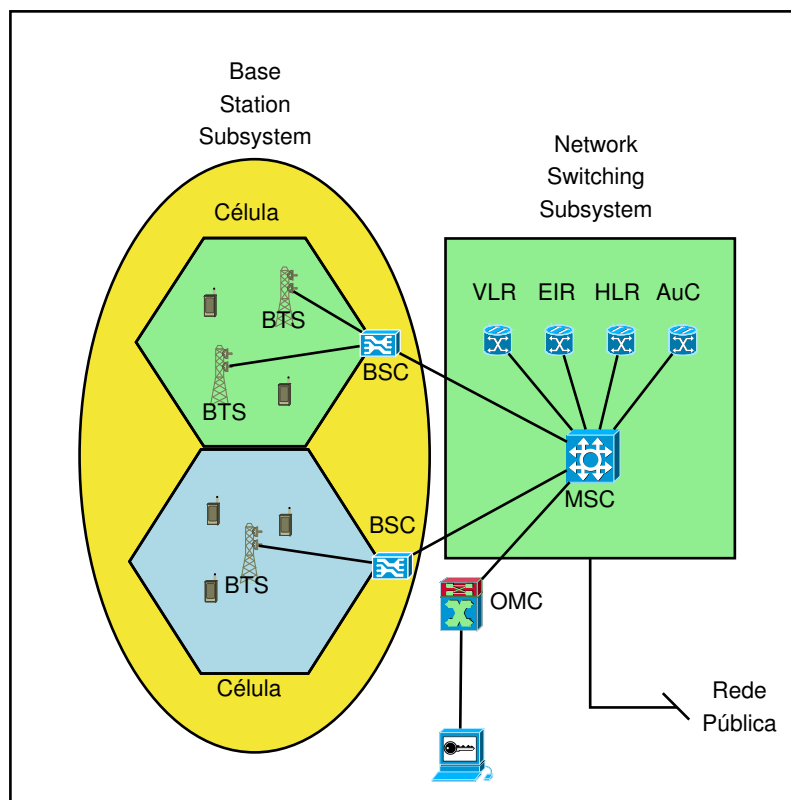
das BTS. Mais concretamente agem como concentradores de rede, reunindo não só os dados, mas também o estado das ligações para com os terminais moveis. Idealmente as BSC controlam um numero considerável de BTS, porém hoje em dia, com o avançar dos componentes eletrónicos e do software de gestão de redes, as BSC começam a desaparecer em prol de BTS “inteligentes”.

O duo BTS e *Base Station Controller* (BSC) chama-se *Base Station Subsystem*. A *Base Station Subsystem* (BSS) é o ponto de ligação entre os MSs e a restante rede. Os dados relativos à comunicação e gestão de rede são enviados a partir de cada BSS para o próximo componente na tipologia: as *Mobile Switching Centers* (MSCs). Por sua vez cada MSC está encarregue dos encaminhamentos de rede, das inicializações necessárias para efetuar uma chamada de voz e de funções básicas de *switching*¹⁰. Está também encarregue de lidar com os procedimentos necessários aquando de uma mudança entre duas BTSs durante uma chamada de voz (exemplificado na figura 2.3 em “trajeto do terminal”). É também capaz de coordenar a mobilidade inter-MSCs. Toda a comunicação inter-rede pode ser efetuada unicamente pelas MSCs. Caso se dê uma tentativa de contacto intra-rede a MSC responsável pela gestão do terminal móvel tratará de anunciar o seu paradeiro na rede e efetuar as ações necessárias para que a comunicação suceda, incluindo com o exterior da sua rede. Existem mais componentes que auxiliam a MSC nas suas funções. Tratam-se de componentes específicos para cada uma das funcionalidades que providenciam. Será apresentada uma breve descrição de cada:

¹⁰Um *switch* de rede é um componente encarregue de transportar cada uma das mensagens de rede e a transportar ao seu destinatário.

Componente	Descrição
HLR (<i>Home Location Register</i>)	Age como uma base de dados para a gestão de MSs. Guarda o IMSI, o identificador para as redes exteriores e o endereço atual do MS no VLR. Guarda ainda a lista de serviços a disponibilizar a cada MS
VLR (<i>Visitor Location Register</i>)	Detém a localização atual do acms e informação administrativa proveniente da HLR, necessários para o controlo de chamadas de voz e para o aprovisionamento dos serviços subscritos. Habitualmente existe uma VLR por MSC, integrada no seu <i>hardware</i>
AuC (<i>Authentication Center</i>)	É uma base de dados protegida que armazena uma cópia da chave secreta incorporada também no SIM de cada MS. A chave é utilizada para efeitos de autenticação e encriptação. O principal objetivo deste componente é prevenir fraudes de utilização.
EIR (<i>Equipment Identity Register</i>)	A EIR é uma base de dados que contém informação para validar os equipamentos moveis dentro da rede perante o respetivo IMEI de cada terminal móvel.
OMC (<i>Operation and Maintenance Center</i>)	O OMC é um sistema de gestão que supervisiona todos os componentes da rede GSM. É uma interface para com o operador de rede, assistindo-o no bom funcionamento da rede. Contém a redundância de <i>hardware</i> e mecanismos de deteção de erros para prevenir que a rede entre em colapso. É responsável por manter a MSC, a BSC e consecutivamente, as várias BTS.

Podemos então fazer uma ilustração global da tipologia de uma rede GSM:



O protocolo GSM define[WJH97] o modo como um terminal móvel (MS), em conjunto com o cartão SIM, deve aceder à rede: O primeiro passo será pedir ao utilizador o código *Personal Identification Number* (PIN) que permite aceder ao conteúdo de autenticação (entre outros) na rede. Uma vez efetuado o passo um, o terminal móvel poderá proceder ao acesso a uma tabela de redes que reconheçam aquele SIM e continuar para o passo dois: a autenticação dentro da rede. Este segundo passo é composto por um pedido de acesso à rede GSM (entre as disponíveis na área em que se encontra). A MSC responsável pela BTS que recebeu esse pedido fica encarregue de enviar o mesmo para uma HLR capaz de identificar o terminal móvel. O pedido de autenticação que é recebido pela HLR contém o IMSI do novo terminal a autenticar. A HLR deve, por sua vez, verificar a validade desse IMSI e se de facto o mesmo pertence à rede. Para isso consulta a sua base de dados e, em caso de sucesso de validação, devolve o pedido de autenticação validado ao AuC. O AuC por sua vez gera um conjunto de elementos necessários para os processos de encriptação subsequentes chamados *Triplets* e envia-os para a HLR. Esta, por sua vez, envia-os para a MSC. O terceiro passo é composto pelo envio de dados de autenticação para o MS para que este se autentique. Caso a resposta ao pedido de autenticação por parte da MSC com os dados gerados pelo AuC seja fidedigna o MS é considerado válido dentro da rede.

Existem eventos que podem obrigar o terminal móvel a autenticar-se novamente na rede. Cabe ao operador de rede optar por configurar quando ou quais os eventos. O operador escolherá as configurações que achar necessárias consoante o nível de segurança que pretender para as operações na rede, tendo em conta a latência e o volume de tráfego adicionais que estes processos possam implicar. Os eventos que poderão despoletar uma nova autenticação podem ser desde mudanças de localização (de célula), inícios ou finalizações de chamadas telefónicas, *smss*, etc. O operador deverá pedir nova autenticação sempre que um MS mude para uma nova célula (considerando o trajeto da seta a cinzento da figura 2.3 o MS efetua três mudanças de célula).

2.2.5 Outras Tecnologias para Localização *Indoor*

Embora o objetivo deste trabalho seja a localização *indoor tagless*, iremos fazer um breve apanhado de tecnologias que embora permitam localizar um dispositivo móvel dentro de um edifício requerem um emissor/recetor específico para o efeito.

2.2.5.1 Infravermelhos

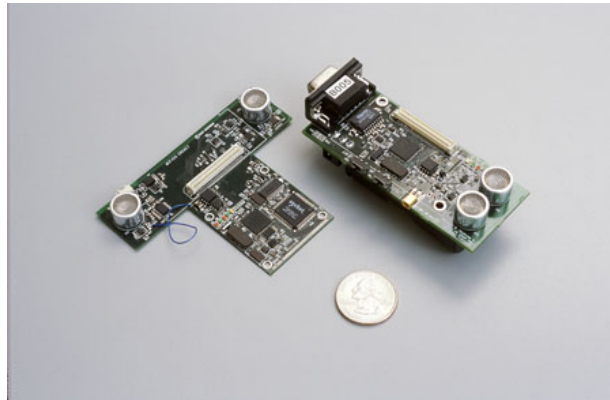
A tecnologia de comunicação por infravermelhos é das mais comuns hoje em dia, devido à quantidade de dispositivos que a adotaram (TVs, impressoras, telemóveis, PDAs, etc.)[KB97]. Um sistema que utilize esta tecnologia para localização terá forçosamente que garantir linha de vista entre emissor e transmissor para que o sistema comunique. Isto porque os infravermelhos não atravessam paredes, sendo técnica comum para este problema a instalação de pontos de acesso que liguem as várias divisões de um mesmo local. Para além deste obstáculo[KB97], que muitas vezes é difícil de cumprir, esta tecnologia está sujeita a sofrer interferências por fontes de luz. Embora os emissores e recetores possam ser relativamente baratos, o sistema pode ser dispendioso de instalar pois será necessário ter dispositivos de comunicação por infravermelhos em pontos estratégicos. Note-se ainda que será um investimento, que à partida servirá o único propósito de localização. Quanto à autonomia, sistemas baseados neste tipo de comunicação, como por exemplo o *Active Badge System*[WHFG92], com uma frequência de transmissão de *beacons* de localização espaçados entre si por 15 segundos durará cerca de meio ano. Quanto à ergonomia, os sensores portáteis, do mesmo sistema, são pequenos e leves, logo fáceis de transportar.

2.2.5.2 Ultrassons

Obter medições de tempo de chegada da mensagem por ultrassons[WJH97] é outro método de localização *indoor*. Normalmente, são instalados sensores recetores de ultrassons nas paredes e/ou tetos que detetam o tempo de chegada de ultrassons emitidos pelas etiquetas a localizar. A estratégia comum é utilizar a técnica de multilateração, que consiste na recolha de várias medições da mesma mensagem conhecida pelo sistema e através destas estimar a posição da etiqueta. Devido às características dos ultrassons, principalmente a sensibilidade dos detetores, é possível obter uma posição com 3 dimensões ($O(x,y,z)$). Também nesta alternativa a ergonomia para o utilizador não é comprometida, pois os sensores não tem dimensões que o tornem desconfortáveis, no entanto, a instalação do sistema tem um ponto fraco em comum com a tecnologia de infravermelhos: custos elevados de instalação. É no entanto possível utilizar uma implementação deste modo de localização para transmissão de mensagens curtas. O sistema *cricket* implementa um mecanismo de sincronização do relógio das etiquetas (emissores) com os recetores (nos tetos) por rádio frequência. Esse mesmo mecanismo é também utilizado para dar conhecimento ao sistema do estado da bateria da etiqueta emissora [WJH97].

2.2.5.3 RFIDs

Os sistemas de localização baseados em *Radio Frequency Identifications*[Jue06] providenciam vantagens em relação às duas tecnologias supra mencionadas, na medida em que os dispositivos moveis a detectar podem estar situados em divisões bem mais distantes do recetor. Isto deve-se ao facto de os sistemas de infravermelhos e de ultrassons requererem linha de vista ou não propagarem através de estruturas (como as paredes, por exemplo). As implicações deste facto obrigam a colocação de dispositivos de deteção em todas as divisões de modo a garantir a cobertura destas por parte do sistema de localização.

Figura 2.4: Sensores *cricket*

2.3 Detecção e Localização de Sinais de Rádio

Serão então utilizadas para os trabalhos desta tese tecnologias de comunicação sem fios que aproveitem as infraestruturas de rede já em uso de modo a minimizar os custos de implementação do sistema. Para cumprir esse objetivo serão adotadas três estratégias de recolha de ondas rádio, mais especificamente, BT, WiFi e GSM. Esta opção prende-se com a disponibilidade atual destas tecnologias tanto nos dispositivos a localizar (nos telemóveis), como no fornecimento destas redes (embora o BT esteja em desvantagem, pois não foi idealizada para comunicações a grandes distancias, ou de grande mobilidade). Estas tecnologias permitem otimizar o número de antenas recetoras necessárias por divisão (ver exemplo de propagação de sinal da figura 2.5) e têm um gasto energético suficientemente baixo para que os dispositivos moveis funcionem durante um período aceitável.

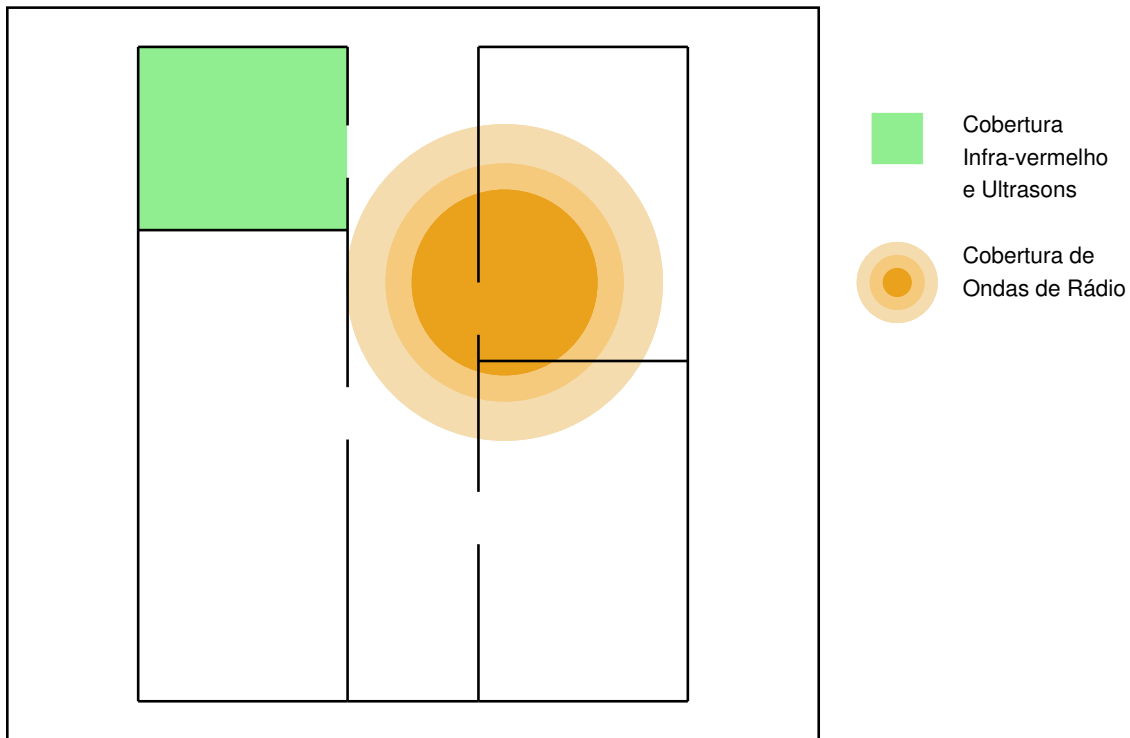


Figura 2.5: Propagação de Sinais Rádio *versus* Infravermelhos ou Ultrassons dentro de um edifício

Encontramos duas alternativas para recolher os sinais rádio de cada tecnologia: a recolha ativa e a recolha passiva. Para cada uma destas abordagens é possível idealizar dois métodos de localização: com ou sem coordenadas determinísticas. Na figura 2.6 é apresentado um esquema representativo das alternativas.

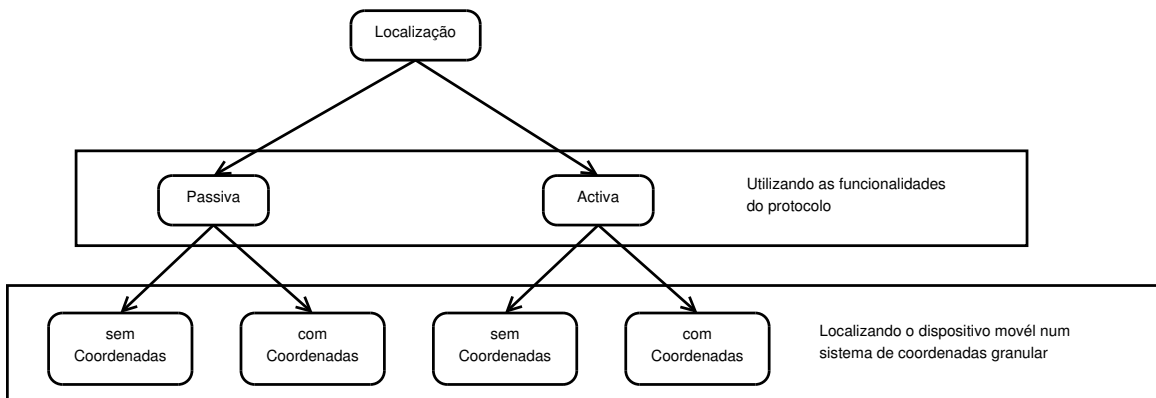


Figura 2.6: Alternativas para a localização de dispositivos *indoor*

Um sistema de localização ativo distingue-se de um sistema passivo por processar a informação trocada de acordo o protocolo de comunicação. Do ponto de vista operacional, torna a implementação do *software* mais fácil pois permite controlar a granularidade dos dados a utilizar e descartar. Por outro lado, um sistema passivo não tem conhecimento dos detalhes do dispositivo nem dos dados que são transmitidos. Apenas escuta se determinada frequência teve um aumento de potência num instante no tempo. Com isto, partindo do princípio que se sabe a frequência de transmissão, podemos observar as diferenças no espectro e deduzir, perante a localização da antena, onde se situa o dispositivo móvel a localizar. Caso se pretenda a localização concreta do terminal será necessário recorrer a uma abordagem com coordenadas. No caso da recolha ativa, tal informação pode ser facilmente adquirida através dos campos potência de sinal recebido (RSSI), qualidade do canal ou da relação sinal com o ruído do canal de comunicação. Caso se opte pela abordagem passiva, a aquisição de potência de sinal terá de ser feita recorrendo a antenas direcionais e a dedução das coordenadas terá de ser efetuada consoante o ângulo de direção na altura da captura.

2.3.1 Propagação de Sinal

Quando um sistema de comunicação rádio necessita de enviar ou receber mensagens procede à transmissão ou interceção de ondas eletromagnéticas em espaço aberto. Este processo consiste em enviar dados para o transmissor, em sinais elétricos com determinada frequência, que por sua vez os converte numa onda eletromagnética a ser transmitida pela sua antena. Esta onda é propagada através do espaço aberto, que é o meio de transmissão das mensagens. Do outro lado, o recetor interceta a onda eletromagnética também através da sua antena e converte-a de volta para um sinal elétrico na devida frequência. Este processo também é chamado de modulação e demodulação.

