

**CLASSIFICAÇÃO DE AMBIENTES PREDIAIS PARA USUÁRIOS
UTILIZANDO TAGS RFID ATIVAS E FILTRO DE KALMAN**

CRISTOVAM ALBANO DA SILVA JÚNIOR

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**CLASSIFICAÇÃO DE AMBIENTES PREDIAIS PARA USUÁRIOS
UTILIZANDO TAGS RFID ATIVAS E FILTRO DE KALMAN**

CRISTOVAM ALBANO DA SILVA JÚNIOR

Orientador: Adolfo Bauchpiess

Dissertação de Mestrado em
Engenharia Elétrica

Publicação: PGEA-DM 501/2012

BRASÍLIA, 21 DE AGOSTO DE 2012.

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**CLASSIFICAÇÃO DE AMBIENTES PREDIAIS PARA USUÁRIOS
UTILIZANDO TAGS RFID ATIVAS E FILTRO DE KALMAN**

CRISTOVAM ALBANO DA SILVA JÚNIOR

**DISSERTAÇÃO DE MESTRADO ACADÊMICO SUBMETIDA AO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE
TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM
ENGENHARIA ELÉTRICA**

APROVADA POR:

Prof. Dr. Ing. Adolfo Bauchpiess (orientador)
Dpto de Engenharia Elétrica / Faculdade de Tecnologia/ Universidade de Brasília

Prof. Dr. Marcelino Monteiro de Andrade (examinador externo)
Faculdade do Gama / Universidade de Brasília

Prof. Dr. Renato Alves Borges (examinador interno)
Dpto de Engenharia Elétrica / Faculdade de Tecnologia/ Universidade de Brasília

BRASÍLIA, 21 DE AGOSTO DE 2012.

FICHA CATALOGRÁFICA

SILVA JR, CRISTOVAM ALBANO DA

Classificação de Ambientes Prediais para Usuários Utilizando *Tags* RFID Ativas e Filtro de Kalman [Distrito Federal] 2012.

94p., 210 x 297 mm (ENE/FT/UnB, Mestre, Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia).

Departamento de Engenharia Elétrica

1. RFID *Tag* Ativa

2. Filtro de Kalman

3. Redes Neurais

4. SVM

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

SILVA JR., C. A. (2012). Classificação de Ambientes Prediais para Usuários Utilizando *Tags* RFID Ativas e Filtro de Kalman. Dissertação de Mestrado em Engenharia Elétrica. Publicação PGEA-DM 501/2012, Departamento de Engenharia Elétrica, Universidade de Brasília. Brasília – DF, 95p.

CESSÃO DE DIREITOS

AUTOR: Cristovam Albano da Silva Júnior

TÍTULO: Classificação de Ambientes Prediais para Usuários Utilizando *Tags* RFID Ativas e Filtro de Kalman.

GRAU: Mestre

ANO: 2012

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente com propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Cristovam Albano da Silva Júnior

Universidade de Brasília – UNB

Campus Darcy Ribeiro

Faculdade de Tecnologia – FT

Departamento de Engenharia Elétrica

CEP: 70910-900 – Brasília – DF

*À memória de minha avó
Arlinda Gomes da Silva,
matriarca da família Albano,
que de muitos foi referência. E
ao meu filho Frederico
Spolidoro Albano, de quem
espero sempre ser especial
referência.*

AGRADECIMENTOS

Não posso iniciar agradecendo senão a Deus, que em sua bondade paterna nos cerca de anjos para nos guardar em nossa caminhada. Alguns desses anjos têm nome, que agora cito, para de alguma forma poder retribuir toda ajuda e gentilezas proporcionadas nesta caminhada.

Aos professores que estiveram em Cuiabá para ministrar os créditos: João Ishihara, Leonardo Menezes, Franklin Silva, Francisco Damasceno, Kleber Melo, Adson Rocha, Fabiano Soares, Alexandre Romariz, Ivan Camargo, Rafael Shayani, todos capitaneados pelo professor Marco Aurélio, coordenador do projeto Minter UnB/IFMT. À equipe da secretaria do Departamento e da Pós Graduação.

Ao professor Adolfo, que me acolheu como a um amigo da sua família. E que sempre fora minha escolha pessoal para orientação neste trabalho, mesmo antes da confirmação de minha seleção para este mestrado. Meu sincero agradecimento professor Adolfo, pelo crédito a minha capacidade de empreender uma pesquisa acadêmica. Continuarás sendo minha referência de pesquisador, onde quer que eu esteja trilhando meu caminho.

Ao professor Marcelino, por aceitar o convite de participar de minha banca avaliadora e por contribuir com sugestões ao meu trabalho.

A toda equipe do LARA, liderada pelo professor Geovany Borges, em especial ao amigo Henrique Menegaz, que desde o primeiro dia em que aportei no LARA me recebeu como a um irmão. Também ao Felipe Brandão, ao Roberto, à Cláudia Uchoa, ao professor Renato Borges, ao professor Antonio Padilha, ao professor Henrique Ferreira, sempre solícitos no período em que por lá permaneci.

Aos companheiros Figueiró, Gama e Vinicius, com quem tive o prazer de “perder” algumas noites de sono, fazendo experimentos com RFID, e também o prazer de “ganhar” muito conhecimento. A dedicação de vocês na pesquisa de seus trabalhos, me serviu de exemplo e força para superar minhas adversidades e limitações.

Ao colega Engenheiro Lucas Fonseca, precursor dos trabalhos com RFID no LARA, quando aluno da UnB, e que sempre foi prestativo às minhas solicitações e indagações acerca do assunto.

Ao amigo Nonato, que por 4 meses me aturou na Colina (apartamento de trânsito), sendo muitas vezes também um anjo da guarda a olhar por mim, enquanto estive em Brasília.

Aos amigos que fiz na Colina, cujos nomes não vou citar porque são muitos, mas que todos saibam que estão no meu coração, nas minhas orações e até no meu Facebook.

Ao Claudionel Campos Leite, que me proporcionou um apoio crucial nos primeiros dias em Brasília. Espero um dia poder lhe retribuir tamanha disponibilidade meu amigo. Que Deus o abençoe.

Aos colegas de Minter... uma lágrima me vem aos olhos, quando neste momento me lembro do que juntos passamos. Obrigado por toda a força.

Ao amigo Everaldo, que sempre me arrastou com seu empenho e dedicação a este mestrado e, além disso, sempre disponibilizou um espaço em sua sala para eu poder ler os artigos e escrever minha dissertação.

Um obrigado especial a Márcia Jani Cícero, que no auge de meu desespero, por não conseguir escrever a dissertação, me trouxe a luz, pronunciando palavras que eu vivia a repetir a todos, sem contudo assimilar o que me afligia: “Vai para Brasília, procure o seu orientador, ele vai te direcionar”. Muito obrigado Márcia.

Aos amigos Tony Inácio e Mário Anderson, irmãos de caminhada.

Aos meus pais Cristovam e Elzita, que sempre apoiaram também minha caminhada acadêmica. Às minhas tias Teolinda e Anaídes, que só pela torcida me encheram de coragem para avançar na empreitada.

A minha esposa Marli, por todo apoio, desprendimento e coragem de encarar um marido chato, desmotivado, travado (sem conseguir escrever), sempre dedicando amor e carinho, buscando motivação a todo instante para estar junto comigo nesta caminhada. Querida este trabalho tem muito de você, de suas orações, de suas abstenções, de nossa vida. Ele é parte de nossa meta para 2026. Amo você e nosso filho Frederico.

Senhor Jesus, eis o trabalho que é seu, e não meu. Amém!

RESUMO

CLASSIFICAÇÃO DE AMBIENTES PREDIAIS PARA USUÁRIOS UTILIZANDO TAGS RFID ATIVAS E FILTRO DE KALMAN

Autor: Cristovam Albano da Silva Júnior

Orientador: Prof. Dr. Ing. Adolfo Bauchspiess, Departamento de Engenharia Elétrica / Faculdade de Tecnologia / Universidade de Brasília

Programa de Pós-Graduação em Engenharia de Sistemas Eletrônicos e de Automação Brasília, 21 de agosto de 2012.

RFID e outras tecnologias sem fio têm sido objeto de estudo em aplicações de localização de usuários em ambientes internos de edificações. Existem diferentes formas de se prover a localização com as variáveis disponibilizadas por estes sistemas, sendo uma das mais populares, a utilização do RSSI como referência, pois esta grandeza possui relação com a distância entre o transmissor e o receptor do sinal. Entretanto, promover a localização em ambientes internos de edificações com o RSSI não é trabalho tão trivial, pois o sinal RF sofre interferências diversas, próprias do canal de comunicação RF em ambientes internos e de difícil modelagem, durante seu trajeto, ocasionando muita incerteza nos valores registrados pelos equipamentos que compõe o sistema RFID. Elas podem ser minimizadas através da aplicação de ferramentas computacionais como o filtro de Kalman e as redes neurais MLP e LVQ ou as SVM, treinadas com amostras de sinal que representem de forma qualitativa a área em que será provida a localização. O presente trabalho compara o desempenho destas ferramentas computacionais aplicadas a um banco de dados coletado de um sistema de RFID com *tags* ativas disponibilizado no LARA para experimentos. Tal sistema será usado para a classificação de ambientes delimitados do LARA, onde estarão usuários portadores das *tags* RFID, para assim controlar o sistema de climatização daquele laboratório.

Palavras chave: RFID, Tag Ativa, Filtro de Kalman, Classificação de Ambientes, MLP, LVQ, SVM

ABSTRACT

CLASSIFICAÇÃO DE AMBIENTES PREDIAIS PARA USUÁRIOS UTILIZANDO TAGS RFID ATIVAS E FILTRO DE KALMAN

Author: Cristovam Albano da Silva Júnior

Advisor: Prof. Dr. Ing. Adolfo Bauchspiess, Departamento de Engenharia Elétrica / Faculdade de Tecnologia / Universidade de Brasília

**Programa de Pós-Graduação em Engenharia de Sistemas Eletrônicos e de Automação
Brasília, august 21, 2012**

RFID and others technologies have been object of study in applications of indoor localization. There are different ways of providing localization with the variables provided by these systems, one of the most popular is the use of RSSI as the reference, since it has relation with the distance between the transmitter and the receiver of the signal. However, to provide the indoor localization with RSSI is not so trivial, because the RF signal suffers from various interferences, own of RF indoor communication channel and its very difficult modeling, during its course, causing much uncertainty on the registered values of RSSI by the equipments of RFID system. These uncertainties can be minimized by the application of some computational tools such as Kalman Filter and classifiers as MLP or LVQ neural networks, or the SVM, trained with signal samples that represent in a qualitative form the area in that the localization will be provided. This work compares the performance of this computational tools applied to the data collected from an RFID system with active tags, available in LARA to the experiments, and that will be used to classify some delimited environment of LARA, where users with RFID tags will be, and then control the HVAC system of that laboratory.

Word keys: RFID; Active tag; Kalman filter; Environment Classification; MLP; LVQ; SVM

SUMÁRIO

RESUMO	vii
ABSTRACT	viii
GLOSSÁRIO	xi
LISTA DE FIGURAS	xv
LISTA DE TABELAS	xvii
1 – INTRODUÇÃO	1
1.1 – MOTIVAÇÃO	2
1.2 – PROBLEMÁTICA	3
1.3 – OBJETIVOS	5
1.3.1 – Objetivo Geral	5
1.3.2 – Objetivos Específicos	5
1.4 – APRESENTAÇÃO DO TRABALHO	5
2 – MÉTODOS DE LOCALIZAÇÃO INDOOR	7
2.1 – SISTEMA RFID	7
2.1.1 – As incertezas do RSSI	10
2.2 – SISTEMA GPS-Indoor	13
2.3 – SISTEMA UWB	17
2.4 – OUTROS SISTEMAS DE LOCALIZAÇÃO INDOOR	18
2.4.1 – Sistemas de Localização Baseada em Redes Sem Fio padrão 802.11 (WIFI)	19
2.4.2 – Sistemas de Localização Baseada em Redes de Sensores Sem Fio (WSN)	21
2.4.3 – Sistemas de Localização Baseada em Redes Bluetooth	22
3 – MÉTODOS DE FILTRAGEM E CLASSIFICAÇÃO DE DADOS	24
3.1 – FILTRO DE KALMAN ESTENDIDO (EKF)	25
3.2 – MULTI LAYER PERCEPTRON (MLP)	29
3.3 – LEARNING VECTOR QUANTIZATIONS (LVQ)	32
3.4 – SUPPORT VECTOR MACHINE (SVM)	33
4 – MATERIAIS E MÉTODOS	37
4.1 – MATERIAIS	37
4.1.1 – Hardware do Sistema RFID Proposto	38
4.1.2 – Middleware do Sistema RFID Proposto	41
4.1.3 – Neural Network Toolbox do MatLab®	43
4.1.4 – LIBSVM: Biblioteca de Máquina de Vetor de Suporte	45
4.2 – MÉTODOS	46
4.2.1 – Definição de Espaços, Ambientes e Pontos de Coleta de Dados	46
4.2.2 – Definição de Posições e Formas de Instalação das Leitoras	47

4.2.3 – Leitura e Pré-tratamento dos Dados.....	50
4.2.4 – Preparação dos Arquivos de Treinamento e Teste dos Classificadores.....	57
4.2.5 – Treinamento dos Classificadores	59
5 – APRESENTAÇÃO E ANÁLISE DOS RESULTADOS	62
5.1 – ORIENTAÇÃO VERTICAL DA ANTENA DA LEITORA.....	62
5.2 – APLICAÇÃO DO FILTRO DE KALMAN ESTENDIDO	64
5.3 – TREINAMENTO DAS REDES MLP.....	73
5.4 – TREINAMENTO DAS REDES LVQ.....	75
5.5 – TREINAMENTO DAS SVM.....	79
5.6 – COMPARAÇÃO DO DESEMPENHO DE TODOS OS CLASSIFICADORES...	82
6 – CONCLUSÃO	86
6.1 – TRABALHOS FUTUROS	89
REFERÊNCIAS BIBLIOGRAFICAS.....	90

GLOSSÁRIO

- A-GPS – Do inglês, *Assisted – Global Positioning System*. É um sistema de suporte ao sistema GPS, que auxilia os aparelhos de telefonia celular a determinar sua localização em ambientes onde não possuem visadas diretas aos satélites do sistema GPS.
- Age Counter* – Termo utilizado para designar um contador de pulsos que relativiza o tempo de vida útil da *tag*.
- AOA – Do inglês, *Angle Of Arrive*. É uma metodologia de triangulação aplicada a localização de usuários de sistemas RF.
- Arquivo de log – Arquivo onde são armazenados dados coletados de forma *online*, em tempo real.
- AP – Do inglês, *Access Point*. É um equipamento roteador que fornece acesso a uma rede IP de forma não cabeada (*wireless*).
- Bluetooth* – Tecnologia de comunicação de dados aplicável às denominadas rede WPAN (*Wireless Personal Area Network*).
- CDMA – Do inglês, *Code Division Multiple Access*. É uma técnica de multiplexação de dados aplicada na comunicação digital.
- COMPASS – É a denominação do sistema chinês de navegação por satélite. É uma evolução de um antigo sistema denominado BeiDou.
- CRC – Do inglês, *Cyclic Redundance Check*. É um algoritmo que permite fazer a checagem de um pacote de dados armazenado ou transmitido, evitando assim a interferência de ruído no canal.
- CSV – Do inglês, *Comma-separated Values*. É um formato de arquivo de texto que armazena dados tabelados.
- dB – É uma unidade logarítmica que indica a proporção de uma quantidade de energia como, por exemplo, a potência de determinado sistema, em relação a um nível de referência especificado, ou implícito.
- DOS – Do inglês, *Disk Operating System*. É um antigo sistema operacional de computadores do tipo IBM-PC.
- E-911 – É a denominação do serviço de atendimento de emergência norte americano, cujo acesso é disponibilizado através dos serviços de telefonia fixa e móvel.
- EKF – Do inglês, *Extended Kalman Filter*. É uma variante do filtro de Kalman aplicado a sistemas não lineares.

- EPE – Empresa de Pesquisa Energética, instituição vinculada ao Ministério de Minas e Energia, responsável pela pesquisa e planejamento do consumo de energia do Brasil.
- EZ – É a denominação de um algoritmo, cuja finalidade, é de fazer a comunicação entre um dispositivo móvel e um servidor de localização, em um sistema de localização baseado em WIFI.
- FCC – Do inglês, *Federal Communications Committee*. É um órgão Norte Americano, responsável pela normatização do setor de comunicações, naquele país.
- Fingerprint* – Denominação dada às *tags* distribuídas em posições previamente determinadas, para serem utilizadas como referências dentro de um ambiente em que se quer fazer localização.
- GALILEO – É a denominação do sistema europeu de navegação por satélite.
- GLONASS – É a denominação do sistema russo de navegação por satélite.
- GNSS – Do inglês, *Global Navigation Satellite System*. É uma denominação genérica para se referir aos sistemas de localização baseados redes de satélites estacionários.
- GPS – Do inglês, *Global Positioning System*. É a denominação do sistema norte americano de navegação por satélite.
- ID – Do inglês, *Identification*. Termo que se refere a um número de registro, para se identificar um usuário ou objeto de determinado sistema.
- IEEE – Do inglês, *Institute of Electrical and Electronics Engineers*. É uma entidade profissional sem fins lucrativos que congrega engenheiros eletricitas de todo o mundo.
- Indoor* – Termo utilizado para denominar ambientes internos de edificações.
- ITU – Do inglês, *International Telecommunication Union*. É uma organização internacional destinada a padronizar equipamentos e processos de telecomunicações.
- KF – Do inglês, *Kalman Filter*. É um algoritmo que opera de forma recursiva, utilizado para filtrar sinais de sistemas lineares.
- LANDMARC – É um sistema de localização baseado em RFID com *tags* ativas.
- LARA – Laboratório de Automação e Robótica. É um laboratório de pesquisa, do Departamento de Engenharia Elétrica – Faculdade de Tecnologia da Universidade de Brasília.
- LIBSVM – É uma biblioteca de software livre, composta por aplicativos de treinamento, teste e aplicação de SVM.

- LOCATA – Denominação de um sistema de localização para ambientes *indoor*, que simula o sinal de satélites da rede GPS, dentro destes ambientes.
- LVQ – Do inglês, *Learning Vector Quantization*. É um tipo de rede neural auto organizável com aprendizagem supervisionada.
- MAC *address* – Termo utilizado para designar um número de identificação de uma interface de rede de comunicação IP.
- MatLab® – É um software interativo de alto desempenho voltado para o cálculo numérico de matrizes.
- Middleware* – Termo utilizado para designar um software com função de interligar as leitoras RFID com o sistema de processamento dos dados fornecidos por estas leitoras, num sistema de localização por RFID.
- MLP – Do inglês, *Multi Layer Perceptron*. É um tipo de rede neural, com aprendizagem supervisionada, aplicável em reconhecimento de padrões e classificação de dados.
- MME – Ministério de Minas e Energia da República Federativa do Brasil.
- NIC – Do inglês, *Network Interface Controller*. Em um computador, é a parte hardware da interface de rede.
- Off line* – Termo que significa desligado.
- PDA – Do inglês, *Personal Digital Assistant*. É um computador de dimensões reduzidas, que facilita sua portabilidade. É o precursor dos Smart Phones.
- PNE-2030 – Plano Nacional de Energia 2030. É um documento emitido pela EPE, onde são feitas projeções de consumo de energia do Brasil e a previsão de atendimento destas demandas de consumo.
- RBF – Do inglês, *Radial Base Function*. Neste trabalho é uma forma de configuração do núcleo interno de uma SVM.
- RF – Do inglês, *Radio Frequency*. É uma faixa de frequência que abrange 3 KHz até 300 GHz e que corresponde a frequência das ondas de rádio.
- RFID – Do inglês, *Radio Frequency Identification*. É um sistema de identificação baseado em emissão de ondas de rádio.
- RSSI – Do inglês, *Received Signal Strength Indicator*. É um indicador do nível de sinal percebido na antena do receptor de rádio.
- SCDM – Do inglês, *Signal Coverage Density Method*. É um algoritmo aplicado a técnicas de localização para melhorar o desempenho destas técnicas.

- SPKF – Do inglês, *Sigma Point Kalman Filter*. É uma técnica de filtragem recursiva baseada no UKF.
- SVM – Do inglês, *Support Vector Machine*. É uma ferramenta computacional para classificação de dados.
- Tag – Termo utilizado para designar a parte móvel do sistema RFID, que geralmente acompanha o ente ou item a ser localizado.
- TOA – Do inglês, *Time Of Arrive*. É uma metodologia de triangulação aplicada a localização de usuários de sistemas RF.
- TWh – Tera Watt hora. É uma unidade de medida do consumo de energia elétrica.
- UKF – Do inglês, *Uncented Kalman Filter*. É uma técnica de filtragem recursiva baseada na transformada *unscented*.
- UWB – Do inglês, *Ultrawide Band*. É uma técnica transmissão de dados baseada no espalhamento espectral da mensagem.
- VIRE – Do inglês, *Virtual Reference Elimination*. É um sistema de localização derivado do LANDMARC.
- WIFI – É uma marca registrada da *Wi-Fi Alliance* utilizada por produtos certificados que pertencem a classe de dispositivos de rede local sem fio (WLAN) baseados no padrão IEEE 802.11.
- Wireless – Termo utilizado para denominar tecnologias de conexão de dados sem fio.
- WLAN – Do inglês, *Wireless Local Area Network*. É uma de rede sem fio baseada no padrão IEEE 802.11.
- WPAN – Do inglês, *Wireless Personal Area Network*. É uma rede sem fio baseada no padrão IEEE 802.15.
- WSN – Do inglês, *Wireless Sensor Network*. É uma rede de comunicação de dados formada por vários nós com sensores acoplados a estes nós.

LISTA DE FIGURAS

Figura 2.1 – Enfraquecimento do sinal em pequena e larga escala de deslocamento	11
Figura 2.2 – Comportamento do RSSI em teste de longa duração (48 horas)	13
Figura 3.1 – Arquitetura de uma rede MLP com duas camadas ocultas	30
Figura 3.2 – Arquitetura de uma rede neural do tipo LVQ	32
Figura 3.3 – Exemplo de classificação de dados utilizando LVQ	33
Figura 3.4 – Padrões linearmente separáveis e hiperplano definido pelos vetores de suporte	34
Figura 3.5 – Arquitetura de uma SVM com núcleo de produto linear	35
Figura 4.1 – Equipamentos RFID da Wavetrend: (a) leitora modelo L-RX900 e (b) antena modelo L-AN200	39
Figura 4.2 – Curva do comportamento do RSSI pela distância em campo aberto ..	40
Figura 4.3 – <i>Tag</i> Ativa modelo L-TG501 MS da Wavetrend	41
Figura 4.4 – Tela de apresentação do <i>middleware</i> utilizado na coleta de dados	42
Figura 4.5 – Tela da interface gráfica de treinamento e teste da <i>Neural Network Toolbox</i> TM	44
Figura 4.6 – Leiaute da área experimental definida no LARA, com indicação dos pontos de coleta de dados	47
Figura 4.7 – Leiaute da área experimental definida no LARA, com indicação do posicionamento das leitoras	48
Figura 4.8 – Fotos das leitoras instaladas nos ambientes do LARA	49
Figura 5.1 – Comparativo dos valores médios de RSSI coletados com antenas posicionadas verticalmente em sentidos inversos	63
Figura 5.2 – Comportamento do RSSI nos dados coletados no ponto 04 pela leitora 1, (a) dados sem tratamento metodologia 1, (b) dados tratados pela metodologia 2 e (c) dados tratados pela metodologia 3	65
Figura 5.3 – <i>Curve Fitting</i> do modelo de propagação utilizado no EKF	67
Figura 5.4 – Valores de RSSI coletados pela leitora 3 nos pontos (a) 10, (b) 11 e (c) 14	69

Figura 5.5 – Valores de RSSI coletados pela leitora 0 nos pontos (a) 05 e (b) 06 ..	70
Figura 5.6 – Exemplo de comportamento do RSSI nas seis bases de dados disponibilizadas para o treinamento dos classificadores. (a) Base de dados da metodologia 1, (b) Base de dados da metodologia 2 e (c) Base de dados da metodologia 3	72
Figura 5.7 – Desempenho do treinamento das redes MLP	74
Figura 5.8 – Desempenho do treinamento das redes LVQ	76
Figura 5.9 – Comportamento do erro quadrático médio durante o treinamento (a) Rede LVQ com 100 neurônios na camada competitiva, base de dados Método_3E, (b) Rede LVQ com 10 neurônios na camada competitiva, base de dados Método_1E e (c) Rede LVQ com 200 neurônios na camada competitiva, base de dados Método_1E	78
Figura 5.10 – Tela de um procedimento de treinamento e teste das SVM através da LIBSVM com os aplicativos para MS-Windows™	80
Figura 5.11 – Desempenho do treinamento das SVM	81
Figura 5.12 – Melhor desempenho dentre os classificadores para cada base de dados de treinamento e teste	85

LISTA DE TABELAS

Tabela 2.1 – Possíveis aplicações da tecnologia RFID	08
Tabela 2.2 – Possíveis aplicações da tecnologia UWB	18
Tabela 4.1 – Exemplo de dados gerados no arquivo de log pelo <i>middleware</i>	50
Tabela 4.2 – Exemplo de falha do <i>middleware</i> no registro do ID da leitora	51
Tabela 4.3 – Exemplo de correção da falha do <i>middleware</i> no registro do ID da leitora	52
Tabela 4.4 – Exemplo de pulso não registrado em determinada leitora mesmo com nível de sinal acima do limite inferior de corte	53
Tabela 4.5 – Metodologias para definição do valor de RSSI não registrado pelas leitoras	54
Tabela 4.6 – Exemplo de disposição dos dados nos arquivos de treinamento das redes neurais	58
Tabela 4.7 – Exemplo de disposição dos dados nos arquivos de treinamento da SVM pela ferramenta computacional LIBSVM	59
Tabela 5.1 – Valores médios de RSSI apurados nas leitoras 2 e 4	63
Tabela 5.2 – Conjunto de dados gerados para o treinamento e teste dos classificadores e os tratamentos recebidos pelo seu conteúdo	68
Tabela 5.3 – Configurações das redes MLP utilizadas nos experimentos	73
Tabela 5.4 – Desempenho das redes MLP treinadas e quantidade de épocas necessárias para alcançar este desempenho	75
Tabela 5.5 – Configurações das redes LVQ utilizadas nos experimentos	76
Tabela 5.6 – Desempenho das redes LVQ treinadas	79
Tabela 5.7 – Desempenho das SVM treinadas e quantidade de vetores de suporte separados	82
Tabela 5.8 – Comparação do desempenho de todos os classificadores testados	83

1 – INTRODUÇÃO

A localização de usuários em ambientes prediais se faz necessária em muitas aplicações do chamado ambiente inteligente. Com a informação da localização dos usuários destes espaços, seja ela expressa em coordenadas ou simplesmente expressa em porções previamente delimitadas do local, é possível prover aplicações que ofereçam ao usuário conforto, segurança e até mesmo auxílio em casos de emergência.

Para se prover a localização de tais usuários encontra-se disponível hoje varias tecnologias, todas com princípios de funcionamento, necessidades de infraestrutura e precisão de resultados diferentes umas das outras. Isto implica numa difícil escolha, posto que torna-se necessário conhecer muitas destas tecnologias disponíveis para melhor avaliar um bom atendimento a demanda de localização.

Adicione a este fato a possibilidade de utilização de diferentes ferramentas computacionais para se fazer o tratamento dos dados dispostos pelos sistemas de localização, daí se obtém novamente outra gama de resultados de desempenho, possibilitando o surgimento de outras aplicações antes não pensadas.

No âmbito deste contexto o LARA vem promovendo pesquisas em torno de tecnologias aplicáveis a localização de usuários em ambientes prediais, para se prover o conforto térmico destes locais e também a racionalização de consumo de energia pelos sistemas de climatização de edifícios. Neste caminho, em 2010 o Prof. Adolfo Bauchspiess, pesquisador do LARA, conseguiu a aprovação de um projeto de pesquisa junto a Fundação de Apoio a Pesquisa do DF – FAPDF, que tem por finalidade estudar e aplicar sistemas sem fio baseados em RFID, na racionalização do uso da energia elétrica e na busca do conforto térmico de usuários de edificações, de forma automática.

Os trabalhos de pesquisa deste projeto se iniciaram com Fonseca [1], que começou pelo dimensionamento dos equipamentos do sistema de RFID. O planejamento inicial do projeto previa que tal sistema deveria prover dupla função, sendo a primeira o controle de acesso dos usuários ao recinto do LARA e segundo, prover dados para um sistema

de localização que controlaria o uso do sistema de climatização do ambiente, podendo também se estender ao sistema de iluminação.

O sistema escolhido para tais funções foi o de fabricação da Wavetrend, que possui dentre outras facilidades, o formato de crachá para as *tags* ativas RFID. Tal formato facilita aos usuários seu uso, posto que dependerão de umas destas *tags* para manterem seu ambiente de trabalho devidamente climatizado.

O trabalho de Fonseca [1] encerrado em 2010 limitou-se a um estudo mais aprofundado dos equipamentos que compunham o Sistema de RFID, buscando seu entendimento e identificação de propriedade e limitações. Fonseca [1] desenvolveu ainda uma interface de comunicação para coleta de dados do Sistema de RFID, um *middleware*, e o utilizou para seus experimentos, que proporcionaram alguns ensaios de localização que à época se pretendeu fazer em coordenadas cartesianas.

O presente trabalho se inicia logo após o período empreendido por Fonseca [1], dando continuidade ao projeto RFID do LARA e buscando contribuir com os estudos necessários à implementação de um sistema de tal complexidade. Junto com as atividades referentes ao presente trabalho, ocorreram também outras atividade ligadas ao projeto RFID, como o caso do trabalho de Souza e Oliveira [2], que fizeram ricas contribuições ao mesmo.

1.1 – MOTIVAÇÃO

O Plano Nacional de Energia (PNE-2030), documento base que dispõe sobre o planejamento energético do Brasil, elaborado pela Empresa de Pesquisa Energética (EPE), empresa pública, vinculada ao Ministério de Minas e Energia, prevê uma evolução no consumo de energia elétrica de 375 TWh, em 2005, para valores entre 850 e 1.250 TWh, em 2030, dependendo somente dos cenários macro-econômicos que ocorrerão neste período, e da postura do Brasil perante estes cenários [3].

É sem dúvida um crescimento expressivo que, dentre outras questões, considera um aumento considerável também da eficiência no consumo da energia, principalmente no consumo da energia elétrica, único segmento que teve destaque para este

item no documento citado. Isto denota a importância de estudos que venham prover o conhecimento, o desenvolvimento e a aplicação de técnicas que buscam alcançar o uso mais eficiente da energia elétrica, seja em ambientes industriais, prediais ou residenciais.

Neste contexto, o presente estudo buscou, através da aplicação de técnicas de localização em ambientes internos de edificações, desenvolver subsídios para se implementar o controle de sistemas de climatização e de iluminação de ambientes prediais, provendo assim maior eficiência no uso destes sistemas.

O trabalho foi concebido utilizando-se um sistema de localização baseado em tecnologia RFID com *tag* ativa, sistema este fabricado pela empresa Wavetrend e dimensionado por Fonseca [1], em seu trabalho de graduação, para atuar como parte do sistema de controle de climatização do LARA. A idéia de se usar o RFID com *tags* ativas partiu da premissa de se aproveitar a utilização do sistema de controle de acesso, para prover também a localização dos usuários dentro do laboratório. Sobre este fato pesa também que os usuários naturalmente já portam as *tags* para prover seu acesso ao laboratório, o que facilita então a sua localização através de um sistema que aproveite as informações geradas por esta *tags*.

A literatura sobre localização *indoor* possui muitos trabalhos apresentando resultados para sistemas utilizando RSSI provenientes de redes WIFI (protocolo 802.11) [33, 34, 35, 59], ou de sistemas baseados em RFID, principalmente do tipo *tag* passiva [11, 39, 57, 58], cuja aplicação é a localização de objetos ou “sinalização” de pessoas. Sistemas com *tags* ativas geralmente possuem aplicação diferente de localização, geralmente controle de acesso [56], e quando são utilizados para o fim de localização, são combinados com outras técnicas e/ou redes de sensores para habilitar a localização [4, 9, 12].

1.2 – PROBLEMÁTICA

A localização pretendida pelo sistema proposto no projeto de pesquisa, no qual este trabalho está inserido, é baseada na relação existente entre o valor do RSSI e a distância entre o emissor e o receptor do sinal, neste caso *tag* e leitora RFID. Entretanto, esta variável para canal de comunicação *indoor* é influenciada por fatores diversos, como os caminhos

múltiplos, que alteram sua relação com a distância, e por consequência dificultam sua modelagem.

Fonseca [1] em seu trabalho mapeou algumas dificuldades concernentes à variável RSSI, necessária ao funcionamento do sistema de localização. Dentre elas pode-se citar a influência direta da movimentação de pessoas próximas à *tag* RFID que está sob a observação do sistema de coleta, no nível de RSSI reconhecido pelas leitoras. Este fator é reconhecido também por Zhao, Liu e Ni [4], que em seu trabalho acadêmico atestaram ser este o motivo de instabilidade.

Uma vez conhecido que os próprios usuários causam influência no valor do RSSI percebido pelas antenas das leitoras, é prudente que a pesquisa seja direcionada ao registro e análise do comportamento desta variável. E ainda que as *tags* RFID sejam submetidas a situações que serão encontradas durante sua operacionalidade, ou seja, situações que envolvam a interação dos usuários com as *tags* RFID no dia a dia dentro do LARA, e não em situações controladas. Esta situação ainda não foi experimentada com o sistema disposto no LARA.

