

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL**  
**FACULDADE DE INFORMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**PREDIÇÃO DO FUNCIONAMENTO DE  
SISTEMAS RFID APLICADO A  
CRACHÁS INTELIGENTES**

**EVERTON LUÍS BERZ**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciência da Computação na Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Fabiano Passuelo Hessel  
Co-Orientador: Prof. César Augusto Missio Marcon

Porto Alegre  
2011



## Dados Internacionais de Catalogação na Publicação (CIP)

B553p Berz, Everton Luís.  
Predição do funcionamento de sistemas RFID aplicado a crachás inteligentes / Everton Luís Berz. – Porto Alegre, 2011.  
96 p.

Diss. (Mestrado) – Fac. de Informática, PUCRS.  
Orientador: Prof. Fabiano Passuelo Hessel

1. Informática. 2. RFID (Identificação por Rádio Frequência).  
3. Redes Neurais (Informática). 4. Regressão Linear. 5. Sistemas Automáticos de Identificação. I. Hessel, Fabiano Passuelo. II. Título.

CDD 006.3

**Ficha Catalográfica elaborada pelo  
Setor de Tratamento da Informação da BC-PUCRS**





Pontifícia Universidade Católica do Rio Grande do Sul  
FACULDADE DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

## TERMO DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO

Dissertação intitulada "**Predição do Funcionamento de Sistemas RFID Aplicado a Crachás Inteligentes**", apresentada por Everton Luís Berz, como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação, Sistemas Embarcados e Sistemas Digitais, aprovada em 24/01/2011 pela Comissão Examinadora:

Prof. Dr. Fabiano Passuelo Hessel -  
Orientador

PPPGCC/PUCRS

Prof. Dr. César Augusto Missio Marcon -  
Coorientador

PPPGCC/PUCRS

Prof. Dr. Fernando Gehm Moraes -

PPGCC/PUCRS

Prof. Dr. Fernando César Comparsi de Castro -

PPGEE/PUCRS

Homologada em 29/03/2011, conforme Ata No. 004..... pela Comissão Coordenadora.

Prof. Dr. Fernando Luís Dotti  
Coordenador.

**PUCRS**

**Campus Central**

Av. Ipiranga, 6681 - P32- sala 507 - CEP: 90619-900

Fone: (51) 3320-3611 - Fax (51) 3320-3621

E-mail: [ppgcc@pucrs.br](mailto:ppgcc@pucrs.br)

[www.pucrs.br/facin/pos](http://www.pucrs.br/facin/pos)



## DEDICATÓRIA

Dedico este trabalho aos meus pais e a minha esposa.

“Para mim, é muito melhor compreender o universo como ele realmente é do que persistir no engano, por mais satisfatório e tranquilizador que possa parecer.”

(Carl Sagan)



## AGRADECIMENTOS

Agradeço a todos que participaram e se envolveram de alguma forma com este trabalho.

Aos professores do PPGCC pelo apoio e disposição para dirimir eventuais dúvidas.

Ao orientador Prof. Fabiano Passuelo Hessel pela compreensão e direcionamento nesta jornada.

Ao colega e professor Marcelo Azambuja nas inúmeras sessões de experimentos em campo, cansativas, mas muito produtivas.

Aos professores César Marcon, Reginaldo Caetano e Carlos Fernando Jung pela ajuda na resolução de problemas e análise dos dados coletados.

Por fim agradeço aos meus pais e a minha esposa Daniela pelo apoio incondicional, suporte, e participação ativa na realização desta dissertação de Mestrado.



# PREDIÇÃO DO FUNCIONAMENTO DE SISTEMAS RFID APLICADO A CRACHÁS INTELIGENTES

## RESUMO

A demanda por sistemas RFID tem crescido muito devido à maneira simples e eficiente de identificação automática com que tal tecnologia trabalha. A tecnologia RFID tem um grande potencial e traz muitos benefícios para organizações que precisam identificar qualquer tipo de bem e ampliar sua automação em cadeias de suprimentos. Muitas vezes sistemas RFID precisam ser implantados em ambientes onde o custo de instalação pode ser muito alto, e uma simulação de tal ambiente traria muitos benefícios, como economia de tempo e de recursos financeiros. A simulação funcionaria como um método de predição que verifica se a configuração do cenário desejado funciona corretamente e atende os requisitos do sistema. Este trabalho apresenta uma pesquisa acerca de modelos estatísticos de predição para o funcionamento de sistemas RFID em um cenário específico. O cenário escolhido aborda o uso da tecnologia RFID aplicado a crachás inteligentes. É apresentado um modelo estatístico utilizando Regressão Linear Múltipla e uma segunda abordagem utilizando Redes Neurais Artificiais na inferência dos resultados. O uso de modelos de predição pode trazer vantagens na implantação da tecnologia bem como diminuição de tempo e recursos no projeto de instalação do sistema RFID.

**Palavras Chave:** RFID, Simulação de Ambientes RFID, Regressão Linear, Redes Neurais Artificiais.



# **PREDICTION OF RFID SYSTEMS' FUNCTIONING APPLIED TO INTELLIGENT CARDS**

## **ABSTRACT**

Demand for RFID systems has risen greatly due to the simplicity and efficiency of its automatic identification system. RFID has great potential and brings many improvements to organizations that need to identify any kind of goods as well as increase their supply chain's automation level. RFID systems usually need to be applied where the installation cost can be very high, and thus, running a simulation in such environments would bring many benefits such as saving time and money. In this case, the simulation could be considered as a prediction method which verifies that the configuration for the desired scenario works correctly and meets the system's requirements. This work presents a research about prediction statistical models for RFID systems' functioning in a specific scenario. In the chosen scenario, RFID technology is applied to intelligent cards. This research presents a model using Multiple Linear Regression and another approach using Artificial Neural Networks on the prediction of the results. Prediction models can bring advantages to the adoption of RFID, like the reduction of time and resources necessary for the project installation.

**Keywords:** RFID, RFID Environment Simulation, Linear Regression, Artificial Neural Networks.



## LISTA DE FIGURAS

Figura 1 – Componentes de um sistema RFID [GLO06].	25
Figura 2 – Componentes básicos de uma etiqueta RFID [GLO06].	26
Figura 3 – Normas da tecnologia RFID e suas características [AZA08a].	29
Figura 4 – Disposição dos elementos em um ambiente com crachá inteligente.	39
Figura 5 – Sentido que a pessoa está se locomovendo pelo ambiente.	42
Figura 6 – Experimento no cenário móvel com distância de 0,8m.	44
Figura 7 – Plano 2D do local de realização dos ensaios. As elipses representam a posição do indivíduo no sistema estacionário.	44
Figura 8 – Exemplo de variância das leituras nas repetições dos experimentos para determinada situação (cenário móvel).	49
Figura 9 – Gráfico de dispersão e a linha de regressão linear [LAR10].	53
Figura 10 – Resultados dos modelos desenvolvidos utilizando regressão linear (cenário móvel).	59
Figura 11 – Relação entre a distância e o ângulo de cobertura da antena. Distâncias mais próximas da antena possuem um ângulo menor de cobertura.	60
Figura 12 – Relação entre número de leituras medido e inferido na validação do modelo linear (cenário móvel). As elipses representam o erro absoluto de cada teste.	63
Figura 13 – Modelo não-linear de um neurônio artificial [HAY01].	66
Figura 14 – Rede <i>feedforward</i> totalmente conectada com uma camada oculta e uma camada de saída [HAY01].	68
Figura 15 – Parte de uma rede recorrente sem neurônios ocultos [HAY01].	68
Figura 16 – RNA <i>feedforward</i> multi-camadas (cenário móvel).	73
Figura 17 – Relação entre leituras reais e leituras inferidas pela RNA ao final do treinamento (cenário móvel).	75
Figura 18 – Resultados da RNA para o cenário móvel.	76
Figura 19 – Relação entre o número de leituras medido e inferido na validação da RNA (cenário móvel).	78
Figura 20 – RNA <i>feedforward</i> multi-camadas (cenário estacionário).	80
Figura 21 – Relação entre leituras reais e leituras inferidas pela RNA (cenário estacionário).	81
Figura 22 – Cobertura de leituras inferida no cenário estacionário (potência: 30 dBm).	82

Figura 23 – Relação da inferência entre material, sentido, distância (Y) e potência no cenário estacionário.....	83
Figura 24 – Relação entre o número de leituras medido e inferido na validação da RNA (cenário estacionário).....	84
Figura 25 – Diagrama de radiação no plano H (horizontal) da antena MT-262010/TRH/A [MTI10].....	95

## LISTA DE TABELAS

Tabela 1 – Comparação entre as categorias de etiquetas RFID [JON07].....	27
Tabela 2 – Características das freqüências de operação da tecnologia RFID [GLO06]......	28
Tabela 3 – Fatores utilizados em ambos cenários e suas características.....	41
Tabela 4 – Fatores selecionados como variáveis de influência e o critério empregado para medição dessas variáveis.....	41
Tabela 5 – Dados coletados para a situação “material: vidro, sentido: entrando, distância: 0,8” do cenário móvel .....	48
Tabela 6 – Resultados do $R^2$ e do erro padrão na regressão linear múltipla .....	57
Tabela 7 – Coeficientes e modelos de predição para cada situação de material e sentido do cenário móvel.....	58
Tabela 8 – Amostras da validação para o modelo de regressão linear do cenário móvel.....	62
Tabela 9 – Média e desvio padrão dos erros absolutos (cenário móvel) .....	63
Tabela 10 – Amostras da validação para o modelo de RNA do cenário móvel.....	77
Tabela 11 – Diferença entre o valor inferido e o valor medido na validação do modelo utilizando RNA (cenário móvel).....	78
Tabela 12 – Amostras da validação para o modelo de RNA do cenário estacionário .....	85
Tabela 13 – Diferença entre o valor inferido e o valor medido na validação do modelo utilizando RNA (cenário estacionário).....	85



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>21</b>
1.1	Motivação.....	22
1.2	Objetivo.....	22
1.3	Organização do texto.....	24
<b>2</b>	<b>RFID</b> .....	<b>25</b>
2.1	Etiquetas.....	25
2.1.1	Frequência de operação.....	27
2.1.2	Padrões e normas.....	28
2.2	Leitor.....	29
2.3	Colisão dos sinais emitidos.....	30
2.4	Simulação de Ambientes RFID.....	31
2.4.1	Modelos estatísticos e modelos determinísticos.....	31
<b>3</b>	<b>TRABALHOS RELACIONADOS</b> .....	<b>33</b>
3.1	Floerkemeier e Sarma: RFIDSim.....	33
3.2	Nešković.....	34
3.3	Minho Jo.....	35
3.4	Conclusão sobre os trabalhos.....	36
<b>4</b>	<b>CENÁRIOS DE USO DOS CRACHÁS INTELIGENTES</b> .....	<b>39</b>
4.1	Projeto e execução dos experimentos.....	40
4.2	Software de coleta dos dados.....	46
4.3	Análise estatística dos dados.....	47
4.4	Técnicas de eliminação de dados espúrios.....	49
<b>5</b>	<b>MODELO DE PREDIÇÃO UTILIZANDO REGRESSÃO LINEAR</b> .....	<b>53</b>
5.1	Regressão Linear.....	53
5.2	Modelo.....	55
5.3	Resultados.....	57
5.4	Validação.....	60
<b>6</b>	<b>MODELO DE PREDIÇÃO UTILIZANDO REDES NEURAS ARTIFICIAIS</b> ....	<b>65</b>
6.1	Redes Neurais Artificiais.....	65
6.1.1	Arquiteturas de RNA.....	67
6.1.2	Aprendizagem da Rede Neural.....	69
6.1.2.1	Algoritmo de retro-propagação.....	70
6.1.2.2	Método do gradiente conjugado.....	70

6.1.3	Software para simulação de RNA.....	71
6.1.4	Normalização dos dados .....	72
<b>6.2</b>	<b>Cenário móvel.....</b>	<b>72</b>
6.2.1	Modelo.....	73
6.2.2	Resultados.....	74
6.2.3	Validação.....	77
<b>6.3</b>	<b>Cenário estacionário.....</b>	<b>79</b>
6.3.1	Modelo.....	79
6.3.2	Resultados.....	80
6.3.3	Validação.....	84
<b>7</b>	<b>CONCLUSÃO.....</b>	<b>87</b>
7.1	Trabalhos futuros.....	88
	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>91</b>
	<b>ANEXO A – Diagrama de radiação da antena MT-262010/TRH/A.....</b>	<b>95</b>

# 1 INTRODUÇÃO

As tecnologias RFID (*Radio Frequency Identification*) oferecem benefícios para quase todos que precisam manter a rastreabilidade de bens físicos. Implantando tecnologias RFID os fabricantes melhoram o planejamento e a execução de suas cadeias de suprimento. Já no varejo, sistemas RFID são utilizados para prevenir roubos, aumentar a eficiência das vendas, e melhorar o planejamento do estoque. Governos de determinados países já estão emitindo passaportes utilizando a tecnologia RFID para identificação automática dos passageiros nos aeroportos. Enfim, o interesse das organizações em utilizar RFID é cada vez maior, devido ao seu potencial que torna mais simples e eficiente a identificação automática de qualquer tipo de objeto.

Conforme [XAV09], uma das principais tarefas na implantação de uma solução RFID é projetar e instalar um ambiente de testes e experimentação na empresa. Esse ambiente serve para os gestores de negócio e de TI perceberem, como, de fato, uma solução RFID poderia atender as necessidades de negócio da empresa e estudar o impacto da implantação da tecnologia RFID nos processos de negócio. No ambiente de testes é possível a execução de diferentes experimentos para vislumbrar, de forma mais precisa, onde e como os equipamentos deverão ser instalados, permitindo assim estimar os gastos envolvidos na implantação de um sistema deste tipo. Entretanto, o valor do investimento associado à implantação de um ambiente de teste pode chegar a dezenas de milhares de reais em uma empresa de médio porte, dado o valor dos equipamentos e mão de obra especializada necessária para montar um case experimental.

Com o intuito de diminuir os custos associados à implantação do ambiente de teste, uma alternativa é realizar simulações do ambiente onde os testes seriam realizados. A simulação de ambientes onde a tecnologia RFID venha a ser utilizada é um dos principais desafios da área. Os métodos atuais de verificação de funcionamento do sistema consistem na realização de experimentos de campo baseados na técnica de tentativa e erro, o que consome muito tempo e recursos financeiros. Entende-se por verificação de funcionamento a possibilidade dos componentes de um sistema RFID (antena, leitor e etiqueta, detalhados no Capítulo 2) se comunicarem ou não adequadamente, tendo como principal dificuldade a detecção da etiqueta RFID no ambiente. Uma comunicação é adequada quando a etiqueta é detectada no ambiente, sendo possível realizar a leitura dos dados nela armazenados. O processo completo de detecção e leitura de etiquetas é descrito no Capítulo 2.

## 1.1 Motivação

O uso de simuladores de ambientes RFID tem o objetivo de diminuir os valores investidos no projeto de implantação desta tecnologia. Através da predição do funcionamento destes sistemas pode ser possível estimar com mais facilidade o custo associado a um projeto de sistema RFID, não necessitando realizar testes em campo, baseados na técnica de tentativa e erro.

A predição do funcionamento de sistemas RFID não é uma tarefa nem um pouco trivial, pois ambientes RFID possuem o desafio de simulações envolvendo diversas camadas de comunicação (por exemplo, camada física e camada lógica). Além disso, os componentes do sistema RFID possuem uma alta sensibilidade em relação aos fenômenos interferentes do ambiente, como, por exemplo, a reflexão do sinal de comunicação oriunda de outros objetos presentes no local. Desta forma, encontrar a posição e a configuração mais apropriada dos dispositivos pode se tornar uma tarefa árdua caso não exista um sistema de apoio. Uma ferramenta que possa simular o funcionamento do sistema RFID em determinado ambiente pode trazer diversos benefícios, tais como a facilidade de especificação e alteração do cenário, menor tempo de verificação e economia de recursos financeiros.