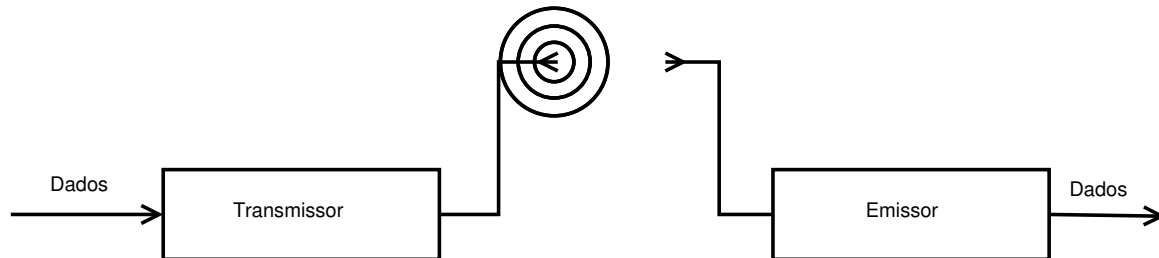


Figura 2.7: Ilustração do envio de dados por sinais rádio

Iremos de seguida abordar alguns dos termos técnicos usados em transmissão de dados por rádio frequência e respetiva definição:

Décibel (dB): é a relação matemática entre dois valores exprimida em logaritmo

Potência do Sinal Rádio: pode ser medido tanto o valor de emissão como o de receção. A relação pode ser expressa da seguinte forma:

$$P_{sinalradio} = 10 \times \log P_{otência}(mW)$$

. Por exemplo, para uma potência de emissão de 0.5 *Watts* temos:

$$0.5 \text{ Watt} = 500mW, \text{ logo}$$

$$P_{sr} = 10 \times \log 500 = 26.9dBm$$

Perda de caminho: a perda de potência de sinal que se propaga pelo espaço aberto. É exprimida em dB e depende de:

- distância entre as antenas emissora e recetora
- obstáculos existentes entre as antenas
- altura da antena

Atenuação em espaço livre: é a atenuação do sinal enquanto se propaga segundo a sua frequência. Pode ser calculada através da fórmula

$$Aten = 32.4 + 20 \times \log F + 20 \times \log R$$

onde F é a frequência do sinal em MHz e R a distância em Quilómetros.

Desvanecimento do Sinal: o desvanecimento do sinal pode ocorrer devido a vários fatores:

Multi-caminho: o sinal transmitido atinge o recetor a partir de várias direções devido a diferentes caminhos percorridos. Para cada um desses caminhos sofrerá ao longo da distância percorrida atenuações diferentes.

Fraca linha de visibilidade: uma linha de vista imaginária que conecta ambas as antenas

Condições climatéricas: Dependendo da temperatura, humidade relativa e velocidade do vento. A 150mm/hr de chuva a atenuação a 2.4GHz poderá atingir um máximo de 0.02 dB/Km. O vento pode afetar as antenas devido ao movimento que pode provocar.

Ganho: é expresso em dBi e definido através da razão entre a intensidade do sinal numa dada direção com a intensidade da radiação que seria obtida se a potência fosse pela antena fosse irradiada igualmente por todas as direções.

antena omni-direcional: é um tipo de antena que transmite e recebe igualmente por todas as direções (ver figura 2.8)

antena direcional: é um tipo de antena que envia e recebe a maior parte das ondas eletromagnéticas em uma só direção (ver figura 2.8)

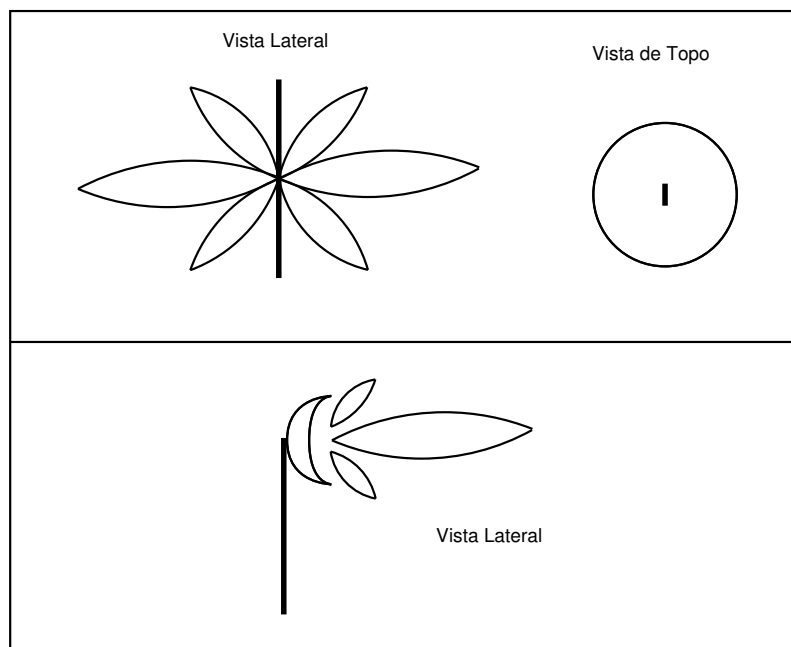


Figura 2.8: Ilustração de uma antena omni-direcional e uma antena direcional (em cima e em baixo, respetivamente)

2.3.2 *Software Defined Radio*

Software Defined Radio ou Rádio Definido por *Software* é um sistema composto por um computador e um dispositivo capaz de modular e demodular sinais elétricos[TW02]. Neste sistema os componentes de *hardware* que tipicamente compõem um rádio são

substituídos por funções matemáticas escritas em *software*. A vantagem de tal configuração é a adaptabilidade, pois permite a flexibilidade suficiente para ultrapassar as “limitações” que um rádio em *hardware* tem - tais como componentes específicos para um dado espectro - embora não garanta o mesmo tipo de precisão. Cabe também ao programador programar as implementações dos componentes outrora físicos, em código que um computador pessoal possa processar, imitando exatamente o comportamento “real”. Tendo como base vários componentes implementados é possível combinar os mesmos de modo a construir um sistema complexo de rádio.

Um dos principais obstáculos para *Software Defined Radio* (SDR) é a necessidade de poder computacional necessário para processar o *Input/Output* dos vários componentes virtuais que integrados fazem a operação necessária. No entanto, com o evoluir da tecnologia, os recursos computacionais de um computador pessoal começam a ser suficientes para simular redes complexas tais como uma BTS do GSM.

Como exemplo de ferramenta de desenvolvimento de SDR temos o GNU Radio[Blo04]. Como exemplo de dispositivo que implemente SDR temos o URSP1 da *Ettus Research*[Ett08][DDN08], demonstrado na figura 2.9.



Figura 2.9: USRP1 da *Ettus Research*, com um rádio para envio e outro rádio para recepção

2.3.3 Sistemas Ativos e Sistemas Passivos

2.3.3.1 Sistemas Ativos

Para recolher dados úteis para a localização de um terminal móvel à custa da informação disponível sobre a rede é necessário ter conhecimento do protocolo e processar cada uma das suas camadas. Existem vários tipos de parâmetros utilizáveis para o efeito. Embora estes variem conforme a tecnologia em uso, existe um que é comum ao Bluetooth, ao WiFi e ao GSM: o RSSI. Tal como já foi mencionado, este parâmetro mede a potência de sinal recebida pela antena do recetor. A sua variação pode dever-se a vários fatores, também já aqui introduzidos (ver sub-secção 2.3.1). Porém existe um outro fator, cuja origem torna a potência dos vários sinais recebidos menos previsíveis. Esse fator é o fabricante do dispositivo móvel. Uma das características mais apreciável num terminal móvel é a sua autonomia, o que faz com que a relação entre a potência de emissão das comunicações sem fio e a qualidade do canal de comunicação (largura de banda máxima, tempo de resposta, etc.) seja um parâmetro dependente do fabricante. Ou seja, a especificação dos protocolos não impõe limites quanto ao comportamento que um dispositivo móvel deve ter para além do necessário para existir uma canal que sirva o propósito desse dispositivo. Além disso também o utilizador, através de aplicativos fornecidos pelo fabricante, pode aumentar ou reduzir a potência de sinal se assim o desejar. Na prática isto significa que nem sempre para os mesmos valores médios de RSSI o dispositivo móvel vai estar a uma mesma distância em relação à mesma antena recetora.

Através dos protocolos que implementam o canal de comunicação é possível intervir ativamente na operação de medição da força de sinal. Isto porque ao utilizar os parâmetros de configuração e funcionamento da tecnologia é possível interagir com os dispositivos vizinhos. Isto é, pode-se intervir na comunicação através de mensagens específicas para a descoberta da rede em redor da antena ou fazer pedidos a terminais em específico ou ainda pedidos que todos recebam e consequentemente respondam. Esta abordagem permite ainda saber quando certos eventos vão acontecer e por isso prevê-los ou espera-los de modo a adaptar a estratégia de localização.

Para uma determinação o mais próximo do real possível deve-se optar por uma estratégia que estime a posição do terminal num sistema de coordenadas obtido a partir da planta do local. Um sistema deste tipo irá necessitar de um maior número de antenas dispersas uniformemente pelo local. Isto prende-se com o facto de serem necessárias obter, pelo menos, três medições de antenas distintas num curto intervalo de

tempo para se estimar uma coordenada de dois componentes (ou seja, num referencial $O(x,y)$) [Fan86].

Caso tal não suceda, ou tal precisão não seja necessária existe a possibilidade de determinar se o terminal se encontra numa área em concreto recorrendo a uma medição com duas antenas distintas num instante do tempo. Utilizando apenas uma antena é possível saber se um terminal se encontra nas redondezas, ou numa dada direção caso se trate de uma antena omni-direcional ou direcional (respetivamente).

Para melhor demonstrar utilizações possíveis para tais sistemas (de uma, duas ou três antenas) será dado um exemplo de utilização para um sistema ativo de localização *indoor*. A descrição deste será feita no paragrafo seguinte com recurso à figura 2.10 onde se utiliza uma antena de deteção para ativar cada duas iluminárias:

Imagine-se que um técnico de reparações necessita de efetuar uma operação de manutenção em todo o comprimento de uma tubagem ao longo de um corredor. Como se trata de um corredor que serve apenas para possibilitar o acesso à tubagem, as lâmpadas que o iluminam encontram-se desligadas a menos que alguém autorizado se encontre presente. Dado que o técnico tem permissão para efetuar o trabalho de manutenção, este liga o BT do seu telemóvel e dirige-se para os locais necessários. O sistema de localização ao detetar a presença de um terminal cujo endereço de rede está registado determina a sua posição no corredor e acende, progressivamente, apenas as luzes necessárias para que o técnico caminhe confortavelmente e em segurança pelo corredor. As luzes não necessárias são também apagadas. Em caso de o sistema não ter detetado atempadamente três medições de antenas diferentes no mesmo instante de tempo, mas apenas duas, é possível que acenda apenas secções do corredor. Embora os gastos energéticos sejam maiores, continuam a ser inferiores comparando com o acender as lâmpadas de todo o corredor. No extremo, se apenas uma antena recolher sinais BT do técnico, ainda que esporádicos, o sistema poderá ser programado para iluminar todo o corredor mas inferir secções por onde o técnico já tenha passado e progressivamente desligar as lâmpadas desnecessárias. Ao longo do tempo o sistema acabará por compensar os custos da sua instalação reduzindo os custos com a iluminação e troca de luminárias.

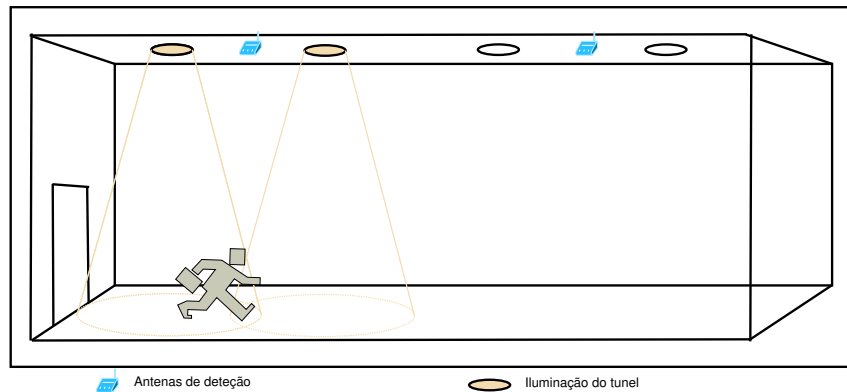


Figura 2.10: Exemplo de Utilização do Sistema Para Iluminação

Neste exemplo admite-se que o técnico não está preocupado com a privacidade da sua localização pois o endereço de rede do seu dispositivo foi dado a conhecer ao sistema. Como consequência será possível traçar o seu percurso e quanto tempo demorou em cada secção do corredor. Em último caso esta informação poderá servir para avaliar a sua *performance* sem que este tenha conhecimento.

2.3.4 Utilizando a força de sinal observada

Utilizando a abordagem passiva, ou seja, sem utilizar o protocolo como meio para inferir informações à cerca da rede em sua volta, reduz-se a necessidade de poder de processamento no sistema de localização. Isto deve-se ao facto de não ser necessário processar os dados de cada uma das camadas do protocolo a utilizar. A abordagem passa por utilizar SDR para analisar a sensibilidade do espectro numa frequência exata e assim determinar diferenças na potência de sinal observado. Uma alteração na potência observada implicará o uso de dispositivos sem fios e consequentemente a sua presença.

2.3.5 Algoritmos matemáticos

2.3.5.1 Trilateração

A trilateração é a obtenção da localização de um ponto utilizando distâncias entre pontos e que envolvam cálculos geométricos. O algoritmo da trilateração consiste, por isso, na localização de um ponto usando pelo menos dois pontos de referência e a distância entre o ponto a determinar e cada um dos pontos de referência[Fan86]. É

ainda necessário saber as coordenadas dos pontos de referência. Quanto mais pontos de referência forem utilizados, mais precisa é a localização do ponto desconhecido. Utilizando dois pontos de referência é possível limitar a localização a dois pontos, sendo necessário um terceiro para inferir sobre qual dos pontos é o correto. Por este motivo, considera-se que três é o número mínimo para obter uma localização confiável. Por vezes, e quando não é possível obter a distância exata dos pontos de referência ao ponto a ser calculado, é obtida uma zona de intersecção ao invés de um ponto. É apenas em casos de medições imprecisas que o uso de mais de três pontos de referência se tornam vantajosos.

O algoritmo na sua essência resolve o seguinte problema: Num plano de duas dimensões, um certo ponto p terá de ser localizado dadas as suas distâncias às referências conhecidas. As coordenadas de cada referência são conhecidas (x,y) assim como as distâncias desde o ponto p até às referências (no exemplo que será dado de seguida apresenta quatro referências: A_r , B_r , C_r , D_r). Ora então, com estas informações o algoritmo da trilateração irá calcular quais as coordenadas do ponto p no plano recorrendo a três referências conhecidas. Podem, no entanto, existir erros nas medidas conhecidas. Mais concretamente, neste trabalho, as distâncias entre o ponto a localizar e as referências conhecidas são medidas recorrendo ao valor da potência de sinal (RSSI) observado pelas referências e extrapoladas a partir da distancia entre cada uma. Isto impõe um problema, já mencionado no subcapítulo 2.3.1, explicado pela variação constante do RSSI. De modo a atenuar este erro podemos optar por escolher as três referências cujo RSSI é melhor em relação a p . De seguida, para resolver o problema da trilateração podemos optar por uma das duas soluções: *i*) utilizar o método dos mínimos quadrados não-lineares [PTVF92], ou *ii*) calcular a intersecção entre círculos.

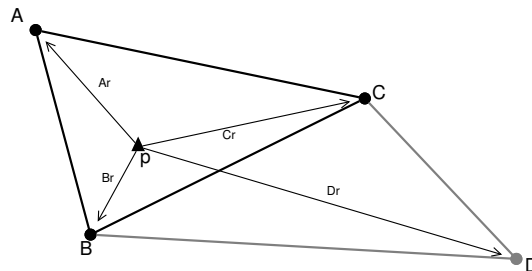


Figura 2.11: Demonstração do processo de trilateração

Passemos a exemplificar a resolução utilizando a alternativa *i*), a figura 2.11 e os dados conhecidos:

$$A = (a_x, a_y), B = (b_x, b_y), C = (c_x, c_y)$$

$$a_r = \overline{Ap}, b_r = \overline{Bp}, c_r = \overline{Cp}$$

A descrição do problema pode ser vista da seguinte forma:

$$(p_x - a_x)^2 + (p_y - a_y)^2 - a_r^2 = 0$$

$$(p_x - b_x)^2 + (p_y - b_y)^2 - b_r^2 = 0$$

$$(p_x - c_x)^2 + (p_y - c_y)^2 - c_r^2 = 0$$

Neste sistema de equações é relacionado cada componente de p (p_x, p_y), com cada componente do ponto de referência conhecido ($(a_x, a_y), (b_x, b_y), (c_x, c_y)$ no caso do nosso exemplo) e com a distância desde p até cada referência (a_r, b_r, c_r). Como se pode notar estamos perante um sistema de três equações com duas incógnitas, ou seja, um sistema sobredeterminado. Por esse motivo existe, de modo geral, mais do que uma solução única ([PTVF92], capítulo 15). Para encontrarmos as possíveis soluções poderemos recorrer ao método dos quadrados não-lineares. Para isso re-escrevemos o sistema substituindo o lado direito de cada equação por um valor residual não nulo:

$$(p_x - a_x)^2 + (p_y - a_y)^2 - a_r^2 = a_\Delta^2$$

$$(p_x - b_x)^2 + (p_y - b_y)^2 - b_r^2 = b_\Delta^2$$

$$(p_x - c_x)^2 + (p_y - c_y)^2 - c_r^2 = c_\Delta^2$$

Com isto será possível chegar a uma solução única (p_x, p_y) que minimiza a soma dos quadrados residuais ($a_\Delta^2 + b_\Delta^2 + c_\Delta^2$). Porque a solução do sistema é resultado da soma de termos quadráticos é justificado o uso do algoritmo do método dos quadrados não-lineares. Este começa com uma aproximação inicial para a solução e itera sobre esta o número de vezes necessário até encontrar a solução. Em cada iteração cada valor residual é corrigido de modo a minimizar a sua soma[PTVF92].

A segunda solução, a intersecção de círculos, consiste em determinar o centróide da intersecção de pares de círculos. O resultado final será, no caso ótimo um único ponto, tal como demonstrado na figura 2.12 do lado esquerdo. Caso a medição das distâncias a cada referência não seja precisa (conter um erro de medição pequeno o suficiente) o resultado será uma pequena área resultante da intersecção das três, tal como demonstrado no lado direito da mesma figura. O centróide da área resultante é o ponto onde se pode afirmar com um certo grau de confiança - determinado pela área resultante da intersecção de todas as referências - que se encontra o ponto cujas coordenadas não são conhecidas.