Uma sugestão de trabalho futuro proposto por Fonseca [1], é o de se utilizar outros tipos de classificadores para prover a localização baseada nos níveis de RSSI coletados do sistema de RFID. Outra é da aplicação do filtro de Kalman estendido, utilizando para isto o modelo de propagação da onda eletromagnética emitida pelos equipamentos que compõem o sistema de RFID.

Todas estas questões são consideradas no presente trabalho, que inserido num projeto maior, busca contribuir com questões antigas e já conhecidas, assim como, levanta questões novas, que surgiram de novos caminhos que foram trilhados.

1.3 – OBJETIVOS

1.3.1 – Objetivo Geral

De uma forma geral, é objetivo deste trabalho, investigar algoritmos de localização de usuários em ambientes prediais, que permitam a racionalização de energia elétrica. Em particular, considerando sistemas de refrigeração com compartilhamento de carga térmica dos ambientes.

1.3.2 – Objetivos Específicos

Investigar o desempenho do EKF em Sistemas de Localização baseados em RSSI de um Sistema de RFID.

Investigar e comparar o desempenho dos classificadores MLP, LVQ e SVM submetidos a diferentes situações de pré-tratamento e filtragem de dados, provenientes de sistemas RFID.

1.4 – APRESENTAÇÃO DO TRABALHO

A dissertação está dividida em seis capítulos. O primeiro contextualiza a pesquisa desenvolvida, justificando sua importância e aplicabilidade, no contexto do projeto de conservação de energia em desenvolvimento no LARA. Este capítulo trás ainda a descrição dos objetivos do trabalho.

No segundo capítulo são abordados os métodos de localização *indoor*, em seu estado de arte. Faz-se nesta oportunidade, uma avaliação dos vários métodos de localização aplicados em diferentes necessidades de localização de usuários em ambientes internos de edificações. Cabe ressaltar que maior ênfase foi dada aos métodos baseados em sistemas RFID ou em métodos baseados no relacionamento do RSSI com a posição do usuário.

As ferramentas computacionais aplicadas neste trabalho são descritas no capítulo três, que através de autores e bibliografias mais solidificadas, justifica sua aplicação na busca do objetivo proposto ao presente trabalho. Dentre as ferramentas computacionais eleitas para os experimentos encontram-se o filtro de Kalman estendido, as redes neurais MLP e LVQ e, por último, as SVM.

O capítulo quatro descreve todos os materiais aplicados nos experimentos conduzidos no LARA. Suas funcionalidades, particularidades, formas de aplicação e necessidades estão descritas, nos vários subitens deste capítulo. Também os métodos aplicados nos procedimentos, desde o planejamento dos experimentos, onde foram pensadas a disposição dos equipamentos do sistema de RFID, as formas de coleta dos dados gerados por este sistema, até o tratamento final dos dados, com a aplicação das ferramentas computacionais adequadas.

O quinto capítulo apresenta os resultados de cada etapa dos experimentos efetuados, desde a coleta, passando pelo tratamento dos dados e definição das bases de treinamento e finalizando com o resultado dos treinamentos aplicados aos classificadores. Concomitante à apresentação, são feitas análises e são tecidos comentários acerca destes resultados. Para encerrar este capítulo é feita uma comparação de desempenho dos três tipos de classificadores aplicados na identificação dos ambientes.

O sexto capítulo encerra o texto, indicando a que conclusão, os experimentos, os trabalhos advindos destes experimentos e seus resultados, conduziram o autor. Ainda neste capítulo são feitas algumas referências de possibilidades de continuação da pesquisa partindo dos resultados aqui registrados.

2 – MÉTODOS DE LOCALIZAÇÃO INDOOR

A necessidade de promover a localização de usuários dentro de um ambiente predial pode advir de várias aplicações que variam com a forma de interação deste ambiente com tais usuários. Nos chamados ambientes inteligentes a localização se faz necessária para prover serviços ao usuário que visam desde o seu conforto, até alguns procedimentos de segurança ou auxílio em casos de emergência.

É certo que, dependendo de cada aplicação, determinada técnica poderá ter maior vantagem de uso, seja pela facilidade, seja pela interface desenvolvida para a aplicação. Também é certo que todas as técnicas desenvolvidas buscam executar o melhor no quesito precisão, custo e aproveitamento de infra-estrutura já existente no ambiente a ser explorado. Os itens que seguem apresentarão algumas das técnicas mais comuns de localização em ambientes prediais, assim como algumas de suas características e resultados apresentados por trabalhos desenvolvidos em torno de cada técnica.

2.1 – SISTEMA RFID

A identificação por radio frequência ou RFID, tem uma história relativamente longa, de mais de 50 anos no campo das comunicações sem fio, mas somente na última década é que recebeu uma atenção considerável e tornou-se uma tecnologia de uso geral [5].

Quer percebamos ou não, RFID é uma parte integrante de nossas vidas [6]. Sistemas RFID aumentam a produtividade e a conveniência das empresas ou instituições que os utilizam, seja pelo conforto proporcionado aos seus usuários, ou seja pelos serviços e produtos, baseados em RFID, que ofertam estas empresas e instituições.

Para exemplificar algumas das centenas de aplicações, listamos na tabela 2.1 alguns usos citados por Landt [6].

Tabela 2.1 – Possíveis aplicações da tecnologia RFID

Aplicações da Tecnologia RFID
Prevenção de roubos de automóveis e mercadorias
Pagamento de pedágios sem paradas
Gerenciamento de tráfego
Controle de acesso a edifícios
Automatização de estacionamentos
Controle de acesso para veículos em condomínios fechados, corporações, aeroportos, etc
Gerenciamento de ativos
Distribuição de suprimentos
Localização e controle de empréstimos de livros em bibliotecas

A premissa básica por trás de um sistema RFID é marcar itens com uma *tag*. As *tags* RFID são dispositivos simples, cuja principal finalidade é guardar um número próprio chamado de ID e transmiti-lo às leitoras, em intervalos de tempo definidos, no caso de *tags* ativas, ou quando solicitadas, no caso de *tags* passivas [5]. Tais transmissões se dão através de transmissores de curto alcance. Além do número de identificação, as *tags* podem também transmitir outras informações, conforme sua tecnologia ou necessidade da aplicação, que são captadas pelas leitoras destes sistemas [7]. É justamente este processo de troca de informações entre *tags* e leitoras que permitem as aplicações listadas anteriormente.

Os sistemas RFID, hoje estabelecidos, se dividem em duas classes, os de *tags* passivas e os de *tags* ativas. Finkenzeller [8] explica que a diferença entre estas duas modalidades está na disponibilidade de fonte de energia, pois as do tipo passivas não possuem fonte de energia própria, sendo sua operação provida pela energia captada dos campos magnéticos ou eletromagnéticos emitidos pelas leitoras, enquanto que as do tipo ativas possuem fonte de energia própria, geralmente uma bateria, para prover sua operação.

A localização *indoor* baseada em sistemas RFID pode ser executada com qualquer uma das classes de *tags* ou ambas, dependendo da aplicação específica e dos resultados esperados. Ni et al. [9] propuseram o sistema LANDMARC, que se baseava em RFID com *tags* ativas para fazer localização *indoor*. Tal sistema empregava a idéia de fixar algumas *tags* para referenciar a localização e ajudar na calibração do sistema. Este sistema alcançou a marca de 1m de erro na localização para até 50% das medidas e um erro máximo de 2m. Os experimentos foram feitos dentro de uma sala num espaço cujas dimensões são 4m de largura por 9m de comprimento, aparentemente controlados e sem divisórias. Não há também a indicação de mobiliário no espaço reservado ao experimento.

Zhao, Liu e Ni [4] trouxeram outra proposta para a abordagem do LANDMARC, que chamaram de VIRE (do inglês, *Virtual Reference Elimination*). Desta vez foi proposto utilizar *tags* virtuais e não mais *tags* físicas. Tais *tags* virtuais atuam como um mapa de referência para cada leitora do sistema. A implementação do VIRE melhorou a performance do LANDMARC, diminuindo o erro para cerca de 0,29m em até 73% dos casos. Mesmo as publicações mais atuais sobre sistemas de localização baseados em RFID com *tags* ativas, como o trabalho de Jin, Lu e Park [10], partem do LANDMARC, ou buscam aprimorá-lo, alcançando resultados pouco melhores.

Willis e Helal [11] propõem um sistema de navegação e determinação de localização com *tags* passivas para usuários cegos. O sistema é baseado em um *grid* de *tags* passivas que são distribuídas em determinado espaço, e relacionadas às coordenadas espaciais do local e também às características do local em sua vizinhança. O usuário carrega uma leitora de dimensões extremamente reduzidas, que pode ser integrada a um sapato, coletando assim o sinal das *tags* situadas junto ao solo e depois transmitindo via *Bluetooth* a um PDA ou um *Smart Phone*, que irá processar os dados relativos a localização.

Um sistema baseado em *tags* ativas e passivas simultaneamente foi implementado por Tesoriero et al. [12] para prover um serviço de posicionamento de equipamentos móveis em espaços fechados e ao mesmo tempo disponibilizar informações contextuais sobre determinados objetos expostos em museus e galerias de arte. As *tags* ativas são utilizadas para prover a localização e as *tags* passivas são utilizadas para prover as informações dos objetos.

2.1.1 – As incertezas do RSSI

O RSSI é um valor adimensional utilizado para referenciar o nível de sinal percebido pela antena da leitora. Este conceito foi introduzido pela norma IEEE 802.11, que o definiu como um mecanismo pelo qual a energia da rádio frequência deve ser mensurada nos circuitos dos controladores de interface de rede *wireless* [13]. Ainda de acordo com esta norma, o RSSI é expresso por um número inteiro que pode variar entre 0 e 255, pois é expresso por 8 bits (1 byte), entretanto cada fabricante pode adotar um valor para RSSI máximo menor que 255, buscando representar melhor o sinal lido.

O valor de RSSI não está associado a nenhuma unidade que expresse valor de energia ou potência elétrica, desta forma, os fabricantes ficam livres para assumir qualquer granularidade desde que o RSSI seja sempre expresso por números inteiros. Esta definição partiu do pressuposto de que o valor de RSSI seria de interesse somente de alguns processos internos dos controladores de interface de rede, como reconhecimento de uso do canal por outros controladores.

Extrapolando esta norma, outros equipamentos de rádio frequência adotaram também o uso do RSSI como parâmetro disponibilizado para aplicações diversas, sempre com a idéia de expressar o nível do sinal percebido pela antena do receptor. Como o nível de sinal recebido se relaciona com a distância entre transmissor e receptor, não tardaram aparecer aplicações buscando relacionar o RSSI com a distância entre estes entes, mesmo tendo uma relação imprecisa entre o RSSI e a real potência do sinal recebido. Este interesse se firmou principalmente em aplicações de localização em ambientes internos de prédios e edifícios, onde técnicas de localização tradicionais como o GPS não tem aplicação por limitação óbvia de falta de sinal.

Rappaport [14] destaca que o canal de comunicação sem fio *indoor* se difere do canal de comunicação sem fio tradicional em dois aspectos: primeiro a distância de cobertura é bem menor para o *indoor* e a variabilidade do ambiente é muito grande para uma pequena faixa de distância de separação entre transmissor e receptor. Diz-nos ainda ter sido observado que a propagação da onda eletromagnética dentro de edificações é fortemente influenciada por características construtivas destas edificações, como leiaute, materiais aplicados e tipo de

utilização, pois estas características causam um efeito conhecido como caminhos múltiplos (do inglês, *multipath*).

Todas estas características refletem a dificuldade de se modelar um canal de comunicação móvel, em especial para ambientes *indoor*, que tem sido, historicamente, uma das partes mais difíceis do projeto de sistemas de rádio móveis. Aos modelos de propagação que caracterizam as flutuações rápidas da intensidade do sinal recebido para distância muito curtas (alguns comprimentos de onda) ou para curtas durações denominam-se modelos em pequena escala ou modelos de atenuação [14].

Na atenuação em pequena escala, a potência do sinal recebido pode variar em até três ou quatro ordens de grandeza (30 ou 40 dB) quando um receptor, ou transmissor, é movimentado apenas por uma fração do comprimento de onda. Tal variação ocorre em função de que o sinal recebido é o resultado de uma soma de muitas contribuições vindas de diferentes direções, ocasionadas pelo efeito dos caminhos múltiplos. A figura 2.1 mostra esta variabilidade acentuada no sinal para pequenos deslocamentos, em comparação com o comportamento do nível de sinal para deslocamento em larga escala, que varia de forma mais suave.

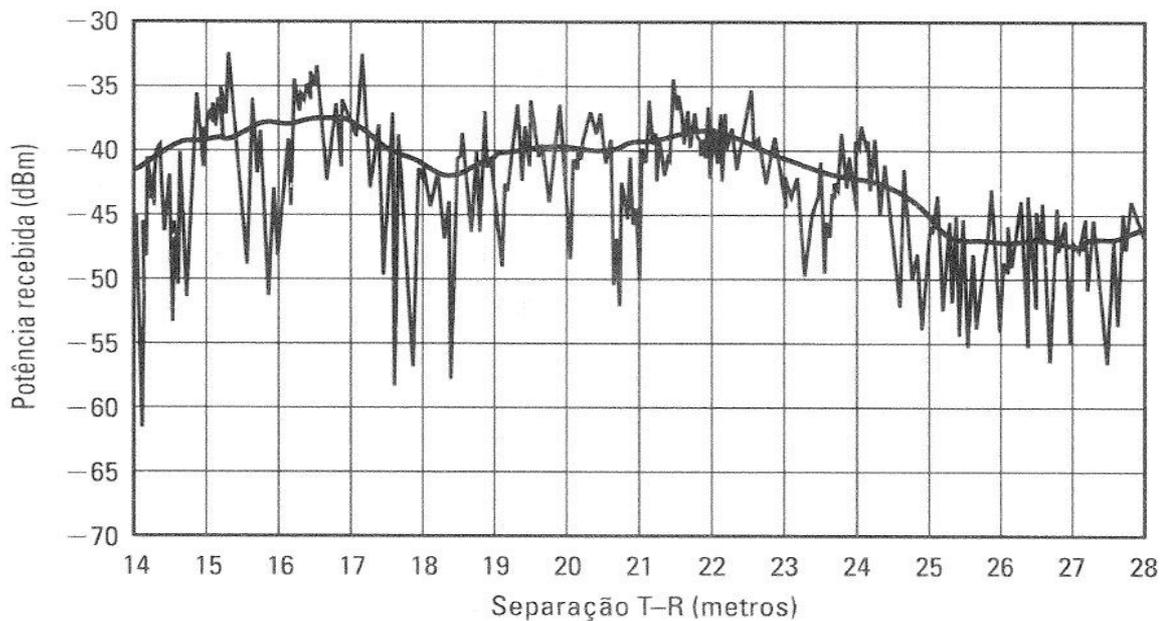


Figura 2.1 – Enfraquecimento do sinal em pequena e larga escala de deslocamento [14]

Lymberpoulos, Lindsey e Savvides [15] após conduzirem um estudo empírico sobre a variabilidade do RSSI, dentro de um espaço amplo e minimamente influenciado pelo efeito de caminhos múltiplos, concluíram que a orientação da antena influencia no nível do sinal percebido por esta. Logo, isto altera significativamente a qualidade da informação que o RSSI fornece para se estabelecer uma relação espacial entre transmissor e receptor. Os autores afirmam ainda que a localização utilizando RSSI funciona em ambientes controlados, entretanto em outros cenários é uma tarefa extremamente desafiadora.

Ni et al. [9] afirma que um dos problemas em usar RFID com *tags* ativas para localização de objetos é a inconsistência na recepção do RSSI, e que isto pode estar associado ao ambiente ou ao próprio equipamento, sendo na maioria dos casos o ambiente responsável por estas inconsistências. Fatores como posicionamento de móveis, presença de pessoas se movimentando podem impactar na precisão da localização e inclusive na distância máxima para se detectar algo.

De fato os trabalhos de Souza e Oliveira [2] e Fonseca [1], desenvolvidos no LARA corroboram com esta afirmação, pois os resultados apurados pelos primeiros em campo aberto e pelo segundo nos manuais dos equipamentos estão muito além dos seus resultados mensurados no ambiente fechado do LARA, onde são notórios os dois fatores de impacto mencionados por Ni et al. [9].

A figura 2.2 mostra o comportamento do RSSI de uma *tag*, coletado durante 48 horas por Fonseca [1]. Neste experimento a *tag* ficou parada, mas os usuários do laboratório tiveram pleno acesso às instalações. Conforme demonstra esta figura, houve um período em que os valores de RSSI sofreram grandes alterações. Segundo Fonseca [1], tal fato ocorreu num período em que um grupo de alunos estava presente no ambiente, estudando próximo à mesa em que a *tag* estava posicionada. Por outro lado, os períodos de grande estabilidade, geralmente à noite, representam momentos em que não houve presença de pessoas dentro do laboratório.

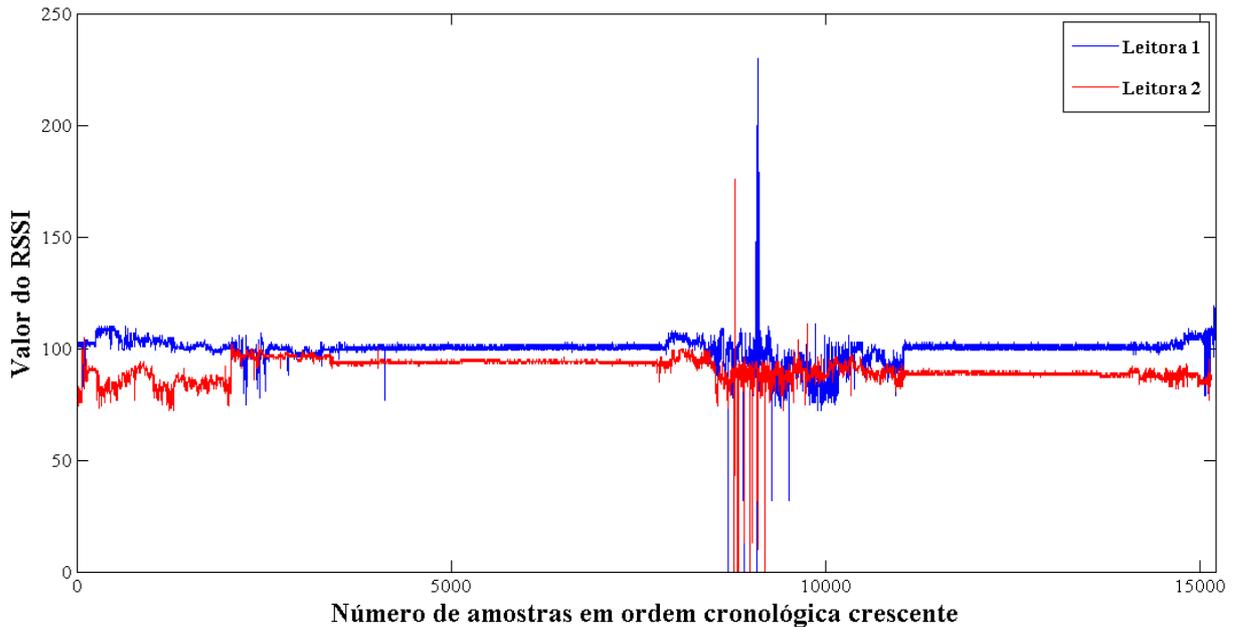


Figura 2.2 – Comportamento do RSSI em teste de longa duração (48 horas) [1]

Zhao, Liu e Ni [4] também destacam que a movimentação humana entre as *tags* e leitoras é a principal causa de instabilidade nos valores apurados de RSSI. Os experimentos feitos por estes autores mostraram mudança repentina nos valores de RSSI quando uma pessoa caminhava na área em que fazia tais experimentações. Zhao, Liu e Ni alertam ainda que este fator deve ser evitado ou ainda filtrado quando se projetar sistemas desta natureza. Fonseca [1] também destaca esta observação em seu trabalho, num experimento em que deixou uma *tag* parada por 48 horas no LARA. Assim, o período em que os níveis de sinal mais variaram foi o mesmo em que os usuários do laboratório estiveram próximos da mesa onde a *tag* fora posicionada.

2.2 – SISTEMA GPS-INDOOR

O GPS – *Global Positioning System* foi o primeiro sistema de posicionamento mundial a ser desenvolvido e efetivamente implantado. Desenvolvido pelos Estados Unidos, caracteriza-se por uma rede de 24 (vinte e quatro) satélites geostacionários que transmitem códigos telemétricos e dados de navegação em duas frequências, L1 (1.575,42 MHz) e L2 (1.227,6 MHz), usando uma técnica de multiplexação denominada CDMA – *Code Division Multiple Access* [16]. Com equipamentos receptores apropriados, usuários podem obter sua localização em coordenadas latitudinais e longitudinais, quando recebem sinais

disponibilizados por três ou mais satélites desta rede. Quanto mais satélites estiverem disponíveis ao receptor, maior será a precisão da localização informada por ele.

Atualmente o sistema GPS compõe o chamado GNSS - *Global Navigation Satellite Systems*, junto com outros sistemas que estão em fase de ampliação ou desenvolvimento. São eles o russo GLONASS, que já está em operação, mas tem cobertura restrita, o GALILEO da União Européia que ainda está em fase de projeto, e é previsto para estar operacional em 2020 e o chinês COMPASS, também em fase de projeto, que é uma evolução de outro projeto chinês chamado BeiDou [17].

Na forma em que foi concebido, o sistema GPS exige que os receptores tenham visada direta aos satélites, pois obstruções de qualquer que seja a natureza, atenuam os sinais a níveis que impedem seu reconhecimento pelos equipamentos receptores. Este fato impede o uso destes receptores em ambientes internos de edificações ou mesmo externo obstruído.

A utilização deste sistema para localização *indoor* surgiu da evolução de um projeto do governo americano, que previa a localização de usuários dos serviços de telefonia móvel que discassem ao número de emergência 911. Este sistema ficou conhecido como E911 – *Enhanced 911*, como explica Dedes e Dempster [18]. Este projeto iniciado em 1996 determinou, através da FCC – *Federal Communications Committee*, que as operadoras dos serviços de telefonia móvel, informassem com precisão a localização de chamadas feitas ao serviço de emergência 911. As metas previam para a primeira fase, que as operadoras deveriam informar o número do telefone que originara a chamada e ainda a localização da torre a que este usuário estava conectado no momento da chamada. Na segunda fase era necessário informar a localização do usuário, baseada no sinal do aparelho móvel utilizado para estabelecer a chamada, com uma precisão de pelo menos 100 m em 67% dos casos e de até 300 m para 95% dos casos.

Em 1999, estas metas foram alteradas para 50m em 67% dos casos e de até 150 m para 95% dos casos, para usuários de aparelhos equipados com A-GPS – *Assisted GPS*. Isto criou uma expectativa no mercado de aparelhos celulares que deveriam ser equipados com mais esta facilidade e criou também uma demanda no desenvolvimento de uma solução quanto ao sinal GPS, que notoriamente não podia ser recuperado dentro de edificações, o que impedia a localização dos usuários, por este fato [18]. Dentro de edificações os usuários estão

sujeitos às várias intempéries como recepção do sinal fraco, efeitos de caminhos múltiplos (próprios de propagação em canal indoor) e interação entre os diferentes sinais transmitidos na mesma frequência pelos satélites, em função do fraco nível de sinal recebido.

Diante do desafio de prover a localização *indoor* através de um sistema já firmado como ferramenta de localização *outdoor*, surgiram alternativas que aproveitaram a estrutura do sistema GPS, a exemplo do citado A-GPS, assim como outros sistemas baseados em suas características, mas não utilizando propriamente sua estrutura, todos com a finalidade de se prover a localização em ambientes fechados.

O A-GPS é constituído basicamente por uma rede de servidores que tem acesso ao sinal de todos os satélites que compõem o sistema GPS, logo possuem os dados de telemetria de cada um destes satélites *online*. Estes dados são disponibilizados à rede de telefonia móvel, que por sua vez repassa aos aparelhos móveis que tenham o serviço de localização implementado, para que possam prover sua localização, quando estiverem em ambientes onde o sinal do sistema GPS não possuir níveis suficientes para tal [19].

Isto pode acontecer de duas formas, primeiro com o processamento ocorrendo no próprio aparelho que precisa estar equipado com um chip de nova geração, denominado processador de convolução, que permite a utilização do próprio processador do aparelho de telefonia móvel para os cálculos sem, contudo prejudicar suas outras funções. Isto é possível porque a rede de servidores do A-GPS disponibiliza os dados de cada satélite cujo sinal possa ser detectado pelo receptor em questão, enquanto que para os receptores convencionais estes dados são baixados do próprio satélite após a identificação do sinal. Uma vez que o receptor possui os dados referentes a localização do satélite, basta que este compare com o sinal que está recebendo, o que diminuí a carga de processamento no receptor. Neste caso a informação é processada em tempo real.

Segundo, o receptor acumula as informações coletadas dos satélites e depois as repassa a rede A-GPS, que então faz o processamento do posicionamento através de uma convolução no domínio da frequência. Este tipo de localização não se dá em tempo real, pois o processamento somente é feito após o acúmulo de um volume de dados provenientes dos satélites.

Fluerasu et al. [20] propõem a localização em ambientes *indoor*, utilizando-se de repetidores para reforçar o sinal emitido pelos satélites dentro destes ambientes. Na configuração por eles testada quatro repetidores foram instalados dentro de um estacionamento cuja estrutura toda em aço não permitia a recepção direta dos sinais emitidos pelos satélites. Prevendo a ocorrência de caminhos múltiplos, com a emissão simultânea dos sinais pelos repetidores, os autores adotaram um método de transmissão sequencial, onde cada repetidor transmite isoladamente em uma janela de tempo, até que se complete um ciclo, retornando ao primeiro repetidor que deu início ao processo.

Nos experimentos práticos conduzidos por Fluerasu et al. [20], a precisão alcançada foi de aproximadamente 10 metros, isto sem qualquer tipo de tratamento nos sinais. Aplicando uma técnica denominada SMICL (do inglês, *Short Multipath Insensitive Code Loop*), que busca mitigar o efeito dos caminhos múltiplos, na demodulação do sinal emitido pelo satélite, a precisão aumenta para valores entre 2 a 3 metros.

Ozsoy, Bozkurt e Tekin [21] também propõem um sistema de localização *indoor* que utiliza repetidores de sinal para prover o sinal dos satélites dentro das edificações e cujo desempenho, segundo os autores é muito próximo à do GPS para ambientes externos. Entretanto os autores utilizam antenas direcionais, que são posicionadas na parte superior das edificações, para captar o sinal dos satélites, em linha de visada e, então, dispor este sinal aos repetidores. As antenas são direcionais para que possam limitar a área de atuação de cada repetidor, evitando que o sinal de um determinado satélite seja captado, amplificado e repetido por dois equipamentos distintos, desta forma os repetidores atuam simultaneamente. O sistema proposto ainda utiliza receptores convencionais, bastando que passem por uma atualização de *firmware*.

Rizos et al. [22] apresentam uma tecnologia de localização, inclusive já registrada comercialmente, denominada LOCATA. Esta tecnologia cria uma réplica autônoma terrestre de um sistema GNSS, permitindo que receptores comuns deste último sistema possam prover a localização de usuários mesmo em ambientes onde sabidamente os sinais originais de um sistema GNSS têm pouca, ou nenhuma penetração. Nos testes apontados no trabalho dos autores, é indicada uma precisão da ordem de centímetros, em um ambiente de um armazém todo construído em aço, apontado como ambiente repleto de

caminhos múltiplos, onde sistemas convencionais de localização baseados em Rádio Frequência não seriam bem sucedidos.

2.3 – SISTEMA UWB

O UWB tem o princípio de funcionamento similar ao radar. Os transmissores emitem pulsos dentro de uma banda de frequência muito larga, geralmente 0,5 GHz ou cerca de 20% da frequência da portadora. O pulso enviado tem potência muito pequena, mas suficiente para excitar o receptor que faz a varredura do ambiente de acordo com o sinal recebido, provendo assim o leiaute e a localização das pessoas neste ambiente.

Nekoogar [23] afirma que esta tecnologia não é nova, pois foi utilizada por Marconi em 1901 para transmitir sequências de Código Morse, entre a Europa e a América. Mais tarde, na década de 60, este tipo de transmissão encontrou aplicação nos radares militares e em comunicações secretas, que exigiam alto nível de segurança. Além da transmissão pulsada, outra característica do sistema UWB é a largura de banda, que comumente é maior que 0,5 GHz ou maior que 20% da frequência da portadora [24].

Somente em 2002 o FCC, órgão Norte Americano responsável por regular as comunicações via rádio frequência ou via cabo, regulamentou o uso desta tecnologia para propósitos comerciais [23]. Após esta regulamentação, muitas pesquisas em torno desta tecnologia despontaram, como o trabalho de Alavi, Alsindi e Pahlavan [25], que buscou a definição de um modelo de propagação do sinal UWB em ambiente *indoor*, sob influência de caminhos múltiplos, através de medições de sinais UWB em quatro tipo de edificações diferentes. Outros trabalhos também contribuíram para desenvolver propostas de aplicações que podem ser observadas a tabela 2.1.

Dentro deste contexto apresentado, Galler et al. [26] propõem a combinação de duas técnicas de localização, AOA e TOA, que adicionadas a um sistema UWB, permitem promover a localização de usuários em determinado ambiente *indoor*. Segundo os autores, a combinação destas técnicas de localização, permite a eliminação da principal causa de imprecisão de sistemas de localização baseados em sinais de rádio frequência, os caminhos

múltiplos. O resultado apurado pelos autores destaca uma precisão da ordem de decímetros em condições de visada direta entre transmissor e receptor.

Sua aplicação como radar é o principal motivo de pesquisa dentro da academia onde se consegue encontrar publicações como as de Yu et al. [27], Petroff [28], Huseth, Dewberry e McCroskey [29] e Saidi et al. [30]. Aplicações em localização de robótica móvel também são encontradas a exemplo de Gonzáles et al. [31].

Tabela 2.2 – Possíveis aplicações da tecnologia UWB, Nekoogar [23]

Aplicação		
	Militar e Governamental	Comercial
Comunicação de Dados	Comunicação Segura – LPI/D (baixa probabilidade de interceptação/detecção)	Redes Locais e Pessoais Redes Domésticas (distribuição de sinal de vídeo)
	Redes de Sensores sem Fio Secreta (operações de guerra)	Redes de Sensores sem Fio (monitoramento de saúde, automação residencial, etc)
Radar	Imagem Através de Parede (aplicação da lei, bombeiros)	Imagem Médica (monitoramento remoto coração)
	Radar de Penetração no Solo (para operações de resgate)	Radar de Penetração no Solo (detecção de fiação elétrica, pregos, etc, nas construções civis)
	Vigilância e Monitoramento	Indústria Automotiva (sensor anti colisão) Segurança Residencial (sensor de proximidade)
Localização	Identificação Pessoal	Controle de Estoque
	Crianças Desaparecidas	Etiquetagem e Identificação
	Controle de Prisioneiros	Gerenciamento de Ativos

2.4 – OUTROS SISTEMAS DE LOCALIZAÇÃO INDOOR

Pelo exposto até aqui, pode-se observar que a localização em ambientes *indoor* pode ser implementada através de sistemas bastante diferentes, sendo que a escolha entre eles basear-se-á na disponibilidade de determinada plataforma, ou na necessidade específica de se

promover a localização. Neste item apresentaremos mais alguns sistemas, que também podem ser utilizados na localização em ambientes *indoor*.

2.4.1 – Sistemas de Localização Baseada em Redes Sem Fio padrão 802.11 (WIFI)

Sistemas baseados em RSSI de redes sem fio padrão IEEE 802.11, mais comumente chamada WIFI, geralmente têm aplicação na computação ubíqua, onde a localização de determinado usuário contribui na contextualização de um aplicativo que oferecerá iterações a este usuário. Para promover este tipo de localização é necessário que o ambiente *indoor* a ser explorado possua uma rede sem fio do padrão 802.11 com pelo menos três APs – *Access Points*, pois este número já permite a utilização da técnica mais clássica de localização que é a triangulação. Loiola [32] chama a atenção para o fato de que os APs que constituirão o sistema de localização devem prover o seu sinal na maior parte do ambiente a ser explorado, pois isto evitará a inconsistência nos dados de localização do pretendo usuário.

A despeito das técnicas de localização para este tipo de sistema, existem muitas variações que são objeto de estudo e de novas propostas de implementação por parte do meio acadêmico. Fang e Lin [33], por exemplo, propõem uma nova abordagem da técnica denominada *fingerprint* que em português significa impressão digital, buscando aprimorá-la com uma ferramenta chamada Componentes Discriminativos. A esta nova abordagem nomearam DANN (do inglês, *Discriminant-Adaptive Neural Network*). A técnica *fingerprint* baseia-se na coleta dos níveis de sinal na maior parte possível do ambiente onde o usuário poderá se deslocar, criando a partir daí um mapa destes níveis de sinal, que é utilizado para comparar com o sinal coletado do usuário quando este estiver por ser localizado. A principal aplicação deste mapa é o treinamento de uma rede neural por exemplo. Os autores não indicam se ocorre melhor precisão na localização com a nova abordagem proposta, entretanto destacam que ocorre maior facilidade de treinamento da rede neural que irá decodificar os níveis de sinal em posicionamento do usuário.

Chintalapudi, Iyer e Padmanabhan [34] apresentam um algoritmo que realmente suscita uma forma diferente de prover a localização de determinado usuário, baseado em três pressupostos: primeiro que, sendo o projeto baseado em sistemas de rede sem fio padrão 802.11 (WIFI), deve existir no ambiente a ser explorado uma quantidade de

APs suficiente (não cita números) para uma boa cobertura de sinal destes AP em toda área deste ambiente; segundo que os usuários estejam portando dispositivos móveis, como *Smartphones* ou *Netbooks*, equipados com WIFI; e terceiro que ocasionalmente o dispositivo móvel obtenha a informação absoluta de sua localização por GPS, nos limites do ambiente, onde mais facilmente conseguirá sinal dos satélites, como em janelas ou entradas/saídas.