Os dispositivos fundamentais que compõem os sistemas RFID (leitor, antena e etiqueta) possuem uma grande variedade de tipos com diferentes características técnicas. Além disso, existem diversos tipos de ambientes onde sistemas RFID podem ser instalados. Essa realidade, somada ao fato de que um dos maiores desafios dos sistemas RFID se refere à configuração física do ambiente onde os equipamentos serão instalados [JO09a], explica o motivo pelo qual ainda existem poucas propostas de sistemas capazes de realizar a predição do funcionamento, como por exemplo simuladores de ambientes RFID.

## 1.2 Objetivo

Este trabalho tem como objetivo principal o desenvolvimento de um modelo estatístico de predição para o funcionamento de sistemas RFID em um cenário específico. O modelo estatístico, partindo de uma base de dados com informações obtidas em experimentos reais, deverá ser capaz de inferir o funcionamento do sistema RFID e levar em consideração a maioria dos aspectos significativos de um ambiente real. Uma das principais características dos modelos estatísticos diz respeito ao fato das influências essenciais do ambiente serem implicitamente levadas em consideração [NEŠ00]. Para obter a base de dados necessária no desenvolvimento do modelo estatístico foram

realizados experimentos reais em campo, sendo que esta coleta de dados foi um dos principais desafios deste trabalho (Seção 4.1).

O cenário específico de uso abordado neste trabalho contempla os sistemas RFID aplicados a crachás inteligentes. O objetivo deste tipo de sistema é o registro e a identificação de pessoas no acesso a determinados locais. Neste tipo de cenário uma etiqueta RFID é afixada em um crachá, não sendo necessário nenhum tipo de intervenção humana no processo de identificação. Desta forma, este tipo de solução ficou sendo conhecida como “crachá inteligente” [RFI09]. O Capítulo 4 descreve em detalhes o cenário de uso dos crachás inteligentes, bem como os fatores de influência deste cenário que irão fazer parte do modelo.

Visando inferir o funcionamento do sistema RFID, o modelo de predição desenvolvido deverá resultar a intensidade de leituras da etiqueta RFID, de acordo com os valores das variáveis informadas no modelo e levando em consideração todas as influências presentes no ambiente. Tais variáveis foram selecionadas dentre os fatores mais relevantes encontrados no cenário de uso específico utilizado neste trabalho, sendo elas: localização da etiqueta no ambiente, material da parede da sala de testes, sentido de movimentação da pessoa dentro do ambiente e potência configurada no equipamento leitor. A descrição completa e o critério de seleção destas variáveis estão descritos na Seção 4.1

Uma das variáveis do modelo diz respeito à localização da etiqueta no ambiente. A partir desta variável, o modelo deverá ser capaz de inferir a intensidade de leituras de etiqueta para determinado ponto no ambiente onde o sistema RFID está atuando. Aplicando o modelo em todo o cenário, uma simulação pode resultar em uma série de informações capazes de identificar a intensidade de leituras de cada ponto do cenário de uso. Por exemplo, em um cenário de 3 m de largura por 2 m de comprimento pode-se aplicar o modelo de predição em cada ponto no intervalo de 0,1 m. Desta forma, é possível saber previamente em quais locais uma etiqueta terá uma probabilidade maior de ser lida, e, principalmente, conhecer os pontos do ambiente mais propícios para realizar a leitura com sucesso.

O trabalho busca uma abordagem inovadora em relação à predição do funcionamento de sistemas RFID, sendo utilizado o número de leituras de uma etiqueta para verificar a cobertura do sistema RFID no ambiente. Também não é de conhecimento algum outro trabalho que considere o material da parede dos ambientes internos como um fator a ser considerado na verificação de sistemas RFID. Além disso, este trabalho considera cenários com pessoas portando uma etiqueta, inclusive em movimento. A única

publicação encontrada, relacionada à influência de seres humanos no uso da tecnologia RFID, está concentrada na análise da dispersão dos sinais de RF próximos à cabeça de um indivíduo parado [JAI10].

Ao conquistar o objetivo, o trabalho visa apoiar a implantação da tecnologia RFID, reduzindo o tempo e recursos financeiros necessários no projeto de um sistema deste tipo.

### **1.3 Organização do texto**

O trabalho está estruturado da seguinte forma: no Capítulo 2 há uma explanação sobre RFID e seus componentes, sendo também abordados os métodos de simulação para ambientes RFID. No Capítulo 3 são listados os trabalhos correlatos. No Capítulo 4 é apresentado o cenário de uso da tecnologia RFID abordado neste trabalho, incluindo o projeto e execução dos experimentos realizados em tal cenário. Nos Capítulos 5 e 6 são desenvolvidos os modelos de predição utilizando Regressão Linear e Redes Neurais Artificiais, respectivamente. Finalmente, no Capítulo 7 é apresentada a conclusão e os trabalhos futuros que podem vir a ser realizados.

## 2 RFID

A sigla RFID significa *Radio Frequency Identification*, ou Identificação por Rádio Freqüência, sendo a expressão que representa qualquer sistema de identificação onde existe um dispositivo eletrônico que usa rádio frequência (RF) ou variações de campo magnético para efetuar a comunicação entre os elementos do sistema [GLO06].

Segundo [GLO06], os dois principais componentes de um sistema RFID são:

- Etiqueta (também chamada de *tag* ou *transponder*): dispositivo de identificação afixado ao objeto que se deseja rastrear;
- Leitor (ou interrogador): é o dispositivo que pode detectar a presença de etiquetas RFID e ler as informações nelas armazenadas. Dependendo da tecnologia ou do projeto, um leitor também pode escrever informações nas etiquetas.

Geralmente o leitor está conectado em algum sistema que possui um software cuja função é intermediar a transmissão de dados entre o leitor e a aplicação. Este software é denominado *middleware* RFID. O computador que executa o software *middleware* geralmente é conhecido como servidor do *middleware* RFID. O leitor gera sinais eletromagnéticos que são transmitidos por uma ou mais antenas. Como segunda função, o leitor também monitora sinais eletromagnéticos oriundos das etiquetas, usualmente utilizando o mesmo conjunto de antenas [FIN03][GLO06]. A Figura 1 ilustra como se relacionam os componentes do sistema RFID.

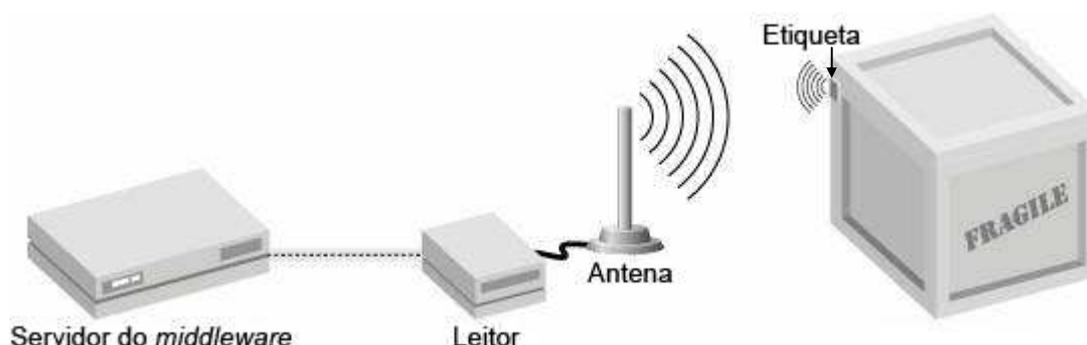


Figura 1 – Componentes de um sistema RFID [GLO06].

### 2.1 Etiquetas

As etiquetas são os dispositivos que, afixados ao que se deseja identificar, respondem aos sinais emitidos pelo equipamento leitor. As etiquetas podem ser colocadas diretamente em cada objeto individualmente (como em um determinado

produto ou cartão de identificação), ou em containers ou *pallets* que carregam vários objetos. A principal função de uma etiqueta é a transmissão de dados para o restante do sistema RFID. As etiquetas geralmente contêm três partes básicas: (i) um circuito integrado eletrônico; (ii) uma mini antena; (iii) um substrato que mantém o circuito e a antena juntos e permite que seja colocado um adesivo (para afixar a etiqueta em algum objeto) [JON07]. A Figura 2 exibe o desenho de uma etiqueta RFID e seus componentes básicos. Etiquetas RFID podem conter outros componentes, dependendo da categoria em que elas se enquadram, sendo estes descritos a seguir.

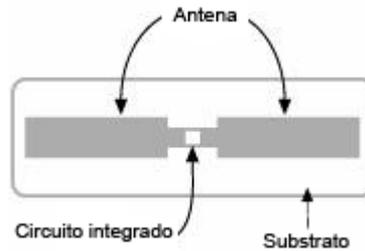


Figura 2 – Componentes básicos de uma etiqueta RFID [GLO06].

As etiquetas de um sistema RFID estão enquadradas em três categorias: ativas, passivas e semi-passivas. As etiquetas ativas possuem uma fonte própria de energia, tal como uma bateria, e podem tomar a iniciativa de iniciar a transmissão de dados. Etiquetas ativas podem conter componentes mais sofisticados como GPS<sup>1</sup>, monitores de umidade, temperatura, vibração ou movimento. As etiquetas passivas obtêm sua energia do sinal transmitido pelo leitor, emitindo informações somente após terem sido interrogadas por um equipamento leitor. As etiquetas semi-passivas possuem fonte própria de energia para alimentar o circuito da etiqueta, mas não possuem funcionalidades de transmissão ativa, apenas respondendo os seus dados quando questionadas pelo leitor [JON07]. Devido à grande quantidade de modelos e tecnologias utilizadas na fabricação de etiquetas RFID, não é possível especificar em definitivo um tamanho, custo ou capacidade de armazenamento. Entretanto, as categorias de etiquetas possuem padrões que tornam as características mencionadas distintas entre cada uma delas. A Tabela 1 faz uma comparação entre as características das categorias de etiquetas existentes.

---

<sup>1</sup> GPS: *Global Positioning System* ou Sistema de Posicionamento Global.

Tabela 1 – Comparação entre as categorias de etiquetas RFID [JON07].

	<b>Etiqueta Ativa</b>	<b>Etiqueta Passiva</b>	<b>Etiqueta Semi-passiva</b>
<b>Fonte de energia</b>	Bateria interna	Energia transferida pelo leitor	Circuito usa bateria interna; Transmissão de dados usa a energia transferida pelo leitor
<b>Faixa de distância de leitura</b>	Medido em milhares de metros	Medido em metros	Medido em metros
<b>Tamanho</b>	Maior	Menor	Menor
<b>Armazenamento de dados</b>	Mais espaço	Menos espaço	Mais espaço
<b>Custo</b>	Mais alto	Mais baixo	Mais alto

As etiquetas RFID funcionam da seguinte maneira: o equipamento leitor gera um campo eletromagnético que induz uma corrente para a antena da etiqueta. A corrente é utilizada para ativar o circuito integrado. No caso de etiquetas passivas, a corrente energiza um capacitor existente neste tipo de etiqueta, o qual fornece energia para o circuito integrado. Nas etiquetas ativas existe uma bateria própria para realizar essa função. Uma vez ativa, a etiqueta recebe comandos do leitor e responde enviando seu identificador ou a informação solicitada. De maneira geral, a etiqueta não tem energia suficiente para criar seu próprio campo eletromagnético, ao invés disto ela utiliza a reflexão (*backscattering*) para modular<sup>1</sup> o sinal enviado pelo leitor. Cabe observar que muitos fluídos absorvem sinais eletromagnéticos e muitos metais refletem estes sinais, sendo a leitura de etiquetas na presença destes materiais um problema no uso da tecnologia [JEC06].

### 2.1.1 Freqüência de operação

Os sistemas RFID geram e irradiam ondas eletromagnéticas e por este motivo são classificados como sistemas de rádio. Os sistemas RFID não podem interferir em outros serviços de rádio, como telefonia móvel e transmissões de televisão. Assim, existem freqüências de operação definidas para os sistemas RFID. A freqüência de operação é a freqüência eletromagnética que a etiqueta usa para se comunicar e para obter energia. Sistemas RFID tipicamente operam em quatro faixas de freqüência: (i) LF: *low frequency*; (ii) HF: *high frequency*; (iii) UHF: *ultra-high frequency*; (iv) microondas. Devido ao cuidado

---

<sup>1</sup> Modulação é o processo que altera as características de um sinal de alta freqüência (por exemplo, a amplitude, freqüência ou fase) [FIN03].

necessário para assegurar que sistemas RFID não interfiram em outros sistemas de rádio, somente é possível utilizar RFID em faixas de frequência reservadas especificamente para aplicações médicas, científicas e industriais (*ISM frequency ranges, Industrial-Scientific-Medical*) [FIN03][GLO06]. A Tabela 2 exhibe as faixas de frequência de sistemas RFID.

Tabela 2 – Características das frequências de operação da tecnologia RFID [GLO06].

<b>Frequência</b>	<b>Faixa</b>	<b>Frequência ISM</b>	<b>Distância máxima de leitura para etiquetas passivas</b>
LF	30 a 300 kHz	< 135 kHz	50 cm
HF	3 a 30 MHz	6,78 MHz; 13,56 MHz; 27,125 MHz; 50,680 MHz	3 m
UHF	300 MHz a 3 GHz	433,92 MHz; 869 MHz; 915 MHz	9 m
Microondas	> 3 GHz	2,45 GHz; 5,8 GHz; 24,125 GHz	> 10 m

### 2.1.2 Padrões e normas

Atualmente existem normas que padronizam e classificam as diversas faixas de frequência, protocolos anti-colisão (Seção 2.3) e codificação do sinal utilizados na tecnologia RFID. As principais organizações de padronização para RFID são a ISO (*International Organization for Standardization*) e a EPCGlobal, esta última uma instituição formada para controlar a numeração dos números seriais únicos para cada etiqueta ao redor do mundo (a classe de codificação EPC - *Electronic Product Code*), bem como criar padrões para a tecnologia (incluindo protocolos e frequências), e promover pesquisas sobre RFID. O padrão ISO para RFID é a série de normas ISO 18000, cobrindo também os protocolos para a comunicação sem fio. Estes padrões abrangem as atuais frequências utilizadas para RFID ao redor do mundo. Na Figura 3 são relatadas as sete partes da ISO 18000, a classificação da EPCGlobal e outras tecnologias que são discutidas ao longo deste trabalho.

Devido à combinação de tamanho da etiqueta, baixo custo, arquitetura simples e baixo consumo de energia, grande parte dos esforços para a utilização de RFID são direcionados para as etiquetas passivas UHF, padronizadas pelas normas ISO 18000-6 C e EPCGlobal Gen2 [RID10][GLO06][AZA08a]. Por este motivo, este trabalho está concentrado nas características físicas e sistema de comunicação utilizados por estas etiquetas.

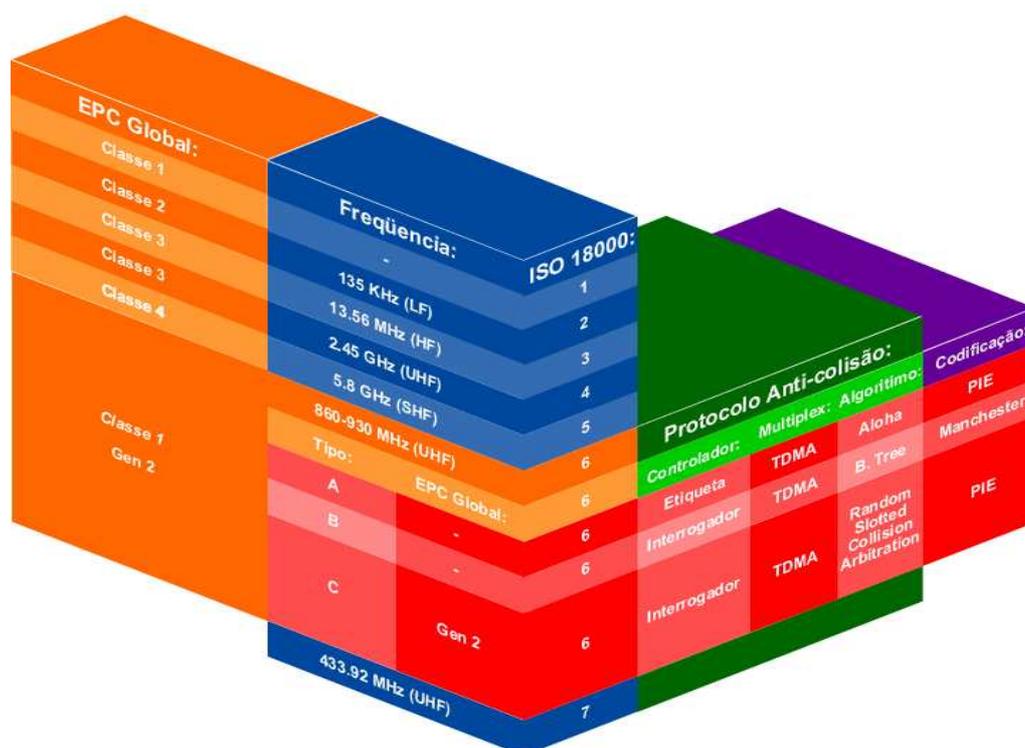


Figura 3 – Normas da tecnologia RFID e suas características [AZA08a].