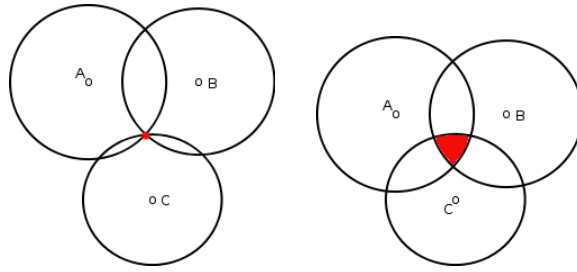


Figura 2.12: Exemplo da intersecção de três círculos consoante a precisão das medições obtidas para a distância

Para melhor percebermos como podemos obter a distância através deste algoritmo, passaremos a explicar como obter o centróide da intersecção entre os dois círculos da figura 2.13:

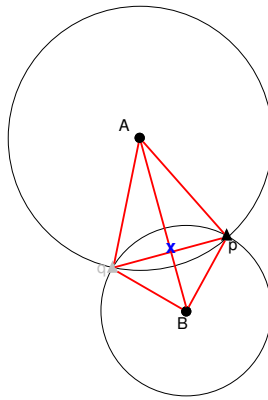


Figura 2.13: Determinação do centróide da intersecção entre dois círculos

$$A_{\odot} = (a_x, a_y), B_{\odot} = (b_x, b_y) \quad (2.1)$$

$$A_r = \overline{Ap}, B_r = \overline{Bp} \quad (2.2)$$

Devemos primeiro obter os pontos resultantes da interseção p e q de coordenadas (p_x, p_y) e (q_x, q_y) respetivamente. Para tal encontramos primeiro a distancia (Δ) entre o centro ($\odot$) de cada círculo:

$$\Delta_x = b_x - a_x \quad (2.3)$$

$$\Delta_y = b_y - a_y \quad (2.4)$$

$$\Delta^2 = \Delta_x^2 + \Delta_y^2 \quad (2.5)$$

$$\Delta = \sqrt{(\Delta^2)} \quad (2.6)$$

A partir desta distancia podemos verificar se os círculos se intercetam testando se a soma dos seus raios é menor ($\Delta < A_r + B_r$). Caso tal não se verifique, então podemos concluir que os círculos não se intercetam ou estão contidos um dentro do outro. Posto esta verificação podemos avançar para calcular o centróide (X): Tomando s como sendo a distancia $\overline{XA_\odot}$, t a distancia $\overline{XB_\odot}$ e u a distancia $\overline{qX}$ (ou $\overline{pX}$) temos:

$$s^2 + u^2 = A_r^2 \quad (2.7)$$

$$t^2 + u^2 = B_r^2 \quad (2.8)$$

Subtraindo as duas equações temos:

$$s^2 - t^2 = A_r^2 - B_r^2 \quad (2.9)$$

Fatorizamos o lado esquerdo:

$$(s - t)(s + t) = A_r^2 - B_r^2 \quad (2.10)$$

Como $s + t = \Delta$ e $t = \Delta - s$ podemos deduzir que s pode ser obtido a partir de:

$$(s - (\Delta - s))\Delta = A_r^2 - B_r^2 \quad (2.11)$$

$$2s\Delta - \Delta^2 = A_r^2 - B_r^2 \quad (2.12)$$

$$s = \frac{\Delta^2 + A_r^2 - B_r^2}{2\Delta} \quad (2.13)$$

Com isto podemos escrever que os componentes x e y das coordenadas de X são:

$$X_x = A_{\odot x} + \frac{\Delta_x s}{\Delta} \quad (2.14)$$

$$X_y = A_{\odot y} + \frac{\Delta_y s}{\Delta} \quad (2.15)$$

Finalmente para obter as coordenadas de q e p , substituindo a equação 2.13 em 2.7 temos:

$$u = \sqrt{A_r^2 - s^2} \quad (2.16)$$

, logo

$$q_x = X_x - \frac{\Delta_y u}{\Delta} \quad (2.17)$$

$$q_y = X_y - \frac{\Delta_x u}{\Delta} \quad (2.18)$$

$$p_x = X_x - \frac{\Delta_y u}{\Delta} \quad (2.19)$$

$$p_y = X_y - \frac{\Delta_x u}{\Delta} \quad (2.20)$$

$$(2.21)$$

Com isto obtemos as coordenadas de p e de q , em que X_x e X_y são as coordenadas do centro da interseção. Este processo deverá ser repetido n vezes, onde n é o número de círculos que se intersejam. Deverá ser iniciado pelos dois pontos de interseção mais próximos entre si e os resultados são confiáveis se estes se encontrarem mais próximos entre si do que os restantes pontos mais distantes. Então cada par de interseções irá gerar um aglomerado de pontos centróides das sua interseções. O centro deste aglomerado de pontos será a posição central, cujas coordenadas eram desconhecidas no início do processo.

2.3.5.2 Triangulação

A triangulação é a obtenção da localização de um ponto utilizando o ângulo formado entre pontos[Jic79][ECC03]. O algoritmo da triangulação consiste, por isso, na localização de um ponto usando dois pontos de referência, o ângulo medido entre o ponto

a determinar e cada um dos pontos de referência. Para que se possa determinar o ponto final é ainda necessário saber a localização dos pontos de referência, de modo a se saber a distância entre os dois.

Esta técnica é conceptualmente idêntica à Trilateração. Diferencia-se por utilizar não as distâncias desde o ponto a localizar até às referências conhecidas, mas o ângulo com o qual cada referência observa o ponto a localizar¹¹. Tomando uma referência angular em comum (o norte magnético, por exemplo), cada referência calcula o ângulo a partir do qual é possível comunicar com o ponto a localizar. Dadas três referências a comunicar com o ponto, e as suas coordenadas é possível estabelecer uma relação entre o ângulo de comunicação, as distâncias de cada referência e as coordenadas do ponto.

Esta técnica requer, no entanto, um investimento grande em *hardware* de deteção. Especificamente, requer que cada referência possua um conjunto de antenas direcionais cuja amplitude angular de receção individual cubra a área a observar. Pode também ser conseguido o mesmo efeito com apenas uma antena a observar em cada referência, mas cujo ângulo de observação varie em cada instante de tempo, numa amplitude de 360° . É também conhecido que o seu comportamento é negativamente afetado em ambientes propícios a aumentar a propagação por multi-caminhos[FMMSG09]. Isto porque ao recorrer ao ângulo de deteção dos sinais de rádio, em ambientes pouco amplos um sistema deste género irá obter um grande número de falsas deteções.

Dados estes pontos, uma instalação deste tipo de sistema vai completamente contra os objetivos desta Tese. Por este motivo será apenas mencionado neste sub-capítulo.

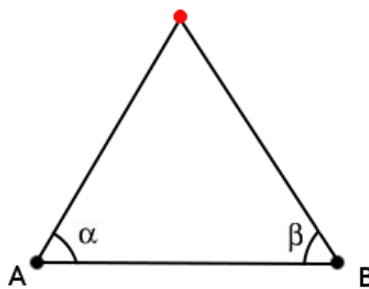


Figura 2.14: Triangulação usando os ângulos entre o ponto a ser calculado e as referências

¹¹Também o ponto a localizar pode utilizar o ângulo no qual observa as referências conhecidas. No entanto o objetivo deste trabalho é efetuar as medições no sistema de localização e não no terminal móvel. Nesse sentido esta opção não será explorada

2.3.5.3 Trilateração com Recurso a Extrapolação

A Trilateração com recurso a extrapolação tem como base a aproximação utilizada na Trilateração acima descrita (2.3.5.1) com um acrescento nos cálculos. Este acrescento torna-se útil quando a distância estimada desde uma antena de recolha até ao terminal emissor dos sinais rádio é maior do que a distância entre esta e as antenas vizinhas. Apesar de contra intuitivo, segundo [YR02], este caso pode acontecer enquanto o terminal emissor se encontra dentro da área abrangida pelo sistema de triangulação. A justificação para este tipo de comportamento encontra-se no facto do algoritmo utilizar como referência de distância duas antenas do próprio sistema de localização. Como consequência teremos para cada par de antenas utilizadas, a primeira servindo-se do valor de RSSI da segunda como referência para a distância entre ambas as antenas. Num caso ideal, este pormenor não introduziria qualquer tipo de erro nas estimativas. No entanto, dado o carácter móvel de um terminal, cuja natureza implica um funcionamento com baixos consumos energéticos, os fabricantes tendem a minimizar as potências de transmissão em prol da longevidade operacional em detrimento do alcance de comunicação por ondas rádio. Com esta ideia em foco, devemos também compreender que as ondas rádio de frequência de funcionamento que ronde os 2.4GHz tendem a propagar-se com dificuldade através de massas compostas principalmente por líquidos, como é o caso do corpo humano. Se imaginarmos a utilização normal de um dispositivo móvel, este permanecerá praticamente, ou protegido no interior da roupa do seu dono, ou muito próximo dele caso esteja em utilização.

Posto isto, facilmente conseguiremos entender que mesmo em condições ótimas de receção, um terminal móvel situado no centro de um triângulo equilátero imaginário, cujos vértices são compostos pelas antenas do sistema de localização, uma terá maior perda de sinal em relação às suas vizinhas. Fazendo por isso parecer que o terminal móvel se encontre mais próximo das antenas que tenham uma linha de vista livre, contra a antena que tem o utilizador entre o terminal e si mesma. Ilustrarmos este efeito de propagação na figura 2.15:

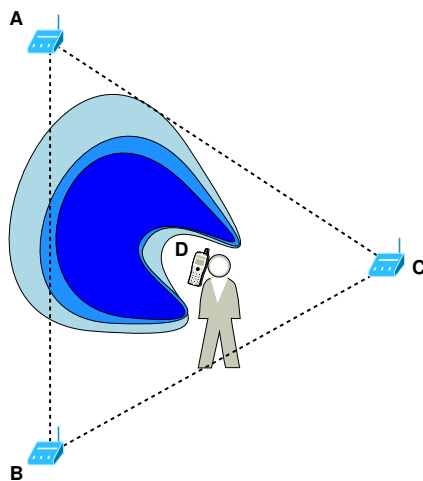


Figura 2.15: Exemplo de possível propagação de sinais rádio

No exemplo apresentado pretendemos demonstrar que embora os valores de RSSI entre as antenas A, B e C sejam sensivelmente iguais, assim como as três distâncias entre as antenas, os valores de RSSI entre as antenas A e B e o dispositivo D são substancialmente melhores do que entre a antena C e o mesmo dispositivo. Devido a este efeito, torna-se por vezes, impossível de efetuar uma boa estimativa da posição real dos dispositivos utilizando apenas a trilateração clássica tal como a descrita em 2.3.5.1.

Como solução possibilitaremos a extrapolação das distâncias entre as antenas e os dispositivos móveis sempre que a estimativa de coordenadas assim o requeira. Para tal habilitaremos a solução clássica para o problema da trilateração com interseção de círculos da capacidade de calcular o centróide final a partir de pontos cujos cálculos intermédios ultrapassem os limites do triângulo criado pelas antenas de deteção. Na próxima figura (2.16) faremos uma ilustração do resultado da aplicação desta solução.

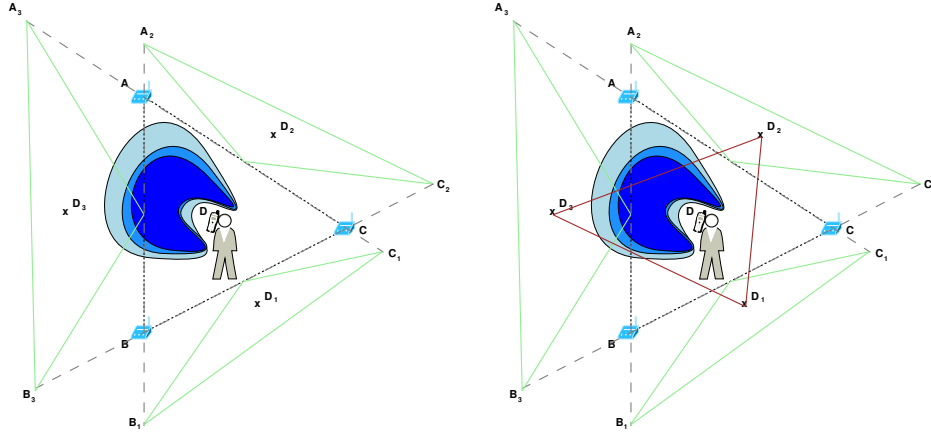


Figura 2.16: Solução para a Trilateração com erro através de extrapolação

Na ilustração acima pretendemos demonstrar o erro da potência de sinal recebido por cada antena devido à interferência provocada pelo utilizador. Esta é mais acentuada para a antena C, razão das distâncias $\overline{AA_3}$ e $\overline{BB_3}$ serem maiores que $\overline{BB_1}$ e $\overline{CC_1}$, $\overline{AA_2}$ e $\overline{CC_2}$. Pode-se verificar que estamos perante um caso em que todas as antenas do sistema detetam o dispositivo móvel como estando fora da área abrangida pelas antenas. Este facto irá gerar um erro de medição, já que o sistema de localização será levado a estimar a distância entre cada antena e o terminal móvel como sendo maior do que a distância entre a antena que serviu de referência para o calculo desse comprimento. Como resultado, os pontos médios resultantes do cálculo dos centróides (D_1 , D_2 e D_3) terão coordenadas situadas fora do triângulo de localização formado por ABC . No entanto, tendo conhecimento desta possibilidade podemos avançar para a estimativa final, sabendo que apesar da existência de um erro de medição o centróide final está aproximado da realidade.

2.4 Sistemas de Bases de Dados *NoSQL*

Com a evolução dos sistemas distribuídos e as aplicações *Web* foi necessário que os sistemas de suporte de dados crescessem de modo a acomodar quantidades de dados cada vez maiores. Os sistemas de base de dados tradicionais, de sua natureza relacionais, existentes antes da era *Web 2.0* apesar de capazes na sua altura, foram rapidamente perdendo eficiência perante o tipo de operações que as novas aplicações estavam a pedir. Embora estes sistemas tradicionais fossem aptas soluções para transações de escrita e leitura frequentes mas de pequenas dimensões, ou para grandes transações com um raro número de escritas, não conseguiam manter a sua eficiência perante

transações com grandes volumes de escritas e leituras em simultâneo[HHL11][Sto10].

Dada a necessidade deste tipo de operações, vários sistemas de base de dados surgiram como solução abdicando do princípio relacional das soluções existentes. Cada um com características distintas, que melhor serviram o “boom” das aplicações *Web 2.0*. Estas soluções implementam o conceito chamado de *Not Only SQL* (NoSQL). Este pretende endereçar os problemas que os sistemas de base de dados relacionais não conseguiam cumprir:

Escalabilidade: dado o crescente volume de dados torna-se necessário suportar incrementos de capacidade que sejam eficientes e de baixo custo

Disponibilidade: o acesso aos dados deve ser o mais rápido possível de modo a diminuir o tempo de resposta aos pedidos efetuados

Paralelização: o sistema de base de dados terá de processar várias transações em simultâneo mantendo a integridade dos dados

Manutenção: já que o volume de dados tende a crescer rapidamente os custos de expansão e manutenção não deverão aumentar linearmente com a dimensão do sistema

Com estes pontos em foco surgiram várias alternativas NoSQL capazes de servir grandes quantidades de dados mantendo a arquitetura estrutural de uma base de dados. Dadas as especializações de cada um destes novos sistemas de armazenamento e consulta de dados, faremos na tabela 2.4, um apanhado destas [Mon12][TB11][Cat11]:

¹²O *Hadoop Distributed File System (HDFS)* é um sistema de ficheiros distribuído concebido para funcionar em *hardware* acessível a utilizadores pessoais. http://hadoop.apache.org/docs/hdfs/current/hdfs_design.html

Nome	Propriedades
MongoDB	- mantém algumas propriedades das soluções tradicionais como <i>queries</i> e índices - as <i>queries</i> são escritas em <i>Javascript</i> - permite executar rotinas no lado do servidor - porque mapeia dados e índices em memória existe um limite de $\sim 2.5GB$ sobre a totalidade do sistema em máquinas <i>32-bit</i> - suporta dados geo-espaciais
CouchDB	- suporta duplicação do sistema - as operações de escrita não bloqueiam as de leitura - necessita de manutenção assistida - suporta versionamento de documentos - inclui biblioteca própria escrita em <i>jQuery</i> - é capaz de validar documentos <i>server-side</i> - suporta autenticação
HBase	- segue o modelo da <i>BigTable</i> da <i>Google</i> - usa o sistema de ficheiros HDFS ¹² para armazenamento de dados - permite pré-otimizar <i>queries</i> cuja resposta tenha que ser em tempo real - performance de acesso aleatório semelhante ao <i>MySQL</i> - usa o modelo Map/reduce
CassandraDB	- permite personalizar o comportamento balanceando entre a distribuição ou a replicação dos dados - os dados são serializados por coluna - permite usar índices secundários - as escritas são consideravelmente mais rápidas que as leituras - suporta o modelo Map/reduce se necessário - os nós do sistema são polivalentes (podem gerir ou ser geridos)

Tabela 2.4: Resumo de Sistemas de Base de Dados *NoSQL*

Capítulo 3

Sistema

O sistema de localização *indoor* proposto divide-se em três componentes distintos. O primeiro e segundo situam-se no local onde o sistema deverá proporcionar localização *indoor*. O terceiro componente encontra-se distribuído de preferência por vários locais (*datacenters*).

O primeiro componente é responsável por detetar os sinais rádio provenientes de terminais móveis que suportem comunicação sem fios por BT, WiFi ou GSM. É composto pelas antenas de deteção e respetivo computador onde estas estão conectadas. O segundo componente, o Servidor Local (SL) - assim denominado por se encontrar no local onde o sistema estará instalado - irá agregar as deteções efetuadas pelos componentes de deteção. O primeiro componente, por ser responsável pela recolha de sinais rádio dos terminais móveis, será composto por várias antenas formando grupos. Será necessário instalar vários destes componentes para que seja possível cobrir toda a área do local onde o sistema atuará. Dentro de cada grupo existiram dois papéis a desempenhar: o de *Master* e o de *Slave*. A mesma antena pode ser *Master* dentro de um determinado grupo mas agir como *Slave* dentro de um outro grupo. Só existirá um *Master* dentro de cada grupo. Na próxima figura ilustramos uma visão geral do sistema (figura 3.1).

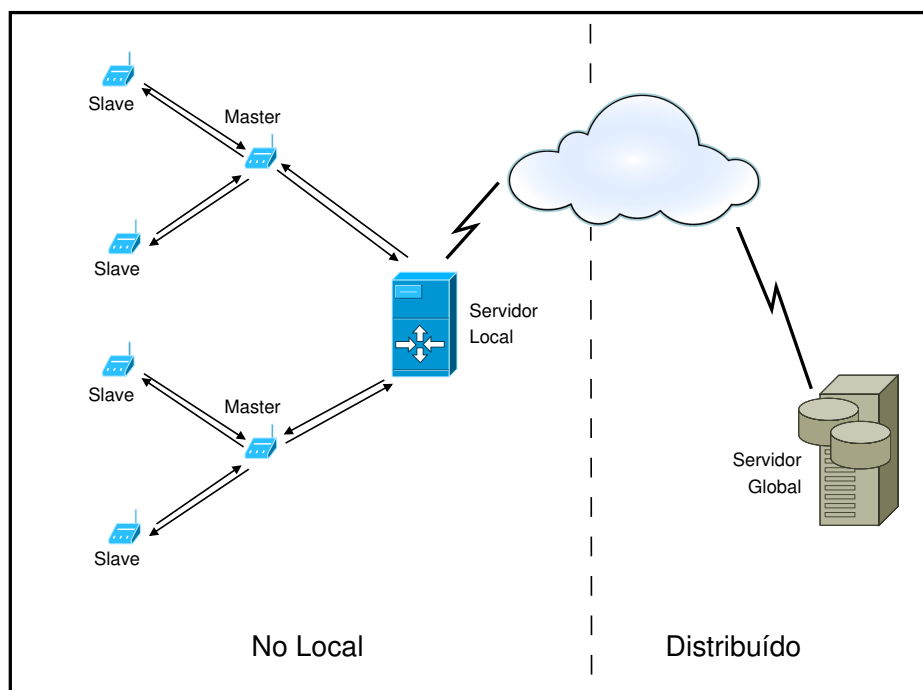


Figura 3.1: Visão Global dos Componentes do Sistema

O SL é responsável por aglomerar todas as detecções captadas pelos grupos de antenas instaladas no local. Quando for conveniente irá enviar as detecções para o SG do sistema. Este servidor tem também a responsabilidade de fornecer as coordenadas geográficas de cada grupo de antenas. Estas serão enviadas pelo segundo componente do sistema para o Servidor Local que as distribuirá pelas antenas.

O segundo componente do sistema de localização, o SG, tem a capacidade de receber todas as recolhas de detecção por parte das antenas, provenientes dos respectivos SLs. Este componente tem também a capacidade de configurar a localização de cada antena dentro do local onde será instalada. Isto será feito com recurso a um sistema de base de dados com suporte para dados geospaciais. Para além disso, é responsável por providenciar uma interface para que seja possível observar a localização de cada terminal móvel detetado pelo sistema.

A troca de dados entre as antenas e o SL, assim como entre este e o SG, é efetuada através de uma conexão segura por *Secure Sockets Layer* (SSL) utilizando *sockets* TCP/IP[Bhi02]. A *interface* que permite visualizar as detecções em tempo real é composta por uma página *web* que comunica com o SG por *http* e *web sockets*.

Quanto ao tipo de detecção a realizar (das descritas no capítulo 2.3) optamos por utilizar as funcionalidades do protocolo com coordenadas. Isto para as três tecnolo-

gias suportadas. Esta escolha prende-se com o facto de ser bastante mais complexo distinguir as recolhas efetuadas sem recorrer à informação contida nas camadas do protocolo. Seria também necessário recorrer a antenas direccionais, no género das necessárias para implementar o algoritmo da triangulação.