Isto basta ao algoritmo, denominado pelos autores de EZ (que seria um acrônimo para a palavra em inglês *easy*), que roda em um servidor que se comunica com o dispositivo móvel do usuário, para que seja promovida a localização deste usuário em coordenadas absolutas (latitude e longitude). A precisão conseguida nos experimentos efetuados pelos autores ficou entre 2m e 7m. Segundo os próprios autores, estes valores não são muito atrativos quando comparados a outras técnicas, entretanto a grande relevância do trabalho encontra-se no fato de que não há a necessidade prévia de se conhecer o ambiente, ou ainda a localização ou níveis de sinal de cada um dos APs para calibragem do sistema de localização.

Dentre as técnicas baseadas em redes sem fio padrão 802.11, alguns trabalhos fazem uma abordagem baseada em sistemas múltiplos, como Paul e Wan [35], que utilizam uma versão do filtro de Kalman *unscented* para fundir as informações dos vários sistemas em prol da localização, em vez de lidar com as *fingerprints*. O sistema proposto pelos autores é baseado num SPKF, que faz a fusão do modelo dinâmico de marcha de uma pessoa, que considera ainda forças de repulsão para se evitar deslocamentos através de paredes ou móveis, por exemplo, com certo número de sensores de movimento de baixo custo, permitindo o rastreamento do usuário, incluindo sua velocidade de deslocamento.

Os resultados apurados pelos autores revelaram uma precisão de até 1,9m para localização considerando somente o sistema WIFI e precisão de até 0,46m com a fusão dos sensores de movimento. Paul e Wan [35] destacam ainda em suas conclusões que apesar de a abordagem realizada ser capaz de prover boa precisão, muitos fatores podem limitar esta performance, tais como ruído na medição do RSSI, variações no ambiente, taxa de amostragem utilizada na coleta de dados, além das complicações comuns da propagação da onda eletromagnética dentro de um ambiente *indoor*.

2.4.2 – Sistemas de Localização Baseada em Redes de Sensores Sem Fio (WSN)

Redes WSN também encontram aplicações de localização, como demonstrado por Figueredo, Couto e Bauchspiess [36] que propuseram testar o desempenho de dois métodos de localização numa rede de sensores sem fio, baseada na plataforma ZigBee, inserida dentro de um ambiente inteligente. O primeiro método, o da triangulação hiperbólica, requer pelo menos três sensores atuando como receptor de sinais, para se estimar a posição do usuário. Para cada sensor calcula-se um raio de distância, com base no RSSI medido pelo sensor, onde provavelmente se localiza o usuário, então a localização deste usuário é definida através da interseção das circunferências formadas pelos raios de todos os sensores.

O segundo método, também se baseia na medição de RSSI pelos sensores, mas desta vez, se constrói um mapa de níveis de RSSI em todo ambiente a ser localizado, criando assim um perfil do ambiente com relação à propagação eletromagnética do sinal emitido pela rede de sensores sem fio. Este mapa é então utilizado para treinar uma rede neural, neste caso utilizou-se uma rede neural do tipo MLP e treinamento *Backpropagation*. Na entrada da rede neural foram implementados os níveis de RSSI mensurados pelos sensores e na sua saída foram apuradas as coordenadas cartesianas do usuário, relativas ao ambiente em estudo.

Os testes efetuados por Figueredo, Couto e Bauchspiess [36] foram realizados dentro de um ambiente de 7,50m de comprimento por 3,00m de largura, e permitiram chegar numa precisão de 1,50m para o método da triangulação hiperbólica e 1,05m para o método do mapa de propagação eletromagnética com treinamento de rede neural MLP. Em suas conclusões os autores indicam que o melhor desempenho da rede neural deve-se ao fato de que o treinamento com os dados coletados diretamente no ambiente de aplicação já considera algumas perturbações ambientais comuns à propagação em ambientes *indoor*.

Outro trabalho utilizando redes WSN foi proposto por Larranaga et al. [37], que através do relacionamento do RSSI com as distâncias euclidianas entre os nós da rede WSN buscam prover a localização de usuários. A cada solicitação de localização o sistema proposto faz uma calibragem (nova leitura) dos valores de RSSI entre os nós chamados de referência, pois possuem sua localização fixa e conhecida. Desta forma, a cada solicitação se restabelece uma nova relação entre as distâncias euclidianas e os valores de RSSI apurados,

buscando-se eliminar a influência de eventuais mudanças de cenário dentro do ambiente a ser explorado. Todo este processamento é feito em um servidor que apura a localização do referido nó e depois a disponibiliza para a rede. Nos testes experimentais do algoritmo proposto pelos autores, o erro médio das localizações ficou próximo de 3 metros. Destaca-se no trabalho de Larranaga et al. [37] o fato de não necessitar pré-processamento de dados nem mesmo para a implantação do sistema, e também de que já na sua concepção foi preparado para considerar qualquer mudança significativa do ambiente, a cada evento de localização.

2.4.3 – Sistemas de Localização Baseada em Redes Bluetooth

As redes *Bluetooth*, consideradas redes de aplicação para comunicação pessoal (WPAN), são geralmente de aplicação restrita as comunicações de pequenos equipamentos como aparelhos celulares, máquinas fotográficas, *notebooks* e acessórios para estes equipamentos, que não distam mais que 10 metros uns dos outros. A comunicação através desta tecnologia se dá através de emissão de RF, logo possui também todas as características descritas para as demais tecnologias anteriormente citadas.

Utilizando-se da tecnologia *Bluetooth*, Diaz et al. [38], propuseram uma aplicação de localização que nomearam Bluepass. Tal proposta fixa algumas estações pelo ambiente em que se quer prover a localização, para que estas possam coletar o RSSI de equipamentos móveis encontrados em sua vizinhança. Uma vez coletados, os dados relativos ao equipamento móvel encontrado são enviados a um servidor central que irá computar a localização deste equipamento e armazenar estas informações disponibilizando-as a quem interessar. Para estimar a localização dos usuários, os autores utilizam a técnica da triangulação, calculando o comprimento do raio a partir de um modelo de propagação definido pelo ITU – *International Telecommunication Union*, e também utilizam uma técnica denominada SCDM – *Signal Coverage Density Method*.

A SCDM divide toda a área em que se quer prover a localização em pequenas áreas quadradas, então calcula uma distância a partir de cada estação leitora, onde mais provavelmente encontra-se o usuário e define um peso para as pequenas áreas anteriormente definidas. As áreas sobrepostas terão mais peso atribuído, logo representam a área mais provável de localização do usuário. Esta técnica é aplicada para definir em qual ambiente está o usuário e para minimizar eventuais erros de localização na aplicação da técnica de triangulação.

Os resultados apurados pelos autores em seu trabalho mostraram uma precisão de 3,23m para localização num ambiente de 195m², dividido em 4 ambientes distintos. Após estes resultados os autores ainda fizeram mais alguns testes, tentando relacionar o acerto do ambiente em que estava o usuário quando localizado. Nestes últimos testes fizeram coletas de dados em três ambientes e em três pontos distintos em cada ambiente. A localização através do método SCDM se mostrou mais eficaz na identificação do ambientes em que o usuário se encontrava presente, independente do ponto em que a coleta dos sinais ocorrera dentro do ambiente, enquanto que o método da triangulação se demonstrou muito debilitado nesta função, por depender mais diretamente do resultado do modelo de propagação do sinal RF.

3 – MÉTODOS DE FILTRAGEM E CLASSIFICAÇÃO DE DADOS

Dados produzidos por experimentos científicos muitas vezes trazem consigo uma infinidade de variações e ruídos que prejudicam sua compreensão. Em primeira vista, um conjunto de dados nesta condição, ou seja, misturados a ruído ou sem a devida ordenação, pode não fazer sentido ao objetivo almejado pela pesquisa. Daí surge a necessidade de um tratamento destas informações, fazendo-se uma seleção de tudo que possa interessar ao resultado esperado ou ainda fazendo uma ordenação destes dados para que possam melhor expressar a realidade contida em cada parâmetro ou grandeza associada a estes dados.

Existem muitas técnicas que podem ser aplicadas nesta fase de pré tratamento dos dados, obtidos em determinada circunstância de pesquisa. A escolha de uma destas técnicas deve ser sempre pautada nos enfoques e benefícios que se busca dentro do universo de informações contidas no conjunto dos dados. Dentre as técnicas de filtragem disponíveis, foi escolhida para aplicação no presente trabalho, como elemento de correção de eventuais erros de leituras do RSSI, o filtro de Kalman estendido. O filtro de Kalman apresenta aplicações já solidificadas dentro da indústria, como nos receptores GPS, e também é de extensa aplicação ou objeto de estudo nos projetos de pesquisa desenvolvidos nos laboratórios de pesquisa das universidades pelo mundo, como exemplo podem ser citados os trabalhos de Fonseca [1], Souza e Oliveira [2], Paul e Wan [35], Bekkali, Sanson e Matsumoto [39], Vaidehi et al. [40], Teixeira, Tôrres e Aguirre [41] e Santana [42].

Também foram aplicadas neste trabalho três técnicas de classificação de dados, muito conhecidas da engenharia de controle, as redes neurais artificiais MLP e LVQ, e as SVM. Exemplos de aplicações destes classificadores são encontrados em Fonseca [1], Fang e Lin [33], Loiola [32] e Hata e Wolf [43] para aplicações de redes neurais artificiais, sejam elas do tipo MLP ou LVQ, e em Zhen e Jia [44], Hsu, Chang e Lin [45], Yamano et al. [46] e Hata e Wolf [43] para aplicações das SVM. Estas técnicas serão apresentadas a seguir.

3.1 – FILTRO DE KALMAN ESTENDIDO (EKF)

O filtro de Kalman nada mais é que um estimador recursivo ótimo [47] que, em diversos aspectos se assemelha a outros como o Estimador Recursivo de Mínimos Quadrados. A estimação recursiva é o procedimento em que os dados coletados e disponibilizados de forma sequencial, são utilizados para atualizar o vetor de parâmetros de um determinado modelo. Uma das principais utilidades das técnicas recursivas é a estimação de parâmetros de um determinado modelo à medida que os dados do processo são disponibilizados [47].

Originalmente, o filtro de Kalman foi proposto para atuar em sistemas lineares, entretanto surgiram variações, como o filtro de Kalman estendido (EKF) e o filtro de Kalman *unscented* (UKF). Estas duas novas técnicas permitiram a aplicação do Filtro de Kalman em sistemas não lineares, através de processos como a linearização analítica e a transformada *unscented*, que ajudam no cálculo da propagação temporal da média e da covariância, fatores preponderantes na determinação do ganho de Kalman [47].

Teixeira, Tôrres e Aguirre [41] enfatizam que o EKF é a solução mais utilizada nas últimas quatro décadas, entretanto, apresenta algumas limitações relativas à definição dos estados iniciais das variáveis do sistema e das matrizes de covariância de ruído do sistema e das medições. Isto tudo piora para sistemas fortemente não lineares. Os autores informam ainda que a teoria proposta por Kalman baseia-se no uso de duas fontes de informações que produzem valores incertos. A primeira fonte é o modelo matemático do sistema, que nem sempre reflete a totalidade do comportamento apresentado por este sistema, e a segunda as medições, que também apresentam incertezas em suas covariâncias de difícil definição.

Welch e Bishop em seu relatório técnico [48] fazem uma apresentação sucinta, mas bastante precisa sobre o filtro de Kalman, onde apresentam passo a passo, as principais condições assumidas na formulação deste algoritmo. Começam por definir a formulação básica do sistema, em que um filtro de Kalman Discreto irá, por definição, tentar estimar $x \in \mathfrak{R}^n$ de um processo controlado discreto no tempo. Tal sistema é regido pela equação estocástica linear mostrada em (3.1). A medição das grandezas disponibilizadas pelo sistema representada por $z \in \mathfrak{R}^m$ é dada pela equação (3.2).

$$x_k = \mathbf{A}x_{k-1} + \mathbf{B}u_{k-1} + w_{k-1} \quad (3.1)$$

$$z_k = \mathbf{H}x_k + v_k \quad (3.2)$$

No sistema apresentado pelas equações (3.1) e (3.2), as variáveis w_k e v_k representam, respectivamente, o ruído branco do processo e da medição. Elas são independentes uma da outra e apresentam distribuição de probabilidade normal, com média zero e covariância \mathbf{Q} e \mathbf{R} (respectivamente). A matriz \mathbf{A} tem dimensão $n \times n$ e relaciona o estado do sistema entre os instantes a priori $k-1$ e o instante atual k . A matriz \mathbf{B} tem dimensão $n \times l$ e relaciona a entrada de controle do sistema $u \in \mathfrak{R}^l$ ao estado x . A matriz \mathbf{H} tem dimensão $m \times n$ e relaciona o estado do sistema à medição das grandezas disponibilizadas por este a z_k .

Santana [42] nos anexos A3 e A4 de seu trabalho também explica sobre o Filtro de Kalman Puro e Estendido e expõe o algoritmo destes filtros dividido em duas etapas. A primeira é chamada de “Predição de Estimativas”, na qual basicamente o filtro aplica o equacionamento do modelo para estimar o estado futuro do sistema. Faz isto somando ao estado do instante anterior, que deve ser conhecido, a entrada de controle do sistema, dando origem ao estado chamado de a priori. Neste passo é feita ainda a predição da covariância do estado a priori, também considerando a covariância deste estado conhecida no instante anterior, a covariância da entrada de controle do sistema e, ainda, a covariância do ruído do sistema. Abaixo se observa o equacionamento deste passo do algoritmo do filtro de Kalman puro.

$$\bar{x}_k = \mathbf{A}_{k-1}\hat{x}_{k-1} + \mathbf{B}_{k-1}\bar{u}_{k-1} \quad (3.3)$$

$$\bar{\mathbf{P}}_k = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{B}_{k-1}\mathbf{P}_{k-1}\mathbf{B}_{k-1}^T + \mathbf{Q}_{k-1} \quad (3.4)$$

A segunda etapa é chamada de “Correção das Estimativas”, e apresenta o equacionamento que segue caso haja disponibilidade de medição do sistema:

$$\mathbf{K}_k = \bar{\mathbf{P}}_k\mathbf{H}_k^T(\mathbf{R}_k + \mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T)^{-1} \quad (3.5)$$

$$\hat{x}_k = \bar{x}_k + \mathbf{K}_k(z_k - \mathbf{H}_k\bar{x}_k) \quad (3.6)$$

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\bar{\mathbf{P}}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \quad (3.7)$$

A equação (3.5) calcula o ganho de Kalman, que é na verdade o fator de correção entre valor estimado e valor efetivamente medido. A sintonia do filtro, que é

baseada nas covariâncias das grandezas do sistema/modelo e na covariância das medições, é que indicará a ordem de grandeza deste ganho, pois dependendo destas covariâncias, indica-se maior confiabilidade no modelo ou na medição. Daí vem a necessidade de se conhecer bem o sistema em que se pretende aplicar a filtragem de Kalman. Isto implica inclusive numa boa definição do modelo.

Uma vez definido o ganho que será aplicado na correção do valor estimado, a equação (3.6) irá atualizar o valor dos estados do sistema e a equação (3.7) irá atualizar a covariância destes estados. Santana [42] explica ainda que, caso não haja a disponibilidade de medições, como por exemplo, em falhas de sensores, os valores dos estados e da covariância destes estados a posteriori devem ser igualados aos valores a priori, ou seja, somente a estimativa será considerada.

Para o filtro de Kalman estendido o equacionamento é muito parecido, entretanto existem alguns passos a mais para se obter o resultado final. Estes passos estão relacionados à linearização analítica do sistema, que dever ser processada antes de qualquer outro passo do processo de filtragem. A linearização é feita através do cálculo da Jacobiana (derivada parcial de primeira ordem) do modelo do sistema, a cada instante em que se pretende aplicar a referida filtragem.

Desta forma, e seguindo a exposição efetuada para o filtro de Kalman puro, se apresenta abaixo a formulação requerida para o filtro de Kalman estendido, a começar pela definição do modelo do sistema não linear, no qual temos que o estado atual é função do estado no instante anterior e de uma entrada de controle do sistema também no instante anterior. A isto se soma o ruído branco do sistema, como demonstra a equação (3.8). A medição do estado no instante atual será função do estado no instante atual mais um ruído característico da medição, como se vê na equação (3.9).

$$x_k = f(x_{k-1}, u_{k-1}) + w_{k-1} \quad (3.8)$$

$$z_k = h(x_k) + v_k \quad (3.9)$$

Para os cálculos de predição de estimativas, na primeira etapa de filtragem do filtro de Kalman estendido, é preciso antes efetuar o cálculo das Jacobianas relativas à

predição do estado, considerando o modelo que define o comportamento do sistema. Este passo é efetuado com as equações dispostas em (3.10).

$$\mathbf{J}_{f,x} = \left. \frac{\partial f(x_{k-1}, u_{k-1})}{\partial x_{k-1}} \right|_{\hat{x}_{k-1}, \bar{u}_{k-1}} \quad (3.10)$$

$$\mathbf{J}_{f,u} = \left. \frac{\partial f(x_{k-1}, u_{k-1})}{\partial u_{k-1}} \right|_{\hat{x}_{k-1}, \bar{u}_{k-1}}$$

Uma vez calculadas as jacobianas, o próximo passo é calcular a predição dos estados, começando pela equação (3.11), que estima o estado atual em função do estado anteriormente conhecido e da entrada de controle também anteriormente conhecido, e da equação (3.12), responsável pela predição da covariância, na qual se observa a aplicação das Jacobianas recém calculadas.

$$\bar{x}_k = f(\hat{x}_{k-1}, \bar{u}_{k-1}) \quad (3.11)$$

$$\bar{\mathbf{P}}_k = \mathbf{J}_{f,x} \hat{\mathbf{P}}_{k-1} \mathbf{J}_{f,x}^T + \mathbf{J}_{f,u} \mathbf{P}_{u,k-1} \mathbf{J}_{f,u}^T + \mathbf{Q}_{k-1} \quad (3.12)$$

Na segunda etapa de filtragem também ocorre a necessidade de se calcular primeiro a Jacobiana relativa à medição das grandezas do sistema, como se vê na equação (3.13), pois esta será utilizada na definição do ganho de Kalman, na atualização da predição dos estados efetuada na etapa anterior e na atualização da covariância do sistema.

$$\mathbf{J}_{h,x} = \left. \frac{\partial h(x_k)}{\partial x_k} \right|_{\bar{x}_k} \quad (3.13)$$

$$\mathbf{K}_k = \bar{\mathbf{P}}_k \mathbf{J}_{h,x}^T (\mathbf{R}_k + \mathbf{J}_{h,x} \bar{\mathbf{P}}_k \mathbf{J}_{h,x}^T)^{-1} \quad (3.14)$$

$$\hat{x}_k = \bar{x}_k + \mathbf{K}_k (z_k - h(\bar{x}_k)) \quad (3.15)$$

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{h,x}) \bar{\mathbf{P}}_k (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{h,x})^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad (3.16)$$

A utilização do filtro de Kalman estendido pressupõe a utilização de um modelo matemático que represente o sistema de forma mais fiel possível, pois isto minimizará o acúmulo de erros na linearização analítica, em cada ciclo de filtragem. Como se percebe, a cada ciclo de filtragem, a Jacobiana é recalculada para que possa estar representando a linearização do modelo para aquele instante em que se pretende fazer a estimativa do estado.

Neste trabalho em função da não linearidade apresentada pela propagação da onda eletromagnética do sistema de localização *indoor* baseado em RFID com *tags* ativas, foi proposto o uso de filtro de Kalman estendido como ferramenta de filtragem dos níveis de RSSI coletados pelas antenas dos receptores. O intuito de tal filtragem era de minimizar as incertezas do RSSI causadas pela variabilidade do ambiente em que estes equipamentos estão submetidos.

3.2 – MULTI LAYER PERCEPTRON (MLP)

O MLP é um dos reconhedores de padrões mais utilizados para classificação de dados. Sendo por princípio uma rede neural artificial, sua configuração é baseada em nós computacionais denominados neurônios. O neurônio computacional é o elemento unitário de uma rede neural artificial. Sua formatação mais conhecida, o Perceptron, foi proposta por Rosenblatt no final dos anos 50 do século XX e tem por finalidade básica transmitir para sua saída, os sinais inferidos em sua entrada, aplicando neles os pesos sinápticos que lhes são devidos [49].

A aprendizagem do Perceptron se dá através de um processo supervisionado, no qual um conjunto de dados que resulta em saídas previamente conhecidas é apresentado nas entradas. Conforme estes dados são apresentados nas entradas, a saída tem seu resultado comparado com os valores desejados das saídas e previamente conhecido, e na ocorrência de erro é aplicada uma correção aos pesos sinápticos, para que estes possam melhor representar os modelos que estão sendo apresentados.

Apesar de ter causado certa euforia entre os pesquisadores da inteligência artificial, na época em que foi proposto, por sua capacidade de classificação de padrões linearmente separáveis, o Perceptron também frustrou outros tantos, pois não é aplicável a problemas de classificação onde os padrões não são separáveis por um único hiperplano, como por exemplo, o problema do OU exclusivo. A solução deste problema aconteceu somente após a proposição de um algoritmo de treinamento, que permitiu a configuração de Perceptrons em mais de uma camada, o Algoritmo *Backpropagation*.

Haykin [49] diz que as redes perceptron multicamadas representam uma generalização das redes perceptron de camada única, e são tipicamente formadas por um conjunto de unidades sensoriais que constituem a camada de entrada, e por uma ou mais camadas ocultas de nós computacionais (neurônios), e por fim uma camada de nós computacionais alinhadas na saída destas redes. A figura 3.1 mostra a arquitetura de uma rede MLP.

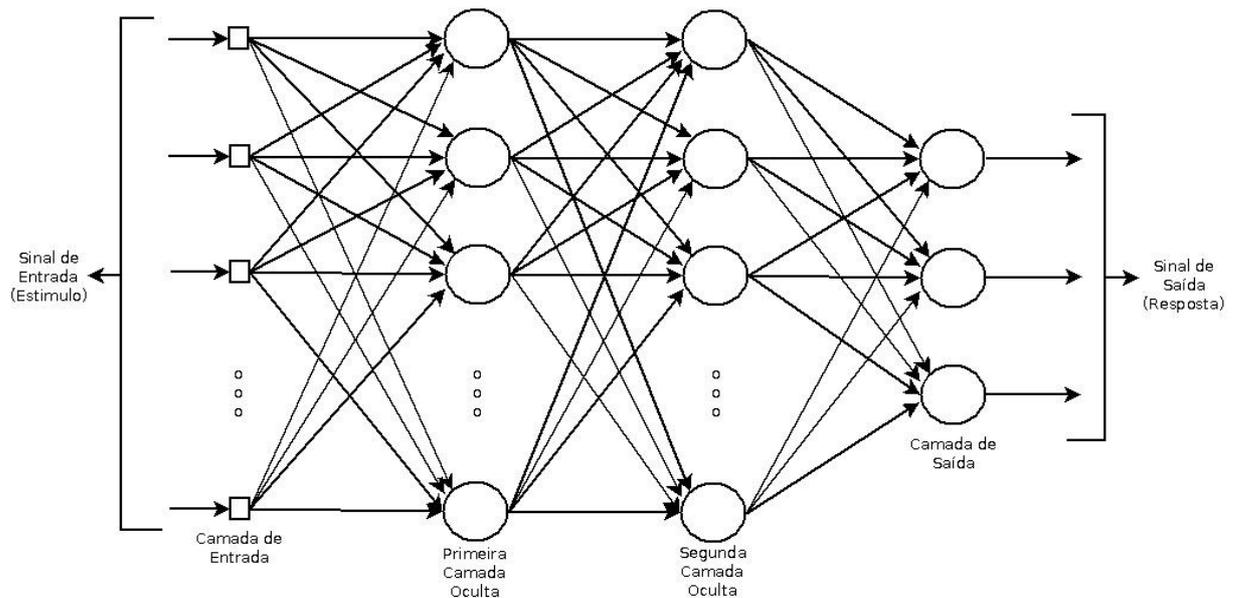


Figura 3.1 – Arquitetura de uma rede MLP com duas camadas ocultas, adaptado de Haykin [49]

O algoritmo *Backpropagation* é realizado basicamente em dois passos, sendo o primeiro referente à propagação do sinal de entrada, constituído por um vetor que define um padrão de atividade até a saída, onde produzirá uma resposta da rede, baseada em seus nós computacionais e nos pesos sinápticos de suas ligações, a este estímulo aplicado na entrada. O segundo passo, denominado propriamente de retropropagação, consiste em subtrair a resposta obtida na saída da rede do valor de resposta desejada, de forma a se obter um valor quantitativo do erro produzido pela rede. Este erro é aplicado na correção dos valores dos pesos sinápticos, que ocorre da saída para entrada, propagando-se através das várias camadas, daí surge o nome do algoritmo de *Backpropagation* [49].

Três características básicas da rede MLP, que permitiram sua aplicação como classificadora de padrões não separáveis por um hiperplano linear são descritas por Haykin [49]: Primeiro, cada neurônio da rede possui uma função de ativação não linear. Existem

vários tipos de funções de ativação, mas geralmente esta função é do tipo sigmóide, como por exemplo, a tangente hiperbólica, cuja característica de não linearidade é suave o suficiente para que possa ser diferenciável em qualquer ponto. A não linearidade desta função é importante ainda para que se possa promover a separação de padrões não separáveis por hiperplanos lineares.

Segundo, a rede possui uma ou mais camadas ocultas entre a entrada e a saída da rede. Isto permite a aprendizagem de tarefas complexas, extraindo de forma progressiva, durante o treinamento, as características mais significativas dos padrões de entrada. Terceiro, a rede possui um alto grau de conectividade, que pode ser observado nas ligações sinápticas de seus neurônios. Isto implica que todos os sinais de entrada influenciam em todos os neurônios das camadas escondidas, e estes, por sua vez, influenciam nos valores de cada saída, através dos pesos sinápticos de suas ligações.

Uma importante preocupação deve ocorrer quando do treinamento da rede, a generalização da aprendizagem, ou seja, a garantia de que a rede apresente um bom índice de acerto para dados nunca visto por ela. Haykin [49] aponta que uma rede generaliza bem, quando o mapeamento de entrada-saída computado por ela é correto, para dados de teste não utilizados para a criação ou treinamento dela. Estes dados de teste são geralmente separados da própria base de dados antes do treinamento da rede.

A generalização final da rede é sem dúvida influenciada pela arquitetura da rede escolhida para resolver o problema e pelo tamanho e representatividade das classes do conjunto de dados a ser utilizadas no treinamento. Há, ainda, a influência da complexidade física que envolve o problema, sendo que nesta questão não há como interferir. Entretanto, para os outros dois fatores apontados, é possível que o projetista da rede possa buscar uma combinação que melhor satisfaça o desempenho final da rede. Para o caso de se ter a arquitetura fixa da rede, é preciso então determinar o tamanho ideal do conjunto de dados de treinamento, enquanto que para o caso inverso, ter o conjunto de dados de treinamento limitado, é preciso que se busque a arquitetura de rede que melhor generalize ao final do treinamento.

3.3 – LEARNING VECTOR QUANTIZATIONS (LVQ)

As LVQs são redes neurais do tipo mapas auto organizáveis, que têm a aprendizagem supervisionada, ou seja, fazem a separação das classes conforme os padrões apresentados, mas também consideram as saídas esperadas para este conjunto de dados. Isto é possível porque em sua arquitetura possuem uma primeira camada competitiva, e uma segunda camada linear, figura 3.2, cujas saídas podem ser comparadas com valores esperados durante o treinamento.

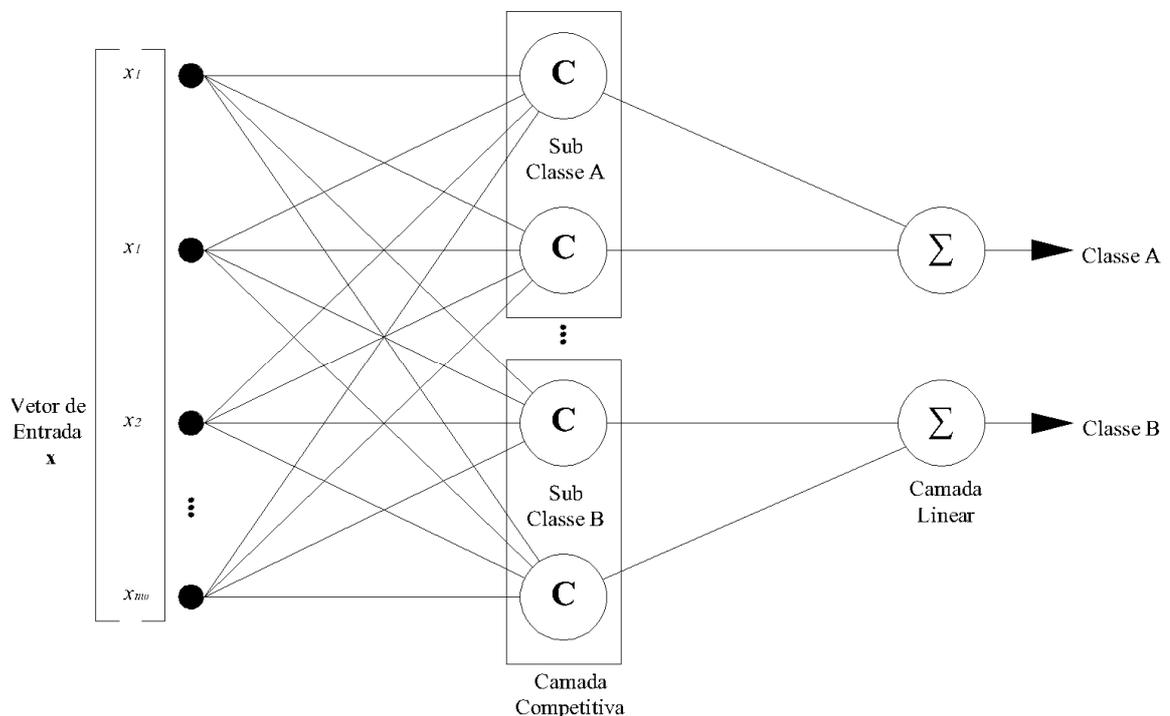


Figura 3.2 – Arquitetura de uma rede neural do tipo LVQ [50]

Numa rede LVQ, a camada linear transforma a classificação dos dados pela camada competitiva em padrões previamente definidos pelo usuário da rede durante seu treinamento. Desta forma, o resultado obtido na camada competitiva é denominado de subclasse e o resultado obtido na camada linear é denominado classe [50].

A quantidade de neurônio nas camadas será igual a quantidade de subclasses para a camada competitiva e igual a quantidade de classes para a camada linear. Isto implica que a quantidade de neurônios na camada competitiva é sempre maior que na camada linear. Analisando esta configuração, percebe-se que cada neurônio de saída (camada linear), será

ativado por qualquer neurônio dentro de um grupo específico de neurônios da camada competitiva. Estes, por sua vez, são ativados através da apresentação de um vetor de entrada que despontará em um neurônio vencedor.

A figura 3.3 mostra um exemplo de classificação efetuada com uma rede LVQ. Os dados utilizados para o treinamento da rede estão divididos em duas classes, representadas na figura 3.3 pelas cores azul e vermelho. O treinamento desta rede resultou na geração de quatro subclasses, duas para cada classe. Os demais pontos na figura 3.3 representam pontos aleatórios submetidos à rede LVQ para que esta pudesse fazer a classificação destes pontos.

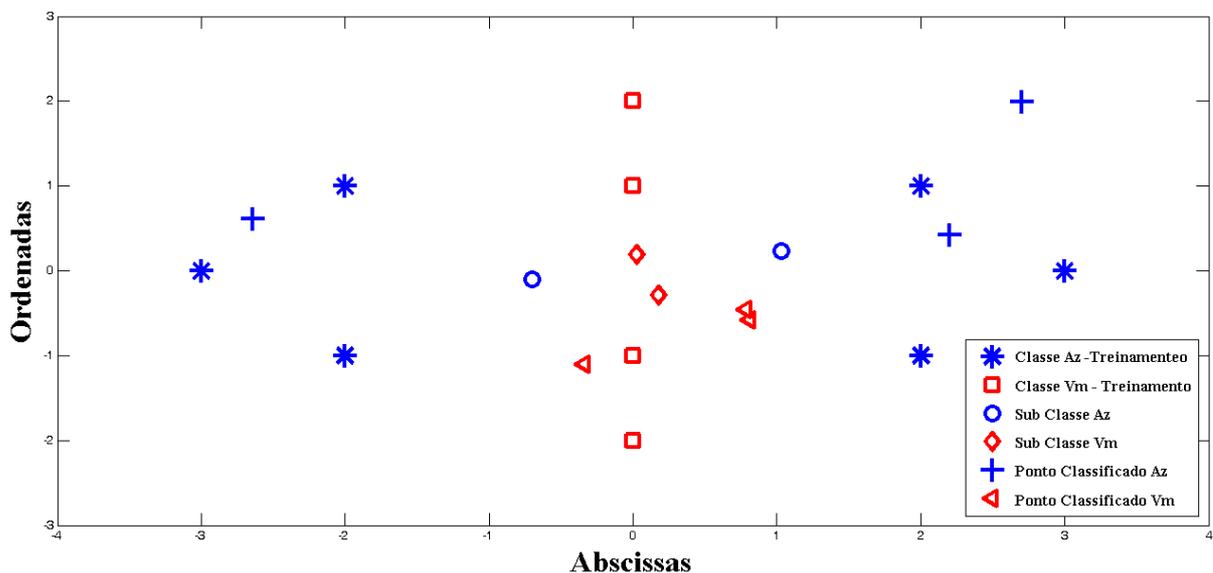


Figura 3.3 – Exemplo de classificação de dados utilizando LVQ, adaptado de [50]

3.4 – SUPPORT VECTOR MACHINE (SVM)

As SVM foram concebidas basicamente para classificar padrões linearmente separáveis. Em sua configuração padrão, ou seja, em sua concepção inicial, são utilizadas para classificar dois padrões, e faz isto estabelecendo um hiperplano, como superfície de decisão, de tal forma que a margem de separação entre exemplos destes dois padrões seja máxima [49].