## 2.2 Leitor

O leitor é um dispositivo que gera sinais eletromagnéticos, transmitidos para as etiquetas RFID através de uma ou mais antenas. Em uma operação normal (sem interferências), o leitor está continuamente transmitindo os sinais eletromagnéticos com o intuito de buscar etiquetas no ambiente. O leitor RFID também é responsável por monitorar sinais eletromagnéticos provenientes das etiquetas RFID.

O leitor contém um módulo de rádio frequência (transmissor e receptor), uma unidade de controle e um elemento de conexão (normalmente uma antena) com a etiqueta RFID. Além disso, muitos leitores são produzidos com uma interface adicional (Ethernet, RS 232, RS 485, etc) para possibilitar o direcionamento dos dados recebidos para o software *middleware*, geralmente instalado em um PC [FIN03].

Segundo [KIM08], em um sistema RFID o canal de comunicação é compartilhado entre todos os dispositivos do sistema, sendo dividido em um canal de envio e outro de retorno. No canal de envio, o transmissor do leitor envia um sinal eletromagnético para a etiqueta. No canal de retorno, o leitor recebe um sinal refletido (*backscattered*) pela etiqueta. Diferentemente de outros sistemas de comunicação sem fio, na tecnologia RFID não existe reciprocidade entre os canais de envio e retorno, ou seja, o canal de envio possui um alcance de distância muito maior do que o existente no canal de retorno.

A comunicação na camada física de um sistema RFID sofre uma grande quantidade de tipos de interferência, tais como aquelas causadas por barreiras físicas entre os elementos comunicantes, sinais que geram ruídos na comunicação original, reflexões, entre outras. No caso da tecnologia RFID, em específico, ainda existe a particularidade da baixa potência do sinal emitido pelas etiquetas do tipo passivas, aumentando a dificuldade da comunicação. Por fim, as características dos componentes da etiqueta, como antena, circuito e materiais utilizados na sua confecção, tornam as questões físicas sobre o ambiente onde o sistema estará instalado um fator determinante para o bom funcionamento da tecnologia [KIM08].

### 2.3 Colisão dos sinais emitidos

A colisão de sinais provenientes de etiquetas é um fator importante que atualmente limita a melhora do desempenho dos sistemas RFID. Este problema exige a aplicação de protocolos anti-colisão para a identificação do sinal individual de cada etiqueta. Estes sinais utilizam a mesma frequência, e quando emitidos simultaneamente por mais de uma etiqueta se tornam não identificáveis para os equipamentos interrogadores. Os protocolos anti-colisão atualmente padronizados utilizam algoritmos que tentam identificar as mensagens individuais de cada etiqueta eletrônica [FIN03].

Os protocolos anti-colisão utilizados por dispositivos RFID são baseados em algoritmos executados nas etiquetas eletrônicas e no leitor. O leitor percebe a presença de mais de uma etiqueta no ambiente através do recebimento de um sinal não identificável, composto por sinais sobrepostos devido respostas simultâneas de mais de uma etiqueta. Quando isto acontece, o leitor inicia a execução do algoritmo anti-colisão, envolvendo rotinas executadas também pelas etiquetas [AZA08a][FIN03].

Conforme pode ser verificado na Figura 3, o protocolo anti-colisão utilizado pelas etiquetas EPCGlobal Gen2 (foco deste trabalho) é o Random Slotted (Algoritmo Q). Neste protocolo alguns controles são realizados pela etiqueta, embora o controlador geral do processo seja o leitor. Segundo [AZA08a], este protocolo utiliza um mecanismo que aloca as transmissões das etiquetas em etapas (*rounds*) e espaços (*slots*). Um *round* consiste em um número de *slots*. Cada *slot* possui uma duração com tempo suficiente para o leitor receber a resposta de uma etiqueta, sendo que o leitor determina este tempo.

O leitor, quando inicia o processo de identificação Random Slotted, envia para as etiquetas um parâmetro denominado parâmetro Q. As etiquetas participantes do processo de identificação sorteiam um valor na faixa entre zero e  $2^Q - 1$  e armazenam este valor no respectivo contador de *slot*. A etiqueta que sortear o valor zero neste momento pode

responder imediatamente. Etiquetas que sortearam valores maiores que zero aguardam instruções do leitor para decrementar o valor do seu contador, e cada etiqueta que chegar a zero poderá responder [AZA08a].

## 2.4 Simulação de Ambientes RFID

Segundo [CHW07], um modelo de simulação procura capturar o comportamento de um dado sistema e representá-lo como um modelo computacional. Todo modelo de simulação deve ser capaz de representar os diversos fenômenos aleatórios existentes no sistema, construídos a partir da observação e levantamento de dados do fenômeno, de modo que este modelo de simulação seja o mais representativo possível do sistema real.

A dificuldade, e até o impedimento, em conceber um projeto de implantação de um sistema RFID se deve a fenômenos como a distância máxima de leitura das etiquetas, interferência eletromagnética e colisões causadas pela leitura de várias etiquetas simultaneamente [KIM08]. A partir disso conclui-se que para uma simulação de sistema RFID ser completa, uma simulação deve atender a camada física e lógica da tecnologia.

Este trabalho está direcionado para a simulação da camada física da tecnologia RFID, isto se deve ao fato de que já existia, no grupo de estudos ligado a este trabalho, uma pesquisa relacionada à simulação da camada lógica. Além disso, a simulação da camada física, relacionada especificamente à tecnologia RFID, é um assunto de maior complexidade e com poucos resultados encontrados na literatura até o momento.

Basicamente os métodos de simulação da camada física do sistema RFID podem ser encontrados nas pesquisas feitas acerca dos métodos de predição em sistemas RF, pois este é o método de comunicação presente nos principais dispositivos RFID.

Como este trabalho tem um foco maior nas dificuldades de projeto da camada física, os experimentos e modelos desenvolvidos atuaram sempre sobre uma etiqueta única. Desta forma, a simulação da camada lógica ficou bastante simplificada, sendo que não existe atuação direta dos algoritmos anti-colisão (Seção 2.3) no modelo gerado. A configuração do sistema foi preparada para não fazer uso de tais algoritmos, conforme especificado na Seção 4.2.

### 2.4.1 Modelos estatísticos e modelos determinísticos

Segundo [NEŠ00], modelos de predição podem ser tanto estatísticos (baseados em dados empíricos) ou determinísticos (específicos de determinado ambiente), ou a combinação de ambos.

Modelos estatísticos, ou empíricos, usam a observação direta dos dados operacionais do processo obtidos através de experimentação (relações causa/efeito

correlacionando dados de entrada/saída do processo). Neste caso, sinais de entrada e de saída do sistema são registrados e submetidos a uma análise para se inferir um modelo. Na construção de modelos empíricos, os modelos são determinados efetuando-se pequenas alterações nas variáveis de entrada em torno de uma condição nominal de operação. A resposta dinâmica resultante é usada para determinar o modelo. Esse procedimento gera modelos experimentais do processo válidos em alguma região em torno das condições experimentais em que os dados de entrada e saída do processo foram medidos [GAR09].

Nos modelos estatísticos todas as influências do ambiente são implicitamente levadas em consideração. A precisão destes modelos não depende somente da exatidão dos experimentos, mas também das semelhanças entre o ambiente simulado e o ambiente onde os experimentos foram realizados. Por outro lado, o desempenho computacional destes modelos é satisfatório na maioria das vezes [NEŠ00].

Os modelos determinísticos são baseados nos princípios da física, e devido a isso eles podem ser aplicados a diferentes ambientes sem afetar a precisão dos resultados. A implementação de modelos determinísticos normalmente necessita de uma base de conhecimento enorme acerca do ambiente, o que às vezes é impossível de se obter. Os algoritmos destes modelos geralmente são muito complexos e não possuem um bom desempenho no processamento computacional. Por esse motivo, a implementação de modelos determinísticos é usualmente restrita a áreas e problemas menores. Se os modelos determinísticos forem aplicados corretamente, a precisão dos resultados de uma predição pode ser superior aos resultados de um modelo estatístico [NEŠ00].

### 3 TRABALHOS RELACIONADOS

Na pesquisa sobre modelos de predição para sistemas RFID podem ser encontradas publicações acerca da simulação da camada lógica e da camada física da tecnologia, bem como trabalhos que detalham o funcionamento de ambas as camadas, como o estudo dos protocolos anti-colisão e da propagação de sinais de RF.

O presente trabalho está engajado em um grupo de pesquisa cuja preocupação inicial era o problema das colisões de mensagens e o estudo dos protocolos anti-colisão de sistemas RFID, assunto com publicações em [AZA08a], [AZA08b] e [AZA10].

Abordando a camada física da tecnologia, [NIK08] apresenta um estudo detalhado sobre propagação de sinais RFID, considerando diversas características técnicas e físicas dos dispositivos, porém não são apresentados modelos de ambiente de uso.

Em [JAI10], é apresentado um estudo focado na propagação de RF em dispositivos RFID com o corpo humano localizado próximo de uma etiqueta. O teste é específico para etiquetas da marca Alien Squiggle [ALI10]. Na pesquisa o ambiente não é considerado, sendo analisados os fatores potência e frequência do sistema RFID.

Apesar de estar focada somente na predição de sinais de RF, a pesquisa de [NEŠ00] representa o passo inicial para a idéia utilizada neste trabalho. Devido a sua importância, esta pesquisa está detalhada nas seções seguintes.

Dentre os trabalhos encontrados, possuem destaque as publicações dos cientistas Floerkemeier, Sarma [FLO09] e Minho Jo [JO07][JO09a][JO09b]. Tais trabalhos foram os únicos encontrados que se concentram no desenvolvimento de modelos de predição para sistemas RFID. Estes trabalhos serão detalhados nas seções seguintes.

#### 3.1 Floerkemeier e Sarma: RFIDSim

Em [FLO09], Floerkemeier e Sarma, dois cientistas do Auto-ID/MIT, entre os principais autores na área de RFID, apresentam o RFIDSim (RFID *simulation engine*). O RFIDSim é uma ferramenta que simula em uma máquina de eventos discretos o funcionamento dos algoritmos anti-colisão e os aspectos físicos de funcionamento de um sistema RFID. Entretanto, a simulação dos aspectos físicos é realizada de forma unicamente teórica, sem testes reais e considerando os modelo de propagação de RF *pathloss* e *Rician-fading*.

Um dos modelos de propagação RF apresentados é o modelo básico de perdas do caminho (*pathloss*), onde a potência recebida diminui de acordo com a relação entre a distância e o comprimento de onda, elevado à potência  $n$ . Na ferramenta proposta, a

potência  $n$  deve ser configurada pelo usuário. Já o *Rician-fading* é um modelo estatístico, sendo as amostras geradas de acordo com a distribuição Rician, tendo um foco maior na propagação das ondas no espaço em relação ao tempo.

Neste trabalho não são consideradas as influências do ambiente sobre o modelo de propagação RF, tais como objetos interferentes que podem atenuar ou refletir os sinais, e o problema das etiquetas em movimento diante da antena do leitor também é discutido de forma teórica. As simulações e resultados apresentados consideraram o funcionamento dos protocolos anti-colisão e resultados teóricos sobre o modelo de propagação de RF, mas sem citações sobre como os protocolos anti-colisão foram implementados, tendo em vista que partes fundamentais das rotinas desses não são especificadas nos documentos de padronização

### 3.2 Nešković

Em [NEŠ00] é apresentado um modelo de predição para o nível de intensidade de campo de sinais de RF em ambientes internos. Apesar do modelo não ser focado diretamente na tecnologia RFID, ele aborda os princípios básicos de propagação de sinais de RF, mesmo método de comunicação presente nos dispositivos RFID. O modelo é baseado nos princípios das Redes Neurais Artificiais (RNA) *feedforward* com retro-propagação. A análise inicial do modelo se dá através da construção de uma base de dados com experimentos realizados em salas de dois prédios distintos. Os locais onde ocorreram os ensaios tiveram seus ambientes classificados em 11 categorias: parede, corredor, área externa, laboratório, sala de aula, auditório, porta, janela, elevador, escadas para o andar de cima, e escadas para o andar de baixo. Um banco de dados para armazenar a informação de intensidade de campo foi criado utilizando uma escala de 10 x 10 cm.

O modelo de RNA desenvolvido baseou-se em uma rede *perceptron feedforward* multi-camadas. A rede neural possui 12 entradas, sendo que o primeiro nodo recebeu a distância normalizada entre o transmissor e o receptor. Também existe uma entrada para cada categoria de ambiente definido. Cada entrada representa o número de ocorrências (janelas, portas, etc) ou um percentual (parede, corredor, etc) da categoria que está posicionada entre o transmissor e o receptor (traçando uma linha reta entre eles).

Mais de 600.000 amostras de valores de intensidade de campo em cada posição dos ambientes foram coletadas. Os experimentos foram realizados com o transmissor em cinco locais diferentes em cada um dos dois prédios. As amostras coletadas nos dois primeiros locais de transmissão do primeiro prédio serviram de treinamento para a rede. A

verificação do modelo se deu através das amostras coletadas em cinco diferentes transmissores de ambos os prédios, onde foi comparado o valor de intensidade inferido com o valor respectivo obtido no experimento real. A exatidão do modelo desenvolvido é parecida (ou até melhor) do que os modelos de predição mais precisos conhecidos. O modelo de RNA desenvolvido pode ser comparado com os modelos mais precisos projetados para ambientes internos.

### 3.3 Minho Jo

Concentrado em modelos para predição de sistemas RFID, em [JO07] é desenvolvido uma RNA com retro-propagação para a inferência da detecção de etiquetas RFID próximas a garrafas de água. O cenário de uso definido neste trabalho consiste em um ambiente com uma esteira transportando uma caixa de papelão com garrafas de água dentro. Segundo [RAM03], é bastante comum as etiquetas RFID serem afixadas em produtos que contém água, sendo que este tipo de material provoca efeitos consideráveis no desempenho do sistema. As antenas RFID foram instaladas em forma de pórtico sobre a esteira, tendo o objetivo de detectar as etiquetas afixadas na caixa de papelão em movimento. Dentre todos os fatores existentes no cenário, a posição e a velocidade da etiqueta foram consideradas as variáveis de maior relevância, sendo representados por nodos de entrada da rede neural modelada. O treinamento da rede foi executado utilizando 800 amostras provenientes de experimentos reais. A rede neural modelada possui somente um nodo na camada de saída, o qual representa um valor lógico que informa se a etiqueta foi lida ou não. Os resultados de predição foram verificados alterando-se o número de camadas ocultas da rede, tendo um percentual mínimo de predição na faixa de 70% e o percentual máximo de aproximadamente 96%.