3.1 No local

Para providenciar dados de localização *indoor* através do sistema de localização proposto é necessário primeiro observar o ambiente estrutural do local. Desta observação resultará os locais a instalar cada antena de modo a otimizar as capturas de sinais rádio. Este primeiro passo é essencial para garantir um bom funcionamento do sistema e implica um conhecimento não só da planta estrutural do edifício mas também da infraestrutura de rede com e sem fios existente. Quanto à parte estrutural do local será boa prática evitar elementos que interfiram com o funcionamento de sinais rádio na frequência dos 2.4GHz (como fossos de elevadores, paredes espessas ou cozinhas equipadas com micro-ondas, por exemplo). Quanto à infraestrutura de rede existente é conveniente no caso do WiFi que as antenas captem as mensagens de *broadcast* dos SSIDs instalados, por exemplo.

Por questões de redução de custos, complexidade de instalação e configuração do sistema é preferível que as antenas estejam conectadas à mesma rede *ethernet* existente. Uma vez instaladas as antenas, será necessário que estas comuniquem com o SL para que enviem as deteções captadas. É através do Servidor Local que as antenas comunicam com o resto do sistema de localização.

3.1.1 Servidor Local

O Servidor Local tem como principal responsabilidade agregar as deteções realizadas pelos grupos de antenas instalados. Para que tal aconteça o SL comunica aos grupos de antenas as suas coordenadas relativas. Isto é, cada antena tem as suas próprias coordenadas dentro do grupo apesar de ser indiferente para o algoritmo de cálculo de posições as coordenadas de cada antena (desde que sejam respeitadas as distâncias reais entre elas). Esta decisão de modificar as coordenadas para cada antena deve-se à previsão de que talvez não seja do interesse de um proprietário dar a conhecer a planta completa de um edifício. Assim, poderá ser publicamente conhecida uma planta parcial do local onde é disponibilizado um sistema de localização *indoor* salvaguardando o

completo conhecimento do edifício. Dada esta funcionalidade cabe ao SL normalizar as coordenadas recebidas das antenas para as verdadeiras em relação ao edifício. Com isto, torna-se ainda possível existirem instalações em diferentes alturas e independentes entre si.

3.1.2 Master e Slave

Os elementos do sistema denominados Master e Slave são compostos por um computador e a antena respetiva à tecnologia a utilizar. A sua função é detetar pacotes emitidos pelos terminais moveis que contenham dados úteis para a sua localização. Estes elementos encontram-se instalados em grupos de três antenas. Cada grupo será portanto, composto por uma antena Master e duas antenas Slaves. Esta distinção entre as antenas é utilizada para orientar o algoritmo de localização sobre quem e quando será feita a tradução das deteções em coordenadas geográficas. Uma antena que aja como Master dentro de um grupo tem então, para além da tarefa de detetar terminais moveis, um acréscimo na informação a processar. Este acréscimo consiste em:

- pedir às restantes antenas que operem em modo *offline*
- efetuar os cálculos necessários para a triangulação de deteções
- enviar os pontos calculados para o SL

Comecemos por descrever a fase *offline* e explicar a sua necessidade. Existem duas justificações para esta. A primeira tem que ver com o facto do algoritmo de triangulação necessitar de saber a estimativa da distância dada a potência de sinal recebida (RSSI) de um dado terminal móvel detetado (ver dados conhecidos para iniciar os cálculos da triangulação por interseção de círculos em 2.3.5.1). Isto implica saber com alguma precisão qual o RSSI com que cada antena deteta as restantes do grupo. Sabendo esta informação e as coordenadas de cada uma das antenas do grupo o Master pode estimar a distância ao terminal móvel através da seguinte equação:

$$\text{Distância} = \frac{d_a \cdot rssi_d}{rssi_{ab}},$$

em que d_a é a distância entre as duas antenas,
 $rss_{i_{ab}}$ e rss_{i_b} é o rssi com que a antena A detetou a antena B,
e rss_{i_d} o rssi com que a antena A captou o terminal móvel

O protocolo BT não permite que antenas em modo *inquiry scan* respondam a *inquiry scans* de outras antenas. A biblioteca de *C* utilizada para ler aos dados das camadas do protocolo WiFi não permite aceder ao tipo de pacote de um modo direto. Por estas duas razões a fase *offline* é feita individualmente por cada antena antes de se iniciarem as deteções caso se esteja a detetar dispositivos BT, ou em paralelo com as deteções caso se esteja a trabalhar em WiFi.

O segundo propósito do modo *offline* é a alteração das condições do ambiente em que o sistema de localização está a operar. Ou seja, caso a temperatura, vento, humidade, disposição de objetos, etc. se altere com o passar do tempo, pedir às antenas que façam uma fase *offline* pode corrigir potenciais desvios na potência de sinal recebido com que cada antena conhece - da última fase *offline* - as restantes antenas do grupo. Enquanto que o envio dos pontos calculados para o SL não passa de uma simples transmissão de dados, efetuar os cálculos necessários para a triangulação exige mais ciclos ao *CPU* do Master e o triplo da memória *RAM*. Isto porque, após a antena Master efetuar uma escuta por sinais rádio para detetar os dispositivos móveis na sua vizinhança, vai pedir às restantes antenas as deteções por elas capturados, ordena-los, filtra-los, aplicar o algoritmo da triangulação sobre estes e por fim envia-los para o SL. A utilização de ciclos de *CPU* extra é consumida principalmente pela ordenação e filtragem dos dados. A ordenação é efetuada utilizando o método de ordenação *QuickSort*¹ por tempo de captura, sobre os dados de todas as antenas. A filtragem por sua vez, é necessária para garantir que pelo menos três deteções distintas (uma de cada antena) para o mesmo dispositivo móvel são enviadas para o algoritmo de trilateração. Este processo inclui ainda um parâmetro, a que chamamos “*intervaloMaximo*” cuja finalidade é garantir que não existe um intervalo maior que o valor deste parâmetro entre pelo menos três deteções para o mesmo dispositivo (originárias de antenas distintas). A resolução deste parâmetro é em milissegundos.

Na figura 3.2 ilustramos o processo de filtragem onde podemos encontrar várias deteções distintas para o mesmo dispositivo num determinado intervalo de tempo. As deteções relevantes para a descrição do processo de filtragem estão diferenciadas a cor. Cada cor corresponde a uma antena distinta. O processo de filtragem irá iterar

¹O algoritmo de ordenação *QuickSort* tem uma complexidade temporal de $O(n \log n)$ explicada pela sucessiva divisão em subconjuntos do conjunto inicial de elementos a ordenar.

sobre as deteções recolhidas pelo Master e Slaves alterando o tempo de deteção das recolhas. O objetivo é criar conjuntos inicializados por deteções do Master que ainda não tenham sido filtradas e se encontrem separadas entre si por mais de $t_{intervaloMaximo}$ (ms).

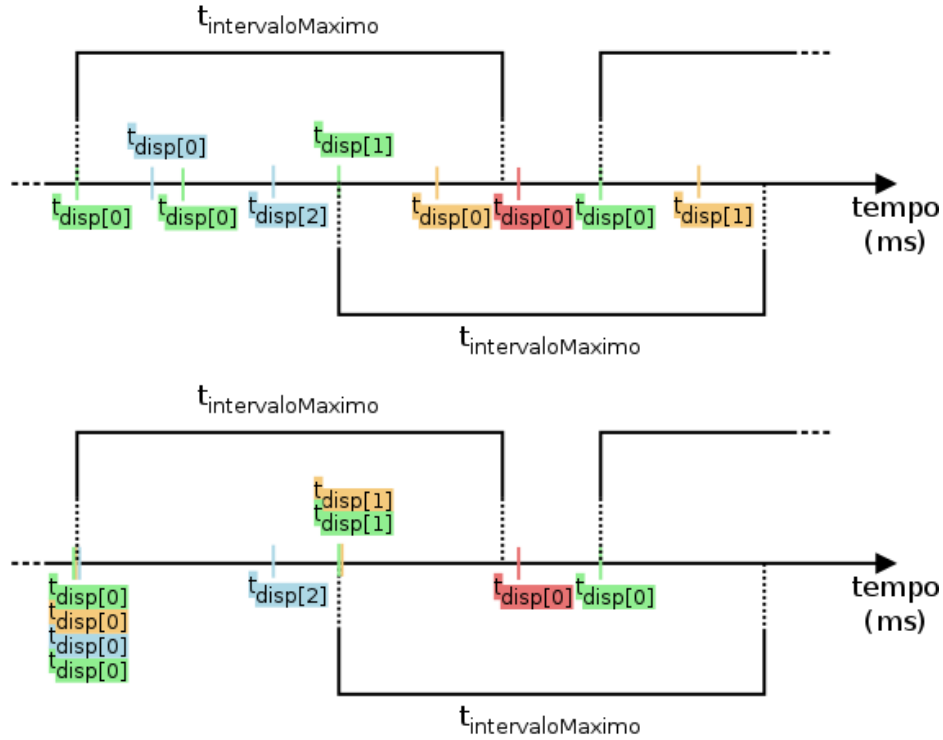


Figura 3.2: Visão global da implementação dos sistemas Master e Slave

No intervalo de tempo exemplificado na figura em cima, o processo de filtragem criou dois conjuntos de deteções. O primeiro para o dispositivo_[0] com 4 deteções e um segundo para o dispositivo_[1] com duas deteções. Um outro conjunto será ainda criado, novamente para o dispositivo_[0], dado que a próxima deteção não se encontra dentro da janela de $t_{intervaloMaximo}$ do primeiro conjunto. Note-se ainda que existe uma deteção - a vermelho - que irá acabar por ser descartada pois não cumpre as condições necessárias para pertencer a um conjunto. O segundo conjunto é composto por duas deteções para o dispositivo_[1] e também este acabará por não ser processado pelo algoritmo de localização. Isto acontecerá pois não contém o número mínimo de deteções por antenas distintas.

Podemos então resumir a lógica de criação de conjuntos de deteções, efetuada no processo de filtragem, ao seguinte esquema:

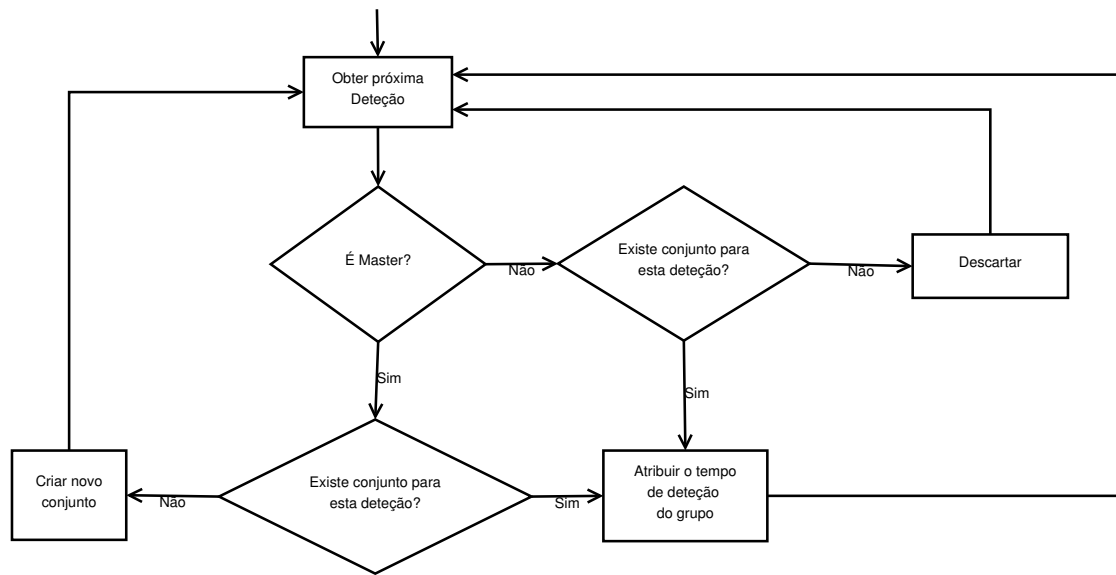


Figura 3.3: Lógica do processo de filtragem

Para concluirmos a descrição dos processos de funcionamento interno das antenas apresentaremos o algoritmo de cada rotina principal em pseudo-código: o algoritmo 3.1 para o Master e o algoritmo 3.2 para os Slaves.

Algorithm 3.1 Pseudocódigo da rotina principal dos Masters

```

if offlineFlag then
    Vizinhos  $\leftarrow$  descobreVizinhos()
end if

Detecoes  $\leftarrow$  obterDetecoes()

Slave1  $\leftarrow$  receberDados(idSlave1)
Slave2  $\leftarrow$  receberDados(idSlave2)

tdsDetecoes  $\leftarrow$  Detecoes
tdsDetecoes  $\leftarrow$  Slave1
tdsDetecoes  $\leftarrow$  Slave2

OrdenaPorTempo(tdsDetecoes)
tdsDetecoes  $\leftarrow$  normalizaEtiquetasTempo(tdsDetecoes, intervaloMaximo)

Pontos  $\leftarrow$  CalculaTri(tdsDetecoes)
Pontos  $\leftarrow$  criaMensagem(Pontos)
socketSL  $\leftarrow$  Pontos

```

Algorithm 3.2 Pseudocódigo da rotina principal dos Slaves

```

if offlineFlag then
    Vizinhos  $\leftarrow$  descobreVizinhos()
end if

Detecoes  $\leftarrow$  obterDetecoes()
Detecoes  $\leftarrow$  criaMensagem(Detecoes)
socketMaster  $\leftarrow$  Detecoes

```

Iremos de seguida falar sobre os dados contidos nos pacotes de rádio recolhidos. Esta informação contem diferentes campos consoante a tecnologia em uso. Faremos por isso uma breve exposição à cerca das diferenças.

No caso do BT os pacotes de interesse são as respostas aos *inquiry scans*, as *inquiry responses*. Estas respostas originam dos dispositivos BT que respondem com: o seu endereço MAC (endereço de rede) e a classe de dispositivo. A estes dados o controlador de *hardware* do recetor acrescenta ainda o *clock-offset* entre os dois relógios, uma etiqueta em milissegundos cujo valor é a data e hora em que o pacote foi recebido e o RSSI. Para a localização do terminal iremos utilizar o endereço de rede, o RSSI e o tempo de receção do pacote.

Caso estejamos a trabalhar em WiFi existem dois tipos de pacotes interessantes. O primeiro acontece no caso de *broadcast* de SSID (anúncio de rede por parte de um *PA*) e contem o campo *bssid* igual ao campo *source* e o campo *destination* igual a “FF:FF:FF:FF:FF:FF” (significando todos). Este tipo de pacote de dados será utilizado para acumular informações sobre as antenas Master e Slave do sistema de localização (fase *offline*). O segundo pacote de dados contendo informação relevante contém no campo *source* o endereço de rede do dispositivo que o emite e ambos os campos *bssid* e *destination* com o valor “FF:FF:FF:FF:FF:FF”, significando um pedido de *broadcast* por parte de um dispositivo WiFi aos *PAs* nas suas redondezas. Para obter o valor de potência de sinal recebido é necessária uma estratégia um pouco mais trabalhada que o BT apesar de tal como no primeiro, este valor ser etiquetado ao pacote recebido. Isto porque é necessário ter em conta dois pormenores técnicos. O primeiro é que um dispositivo WiFi em funcionamento normal, apenas expõe às camadas do protocolo passíveis de serem interpretadas por *software*, pacotes de dados que a ele sejam destinados. O segundo pormenor a ter em conta tem a ver com a obtenção do *rssi* das antenas do sistema de localização. Estas terão obrigatoriamente que estar a funcionar em modo *PA* para que emitam pacotes de *broadcast* de SSID para estes que possam ser captados pelas restantes antenas do sistema. Enquanto que agir como *Access Point* resolve o segundo problema, para resolver o primeiro é necessário que o *hardware* da antena WiFi suporte operar em modo promiscuo. Neste modo, os dados destinados a outros *PAs*, ou a *PAs* da mesma rede mas que funcionem em canais diferentes e que em funcionamento normal do *hardware* eram omitidos ao *software* são entregues às camadas superiores, possibilitando assim o seu processamento. Note-se que estes tipo de pacotes não terão qualquer interesse para um terminal WiFi ligado a uma rede e em comunicação com os restantes dispositivos a ela conectados. Logo, o facto de serem descartados é do interesse do terminal que os recebe.

No caso do terceiro protocolo, o GSM, os dados recolhidos não estão aglomerados num mesmo pacote. Por isso, será necessário observar os canais de controlo de ligação do protocolo e extrair a informação pretendida. Concretamente iremos necessitar dos dados de dois canais distintos. O primeiro denominado *Standalone Dedicated Control Channel* (SDCCH) pertence aos canais dedicados a controlo (*Dedicated Control Channel* (DCCH)). O segundo de nome *Access Grant Channel* (AGCH), pertence ao grupo dos canais para *Common Control Channels* (CCCH). Ambos os canais pertencem à camada superior do protocolo GSM (L3). A informação que será útil extrair pode acontecer quando um terminal móvel pretende executar uma das três ações ([TC-96], secção 4.4.4):

actualização de localização normal: atualizar a sua localização através de uma mensagem *Location Updating Request* (LUR)

actualização periódica: atualizar a sua localização após um determinado período de tempo²

IMSI *attach*: dar-se como ativo na rede da operadora

Do primeiro canal, o SDCCH um canal bidirecional - utilizado nos sentidos *uplink* e *downlink*, podemos extrair o IMSI que servirá como identificador único de cada dispositivo móvel. A extração do IMSI torna-se possível pois qualquer uma das três ações acima descritas implica o envio do identificador único do MS. Embora por vezes, consoante a configuração da operadora, o terminal emita apenas um identificador temporário (*Temporary Mobile Subscriber Identity* (TMSI)), este servirá de igual modo para a finalidade pretendida. Através do segundo canal podemos extrair a potência de sinal recebida, por parte da BTS, observando a informação produzida pela ação de atualização realizada pelo terminal móvel. Este canal, AGCH, comunica apenas da BTS para o MS e serve para anunciar qual dos os canais SDCCH disponíveis o dispositivo GSM poderá utilizar.

²Este tempo é determinado pelo contador T3212. Caso este não seja inicializado pela BTS o terminal móvel deverá (re-)inicializa-lo sempre que entre em modo *NORMAL STATE* ou *ATTEMPT TO UPDATE*. Este é parado sempre que o terminal abandone o qualquer sub-estado pertencente ao modo *MM IDLE STATE*.

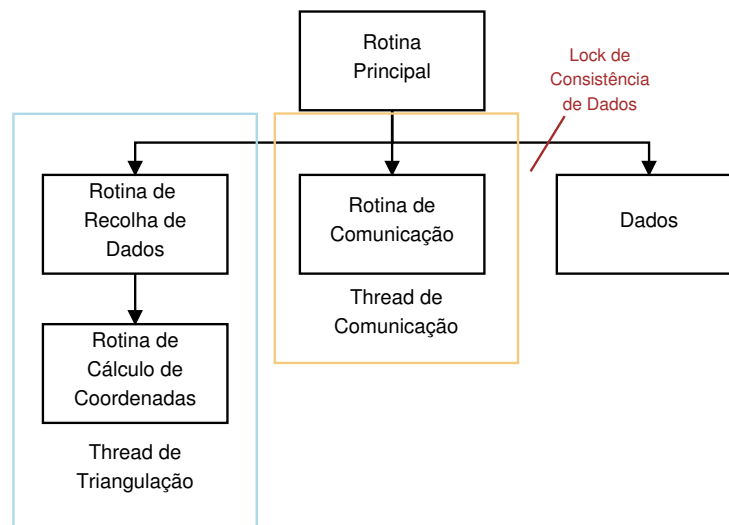


Figura 3.4: Visão global da implementação dos sistemas Master e Slave

Para a comunicação com o SL foi implementado um pequeno protocolo de troca de mensagens que está descrito na subsecção 3.1.5. É importante dizer que a rotina principal, tanto do Master como do Slave, é processada em duas *threads* (ver figura 3.4, sendo a primeira dedicada às deteções e a segunda às comunicações. A principal razão para este tipo de implementação é explicada pelo facto de tanto a sub-rotina que obtém deteções como a sub-rotina de comunicações serem bloqueantes. Dado que a rotina principal implementa uma *thread* para cada sub-rotina bloqueante esta não necessita de parar a sua execução enquanto não terminar uma das referidas sub-rotinas. Caso contrário, se determinado grupo se deparasse com uma falha de comunicação temporária, ou ausência prolongada de terminais móveis na sua vizinhança, implicaria uma falha em todo o sistema de localização. Então, caso se verifiquem vários pedidos de dados por parte de uma antena Master a uma Slave cujas respostas não contenham deteções poder-se-à concluir que ou não existem terminais móveis na redondeza dessa antena vizinha, ou que pode ter ocorrido um qualquer problema com o *hardware* de deteção. De igual modo, caso exista uma demora na comunicação de deteções pode-se concluir que determinada antena terá ficado sem acesso á rede *ethernet*. Neste último caso as deteções não enviadas serão mantidas em memória até o seu tamanho perfazer um limite que pode ser parametrizado.