A figura 3.4 mostra um exemplo de padrões linearmente separáveis e um hiperplano que é definido pelos vetores localizados nas bordas do agrupamento do

subconjunto de dados relativos a cada classe. Estes vetores são denominados vetores de suporte e são definidos, dentre os dados de treinamento, durante o processo de aprendizagem através do algoritmo da máquina de vetor de suporte. Uma vez definido o hiperplano ótimo de separação entre as classes, a máquina de vetor de suporte pode classificar qualquer dado apresentado em sua entrada, calculando a diferença entre a distância da origem ao hiperplano e a distância entre a origem e o ponto a ser classificado. O resultado desta operação define como classe positiva e negativa o padrão apresentado na entrada.

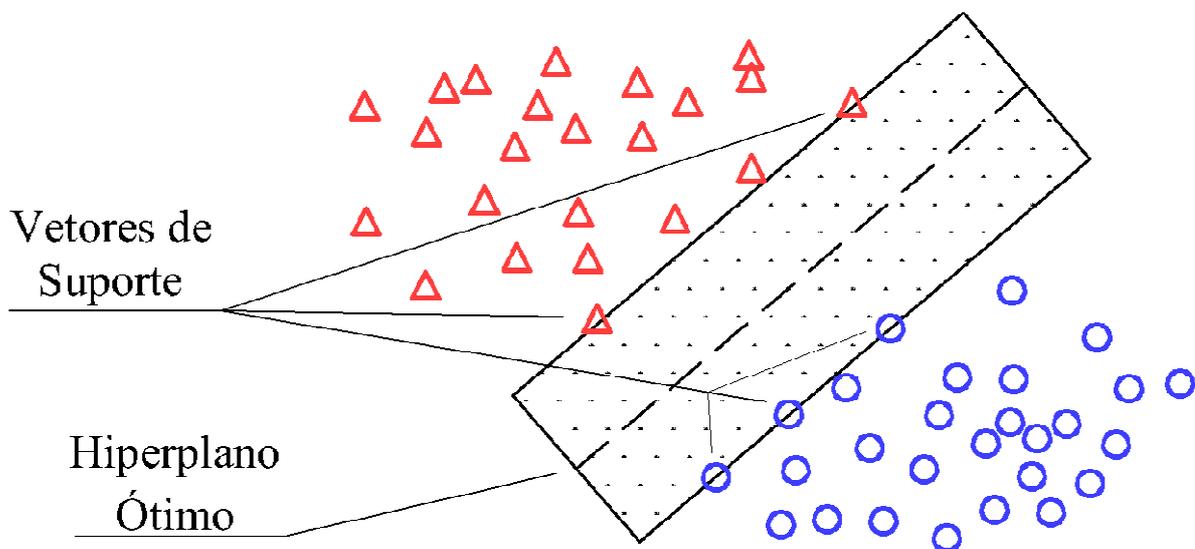


Figura 3.4 – Padrões linearmente separáveis e hiperplano definido pelos vetores de suporte, adaptado de Haykin [49]

O hiperplano ótimo de separação das classes numa SVM é definido pelo somatório do produto interno entre o vetor de entrada e os vetores de suporte, transformados por uma função de transferência que os transporta do espaço dimensional da entrada para outro espaço dimensionalmente muito superior, denominado espaço de características. Esta transformação para o espaço de característica é feita para proporcionar a separabilidade dos padrões de entrada de forma linear [49]. O conjunto responsável por estas transformações e definições é chamado de núcleo de produto interno.

Na sua forma mais simples, uma SVM possui núcleo de produto interno linear, ou seja, o somatório do produto interno é direto entre os vetores de entrada e os vetores de suporte. Entretanto, Haykin [49] enfatiza que é possível implementar outras máquinas de aprendizagem quando se implementa o núcleo de produto interno de uma SVM, podendo

inclusive serem caracterizadas por superfícies de decisão não lineares, como as máquinas de aprendizagem polinomial, as redes de função base radial e redes perceptron de duas camadas (uma camada oculta). A figura 3.5 mostra a arquitetura de uma SVM com núcleo de produto interno linear.

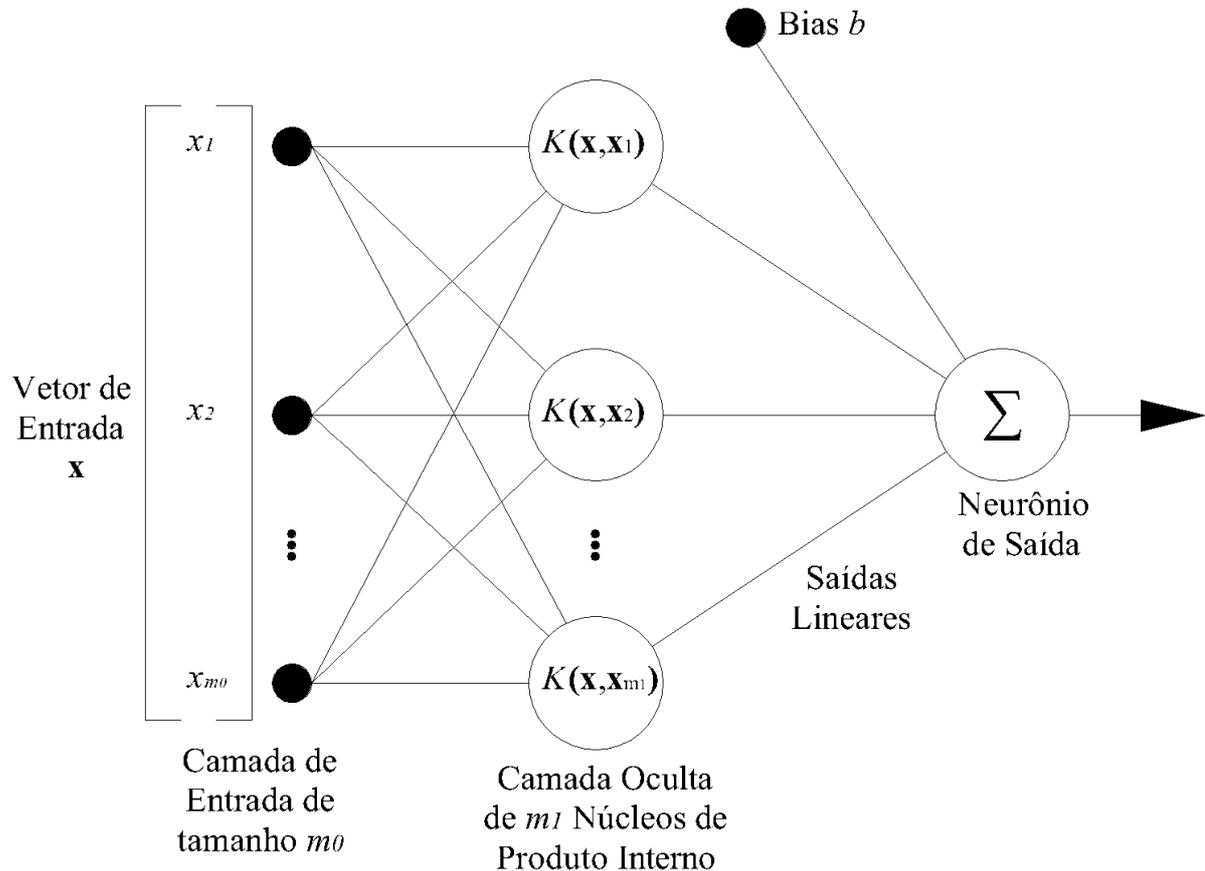


Figura 3.5 – Arquitetura de uma SVM com núcleo de produto linear, adaptado de Haykin [49]

As SVM até aqui citadas trabalham classificando somente duas classes, entretanto é comum encontrar problemas que demandam a classificação de elementos em três ou mais classes. Nestes casos, destaca Lorena e Carvalho [51], existem duas abordagens para a generalização das SVM, sendo a mais comum, a decomposição do problema em vários subproblemas binários, ou ainda numa segunda abordagem, é possível a modificação do algoritmo original de treinamento das SVM, criando versões multiclases.

Hsu, Chang e Lin [45], também enfatizam que a abordagem da decomposição do problema original de classificação em subproblemas seja de mais fácil lida, e propõe uma ferramenta de classificação denominada LIBSVM que trabalha os problemas multiclases

desta forma. Os autores informam que para um número de classes igual k , são construídos $k(k-1)/2$ classificadores que serão treinados para identificar as classes de duas em duas. Desta forma para a classificação é utilizada a estratégia de votação, na qual cada classificador irá votar em uma das classes do problema, sendo a classe vencedora aquela que mais votos obtiver entre os classificadores. Caso haja empate entre duas classes, o algoritmo escolhe a que primeiro estiver relacionada no rol de classes.

Os autores da LIBSVM alertam que muitos usuários das SVM não conseguem alcançar bom resultados nas primeiras experiências, pois tendem a fazer os testes de forma aleatória, experimentando algumas configurações dos parâmetros dos principais núcleos de produto interno. Isto comumente não produz bons resultados e pode no máximo conseguir desempenho igual a outros tipos de classificadores.

Hsu, Chang e Lin [45] sugerem que os testes da LIBSVM sejam iniciados com o núcleo de produto interno do tipo RBF, pois constitui uma configuração que permite maior gama de aplicação. Justificam isso afirmando que o núcleo Linear comporta-se como um caso particular do núcleo RBF, dependendo somente de ajustes em seus parâmetros. Comparam ainda o núcleo RBF ao núcleo de função sigmoïdal, afirmando que este segundo também tem comportamento similar ao primeiro, em determinadas configurações de parâmetros. Os autores finalizam enfatizando que o núcleo RBF apresenta pouca demanda computacional para os processamentos numéricos que, muitas vezes, se apresentam nas configurações de núcleo polinomial, pois este último pode alcançar valores próximos de infinito ou próximos de zero, quando os expoentes forem muito grandes.

Com base em todos os fatores considerados, Hsu, Chang e Lin [45] indicam o uso do núcleo RBF como primeira escolha dos testes de treinamento de uma SVM. A partir dos melhores resultados obtidos, deve-se então adotar os valores dos parâmetros de ajuste, como o controle do compromisso entre a complexidade da rede e o percentual de pontos não separáveis dos dados de treinamento, e outros parâmetros.

4 – MATERIAIS E MÉTODOS

Os equipamentos disponibilizados no LARA para implementação do sistema de localização utilizado neste trabalho, foram dimensionados por Fonseca [1] que, em seu trabalho, propiciou as primeiras informações acerca do comportamento destes equipamentos. Também Souza e Oliveira [2] efetuaram estudos com estes equipamentos disponibilizando outros importantes dados de seu comportamento. O sistema proposto neste trabalho busca atender a necessidade de localização dos ambientes internos do LARA para integrar um projeto ainda maior denominado Ambientes Inteligentes que, dentre outros assuntos, lida com a racionalização da energia em ambientes prediais automatizados.

Um dos motivos para a escolha do sistema RFID com *tags* ativas foi a compatibilidade com a aplicação no controle de acesso dos usuários ao LARA. Tal aplicação em controle de acesso imprime nos usuários a necessidade de portar uma *tag* RFID sempre que este necessitar adentrar ao laboratório. Esta dupla aplicação, somada ao formato das *tags*, como veremos a seguir, auxilia também na adaptação dos usuários no que diz respeito a portar as *tags* consigo, durante sua permanência dentro do laboratório.

Os equipamentos do sistema idealizado para a localização dos usuários do LARA geram dados que necessitam de um tratamento adequado, para se chegar ao efetivo resultado de localização. Este tratamento é feito primeiro através de um processo de filtragem, no qual os dados são depurados, retirando-se incertezas e eventuais erros grosseiros de medição e registro. Posteriormente, os dados passam por um processo de classificação, em que, de acordo com padrões previamente registrados, são separados em três classes distintas, definidas por três ambientes fisicamente distintos, dentro do LARA. Nos itens que seguem, os materiais e métodos aplicados no sistema de localização idealizado serão detalhados.

4.1 – MATERIAIS

Como citado no item 2.1, um sistema RFID é composto de leitoras, *tags* e um servidor no qual são acumuladas e processadas as informações coletadas pelas leitoras,

chamado *middleware*. O sistema implantado no LARA e, portanto, utilizado nas experiências, é composto por leitoras do modelo L-RX900 e *tags* ativas modelo L-TG501, do fabricante Wavetrend, cuja frequência de operação é 433MHz. Nas leitoras foram utilizadas ainda antenas do tipo onidirecionais no plano horizontal, do modelo L-N200 *Stub*. No item 4.1.1 estes equipamentos serão detalhados. O *middleware* utilizado neste trabalho foi um software aplicativo desenvolvido por Fonseca [1], que teve por finalidade coletar os registros dos pulsos efetuados em cada leitora e armazenar em um arquivo do tipo data *log*. O *middleware* será devidamente explorado no item 4.1.2.

A filtragem dos dados foi feita através da implementação do algoritmo do filtro de Kalman estendido e a classificação dos dados de acordo com os ambientes, foi efetuada aplicando-se três ferramentas computacionais distintas para se comparar o desempenho delas, as redes neurais do tipo MLP e LVQ, e também as SVM. Todas estas ferramentas computacionais serão exploradas nos itens 4.1.3 e 4.1.4, que seguem.

4.1.1 – Hardware do Sistema RFID Proposto

4.1.1.1 – Leitora L-RX900 e Antena L-AN200 stub

As leitoras do sistema implantado no LARA, cujo exemplar pode ser visualizado na figura 4.1 (a), são do modelo L-RX900 de fabricação da Wavetrend e operam em 433 MHz, faixa que no Brasil é de utilização aberta, ou seja, não possui restrição quanto ao seu uso ou necessidade de registro na ANATEL – Agência Nacional de Telecomunicações, segundo normatização emitida por esta agência.

Ao identificar um pulso emitido pelas *tags* ativas, este equipamento faz um pré tratamento dos dados recebidos que compreende a decodificação dos dados implícitos no sinal enviado pela *tag* e a validação destes dados. Estas leitoras possuem uma característica muito interessante para a aplicação desejada, elas possuem a capacidade de fazer a leitura simultânea de várias *tags*, depurando através de seu *firmware* eventuais colisões de dados provenientes de pulsos de diferentes *tags*.

Tais dados podem ser armazenados em uma pilha, ou ainda enviados a um servidor, dependendo somente de uma configuração prévia desta habilidade. No caso do armazenamento em pilha, os dados serão guardados até que sejam solicitados por um servidor através de comando específico para isto, ou até que a pilha atinja o seu máximo volume de armazenamento – dados referentes a 40 pulsos – quando começará a descartar os dados mais antigos.



(a) (b)
Figura 4.1 – Equipamentos RFID da Wavetrend: (a) leitora modelo L-RX900 e (b) antena modelo L-AN200

Nestas leitoras foram acopladas antenas do modelo L-AN200 *stub*, figura 4.1(b), também projetadas para a frequência de operação de 433 MHz e com diagrama de radiação omnidirecional no plano radial. Com esta antena, segundo o fabricante, as leitoras podem detectar sinais emitidos pelas *tags* em distâncias de até 40m. A figura 4.2 mostra o comportamento do RSSI em campo aberto.

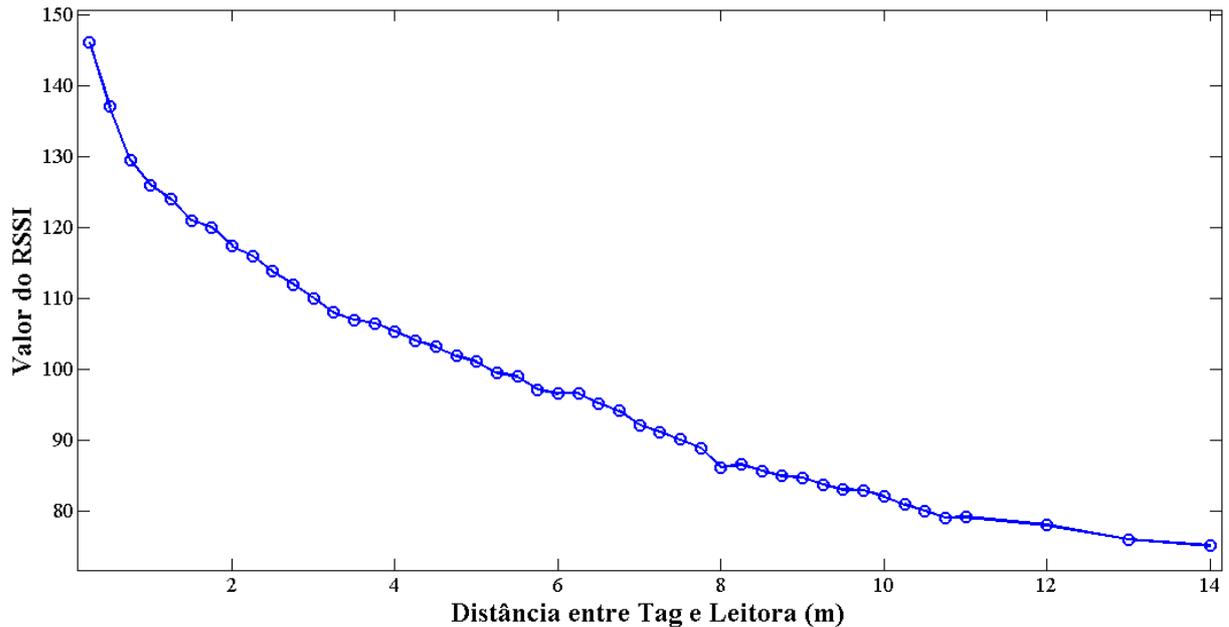


Figura 4.2 – Curva do comportamento do RSSI pela distância em campo aberto [2]

Souza e Oliveira [2] levantaram a curva de comportamento do nível de sinal em campo aberto, que está disposta na figura 4.2, e observaram que os níveis de RSSI para distâncias superiores a 14m já não permitiam que as leitoras fizessem o devido registro dos pulsos emitidos pelas *tags*. Os experimentos efetuados no LARA mostraram que o RSSI captado por estas leitoras em ambientes fechados, se comporta de maneira ainda mais crítica, pois o sinal caiu a níveis comprometedores em distâncias bem menores, da ordem de 7 metros.

Cada leitora sai de fábrica com um MAC *address* único, e com um endereço IP padrão (192.168.0.10) já programado. Este endereço, assim como outras configurações, pode ser alterado através da porta Ethernet ou da porta serial RS232. Já a coleta de dados ocorre somente pela porta Ethernet.

A interface Ethernet se configura como uma importante facilidade de uso, pois permite a fácil conexão com o *middleware* ou qualquer outro software que se proponha a coletar os dados gerados pelas leitoras através dos pulsos emitidos pelas *tags*. Desta forma, as leitoras podem ser conectadas diretamente em qualquer rede do tipo Ethernet 10/100Mbps, sem a necessidade de se implementar uma rede única para o sistema.

4.1.1.2 – Tag Ativa L-TG501 MS

A *tag* ativa modelo L-TG501 MS mostrada pela figura 4.3 tem o formato de um crachá, o que facilita sua portabilidade e uso em aplicações de controle de acesso. Possui bateria interna e sensor de movimento, o que permite, por um tempo de aproximadamente 5 anos, gerar pulsos na frequência de 433 MHz, em intervalos de 15s quando está parada e intervalos de 1,5s quando em movimento. O intervalo de tempo entre os pulsos pode ser ajustado pelo fabricante, quando solicitado pelo cliente, mas via de regra, os valores são os informados acima.

A cada pulso emitido a *tag* incrementa um contador que é chamado de *Tag Age Counter*, isto auxilia a avaliar o tempo vida útil da bateria. Este contador também pode ser útil em outras aplicações como o controle de perda de pulso pelas leitoras. Dentre os dados enviados nos pulsos emitidos por este modelo de *tag*, podemos destacar o seu próprio ID, o valor do *Age Counter*, o tipo de pulso emitido, se de *tag* parada (15s) ou de *tag* em movimento (1,5s), etc.



Figura 4.3 – Tag Ativa modelo L-TG501 MS da Wavetrend

4.1.2 – Middleware do Sistema RFID Proposto

Para Oztaysi, Baysan e Akpınar [52], o *middleware* é um software que gerencia as leitoras e os dados provenientes das *tags*, passando estes dados para o aplicativo final do

sistema, ou para um banco de dados, pois fica no meio do fluxo de dados. Baudin e Rao [52] descrevem o *middleware* como uma interface que além de proporcionar a comunicação do aplicativo final do sistema de RFID com as leitoras, faz ainda a filtragem das várias informações coletadas pelas leitoras, informações estas que muitas vezes são irrelevantes ao sistema final de aplicação, ou ainda são coletadas de forma redundante.

O fabricante dos equipamentos do sistema RFID disponibiliza um Software para supervisão e configuração dos equipamentos, capaz até acompanhar o registro de pulsos pelas leitoras, entretanto não disponibiliza o log dos registros efetuados pelas leitoras. Desta forma, o *middleware* utilizado para coleta de dados foi desenvolvido por Fonseca [1] na linguagem de programação Delphi. Este Software comunica-se diretamente com as leitoras, fazendo a coleta dos dados apurados por estas e filtra dentre os dados as informações referentes à identificação da *tag*, o nível do RSSI percebido pelas leitoras, o *Age Counter* do pulso emitido pela *tag* e o período a que se refere o pulso emitido pela *tag* (se período de 1,5s ou de 15s). A tela de apresentação do *middleware* desenvolvido por Fonseca [1] pode ser observada na figura 4.5.

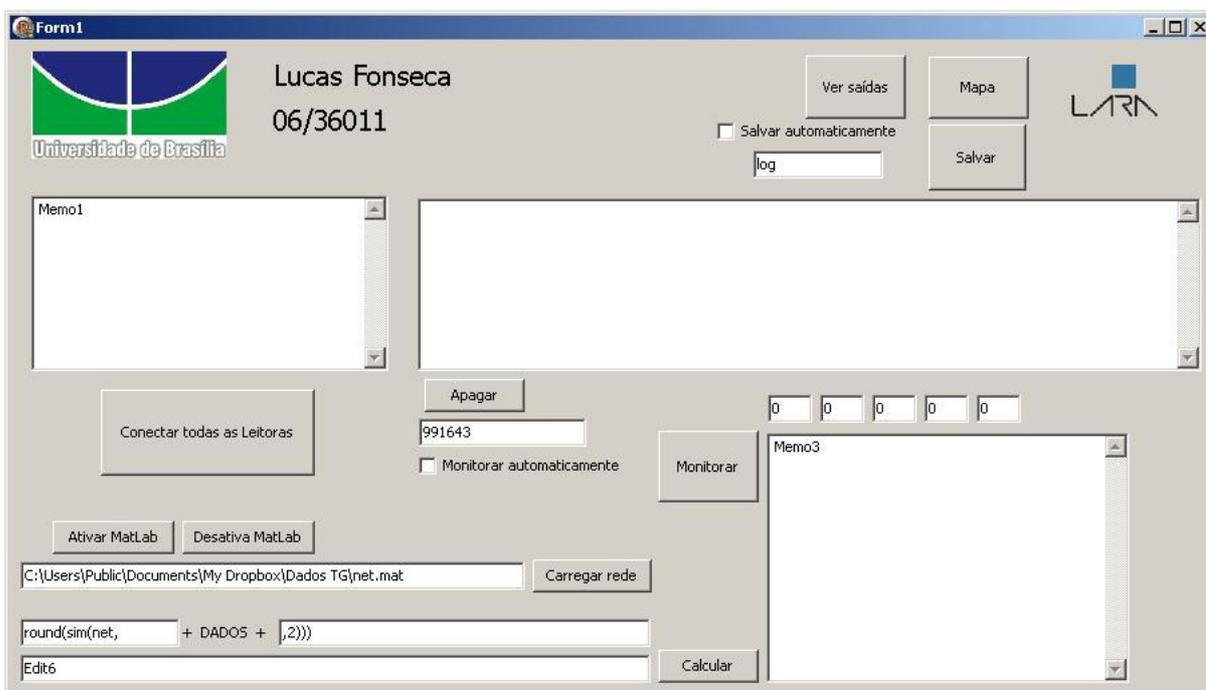


Figura 4.4 – Tela de apresentação do *middleware* utilizado na coleta dos dados [1]

Os dados coletados podem ser visualizados diretamente na tela de apresentação e ainda podem ser armazenados em um arquivo do tipo log de texto, onde cada linha do

arquivo refere-se ao registro de um pulso de uma *tag* por uma leitora. O armazenamento destes dados em arquivos do tipo log pode ser feito de forma manual ou de forma automática, bastando para tal que se sinalize em uma *check box*. Na função de salvamento automático, a cada 5 minutos, o arquivo de log será atualizado com os dados coletados nas leitoras.

Originalmente este programa pode ler dados em até cinco leitoras simultaneamente, número este de leitoras disponíveis no LARA para os experimentos com RFID. Através deste programa pode-se ainda fazer o monitoramento de uma *tag*, ou seja, ele pode destacar os dados relativos a uma única *tag*, em janela própria para isto, bastando para tal informar o número da *tag* a ser monitorada e ativar esta função através do botão monitorar. Esta função de monitoramento também pode ser efetuada de forma automática, bastando para tal sinalizar numa *check box* criada para isto.

Outra função interessante deste software é que ele pode apresentar os dados lidos pelas leitoras separadamente, ou seja, mostra em tela específica o que cada leitora está coletando junto as *tags*. Isto permite que o operador do sistema possa observar o comportamento do sinal em cada uma das leitoras de forma instantânea, através de uma tela de apresentação. Esta função pode ser útil inclusive na detecção de falhas do sistema.

Fonseca [1] desenvolveu também neste software uma interface para comunicação com o MatLab[®], onde se pretendia fazer a localização de determinada *tag* através de uma aplicação de rede neural MLP. Em seu trabalho Fonseca [1] chegou a treinar uma rede neural para tal aplicação, mas o fez com amostras de sinal coletadas a partir de uma *tag* parada em posições estrategicamente determinadas, que formaram um *grid* nas principais áreas de permanência de usuários do LARA. Todos os experimentos conduzidos por Fonseca [1] visando a localização de *tags* foram feitos com *tags* estáticas.

4.1.3 – Neural Network Toolbox do MatLab[®]

Como já foi referido, o MatLab[®] se apresenta como uma importante ferramenta computacional para simulações em engenharia e outras áreas que necessitam trabalhar com cálculos matriciais. O MatLab[®] é comercializado em pacotes chamado *toolboxes*, que rodam num console básico que executa comandos básicos de operações com matrizes. O LARA

possui licenças deste Software, que é utilizado em simulações e diversas outras aplicações nas pesquisas desenvolvidas neste laboratório.

Este software possui várias funções pré programadas, que são agrupadas nas *toolboxes*. A *toolbox* que possui funções relativas às redes neurais artificiais é denominada *Neural Network Toolbox™*, que em uma tradução grosseira significa “Caixa de Ferramentas para Redes Neurais”. Esta *toolbox* possui funções de treinamento, teste e operação de vários tipos redes neurais, que podem ser acessadas através de linhas de comandos, no console do próprio MatLab®, ou para algumas funções específicas através de uma interface gráfica como mostra a figura 4.5.

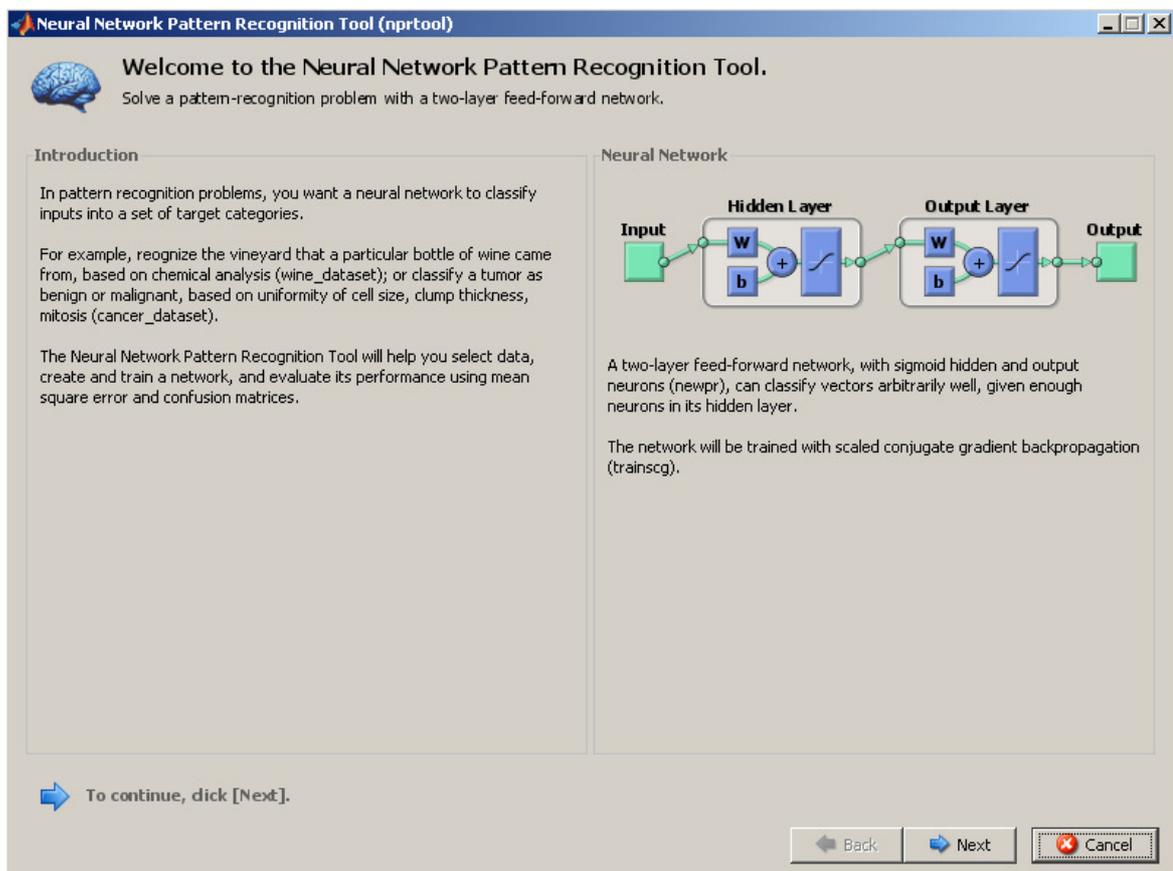


Figura 4.5 – Tela da interface gráfica de treinamento e teste da *Neural Network Toolbox™* [54]

Dois dos métodos de classificação aplicados neste trabalho foram implementados através da *toolbox* de redes neurais do MatLab®, as redes MLP e as LVQs, ambas possuem funções pré programadas nesta *toolbox*. As redes MLP podem ser configuradas e treinadas via comando a partir do console do MatLab®, ou a partir de uma

interface gráfica como a mostrada na figura 4.5. As redes LVQ são configuradas e treinadas a partir de comando do console. Para ambas as redes, caso o processo de treinamento seja iniciado via comando do console, a *toolbox* irá ativar uma interface para acompanhamento da evolução do treinamento, onde será possível inclusive ver a evolução de seu desempenho com os conjuntos de dados de teste e checagem.

O treinamento de uma rede neural, quer seja utilizando linhas de comando no console do MatLab[®], ou utilizando as interfaces gráficas da *toolbox*, irá criar uma variável do tipo *'net'*, na qual são armazenados todos os dados relativos à rede neural treinada (configuração de entradas e saídas, quantidade de camadas ocultas e neurônios nestas camadas, pesos sinápticos das ligações entre os neurônios da rede, etc). Esta variável, que é a rede neural propriamente dita, deve ser salva para que possa ser utilizada na aplicação da rede na classificação dos padrões desejados.

4.1.4 – LIBSVM: Biblioteca de Máquina de Vetor de Suporte

A SVM também foi escolhida como ferramenta computacional de classificação. Desta ferramenta foi encontrado um pacote de software livre denominado LIBSVM [55], que foi inicialmente desenvolvido por pesquisadores chineses do Departamento de Ciência da Computação e Engenharia da Informação da Universidade Nacional de Taiwan. Por se tratar de software livre acabaram por receber colaborações de outros pesquisadores de outras tantas instituições, expandindo sua aplicação e diversificando sua plataforma.

Atualmente é composta por aplicativos que rodam em sistemas operacionais como Linux, Windows, Unix, baseados em linguagem de programação como Java[™], C e Python[™], possuindo ainda interfaces para ferramentas computacionais como o MatLab[®], o Scilab e o Labview. A LIBSVM faz classificação e regressão de dados, podendo no caso de classificação trabalhar com classes binárias e também multi-classes. Para isso, conta com diferentes tipos de núcleos: de função linear, de função polinomial, de função de base radial e de função sigmóide.

A LIBSVM possui um aplicativo para treinamento da SVM e outro distinto para fazer efetiva classificação dos dados. Do treinamento surge um arquivo cuja extensão será *'model'*. Este arquivo conterá os dados estruturais da rede SVM e os parâmetros de classificação, ou seja, os vetores de suporte que permitirão fazer a separação entre as classes previamente treinadas.