Como continuação desse trabalho, [JO09b] apresenta um modelo de detecção de etiquetas RFID utilizando Máquinas de Vetor de Suporte (*Support Vector Machines - SVM*). Segundo [HAY01], SVM se trata de uma outra categoria de redes *feedforward*, sendo reconhecido que as SVM executam os mesmos métodos que uma RNA usa para solucionar problemas de classificação. Assim como os perceptrons de múltipla camadas e as redes de função radial, as SVM podem ser usadas para classificação de padrões e regressão linear. Fundamentalmente, a máquina de vetor de suporte é uma máquina linear, podendo fornecer um bom desempenho de generalização em problemas de classificação de padrões.

Em [JO09b], são analisadas as características de detecção de etiquetas afixadas em caixas de papelão em um ambiente estacionário. As diversas caixas de papelão estão

sobre um *pallet*, caracterizando um cenário típico de uso de sistemas RFID. O objetivo do modelo é a predição de intensidade do sinal de cada etiqueta afixada nas caixas de papelão. Para medir a intensidade do sinal foi utilizado o software EPC Hotspot [QTR10]. Como solução de predição, inicialmente foi testado um modelo de regressão linear, o qual não teve resultados satisfatórios, sendo então aplicado um modelo baseado em SVM. A técnica de SVM tem sido aplicada com sucesso em muitos problemas de classificação não-linear. Os resultados de predição se deram através da comparação do valor de intensidade inferido com o valor respectivo obtido no experimento real, sendo o percentual de precisão próximo a 95%.

Em [JO09a], a mesma técnica de SVM foi aplicada a outro cenário de uso típico RFID. Neste cenário as etiquetas foram afixadas no vidro da frente de uma caminhonete e as antenas foram posicionadas em um pórtico (sistema semelhante aos utilizados em praças de pedágio). Durante a execução dos experimentos o veículo atravessava o pórtico em cinco diferentes velocidades. O objetivo do trabalho era a inferência da detecção de etiquetas em determinada velocidade de movimentação no ambiente descrito. Também foram realizadas análises sobre a posição e o ângulo da antena afixada no pórtico. Os resultados de predição se deram através da comparação da detecção de etiquetas inferidas com o valor respectivo obtido no experimento real, sendo o percentual de precisão entre 94,4 e 100%.

### **3.4 Conclusão sobre os trabalhos**

O trabalho de [FLO09] apresentou uma ferramenta de simulação que leva em consideração somente alguns dos vários aspectos que influenciam a comunicação de sistemas RFID. Ainda, a única parametrização encontrada para a ferramenta está baseada na informação da potência  $n$  utilizada na equação de perdas do sinal. Tal informação é extremamente técnica e precisa ser medida a partir de experimentos em campo no ambiente que se deseja simular. Também não são apresentados testes em campo para a validação dos resultados obtidos na ferramenta. Este trabalho poderia ser melhorado a partir do uso de outros modelos de propagação, onde as informações para a parametrização do ambiente pudessem ser mais fáceis de serem obtidas.

O trabalho de [NEŠ00] apresentou resultados bastante satisfatórios para a predição de sinais de RF utilizando Redes Neurais Artificiais. Neste trabalho, a variável resposta da predição é representada pela intensidade do sinal de cada ponto do ambiente, sendo coletada a partir de instrumentos específicos para esta tarefa. Infelizmente, no desenvolvimento desta dissertação, não foi possível o uso de tais instrumentos de

medição. Entretanto, a idéia de realizar a predição da cobertura do sinal inspirou o desenvolvimento dos modelos desta dissertação, onde também se propõe o uso de modelos estatísticos aplicados à ambientes internos. No caso, utilizou-se uma abordagem inovadora para a variável resposta da predição: o número de leituras capturado pelo sistema RFID.

Minho Jo apresenta, em três publicações, modelos de predição estatísticos para cenários específicos de uso da tecnologia RFID. Em [JO09b], a variável inferida pelo modelo é a intensidade do sinal de cada etiqueta RFID, obtida através do software EPC Hotspot. Não é de conhecimento a forma como o EPC Hotspot obtém o valor da intensidade do sinal, sendo provável que a intensidade do sinal seja inferida por algum modelo de propagação. Cabe ressaltar que no software não existe interface para nenhum instrumento de medição externo. Assim, é possível que [JO09b] realize estimativas sobre um valor que já é resultado de uma inferência de outro modelo. Nos trabalhos [JO07] e [JO09a], tanto as estimativas realizadas utilizando Redes Neurais Artificiais, quanto os modelos de SVM, utilizam a informação de leitura (se houve ou não leitura) da etiqueta na inferência dos resultados, obtendo resultados satisfatórios na etapa de validação.

A metodologia desenvolvida por Minho Jo para o desenvolvimento dos modelos serve de suporte para esta dissertação, principalmente no que diz respeito ao projeto dos experimentos e fatores que são considerados no cenário de uso específico proposto.

Cada um dos trabalhos apresentados possui uma abordagem diferenciada e contribuem de forma considerável para esta dissertação. Em destaque, as publicações de Nešković e Minho Jo servem de suporte para o desenvolvimento dos modelos de predição deste trabalho.



## 4 CENÁRIOS DE USO DOS CRACHÁS INTELIGENTES

Um cenário típico de uso de sistemas RFID pode ser encontrado na entrada de determinados ambientes que necessitam registrar a atividade de entrada e saída de pessoas, ou até identificar os indivíduos que estão fazendo acesso ao local. Os exemplos são vários, dentre eles destacam-se os seguintes: sistemas RFID instalados na porta de entrada de determinada empresa quando se deseja registrar o horário de entrada e saída dos funcionários; em determinado evento quando é necessário saber a quantidade de pessoas que entraram e saíram do ambiente, possibilitando calcular a média de permanência das pessoas no evento; ou então o apoio do sistema RFID na realização do controle de acesso de pessoas a um ambiente, permitindo ou não o acesso ao local.

Uma das formas conhecidas de uso da tecnologia RFID consiste em afixar uma etiqueta passiva RFID no crachá dos indivíduos e posicionar a antena do leitor RFID ao lado da porta, em paralelo à parede, conforme é exibido na Figura 4. De acordo com [RFI09], a adoção de um crachá inteligente é especialmente indicada em instalações como hospitais, hotéis, escolas, parques de diversões, cruzeiros e edifícios governamentais, bem como em eventos, garantindo mais segurança e precisão no controle de acesso do público e agilidade no atendimento. Em destaque na Figura 4 estão a antena do leitor (1) e a etiqueta afixada em um crachá de plástico (2).



Figura 4 – Disposição dos elementos em um ambiente com crachá inteligente.

Como este trabalho está focado em modelos estatísticos, uma das primeiras atividades a ser realizada para a construção do modelo diz respeito à coleta dos dados em campo, sendo executada através de experimentos reais.

#### **4.1 Projeto e execução dos experimentos**

Durante o projeto dos experimentos foram preparados dois cenários bastante realísticos para a realização dos mesmos, sendo o objetivo do primeiro cenário a execução de ensaios com a etiqueta em movimento (cenário móvel) e no segundo cenário os experimentos são realizados com o indivíduo parado em diversos pontos do ambiente (cenário estacionário).

No uso de sistemas RFID com crachás inteligentes, os principais fatores que influenciam na detecção das etiquetas RFID são: material onde a etiqueta está afixada; modelo do leitor e etiqueta; orientação da etiqueta (vertical ou horizontal); sentido de locomoção do indivíduo (entrando ou saindo da sala); distância entre a etiqueta e o leitor; potência de transmissão do leitor; ganho das antenas; frequência de operação; altura da antena do leitor em relação ao solo; velocidade de movimentação da etiqueta (apenas no cenário móvel); número de etiquetas no sistema; e por fim as características originais do ambiente (dimensões da sala e da porta de acesso, materiais de construção da sala e de objetos interferentes, etc), destacando neste cenário o material da parede que está ao lado da antena, bem como a largura e altura da porta. Os fatores que foram utilizados na realização dos testes em campo e suas características estão descritos na Tabela 3.

Para os fatores que possuem propriedades fixas foram utilizadas características que mais se assemelham ao caso dos ambientes típicos relatados no início deste Capítulo. Por exemplo, a altura da etiqueta em relação ao solo equivale à altura do crachá utilizado por uma pessoa de estatura média (1,75 m), a velocidade da etiqueta em movimento é o equivalente a uma pessoa caminhando, o material do crachá geralmente é de plástico, e a altura da antena foi fixada na mesma altura da etiqueta. No caso do fator “Número de objetos com etiqueta” houve uma simplificação do mundo real (se sabe que muitas vezes existem diversos indivíduos e etiquetas simultâneas próximos do leitor), os motivos que levaram a utilizar somente uma etiqueta foram explicados na Seção 2.4.

Levando em consideração a criação dos dois cenários e todos os fatores citados anteriormente, foram selecionados para cada cenário os fatores considerados mais importantes, definindo-os como as variáveis de maior influência para o funcionamento do sistema. Tais variáveis serão os itens que definirão o modelo proposto neste trabalho. A definição das variáveis e o critério de medição são apresentados na Tabela 4.

Tabela 3 – Fatores utilizados em ambos cenários e suas características

Fatores		Característica
<b>Etiqueta</b>	Material onde está afixada	Plástico
	Objeto	Crachá afixado em uma pessoa de estatura média (1,75m)
	Modelo	UPM Rafsec UHF DogBone 3000825
	Altura em relação ao solo	1,25 m
	Velocidade	Cenário móvel <sup>1</sup> : 1,5 m/s Cenário estacionário: estático
	Orientação	Horizontal
	Sentido	Variável
<b>Leitor</b>	Modelo do leitor	Mercury 4
	Modelo da antena	MTI MT-262010/TRH/A
	Ganho da antena	8 dBi <sup>2</sup>
	Altura da antena (centro)	1,25 m
	Frequência	915,25 MHz
	Número de antenas	1
	Potência	Variável
<b>Ambiente</b>	Número de objetos com etiqueta	1
	Distância entre leitor e etiqueta	Variável
	Altura x largura da porta	2,20 x 0,82
	Material da parede	Variável (salas diferentes)

Tabela 4 – Fatores selecionados como variáveis de influência e o critério empregado para medição dessas variáveis

Fatores		Cenário móvel	Cenário estacionário
<b>Etiqueta</b>	Sentido	Indivíduo entrando ou saindo da sala, conforme Figura 5.	Crachá direcionado para o interior da sala (Entrando) ou crachá direcionado para o lado externo da sala (Saindo), conforme Figura 5.
<b>Leitor</b>	Potência (em dBm <sup>3</sup> )	Mínima: 24,5 Máxima: 32,5 Incremento: 2	Mínima: 24,5 Máxima: 32,5 Incremento: 1
<b>Ambiente</b>	Distância entre leitor-etiqueta	0,40 m e 0,80 m (em linha reta)	Dez pontos de leitura, distribuídos conforme exibido na Figura 7.
	Material da parede	Local sala de aula: alvenaria <sup>4</sup> Local biblioteca: vidro	Local sala de aula: alvenaria Local biblioteca: vidro

<sup>1</sup> Velocidade média de uma pessoa caminhando [KNO96]. Esta velocidade foi verificada no início dos experimentos utilizando um cronômetro e uma trena.

<sup>2</sup> dBi: dB isotrópico, medida utilizada para expressar o ganho de uma antena comparado a uma antena isotrópica (que distribui energia uniformemente em todas as direções).

<sup>3</sup> dBm: dB miliwatt, medida mais comum para expressar a potência de um equipamento de transmissão de rádio.

<sup>4</sup> No caso se trata de uma construção de tijolos assentados com argamassa.

A variável “Sentido” se torna importante pelo fato de que, como explicado anteriormente, é comum o uso de crachás inteligentes para o registro de entrada e saída de pessoas em ambientes. Desta forma se faz necessário, em uma possível simulação de ambiente RFID, o projetista informar se o indivíduo está entrando ou saindo do ambiente. A Figura 5 ilustra as duas situações possíveis desta variável, para o valor “entrando” o indivíduo está se locomovendo do lado externo para o lado interno da sala, já no valor “saindo” ocorre o inverso. Cabe ressaltar que os sentidos “entrando” e “saindo” se aplicam tanto para o cenário móvel quanto para o cenário estacionário. No caso do cenário móvel este é o sentido em que a pessoa se movimenta (de fora para dentro ou de dentro para fora da sala). No cenário estacionário este é o sentido em que a pessoa está posicionada estaticamente, ou seja, no sentido “entrando” a pessoa está com a parte da frente do corpo (onde está o crachá) direcionada para o lado de dentro da sala, já no sentido “saindo” a pessoa está com a parte da frente do corpo direcionada para o lado externo da sala.

A informação de número de leituras para um indivíduo entrando na sala pode ser diferente do número de leituras obtido em um indivíduo que está saindo da sala, pois o corpo da pessoa portadora do crachá pode estar encobrindo o sinal de emissão da antena em pontos diferentes para cada opção da variável “Sentido”. Por exemplo, é possível que no sentido “entrando”, na posição X;Y(0,8; 0,8) da Figura 5, a comunicação entre leitora e etiqueta seja ruim, pois entre a antena e a etiqueta está o corpo do indivíduo portador do crachá. Já no sentido “saindo”, para a mesma posição, ocorre a visada direta entre a antena e a etiqueta, causando uma possível melhora na comunicação.

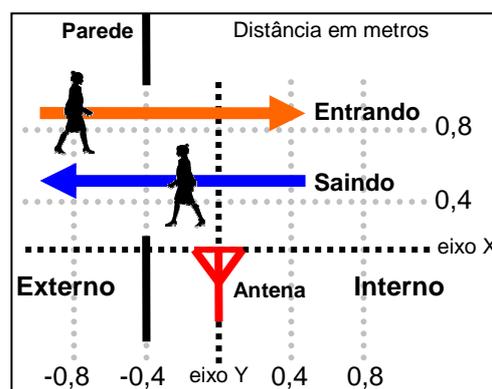


Figura 5 – Sentido que a pessoa está se locomovendo pelo ambiente.

A potência configurada no leitor é um importante fator a ser considerado no projeto de implantação de sistemas RFID. Conforme visto no Capítulo 2, uma etiqueta passiva recebe a energia proveniente do equipamento leitor, assim a intensidade desta energia

depende diretamente da configuração da potência do leitor. Desta forma, uma potência baixa não fornecerá energia suficiente para ativar a etiqueta, ou no caso de ser ativada, a etiqueta poderá não ter potência suficiente para retornar a informação ao leitor. Cabe ainda ressaltar que a potência máxima pode estar condicionada à legislação vigente do país ou região. Por exemplo, no Brasil a potência máxima de uma comunicação RFID não pode ultrapassar 36 dBm EIRP na frequência de 915 MHz, já em Israel, nesta mesma frequência, a potência máxima não pode ser maior que 33 dBm EIRP [EPC10]. Segundo [RID10], EIRP (*Effective isotropically radiated power*) é a potência efetiva irradiada isotropicamente<sup>1</sup>, ou seja, é a potência transmitida pelo leitor menos as perdas dos cabos mais o ganho da antena. No caso dos experimentos realizados, a antena utilizada possuía ganho de 8 dBi [MTI10] e a atenuação dos cabos era de aproximadamente 4,5 dBm<sup>2</sup>. Desta forma, a potência máxima configurada no leitor foi de 32,5 dBm ( $32,5 - 4,5 + 8 = 36$ ).

Conforme pode ser visualizado na Tabela 4, no cenário móvel a potência foi configurada entre 24,5 dBm e 32,5 dBm, com incremento de 2 dBm, ou seja, haverá medição nas potências 24,5, 26,5, 28,5, 30,5 e 32,5 dBm. Já no cenário estacionário a potência teve incremento de 1 dBm dentro dos mesmos limites utilizados no cenário móvel. No cenário estacionário foi possível realizar a automatização via software do processo de reconfiguração da potência do leitor (detalhado na Seção 4.2). Assim, um número maior de experimentos para a variável potência não necessitou de um aumento no tempo e esforço despendido na realização dos experimentos deste cenário. Já no cenário móvel foi selecionado um incremento de 2 dBm, pois nos testes desse cenário não foi possível realizar a automatização do processo de alteração da potência do leitor (o software coletor não conseguiria saber, sem intervenção humana, se o indivíduo já havia finalizado a movimentação no ambiente). Assim, o tempo e esforço despendido na coleta de dados do cenário móvel aumentariam caso fossem realizados experimentos com um incremento de potência menor.