Dado que ambas as rotinas necessitam de acesso à memória para operações de leitura e escrita de dados, e visto que ambas as rotinas estão a ser executadas em simultâneo foi necessário implementar um mecanismo de acesso exclusivo aos dados. Isto é obrigatório pois pode acontecer que num determinado instante ambas as rotinas necessitem

de realizar operações de leitura e/ou escrita. Como o acesso aos dados é sequencial e não atômico³ o que uma rotina vê pode vir a ser alterado e vice-versa. Então para assegurar a consistência das leituras e escritas foi implementado um mecanismo de acesso exclusivo. Este faz com que a primeira rotina a pedir acesso aos dados acione o mecanismo de exclusividade que não permite que outras tentativas de acesso sucedam, liberta-o quando a primeira operação de acesso terminar e avisa as restantes rotinas que estão a aguardar pelo acesso aos dados que estes se encontram livre. Deste modo garantimos que apenas uma das rotinas tem acesso aos dados em memória, seja para leitura ou para escrita. Na seguinte figura demonstramos uma visão global das rotinas dos Masters e Slaves.

A troca de mensagens entre Master e Slave, implementado pelas rotinas supra mencionadas consiste em três fases distintas. À primeira chamamos de *handshake* dado que é através desta que o canal de comunicação entre cada antena Slave e a Master é criado. Este canal servirá para troca de mensagens ente os dois componentes, inclusive os dados recolhidos das deteções efetuadas pelas antenas Slave. Devido à necessidade de comunicação constante, este canal é persistente. Isto é, só é fechado caso o sistema de localização vá ser desligado.

A segunda fase de funcionamento - a fase *offline* - servirá para que cada Master, de cada grupo de antenas tenha conhecimento da potência de sinal recebido de cada um dos Slaves do seu grupo. Esta fase é particularmente necessária caso o sistema esteja a operar em BT, pois neste caso cada antena irá na sua devida vez (coordenada pelo Master) efetuar deteções procurando unicamente pelas restantes antenas do grupo. Este processo de descoberta ocorre pelo menos uma vez em cada antena antes de se iniciar o modo de funcionamento normal. Caso a antena Master não tenha recebido um número mínimo de deteções de cada antena, irá pedir novamente aos Slaves em falta que repitam a fase *offline*. Se o número de deteções entre as antenas do grupo for suficiente, o Master irá imitar um pedido a cada Slave para que cada um inicie o modo de funcionamento normal.

No modo normal cada antena irá recolher as deteções de terminais móveis na sua redondeza. O tempo que este processo demora é parametrizável dependendo do tipo de aplicação. De um modo geral caso o tipo de movimento alvo seja um utilizador a andar o tempo de recolha poderá ser mais relaxo em comparação com, por exemplo, um utilizador a correr. Por outras palavras, o tempo de recolha deverá ser ajustado tendo em conta que o resultado final será mais preciso quanto mais tempo demorar

³não se realiza em apenas uma operação, mas num conjunto de sub-operações

a recolha - partindo do princípio que o utilizador se deslocou muito pouco ou nada. Caso se esteja perante uma aplicação, cujos deslocamentos alvo sejam bastante mais rápidos, como por exemplo um carro percorrendo uma rua, recomenda-se que o tempo de recolha seja o mais breve possível e o resultado final expandido de coordenadas que definem um ponto para uma secção contida dentro do espaço abrangido pelo sistema de localização.

Abaixo está representado um diagrama UML de sequência com os eventos mencionados.

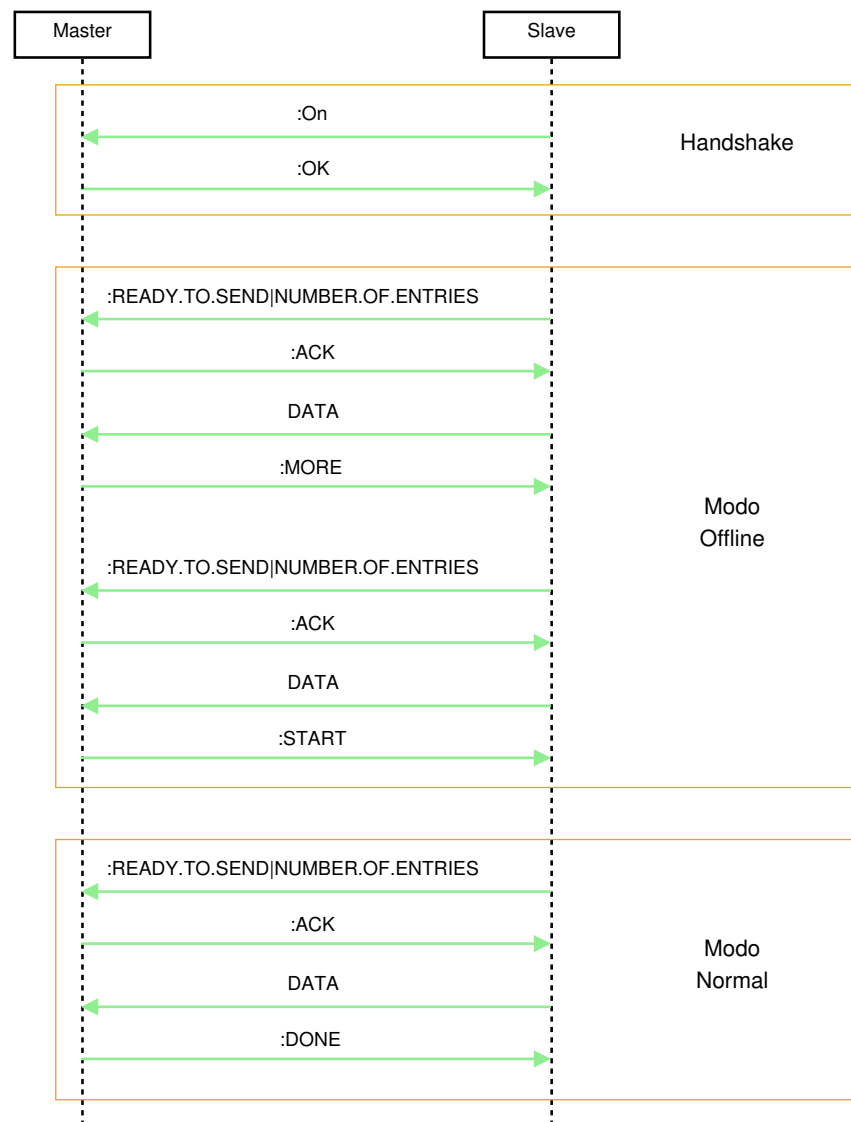


Figura 3.5: Diagrama de sequência das rotinas dos componentes Master e Slave do Sistema de Localização

Em prol da redução de custos com o sistema de localização, foram testadas várias plataformas para as máquinas às quais as antenas estão conectadas. Isto foi possível graças às capacidades de *cross compiling*⁴ do compilador de *C++*, o GCC⁵. Foi necessário recorrer a *cross compiling* (e a compilar as bibliotecas necessárias) para a plataforma *Advanced RISC Machine* (ARM). Isto porque foram testadas várias máquinas cuja arquitetura do *CPU* é ARM. Nomeadamente testamos a execução prolongada - até dois dias - em duas máquinas desta plataforma *DreamPlug* e *Raspberry Pi*⁶. No entanto, estas acabavam por eventualmente falhar durante a execução do sistema. A causa provável deste comportamento será o aquecimento do equipamento devido ao funcionamento contínuo por longos períodos de tempo. Dado que o estudo aprofundado desta Tese não ser a razão destas falhas, resolvemos adotar a plataforma x86 devido à robustez de execução. O *CPU* eleito foi o Atom⁷ da Intel, devido à relação velocidade/preço. Com esta solução o maior período de execução consecutivo foi 1 semana. Após esse período resolvemos parar a execução do sistema de localização e classificamos a solução como fiável.

3.1.3 Antenas Captura (BT/WIFI/GSM)

⁴*cross compiling* é a capacidade de gerar código executável para uma certa plataforma alvo a partir da plataforma original, onde o código é compilado

⁵<http://gcc.gnu.org>

⁶<http://www.raspberrypi.org>

⁷<http://www.intel.com/content/www/us/en/processors/atom/atom-processor-details.html>

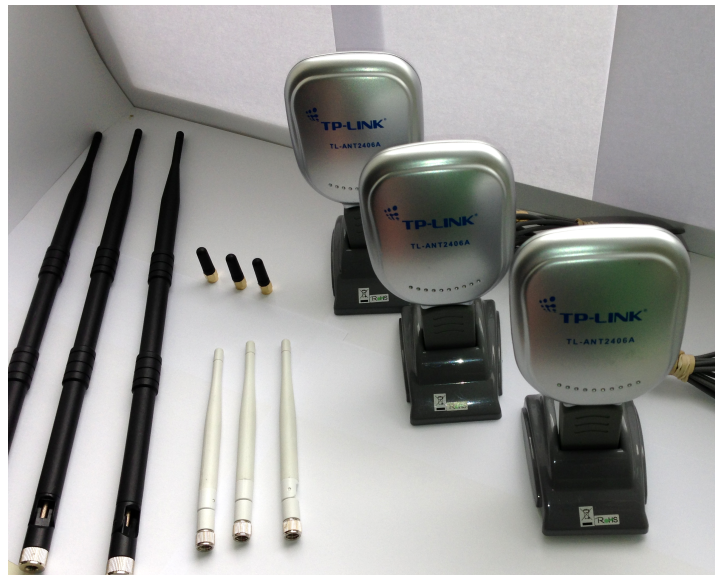


Figura 3.6: Antenas de 2.4 GHz utilizadas no sistema

As antenas utilizadas para a captura de deteções BT e WiFi foram as mesmas. Isto é possível pois a frequência de operação de ambas as tecnologias é a mesma: 2.4GHz. O comportamento do sistema foi testado utilizando antenas omni-direcionais e antenas direcionais. Obtivemos para os testes, três tipos de antenas omni-direcionais cujos ganhos são de 0, 3 e 9 dBi. Na figura 3.6, do lado esquerdo, estão demonstradas as antenas omni-direcionais utilizadas. No que toca a antenas direcionais apenas um modelo estava disponível: a TP-LINK TL-ANT2406A (lado direito da mesma figura), cujo ganho é de 6 dBi. Dada a necessidade de valores de potência de sinal o mais determinísticos possível dentro da área a operar, optamos por utilizar as antenas direcionais para o funcionamento em BT e as antenas de 3 dBi omni-direcionais para o WiFi. Embora contra-intuitiva, a escolha deve-se ao facto de para a mesma configuração do sistema os sinais de rádio BT propagarem-se com notável dificuldade em pontos longínquos das antenas e os sinais WiFi demonstrarem muito pouca ou nenhuma diferença para deslocações relativamente distintas (1 a 2 metros). Esta explicação é facilmente aceite dada a potência de transmissão de cada uma das tecnologias 2.4Ghz, que correspondem aos propósitos de utilização das mesmas. Na figura 3.7 pode ser observado o comportamento de propagação de sinal das antenas de 9 dBi omni-direcionais e das antenas direcionais, na metade esquerda e metade direita, respetivamente.

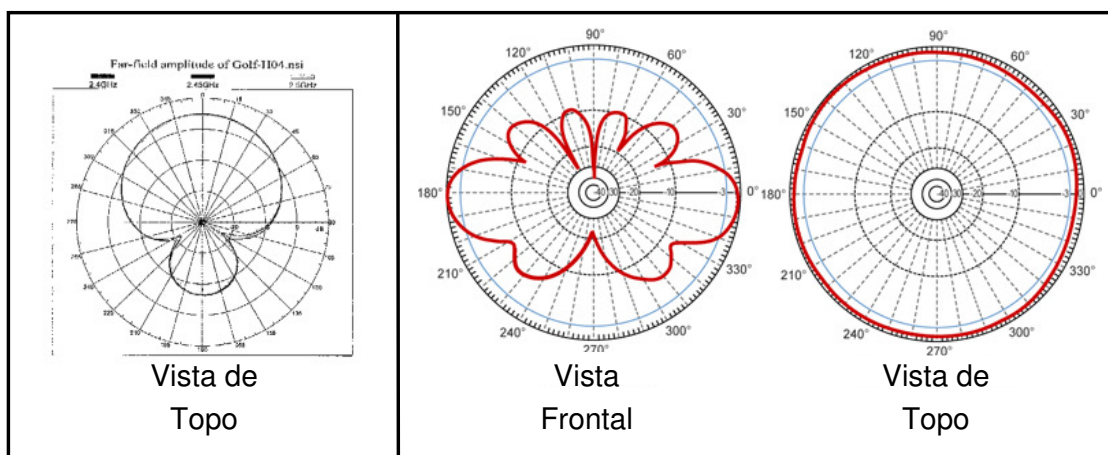


Figura 3.7: Padrão de propagação das antenas 2.4GHz utilizadas

As antenas utilizadas para o GSM são também específicas para o efeito. O intervalo de frequências em que operam situa-se entre os 824MHz e os 960MHz e entre os 1710MHz e os 1990MHz. Dado que o protocolo opera entre os 890MHz e os 915MHz para *uplink* e entre os 935MHz e os 960MHz para *downlink* na sua versão GSM-900 e entre os 1710MHz e os 1785MHz para *uplink* e entre os 1805MHz e os 1880MHz para *downlink*, estas antenas são ideais para a finalidade já que irão minimizar a ruído (neste caso ruído será qualquer sinal rádio foras das frequências que serão utilizadas). Dado que o GSM utiliza dois canais para comunicar, necessitando efetivamente de dois rádios para comunicar, são necessárias duas antenas para cada componente de deteção. As antenas podem ser vistas na figura 3.8.



Figura 3.8: Antenas de GSM-900/DCS-1800 utilizadas no sistema

3.1.4 Kinect

Durante o processo de testes ao sistema foi idealizado um mecanismo capaz de agir como interruptor do sistema de localização. Este serviria para ativar o sistema de localização em locais que não necessitem de um funcionamento constante, mas cuja

inicialização não dependesse da interação do utilizador. Para cumprir este fim, tal mecanismo foi implementado recorrendo a um sensor de movimentos munido com uma câmara de profundidade de alta resolução. O sensor utilizado, é fabricado pela Microsoft, chama-se Kinect⁸ e serve principalmente como dispositivo para jogar jogos na consola Xbox do fabricante.

Acedendo aos dados de profundidade utilizando biblioteca *libfreenect*⁹ pertencente ao projeto *OpenKinect*¹⁰ é possível medir a distância de objetos situados em frente à câmara de profundidade e inclusive segui-los. Foi exatamente essa a funcionalidade utilizada (entre a câmara VGA, os motores de orientação, e o microfone) para implementar um interruptor para o sistema. Dada a facilidade de programação proporcionada pela biblioteca mencionada foi implementado um algoritmo que segue um determinado corpo que se achesse em frente ao sensor e no evento deste ultrapassar uma meta previamente definida é disparado um segundo evento que pode ser interpretado como sendo uma pessoa que necessita do sistema de localização. O evento a disparar foi programado enviar ao Servidor Global uma mensagem contendo a data e a hora em que um objeto passou a meta definida. Com esta informação é possível o SG enviar uma mensagem ao SL para que ative conjunto de antenas responsável por providenciar localização *indoor* na área desejada.

Idealmente o sensor deve ser instalado no teto de um andar e de modo a que a câmara de profundidade cubra a entrada para a área abrangida pelo sistema de localização. Os parâmetros de deteção deste sistema são:

Tamanho máximo do objeto: tamanho máximo que um objeto a seguir pode ter antes que o sistema desista de o seguir

Distancia mínima de deteção: distância mínima desde o sensor até ao objeto para que este seja seguido

Distancia máxima de deteção: distância máxima desde o sensor até ao objeto para deixe de ser seguido pelo sistema

Vida útil de cada objeto: número máximo de *frames* que um objeto a ser seguido é memorizado em memória

⁸<http://www.xbox.com/en-US/KINECT>

⁹<https://github.com/OpenKinect/libfreenect>

¹⁰<http://openkinect.org/wiki/>

O algoritmo de detecção e rastreamento de objetos executa a sua tarefa recolhendo dados de infravermelho através de uma câmara monocromática para uma matriz de resolução *VGA* para posteriormente os processar. O processamento implica identificar os *pixels* com profundidade inferior ao alcance máximo do sensor, criar grupos destes elementos que sejam vizinhos (na matriz) e cuja profundidade seja idêntica. Após esta semi-identificação de objetos, chamados de *blobs*, o algoritmo cria uma caixa para cada um destes, chamadas de *bounding boxes*. Após ter processado toda a matriz - que representa um fotograma - esta informação é guardada em memória para ser comparada com a informação do próximo fotograma. Para rastrear cada objeto o algoritmo verifica se entre dois fotogramas consecutivos duas *bounding boxes* se intercetam. Caso isto se verifique assume-se que se trata do mesmo objeto, presente em ambos os fotogramas, cujo movimento foi o deslocamento dessa *bounding box* ente um fotograma e outro.

Após um objeto ultrapassar a meta definida a respetiva *bounding box* é eliminada da memória e um evento de novo utilizador na área de detecção é lançado. A implementação do algoritmo é descrita na figura 3.3 em pseudo-código.

A informação de profundidade é processada a cerca de 15 fotogramas por segundo. O que permite um correto rastreamento dos objetos que se atravessam em frente ao sensor. No entanto, também a implementação deste algoritmo teve em conta os recursos necessários para o seu funcionamento. Por isso foi integrado um mecanismo que automaticamente salta fotogramas consecutivos caso o CPU não esteja a processar os dados de profundidade a uma velocidade suficiente. Caso o sistema não tenha memória RAM suficiente para a memorização dos objetos já detetados, o parâmetro “Vida útil de cada objeto” pode ser reduzido para diminuir o consumo de memória.

Algorithm 3.3 Pseudo-código do algoritmo de rastreamento de objetos

matrizProfundidade $\leftarrow$ *obterInformacaoProfundidade()*

fotogramaAnterior $\leftarrow$ *obterBlobs(matrizProfundidade)*

fotogramaAnterior $\leftarrow$ *criarBoundingBoxes(fotogramaAnterior)*

while !*terminar* **do**

matrizProfundidade $\leftarrow$ *obterInformacaoProfundidade()*

fotogramaAtual $\leftarrow$ *obterBlobs(matrizProfundidade)*

fotogramaAtual $\leftarrow$ *criarBoundingBoxes(fotogramaAtual)*

movsDetetados $\leftarrow$ *intersetaBBboxes(fotogramaAtual, fotogramaAnterior)*

for all *movimento* $\in$ *movsDetetados* **do**

if *movimento.deslocamentoVertical* $\geq$ *meta.deslocamentoVertical* **then**

Eventos $\leftarrow$ *obtemDataEHora()*

end if

end for

socketSG $\leftarrow$ *Eventos*

fotogramaAnterior $\leftarrow$ *fotogramaAtual*

end while

3.1.5 Comunicação entre Componentes

Todas as trocas de dados entre os componentes do sistema de localização, à exceção entre o *Kinect* e o Servidor Global, são feitas através de *sockets* TCP/IP seguros por SSL. Entre *Kinect* e SG não foi implementada a funcionalidade SSL.