Ao final do processo de treinamento da SVM, o aplicativo apresenta o desempenho da máquina treinada, que é medido através dos conjuntos de dados de teste e checagem. O processo de classificação é feito em batelada, ou seja, sempre *off line* e através de um arquivo contendo uma quantidade de dados a serem classificados. Este processo produz um arquivo com as respostas das classes para cada conjunto de variáveis de entrada.

4.2 – MÉTODOS

4.2.1 – Definição de Espaços, Ambientes e Pontos de Coleta de Dados

O experimento que resultou nos dados coletados para o presente trabalho ocorreu entre os dias 13 e 16 de dezembro de 2011. Na oportunidade, foram coletados 14.002 pulsos emitidos por uma das 80 *tags* disponíveis no LARA. Foi delimitada a mesma área adotada por Fonseca [1] e Souza e Oliveira [2] em seus experimentos, para compor a área dos experimentos do presente trabalho.

Desta forma, as bancadas de controle de processos, inspeção de linhas de transmissão e visão computacional, delimitam o espaço definido como ambiente 1, a sala de reuniões delimita o ambiente 2 e a área de acesso pela porta com controle de acesso biométrico, onde ficam os escaninhos do LARA mais algumas bancadas e mesas de trabalho próximas, delimitam o espaço do ambiente 3. Todos estes ambiente estão dispostos em uma área de 6m de largura por 11m de comprimento, cujo leiaute pode ser visto na figura 4.6.

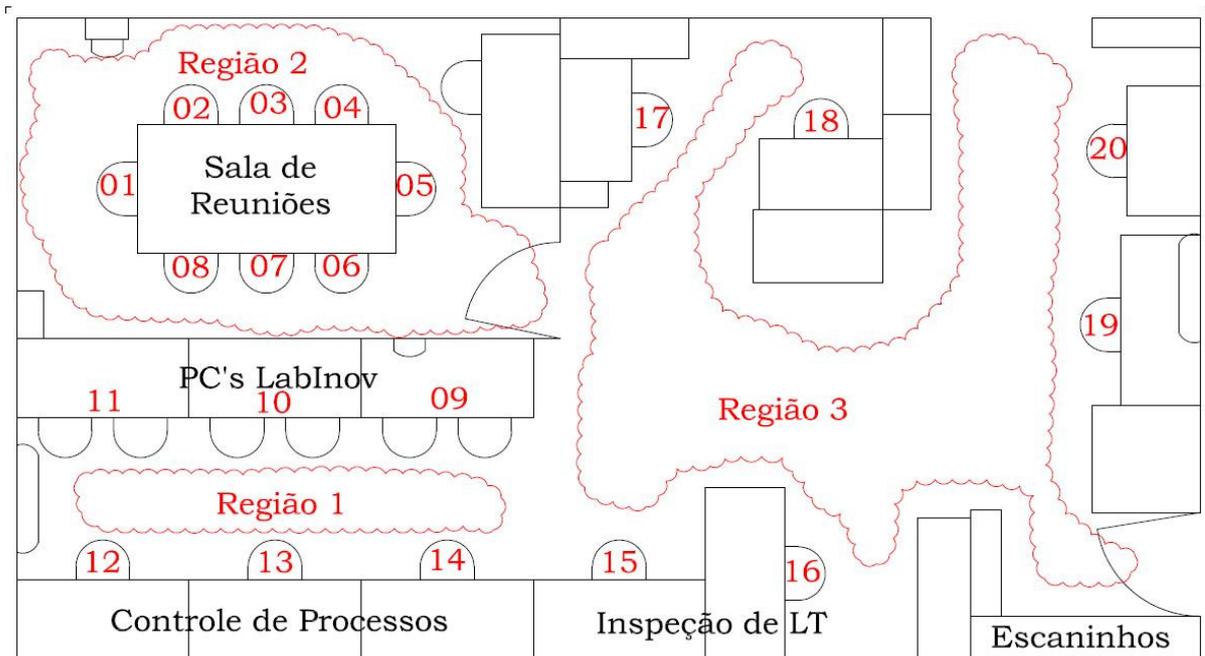


Figura 4.6 – Leiaute da área experimental definida no LARA, com indicação dos pontos de coleta de dados

Em cada um dos pontos indicados na figura 4.6, e mais nas regiões delimitadas em vermelho nesta mesma figura, foram coletados pelo menos 30 minutos de sinal emitido pela *tag*, que estava sempre sendo portada por um usuário do laboratório. Os usuários que portaram a *tag* durante os experimentos a utilizaram como se fosse um crachá pendurado no pescoço, através de cordão próprio. As medições foram efetuadas desta forma, para se propiciar no experimento uma situação mais próxima da realidade a que será submetido o sistema, quando implementado para localizar os usuários do LARA.

Durante a coleta dos dados, o ambiente a que estavam submetidos a *tag* e o usuário não ficou isolado, permitindo o acesso a outros usuários que tendo necessidade se aproximavam e interagiam com o usuário em observação. A única restrição submetida ao usuário foi a de que não deveria deixar a sua posição antes de finalizado o tempo de coleta, podendo entretanto, ter liberdade de movimentos em torno de seu próprio eixo vertical, e até mesmo pequenos deslocamentos no plano horizontal.

4.2.2 – Definição de Posições e Formas de Instalação das Leitoras

O posicionamento definido para as leitoras no plano horizontal, no experimento realizado, pode ser visualizado na figura 4.7. Este posicionamento é diferente do

posicionamento utilizado por Fonseca [1] e por Souza e Oliveira [2], que utilizaram 5 leitoras para fazer seus experimentos, enquanto que neste experimento foram utilizadas somente 4 leitoras. Na aplicação pretendida por Fonseca [1] e Souza e Oliveira [2], esperava-se a localização retornada em coordenadas cartesianas no plano horizontal, utilizando-se o método de triangulação de forma indireta, através de treinamento de redes neurais do tipo MLP, daí a opção de utilização de todas as leitoras disponíveis.

Na abordagem deste trabalho o que se pretende é a localização por ambientes, para se ter o controle de utilização de recursos do sistema de ar condicionado e iluminação destes ambientes. Buscou-se uma configuração que teoricamente permitisse atingir este objetivo e ao mesmo tempo fosse mais enxuta, no que diz respeito a recursos de hardware utilizados.

Numa teoria simplista, como a de que a propagação do sinal eletromagnético é inversamente proporcional ao quadrado da distância, imagina-se que, situando as leitoras de forma simétrica, ou seja, distribuídas em eixos, a proporcionalidade dos níveis de sinal seja mais facilmente percebida, daí a disposição indicada na figura 4.7.

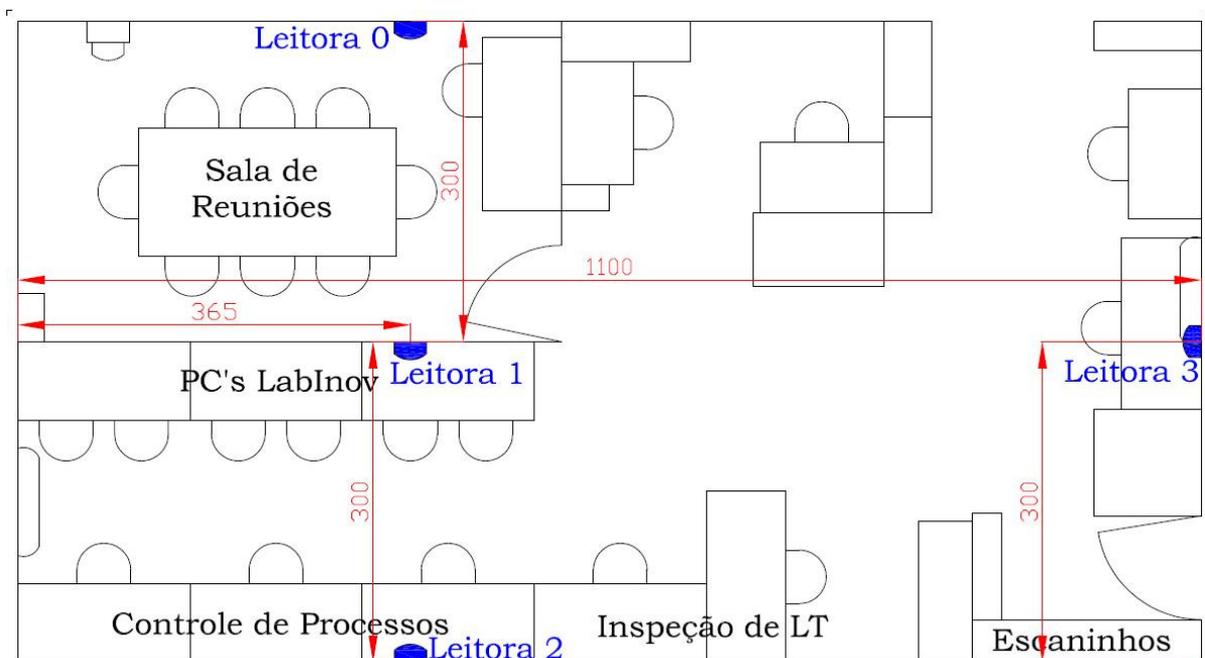


Figura 4.7 – Leiaute da área experimental definida no LARA, com indicação do posicionamento das leitoras

As leitoras foram instaladas, de forma que a extremidade de sua antena ficasse apontada para o teto e posicionada a 1,2m de altura do nível do piso. As leitoras 0, 1 e 2, instaladas dentro dos ambientes 1 e 2, respectivamente bancadas de controle de processos e sala de reuniões, foram alinhadas conforme a figura 4.7, buscando evitar que pontos de ambos os ambientes tivessem simetria nos níveis de sinal apurados por estas três leitoras. Com este posicionamento esperou-se, por exemplo, que pontos no ambiente 1 – bancadas de controle de processos – tivessem maiores níveis de sinal apurados nas leitoras 1 e 2, enquanto que pontos situados no ambiente 2 – sala de reuniões – apresentassem maiores níveis de sinal nas leitoras 0 e 1.

A leitora 3 foi instalada dentro do ambiente 3, que é o maior em área de circulação, entretanto é o menor em número de pontos onde os usuários param para exercer suas atividades. Com a instalação desta leitora neste ambiente pretendeu-se ter um indicador, através do nível de sinal apontado nela, de que o usuário estava no ambiente 3 e não nos ambientes 1 ou 2. A figura 4.8 mostra algumas fotos das leitoras instaladas nos ambientes do LARA. Na figura 4.8(a) temos foto das leitoras 1, em primeiro plano e leitora 0 ao fundo. A figura 4.8(b) mostra novamente a leitora 1 em primeiro plano e a leitora 2 ao fundo. No detalhe mostrado na figura 4.8(c), observa-se a leitora 2, instalada em sentido vertical inverso à leitora de número 4, não utilizada nos experimentos principais deste trabalho, mas utilizada para um teste comparativo, junto a leitora 2, acerca da influência da orientação vertical da antena sobre as leituras do RSSI.

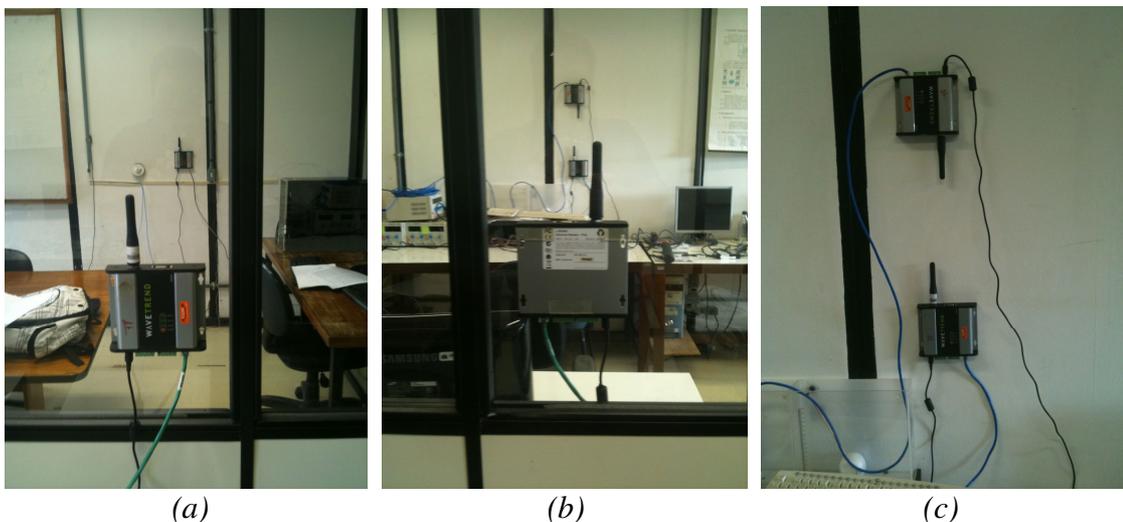


Figura 4.8 – Fotos das leitoras instaladas nos ambientes do LARA

4.2.3 – Leitura e Pré-tratamento dos Dados

Como já informado no item 4.1.2, para a coleta dos dados foi utilizado o software desenvolvido por Fonseca [1], em função da facilidade do seu manuseio e do tipo de arquivo por ele fornecido. Então, para cada ponto e região definida e indicada anteriormente na figura 4.6, foram efetuadas coletas de pulsos emitidos pela *tag* ativa, por um tempo de, pelo menos, 30 minutos. Estas informações foram gravadas em arquivos de texto, sendo um arquivo para cada ponto ou região. Cada linha de arquivo contém o registro de um pulso emitido por uma leitora.

Desta forma, cada linha de arquivo traz a informação da identificação da leitora; do ID da *tag*, que nesta fase não variou, pois foram feitos os experimentos sempre com a mesma *tag*; do nível do sinal percebido pela leitora expresso pelo RSSI; do *Age Counter* que é incrementado a cada pulso; do tipo de pulso, indicado por um *flag* de movimento (pulso de 1,5s ou de 15s). E por último trás o registro da data e hora em que o *middleware* efetuou a leitura do registro. Na tabela 4.1 tem-se uma amostra de dados coletados na disposição em que é feito o arquivamento destes dados.

Tabela 4.1 – Exemplo de dados gerados no arquivo de log pelo middleware

Leitora	ID da TAG	Nível RSSI	Age Counter	Flag de Movimento	Data e Hora da Leitura
3	991698	81	3461690	1,5	14/12/2011 14:43:17
2	991698	93	3461690	1,5	14/12/2011 14:43:17
1	991698	107	3461690	1,5	14/12/2011 14:43:17
0	991698	110	3461690	1,5	14/12/2011 14:43:17
3	991698	82	3461691	15	14/12/2011 14:43:32
0	991698	106	3461691	15	14/12/2011 14:43:32
2	991698	90	3461691	15	14/12/2011 14:43:32
1	991698	106	3461691	15	14/12/2011 14:43:32

Os dados coletados totalizaram 14.002 amostras, armazenadas em 22 arquivos do tipo log de texto, sendo 3.301 amostras referentes ao ambiente número 1, 4.750 amostras referentes ao ambiente de número 2, e 5.951 amostras referentes ao ambiente número 3. Após a filtragem, para que pudessem ser utilizadas no treinamento das ferramentas de classificação de dados, estas amostras foram aglutinadas em um único arquivo, do tipo planilha eletrônicas

do Excel[®], onde cada amostra teve adicionado mais um campo com a informação referente ao ambiente em que fora coletada.

O pré-processamento destas amostras ocorreu praticamente de forma manual, o que permitiu perceber algumas falhas no processo de coletas, bem como as prováveis causas destas falhas, que serão descritas a seguir e exemplificadas nas tabelas que se sucedem.

A primeira falha percebida, vide tabela 4.2, ocorre quando no registro de determinado pulso emitido pela *tag*, o *middleware* registra uma leitura com um número de leitora inexistente (geralmente com dois dígitos) e outra com o campo referente ao ID da leitora vazio.

Tabela 4.2 – Exemplo de falha do middleware no registro do ID da leitora

Leitora	ID da TAG	Nível RSSI	Age Counter	Flag de Movimento	Data e Hora da Leitura
2	991698	101	3462660	1,5	14/12/2011 15:41:38
3	991698	91	3462660	1,5	14/12/2011 15:41:38
0	991698	116	3462660	1,5	14/12/2011 15:41:38
1	991698	103	3462660	1,5	14/12/2011 15:41:38
3	991698	87	3462661	1,5	14/12/2011 15:41:40
	991698	107	3462661	1,5	14/12/2011 15:41:40
12	991698	100	3462661	1,5	14/12/2011 15:41:40
0	991698	112	3462661	1,5	14/12/2011 15:41:40
0	991698	112	3462662	1,5	14/12/2011 15:41:42
2	991698	100	3462662	1,5	14/12/2011 15:41:42
3	991698	88	3462662	1,5	14/12/2011 15:41:42
1	991698	107	3462662	1,5	14/12/2011 15:41:42

Comparando com os registros anteriores e posteriores, foi identificado que o ID de leitora inexistente é, na verdade, a composição do número de duas leitoras, sendo que no registro no qual aparece o ID de leitora inexistente, o dígito correspondente à unidade é o ID da leitora do registro e o número da dezena corresponde ao ID da leitora cujo registro ficou sem número de ID. Mesmo Fonseca [1] (informação verbal)¹ percebeu esta falha no *middleware* quando da execução de seus experimentos, mas não soube atribuir a que se deve a ocorrência deste tipo de erro.

¹ Informação compartilhada pelo Eng^o Lucas Oliveira da Fonseca em reunião de trabalho por telefone, no dia 13/01/2012.

É possível que seja um erro sistêmico do *middleware*, sem relação com os equipamentos do sistema de localização. Na tabela 4.3 observa-se a correta disposição dos dados.

Tabela 4.3 – Exemplo de correção da falha do middleware no registro do ID da leitora

Leitora	ID da TAG	Nível RSSI	Age Counter	Flag de Movimento	Data e Hora da Leitura
3	991698	87	3462661	1,5	14/12/2011 15:41:40
1	991698	107	3462661	1,5	14/12/2011 15:41:40
2	991698	100	3462661	1,5	14/12/2011 15:41:40
0	991698	112	3462661	1,5	14/12/2011 15:41:40

Durante o manuseio dos dados coletados, observou-se que o menor nível de sinal detectado apresentava RSSI igual a setenta, logo se associou que este valor (RSSI = 70) seria algo como um limite inferior de corte, em que as leitoras ou não mais detectam o sinal emitido pelas *tags*, ou ainda, que detectam o sinal, porém devido às inconsistências apresentadas nos dados recebidos, descartam as informações, não registrando assim o pulso em questão e gerando uma lacuna nos valores de RSSI registrados na base de dados.

Na segunda falha percebida não ocorre o registro de determinado pulso em uma das leitoras, mesmo quando a *tag* está a uma distância em que o nível do sinal recebido pela leitora esteja sabidamente acima deste limite inferior identificado, como pode ser observável nos registros anterior e posterior destacados em vermelho na tabela 4.4.

Fonseca [1] (informação verbal)² atribui estas falhas, que ocorrem mesmo quando há níveis de sinal acima do limite inferior de corte, a algum erro de checagem, que no modelo de leitora utilizada é baseado em algoritmo CRC (do inglês, *Cyclic Redundancy Check*).

² Informação compartilhada pelo Eng^o Lucas Oliveira da Fonseca em reunião de trabalho por telefone, no dia 13/01/2012.

Tabela 4.4 – Exemplo de pulso não registrado em determinada leitora mesmo com nível de sinal acima do limite inferior de corte

Leitora	ID da TAG	Nível RSSI	Age Counter	Flag de Movimento	Data e Hora da Leitura
1	991698	108	3461653	1,5	14/12/2011 14:41:41
0	991698	99	3461653	1,5	14/12/2011 14:41:41
2	991698	86	3461653	1,5	14/12/2011 14:41:41
3	991698	78	3461653	1,5	14/12/2011 14:41:41
0	991698	99	3461654	1,5	14/12/2011 14:41:43
2	991698	88	3461654	1,5	14/12/2011 14:41:43
3	991698	77	3461654	1,5	14/12/2011 14:41:43
1	991698	108	3461655	1,5	14/12/2011 14:41:52
3	991698	79	3461655	1,5	14/12/2011 14:41:52
2	991698	95	3461655	1,5	14/12/2011 14:41:52
0	991698	99	3461655	1,5	14/12/2011 14:41:52

A terceira falha percebida é na verdade uma limitação do sistema, pois ocorre em função do sinal estar próximo do limite de corte inferior para determinada leitora. Quando ocorre uma situação em que a *tag* está numa distância de determinada leitora, que resulte num nível de sinal muito próximo do limite inferior de corte, invariavelmente ocorrerão perdas mais frequentes no registro de pulsos nesta leitora, ocasionando as lacunas na base de dados coletada.

Como foi citado no item 4.1.1.1, os experimentos práticos demonstraram que, para ambientes confinados, e principalmente bastante carregados como no caso do LARA, a distância crítica entre *tag* e leitora, para que comece a ocorrer perdas de registro em função do limite inferior de corte do sinal RSSI, é da ordem de 7m. Esta constatação trouxe certa preocupação, em função das dimensões do LARA, pois o ambiente definido para os testes possui dimensão de 6m de largura por 11 metros de comprimento.

A quarta falha observada ocorreu durante a coleta de dados no ponto de número 12, que fica bem próximo de um dos extremos do ambiente 1, no lado oposto ao posicionamento das leitoras. Durante o processo de coleta, num período de aproximadamente 5 minutos, ocorreu a perda de registro de pulsos sucessivos por todas as leitoras. Em média, foram perdidos cerca de 3 a 4 pulsos e, mesmo quando surgia um registro, quase sempre se apresentava somente em uma ou duas leitoras, geralmente as mais próximas. Esta falha foi

denominada de ponto cego que se presume ocorrer sempre que a *tag* estiver à distância da ordem de 7 metros ou mais das leitoras, ou ainda obstruída por algum objeto que impeça a propagação do sinal das *tags*.

A ocorrência de falhas desta natureza, ou seja, a ocorrência de lacunas nos valores de RSSI durante a operação de um sistema de controle, baseado neste sistema de localização, poderá acarretar em um comportamento indesejado do sistema de controle, como desligamento do sistema de ar condicionado, pois poderá sinalizar a ausência do usuário, mesmo quando este estiver presente em um dos ambientes. Logo, o tratamento deste tipo de falha será indispensável num projeto de aplicação, como se pretende fazer no LARA.

Uma vez identificadas as possíveis falhas no registro de pulsos nas leitoras, para o prosseguimento dos experimentos foram então adotadas três diferentes metodologias, que visaram suprir as lacunas nos valores de RSSI, no tratamento de tais falhas. A aplicação destas metodologias permitiu ampliar a quantidade de base de dados, totalizando assim três variantes de dados disponíveis para o treinamento e teste dos classificadores, e assim verificar o desempenho do sistema, sob diferentes formas de tratamento de falhas. A tabela 4.5 mostra a formulação de cada uma destas metodologias para atuar na definição do valor de RSSI não registrado.

Tabela 4.5 – Metodologias para definição do valor de RSSI não registrado pelas leitoras

Metodologia	Formulação
Metodologia 1	Sem alteração (dados conforme coleta)
Metodologia 2	$RSSI = 71$
Metodologia 3	$RSSI_{(i)} = RSSI_{(i-1)}$

Com base nesta definição, os dados coletados formam a primeira base de dados para treinamento dos classificadores. Para a segunda metodologia, o valor que se tem atribuído ao RSSI, na ocorrência de falta de registro, é o valor imediato ao do limite inferior de corte, ou seja, $RSSI = 71$. Já para a terceira metodologia será o valor do registro anterior, ou seja, $RSSI_{(i)} = RSSI_{(i-1)}$.

Estas três bases de dados foram ainda submetidas a um processo de filtragem, baseado no algoritmo do Filtro de Kalman Estendido, o que dobrou a quantidade de bases de

dados passando assim para 06 (seis) diferentes bases de treinamento e teste dos classificadores. Isto proporcionou um maior leque de comparação de desempenho dos classificadores, dentro de perspectivas diferentes que envolvem ações simples no tratamento de falhas.

O filtro de Kalman estendido foi implementado, de acordo com o algoritmo abordado no item 3.1, na ferramenta computacional MatLab[®], em função de sua facilidade de operacionalizar cálculos com matrizes. Como todos os experimentos foram conduzidos de forma *off line*, a filtragem foi implementada em processamento de batelada. Neste procedimento, os dados coletados pelo *middleware* foram salvos em arquivos de log de texto, depois transformados em arquivos de planilha eletrônica, juntando-se vários arquivos de log de texto, e finalmente processados no MatLab[®] através de um programa com o algoritmo do filtro de Kalman estendido.

Como citado no item 3.1, a aplicação do filtro de Kalman pressupõe conhecimento dos estados iniciais do sistema, bem como da covariância do ruído do sistema e da medição. Pressupõe, ainda, o conhecimento de um modelo matemático que represente de forma mais fiel, o sistema sob intervenção do filtro, pois isto implica diretamente no desempenho desta ferramenta computacional.

No caso do RSSI, variável a ser filtrada pelo EKF, a formulação matemática segue o modelo de propagação de onda eletromagnética, vastamente encontrado na literatura acadêmica da área que trata o assunto, a exemplo de Aguirre [47]. Entretanto este modelo, quando aplicado às condições de propagação a que foram submetidos os experimentos do presente trabalho, ou seja, propagação em canal de comunicação *indoor*, fica sujeito a efeitos já conhecidos de interferência, mas de difícil modelagem, como os caminhos múltiplos.

Desta forma, Souza e Oliveira [2] procederam um experimento, cujos dados coletados foram aplicados à *Toolbox Curve Fitting* do MatLab[®], resultando na equação 4.1 que a seguir é mostrada. É preciso ressaltar que os dados geradores deste equacionamento são provenientes de um experimento em ambiente *outdoor*, logo não expressam exatamente o comportamento de propagação nas condições do ambiente de pesquisa do presente trabalho. Entretanto, se aproxima bastante do comportamento esperado, visto que fora efetuado com os

mesmos equipamentos que atenderam aos experimentos aqui descritos, considerando desta forma alguma particularidade que estes equipamentos possam apresentar.

$$RSSI = \frac{12.790}{(d + 12,3)^2} + 56,71 \quad (4.1)$$

Como se verifica na equação 4.1, o RSSI é resultado de uma relação inversa do quadrado da distância entre a *tag* e a leitora RFID. Esta distância entre *tag* e leitora está representada na equação pela letra *d*. Logo, isto implica que a condição inicial desta distância tem que ser conhecida para aplicação do EKF. Como já mencionado anteriormente, todos os experimentos foram conduzidos *off line*, assim esta definição da distância inicial foi de fácil lida. Outra importante informação acerca desta variável é que o processo de filtragem foi aplicado de forma independente para cada ponto de coleta, desse modo a distância entre a *tag* e a leitora RFID não variou durante o período de coleta, visto que os usuários tiveram restrição de deslocamento no plano horizontal. Quanto ao estado inicial do sistema, para o presente trabalho foi considerada como tal, a primeira leitura efetuada em cada ponto de coleta.

Sobre o estado inicial das matrizes de covariância, antes de defini-lo, foi preciso entender algumas considerações sobre o sistema posto em observação. Como já visto, o EKF atuou em uma variável (o RSSI) que dependeu, segundo o equacionamento do modelo aplicado, somente da variável distância. Isto implica que a distância entre a *tag* e uma determinada leitora exerce influência somente no valor de RSSI percebido por esta leitora, não influenciando nos resultados das demais leitoras. Então, baseado nesta fundamentação, se observa, que as matrizes de covariância são diagonais, onde cada elemento da diagonal representa a relação de variância entre a distância da *tag* a uma leitora e o valor percebido de RSSI por esta leitora.

A determinação exata do estado inicial desta matriz de covariância não é tarefa muito fácil, pois depende de fatores não equacionáveis como a influência de caminhos múltiplos no valor do RSSI. Esta determinação foi feita de forma empírica aplicando valores a esta covariância, experimentando o EKF e observando o comportamento dos sinais filtrados. Foi considerada a mesma covariância para as quatro leitoras usadas nos experimento do presente trabalho, sendo o valor numérico de 100 (cem) para a covariância do ruído do

sistema e o valor numérico de 700 (setecentos) para a covariância do ruído de medição. A matriz de covariância dos estados iniciais foi determinada em um valor numérico de 300 (trezentos).

Para o cálculo das Jacobianas foi necessário definir, como se observa na equação 4.2, a derivada de primeira ordem da função que rege o sistema. Como já mencionado, a coleta de dados e a aplicação do filtro de Kalman se deu de forma isolada para cada ponto de coleta, isto implicou que a distância entre *tag* e leitora não variou para estas duas atividades. Por consequência o valor do RSSI calculado e também de sua derivada não variou para um mesmo ponto de coleta.

$$RSSI' = \frac{-2 \times 12.790}{(d + 12,3)^3} \quad (4.2)$$

4.2.4 – Preparação dos Arquivos de Treinamento e Teste dos Classificadores

O treinamento dos classificadores exigiu a preparação dos arquivos das 06 (seis) bases de dados oriundos do pré-tratamento e filtragem dos dados. Fisicamente estas seis bases de dados foram acondicionadas no mesmo arquivo do tipo planilha eletrônica, separados somente por posições de colunas diferentes dentro desta planilha eletrônica. Isto foi feito para facilitar o manuseio destes dados, principalmente no tocante à carga destas bases de dados no MatLab[®], ferramenta utilizada para o treinamento das redes neurais do tipo MLP e LVQ, já citadas como classificadores escolhidos para o teste de desempenho no sistema proposto.

Nesta versão de arquivo com as bases de dados juntas, foram excluídos todos os campos desnecessários ao processo de treinamento, sendo mantidos somente aqueles relativos à leitora de origem do dado registrado, para composição do vetor de entrada, o campo de identificação do ambiente de origem em que o pulso gerador do dado fora emitido, e o *Age Counter* para garantir a referência de origem do dado nos arquivos de coleta. Sendo estes dois últimos campos comuns a todas as bases de dados, não houve necessidade de replicação a cada uma das bases.

A planilha eletrônica contendo as 06 (seis) bases de dado foi, então, submetida a um tratamento no MatLab[®] para fazer o embaralhamento dos dados, de forma a permitir uma apresentação destes dados com as amostras, proporcionalmente bem misturadas, dos três ambientes, situação desejável num processo de treinamento de redes neurais. Ao final deste processo de embaralhamento, o mesmo programa fez a separação de dados de teste numa proporção de 20% dos dados totais disponíveis para o processo de treinamento e teste, originando assim dois arquivos diferentes em planilha eletrônica, sendo um para treinamento e outro para teste das redes neurais.

A tabela 4.6 mostra a configuração de cada linha de um conjunto de dados de interesse no processo de treinamento ou teste das redes neurais. No momento inicial do processo de treinamento, um conjunto de dados, no formato indicado é separado do arquivo de treinamento e disponibilizado à *toolbox* em forma de variáveis no *workspace*. Este processo pode ser efetuado de forma manual no console do MatLab[®] ou, ainda, de forma automatizada via composição de um programa que realiza tais procedimentos.

Tabela 4.6 – Exemplo de disposição dos dados nos arquivos de treinamento das redes neurais

Leitora 0	Leitora 1	Leitora 2	Leitora 3	Ambiente	Age Counter
80	90	103	77	2	3461215
98	97	96	66	3	3484774
85	58	87	117	3	3482947
109	102	96	89	3	3476928
94	104	123	86	1	3472520
98	100	97	82	1	3475457
99	118	94	83	2	3461952
116	122	97	87	2	3459695
109	107	97	52	2	3457959
97	99	96	97	3	3485223

O treinamento das SVM exigiu a preparação de arquivos distintos para o treinamento e para o teste em cada uma das bases de dados disponibilizadas a estes procedimentos. Tais arquivos foram preparados primeiro separando em arquivos do tipo CSV, cada uma das bases de dados, tanto para treinamento quanto para teste, com os campos necessários a estes procedimentos, segundo conforme uma disposição indicada pelos autores da ferramenta de treinamento e aplicação das SVM, a LIBSVM.

Tal disposição teve ainda que ser adequada à formatação, que pode ser observada na tabela 4.7, também indicada pelos autores da LIBSVM. Os moldes exigidos pela ferramenta computacional LIBSVM, pedem que os dados sejam dispostos em arquivo do tipo texto, com os campos apresentados na seguinte ordem: $\langle label \rangle \langle index1 \rangle : \langle value1 \rangle \langle index2 \rangle : \langle value2 \rangle \dots$. Cada linha deste arquivo representa um padrão de entrada a ser classificado, onde $\langle label \rangle$ é um número inteiro que representa a classe do padrão apresentado e os pares $\langle index \rangle : \langle value \rangle$ representam uma característica de entrada, em que, $\langle index \rangle$ é um número inteiro, partindo do número 1, que representa uma das entradas da rede e, $\langle value \rangle$ é o valor apresentado naquela entrada, expresso por um número real. Os pares $\langle index \rangle : \langle value \rangle$ devem ser dispostos em ordem ascendente na linha.

Tabela 4.7 – Exemplo de disposição dos dados nos arquivos de treinamento da SVM pela ferramenta computacional LIBSVM

Linha	Características de Entrada
1	2 1:105 2:117 3:87 4:83
2	2 1:95 2:100 3:97 4:74
3	3 1:94 2:74 3:84 4:86
4	3 1:86 2:77 3:92 4:83
5	3 1:94 2:80 3:91 4:85
6	3 1:83 2:80 3:83 4:73
7	3 1:71 2:76 3:91 4:81
8	1 1:81 2:85 3:108 4:81
9	1 1:90 2:112 3:96 4:79
10	2 1:106 2:92 3:90 4:80

Na divisão dos dados entre treinamento e teste foi guardada a mesma proporção usada no treinamento das redes neurais, pois foram utilizadas as mesmas bases de dados para ambos os procedimentos. Isto garantiu inclusive que o desempenho dos testes fosse aferido com o mesmo grupo de dados.