Na variável “Distância” se tem situações distintas para os dois cenários projetados. No cenário móvel a etiqueta está em movimento, desta forma a variável no modelo está considerando a distância no momento em que a pessoa cruza a frente da antena, traçando uma linha reta imaginária entre o leitor e a etiqueta. Cabe enfatizar que a leitura está sempre ativa dentro de um perímetro de 3 metros, e a pessoa sempre se move de modo retilíneo. Essa situação é ilustrada na Figura 6.

---

<sup>1</sup> Isotropia: que possui as mesmas propriedades físicas independentemente da direção considerada.

<sup>2</sup> Segundo [POL10], o cabo (modelo RG-58) possui atenuação de 0,65 dB por metro na faixa de frequência de 900 MHz. Foi utilizado um cabo de aproximadamente 7 metros.

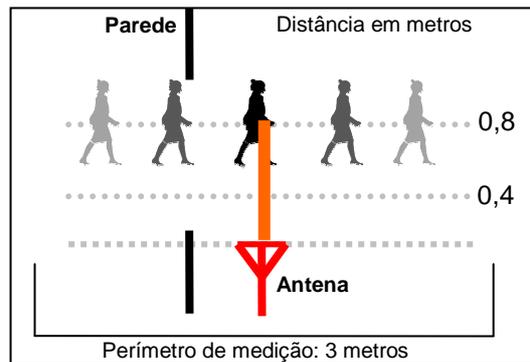


Figura 6 – Experimento no cenário móvel com distância de 0,8m.

No cenário estacionário a etiqueta tem sua posição estática definida em 10 pontos do ambiente, conforme é ilustrado na Figura 7. Neste caso a variável distância pode ser representada utilizando a informação da distância X e distância Y de cada ponto no plano 2D.

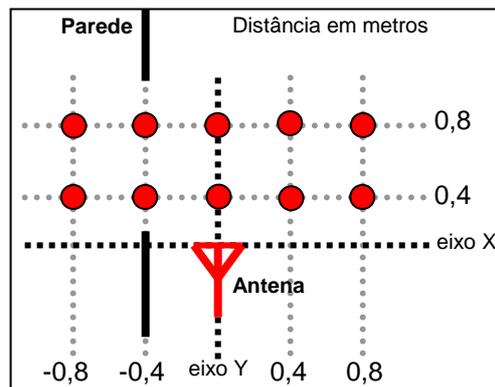


Figura 7 – Plano 2D do local de realização dos ensaios. As elipses representam a posição do indivíduo no sistema estacionário.

A variável “Material da parede” tem como objetivo representar o material de composição da parede dos dois locais em que serão realizados experimentos: sala de aula, que tem uma parede de alvenaria, e a biblioteca, onde a parede é de vidro. A propagação de ondas de rádio em ambientes fechados depende de maneira incisiva de qualquer objeto eletricamente condutor cujas dimensões físicas sejam maiores que  $\frac{1}{4}$  de comprimento de onda. Desta forma, a propagação está relacionada a parâmetros como a geometria das salas e os materiais das paredes e dos móveis. As propriedades dielétricas dos materiais de uma sala determinam as características de transmissão e reflexão, conseqüentemente afetando a propagação de ondas de rádio dentro de ambientes fechados. Materiais diferentes possuem coeficientes de transmissão e reflexão distintos. Deste modo, uma propagação de onda de rádio através de uma parede metálica será

mais atenuada do que uma onda passando em uma parede de tijolos [RAP94][SCH05][ZHA03].

Tendo todas as variáveis definidas, pode-se concluir que no cenário móvel temos um total de quatro variáveis, possuindo a variável “Sentido” dois valores possíveis, a variável “Potência” cinco valores, a variável “Distância” dois valores e a variável “Material da parede” duas opções possíveis. Desta forma se tem 40 situações distintas ( $2 \times 5 \times 2 \times 2 = 40$ ). A realização de experimentos para a coleta dos dados, utilizados no modelo estatístico, deve contemplar todas as 40 situações, sendo que para uma melhor precisão nos resultados, para cada situação devem ser realizadas cinco repetições de experimentos. Desta forma a coleta de dados do cenário móvel necessita um total de 200 experimentos ( $40 \times 5 = 200$ ).

O cenário estacionário possui também quatro variáveis, tendo a variável “Sentido” dois valores, a variável “Potência” nove valores, a variável “Distância” dez posições possíveis e a variável “Material da parede” duas opções. Desta forma se tem 360 situações distintas ( $2 \times 9 \times 10 \times 2 = 360$ ). Para cada situação possível também será executado cinco repetições de experimentos para a coleta de dados, resultando em um total de 1800 experimentos ( $360 \times 5 = 1800$ ) no cenário estacionário.

No cenário móvel a leitura será feita dentro de um perímetro de 3 metros enquanto a pessoa estiver se locomovendo neste ambiente, já no cenário estacionário cada experimento irá ativar o leitor RFID pelo tempo de 5 segundos, registrando o número de leituras realizadas nesse período. Um perímetro ou tempo muito grande poderiam acentuar o esforço e tempo despendidos na realização dos experimentos reais, deste modo os limites de perímetro e tempo foram fixados a partir dos resultados de testes realizados preliminarmente. Nesses testes preliminares foi detectado que não seria possível a etiqueta ter leituras fora do perímetro de 3 metros. Também foi visto que o tempo de 5 segundos era suficiente para a detecção da etiqueta e coleta dos dados.

Tendo os fatores e variáveis apresentados, pode-se resumir que o objetivo do trabalho dentro destes cenários é a inferência da leitura de uma etiqueta, em um sistema que possua as mesmas características dos fatores fixos da Tabela 3, e tendo como itens de configuração do ambiente de simulação as variáveis “Sentido”, “Potência”, “Distância” e “Material da parede”. Entende-se por itens de configuração as características do sistema que se deseja ajustar para realizar a simulação e verificar o funcionamento do sistema RFID.

## 4.2 Software de coleta dos dados

Para realizar a leitura das etiquetas foi desenvolvido um software na linguagem Java. Este software conecta-se no leitor Mercury4 e envia um comando para efetuar a leitura das etiquetas, armazenando as informações resultantes de cada execução. O comando é executado durante o tempo de 5 segundos para o cenário estacionário. No cenário móvel o usuário observa a movimentação do indivíduo portador da etiqueta, ativa a leitura quando ele entra no perímetro pré-definido (Seção 4.1), e desativa a leitura quando o indivíduo sai do perímetro.

Todas as saídas geradas pelos comandos de disparo e leitura da etiqueta são armazenadas em um banco de dados, facilitando assim a consulta posterior a estes registros. No registro do banco de dados são armazenadas as seguintes informações: local dos testes, sentido, distância, potência, frequência, dados armazenados na etiqueta (código EPC<sup>1</sup>), data e horário da leitura.

Independente de ocorrer ou não a leitura da etiqueta durante determinado teste, haverá registro do disparo de leitura. No caso de um teste sem leituras o campo código EPC ficará vazio. Isto permite a consulta de todos os testes realizados, mesmo aqueles em que não houve registro de comunicação entre o leitor e a etiqueta.

Visando aperfeiçoar o processo de realização dos ensaios, o software foi codificado de maneira que a potência do leitor (uma das variáveis do modelo) fosse incrementada automaticamente (sem intervenção humana), tornando os testes mais ágeis e com menor possibilidade de erros no registro dos resultados. O algoritmo inicia o processamento com a potência inicial informada pelo usuário, realiza a leitura, e passa para a próxima potência, executando este procedimento continuamente até chegar à potência final desejada.

A maioria dos leitores possui algumas configurações lógicas que dizem respeito ao protocolo anti-colisão utilizado no sistema RFID. Conforme visto na Seção 2.3, o protocolo utilizado neste trabalho foi o Random Slotted (Algoritmo Q). No software desenvolvido, o protocolo anti-colisão foi configurado de maneira que possibilitasse um maior desempenho na leitura de uma única etiqueta, ou seja, maior número de leituras possível em menor tempo. Para isto, o parâmetro Q (representado no leitor Mercury4 como Gen2InitQ) do algoritmo Random Slotted foi configurado com o valor zero. Conforme pode ser verificado na Seção 2.3, tal configuração faz com que a etiqueta sempre obtenha o

---

<sup>1</sup> Código Eletrônico de Produto (EPC - *Electronic Product Code*): trata-se de um valor único para identificação de produtos com etiquetas RFID.

valor zero ( $2^0 - 1 = 0$ ) no processo de sorteio do algoritmo, ocasionando uma resposta imediata ao leitor.

O software de coleta é executado em modo terminal (console) e os argumentos que devem ser informados na execução são: (i) identificação do teste (configuração atual do cenário); (ii) potência inicial; (ii) potência final; (iii) incremento da potência, padrão 1 dBm; (iv) tempo desejado de leitura para cada potência, padrão 5 segundos.

### **4.3 Análise estatística dos dados**

A execução dos ensaios nos dois cenários projetados resultou em uma base de dados com informações sobre o número de leituras realizadas em cada experimento.

Acerca da faixa de leituras obtida em cada um dos cenários, o cenário móvel se manteve entre 0 e 25 leituras, já o cenário estacionário ficou na faixa de 0 à 96 leituras. O fato de o leitor RFID permanecer com a comunicação ativa durante 5 segundos no cenário estacionário fez com que existisse um número maior de leituras para este caso. Em uma breve análise dos dados percebeu-se que as configurações de ambiente que propiciavam uma melhor comunicação entre os dispositivos possuíam um maior número de leituras, já as configurações ruins (como um cenário com baixa potência do leitor) geralmente resultaram em um baixo número de leituras. Isto demonstra que o número de leituras é uma variável que pode ser utilizada na predição do funcionamento de sistemas RFID, podendo indicar os locais e configurações mais propícias na implantação de um sistema deste tipo.

Analisando estatisticamente os dados coletados, logo se descobriu que, mesmo com o isolamento de várias características que poderiam influenciar na leitura das etiquetas (tais como posicionamento exato dos componentes, retirada dos móveis ou outros objetos interferentes do local), o número de leituras para uma determinada situação teve variância entre uma repetição e outra. Conforme pode ser visto na Tabela 5, em determinada situação do cenário móvel, a primeira repetição com potência 26,5 dBm teve 14 leituras, já a segunda repetição da mesma potência resultou em 3 leituras.

Tabela 5 – Dados coletados para a situação “material: vidro, sentido: entrando, distância: 0,8” do cenário móvel

Potência	Repetição	Leituras	Potência	Repetição	Leituras
24,5	1	9	30,5	1	19
	2	0		2	18
	3	0		3	9
	4	0		4	11
	5	0		5	13
26,5	1	14	32,5	1	21
	2	3		2	15
	3	1		3	19
	4	3		4	17
	5	5		5	15
28,5	1	19			
	2	10			
	3	11			
	4	8			
	5	14			

Esta distribuição do número de leituras se deve principalmente a dois fatores: (i) a alta sensibilidade dos dispositivos do sistema RFID; (ii) variáveis ainda desconhecidas que podem influenciar no funcionamento do sistema. Quanto ao primeiro fator, sempre se tentou posicionar os dispositivos de maneira mais exata possível dentro do cenário, por exemplo, posicionando a antena do leitor sempre no mesmo local e altura em relação ao solo, ajustando a etiqueta sempre na distância correta e posicionando o crachá na mesma posição e local do corpo do indivíduo. Entretanto, tudo isto nem sempre é praticável, visto que, por exemplo, a alteração de um milímetro na posição ou na orientação da etiqueta já causava uma enorme diferença no número de leituras.

O gráfico da Figura 8 exibe um exemplo de variância do número de leituras entre as cinco repetições executadas (cenário móvel), este exemplo utiliza os dados da Tabela 5 para ilustrar o problema. O gráfico exibido é do tipo *boxplot*, onde a linha central de cada caixa marca a mediana do conjunto de dados, a parte inferior da caixa é delimitada pelo quartil inferior ( $Q_{1/4}$ , 1º quartil) e a parte superior pela quartil superior ( $Q_{3/4}$ , 3º quartil). As hastes inferiores e superiores se estendem, respectivamente, do quartil inferior até o menor valor não inferior a  $Q_{1/4} - 1,5*IEQ$  e do quartil superior até o maior valor não superior a  $Q_{3/4} + 1,5*IEQ$ . O intervalo interquartil (*IEQ*) é o resultado da diferença entre  $Q_{3/4}$  e  $Q_{1/4}$  [SIL10].

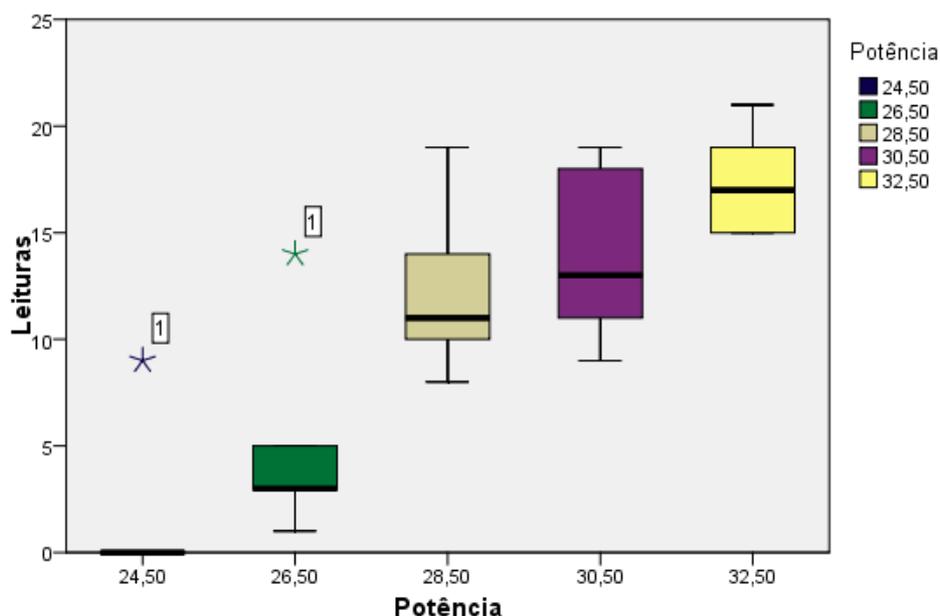


Figura 8 – Exemplo de variância das leituras nas repetições dos experimentos para determinada situação (cenário móvel)

No gráfico *boxplot* da Figura 8, os valores que se encontram além das hastes inferiores ou superiores são destacados com o número da repetição ao lado de um asterisco. Observa-se que, nas potências 24,50 e 26,50 dBm, a primeira repetição (número 1 ao lado do asterisco) obteve um número de leituras completamente diferente das demais repetições. Por exemplo, na potência 26,50 dBm, a primeira repetição registrou o valor de 15 leituras e por estar além da haste superior se encontra destacada com um asterisco. O conjunto de dados das demais repetições está sendo representado pelas hastes e a caixa do gráfico, onde se obteve uma mediana e quartil inferior próximo de 4, quartil superior em 5 e o valor representado pela haste inferior foi próximo de 1. Isto demonstra a variância existente entre as repetições de experimentos onde todos os elementos conhecidos estão fixos.

#### 4.4 Técnicas de eliminação de dados espúrios

Tendo em vista a variância no número de leituras dos experimentos, decidiu-se buscar algum método de exclusão das amostras que continham erros grosseiros de medição.

Segundo [NAR10b], erros grosseiros são aqueles erros que comprovadamente foram causados por imprevistos, fontes involuntárias de erros, problemas em instrumentos de medição, operação, ou do operador que realiza o experimento, ou seja,

todo fator externo, estranho ao experimento, que interfira diretamente no resultado da medição.