O motivo de se ter utilizado *sockets* TCP/IP prende-se com a portabilidade e compatibilidade oferecida. No que toca a compatibilidade podemos dizer que quer a linguagem C++ que implementa as antenas e o Servidor Global, quer o NodeJS que implementa o Servidor Global, suportam por completo este método de comunicação. Da parte da portabilidade, o sistema operativo *Unix* implementa um mecanismo que permite que o destinatário de cada *socket* seja configurado conforme o necessário. Isto permite que cada componente do sistema de localização seja configurado para falar tanto para si mesmo, como para outros componentes na mesma rede, mas também para fora da rede local. Um requisito imposto é adotar a troca de mensagens cujos tipo de dados sejam *strings*. O que implica que tipos de dados (como estruturas ou inteiros) sejam convertidos antes de serem enviados e reconvertidos para serem processados após a receção. Para facilitar este passo optamos por utilizar o formato *JSON*[Cro06] para troca de mensagens. É conveniente dizer que a todos os canais de comunicação se encontram alvos de um processo de autenticação em cada extremo e os dados encriptados. Para este efeito foi utilizada a biblioteca *OpenSSL* disponível em todas os componentes do sistema de localização.

Dentro de cada grupo de antenas cada canal de comunicação seguro é persistente, enquanto que os canais de comunicação - também eles seguros - entre antenas Master e SL, e SL e SG não são persistentes. Isto significa que um canal entre as antenas do mesmo grupo uma vez aberto, assim permanecerá. De modo contrário, cada canal aberto que não seja persistente é fechado já que deixa de ser necessário a sua utilização. Canais persistentes são usados para a comunicação entre os nós de detecção (antenas Master e Slave) e entre Masters e o Servidor Local já que a troca de dados será constante enquanto houverem dispositivos móveis que gerem detecções. Os canais não persistentes são criados sempre que houver uma quantidade de dados aglomerados no SL que justifique o envio para o Servidor Global. Existe porém uma exceção: caso exista uma interface que peça dados em tempo real o SL será notificado e abrirá um canal persistente para o envio das detecções para o SG para que este possa de imediato envia-las para a interface.

A troca de mensagens dentro do sistema de localização é dada através de um protocolo por nós criado. As mensagens trocadas entre antenas do mesmo grupo e entre Masters

e o resto do sistema devem, por isso, segui-lo. Caso contrário uma mensagem com um cabeçalho que não conste no protocolo irá ser ignorada pelo recetor.

Na figura 3.9 está representado um esquema das mensagens que são trocadas entre cada Master e Servidor Local, e desde com o Servidor Global.

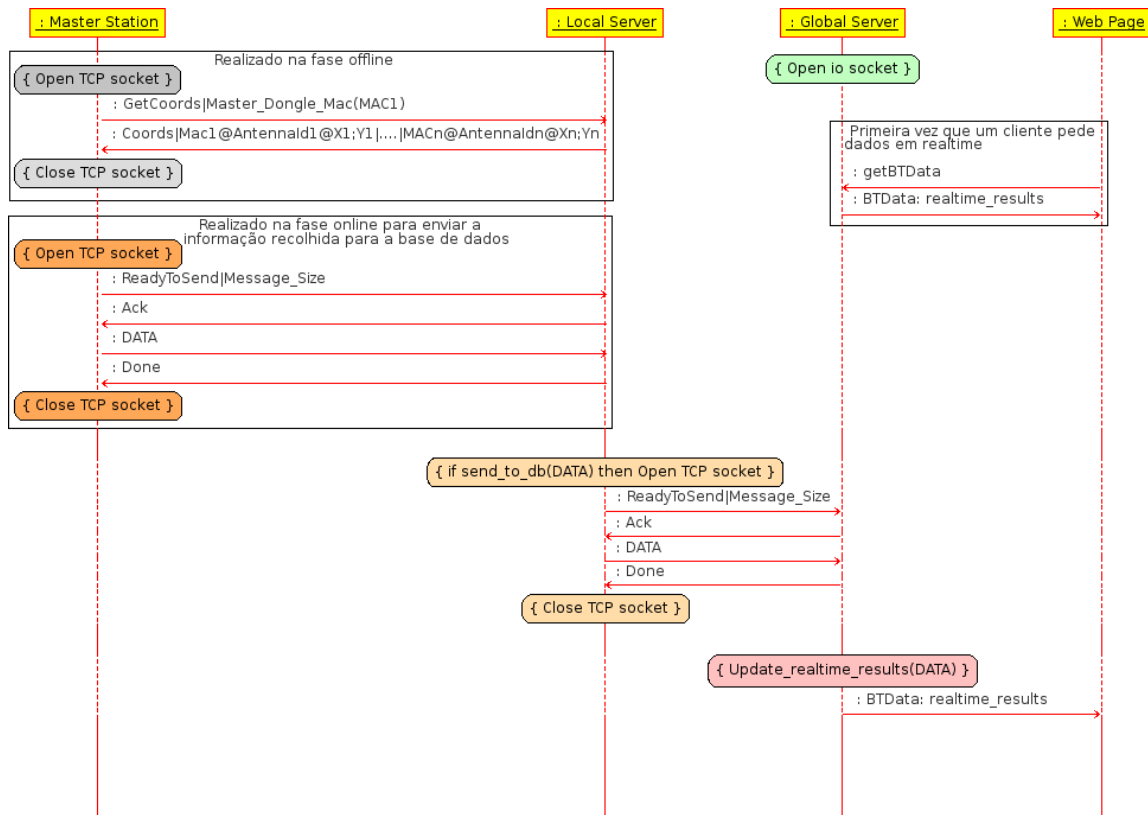


Figura 3.9: Esquema UML de sequência das mensagens trocadas desde uma antena Master até ao Servidor Global

3.1.6 Trilateração

O algoritmo de trilateração implementado segue o modelo teórico exposto em 2.3.5.1. O seu decorrer, tal como descrito na secção 3.1.2, mais concretamente nas listagens 3.1 e 3.2, procedimento *CalculaTri()* é composto por um número de fases iguais ao número de antenas que irão participar na estimativa de coordenadas para um dispositivo móvel detetado. As fases a que nos referimos derivam da necessidade de calcular os centróides relativos a cada par de antenas e de seguida o centróide do ponto final. A implementação desta lógica será apresentada na listagem 3.4 - correspondente ao processo de estimação efetuado com recurso a três antenas.

Algorithm 3.4 Pseudocódigo da rotina de Trilateração com três antenas

Ensure: faseOffline_[], tdsDispositivos, coordsA, coordsB, coordsC

```

1: for all disp  $\in$  tdsDispositivos do
2:   if disp.length() < 3 then continue
3:   end if

4:   ptADB  $\leftarrow$  distRefDispositivo(faseOffline[A][B], distAB, disp.rssi(A))
5:   ptADC  $\leftarrow$  distRefDispositivo(faseOffline[A][B], distAB, disp.rssi(A))
6:   ptMedioBC  $\leftarrow$  obtemPtMedio(coordsB, coordsC)
7:   centroideA = calcCentroide(ptADB, ptADC, ptMedioBC)

8:   ptBDA  $\leftarrow$  distRefDispositivo(faseOffline[B][A], distAB, disp.rssi(B))
9:   ptADC  $\leftarrow$  distRefDispositivo(faseOffline[B][C], distBC, disp.rssi(B))
10:  ptMedioAC  $\leftarrow$  obtemPtMedio(coordsA, coordsC)
11:  centroideB = calcCentroide(ptBDA, ptBDC, ptMedioAC)

12:  ptADB  $\leftarrow$  distRefDispositivo(faseOffline[C][A], distAC, disp.rssi(C))
13:  ptADC  $\leftarrow$  distRefDispositivo(faseOffline[C][B], distBC, disp.rssi(C))
14:  ptMedioAB  $\leftarrow$  obtemPtMedio(coordsA, coordsB)
15:  centroideC = calcCentroide(ptCDA, ptCDB, ptMedioAB)

16:  ptsFinais[disp.obtemId()]  $\leftarrow$  calcCentroide(centroideA, centroideB, centroideC)
17: end for

return ptsFinais

```

Podemos reparar que a rotina de Trilateração ignora grupos de antenas (ver paragrafo referente ao processo de filtragem na secção 3.1.2) que não contenham pelo menos três antenas (linha 2). Nesta listagem é assumida que todos os grupos têm exatamente três detecções de antenas distintas. O centróide relativo a cada par de antenas é calculado entre as linhas 4 e as linhas 15. Na linha 16 podemos verificar que os pontos estimados são concatenados na variável *ptsFinais*, indexados pelo identificador único de cada dispositivo.

3.2 O Sistema Distribuído

O nosso sistema de localização tem uma forte componente distribuída gerida pelo Servidor Global (SG). Este age como uma porta de acesso à parte do sistema de localização que deverá ser repartida fisicamente por vários computadores separados. Até este ponto no nosso trabalho temos feito referências ao SG como se de uma entidade única se tratasse. No entanto, este tem como função dirigir os pedidos que recebe para o componente distribuído responsável ou emitir dados provenientes destes componentes. A cargo do SG estão, uma base de dados NoSQL, uma *ORDBMS* com uma extensão espacial, e um interpretador de *Javascript* orientado para aplicações *web* altamente escalável.

3.2.1 Servidor Global

O Servidor Global tem como função cumprir as seguintes tarefas:

- configurar as antenas de detecção com os respetivos papeis e coordenadas
- receber as detecções dos Servidores Locais
- armazenar as estimativas de localização dos terminais móveis
- alojar a aplicação *Web* com as interfaces de instalação e configuração do sistema de localização
- alojar a aplicação de consulta de detecções em tempo real

Comecemos por descrever a configuração das antenas. As coordenadas reais são dadas a conhecer ao Servidor Local pelo Servidor Global recorrendo a *software* de mapeamento de coordenadas geoespaciais. Optamos por utilizar a aplicação *QuantumGIS*

“Lisboa”¹¹ pertencente ao projeto *OSGeo*¹², pois permite efetuar fácil e eficazmente os mapeamentos sobre uma imagem em formato digital e exportar essa informação geográfica para o sistema de base de dados geoespacial utilizado (ver exemplo na figura 3.10).

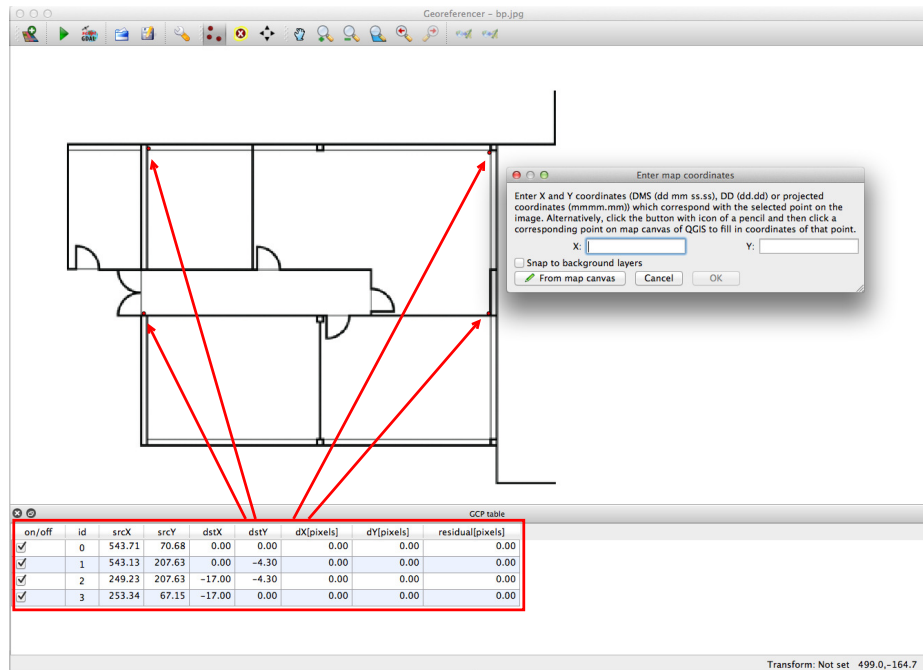


Figura 3.10: Mapeamento do local de trabalho situado em um andar que se encontra dividido entre duas empresas

Uma vez mapeados os locais onde as antenas se irão situar podemos exportar esta informação e posteriormente introduzi-la no componente de georeferenciação. Do mesmo modo podemos alterar a posição das antenas consultando a mesma imagem para depois introduzir as novas coordenadas para uma antena que tenha sido movida de posição. Em ambos os casos não será necessário qualquer outra interação humana para que o sistema gere estimativas para o local onde se encontram os terminais móveis.

A receção das deteções realizadas é feita através de um *socket* dedicado para cada SL. Processa-se em várias mensagens pois nem sempre é possível transportar todas as deteções de uma só vez. Aquando todas as mensagens que contêm as deteções para um dado intervalo de tempo o canal de comunicação é fechado e estas serão enviadas para o componente de armazenamento de dados. O componente responsável pelo envio e

¹¹O *QGIS* é um dos projetos *Open Source* que integra no trabalho da Fundação de Software Livre Geoespacial (OSGeo): <http://www.qgis.org/about-qgis/features.html>

¹²<http://www.osgeo.org/content/foundation/about.html>

pela receção de dados do SG é também responsável pela interface *Web* onde se pode consultar em tempo real as deteções armazenadas.

O componente de armazenamento das deteções será descrito já de seguida no subcapítulo 3.2.2. Os componentes de mapeamento geoespacial, de alojamento e transporte de mensagens serão alvos de uma breve descrição nos subcapítulos 3.2.3 e 3.2.4, respetivamente.

3.2.2 Cassandra

O sistema de base de dados escolhido para armazenar as deteções recolhidas pelo sistema de localização foi o *Apache CassandraDB*[Fea10]. Esta escolha deveu-se ao conjunto de funcionalidades que suporta. A funcionalidade que consideramos mais relevante neste sistema e que justifica a nossa escolha é a capacidade de armazenar dados ao longo de um indeterminado número de nós estejam estes situados numa só máquina, em várias máquinas distintas, ou mesmo em várias máquinas espalhadas pelo globo. Podemos ainda acrescentar que a estratégia de distribuição pode ser alterada em qualquer altura tendo em vista a latência de resposta ou a replicação de dados. Do ponto de vista das características mencionadas no capítulo 2.4 o *Cassandra* implementa o seguinte comportamento[Fea10]:

Disponibilidade: para o utilizador o facto de os dados estarem distribuídos pelo sistema é transparente. Um pedido ao sistema faz com que este seja responsável por descobrir o(s) nó(s) que contêm os dados necessários para o satisfazer.

Consistência nas escritas: o sistema pode ser configurado para diferentes níveis de consistência quanto à velocidade de resposta perante uma escrita. Consoante este parâmetro, uma ação de escrita sucede quando nenhum, um ou vários nós confirmem a sua escrita. Para isso, as escritas são feitas no $\log$ ¹³ de cada nó e só mais tarde realmente efetuadas.

Consistência nas leituras: as leituras são consideravelmente mais demoradas quer a nível de latência de resposta quer a nível de *I/O*. Isto porque o *log*

¹³O termo *log* é utilizado no sentido informático da palavra. Um sistema de *logging* serve para guardar um registo de operações que ocorrem. No caso do *Cassandra* este registo é consultado após cada entrada ter sido registada para ser processado.

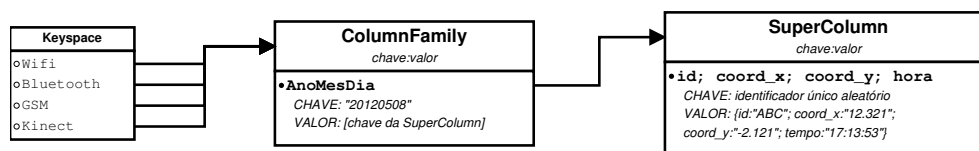
de cada nó deve ser consultado para garantir que os dados mais recentes são devolvidos.

Replicação: para garantir a tolerância a falhas o sistema irá conservar N cópias dos dados, em que N , chamado o fator de replicação, é configurável e tende a aumentar a latência das leituras consoante o número de nós utilizados para replicar os dados. É também possível configurar quais serão os N nós envolvidos na replicação tendo como opção a escolha ser automática, consoante a localização geográfica ou ainda consoante uma estratégia personalizável.

Custos: os custos associados ao aumento de capacidade de armazenamento são inteiramente aplicados no *hardware* adicionado, isto é, o sistema não impõe nenhuma restrição quanto às características de equipamento adicional. Invés, particiona os dados consoante a capacidade de processamento, de armazenamento, largura de banda e distância entre nós.

O comportamento do *Cassandra* no que toca à consistência nas escritas é especialmente interessante para o nosso sistema de localização, já que um local com grandes volumes de deteções irá provocar uma maior afluência de escrita de registos no sistema de armazenamento. Aqui poderemos ajustar a estratégia de consistência nas escritas consoante a relação entre a capacidade de resposta necessária e o número de nós de que dispomos. No que toca às leituras, estas são de menor importância dado que esperamos ter uma quantidade substancialmente menor deste tipo de operações.

Para além da flexibilidade na configuração do comportamento do sistema, também quanto ao formato dos dados este sistema de base de dados se demonstra adaptável. Isto porque através de um modelo de chave-valor permite armazenar dados indexados e livres de esquema permitindo que cada registo contenha mais ou menos campos consoante seja necessário. Este tipo de dados é chamado de *Column*. Uma *SuperColumn*, que alberga as *Columns*, é composta por uma chave e as várias *Columns*. As *SuperColumns* estão contidas em *ColumnFamilies*. Finalmente podemos ter vários *Keyspaces* contendo várias *ColumnFamilies*. De seguida, na figura 3.11 apresentaremos a organização dos dados escolhida para o nosso sistema de localização.

Figura 3.11: Organização dos dados para o *Cassandra*

Tendo em vista a figura que se seguiu, podemos ver quatro *Keyspaces* referentes aos quatro componentes capazes de captar dados. Cada *keyspace* contém uma *ColumnFamily* correspondente à data em que os dados foram recolhidos, cuja chave é composta pelo ano, mês e dia das capturas.

Por sua vez, a cada uma destas chaves é associada a uma *SuperColumn*, cuja chave é um identificador único e o valor o conjunto formado por: *i*) o identificador do dispositivo detetado, *ii*) as coordenadas (x,y) da estimativa para a sua localização e a hora do dia em que a deteção aconteceu.

3.2.3 PostgreSQL/PostGIS

Para a geo-referenciação das antenas Master nos locais onde serão instaladas utilizamos a solução *PostgreSQL + PostGIS*. A nossa escolha deveu-se ao facto do *PostgreSQL* ser um sistema de base de dados relacional que através da extensão de geo-referenciação *PostGIS* fica habilitado de funcionalidades que facilitam o mapeamento das áreas *indoor*[OH11]. O modelo relacional da base de dados é apresentado na figura 3.12:

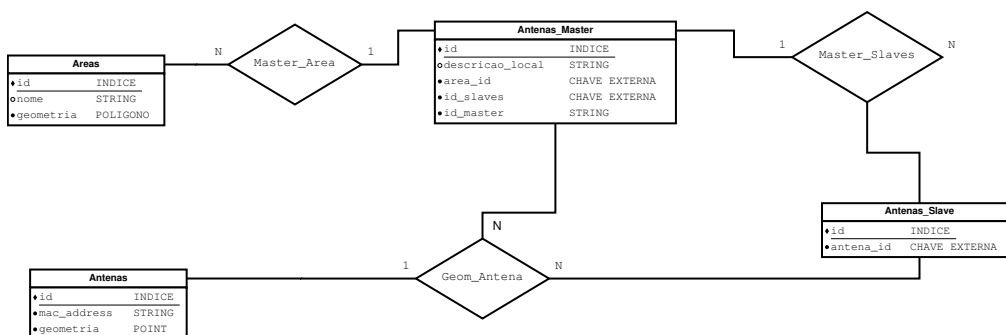


Figura 3.12: Modelo Relacional para dados geo-espaciais

Utilizando este modelos conseguimos consultar quais as antenas em cada local onde o sistema se encontra instalado. Os dados geo-referenciais provenientes do mapeamento efetuado no programa *QGIS* afetam diretamente a tabela *Antenas*. A interface *Web* irá colocar os restantes dados, como endereço *mac* da mesma tabela e *nome* da tabela

Areas. As relações entre antenas Master e Slaves deverão ser configuradas acedendo diretamente à interface *Web* do *PostGres*, o *PhpPgAdmin*.