4.2.5 – Treinamento dos Classificadores

O treinamento da rede MLP foi executado utilizando a interface gráfica da *Neural Network Toolbox™*, que por sua vez aplica o algoritmo *backpropagation* como ferramenta computacional de treinamento da referida rede. Normalmente, em um processo de treinamento de uma rede MLP é comum se dividir os dados disponíveis para o treinamento

em até três diferentes grupos. Um grupo de treinamento, geralmente com cerca de 70% a 80% dos dados; um grupo de teste que pode variar entre 10% e 20% e, se o programa computacional utilizado para treinamento necessitar, pode ainda existir um grupo de dados específicos para checagem, contendo cerca de 10% a 15% dos dados totais.

A interface gráfica da *Neural Network Toolbox*TM possui um dispositivo no qual, durante a configuração da futura rede a ser treinada, é feita a escolha dos percentuais de cada grupo de dados, o que facilita o teste de diferentes configurações de grupos de dados. No caso específico do presente trabalho, foram adotados os percentuais de 90% para o grupo de treinamento, 5% para o grupo de teste e 5% para o grupo de checagem. A checagem desempenha o importante papel de verificar se o treinamento ainda está convergindo para um valor melhor de acerto da rede, ou seja, se ela ainda está generalizando, ou se está somente se especializando no grupo de dados de treinamento.

Cabe ressaltar que esta divisão foi efetuada dentro do grupo de dados separados para treinamento, por se tratar de exigência da interface gráfica de treinamento da Rede MLP. Por outro lado, esta interface também permitiu a implementação de um teste, com a rede já treinada, por um outro grupo de dados, que como foi citado já havia sido separado para tal aplicação. Logo, o desempenho anotado para comparação com as demais técnicas de classificação foi o desempenho resultante do teste com este grupo de dados.

A interface gráfica de treinamento de rede MLP exige ainda que aos dados de treinamento tenham os valores referentes às entradas e saídas guardados em variáveis diferentes, logo tal separação deve ser feita no MatLab[®], antes de se iniciar o processo de treinamento, seja através de linha de comando no console, ou via programação em arquivo de programa, com uma sequência de comandos. Ao final do treinamento da rede é gerada uma variável denominada “*net*” no *workspace* do MatLab[®], com os parâmetros da rede e os valores dos pesos sinápticos das ligações entre os neurônios, que deve ser guardada, pois se trata da rede neural propriamente dita. Esta variável deve ser resgatada para que se possa aplicar a rede na classificação de futuros dados.

O treinamento da rede LVQ é iniciado através de comando no console do MatLab[®], pois a interface gráfica da *Neural Network Toolbox*TM atua somente no acompanhamento da evolução do treinamento/aprendizagem. Na configuração da rede, passo

anterior ao treinamento, é necessário informar a proporção entre as classes do grupo de amostras disponíveis para o treinamento. Assim, como no processo de treinamento da rede MLP, é necessário que os dados de entrada e de saída estejam em variáveis diferentes para que a função de treinamento possa efetivamente ser aplicada.

No caso dos dados relativos às saídas, que são as indicações das classes, estes geralmente são definidos através de índices numéricos que devem ser transformados em vetores binários com a indicação de qual das saídas deve ser ativada. Isto é feito através de comando no console do MatLab[®]. O final do processo de treinamento da rede LVQ também gera uma variável denominada “*net*”, similar a gerada no processo de treinamento da rede MLP, e que deve receber o mesmo tratamento de guarda e recuperação para utilização na classificação de dados futuros.

A SVM foi treinada através do software LIBSVM na versão aplicativos do MS-Windows, pois durante a instalação da interface com MatLab[®] ocorreu uma incompatibilidade na conversão dos arquivos das funções em linguagem C para as funções do MatLab[®], o que inviabilizou o teste desta interface.

O treinamento, neste caso em que foram utilizados os aplicativos do MS-Windows, é iniciado por comando manual através do *prompt* de comando do DOS emulado no MS-Windows. Na execução de tal comando de treinamento são informadas as características da SVM a serem treinadas, como tipo de classificador, qual função de núcleo será utilizada e alguns parâmetros de configuração desta função, e ainda o nome do arquivo em que será guardada a SVM treinada. Este arquivo, produto final do processo de treinamento da SVM, porta todas as informações da referida configuração desta SVM e ainda o conjunto de vetores de suporte definidos durante o treinamento, e deve ser guardado para aplicações na classificação de dados futuros.

5 – APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Neste capítulo são apresentados os resultados dos vários experimentos realizados durante este trabalho. Estes experimentos podem ser divididos em três partes operacionalmente distintas: o planejamento e implementação da sistemática de localização dentro do leiaute definido do LARA; a coleta e prévio tratamento dos dados através de ferramentas como o *middleware*, planilhas eletrônicas e o filtro de Kalman estendido e, por último, o treinamento e teste dos classificadores. Também dentro desta sequência são apresentados os resultados obtidos durante os experimentos, nos itens que seguem.

5.1 – ORIENTAÇÃO VERTICAL DA ANTENA DA LEITORA

Com a definição do posicionamento das leitoras dentro do leiaute do LARA e o início de instalação das mesmas, surgiu uma dúvida acerca da influencia do sentido de instalação na direção vertical. Daí surgiu então o pequeno experimento, cujos resultados são apresentados neste item, quando foi instalada a leitora de número 4 junto à leitora de número 2, conforme foi mostrado na figura 4.8(c), ambas com a mesma configuração.

Para se coletar os dados, foi posicionado um usuário, de aproximadamente 1,70 m de altura, portando a *tag* RFID pendurada ao pescoço, como um crachá, a uma distância de 80 cm da parede onde estavam fixadas as duas leitoras. Na primeira bateria de coleta de dados, o usuário estava em pé, posicionando-se com seu lado esquerdo voltado para as leitoras, depois de frente para as leitoras, em seguida com seu lado direito voltado para as leitoras e por fim de costas para as leitoras. Na segunda bateria, foram feitas coletas de dados nas mesmas orientações, entretanto mudou-se o posicionamento do usuário que passou a ficar sentado durante as coletas.

A figura 5.1 mostra um comparativo entre as médias dos valores de RSSI coletados durante este experimento, sendo que, o valor destacado representa a média de todas as leituras coletadas neste experimento, indicando que há pouca relevância na diferença entre valores coletados entre os diferentes modos de instalação das leitoras.

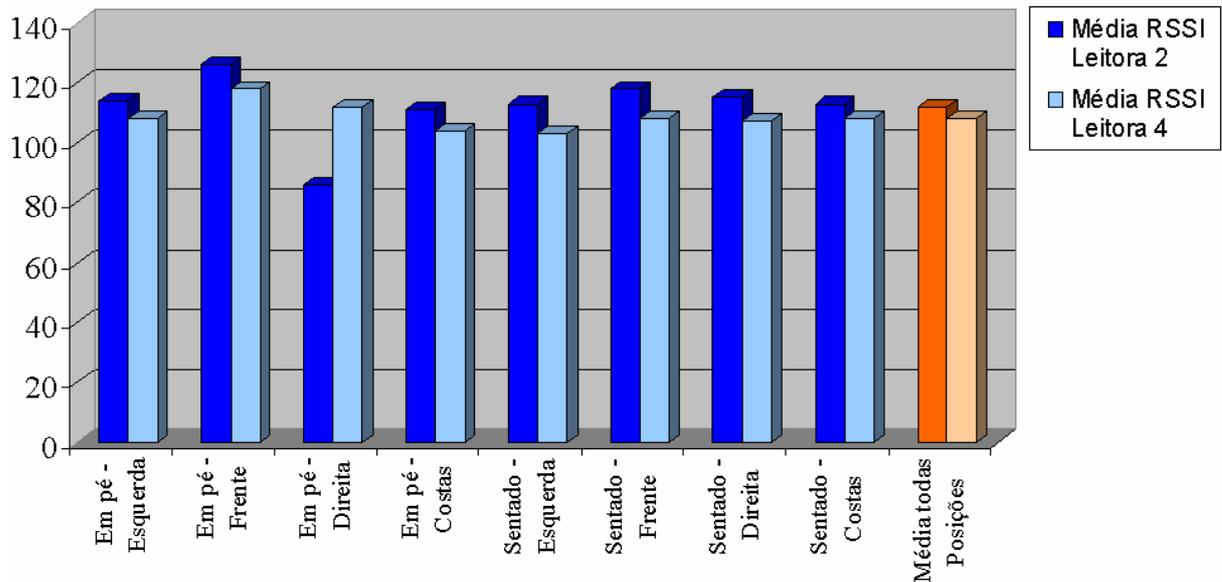


Figura 5.1 – Comparativo dos valores médios de RSSI coletados com antenas posicionadas verticalmente em sentidos inversos

No gráfico mostrado na figura 5.1, o único posicionamento que ofereceu um valor contraditório foi aquele em que o usuário se posicionou em pé, com seu lado direito voltado para as leitoras. Este posicionamento apresentou uma inversão entre os níveis de RSSI percebidos pelas antenas com, inclusive, certa discrepância entre a proporcionalidade destes valores.

Credita-se este comportamento à antena da *tag* que por questões construtivas pode ser influenciada por seu posicionamento em relação à antena da leitora. Na tabela 5.1 são apresentados os valores médios de RSSI coletados durante este experimento.

Tabela 5.1 – Valores médios de RSSI apurados nas Leitoras 2 e 4

Posição do Portador da TAG	Valor Médio do RSSI	
	Leitora 2	Leitora 4
Em pé - Esquerda	114	108
Em pé - Frente	126	118
Em pé - Direita	86	112
Em pé - Costas	111	104
Sentado - Esquerda	113	103
Sentado - Frente	118	108
Sentado - Direita	115	107
Sentado - Costas	113	108
Média dos valores	112	108

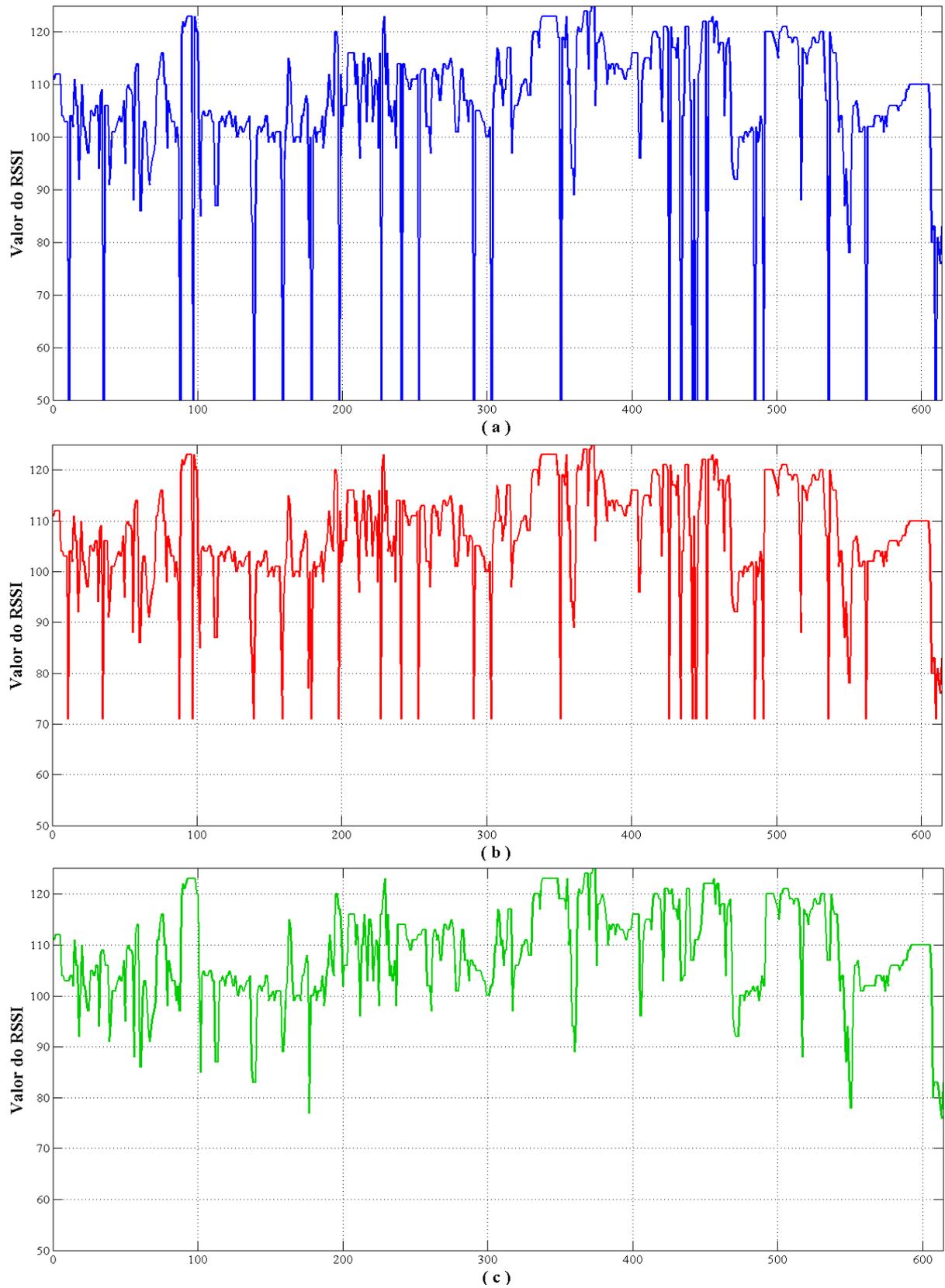
Para os demais posicionamentos foi mantida a proporcionalidade entre os valores percebidos pelas antenas, o que leva a indicação de que tanto um quanto outro direcionamento das antenas não promoveria diferença no resultado final do experimento a que se destinava este trabalho.

5.2 – APLICAÇÃO DO FILTRO DE KALMAN ESTENDIDO

Como definido no capítulo 4, que descreve a metodologia utilizada no presente trabalho, os dados relativos aos níveis de RSSI foram coletados em 20 diferentes pontos de três ambientes distintos dentro de um leiaute previamente limitado no LARA. Somaram ainda duas coletas em regiões de circulação dos ambientes 02 (sala de reunião do LARA) e 03 (sala de acesso ao LARA).

Também já foi explicitado que o modelo de propagação da onda eletromagnética em ambientes internos de edificações sofre forte influência dos efeitos dos caminhos múltiplos. Isto causa instabilidade no sinal e cria variações nos níveis de potência percebidos pela antena do receptor que prejudicam o relacionamento, tal como ocorre no modelo de propagação para ambientes externos, entre nível de potência recebida com a distância entre a fonte e receptor de sinal. A figura 5.2 expõe este tipo de ocorrência nos dados coletados, mostrando que o comportamento do RSSI coletado por uma leitora é comparável com a situação ilustrada por Rappaport [14] e destacada pela figura 2.1 no capítulo 2 deste trabalho. Rappaport [14] indica ainda que tal interferência pode ser advinda da proximidade entre a fonte e o receptor do sinal, na propagação de sinal em ambientes internos de edificações.

Soma-se ainda a todas estas inconveniências de propagação da onda eletromagnética, o fato de que o sistema de registro e coleta dos pulsos e seus respectivos dados associados sofrem perdas, muitas vezes atribuídas ao próprio sistema que, tendo sido projetado para outra aplicação, não necessita uma rigidez neste processo de registro e coleta. E, por fim, tem-se também a inconveniência de que o *middleware* acaba por contribuir com uma pequena parcela de erro de escrita dos dados coletados durante o processo de registro destes dados em um arquivo de log.



*Figura 5.2 – Comportamento do RSSI nos dados coletados no ponto 04 pela leitora 1,
(a) dados sem tratamento metodologia 1, (b) dados tratados pela metodologia 2 e
(c) dados tratados pela metodologia 3*

Todos estes processos geram uma base de dados com deficiências e impurezas e no caso deste trabalho não foi diferente. Logo, os conjuntos de dados coletados necessitaram passar por um pré-processamento, onde foram primeiro depurados alguns resultados incoerentes de coleta, como erros e perdas de registro, seja pelo *middleware* ou pelas leitoras.

Este primeiro pré-processamento foi efetuado de forma manual em que, também, ocorreu a aglutinação dos dados dos 22 conjuntos coletados em um único arquivo de planilha eletrônica, oportunidade em que foi inserida a informação do ambiente a que pertencia cada pulso registrado nestes conjuntos de dados.

Na última etapa do pré-processamento, as lacunas de valores incidentes na base de dados, ou seja, os registros em que determinado valor de RSSI foi atribuído como zero, por qualquer dos motivos elencados no item 4.2.3, foram tratadas como indicado neste mesmo item 4.2.3, gerando assim três grupos de dados para serem filtrados, os quais foram nomeados: Método 1 (dados sem tratamento); Método 2 (onde lacunas de RSSI foram substituídas por valor 71) e Método 3 (onde lacunas de RSSI foram substituídas por valores anteriores).

Finalizado o pré processamento dos dados, os grupos definidos no parágrafo anterior foram submetidos a filtragem através do EKF, implementado em um programa do MatLab[®], baseado em seu algoritmo. Para implementação deste algoritmo, foi necessário ajustar o modelo de propagação, segundo o comportamento de propagação registrado no experimento de Souza e Oliveira [2], pelo qual promoveram o deslocamento controlado de uma *tag*, distanciando-a de uma leitora a passos previamente definidos.

Este conjunto de dados permitiu, através da *Toolbox Curve Fitting* do MatLab[®], levantar uma curva de comportamento da propagação do sinal emitido por uma *tag*. A figura 5.3 mostra o contraste dos pontos medidos por Souza e Oliveira [2], com a curva definida pela *Toolbox Curve Fitting*.

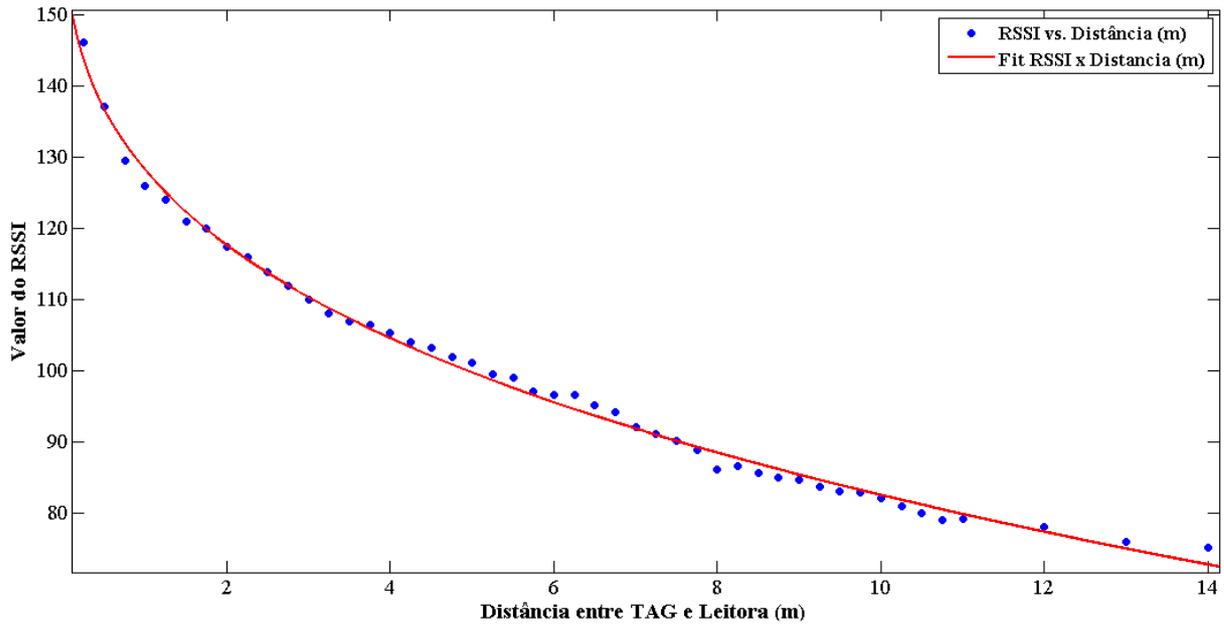


Figura 5.3 – Curve Fitting do modelo de propagação utilizado no EKF [2]

A aplicação da *Toolbox Curve Fitting* nos dados levantados por Souza e Oliveira [2] resultou no equacionamento mostrado abaixo para o modelo de propagação, onde o índice a teve valor definido em 12.790, o índice b teve valor definido em -12,3, o índice c teve valor definido em 56,71 e a variável x correspondeu à distância entre o transmissor e o receptor.

$$RSSI = \frac{a}{(x-b)^2} + c \quad (5.1)$$

Como se observa na equação 5.1, o RSSI possui relação inversa com a distância. Logo para aplicação deste equacionamento no algoritmo do EKF, foi necessário determinar a distância entre cada posição de coleta de dados definida nos três ambientes do experimento às leitoras instaladas também nestes ambientes. Estas distâncias foram determinadas através do leiaute de cada ambiente que já havia sido levantado com a medição dos móveis e também da medição de suas localizações.

Convém destacar que, durante a coleta de dados, os usuários não ficaram necessariamente estáticos, mas sim reproduzindo os movimentos normais de quem estivesse estacionado em um assento trabalhando em seus afazeres. Isto posto, pode-se admitir que a estimação da distância através do método executado não prejudica o processo de filtragem.

Definidas todas as variáveis necessárias, o processo de filtragem foi estabelecido e seus resultados produziram mais três grupos de dados: o MÉTODO 1E, dados tratados somente com EKF; o Método 2E, dados em que as lacunas foram substituídas por valor 71 e, ainda, a aplicação do EKF no conjunto dos dados; e por fim o Método 3E, formado por dados em que as lacunas foram substituídas pelo valor de registro anterior e, também, o tratamento do conjunto de dados pelo EKF. A tabela 5.2 relaciona todos os conjuntos de dados acumulados após o processo de filtragem com o EKF.

Tabela 5.2 – Conjunto de dados gerados para o treinamento e teste dos classificadores e os tratamentos recebido pelo seu conteúdo

Nome do Conjunto	Tipo de Tratamento Recebido pelo seu Conteúdo
Método_1	Conjunto de dados puros sem tratamento de conteúdo. Este conjunto é base para todos os demais
Método_1E	Conjunto de dados cujo tratamento de conteúdo se resume ao EKF. Este conjunto é o resultado da submissão do Método_1 ao EKF
Método_2	Conjunto de dados onde as lacunas do Método_1 foram substituídas por valores de RSSI = 71
Método_2E	Conjunto de dados resultante da submissão do conjunto Método_2 ao EKF
Método_3	Conjunto de dados onde as lacunas do Método_1 foram substituídas pelo valor do último RSSI registrado por aquela leitora
Método_3E	Conjunto de dados resultante da submissão do conjunto Método_3 ao EKF

Mesmo sob a influência do EKF, os dados de alguns pontos de coleta não puderam observar uma melhora considerável, pois a qualidade dos dados nestas ocasiões não contribuiu para tal. Os casos mais críticos são os pontos 10, 11, 12 e 14 que, principalmente por estarem mais distantes da leitora de número 3, tiveram muita perda de informações durante o período de coleta de dados. A figura 5.4 demonstra com mais clareza tais ocorrências.

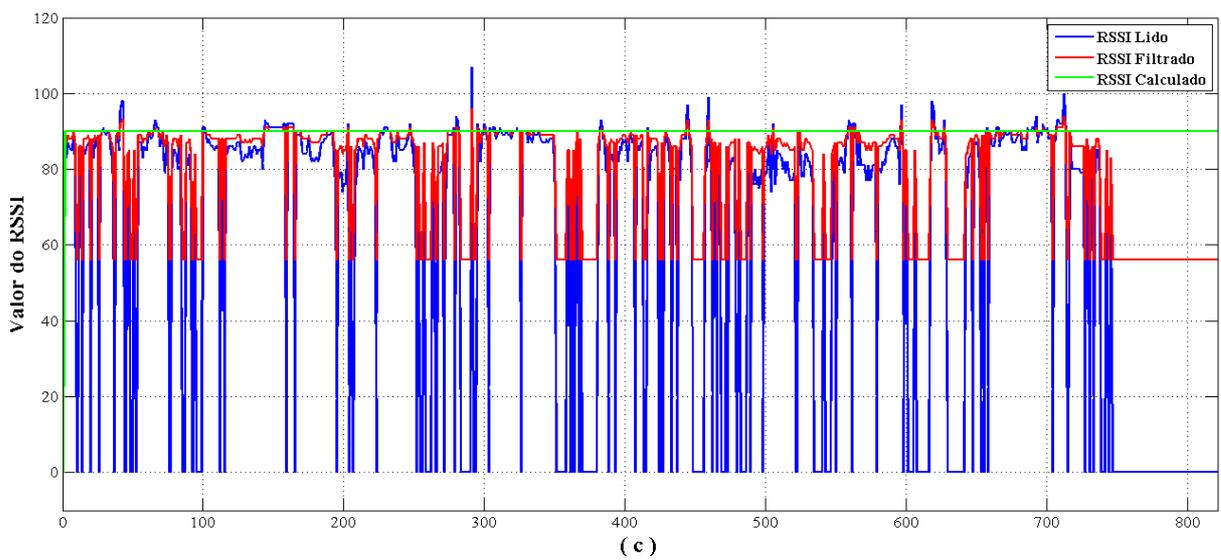
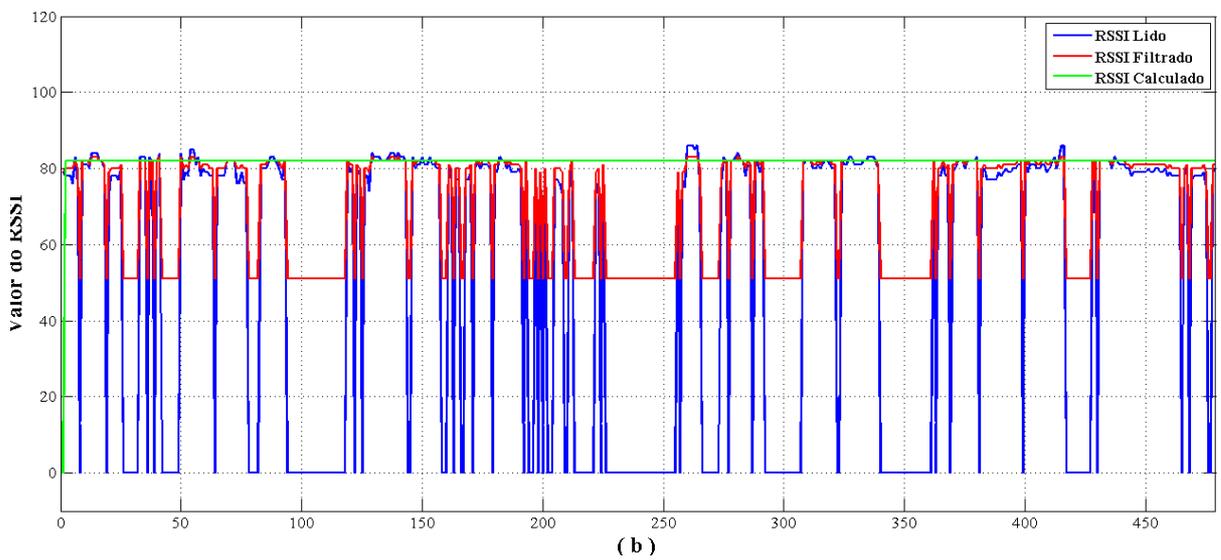
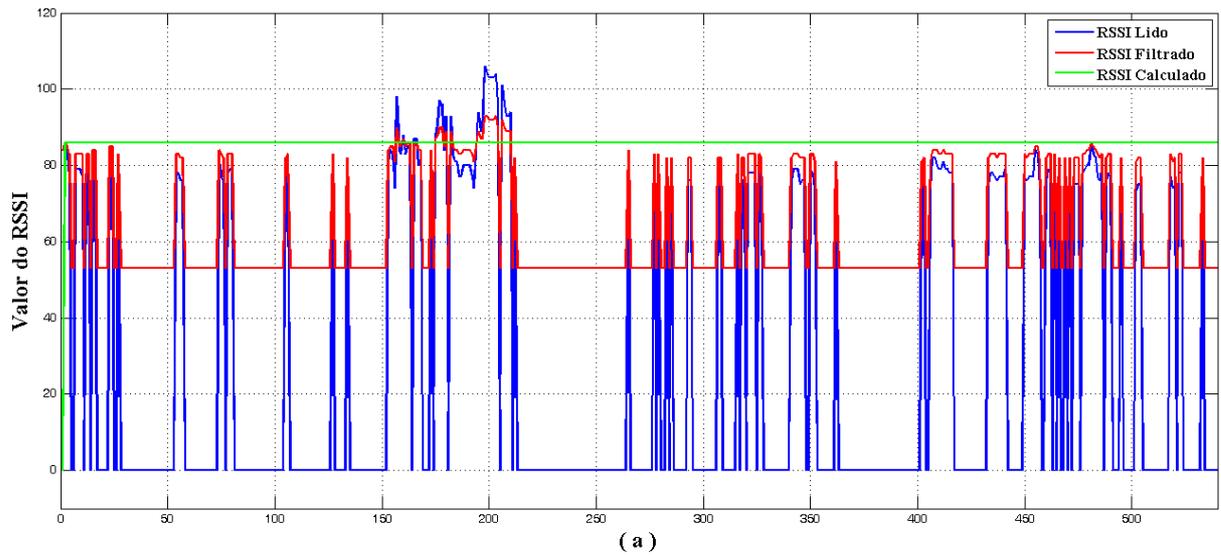


Figura 5.4 – Valores de RSSI coletados pela leitora 3 nos pontos (a) 10, (b) 11 e (c) 14

Das 540 amostras coletadas pelo sistema no ponto 10, a leitora 3 deixou de identificar o pulso amostrado ou, em alguma parte do processo, fora atribuído valor zero para o RSSI em 68,14% das amostras. Ou seja, neste ponto de coleta, mais da metade das amostras não foi registrada de forma completa, com o devido recolhimento dos sinais pelas leitoras, ou pelo *middleware*. Nos pontos 11 e 14 estes índices, foram de 38,41% e 31,43% respectivamente. Em contrapartida, outros pontos de coleta tiveram seus dados beneficiados pela filtragem, como demonstra a figura 5.5.

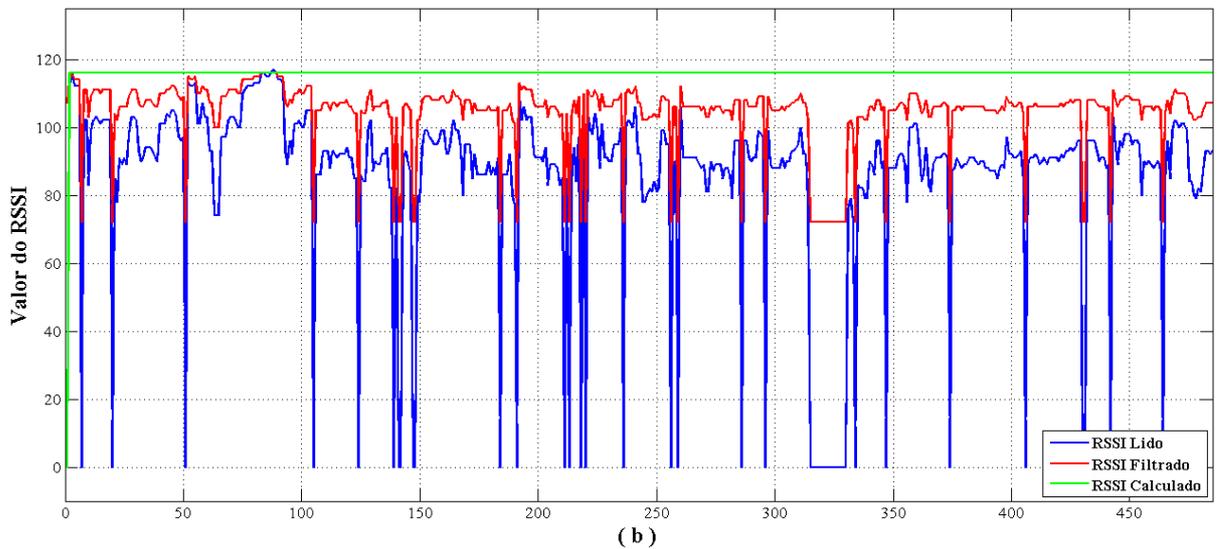
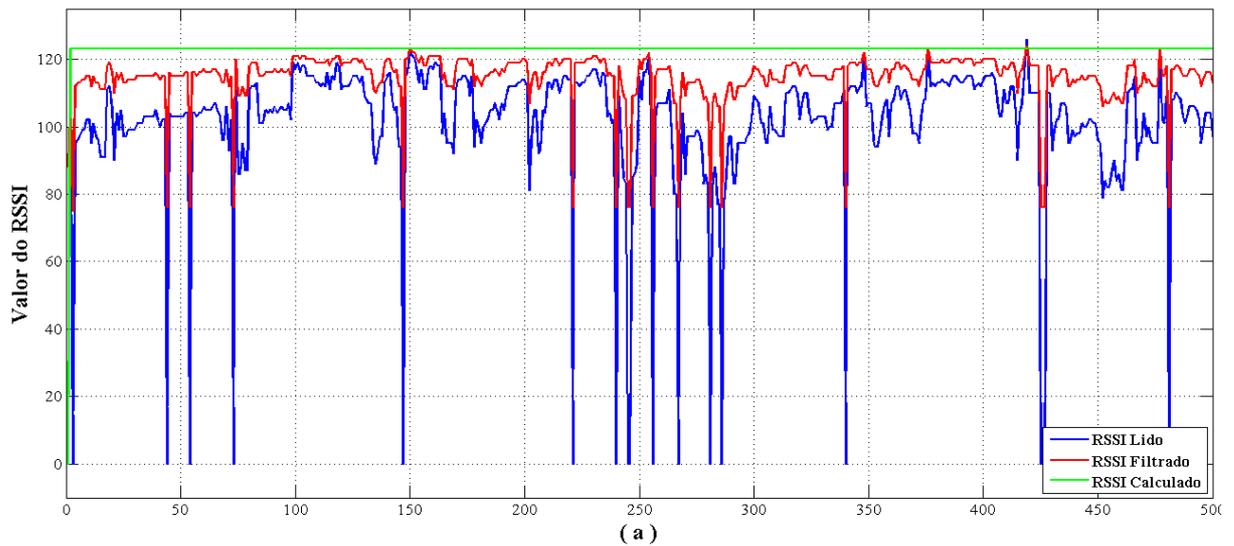


Figura 5.5 – Valores de RSSI coletados pela leitora 0 nos pontos (a) 05 e (b) 06

Os pontos de coleta 05 e 06 são partes deste grupo de beneficiados sem dúvidas. No ponto 05 somente 18 amostras, das 500 coletadas, tiveram valor zero atribuído

ao RSSI pela leitora 0, isto representa 3,6% das amostras. No ponto de coleta 06 este índice passa para 9,28%, sem prejudicar de forma mais ampla a amostragem. Em ambos os pontos a filtragem agiu de forma positiva, pois os valores de RSSI foram corrigidos para valores mais próximos do real.