Com o intuito de eliminar esses erros grosseiros de medição, foi aplicado o Método dos Quartis na base de dados dos experimentos realizados para ambos os cenários. O Método dos Quartis consiste em analisar o quão distantes estão os extremos da base de dados (mínimo e máximo) do 1º e do 3º quartil [NAR10b]. Uma amostra é considerada um dado espúrio, e deve ser removida da base de dados, se:

$$\begin{aligned} x &> Q_{\frac{3}{4}} + 3,0 \cdot IEQ \quad \text{ou;} \\ x &< Q_{\frac{1}{4}} - 3,0 \cdot IEQ \end{aligned} \quad (1)$$

Uma amostra também pode ser considerada um dado espúrio em potencial, cabendo ao operador a decisão de eliminá-la. Tal fato acontece se:

$$\begin{aligned} Q_{\frac{3}{4}} + 1,5 \cdot IEQ < x < Q_{\frac{3}{4}} + 3,0 \cdot IEQ \quad \text{ou;} \\ Q_{\frac{1}{4}} - 3,0 \cdot IEQ < x < Q_{\frac{1}{4}} - 1,5 \cdot IEQ \end{aligned} \quad (2)$$

O gráfico da Figura 8 ilustra de forma prática onde se encontram os quartis e os limites inferiores e superiores para considerar um dado como potencial espúrio. No gráfico *boxplot*, os pontos que estão além das hastes inferiores e superiores são, no mínimo, potenciais pontos a serem eliminados. Isto se deve ao fato da representação das hastes inferiores e superiores do gráfico utilizarem as mesmas restrições encontradas na Equação (2), que apresenta os limites para considerar um dado espúrio em potencial.

Foi criada uma planilha eletrônica para verificar em toda a base de dados a existência de dados espúrios e dados potencialmente espúrios. A verificação utilizando o Método dos Quartis se deu da seguinte forma:

1. Obteve-se o conjunto de dados das repetições realizadas em cada situação distinta de experimento.
2. Neste conjunto de dados foram calculados os quartis ( $Q_{1/4}$  e  $Q_{3/4}$ ) e o intervalo entre eles ( $IEQ$ ).
3. Foram calculados os valores críticos utilizando a Equação (1) e selecionadas para eliminação as amostras que atendiam o critério estabelecido.
4. Foram calculados os limites utilizando a Equação (2), verificando em seguida a existência de dados potencialmente espúrios. Conforme dito

anteriormente, neste caso a decisão de excluí-los depende da observação do conjunto pelo operador.

Para o cenário móvel, foram encontrados na base de dados seis registros considerados espúrios e três registros provavelmente espúrios. Analisando minuciosamente os valores decidiu-se por excluir todos os espúrios e manter na base dois dos três prováveis espúrios, resultando um total de 193 registros.

No cenário estacionário, foram encontrados na base de dados sete registros considerados espúrios e 37 registros com possibilidade de serem espúrios. Após analisar os valores decidiu-se pela eliminação de todos os registros espúrios e manteve-se na base apenas um dos 37 registros potencialmente espúrios. A base de dados final do cenário estacionário contou com um total de 1757 registros.

A decisão de eliminar os dados provavelmente espúrios baseou-se na diferença existente entre o número de leituras obtido nas repetições. Por exemplo, em diversos casos, somente uma das cinco repetições possuía valor superior a zero. Assim, repetições com este comportamento foram eliminadas da base.



## 5 MODELO DE PREDIÇÃO UTILIZANDO REGRESSÃO LINEAR

Com o objetivo de se obter um modelo de predição para o número de leituras utilizando a base de dados dos experimentos, partiu-se inicialmente para um dos modelos básicos de inferência estatística: a regressão linear.

### 5.1 Regressão Linear

A regressão linear é uma técnica que envolve a análise de dados amostrais para investigar e modelar a relação entre variáveis, sendo utilizada principalmente com o propósito de previsão. O modelo obtido através da regressão linear é denominado equação de regressão. Uma equação de regressão é uma expressão utilizada para estimar uma variável dependente  $y$  (resposta), em função das variáveis independentes (explanatórias),  $x_1, x_2, \dots$ , e é denominada *equação de regressão de  $y$  para  $x_1, x_2, \dots$* . Na notação de função, ela é escrita às vezes sob a forma abreviada como,  $y = F(x_1, x_2, \dots)$  que se lê “ $y$  é uma função de  $x_1, x_2$  etc.” [SPI09].

A equação de regressão é uma equação para a linha de regressão (ou linha de melhor ajuste) que melhor modela os dados na correlação entre variáveis. Considerando o gráfico de dispersão e a linha de regressão traçada na Figura 9. Para cada ponto dado,  $d_i$  representa a diferença entre o valor  $y$  observado e o valor  $y$  previsto para um dado valor  $x$  na linha. Essas diferenças são chamadas de resíduos. Se o valor  $y$  observado é igual ao valor  $y$  previsto, então  $d_i = 0$ . De todas as linhas possíveis que podem ser desenhadas através de um conjunto de dados, a linha de regressão é a linha para a qual a soma dos quadrados de todos os resíduos ( $\sum d_i^2$ ) é um mínimo [LAR10].

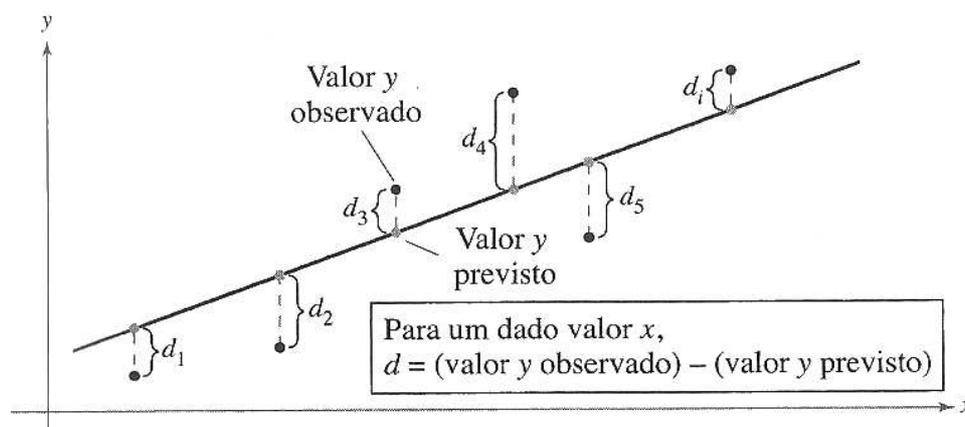


Figura 9 – Gráfico de dispersão e a linha de regressão linear [LAR10].

A equação de uma reta de regressão para uma variável independente  $x$  e uma variável dependente  $y$  é:

$$\hat{y} = mx + b \quad (3)$$

onde  $\hat{y}$  é o valor  $y$  previsto para um dado valor  $x$ . O coeficiente  $m$  (inclinação) e o coeficiente  $b$  (interseção) são dados por

$$m = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} \quad \text{e} \quad b = \bar{y} - m\bar{x} = \frac{\sum y}{n} - m \frac{\sum x}{n} \quad (4)$$

onde  $\bar{y}$  é a média dos valores  $y$  no conjunto de dados e  $\bar{x}$  é a média dos valores  $x$ . A linha de regressão sempre passa pelo ponto  $(\bar{x}, \bar{y})$  [LAR10].

Existem três tipos de variação sobre uma linha de regressão:

- Variação total: é a soma dos quadrados das diferenças entre o valor  $y$  e cada par pedido e a média de  $y$ .  $\text{Variação total} = \sum (y_i - \bar{y})^2$
- Variação explicada: é a soma dos quadrados das diferenças entre cada valor  $y$  previsto e a média de  $y$ .  $\text{Variação explicada} = \sum (\hat{y}_i - \bar{y})^2$
- Variação não explicada: é a soma dos quadrados das diferenças entre o valor  $y$  de cada par pedido e cada valor  $y$  previsto correspondente.  $\text{Variação não explicada} = \sum (y_i - \hat{y}_i)^2$

A soma das variações explicadas e não explicadas é igual a variação total. Como o nome indica, a “variação explicada” pode ser explicada pela relação entre  $x$  e  $y$ . A “variação não explicada” não pode ser explicada por uma relação entre  $x$  e  $y$  e isso é em razão do acaso ou outras variáveis [LAR10].

Uma das medidas de regressão mais encontradas na literatura é o coeficiente de determinação. O coeficiente de determinação, também chamado de  $R^2$ , é igual à relação da “variação explicada” com a “variação total”, podendo ser descrito como:

$$R^2 = \frac{\text{Variação explicada}}{\text{Variação total}} \quad (5)$$

O  $R^2$  representa o percentual de variação dos dados utilizados pelo modelo linear, ou seja, o percentual da variação de  $y$  que pode ser explicado pela relação entre  $x$  e  $y$ . Por exemplo, um  $R^2$  igual a 0,81 significa que 81% da variação de  $y$  podem ser explicados pela relação entre  $x$  e  $y$ ; 19% restante da variação não é explicada, em razão de outros

fatores ou a erro da amostra. Um  $R^2$  de 1,0 (100%) indica que a linha de regressão cruzou perfeitamente o conjunto de dados amostrados [LAR10].

Segundo [LAR10], quando um valor  $\hat{y}$  é previsto de um valor  $x$ , a previsão é um ponto estimado. O erro padrão da estimativa ( $S_e$ ) é o desvio padrão dos valores  $y_i$ , observados sobre o valor  $y_i$  previsto para um dado valor  $x_i$ . O  $S_e$  pode ser descrito como

$$S_e = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2}} \quad (6)$$

onde  $n$  é o número de pares pedidos no conjunto de dados. O resultado do erro padrão da estimativa é informado utilizando a mesma escala dos dados originais, sendo de mais fácil entendimento se comparado ao  $R^2$ .

Em muitos casos, um modelo de previsão necessita mais do que uma variável independente. A quantidade de variáveis experimentais medidas distingue os dois métodos de regressão existentes: a regressão linear simples e a regressão linear múltipla. Para uma regressão simples há somente uma variável independente  $x$  e uma dependente  $y$ . A regressão linear múltipla acontece quando o modelo regredido conta com uma variável dependente  $y$  e várias variáveis independentes  $x$ . A equação de regressão múltipla pode ser descrita como:

$$\hat{y} = b + m_1x_1 + m_2x_2 + m_3x_3 + \dots + m_kx_k \quad (7)$$

onde  $x_1, x_2, x_3, \dots, x_k$  são variáveis independentes,  $b$  é o coeficiente de interseção e  $\hat{y}$  é a variável dependente. Cada coeficiente  $m_i$  é a quantidade de mudança em  $y$  quando a variável independente  $x_i$  é alterada por uma unidade e todas as outras variáveis independentes são mantidas constantes [LAR10][NAR10a].

Devido à complexidade matemática do processo de regressão linear múltipla, o desenvolvimento desses modelos geralmente utiliza algum software de apoio que automatize a regressão, obtendo assim os coeficientes e as métricas de verificação, como o  $R^2$  e o erro padrão da estimativa.

## 5.2 Modelo

Com o objetivo de analisar os dados obtidos nos experimentos reais e efetuar a regressão linear, optou-se por buscar soluções que pudessem automatizar o desenvolvimento do modelo de regressão. Existem atualmente diversos softwares para análise estatística e processamento de regressão linear, tais como o Microsoft Excel [MIC10] e o IBM SPSS Statistics [SPS10a]. Devido à praticidade e robustez do software

SPSS, decidiu-se por utilizar este aplicativo no manuseio e análise dos dados coletados nos experimentos deste trabalho.

Ao inserir os dados no software SPSS, é necessário classificá-los de acordo com os tipos de dados definidos neste software, que são: nominal e escala. O SPSS define uma variável como nominal quando seu valor apresenta categorias que não possuem uma classificação ou ordenação intrínseca. Exemplo: o departamento de uma empresa, regiões de um país, tipos de ambiente de uma universidade (laboratório, sala de aula, biblioteca, auditório, etc). Já uma variável é do tipo escala quando existe algum tipo de métrica associada a ela, sendo possível obter valores que estejam dentro de um intervalo. Por exemplo, a idade de uma pessoa é uma variável do tipo escala, pois existe a métrica (anos) e também é possível obter os valores existentes entre duas idades. Outro exemplo de variável escala é a potência de um leitor RFID, onde a métrica utilizada é o dBm [SPS10b].

Na regressão linear não é possível trabalhar com valores nominais, ou seja, valores que não possuem uma escala numérica definida. Nos cenários especificados neste trabalho existem duas variáveis ditas nominais: material e sentido. Sendo assim, o modelo de regressão, obtido através de uma regressão linear múltipla, somente será aplicado para as variáveis potência e distância, sendo que cada combinação de material e sentido terão coeficientes diferentes no modelo, ou seja, haverá uma equação distinta para cada situação de material e sentido. Baseado na Equação (7) de regressão linear múltipla, foi desenvolvido o modelo para a inferência de leituras no cenário móvel:

$$L(p, d) = m_{pi} \cdot p + m_{di} \cdot d + b_i \quad (8)$$

onde  $L$  é o número de leituras inferido,  $p$  é a potência em dBm e  $d$  é a distância em metros entre o leitor e a etiqueta. Na Equação (8), para cada situação ( $i$ ) de material e sentido, é necessário um coeficiente distinto que multiplica a potência ( $m_{pi}$ ), outro para multiplicar a distância ( $m_{di}$ ) e por fim o coeficiente de ajuste da regressão linear ( $b_i$ ).

No cenário estacionário a função de inferência para leituras pode ser descrita como:

$$L(p, x, y) = m_{pi} \cdot p + m_{xi} \cdot x + m_{yi} \cdot y + b_i \quad (9)$$

onde  $L$  é o número de leituras inferido,  $p$  é a potência em dBm,  $x$  é a distância horizontal (eixo X do plano 2D) em metros entre o leitor e a etiqueta e  $y$  é a distância vertical (eixo Y

no plano 2D) em metros. Da mesma forma que a Equação (8), para cada situação ( $i$ ) de material e sentido é necessário um coeficiente distinto para multiplicar a potência ( $m_{pi}$ ) e as distâncias ( $m_{xi}$ ,  $m_{yi}$ ), bem como o coeficiente de ajuste da regressão linear ( $b_i$ ).

Para utilizar a base de dados dos experimentos na regressão linear, inicialmente foi calculada a média das leituras para cada situação distinta de material, sentido, distância e potência, visto que cada situação possui cinco repetições cada. A base de dados também foi dividida em oito bases de dados menores, uma para cada situação de material e sentido, sendo desta forma executada a regressão linear múltipla para cada uma destas bases de dados.

Tendo as bases de dados separadas e as variáveis de ambos os modelos selecionadas, foi utilizado o software SPSS para a execução da regressão linear múltipla. Utilizou-se o recurso de regressão linear configurando as variáveis potência e distância como independentes e o número de leituras como a variável dependente do modelo. Desta forma, o SPSS executa a regressão e exibe um relatório contendo o  $R^2$ , o erro padrão e os coeficientes da fórmula básica de inferência, que posteriormente podem ser aplicados nas Equações (8) e (9) desenvolvidas.

### 5.3 Resultados

Os resultados do  $R^2$  e do erro padrão para cada situação de material e sentido são apresentados na Tabela 6.

Tabela 6 – Resultados do  $R^2$  e do erro padrão na regressão linear múltipla

Cenário	Material	Sentido	$R^2$	Erro padrão
Móvel	Alvenaria	Entrando	0,831	2,506
		Saindo	0,919	1,621
	Vidro	Entrando	0,727	3,112
		Saindo	0,738	3,582
Estacionário	Alvenaria	Entrando	0,297	21,166
		Saindo	0,413	22,565
	Vidro	Entrando	0,263	28,201
		Saindo	0,368	21,636

Como pode ser visto na Tabela 6, o cenário móvel obteve um  $R^2$  entre 0,7 e 0,92, enquanto o cenário estacionário teve um  $R^2$  muito abaixo de 1, que é o valor ideal para esta métrica (Seção 5.1). Analisando a métrica do erro padrão da estimativa, o cenário

móvel também obteve um desempenho superior ao cenário estacionário. O erro padrão do cenário móvel se manteve entre 1 e 4 leituras, enquanto o cenário móvel teve um erro superior a 20 leituras, sendo este um valor representativo na predição do funcionamento do sistema. Os resultados do cenário estacionário demonstram uma não-linearidade nos dados coletados para este cenário, não sendo o método de regressão linear passível de obter uma boa predição utilizando tais amostras.