3.2.4 NodeJS

O *NodeJS* é uma linguagem de programação baseada em *Javascript* executada do lado do servidor (*server-side*). Esta é baseada numa plataforma da *Google*, de nome “V8”¹⁴ cujo propósito é ter um ambiente de execução independente (*standalone*). Este ambiente desenvolvido na linguagem *C++* compila e executa código escrito em *Javascript*, gere a alocação de memória e respetivo *garbage collector*[Goo12].

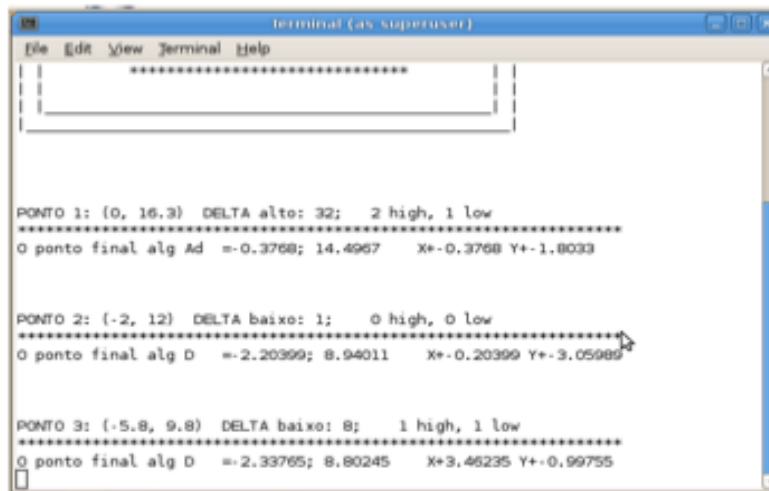
Tendo em foco a performance e o baixo consumo de memória, o *NodeJS* troca o suporte de *multithreading* por eventos de *I/O* assíncronos. Deste modo torna-se possível gerir processos dependentes de *I/O* bloqueante sem recorrer à tradicional solução que usa várias *threads* para tratar pedidos de várias fontes (como é o caso do nosso SG que recebe múltiplas ligações TCP/IP de vários clientes)[TV10]. Dado o peso do ambiente de execução (menos de 100KB no nosso caso) e a facilidade com que lidamos com vários pedidos de *I/O* - quer de rede TCP/IP, quer de acesso ao disco - aliado ao suporte para *Cassandra*, justificamos deste modo a nossa escolha por este sistema. A alternativa seria utilizar o tradicional *Apache PHP* aliado com uma implementação em *C++* para envio e receção de dados, e interação com a base de dados.

¹⁴<https://developers.google.com/v8/embed>

Capítulo 4

Resultados

Para o estudo comportamental do sistema foi desenvolvida uma aplicação de auxílio que apenas implementa o algoritmo de localização escolhido a um conjunto de dados de *input*. Esta aplicação recebe uma linha de texto contendo os valores de RSSI com que cada nó detetou os nós vizinhos, o RSSI com que cada nó detetou o dispositivo móvel e as coordenadas reais deste. O *output* é uma linha de texto com os resultados que o sistema colocaria na base de dados. Na figura 4.1 está um fotograma da aplicação referida.



```
terminal (as superuser)
*****

PONTO 1: (0, 16.3) DELTA alto: 32; 2 high, 1 low
*****
O ponto final alg Ad =-0.3768; 14.4967 X=-0.3768 Y=-1.8033

PONTO 2: (-2, 12) DELTA baixo: 1; 0 high, 0 low
*****
O ponto final alg D =-2.20399; 8.94011 X=-0.20399 Y=-3.05989

PONTO 3: (-5.8, 9.8) DELTA baixo: 8; 1 high, 1 low
*****
O ponto final alg D =-2.33765; 8.80245 X=3.46235 Y=-0.99755
```

Figura 4.1: *Screenshot* da aplicação que implementa o algoritmo de localização a um conjunto de dados

4.1 Local de estudo

O local de estudo é o andar de um edifício com 197 metros quadrados. É composto por 6 salas e o corredor que as serve. As paredes são constituídas por placas de aglomerado de madeira em ambas as faces. O seu interior é composto por espuma cuja finalidade é absorver som e uma placa de alumínio. No local foram instaladas três antenas de modo a cobrir metade do andar. A finalidade era testar a precisão do sistema em dois aspetos: quando os terminais se encontram dentro da área do triângulo criado pelas antenas e quando se encontram fora deste. Uma planta com as coordenadas das antenas de captura está representada na figura 4.2. A escala encontra-se em metros. Decidimos distinguir as áreas abrangidas recorrendo a três cores: a laranja colocamos a área de trabalho, a verde a sala de servidores e a azul a copa. Cremos ser importante distinguir os três espaços pois apresentam características que puderam influenciar os resultados. Isto porque, na sala de trabalho se encontram apenas dispositivos WiFi em modo cliente e vários BTs. Na sala dos servidores encontram-se três *Pontos de Acesso*. Na copa existem 2 micro-ondas.

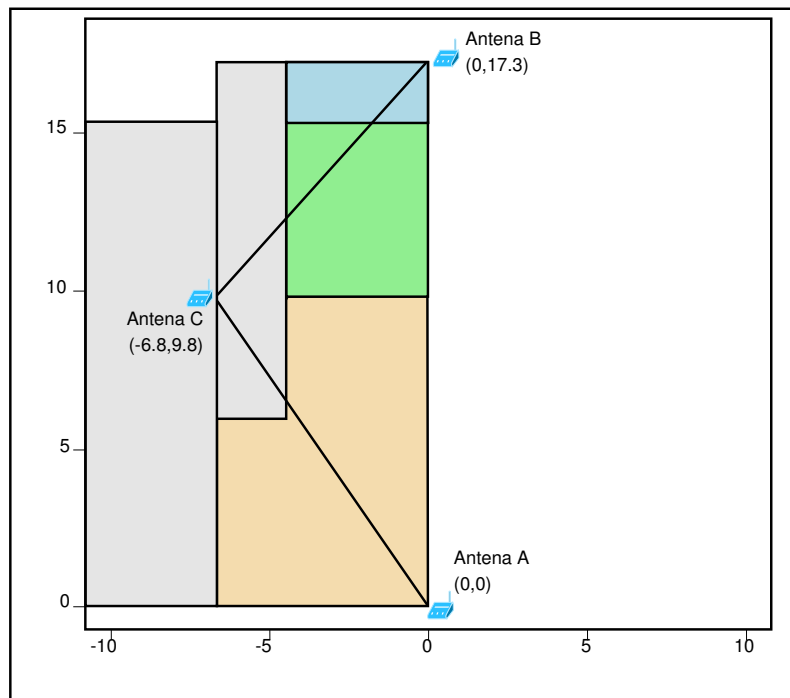


Figura 4.2: Planta do local de estudo

Este local situa-se num ambiente empresarial, no primeiro andar de um edifício de dois andares, partilhados entre várias empresas. Devido a este facto o espectro 2.4GHz está repleto de *APs* (pertencentes às várias empresas instaladas no edifício). Diariamente

podem ser contabilizadas mais de vinte SSIDs. Na parte do andar utilizado para os testes existem cerca de dez estações de trabalho equipadas com computadores maioritariamente *Apple* cujos periféricos de *input* (teclado e rato) comunicam com o computador sem fios através de Bluetooth.

Para realização dos testes foram escolhidos vários pontos de deteção espalhados pela área abrangida pelo alcance das antenas. Na figura 4.3 está uma representação do local de estudo, indicando as paredes a vermelho e evidenciando os pontos de deteção com etiquetas devidamente identificadas. O triângulo visível é a representação da área de deteção.

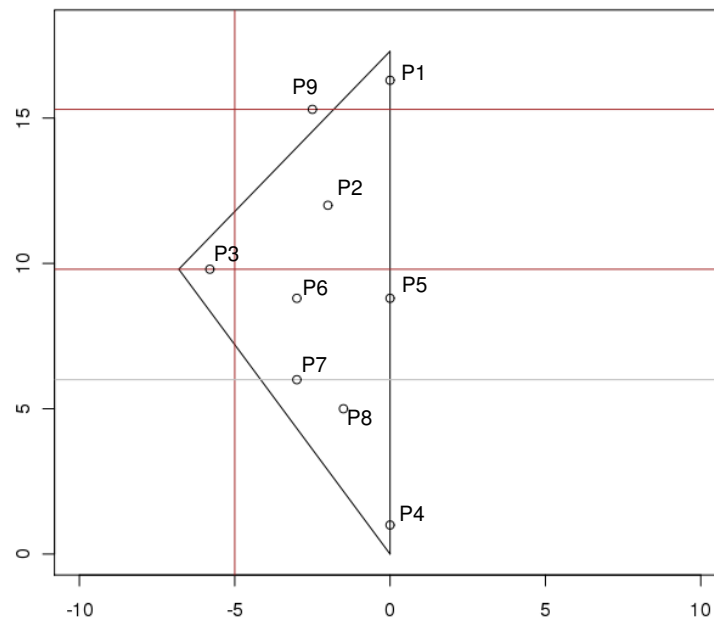


Figura 4.3: Localização dos pontos ($P1$ a $P9$) a estudar no local de teste

A escolha das coordenadas para a posição de cada um dos pontos teve em atenção dois critérios: os pontos $P1$, $P3$, $P4$ e $P5$ medir o comportamento nas faces da área de deteção; os pontos $P2$, $P6$, $P7$ e $P8$ medir o comportamento do algoritmo de localização numa área aproximadamente no centro do triângulo. O ponto $P9$ serviu para observar o comportamento do sistema caso detete pontos no exterior da área abrangida pelas antenas.

4.2 Resultados obtidos com Bluetooth

Utilizando a tecnologia BT observaram-se os seguintes resultados para os pontos de teste:

Ponto (x,y)	RSSI em A (dBm)	RSSI em B (dBm)	RSSI em C (dBm)	x ($metros$)	y ($metros$)	erro ($metros$)
$P1$ (0,16.3)	-76	-48	-80	-0.38	14.5	1.84
$P2$ (-2,12)	-61	-60	-60	-2.08	8.76	3.24
$P3$ (-5.8,9.8)	-67	-70	-62	-2.47	8.38	3.62
$P4$ (0,1)	-54	-79	-73	-0.47	2.14	1.23
$P5$ (0,8.8)	-65	-75	-65	-2.32	7.44	2.73
$P6$ (-3,8.8)	-57	-70	-42	-5.22	8.86	2.22
$P7$ (-3,6)	-55	-75	-66	-0.66	2.45	4.25
$P8$ (-1.5,5)	-57	-75	-68	-0.59	2.43	2.72
$P9$ (-2.5,15.3)	-76	-52	-76	-0.52	14.54	2.12

Tabela 4.1: Resultados Bluetooth

A primeira coluna da tabela contém a etiqueta e as coordenadas correspondentes ao ponto real onde se colocou o terminal móvel. Nas três colunas seguintes colocamos os valores de potência de sinal com que cada antena detetou o dispositivo. De seguida colocamos os valores em metros que o sistema de localização estimou para cada coordenada. A última coluna contém a distância entre o ponto real e o ponto estimado, a que demos o nome de erro, também em metros.

Para a recolha destes resultados a fase *offline* gerou os seguintes valores de RSSI:

	A	B	C
A	-	-50	-40
B	-48	-	-47
C	-37	-46	-

Tabela 4.2: Valores de RSSI da fase *offline* para os resultados BT

Para melhor se compreender os resultados obtidos foi desenvolvida uma interface para com a base de dados em *PHP* que consulta os dados recolhidos pelo sistema e ilustra-os sobre uma imagem de fundo. Os dados apresentados pela interface são compostos pelas recolhas disponíveis - consoante o algoritmo e a data da recolha. É colocado um ponto de cor negra representando o ponto real onde se situava o terminal móvel. A estimativa da localização dada pelo algoritmo escolhido é representado a cor. Na figura 4.4 é demonstrado o resultado para uma recolha efetuada em que se avalia o ponto P_4 utilizando os resultados obtidos com BT.

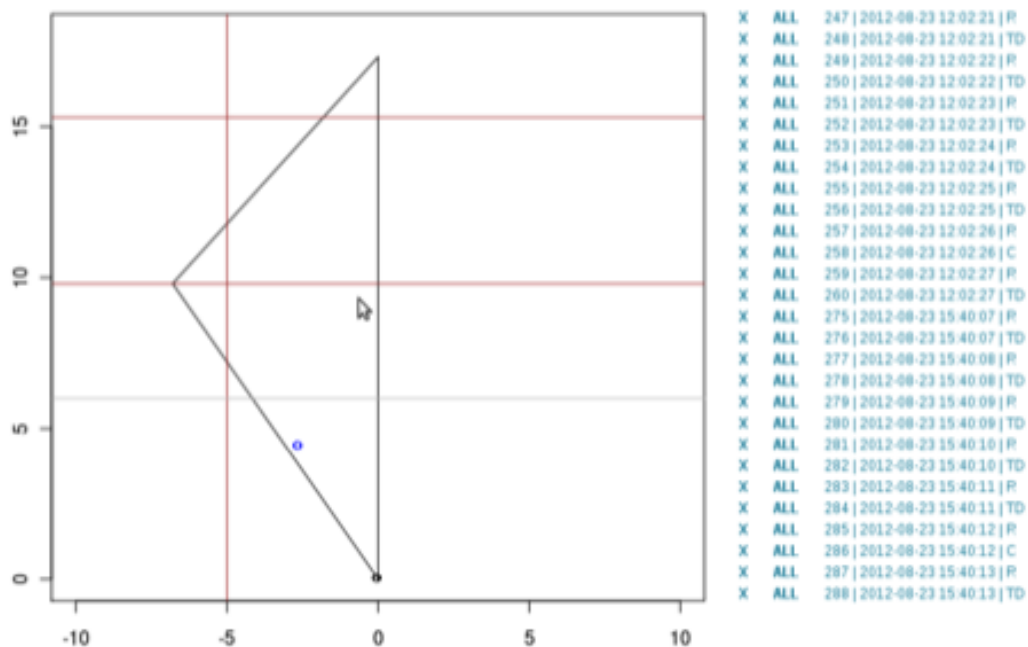


Figura 4.4: *Screenshot* da interface criada para estudar visualmente o comportamento do sistema

4.3 Resultados Obtidos com Wi-Fi

Utilizando os mesmos pontos usados em BT para a tecnologia WiFi observam-se as seguintes estimativas:

Ponto (x,y)	RSSI em A (dBm)	RSSI em B (dBm)	RSSI em C (dBm)	x (metros)	y (metros)	erro (metros)
$P1$ (0,16.3)	-65	-63	-67	-2.12	9.50	7.12
$P2$ (-2,12)	-66	-65	-67	-2.45	9.25	2.78
$P3$ (-5.8,9.8)	-64	-64	-65	-2.39	9.27	3.45
$P4$ (0,1)	-65	-63	-61	-1.09	4.59	3.75
$P5$ (0,8.8)	-65	-65	-66	-2.28	9.15	2.30
$P6$ (-3,8.8)	-66	-64	-64	-2.34	8.97	0.68
$P7$ (-3,6)	-63	-61	-61	-1.19	4.60	2.28
$P8$ (-1.5,5)	-63	-61	-59	-1.18	4.53	0.56
$P9$ (-2.5,15.3)	-64	-62	-65	-2.26	9.54	5.76

Tabela 4.3: Resultados Wi-Fi

A leitura da tabela 4.3 pode ser feita com a mesma interpretação do texto que descreve a tabela 4.1: em que a primeira coluna se pode ver o nome atribuído e respetivas coordenadas em metros do ponto de teste, nas três seguintes colunas o valor da potência de sinal com que cada antena detetou o dispositivo móvel, e as três colunas finais as coordenadas estimadas pelo sistema e o erro em relação às coordenadas reais, respetivamente e em metros.

Note-se que a fase *offline* para esta tecnologia é feita durante o funcionamento normal do sistema.

4.4 Resultados Obtidos com GSM

Devido a diretivas da União Europeia que vigoram em Portugal foi impossível realizar testes durante um período de tempo longo o suficiente para recolher dados que possam ser utilizados para localizar terminais móveis. Faremos de seguida uma descrição dos passos a realizar para obter um identificador único de um terminal GSM e respetivo RSSI e logo após expomos o motivo pelo qual a obtenção destes dados é proibida por

lei.

1. Observar o espectro da banda GSM-900 em procura dos ARFCNs em utilização pelas operadoras GSM
2. Escolher uma operadora alvo e respetivo ARFCN
3. Criar uma BTS fictícia que simule ser da mesma operadora anunciando para isso um canal com os mesmos parâmetros do ARFCN real
4. Recolher os pedidos de LUR que se seguirem

No ponto 3. aquando da criação da BTS fictícia no mesmo ARFCN que a BTS verdadeira estamos a incorrer numa violação da diretiva europeia. Esta diretiva impede qualquer tipo de emissão de sinais rádio na banda GSM. Por esse motivo, a alternativa implementada, apesar de obter dados que permitam localizar os terminais GSM não poderá ser utilizada. O resultado de tal tentativa seria a associação dos terminais móveis à BTS fictícia enquanto estes a avaliassem como tendo uma potência de sinal maior que qualquer outras das BTSs vizinhas. Devido a essa associação os terminais perderiam o acesso à rede da operadora, tornando-se efetivamente inúteis. Incluindo as chamadas de emergência, o que poderia ter consequências graves.

No entanto é possível testar a teoria criando uma rede extraordinária - para além das atualmente existentes - e estudar o processo de associação. Através desta BTS fictícia poderemos ver, conforme demonstrado na figura 4.5, que o telemóvel faz um pedido de ligação, é aceite pela BTS e os dados relevantes para a localização deste são recolhidos. Para o efeito será necessário um terminal cujo funcionamento não esteja bloqueado a uma operadora e forçar a associação à mesma - neste caso à criada para o efeito de localização. Isto pode ser feito através de uma listagem manual das redes GSM à disposição do terminal. A necessidade de forçar a associação deve-se ao facto de os fabricantes configurarem os seus terminais a se ligarem à rede predefinida no cartão SIM.

```

MobilityManagement.cpp:227:LocationUpdatingController: updating IMSI268034100254104 to IMEI:359534007608320
MobilityManagement.cpp:277:LocationUpdatingController: registration ALLOWED: IMSI=268034100254104
RadioResource.cpp:148:AccessGrantResponder: RA=0xb when=0:625391 age=8 delay=2.69922 RSSI=-9
RadioResource.cpp:228:AccessGrantResponder: sending PageMode=(0) DedicatedModeOrTBF=(TMA=0 Downlink=0 DMOrTBF=0) ChannelDescription=(typeAndOff
MobilityManagement.cpp:145:LocationUpdatingController: MM Location Updating Request LAI=(MCC=268 MNC=03 LAC=0x3e8) MobileIdentity=(IMSI=2680341
RadioResource.cpp:148:AccessGrantResponder: RA=0x3 when=0:625636 age=7 delay=2.03125 RSSI=-2
RadioResource.cpp:228:AccessGrantResponder: sending PageMode=(0) DedicatedModeOrTBF=(TMA=0 Downlink=0 DMOrTBF=0)
ChannelDescription=(typeAndOffset=SDCCH/4-2 TN=0 TSC=2 ARFCN=63) RequestReference=(RA=3 T1'=23 T2=24
MobilityManagement.cpp:215:LocationUpdatingController: MM Identity Response mobile id=IMEI=359534007608320
MobilityManagement.cpp:227:LocationUpdatingController: Updating IMSI268034100254104 to IMEI:359534007608320
MobilityManagement.cpp:277:LocationUpdatingController: registration ALLOWED: IMSI=268034100254104
MobilityManagement.cpp:145:LocationUpdatingController: MM Location Updating Request LAI=(MCC=268 MNC=03 LAC=0x2cec) MobileIdentity=(TMSI=0x4004
MobilityManagement.cpp:215:LocationUpdatingController: MM Identity Response mobile id=IMEI=355842020270990
MobilityManagement.cpp:227:LocationUpdatingController: Updating IMSI268031202594678 to IMEI:355842020270990
MobilityManagement.cpp:277:LocationUpdatingController: registration ALLOWED: IMSI=268031202594678
RadioResource.cpp:148:AccessGrantResponder: RA=0x1 when=0:626244 age=8 delay=1.71875 RSSI=-3
RadioResource.cpp:228:AccessGrantResponder: sending PageMode=(0) DedicatedModeOrTBF=(TMA=0 Downlink=0 DMOrTBF=0) ChannelDescription=(typeAndOff

```

Figura 4.5: *Screenshot* dos dados GSM recolhidos para estimação da localização do terminal móvel

Embora se consiga adquirir dados suficientes para a identificação exclusiva de cada terminal e a potência de sinal recebida pela antena de receção GSM da BTS criada, a impossibilidade de utilização mantém-se por parte da diretiva europeia. Mesmo que assim não fosse, consideramos este processo pouco prático e cuja operação se desvia dos nossos objetivos já que implicaria que o utilizador realizasse o processo de associação manualmente e carregasse consigo um dispositivo extra.