Não é possível aqui demonstrar todas as benesses proporcionadas pela filtragem com o EKF na base de dado, pois isto demandaria a plotagem dos dados de todos os pontos de coleta, vinte e dois no total, para as quatro leitoras, nas três variantes derivadas do processo de tratamento. Isto totalizaria 264 (duzentas e sessenta e quatro) figuras para plotagem e comentários.

Então, simplificando um pouco este processo de apresentação dos resultados da aplicação do EKF é apresentado na figura 5.6 um bom exemplo de comportamento do RSSI, após o processo de filtragem, nas 06 (seis) bases de dados disponibilizadas para o treinamento dos classificadores. Nesta figura, é possível verificar que o processo de filtragem aproxima o valor lido do valor de referência, ou seja, do valor calculado do RSSI.

A figura 5.6 ressalta três importantes pontos sobre as bases de dados disponibilizadas ao treinamento dos classificadores. Um deles é que, os valores de RSSI dos sinais coletados nas bases sem a filtragem, apresentam maior amplitude de variação entre os máximos e mínimos. Outro é que, o processo de filtragem com o EKF propicia a diminuição considerável desta variação.

E ainda nota-se que as bases oriundas do processo de filtragem apresentam maior proximidade ao valor de referência, que foi calculado através do modelo que rege a dinâmica da propagação da onda eletromagnética do canal *indoor*. Todas estas características refletirão nos resultados do processo de treinamento e teste dos classificadores.

Uma última análise, comparando as figuras 5.6 (b) e (c), mostra ainda que ocorre uma similaridade dos sinais, após o processo de filtragem pelo EKF, que passaram pelo pré-tratamento das metodologias 2 e 3.

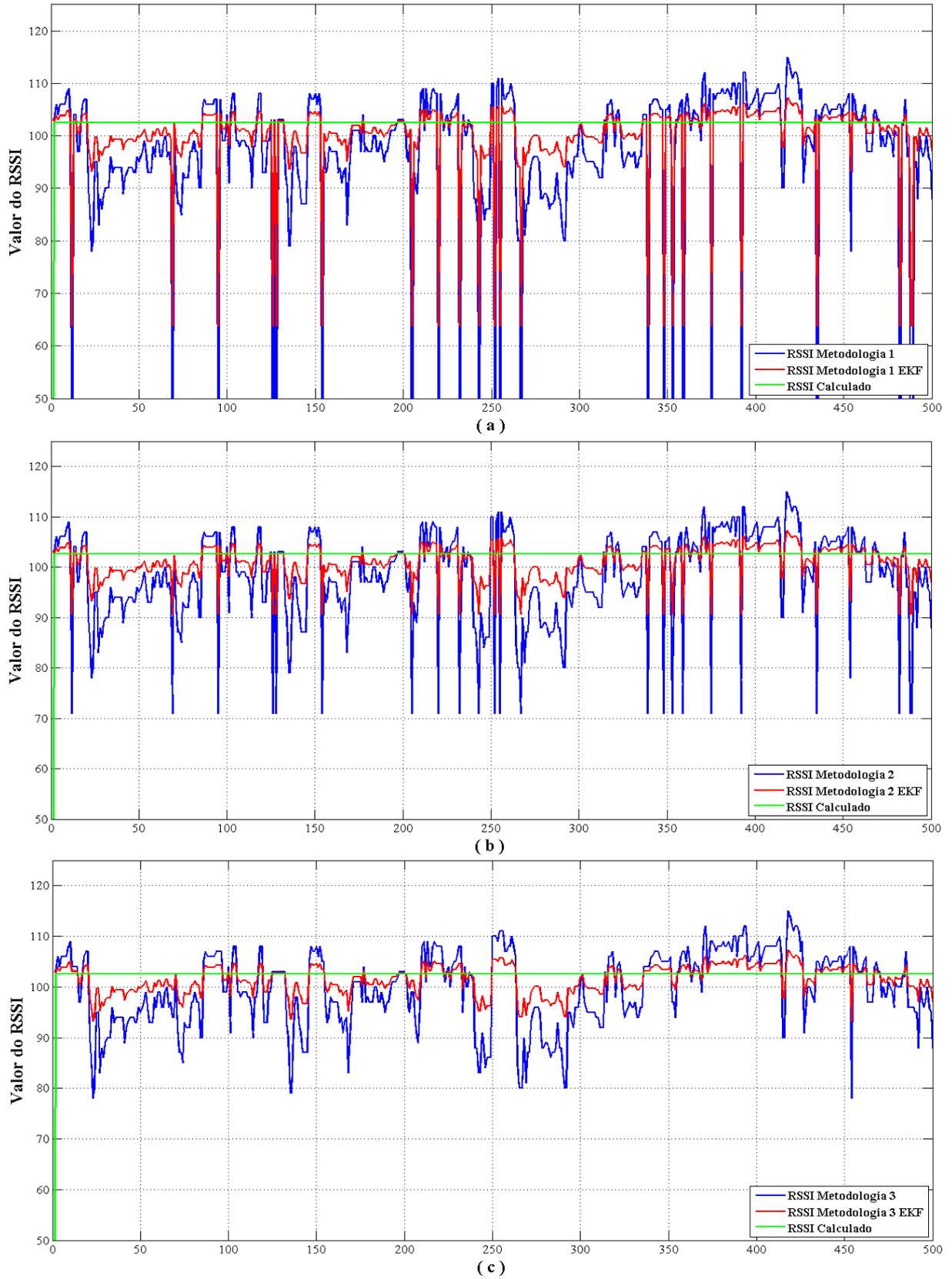


Figura 5.6 – Exemplo de comportamento do RSSI nas seis bases de dados disponibilizadas para o treinamento dos classificadores. (a) Base de dados da metodologia 1 (b) Base de dados da metodologia 2 e (c) Base de dados da metodologia 3

5.3 – TREINAMENTO DAS REDES MLP

Em função da facilidade de manuseio oferecida pela interface gráfica da *Neural Network Toolbox*TM do MatLab[®], e também em função do pequeno tempo demandado para se concluir um processo de treinamento, as redes MLP foram as mais experimentadas com as bases de dados de treinamento e teste.

Várias configurações de redes foram testadas, mas somente 05 (cinco) delas foram experimentadas com todas as bases de dados. As configurações eleitas para os experimentos tentam exprimir o melhor custo benefício, entre desempenho na classificação dos dados e o custo computacional, quer seja este expresso em tempo necessário ao processamento, quer seja na demanda de recursos como memória para armazenamento da rede e dos dados. Em tese, em qualquer computador das gerações atuais, nenhum dos exemplos de custo computacional citado faria sentido, entretanto, se pensarmos em embarcar esta função em equipamentos dotados de microcontroladores, isto já poderá fazer certa diferença. As configurações das redes eleitas constam na tabela 5.3.

Tabela 5.3 – Configurações das redes MLP utilizadas nos experimentos

Nome da Configuração	Configuração da Rede
MLP10	4 entradas, 10 neurônios na camada escondida e 3 saídas
MLP50	4 entradas, 50 neurônios na camada escondida e 3 saídas
MLP100	4 entradas, 100 neurônios na camada escondida e 3 saídas
MLP150	4 entradas, 150 neurônios na camada escondida e 3 saídas
MLP200	4 entradas, 200 neurônios na camada escondida e 3 saídas

Durante os experimentos com as redes MLP foram registrados os melhores desempenhos de cada rede, nos testes com cada uma das bases de dados disponíveis para treinamento e teste dos classificadores. A comparação dos desempenhos registrados pode ser visualizada no gráfico da figura 5.7 que relaciona para cada base de dados separadamente o desempenho das redes eleitas.

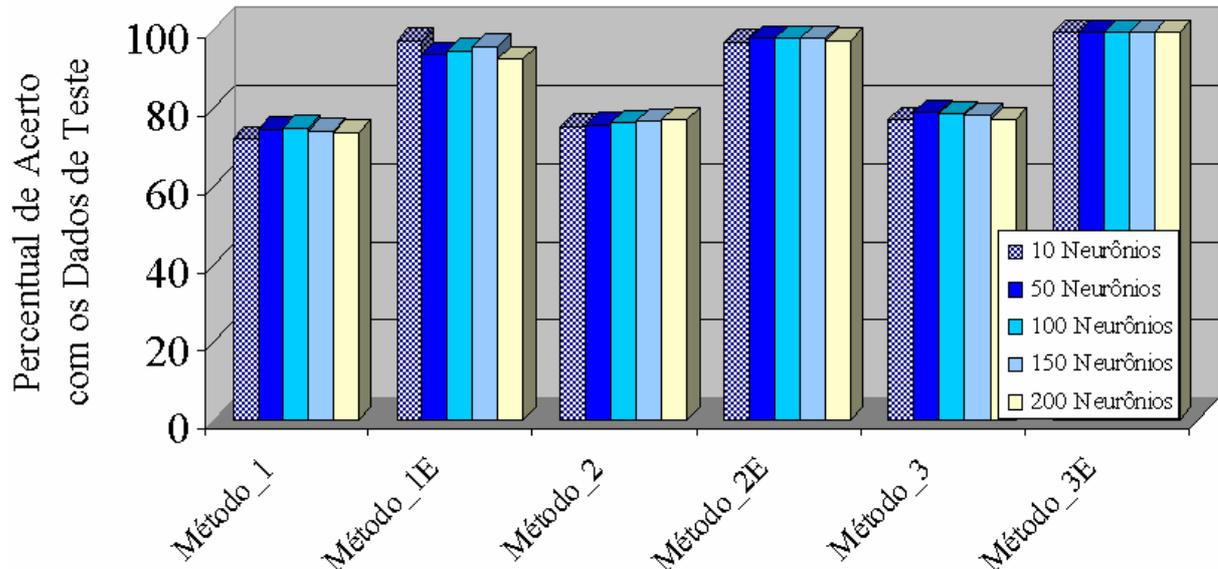


Figura 5.7 – Desempenho do treinamento das redes MLP

A figura 5.7 demonstra claramente que as configurações não apresentam diferença acentuada no desempenho dentro de uma mesma base de dados. A única que apresenta uma pequena discrepância é a base de dados Método_1E, na qual a configuração MLP10 apresentou um desempenho melhor que as demais configurações, o que não ocorreu com as outras bases de dados que mantiveram certo padrão, com os percentuais de desempenho muito próximos uns dos outros.

Outro fator importante de se destacar é que as bases de dados oriundas do processo de filtragem obtiveram percentuais de desempenho acima de 92%, com destaque para a base de dados Método_3E que proporcionou a todas as configurações de rede MLP desempenho superior a 99,1%. Cabe também destacar que as bases não submetidas ao processo de filtragem limitaram o desempenho das redes MLP na casa dos 70%. Na tabela 5.4 pode-se observar estes comportamentos em números.

É com certa curiosidade que se destaca na tabela 5.4 o fato de que para a base de dados Método_1E, cuja única intervenção sofrida foi o processo de filtragem, o processo de treinamento se arrastou por maior quantidade de épocas que nas outras bases que se mantiveram dentro de determinada faixa de ocorrência.

Tabela 5.4 – Desempenho das redes MLP treinadas e quantidade de épocas necessárias para alcançar este desempenho

Grupo de Dados	MLP10		MLP50		MLP100		MLP150		MLP200	
	Desempenho	Épocas								
Método_1	72,2%	87	74,5%	125	74,8%	126	74,0%	108	73,8%	76
Método_1E	97,3%	317	93,6%	220	94,5%	212	95,8%	213	92,8%	149
Método_2	75,3%	85	75,6%	61	76,0%	83	76,6%	76	76,9%	121
Método_2E	97,0%	80	97,7%	71	97,8%	49	97,7%	79	97,4%	44
Método_3	76,9%	63	78,9%	114	78,3%	87	77,9%	64	76,9%	68
Método_3E	99,3%	109	99,3%	59	99,2%	79	99,3%	68	99,1%	47

Convém lembrar que, a base de dado Método_1E, como observado no item 5.2, possui maior amplitude de variação nos valores de RSSI, característica herdada da base de dados que a originou a Método_1, cujos valores de RSSI não sofreram nenhum tipo de intervenção no sentido de promover ajuste nos valores originalmente coletados pelas leitoras do sistema.

5.4 – TREINAMENTO DAS REDES LVQ

Dentro dos classificadores experimentados neste trabalho, as redes LVQ foram as que apresentaram maior demanda de tempo para execução dos processos de treinamento e testes que couberam a cada rede em todas as bases de dados. Cada uma das redes com maior número de neurônios na camada competitiva demorou cerca de três horas para completar o processo de treinamento, cumprindo cada uma delas uma rotina de expor a base de dados à rede dentro do processo de aprendizagem por 100 vezes, ou seja, foram utilizadas 100 épocas de treinamento em cada processo de treinamento. A tabela 5.5 mostra as três configurações escolhidas para os experimentos deste trabalho.

Tabela 5.5 – Configurações das redes LVQ utilizadas nos experimentos

Nome da Configuração	Configuração da Rede
LVQ10	4 entradas, 10 neurônios na camada competitiva e 3 na linear
LVQ100	4 entradas, 100 neurônios na camada competitiva e 3 na linear
LVQ200	4 entradas, 200 neurônios na camada competitiva e 3 na linear

Mesmo com a limitação da quantidade de redes LVQ para experimentação, pelo fator tempo de treinamento, os resultados conseguidos com os experimentos se mostram satisfatórios, apresentando numa análise simplista que, a configuração da quantidade de neurônios na camada competitiva não influenciou muito no resultado final do conjunto deste tipo de classificador.

A figura 5.8 mostra que o desempenho alcançado pelas três configurações praticamente não oferece diferença operacional, e mesmo a diferença numérica do desempenho se mostra inexpressiva, ficando sempre próxima de 1%. Isto demonstra que os sinais gerados pelo sistema RFID seguem um padrão que não necessita ser dividido em muitas subclasses para se conseguir fazer a devida classificação pretendida.

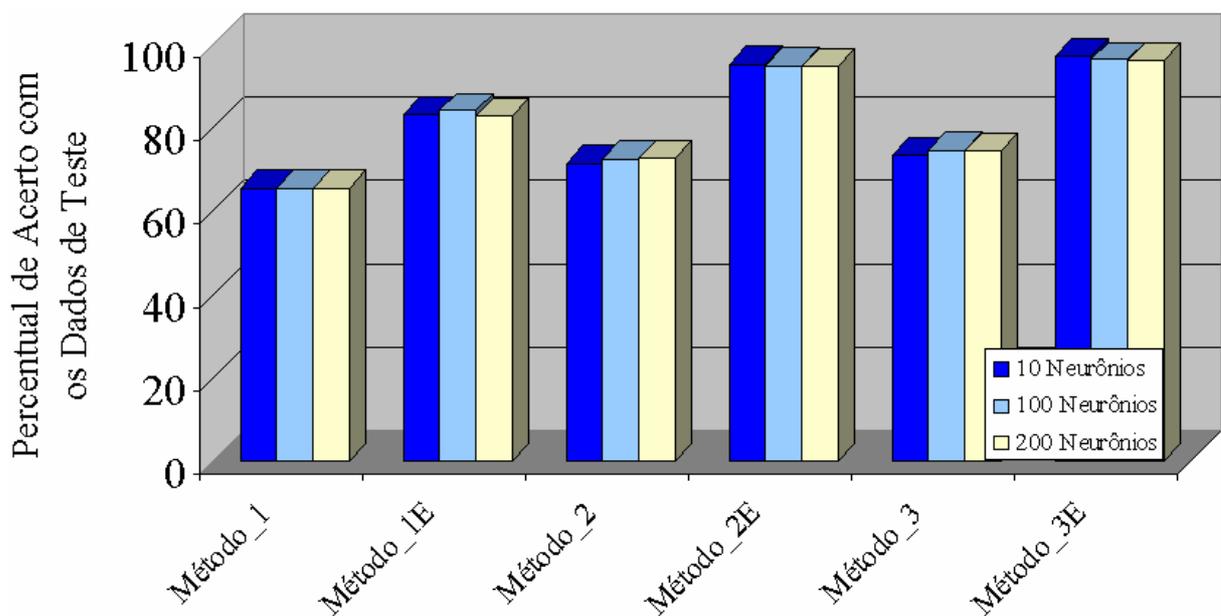


Figura 5.8 – Desempenho do treinamento das redes LVQ

Cada processo de treinamento das redes LVQ foi observado através da interface gráfica do MatLab[®] que permitiu visualizar a evolução destes processos. Dentre os gráficos disponíveis na interface gráfica para observação do processo de treinamento, um informa o erro quadrático médio a cada época de treinamento percorrida, demonstrando assim a evolução da aprendizagem da rede durante o processo de treinamento.

Nos 18 processos de treinamento das redes LVQ se observou que praticamente na primeira época de treinamento a rede já atingia um grau de aprendizagem que não mais se alterava, ou seja, mesmo sendo submetida às novas épocas de aprendizagem a rede não convergia pra valores melhores de aprendizagem.

Este padrão ocorrência pode ser observado na figura 5.9 (a), que mostra o gráfico do comportamento do erro quadrático médio de uma rede que foi treinada por 100 épocas. Em função do exposto, os processos de treinamento, que inicialmente estavam sendo realizados com 100 épocas, passaram a ser realizados com 50 épocas, sem prejuízo ao resultado final dos treinamentos.

Os outros dois gráficos dispostos nas figuras 5.9 (b) e (c) mostram os dois únicos casos em que ocorreu um segundo degrau de diminuição do erro quadrático médio, durante o treinamento das redes LVQ. Apesar desta ocorrência, o desempenho de ambas as redes não apresentaram superioridade, em relação às demais redes LVQ treinadas, nos procedimentos de testes a que foram submetidas.

Há de se observar que, o gráfico disponibilizado pela *Neural Network Toolbox* do MatLab[®] possui escala logarítmica no eixo y, logo, o degrau que aparece nos dois últimos gráficos não possui tamanha influência no resultado porque não significa uma diminuição do erro quadrático médio linear na mesma proporção.

Todas as redes LVQ foram treinadas por mais de uma vez, e ainda assim não apresentaram resultados acentuadamente diferentes que pudessem justificar o retrabalho de uma extensa programação de treinamento destas redes, para se obter o melhor resultado destes treinamentos.

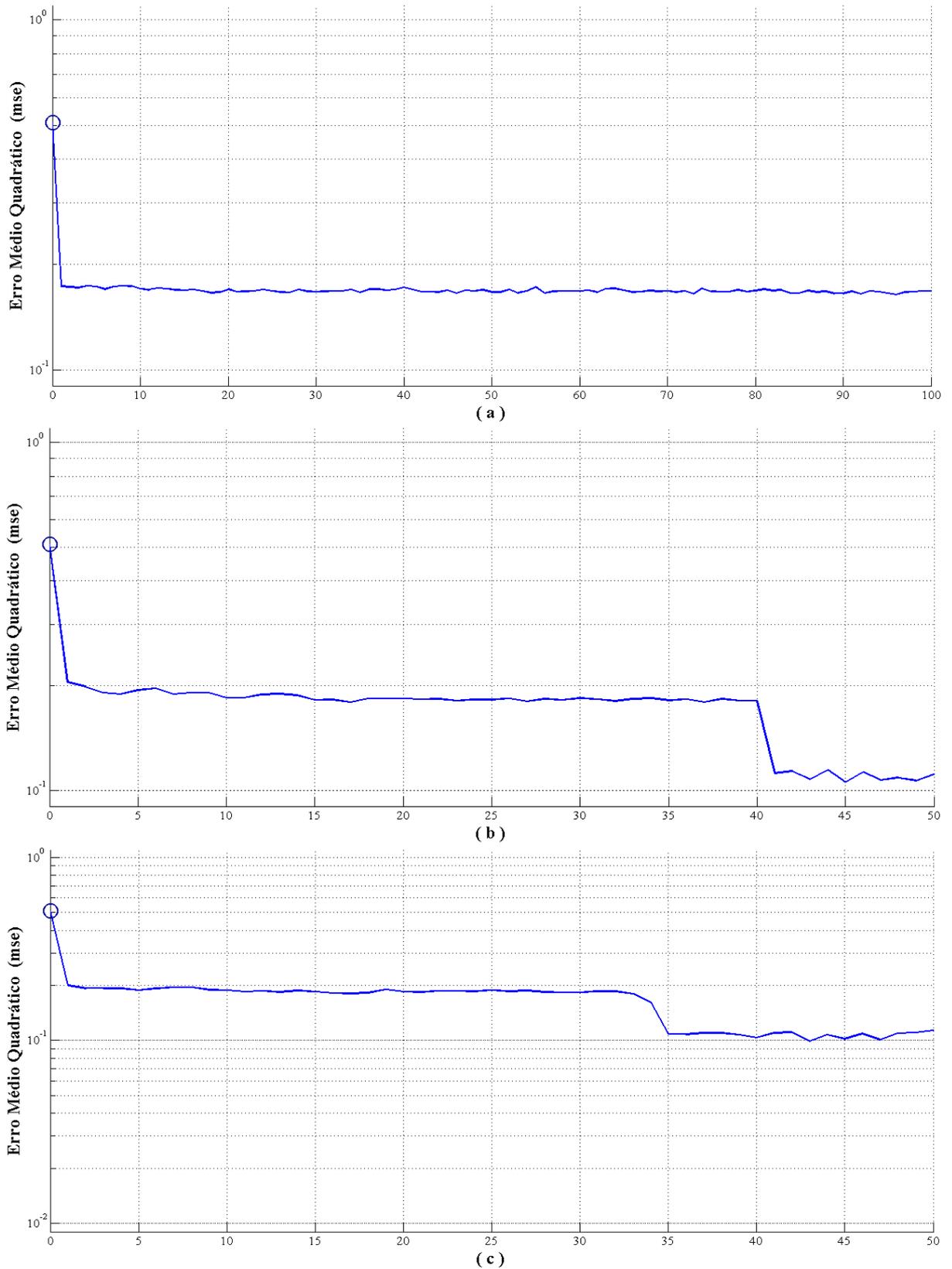


Figura 5.9 – Comportamento do erro quadrático médio durante o treinamento (a) Rede LVQ com 100 neurônios na camada competitiva, base de dados Método_3E, (b) Rede LVQ com 10 neurônios na camada competitiva, base de dados Método_1E e (c) Rede LVQ com 200 neurônios na camada competitiva, base de dados Método_1E.

Finalmente, a tabela 5.6 mostra em números os resultados de desempenho alcançados pelas redes LVQ experimentadas neste trabalho. Assim como as redes MLP, as LVQ apresentaram melhor desempenho com as bases de dados que foram submetidas ao processo de filtragem, com destaque para as duas bases de dados que sofreram pré tratamento dos dados, substituindo as lacunas de RSSI por valores mais próximos dos reais, onde as redes LVQ conseguiram desempenho entre 95% e 97%.

Tabela 5.6 – Desempenho das redes LVQ treinadas

Grupo de Dados	Desempenho		
	LVQ10	LVQ100	LVQ200
Método_1	65,5 %	65,4 %	65,4 %
Método_1E	83,3 %	84,1 %	82,9 %
Método_2	71,3 %	72,5 %	72,7 %
Método_2E	95,4 %	95,0 %	94,8 %
Método_3	73,5 %	74,8 %	74,6 %
Método_3E	97,3 %	96,3 %	96,1 %

5.5 – TREINAMENTO DAS SVM

Como já citado no capítulo anterior, o treinamento das SVM foram realizados através de aplicativos de uma biblioteca denominada LIBSVM. Esta biblioteca possui interfaces com vários aplicativos em diferentes sistemas operacionais, incluindo o MatLab[®], bastando para tal que sejam convertidos alguns arquivos disponibilizados na linguagem C para arquivos de funções do MatLab[®].

Em função de incompatibilidade na versão de compilador da linguagem C, disponível na máquina utilizada para o treinamento e teste das SVM, não foi possível implementar esta interface no MatLab[®] em tempo de se finalizar os experimentos. Desta forma, para as funções de treinamento e teste das SVM, foram utilizados os aplicativos disponíveis para uso direto no sistema operacional MS-Windows[™].

Para o treinamento das SVM foram preparados um arquivo de treinamento e um arquivo de teste, segundo o padrão definido pela biblioteca LIBSVM, para cada um dos 06 (seis) grupos de dados de treinamento previamente preparados após a aplicação do filtro de Kalman estendido. O treinamento, propriamente dito, de uma SVM é feito através do aplicativo 'svm-train.exe' que irá através de seu algoritmo separar dentre os dados apresentados os vetores de suporte para proceder à classificação de dados futuros que forem apresentados na entrada desta SVM. A figura 5.10 mostra a tela de um procedimento de treinamento e teste de uma SVM através destes aplicativos.

```

C:\WINDOWS\system32\command.com

C:\LIBSUM-3.11\WINDOWS>svm-train -s 1 -t 2 LCEKF.train LCEKFrbf.model
.....*
optimization finished, #iter = 6555
C = 0.272474
obj = 394.102209, rho = 0.116330
nSU = 4583, nBSU = 2926
.....*
optimization finished, #iter = 5861
C = 0.309301
obj = 335.633851, rho = -0.288196
nSU = 3609, nBSU = 2101
.....*
optimization finished, #iter = 5318
C = 0.237723
obj = 277.648083, rho = -0.412077
nSU = 3997, nBSU = 2554
Total nSU = 6556

C:\LIBSUM-3.11\WINDOWS>svm-predict LCEKF.train LCEKFrbf.model LCEKFtrain.out
Accuracy = 97.7653% (9581/9800) (classification)

C:\LIBSUM-3.11\WINDOWS>svm-predict LCEKF.test LCEKFrbf.model LCEKFtest.out
Accuracy = 93.4524% (3925/4200) (classification)

C:\LIBSUM-3.11\WINDOWS>_

```

Figura 5.10 – Tela de um procedimento de treinamento e teste das SVM através da LIBSVM com os aplicativos para MS-WindowsTM

No procedimento adotado para o treinamento e teste das SVM, o primeiro passo é o de treinamento, como explicado no parágrafo anterior, que é seguido pela aplicação da SVM construída na classificação dos próprios dados de treinamento para se determinar seu desempenho nos próprios dados de treinamento. Como resultado deste treinamento, a expectativa idealizada é de que ocorra uma classificação correta para 100% dos dados, principalmente nos casos de dados linearmente separáveis. Mas, como se pode ver nos resultados conseguidos, não é o caso dos dados coletados neste experimento.

O último passo dentro do processo de treinamento e teste das SVM é a aplicação da SVM construída na classificação de dados que não participaram do processo de

treinamento, ou seja, dados separados exclusivamente para teste de desempenho do processo de treinamento. No caso da figura 5.10, temos os seguintes resultados do procedimento de treinamento e teste da SVM em questão, $n_{SV} = 6556$, que indica a separação de 6.556 vetores de suporte de um universo de 9.800 vetores de treinamento, que resultaram na correta classificação de 97,77% dos dados de treinamento e 93,45% dos dados exclusivos de teste. O caso apresentado na figura 5.1 utilizou o grupo de dados que recebeu, após a leitura, somente o tratamento da filtragem com o EKF.

Na execução do treinamento de uma SVM, o aplicativo permite que seja escolhido o tipo de função de núcleo que será aplicado na construção da nova SVM. Para o presente trabalho foram escolhidos dois tipos de núcleos com base nas indicações feitas pelos autores da LIBSVM e nas experiências prévias de teste da ferramenta LIBSVM foram definidas as funções de núcleo do tipo função de base radial – RBF e do tipo função linear.

Para cada um das duas funções de núcleo determinadas aos experimentos destes trabalhos foram efetuados os procedimento de treinamento e teste das SVM, totalizando desta forma 12 experimentos de treinamento e teste de SVM. A figura 5.11 mostra a comparação de desempenho das 12 SVM resultantes destes procedimentos. É importante destacar que os resultados apresentados representam o desempenho de classificação dos dados de teste.

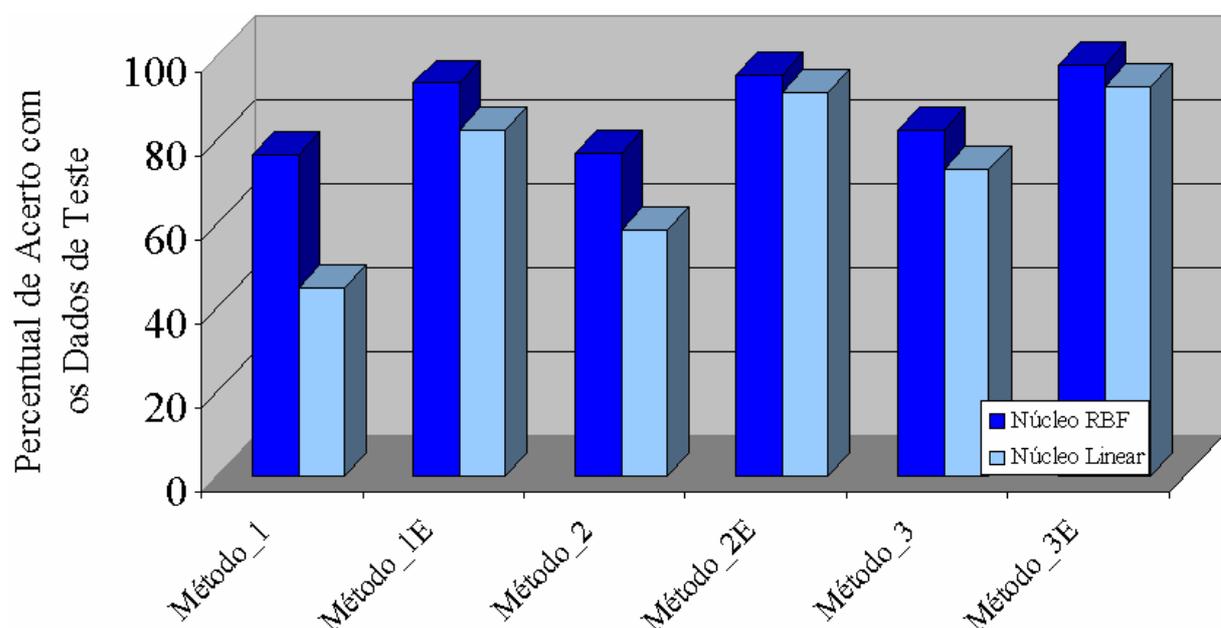


Figura 5.11 – Desempenho do treinamento das SVM

De forma rápida e sem aprofundamentos, a figura 5.11 mostra que as SVM com função de núcleo do tipo RBF são mais eficazes na classificação dos dados do problema em estudo. Mostra ainda que os dados que receberam tratamento através do Filtro de Kalman Estendido permitem maior eficiência no treinamento das SVM, logo permitem melhor desempenho destas SVM na classificação dos dados de teste.

A tabela 5.7 relaciona a quantidade de Vetores de Suporte separados nos procedimentos de treinamento e teste das SVM. Pode-se perceber nesta tabela que, para as SVM cuja função de núcleo seja baseada em RBF, a filtragem dos dados através do EKF contribuiu na diminuição da quantidade de Vetores de Suporte separados pelo treinamento, sem prejudicar seu desempenho e, muito pelo contrário, estas SVM tiveram melhor desempenho em relação às demais.

Tabela 5.7 – Desempenho das SVM treinadas e quantidade de vetores de suporte separados

Grupo de Dados	Função de Núcleo RBF		Função de Núcleo Linear	
	Desempenho	nSV	Desempenho	nSV
Método_1	76,29 %	8.698	44,62 %	6.592
Método_1E	93,45 %	6.556	82,19 %	6.642
Método_2	76,79 %	8.664	58,62 %	6.537
Método_2E	95,02 %	6.539	91,29%	6.684
Método_3	82,12 %	8.527	73,05 %	6.334
Método_3E	97,62 %	6.168	92,52 %	6.659

5.6 – COMPARAÇÃO DO DESEMPENHO DE TODOS OS CLASSIFICADORES

Não seria possível encerrar este capítulo sem fazer a comparação do desempenho de todos os classificadores experimentados neste trabalho. Para tal, foi montada a tabela 5.8 que aglutinou a informação de desempenho das 03 (três) categorias de classificadores utilizadas.

Somadas as diferentes configurações, totalizaram um número de 10 (dez) classificadores que foram experimentados em 06 (seis) versões de bases de dados. Uma destas bases de dados era composta somente pelos valores de RSSI diretamente coletados do Sistema RFID através do *middleware*, outras duas com pré-tratamento, onde foram

substituídas as lacunas, ou seja, valores de RSSI=0, por valores mais próximos do real. As outras três versões originadas do processo de filtragem das três primeiras, pelo algoritmo do filtro de Kalman estendido.