A partir dos valores de  $R^2$  e de erro padrão obtidos no processo de regressão linear múltipla, decidiu-se demonstrar os resultados do modelo de regressão apenas para o cenário móvel, deixando a análise do cenário estacionário para outro método de inferência que possa tratar dados não-lineares.

Sendo encontrados os coeficientes através da regressão linear múltipla executada no SPSS, estes puderam ser aplicados na Equação (8) do cenário móvel, conforme exibido na Tabela 7.

Tabela 7 – Coeficientes e modelos de predição para cada situação de material e sentido do cenário móvel

Material	Sentido	Coeficientes			Equação
		Potência $p$	Distância $d$	Constante $b$	
Alvenaria	Entrando	1,625	3,5	-41,113	$L(p, d) = 1,625.p + 3,5.d - 41,113$
	Saindo	1,4	-11,5	-25,7	$L(p, d) = 1,4.p - 11,5.d - 25,7$
Vidro	Entrando	1,45	-5,5	-27,725	$L(p, d) = 1,45.p - 5,5.d - 27,725$
	Saindo	1,75	4,5	-41,475	$L(p, d) = 1,75.p + 4,5.d - 41,475$

A fim de exibir os resultados completos da intensidade de leituras no ambiente, foi realizada a predição de leituras para o cenário móvel utilizando as equações da Tabela 7. Os modelos de predição foram aplicados em uma base de dados de teste possuindo uma resolução de distância de 0,1 m e intervalo de potência de 1 dBm. A distância foi limitada entre 0,1 e 1 m e a potência entre 24,5 e 32,5 dBm. Os gráficos da Figura 10 exibem os resultados do número de leituras inferido pelos modelos.

Os gráficos da Figura 10 apresentam, para cada situação de material e sentido, a predição do número de leituras obtido através dos modelos da Tabela 7, aplicados à base de testes criada. Os valores de potência e distância informados nos modelos são representados através do eixo X e Y do gráfico, respectivamente. A temperatura (cores mais claras ou mais escuras) do interior dos gráficos representa o número de leituras inferido pelos modelos, sendo que quanto mais escura a cor, maior é o número de leituras.

Como exibido na legenda dos gráficos, o número de leituras inferido pelos modelos para a base de teste resultou valores entre -5 e 20. Apesar de não fazer sentido existir números de leituras negativos, estes podem estar presentes nos resultados, pois os modelos foram desenvolvidos utilizando métodos estatísticos de aproximação linear. Neste caso, os valores negativos devem ser interpretados como zero (nenhuma leitura).

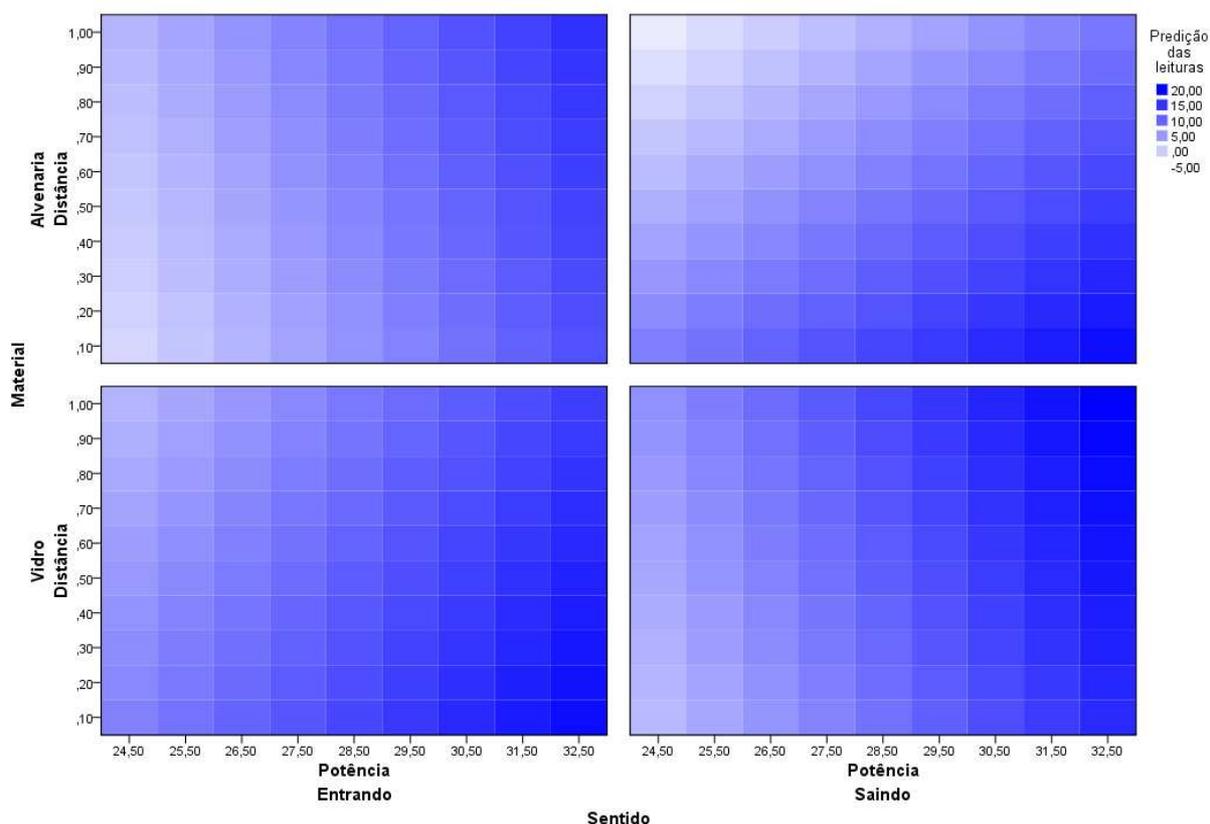


Figura 10 – Resultados dos modelos desenvolvidos utilizando regressão linear (cenário móvel)

Através dos gráficos da Figura 10 é possível visualizar, em cada situação de material, sentido e potência, o impacto no número de leituras em função da distância existente entre o leitor e a etiqueta. Por exemplo, é possível verificar que na situação material “vidro”, sentido “entrando”, potência 28,5 dBm; existe uma diminuição no número de leituras à medida que a distância aumenta.

Analisando os gráficos, percebe-se que, em certas condições de material e sentido, o número de leituras inferido não diminui conforme a distância é incrementada. Os gráficos das situações alvenaria/entrando e vidro/saindo demonstram que, em algumas potências, existe um incremento no número de leituras inferido conforme a distância aumenta (o que não era esperado, seguindo o princípio básico de funcionamento de

sistemas RFID). Apesar de incomum, isto pode ter acontecido devido ao fato da etiqueta, em distâncias longas, ter permanecido exposta por mais tempo dentro da área de cobertura da antena do leitor. Já nos experimentos de 0,4 m, o indivíduo se locomoveu em frente à antena dentro de uma área de abrangência mais curta, ocasionando um menor número de leituras. Segundo o manual [MTI10] da antena utilizada nos experimentos, a irradiação do campo eletromagnético é menor no ângulo mais próximo de  $-90^\circ$  e  $90^\circ$  da antena. O diagrama de radiação da antena pode ser visualizado no Anexo A e o cenário descrito é ilustrado na Figura 11.

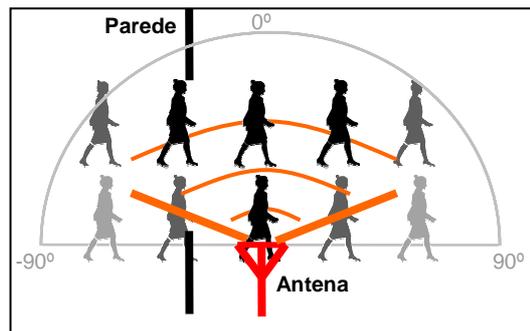


Figura 11 – Relação entre a distância e o ângulo de cobertura da antena. Distâncias mais próximas da antena possuem um ângulo menor de cobertura.

Os gráficos também permitem a análise do comportamento do número de leituras em relação à potência configurada no leitor. Neste caso, é visível que em todas as situações de material, sentido e distância, o número de leituras aumenta à medida que a potência é incrementada, o que era esperado conforme o princípio de funcionamento de sistemas RFID.

Na análise do sistema em relação aos materiais da parede, a principal diferença visível é a de que um material sempre teve comportamento oposto ao do outro. Por exemplo, no sentido “entrando”, para qualquer potência, o material “vidro” tem um decréscimo de leituras conforme a distância aumenta. Já no material “alvenaria”, para o mesmo sentido de locomoção, acontece um crescimento no número de leituras conforme a distância é incrementada.

#### 5.4 Validação

Para a validação do modelo foram realizados novos experimentos com o intuito de avaliar os resultados inferidos pelas equações da Tabela 7 sobre a base de testes. Foram realizadas duas fases de experimentos. Na primeira fase, os experimentos de validação foram realizados em um ambiente com características semelhantes ao local dos

experimentos originais. Na segunda fase, os experimentos de validação utilizaram valores novos de distância e potência, ou seja, valores diferentes dos utilizados nos experimentos iniciais (Seção 4.1).

No projeto dos experimentos da primeira fase, devido à dificuldade em encontrar um local mais apropriado, decidiu-se por simplesmente inverter a posição da antena do leitor para o outro lado da porta. Como explicado na Seção 2.4.1, a precisão dos modelos estatísticos depende inclusive das semelhanças entre o ambiente simulado e o ambiente onde os experimentos foram realizados. Apesar da modificação ter sido simples, o fato de inverter o posicionamento da antena já é um fator importante para os testes de um modelo estatístico desse gênero, considerando o nível de sensibilidade e interferências que atuam os sistemas RFID. Por problemas de tempo e logística na execução dos novos experimentos, esta fase da validação foi realizada somente para o material “alvenaria”, efetuando-se três repetições de experimentos e obtendo-se a média de leituras para cada situação. Assim, a validação do material “vidro” foi realizada somente na segunda fase.

Na segunda fase da validação foram utilizados os mesmos ambientes dos experimentos iniciais, porém com valores de distância e potência nunca antes utilizados. Os dados coletados na segunda fase incluíram a distância 0,6 m e os valores de potência 25,5, 27,5, 29,5 e 30,5 dBm.

Foi obtido um total de 36 amostras para validação, sendo 20 na primeira fase e 16 na fase dois. Tendo os novos experimentos realizados e os dados coletados, foi realizada a comparação entre o número de leituras real e o número de leituras inferido pelo modelo. Alguns dos resultados são apresentados de forma textual na Tabela 8. Os resultados contemplam tanto a primeira quanto a segunda fase da validação.

Na análise dos resultados da validação, deve se levar em consideração o fato de que não é possível realizar a comparação direta entre os valores reais e os valores inferidos para assumir um acerto ou erro do modelo, pois já foi visto que, para uma mesma situação, o número de leituras varia de uma repetição para outra.

Em publicações acerca de modelos para predição de sinal RF, onde a intensidade do campo em dBm é comparada com o valor inferido, sempre são apresentadas a média e o desvio padrão dos erros da validação [NEŠ00]. Devido às semelhanças do modelo desenvolvido com os modelos RF, decidiu-se utilizar a mesma abordagem para apresentar a precisão deste modelo.

Tabela 8 – Amostras da validação para o modelo de regressão linear do cenário móvel

Material	Sentido	Distância (m)	Potência (dBm)	Leituras (real)	Leituras (inferência)
Alvenaria	Entrando	0,4	24,5	4	0
Alvenaria	Entrando	0,4	26,5	5	3
Alvenaria	Entrando	0,4	28,5	6	7
Alvenaria	Entrando	0,4	30,5	6	10
Alvenaria	Entrando	0,4	32,5	6	13
Alvenaria	Entrando	0,6	25,5	6	2
Alvenaria	Entrando	0,6	27,5	5	6
Alvenaria	Entrando	0,6	29,5	12	9
Alvenaria	Entrando	0,6	31,5	15	12
Alvenaria	Entrando	0,8	24,5	1	1
Alvenaria	Entrando	0,8	26,5	5	5
Alvenaria	Entrando	0,8	28,5	8	8
Alvenaria	Entrando	0,8	30,5	11	11
Alvenaria	Entrando	0,8	32,5	12	14
...	...	...	...	...	...

Para se obter a média e o desvio padrão dos erros, foram calculados os erros absolutos de cada teste realizado. O erro absoluto foi obtido através da fórmula  $e = |v - v_{aprox}|$ , sendo  $v$  o número de leituras reais do experimento de validação e  $v_{aprox}$  o valor inferido pelo modelo. Toma-se como exemplo o cálculo do erro absoluto no primeiro resultado da Tabela 8:  $e = |4 - 0|$ ;  $e = 4$ .

O gráfico da Figura 12 exibe a relação entre os valores inferidos e os valores medidos nos experimentos de validação. Neste gráfico, cada elipse está posicionada no eixo Y conforme o valor da leitura inferida pelo modelo e no eixo X conforme o valor medido no experimento de validação. Quanto mais distante da linha diagonal, maior é a diferença entre o valor medido e o valor inferido. No gráfico, a elipse tem um tamanho maior à medida que o erro absoluto também aumenta.

A partir da base de dados contendo os valores medidos e inferidos para cada um dos testes de validação, foi obtida a média e o desvio padrão dos erros absolutos. Os resultados de média e desvio padrão também foram separados por material e sentido. Os resultados são exibidos na Tabela 9.

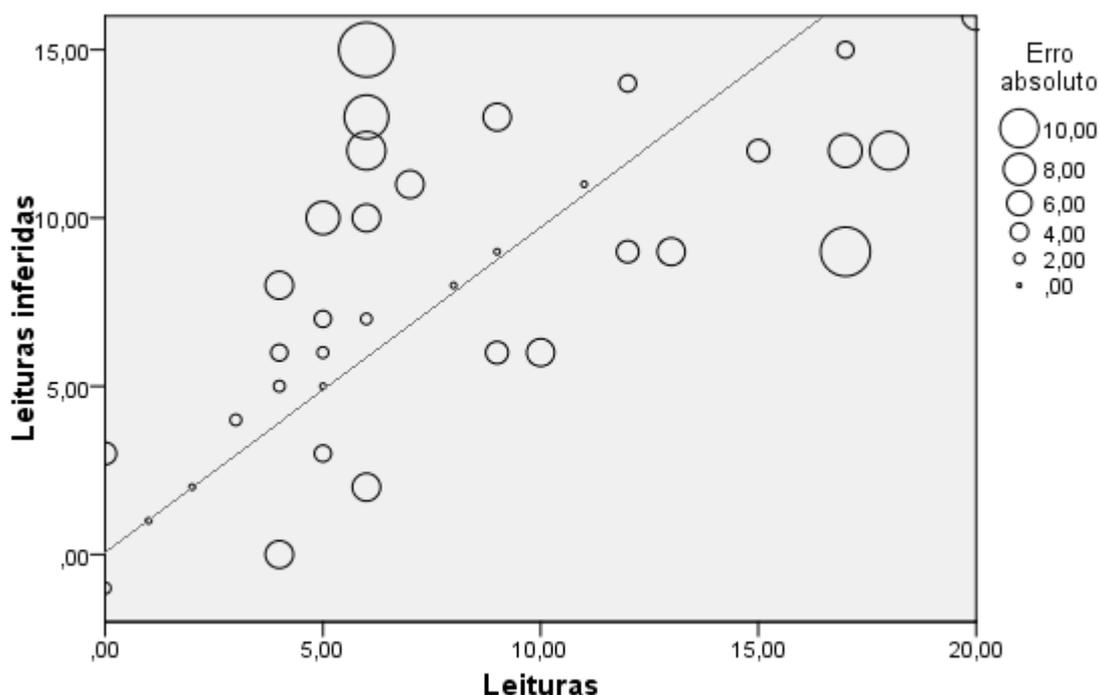


Figura 12 – Relação entre número de leituras medido e inferido na validação do modelo linear (cenário móvel). As elipses representam o erro absoluto de cada teste.