Capítulo 5

Discussão dos Resultados e Trabalho Futuro

Neste capítulo iremos analisar os dados recolhidos pelo sistema, tendo como principal foco as estimativas de localização obtidas. Iremos também referir as considerações que achamos importantes quanto ao *hardware* utilizado para as recolhas de deteções. Por último, tendo em conta estes dois pontos, iremos sugerir tarefas que poderão vir a otimizar o funcionamento deste sistema de localização.

5.1 Discussão dos Resultados

Os dados obtidos através de BT permitiram determinar a localização do dispositivo móvel com um erro que varia entre o valor mínimo de 1.23 metros (ponto P_4) e 4.25 metros (ponto P_7). A média do erro é de 2.66 metros.

Considerando os dados entre os dois conjuntos de pontos: nas faces da área de deteção (P_1 , P_3 , P_4 e P_5) e na área central do triângulo (P_2 , P_6 , P_7 e P_8), a média do erro é de 2.36 metros e de 3.11 metros, respetivamente. Podemos por isso reparar que o algoritmo obtém melhores estimativas para os pontos de deteção quando estes se encontram nas faces da área de deteção.

Será também de notar que - para os resultados do BT - os pontos que distam exactamente o mesmo das antenas, por exemplo, ponto P_1 da antena B, ponto P_3 da antena C e ponto P_4 da antena A, os valores de RSSI foram -48 , -62 , e -52 *dBm*, respetivamente. Este facto sugere que o ambiente onde as antenas são colocadas

influencia o valor do RSSI. Como a antena A se encontra numa sala ampla (a maior do local de estudo), a antena B numa sala pequena e a antena C num corredor, concluímos que a amplitude e organização estrutural dos locais influencia significativamente a propagação do sinal e consequentemente o RSSI.

Quanto ao pior resultado obtido (ponto *P7*, aquele que maior erro contém), acreditamos que este se deve à diferença de RSSI entre as antenas A e C. Dado que o ponto se encontra sensivelmente equidistante de A e C os valores de deteção que estas antenas obtiveram deveriam ter sido mais próximos um do outro. Devido à diferença de 11 *dBm* a estimativa para este ponto ficou distante da posição real onde o terminal móvel foi colocado. Não é fácil inferir sobre a diferença de potencia de sinal, mas julga-se que a movimentação de pessoas e dispositivos moveis no local de estudo, no momento em que ocorreram as deteções, influenciem o ambiente de propagação, piorando as condições de receção de sinal da antena C.

Relativamente à determinação dos pontos usando a tecnologia WiFi, os erros variam entre os 0.56 metros (*P8*) e os 7.12 metros (ponto *P1*). A média de erro, neste caso, é de 3.19 metros. Quanto aos dados entre os dois conjuntos de pontos: nas faces da área de deteção (*P1*, *P3*, *P4* e *P5*) e numa área central do triângulo (*P2*, *P6*, *P7* e *P8*), a média do erro é de 4.16 metros e de 1.58 metros, respetivamente. Temos, portanto, que a estimativa dos pontos de deteção foram, no caso desta tecnologia, melhores nos pontos centrais de deteção.

Observando os valores de RSSI na tabela dos resultados obtidos por BT podemos verificar uma amplitude de 32 *dBm* considerando o valor mínimo de -48 *dBm* e o máximo de -80 *dBm*, ambos do ponto *P9*. Observando as mesmas colunas na tabela dos resultados obtidos através de WiFi existe uma amplitude de apenas 8 *dBm*, considerando o valor mínimo de -59 e o máximo de -67 *dBm* - nos pontos *P8* e *P1* respetivamente. O baixo valor de amplitude de RSSIs no segundo conjunto de dados mostra que existe uma menor sensibilidade na medição da potência de sinal usando o sistema WiFi implementado neste trabalho. Este facto é agravado se considerarmos que o valor de 8 *dBm* tem em consideração todas as captações e que a média da amplitude para o mesmo ponto é de 2.6 *dBm* (num conjunto de valores que varia entre 1 e 4 *dBm*). Esta menor sensibilidade traduz numa maior dificuldade de obter estimativas com baixo erro (na ordem das obtidas com BT, explicando a maior média de erro dos resultados do Wi-Fi em relação aos dados do BT).

O facto de os valores de RSSI da tecnologia Wi-Fi serem tão próximos sugere, sob o ponto de vista do algoritmo, que os pontos estão aproximadamente equidistantes

das três antenas, posicionando-os no centro da área da detecção. Por este motivo, a diferença entre o conjunto de pontos do centro da área do triângulo é baixa (1.58 metros) em relação aos pontos nas faces da detecção (4.16 metros).

5.1.1 Sobre as Capturas

Dados os resultados pouco satisfatórios obtidos com o WiFi, provocados pela diferença - praticamente nula - com que cada antena via os terminais móveis ao longo dos pontos de teste, resolvemos experimentar se uma alteração das antenas poderia melhorar os resultados. Neste sentido, do nosso ponto de vista, achamos que os valores de RSSI obtidos não são determinísticos. Numa tentativa de combater este comportamento, resolvemos experimentar a troca das antenas anteriormente utilizadas (omni-direcionais com 3 *dbi* de ganho) para os testes WiFi pelas antenas direcionais disponíveis. O objetivo era exatamente, aumentar a diferença da potência de sinal recebida, tornando assim o sistema mais preciso. Decidimos também testar se um aumento do tempo de detecções para um valor grande (10 minutos) poderia influenciar os resultados de localização. Com isto estaremos a verificar se a variação do sinal WiFi se mantém ao longo de um período de tempo. Os resultados para estes dois testes foram encontrados na tabela 5.1.

Ponto (x,y) (em metros)	x estimado (em metros)	y estimado (em metros)	Erro (em metros)	Nota
$P6$ (-3,8.8)	-2.29	9.08	0.76	- antenas em paralelo com o limite do triângulo - com média de 10 segundos
$P6$ (-3,8.8)	-2.29	9.05	0.75	- antenas em paralelo com o limite do triângulo - sem média
$P6$ (-3,8.8)	-2.29	8.94	0.72	- antenas direcionadas para o centróide - sem média
$P6$ (-3,8.8)	-2.27	8.95	0.74	- antenas direcionadas para o centróide - com média de 10 minutos
$P6$ (-3,8.8)	-2.31	9.08	0.74	- antenas Omni-direcionais - sem média

Tabela 5.1: Resultados Wi-Fi com diferentes antenas e tempos de recolha

Na tabela supra mencionada relacionamos dois aspetos que, de acordo com os resultados apresentados, alteram o comportamento do algoritmo de localização. O primeiro

passa por utilizar as antenas direcionais *TP-LINK* e alternar a sua direção entre duas posições - no caso de a área de detecção ser composta por três antenas formando um triângulo. Deste aspeto podemos observar que as antenas estando direcionadas paralelamente com o limite do triângulo irão aumentar o erro das estimativas para a localização dos dispositivos móveis (linhas 1 e 2 da tabela). Quanto ao prolongamento do tempo de recolha para cada terminal detetado, podemos observar que se obtêm resultados ligeiramente piores caso este parâmetro seja aumentado (linhas 1 e 4), sendo melhor de um modo geral na não utilização de média (linhas 1, 2 e 3). Devido a estes resultados as detecções com WiFi passaram a ser efetuadas sem recorrer a qualquer tipo de média e utilizando as antenas omni-direcionais de 3 *dBi* (linha 5).

Numa tentativa de melhorar o desempenho da localização através de WiFi também foi experimentada a troca das antenas de 3 *dBi* utilizadas por antenas de 0 *dBi*. Com isto aumentamos a diferença de sinal com que as antenas se detetam, mas pioramos consideravelmente a capacidade de detecção de terminais móveis, sendo por vezes necessário aguardar um longo período de tempo até que os terminais móveis mais longínquos fossem detetados. Por este motivo, demos a alternativa como não sendo fiável para um sistema de localização.

5.2 Conclusões

Do ponto de vista do comportamento do algoritmo de localização utilizado podemos observar as seguintes conclusões:

- sempre que houverem alterações nas posições das antenas de localização as coordenadas das mesmas terão de ser atualizadas
- caso o ambiente mude ao ponto de influenciar largamente propagação de sinal, o algoritmo adaptar-se-à através da fase *offline*
- consoante a velocidade de deslocação esperada por parte dos terminais móveis o parâmetro da média para as detecções poderá ser ajustado, desde que se tenha em conta que quanto maior a velocidade menor terá de ser o parâmetro
- as características do espaço onde o sistema será instalado é um fator determinante na precisão das estimativas, dado que é imutável será aconselhável um estudo prévio para determinar os locais de instalação das antenas de detecção e a densidade destas (se possível)

Quanto à precisão do sistema, do que é do nosso conhecimento, obtivemos resultados a par de soluções de localização por BT que variam entre os 2,5 e os 4 metros. No que toca aos resultados obtidos com WiFi obtivemos um erro superior às que conhecemos, entre os 2 e os 3 metros. No entanto, o erro associado às deteções não deverá ser tido em consideração tendo em conta que as estimativas resultaram do fraco comportamento da fase *offline*, viciando assim os resultados.

5.3 Trabalho Futuro

O algoritmo que determina as coordenadas não está preparado para lidar com dispositivos que estejam fora da área triangular formada pelas três antenas. Isto significa que se um dispositivo estiver fora da área de deteção, mas perto o suficiente para ser detetados pelas três antenas, o software vai determinar o ponto final como estando dentro da área. Por este motivo, é necessário desenvolver uma forma que passa por utilizar uma quarta antena de deteção ou fazer a medição do RSSI das antenas fora da área de deteção de forma a ser excluído.

Face aos resultados do algoritmo usando as captações do WiFi, e principalmente devido à pouca sensibilidade da medição dos valores de RSSI, seria vantajoso o uso de uma área de deteção mais alargada. Com o aumento da área de deteção, haveria uma maior e justa diferença entre os RSSI, tornando as previsões do algoritmo mais precisas. O aumento da área de deteção também permitiria observar a partir de que distâncias as antenas deixariam de receber sinais rádio por parte dos terminais móveis, obtendo-se desta forma o alcance máximo do sistema. Sem aumentar a área de deteção, uma alternativa para melhorar os valores do RSSI na tecnologia WiFi seria utilizar atenuadores de sinais.

Do ponto de vista funcional em caso de falha do Servidor Local, poderia ser implementado um mecanismo de balanceamento de carga no Servidor Global de modo a que, desde que este esteja suficientemente disponível, aceitar as ligações (diretamente) dos vários Masters que estariam ao abrigo do SV em falha.

Do ponto de vista da melhoria da segurança do sistema, seria importante implementar um protocolo mais seguro para as trocas de mensagens. Mais concretamente, para aumentar a segurança ao nível da comunicação, poderia ser implementado um protocolo de troca de mensagens cujos cabeçalhos das mensagens incluíssem um *token* conhecido apenas dentro do sistema. Este serviria como segurança adicional, para além do sal

(*salt*) já incluído no protocolo SSL. No caso de um ataque *man-in-the-middle* este detalhe serviria de segurança adicional.

Referências

- [Bal07] R. Balani. Energy consumption analysis for bluetooth, wifi and cellular networks. *Online*. <http://nesl.ee.ucla.edu/fw/documents/reports/2007/PowerAnalysis.pdf>, 2007.
- [Bhi02] MS Bhigade. Secure socket layer. *Informing Science InSITE-“Where Parallels Intersect*, 2002.
- [Blo04] Eric Blossom. GNU radio: tools for exploring the radio frequency spectrum. *Linux J.*, 2004(122):4–, June 2004.
- [Blu99] BluetoothSIG. Specification of the bluetooth system REV1.00b, 1999.
- [Cat11] Rick Cattell. Scalable SQL and NoSQL data stores. *SIGMOD Rec.*, 39(4):12–27, May 2011.
- [CHRW08] Christer Carlsson, Kaarina Hyvönen, Petteri Repo, and Pirkko Walden. Adoption of mobile services across different technologies. 2008.
- [Cro06] Douglas Crockford. The application_json media type for JavaScript object notation (JSON). <http://tools.ietf.org/html/rfc4627>, 2006.
- [DDN08] M. Dickens, B. Dunn, and L.J. Nicholas. Design and implementation of a portable software radio. *IEEE Communications Magazine*, 46(8):58 –66, August 2008.
- [Dha07] S. Dhawan. Analogy of promising wireless technologies on different frequencies: Bluetooth, WiFi, and WiMAX. In *The 2nd International Conference on Wireless Broadband and Ultra Wideband Communications, 2007. AusWireless 2007*, page 14, August 2007.
- [ECC03] J.S. Esteves, A. Carvalho, and C. Couto. Generalized geometric triangulation algorithm for mobile robot absolute self-localization. In *2003 IEEE*

- International Symposium on Industrial Electronics, 2003. ISIE '03*, volume 1, pages 346 – 351 vol. 1, June 2003.
- [EM99] P. Enge and P. Misra. Special issue on global positioning system. *Proceedings of the IEEE*, 87(1):3 –15, January 1999.
- [Ett08] M. Ettus. Universal software radio peripheral (USRP). *Ettus Research LLC* <http://www.ettus.com>, 2008.
- [Fan86] Bertrand T Fang. Trilateration and extension to global positioning system navigation. *Journal of Guidance Control Dynamics*, 9:715–717, December 1986.
- [Fea10] D. Featherston. *Cassandra: Principles and application*. 2010.
- [FMKK09] J. Fink, N. Michael, A. Kushleyev, and V. Kumar. Experimental characterization of radio signal propagation in indoor environments with application to estimation and control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009. IROS 2009*, pages 2834 –2839, October 2009.
- [FMSG09] S. Fuicu, M. Marcu, B. Stratulat, and A. Girban. Effectiveness and accuracy of wireless positioning systems. *WSEAS Transactions on Computers*, 8(9):1471–1483, 2009.
- [Gar03] PD Garner. Mobile bluetooth networking: technical considerations and applications. In *3G Mobile Communication Technologies, 2003. 3G 2003. 4th International Conference on (Conf. Publ. No. 494)*, page 274–276, 2003.
- [GLN09] Yanying Gu, Anthony Lo, and Ignas Niemegeers. A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials*, 11(1):13–32, 2009.
- [Goo12] Google. Google’s v8 design elements, 2012.
- [GRH⁺06] U. Großmann, C. Röhrig, S. Hakobyan, T. Domin, and M. Dalhaus. WLAN indoor positioning based on euclidian distance and interpolation (isobars). In *Proceedings of the 8th Wireless Technologies Kongress*, page 296–305, 2006.
- [Haa03] Jaap C. Haartsen. Bluetooth radio system. In *Wiley Encyclopedia of Telecommunications*. John Wiley & Sons, Inc., 2003.

- [HHL11] Jing Han, E. Haihong, Guan Le, and Jian Du. Survey on NoSQL database. In *2011 6th International Conference on Pervasive Computing and Applications (ICPCA)*, pages 363 –366, October 2011.
- [Jic79] Todd D. Jick. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24(4):602, December 1979.
- [JP92] G.J.M. Janssen and R. Prasad. Propagation measurements in an indoor radio environment at 2.4 GHz, 4.75 GHz and 11.5 GHz. In *Vehicular Technology Conference, 1992, IEEE 42nd*, pages 617 –620 vol.2, May 1992.
- [Jue06] A. Juels. RFID security and privacy: a research survey. *IEEE Journal on Selected Areas in Communications*, 24(2):381 – 394, February 2006.
- [KB97] J.M. Kahn and J.R. Barry. Wireless infrared communications. *Proceedings of the IEEE*, 85(2):265 –298, February 1997.
- [Mon12] MongoDB Manual. FAQ: MongoDB fundamentals — MongoDB manual. <http://docs.mongodb.org/manual/faq/fundamentals/#what-are-the-32-bit-limitations>, 2012.
- [MP92] Michel Mouly and Marie-Bernadette Pautet. *The GSM System for Mobile Communications*. Telecom Publishing, 1992.
- [MT05] A. Madhavapeddy and A. Tse. A study of bluetooth propagation using accurate indoor location mapping. *UbiComp 2005: Ubiquitous Computing*, page 903–903, 2005.
- [OH11] Regina Obe and Leo Hsu. *PostGIS in Action*. Manning Publications Co., Greenwich, CT, USA, 2011.
- [OP05] B. O’hara and A. Petrick. *IEEE 802.11 handbook: a designer’s companion*. Standards Information Network, 2005.
- [PCB00] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, MobiCom ’00, page 32–43, New York, NY, USA, 2000. ACM.
- [PTVF92] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes in C Book, C version / William H. Press*. Cambridge Univ. Pr., Cambridge, 1992.

- [Sal04] A.E.S. Salazar. Positioning bluetooth reg; and wi-fi trade; systems. *IEEE Transactions on Consumer Electronics*, 50(1):151 – 157, February 2004.
- [Sri05] L Srivastava. Mobile phones and the evolution of social behaviour. *Behaviour & Information Technology*, 24(2):111–129, 2005.
- [Sto10] Michael Stonebraker. SQL databases v. NoSQL databases. *Commun. ACM*, 53(4):10–11, April 2010.
- [TB11] B.G. Tudorica and C. Bucur. A comparison between several NoSQL databases with comments and notes. In *Roedunet International Conference (RoEduNet), 2011 10th*, pages 1 –5, June 2011.
- [TC-96] ETSI TC-SMG. Digital cellular telecommunications system (Phase 2+); mobile radio interface layer 3 specification (GSM 04.08), 1996.
- [TV10] S. Tilkov and S. Vinoski. Node.js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing*, 14(6):80 –83, December 2010.
- [TW02] W.H.W. Tuttlebee and J. Wiley. *Software defined radio: enabling technologies*. Wiley, 2002.
- [VMG⁺01] K. Virrantaus, J. Markkula, A. Garmash, V. Terziyan, J. Veijalainen, A. Katanosov, and H. Tirri. Developing GIS-supported location-based services. In *Proceedings of the Second International Conference on Web Information Systems Engineering, 2001*, volume 2, pages 66 –75 vol.2, December 2001.
- [WHFG92] Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. The active badge location system. *ACM Trans. Inf. Syst.*, 10(1):91–102, January 1992.
- [WJH97] A. Ward, A. Jones, and A. Hopper. A new location technique for the active office. *Personal Communications, IEEE*, 4(5):42 –47, October 1997.
- [YR02] Youngjune Gwon and Ravi Jain. Error characteristics and calibration-free techniques for wireless LAN-based location estimation. In *In Proceedings of the Second international Workshop on Mobility Management & Wireless Access Protocols*, 2002.