Tabela 5.8 – Comparação do desempenho de todos os classificadores testados

Grupo de Dados	Configuração	Conjunto de Dados para Treinamento e Teste					
		Método_1	Método_1E	Método_2	Método_2E	Método_3	Método_3E
MLP	MLP10	72,2%	97,3%	73,3%	97,0%	76,9%	99,3%
	MLP50	74,5%	93,6%	76,5%	97,7%	78,9%	99,3%
	MLP100	74,8%	94,5%	76,0%	97,8%	78,3%	99,2%
	MLP150	74,0%	95,8%	76,6%	97,7%	77,9%	99,3%
	MLP200	73,8%	92,5%	76,9%	97,4%	76,9%	99,1%
LVQ	LVQ10	65,5%	83,3%	71,3%	95,4%	73,5%	97,3%
	LVQ100	65,4%	84,1%	72,5%	95,0%	74,8%	96,3%
	LVQ200	65,4%	82,9%	72,7%	94,8%	74,6%	96,1%
SVM	RBF	76,3%	93,5%	76,8%	95,0%	82,1%	97,6%
	Linear	44,6%	82,2%	58,2%	91,3%	73,1%	92,5%

Dentre os resultados mostrados na tabela 5.8, dois destaque são necessários às SVM. O primeiro é que foi o único classificador, na sua configuração RBF, a conseguir um desempenho superior a 80% nas bases de dados que não foram filtradas. O segundo destaque, também relativo às bases de dados não filtradas, é que foi o classificador com maior desempenho nestas três bases, demonstrando assim sua robustez como classificador de dados.

Nas demais bases de dados, a SVM RBF acompanhou de perto o desempenho das redes MLP, repetindo um bom desempenho, mas insuficiente para acompanhá-las. Por outro lado, a configuração SVM linear obteve o pior resultado em todas as bases de dados experimentadas.

As redes MLP foram um destaque nas bases de dados filtradas, principalmente nas versões de configuração MLP10, pois conseguiram melhor desempenho em duas destas bases, a Método_1E e a Método_3E. Embora a rede MLP10 tenha sido o menor desempenho, dentre as redes MLP na base de dados Método_2E, conseguiu manter-se dentro do mesmo percentual de acerto que as demais redes MLP nesta base de dados, sendo a rede MLP100 o melhor classificador para esta base de dados. Na base de dados Método_3E ocorreu ainda um empate numérico entre as redes MLP10, MLP50 e MLP150, sendo que as demais redes ficaram com desempenho inferior em exatos 0,1% e 0,2%.

As redes LVQ tiveram desempenho mediano, principalmente dentre as bases filtradas, onde acompanhou de perto as redes MLP, principalmente nas bases Método_2E e Método_3E. É fato que, pelo tempo demandado para se efetuar um treinamento das redes LVQ, estas não foram tão exaustivamente treinadas como as MLP.

Entretanto, mesmo as redes LVQ menores, as de 10 neurônios na camada competitiva, que apresentaram menor tempo de treinamento e, portanto, tiveram mais oportunidade de serem submetidas a uma maior quantidade de treinamento, apresentaram desempenho diferenciado das demais.

Finalizando este capítulo, o gráfico da figura 5.12 apresenta o melhor desempenho dentre os classificadores experimentados neste trabalho, para cada base de dados de treinamento e teste.

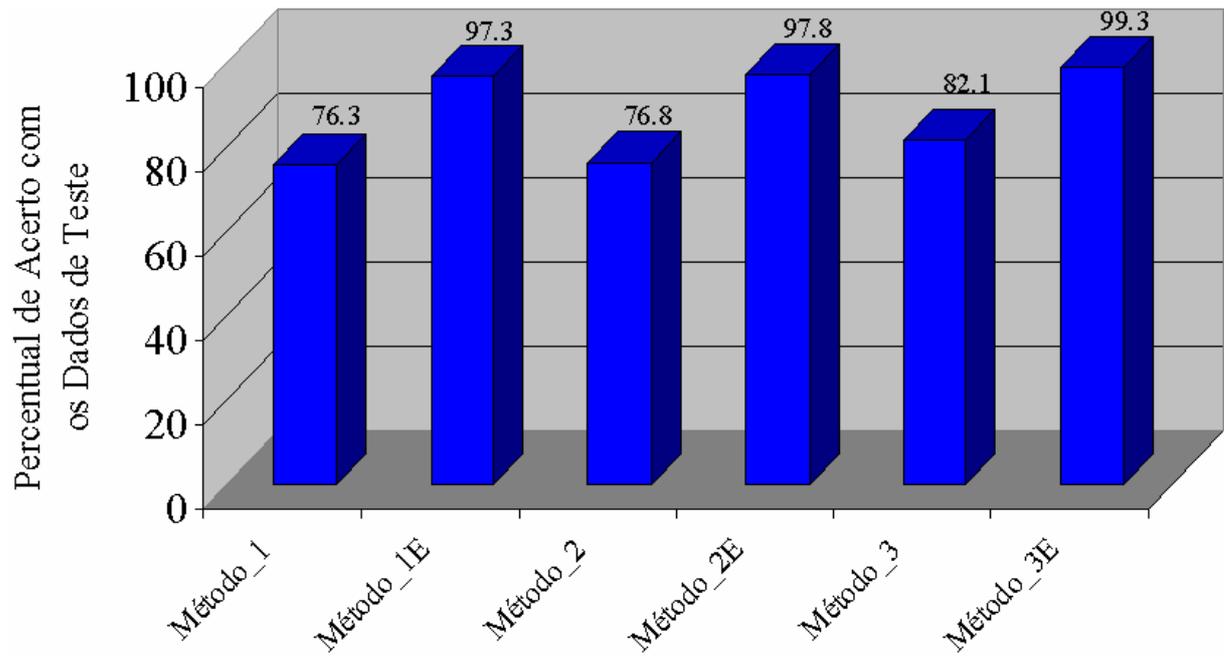


Figura 5.12 – Melhor desempenho dentre os classificadores para cada base de dados de treinamento e teste

6 – CONCLUSÃO

Tendo como objetivo traçado para este trabalho, o de verificar o desempenho de algumas ferramentas computacionais de filtragem e classificação, o sistema de RFID disponibilizado no LARA foi submetido a experimentos que culminaram na coleta de 14.002 pulsos emitidos por uma das *tags* ativas que compõem tal sistema.

A coleta foi procedida dentro de um perímetro delimitado do laboratório, onde foram definidos três ambientes distintos, para efeito de localização. O primeiro ambiente foi composto pela área das bancadas de controle de processos e das bancadas do LabInov. O segundo ambiente foi delimitado pelo perímetro da sala de reuniões. O terceiro ambiente compôs-se pela área das bancadas de inspeção de LT, da mesa do professor Adolfo Bauchspiess, da mesa da impressora, dos escaninhos junto a porta de acesso e das bancadas paralelas ao corredor de acesso à área de robótica móvel.

Antes da coleta foi idealizada para instalação das leitoras, uma disposição de forma que estas ficassem dispostas em eixos ortogonais, locando assim os ambientes nos quadrantes formados por estes eixos. Todo este esforço intencionou a coleta de sinais mais simétricos dentro dos ambientes já citados, e por consequência, maior facilidade na classificação dos ambientes de cada pulso.

Finalizado este processo de definições de infraestrutura, foram coletadas amostras dos pulsos RFID emitidos pela *tag* ativa, com a ajuda do *middleware* desenvolvido por Fonseca [1]. Estas amostras, denominadas ao longo do presente trabalho como conjunto de dados, ou base de dados, foram trabalhadas numa primeira fase de forma manual, para primeiro identificar as suas características mais marcantes e mesmo suas limitações e impurezas, que sabidamente faziam parte do material coletado.

Dentro das limitações encontradas, uma necessitou especial atenção, pois se caracterizou por imprimir valor nulo ao RSSI em determinadas situações, por alguns fatores

levantados no item 4.2.3. Em função desta descoberta, nasceu uma proposta de aplicar três metodologias na formação das bases de dados de treinamento.

A primeira metodologia consistiu em aplicar no treinamento uma base de dados tal como coletados os pulsos, sem qualquer tratamento prévio. A segunda metodologia substituiu os valores nulos de RSSI pelo valor numérico de 71 (setenta e um). Este valor foi escolhido em função de se observar que não ocorriam registros inferiores a este valor. A terceira e última metodologia atribuiu aos valores nulos o valor do RSSI registrado no pulso anterior.

Na etapa seguinte, as bases de dados acumuladas, três até então, foram submetidas a mais um tratamento, desta vez já com uma das ferramentas computacionais, objeto de estudo do presente trabalho, o filtro de Kalman estendido, implementado no MatLab[®], que permitiu trabalhar novamente sua formatação, duplicando então os dados acumulados até então. Com este passo as bases de dados passaram ao número de 06 (seis).

As bases de dados devidamente definidas e trabalhadas foram então submetidas às outras ferramentas de computacionais previstas no objetivo do presente trabalho, os classificadores. Estas bases de dados foram aplicadas no treinamento e teste de classificadores como as redes MLP e LVQ, e ainda as SVM, que intentaram ser aplicadas no sistema de localização proposto no projeto de pesquisa a que este trabalho esteve vinculado.

A metodologia aplicada na coleta dos dados, utilizados nos testes das ferramentas computacionais, buscou a melhor analogia com a futura operação de um sistema de localização, baseado em RFID, a ser implantado no LARA. Por isto não seria exagero afirmar que os resultados apurados em todos estes processos descritos indicam que a variável RSSI coletada dos Sistemas de RFID realmente necessita de intervenções para sua aplicação em sistemas de localização, mesmo que estes sejam mais simplórios, indicando somente o ambiente em que o usuário esteja estacionado.

Os resultados dos experimentos dos classificadores indicaram que os dados coletados puramente, sem nenhum dos artifícios de pré-tratamento dos valores nulos de RSSI, permitiriam somente um desempenho da ordem de 76,3% de acerto. Ou seja, seria bem possível que, eventualmente um sistema de climatização, controlado pelo sistema de

localização, pudesse promover o desligamento de sua máquina mesmo havendo usuário no recinto controlado por estes sistemas, ou em via contrária, promover a operação de máquinas em ambiente vizinho efetivamente vazio.

É possível que, para localização somente de ambiente, não se necessite de classificadores com desempenho na ordem dos 99%, como foi apurado em uma das bases de dados filtradas pelo EKF. Por outro lado, o melhor desempenho para as bases não filtradas ficou em 82,1%. Mesmo neste caso é possível que este índice seja aceitável, mas ainda assim é prematuro afirmar sem um estudo mais contundente, com situações reais, qual é o índice necessário para se prover uma boa localização que sirva ao propósito de controle e automação dos ambientes do LARA.

A escolha do classificador, para aplicação neste futuro sistema de localização também transitará por outras questões ainda indefinidas como o custo computacional de cada tipo classificador, ou seja, sua demanda computacional para o processamento dos dados gerados pelo sistema RFID. Ainda há que ser considerada, na escolha do classificador, a forma com que as informações providas por esta ferramenta impactarão na ferramenta de controle dos sistemas de refrigeração e/ou iluminação, como por exemplo, de forma imediata – identificou acionou – ou se intermediada por um temporizador.

Com todas as considerações expostas, o presente trabalho contribui com importantes informações antes desconhecidas, como a comparação de desempenho dos classificadores MLP, LVQ e SVM, que permitiu inclusive se observar o destaque das SVM como um classificador bastante robusto. As SVM, antes sequer cogitadas para uso nas aplicações pretendidas, agora passam a figurar como importantes ferramentas, inclusive de fácil operação, a serem pelo menos avaliadas antes de uma decisão de uso.

O presente trabalho também ajuda a sedimentar a necessidade de aplicação de um processo de filtragem dos sinais providos pelo RSSI antes que estes possam ser aplicados em um sistema de localização. Há ainda um fruto deste trabalho que não foi finalizado, mas que já foi iniciado, trata-se de artigo científico descrevendo os resultados conseguidos, a ser submetido em evento da área de controle.

6.1 – TRABALHOS FUTUROS

A implementação prática do sistema de localização através do sistema RFID no LARA é sem dúvida o grande desafio para o futuro, mesmo que esta implementação ocorra somente para um teste mais contundente, com situações do dia a dia do laboratório, se ainda lidar com o controle de outros sistemas. Para que ocorra esta implementação prática, algumas ações serão necessárias, dentre elas a implementação do filtro de Kalman de forma dinâmica, ou seja, implementação do filtro dentro do processo da aquisição de dados, pois todos os dados coletados neste para este trabalho ocorreram *off line*, e o tratamento destes dados foi feito por batelada, com o conforto dos dados já terem sido previamente vistoriados e corrigidos de eventuais erros de leitura ou erros sistêmicos do RFID.

O tratamento posterior dos dados coletados do sistema RFID possibilitou inclusive a definição das coordenadas, ou seja, distância entre os pontos de coleta e as leitoras do sistema RFID. A determinação desta grandeza, que tem fundamental importância no modelo de propagação de onda utilizado pelo filtro de Kalman, é bem mais complicada num sistema dinâmico, podendo inclusive necessitar de um estudo específico sobre a modelagem deste sistema, que envolverá movimento cinético das *tags* e a propagação de onda dos pulsos emitidos por estas *tags*. Um bom começo para este assunto pode ser encontrado em Rappaport [14].

As SVM merecem especial atenção também nos trabalhos futuros, pois os resultados encontrados nos experimentos deste trabalho demonstraram que estes classificadores são bastante robustos. Talvez um estudo mais direcionado a esta ferramenta faça aflorar algum diferencial, principalmente no tocante a suas configurações, que possam melhorar seu desempenho, e se isto for conseguido para dados não filtrados, sua aplicação fica ainda mais interessante, pois diminui o custo computacional em aplicações de controle embarcado.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] FONSECA, Lucas Oliveira da, (2010). **Sistema de localização RFID de Usuários Visando a Racionalização de Energia em Ambientes Inteligentes**. Trabalho de Graduação em Engenharia de Controle e Automação, Publicação FT.TG-nº 11, Faculdade de Tecnologia, Universidade de Brasília, Brasília-DF, 2011 - 63p.
- [2] SOUZA, André Luis Gama; OLIVEIRA, Gabriel Figueiró. **Sistema de Localização para Robô Móvel**. Trabalho de Graduação em Engenharia de Controle e Automação, Publicação FT.TG-nº 11, Faculdade de Tecnologia, Universidade de Brasília, Brasília-DF, 2011 - 124p.
- [3] MINISTÉRIO DE MINAS E ENERGIA. Empresa de Pesquisa Energética (EPE). **Balanco Energético Nacional – BEN-2030**. Brasília: MME, 2007. Disponível em http://www.epe.gov.br/PNE/20080111_1.pdf. Acesso em: 08 de agosto de 2011.
- [4] ZHAO, Yiyang; LIU, Yunhao; NI, Lionel M. **VIRE: Active RFID-based Localization Using Virtual Reference Elimination**. Parallel Processing, International Conference on, p. 56, 2007 International Conference on Parallel Processing (ICPP 2007), 2007. DOI: 10.1109/ICPP.2007.84
- [5] PAPAPOSTOULOU, Apostolia; CHAOUCHI, Hakima. **RFID-assisted Indoor Localization and the Impact of Interference on Its Performance**. Journal of Network and Computer Applications. Volume 34, Issue 3, May 2011, Pages 902-913. DOI:10.1016/j.jnca.2010.04.009
- [6] LANDT, Jeremy. **The History of RFID**. IEEE Potentials, vol. 24, no. 4, pp 8-11, Oct/Nov 2005, DOI:10.1109/MP.2005.1549751.
- [7] WEINSTEIN, Ron. **RFID: A Technical Overview and Its Application to the Enterprise**. IT Professional, vol. 7, no. 3, pp. 27-33, May/June 2005, DOI:10.1109/MITP.2005.69.
- [8] FINKENZELLER, Klaus. **RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near Field Communication**. 3ª Ed. tradução: Dörte Müller. Wiltshire-UK: Wiley and sons. 2010.
- [9] NI, Lionel M; LIU, Yunhao; LAU, Yiu Cho e PATIL, Abhishek. **LANDMARC: Indoor Location Sensing Using Active RFID**. Wireless Networks, vol. 10, no. 6, pp. 701-710, Nov 2004, DOI: 10.1023/B:WINE.0000044029.06344.dd
- [10] JIN, Guang-Yao; LU, Xiao-Yi; PARK, Myong-Soon. **An Indoor Localization Mechanism Using Active RFID Tag**. IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, Jun 2006. Volume 1. DOI: 10.1109/SUTC.2006.1636157.
- [11] WILLIS, Scooter; HELAL, Sumi. **A Passive RFID Information Grid for Location and Proximity Sensing for the Blind User**. University of Florida Technical Report Number TR04-009, 2004. Disponível em: <https://www.cise.ufl.edu/tr/DOC/REP-2004-355.pdf>, acesso em 08/03/2012.

- [12] TESORIERO, R.; GALLUD, J. A.; LOZANO, M.; PENICHER, V. M. R. **Using Active and Passive RFID Technology to Support Indoor Location-Aware Systems**. IEEE Transactions on Consumer Electronics, Volume 54, Issue 2, 2008. Pages 578-583. DOI: 10.1109/TCE.2008.4560133.
- [13] WILDPACKETS. **Converting Signal Strength Percentage to dBm Values**. Whitepaper. Walnut Creek, CA-USA: 2002. Disponível em: http://www.wildpackets.com/elements/whitepapers/Converting_Signal_Strength.pdf, acesso em 11/04/2011.
- [14] RAPPAPORT, Theodore S. **Comunicações Sem Fio: Princípios e Prática**. Tradução Daniel Vieira. Revisão Técnica Luiz Carlos Pessoa Albini. 2.ed. Person Prentice Hall – São Paulo – 2009.
- [15] LYMBERPOULOS, Dimitrios; LINDSEY, Quentin; SAVVIDES, Andreas. **An Empirical Characterization of Radio Signal Strength Variability in 3-D IEEE 802.15.4 Networks Using Monopole Antennas**. 3rd. European Conference on Wireless Sensor Networks – EWSN 2006 – LNCS – Volume 3868, pages 326-341. DOI: 10.1007/11669463_24.
- [16] KAPLAN, Elliott D.; HERGARTY, Christopher J. **Understanding GPS: Principles and Applications**. 2nd ed. – Londres: Artech House, 2006.
- [17] WIKIPEDIA. **GNSS**. Disponível em: <http://pt.wikipedia.org/wiki/GNSS>. Acesso em 01/05/2012.
- [18] DEDES, G., & DEMPSTER, A.G. **Indoor GPS positioning: Challenges & opportunities**. IEEE 62nd Vehicular Technology Conference. Dallas, Texas, 25-28 Sep 2005, Pages 412-415. Disponível em: http://www.gmat.unsw.edu.au/snap/publications/dedes&dempster_2005.pdf, acesso em 15/07/2011.
- [19] DIGGELEN, Frank van; ABRAHAM, Charles. **Indoor GPS Technology**. CTIA Wireless Conference, Mar 2001, Las Vegas. Disponível em: http://www.ed-china.com/ARTICLES/2006JUN/5/2006JUN30_IC_RFD_TS_11.PDF, acesso em: 15/07/2011.
- [20] FLUERASU, Anca; PICOIS, Alexandre Vervisch; BOIERO, Gianluca; GHINAMO Giorgio; LOVISOLO, Piero; SAMAMA, Nel. **Indoor Positioning Using GPS Transmitters: Experimental Results**. 2010 IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN). Pages 1-9. DOI: 10.1109/IPIN.2010.5647647
- [21] OSZOY, Kerem; BOZKURT, Ayhan; TEKIN, Ibrahim. **2D Indoor Positioning System using GPS Signals**. 2010 IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN). Pages 1-6. DOI: 10.1109/IPIN.2010.5646234
- [22] RIZOS, Chris; ROBERTS, Gethin; BARNES, Joel; GAMBALE, Nunzio. **Experimental Results of Locata: A High Accuracy Indoor Positioning System**. 2010 IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN). Pages 1-7. DOI: 10.1109/IPIN.2010.5647717
- [23] NEKOOGAR, Faranak. **Ultra-Wideband Communications: Fundamentals and Applications**. Prentice Hall PTR – Upper Saddle River – NJ – 2005.

- [24] TUCHLER, Michael; SCHWARZ, Volker; HUBER, Alexander. **Location accuracy of an UWB localization system in a multi-path environment**. 2005 IEEE International Conference on Ultra-Wideband, Sep. 2005. ICU 2005. Pages 414-419. DOI: 10.1109/ICU.2005.1570023.
- [25] ALAVI, Bardia; ALSINDI, Nayef; PAHLAVAN, Kaveh. **UWB Channel Measurements for Accurate Indoor Localization**. IEEE Military Communications Conference, Oct. 2006. MILCOM 2006. Pages 1-7. DOI: 10.1109/MILCOM.2006.302143.
- [26] GALLER, Stefan; GEROK, Waldemar; SCHROEDER, Jens; KYAMAKYA, Kyandoghere; KAISER, Thomas. **Combined AOA/TOA UWB localization**. International Symposium on Communications and Information Technologies, Oct. 2007. ISCIT '07. Pages 1049-1053. DOI: 10.1109/ISCIT.2007.4392171.
- [27] YU, Kegen; MONTILLET, Jean-philippe; RABBACHIN, Alberto; CHEONG, Paul; OPPERMANN, Ian. **UWB location and tracking for wireless embedded networks**. Signal Processing, Volume 86, Issue 9, Sep. 2006, Pages 2153-2171, DOI: 10.1016/j.sigpro.2005.07.042.
- [28] PETROFF, Alan. **A Practical High Performance Ultra Wideband Radar Platform**. RadarCon 2012 – Atlanta. Disponível em: www.timedomain.com , acesso em 15/05/2012.
- [29] HUSETH, Steve; DEWBERRY, Brandon; McCROSKEY, Robert. **Pulsed-RF Ultrawideband Ranging for the GLANSER GPS-denied Emergency Responder Navigation System**. 2011 International Technical Meeting of the Institute of Navigation, jan 2011, San Diego – CA. Disponível em: www.timedomain.com , acesso em 15/05/2012.
- [30] SAIDI, Kamel S.; TEIZER, Jochen; FRANASZEK, Marek; LYTLE, Alan M. **Static and Dynamic Performance evaluation of a Commercially-Availble Ultra Wideband Tracking System**. Automation in Construction, Vol. 20, Issue 5, Aug 2011, Pag 519-530, DOI: 10.1016/j.autcon.2010.11.018.
- [31] GONZÁLES, J.; BLANCO, J. L.; GALINDO, C.; ORTIZ-DE-GALISTEO, A.; FERNANDEZ-MADRIGAL, J. A.; MORENO, F. A.; MARTÍNEZ, J. L. **Mobile Robot Localization Based on Ultra-Wide-Band Ranging: A Particle filter Approach**. Robotics and Autonomous Systems, Vol. 57, Issue 5, May 2009, Pag 496-507. DOI: 10.1016/j.robot.2008.10.022.
- [32] LOIOLA, Roberto Rodrigues. **Redes Neurais Artificiais Aplicadas ao Problema da Localização em Ambientes Fechados**. Dissertação de Mestrado em Engenharia Elétrica. Departamento de Engenharia Elétrica – Universidade de Brasília. Brasília-DF. 2009. 171p.
- [33] FANG, Shih-Hau; LIN, Tsung-Nan. **Indoor Location System Based on Discriminant-Adaptive Neural Network in IEEE 802.11 Environments**. IEEE Transactions on Neural Networks, Volume 19, Issue 11, Nov 2008. Pages 1973-1978. DOI: 10.1109/TNN.2008.2005494.
- [34] CHINTALAPUDI, Krishina; IYER, Anand Padmanabha; PADMANABHAN, Venkata N. **Indoor Localization Without the Pain**. In Proceedings of the sixteenth annual international conference on Mobile computing and networking (MobiCom '10). Chicago – Illinois – USA. September 20-24 2010. doi: 10.1145/1859995.1860016.

- [35] PAUL, Anindya S.; WAN, Eric A. **RSSI-Based Indoor Localization and Tracking Using Sigma-Point Kalman Smoothers**. IEEE Journal of Selected Topics in Signal Processing, Volume 3, Issue 5, Oct 2009. Pages 860-873. DOI: 10.1109/JSTSP.2009.2032309.
- [36] FIGUEREDO, L.F.C.; COUTO, F.L.; BAUCHSPIESS, A. **An Evaluation of RSSI Based Indoor Localization Systems in Wireless Sensor Networks** In: XI SBAI – Simpósio Brasileiro de Automação Inteligente, 2009, Brasília.
- [37] LARRANAGA, Janire; MUGUIRA, Leire; LOPEZ-GRADE, Juan-Manuel; VAZQUEZ, Juan-Ignacio. **An Environment Adaptive ZigBee-based Indoor Positioning algorithm**. 2010 IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN). Pages 1-6. DOI: 10.1109/IPIN.2010.5647828
- [38] DIAZ, Javier J. M.; MAUÉS, Rodrigo de A.; SOARES, Rodrigo B.; NAKAMURA, Eduardo F.; FIGUEIREDO, Carlos M. S. **Bluepass: an Indoor Bluetooth-based Localization System for Mobile Applications**. 2010 IEEE Symposium on Computers and Communications (ISCC). Pages 778-783. DOI: 10.1109/ISCC.2010.5546506.
- [39] BEKKALI, Abdelmoula; SANSON, Horacio; MATSUMOTO, Mitsuji. **RFID Indoor Positioning Based om Probabilistic RFID Map and Kalman Filtering**. Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMOB 2007. Oct 2007. DOI: 10.1109/WIMOB.2007.4390815.
- [40] VAIDEHI, V.; VASUHI, S.; GANESH, K. S.; THEANAMMAI, C.; NARESH BABU, N. T.; UTHIRAVEL, N.; BALAMURALIDHAR, P.; Chandra, G. **Person Tracking Using Kalman Filter in Wireless Sensor Network**. 2010 Second International Conference on Advanced Computing (ICoAC), 2010. Pages 60-65. DOI: 10.1109/ICOAC.2010.5725362.
- [41] TEIXEIRA, Bruno O. S., TÔRRES, Leonardo A. B., AGUIRRE, Luis A. **Filtragem de Kalman com Restrições para Sistemas Não-Lineares: Revisão e Novos Resultados**. Revista Controle & Automação, Volume 21, Número 10, Mar e Abr-2010. Pag. 127-146.
- [42] SANTANA, Pedro Henrique de Rodrigues Quemel e Assis. **Filtragem Estocástica para Sistemas Híbridos e suas Aplicações em Robótica Aérea**. Dissertação de Mestrado em Engenharia Elétrica. Departamento de Engenharia Elétrica – Universidade de Brasília. Brasília-DF. 2011. 122p.
- [43] HATA, Alberto Yukinobu; WOLF, Denis Fernando. **Mapeamento de Ambientes Externos Utilizando Robôs Móveis**. Dissertação de Mestrado em Ciência da Computação e Matemática Computacional. Instituto de Ciências Matemática e de Computação – Universidade de São Paulo. São Carlos-SP. 2010. 90p. Disponível em : <http://www.teses.usp.br/?&lang=pt-br>. Acesso em: 12/12/2011
- [44] ZHEN, Zi-Ning; JIA, Qing-Shan; SONG, Chen; GUAN, Xiaohong. **An Indoor Localization Algorithm for Lighting Control Using RFID**. IEEE Energy 2030 Conference, 2008. Pages 1-6. DOI: 10.1109/ENERGY.2008.4781041.
- [45] HSU, Chih-Wei; CHANG, Chih-Chung; LIN, Chih-Jen. **A Practical Guide to Support Vector Classification**. Technical Report. Department of Computer Science – National Taiwan University, 2003. Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>. Acesso em: 07/02/2012

- [46] YAMANO, K.; TANAKA, K.; JIRAYAMA, M.; KONDO, E.; KIMUTO, Y.; MATSUMOTO, M. **Self-Localization of mobile Robots with RFID System by Using Support Vector Machine**. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004 – IROS 2004. Pages 3756-3761. DOI: 10.1109/IROS.2004.1389999
- [47] AGUIRRE, Luis Antonio. **Introdução à Identificação de Sistemas: Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais**. 3.ed. Revisada e Ampliada – Belo Horizonte: Editora UFMG, 2007.
- [48] WELCH, Greg; BISHOP, Gary. **An Introduction to the Kalman Filter**. Technical Report TR95-041 – University of North Carolina at Chapel Hill, Department of Computer Science, 1995. Disponível em: <http://www.cs.unc.edu/~welch/kalman/kalmanIntro.html>, acesso em 04/06/2012.
- [49] HAYKIN, Simon. **Redes Neurais: Princípios e Práticas**. Tradução Paulo Martins Engel – 2.ed. – Porto Alegre: Bookman, 2001.
- [50] BEALE, Mark Hudson; HAGAN, Martin T.; DEMUTH, Howard B. **Neural Network Toolbox™ 7 – User’s Guide**. The Mathworks, Inc. 2010.
- [51] LORENA, Ana Carolina; CARVALHO, André C. P. L. de. **Investigação e Estratégias para a Geração de Máquinas de Vetor de Suporte Multiclasses**. XI Simpósio de Teses e Dissertações do Programa de Pós Graduação em Ciência da Computação e Matemática Computacional – STD 2008. Instituto de Ciências Matemática e Computação – USP São Carlos. Disponível em: <http://pet.icmc.usp.br/std2008/artigos/65.Doutorado.pdf>. Acesso em 12/12/2011.
- [52] OZTAYSI, Basar; BAYSAN, Serdar; AKPINAR, Fatma. **Radio Frequency Identification (RFID) in Hospitality**. Technovation, Volume 29, Issue 9, Sep 2009, Pages 618-624. DOI: 10.1016/j.technovation.2009.05.014.
- [53] BAUDIN, Michel; RAO, Arun. **RFID Applications in Manufacturing**. SME Automation & Assembly Summit. Saint Louis, 04/2005. Disponível em: <http://www.mmti-inst.com/contents.htm#Articles>, acesso em 08/11/2011.
- [54] MATHWORKS. MatLab® 7 – Realese 2010b
- [55] CHANG, Chih-Chung; LIN, Chih-Jen. **LIBSVM : a library for support vector machines**. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, acesso em 01/02/2012.
- [56] ROUSSOS, George; KOSTAKOS, Vassilis. **RFID in Pervasive Computing: State-of-the-art and Outlook**. Pervasive and Mobile Computing. Volume 5, Issue 1, February 2009, Pages 110-131. DOI:10.1016/j.pmcj.2008.11.004.
- [57] DEYLE, Travis; KEMP, Charles C.; REYNOLDS, Matthew S. **Probabilistic UHF RFID Tag Pose Estimation with Multiple Antennas and a Multipath RF Propagation Model**. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008 – IROS 2008. Pages 1379-1384. DOI: 10.1109/IROS.2008.4651170.
- [58] HAHNEL, Dirk; BURGARD, Wolfram; FOX, Dieter; FISHKIN, Ken; PHILIPOSE, Matthai. **Mapping and Localization With RFID Technology**. IEEE International Conference on Robotics and Automation, 2004 – ICRA '04. Volume 1, Pages 1015-1020. DOI: 10.1109/ROBOT.2004.1307283.

- [59] REICHENBACH, Frank; TIMMERMANN, Dirk. **Indoor Localization with Low Complexity in Wireless Sensor Networks**. IEEE International Conference on Industrial Informatics, 2006. Pages: 1018 – 1023. DOI: 10.1109/INDIN.2006.275737.
- [60] SUGANO, Masashi et al. **Indoor Localization System Using RSSI Measurement of Wireless Sensor Network Based on Zigbee Standard**. The IASTED International Conference on Wireless Sensor Network (WSN 2006). Banff – Canadá. July 2006.
- [61] AHSAN, Kamran; SHAH, Hanifa; KINGSTON, Paul. **RFID Applications: An Introductory and Exploratory Study**. International Journal of Computer Science Issues. Volume 7, Issue 1, Number 3, Jan 2010. Disponível em: <http://arxiv.org/ftp/arxiv/papers/1002/1002.1179.pdf>, acesso em 08/03/2012.
- [62] LACHAPPELLE, Gerard. **GNSS Indoor Location Technologies**. The 2004 International Symposium on GNSS/GPS. Dec 2004. Sydney-Australia. Disponível em: http://plan.geomatics.ucalgary.ca/papers/gnss04_sydney_dec04_gl%20keynote_footers.pdf, acesso em 15/07/2011.
- [63] WATSON, John Robert A. **High-Sensitivity GPS L1 Signal Analysis for Indoor Channel Modelling**. 2005. MSc Thesis. Department of Geomatics Engineering – University of Calgary. Calgary – Alberta – Canadá. Disponível em: http://www.ucalgary.ca/engo_webdocs/GL/05.20215.JRWatson.pdf, acesso em: 15/07/2011.
- [64] KRANZ, Matthias; FISCHER, Carl; SCHMIDT, Albrecht. **A Comparative Study of DECT and WLAN Signals for Indoor Localization**. 2010 IEEE International Conference on Pervasive Computing and Communications (PerCom). Pages 235-243. DOI: 10.1109/PERCOM.2010.5466970.
- [65] NEVES, Ana Regia de Mendonça; FONSECA, Humphrey; ZOBY, Letícia; RALHA, Célia. **Localização de Usuários em Ambiente Interno Utilizando Abordagem de Sistema Multiagente**. iSys - Revista Brasileira de Sistemas de Informação, Vol. 3 (2010) PPGI – UNIRIO.
- [66] IEEE. **IEEE Std 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications**. Nova Iorque: 2007.
- [67] DEPENTHAL, Claudia. **Path Tracking with IGPS**. 2010 IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN). Pages 1-6. DOI: 10.1109/IPIN.2010.5647501