Tabela 9 – Média e desvio padrão dos erros absolutos (cenário móvel)

Material	Sentido			
	Entrando		Saindo	
	Média	Desvio padrão	Média	Desvio padrão
Alvenaria	2,2	2,1	3,5	2,4
Vidro	4,8	2,8	2,5	1,9
<b>Média geral</b>	3,0			
<b>Desvio padrão geral</b>	2,3			

Não é de conhecimento a existência de um parâmetro na literatura de referência para comparação dos resultados da validação do modelo desenvolvido, especialmente sob as condições apresentadas nos cenários deste trabalho. Devido a isto, cabe uma análise e interpretação dos resultados aqui obtidos. A análise se deu relacionando o desvio padrão dos erros com a faixa de leituras que se obteve nos experimentos iniciais. O número de leituras obtido nos experimentos do cenário móvel ficou entre 0 e 25 (Seção 4.3). Assim, o desvio padrão geral de 2,3 representa apenas 9% do valor máximo de leituras, o que é um percentual baixo e que pode representar pouco impacto na predição do funcionamento do sistema.

Apesar do baixo percentual obtido, é notável na Tabela 8 e no gráfico da Figura 12 que algumas predições tiveram pontos com erro absoluto distantes de zero, quando o ideal seria um valor inferido estar mais próximo do valor que foi medido no experimento real. Tal diferença se deve principalmente ao fato da variância existente entre uma repetição e outra do experimento real, que é afetado diretamente pela sensibilidade dos equipamentos e da comunicação, bem como das variáveis não consideradas no modelo desenvolvido (como as relatadas no final da Seção 4.1).

## 6 MODELO DE PREDIÇÃO UTILIZANDO REDES NEURAS ARTIFICIAIS

A predição do funcionamento de um sistema RFID tem como objetivo verificar se os componentes do sistema se comunicam ou não adequadamente. Se cada leitura de etiqueta RFID é uma comunicação completa do sistema, o número de leituras que determinada configuração do ambiente proporcionou é uma informação bastante valiosa para ser utilizada em uma posterior simulação. Por exemplo, em determinada configuração de sentido e distância, ao diminuir a potência do leitor em 2 dBm, o número de leituras diminui em 5 unidades. Já quando é incrementada a potência do leitor em 2 dBm, o número de leituras aumenta cerca de 7 unidades. Além da variável potência, uma alteração nas outras variáveis (sentido, material da parede e distância) também pode causar algum efeito no número de leituras, sendo que tais alterações podem gerar um padrão de resultados para cada configuração do cenário.

A aplicação de um modelo estatístico de inferência, baseado no reconhecimento de padrões, pode obter um bom resultado para as configurações conhecidas e até para configurações desconhecidas (mas semelhantes àquelas que o modelo já conhece).

Conforme [NEŠ00], as Redes Neurais Artificiais com retro-propagação têm mostrado um bom desempenho para resolver problemas em conjunto de dados não lineares com alternância de valores. Uma característica das Redes Neurais é o paralelismo intrínseco existente, o que permite uma rápida avaliação das soluções. Em uma RNA, o processo de aprendizado pode levar algumas horas, mas o processo de predição dos resultados é extremamente rápido.

Segundo [BAR97], o reconhecimento de padrões é talvez uma das primeiras aplicações de Redes Neurais. Um dos modelos de RNA foi concebido principalmente como instrumento capaz de reconhecer letras. A principal razão é que o reconhecimento de padrões é uma tarefa geralmente desempenhada muito melhor usando as capacidades cognitivas do homem do que executando um algoritmo.

### 6.1 Redes Neurais Artificiais

Redes Neurais Artificiais são ferramentas de aproximação de funções que aprendem a relação entre variáveis dependentes e independentes, à semelhança da regressão linear ou outras abordagens mais tradicionais. A principal vantagem de uma rede neural é a sua capacidade de aproximar relações funcionais, particularmente quando as relações não são bem definidas ou não são lineares, o que torna difícil a utilização de

métodos convencionais para a tentativa de prever as variações futuras dessas relações [HAY01][KOV06].

Uma Rede Neural Artificial (RNA) é um modelo computacional que compartilha algumas das propriedades do cérebro: consiste de muitas unidades simples trabalhando em paralelo sem um controle central. As conexões entre unidades possuem pesos numéricos que podem ser modificados pelo elemento de aprendizado. A essas unidades damos o nome de neurônio e suas conexões sinapses [RUS95].

A Figura 13 mostra um modelo onde podem ser vistos: sinapses representadas pelas entradas e pesos sinápticos, somatório e função de ativação. Cada sinapse é caracterizada por um estímulo de entrada ( $x_j$ ) multiplicado pelo seu peso sináptico ( $w_{kj}$ ) correspondente. Depois desta multiplicação, cada sinal de entrada é somado e o resultado é então aplicado a uma função de ativação ( $\varphi$ ) que define a saída ( $y_k$ ) do neurônio [ROS04]. Tipicamente, o intervalo de saída de uma função de ativação está em um intervalo entre 0 e 1 ou de -1 a +1. O modelo também tem um bias ( $b_k$ ) aplicado externamente, o bias tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente [HAY01].

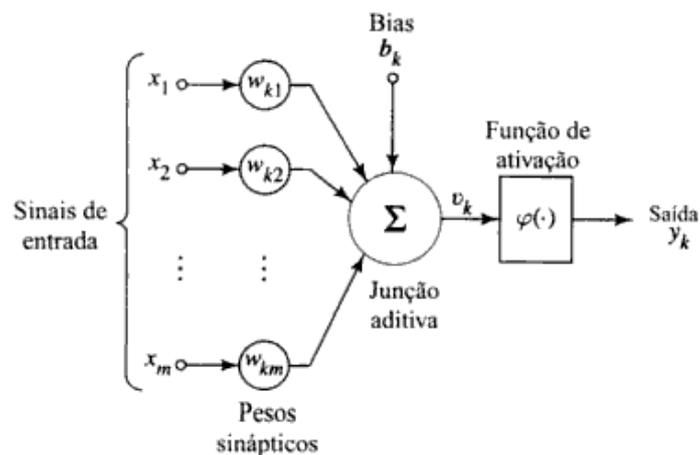


Figura 13 – Modelo não-linear de um neurônio artificial [HAY01].

Em termos matemáticos, o neurônio artificial pode ser descrito pelo seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (10)$$

$$y_k = \varphi(u_k + b_k) \quad (11)$$

onde  $x_1, x_2, \dots, x_m$  são sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos do neurônio  $k$ ;  $u_k$  é a saída do combinador linear;  $b_k$  é o bias;  $\varphi$  é a função de ativação; e  $y_k$  é o sinal de saída do neurônio. Para o resultado da soma  $u_k + b_k$  é dado o nome de “campo local induzido” [HAY01].

A função de ativação define a saída de um neurônio em termos do campo local induzido, tal função deve simular as características não lineares do neurônio biológico. As funções de ativação mais utilizadas são [HAY01][ROS04]:

- Função linear: é uma equação linear.
  - $f(x) = a.x$
- Função degrau: utilizada para valores binários.
  - se  $x > 0$        $f(x) = 1$
  - se  $x \leq 0$      $f(x) = 0$
- Função sigmóide: é a forma mais comum de ativação utilizada na construção de Redes Neurais Artificiais. É definida como uma função estritamente crescente que exhibe um balanceamento adequado entre o comportamento linear e não-linear. A função logística é um exemplo de função sigmóide, definida por:

$$\varphi(v) = \frac{1}{1 + \exp(-\alpha v)} \quad (12)$$

onde  $v$  é o campo local induzido do neurônio e  $\alpha$  é o parâmetro de inclinação da função sigmóide.

### 6.1.1 Arquiteturas de RNA

Basicamente, pode se identificar três classes de arquitetura de RNA: redes *feedforward* com camada única, redes *feedforward* com múltiplas camadas e redes recorrentes.

As redes *feedforward* são geralmente representadas com várias camadas e, neste caso, são chamadas redes *feedforward* multi-camadas. Os neurônios de saída constituem a camada de saída da rede. Os neurônios restantes constituem as camadas ocultas da rede. As unidades ocultas não são parte da saída ou da entrada da rede. A primeira camada oculta é alimentada pela camada de entrada (nós de fonte), as saídas resultantes da primeira camada oculta são por sua vez aplicadas à próxima camada oculta ou para a camada de saída da rede. A rede neural da Figura 14 é dita totalmente conectada, pois cada um dos nós de uma camada de rede está conectado a todos os nós da camada

adjacente seguinte. Caso alguns dos elos de comunicação estiverem faltando, se diz que a rede é parcialmente conectada [HAY01].

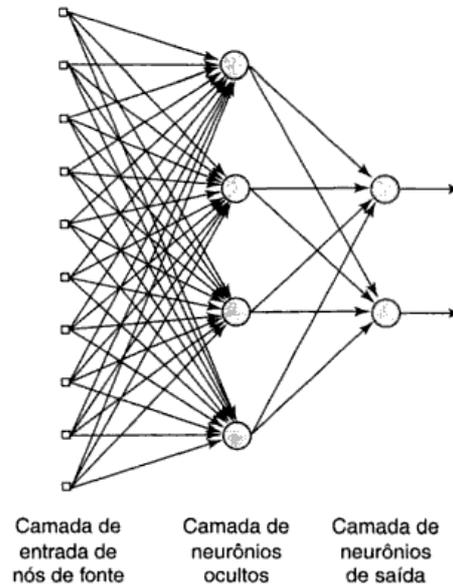


Figura 14 – Rede *feedforward* totalmente conectada com uma camada oculta e uma camada de saída [HAY01].

A principal diferença entre uma rede neural *feedforward* e uma rede recorrente está no fato desta última ter pelo menos um laço de realimentação. Uma rede recorrente pode consistir de uma camada única onde cada neurônio esteja alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios, conforme ilustrado na Figura 15. A aplicação de realimentação permite que as redes recorrentes adquiram representações de estado, o que as torna dispositivos apropriados para aplicações tão diversas como previsão não-linear e modelagem, equalização de canais de comunicação, processamento de voz, entre outros [HAY01].

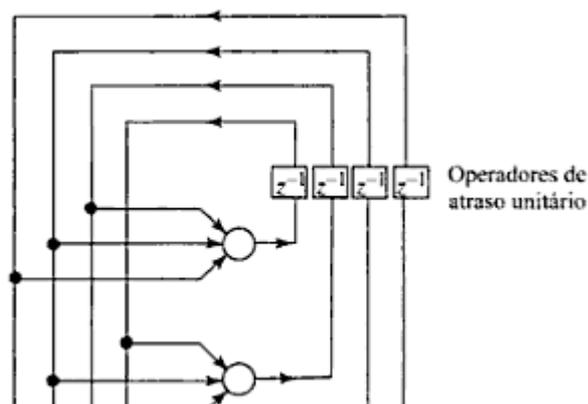


Figura 15 – Parte de uma rede recorrente sem neurônios ocultos [HAY01].

### 6.1.2 Aprendizagem da Rede Neural

Uma rede neural aprende acerca do seu ambiente através de um processo interativo de ajustes aplicados a seus pesos sinápticos e níveis de bias. Idealmente, a rede se torna mais instruída sobre o seu ambiente após cada iteração do processo de aprendizagem [HAY01].

O aprendizado no contexto das Redes Neurais é definido como um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. Os paradigmas de aprendizagem são:

- **Supervisionado:** o paradigma supervisionado também é conhecido por aprendizado com professor, isto porque se pode considerar o professor como tendo conhecimento sobre o ambiente, com este conhecimento sendo representado por um conjunto de exemplos de entrada-saída. A partir da entrada a rede realiza seu processamento e a saída obtida é comparada com a saída esperada. Caso não sejam iguais, um processo de ajuste de pesos é aplicado buscando-se um erro mínimo ou aceitável. Este processo de ajuste é conhecido como método de otimização ou algoritmo de treinamento [HAY01][ROS04].
- **Não supervisionado:** como o nome implica, não há um elemento externo para supervisionar o processo de aprendizagem. Isto significa que não há exemplos rotulados da função a ser aprendida pela rede. Este paradigma possui duas divisões: aprendizado por reforço que consiste no mapeamento entrada-saída através da interação com o ambiente, e aprendizagem não-supervisionada ou auto-organizada onde, a partir de métricas de qualidade do aprendizado ocorre a otimização dos parâmetros livres da rede [HAY01][ROS04].

O paradigma mais comum de aprendizado, no caso do reconhecimento de padrões, é o supervisionado, associado a uma rede direta multi-camadas. Devido a sua disponibilidade, o algoritmo de retro-propagação é frequentemente usado [BAR97].

O treinamento de redes neurais com várias camadas pode ser entendido como um caso especial de aproximação de funções, onde não é levado em consideração nenhum modelo explícito dos dados. Em [SIL98] pode ser encontrada uma lista com doze métodos de otimização para treinamento supervisionado. Devido à popularidade e disponibilidade, os algoritmos de treinamento denominados retro-propagação e uma de suas variantes, o

método do gradiente conjugado, tiveram um foco maior no desenvolvimento deste trabalho.

#### 6.1.2.1 Algoritmo de retro-propagação

O algoritmo de retro-propagação (*Backpropagation*) é um algoritmo de aprendizado supervisionado baseado na correção de erros. O objetivo do algoritmo é encontrar na superfície de erro um mínimo global, ou seja, valores para os pesos sinápticos que minimizam o erro da rede [HAY01]. A medida mais comum utilizada para verificação de erros em Redes Neurais é o RMSE (*Root Mean Square Error*), que pode ser descrito como:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (13)$$

onde  $x$  é o valor que está sendo ensinado (valor real),  $y$  é o valor previsto pela rede e  $n$  é o número de nodos  $i$  que possuem um valor previsto para a saída da rede. Segundo [HEA05], um RMSE abaixo de 0,1 é aceitável na maioria dos casos.

Em um algoritmo de aprendizado supervisionado são utilizados pares entrada-saída para ajustar os pesos da rede. No algoritmo, um sinal de erro se origina em um neurônio de saída da rede e se propaga para trás (camada por camada) através da rede. Sua computação por cada neurônio envolve uma função dependente do erro, de uma forma ou de outra [HAY01].

A frequência com que os ajustes são feitos na rede depende do modo de treinamento, que pode ser: seqüencial (on-line) onde a atualização é realizada após cada exemplo apresentado; ou por lote, onde a atualização é realizada após todos os exemplos do conjunto de treinamento serem apresentados a rede [HAY01][ROS04].

Em geral, não se pode demonstrar que o algoritmo de retro-propagação convergiu, sendo que não existem critérios bem definidos para encerrar seu processamento. Existem alguns critérios razoáveis, que podem ser: quando, percorrendo-se a superfície de erro, for encontrado um mínimo; a variação do erro for infinitamente pequena sendo insuficiente para melhorar o desempenho da rede; ou então baseado no número de épocas (número de vezes que foi processado) [HAY01].

#### 6.1.2.2 Método do gradiente conjugado

O processo de treinamento supervisionado de Redes Neurais Artificiais multi-camadas é equivalente a um problema de otimização não-linear, onde uma função de erro global é minimizada a partir do ajuste de parâmetros da rede neural (pesos). Esta

perspectiva do processo de treinamento supervisionado permite desenvolver algoritmos de treinamento baseados em resultados bem fundamentados da teoria de análise numérica convencional [SIL98].

Segundo [SIL98], os métodos de segunda ordem são considerados a maneira mais eficiente de se fazer o treinamento de redes neurais multi-camadas. Estes algoritmos recorrem a um rigor matemático baseado em modelos de otimização não-linear irrestrita bem definidos, não apresentando assim um vínculo natural com a inspiração biológica inicialmente proposta para as RNA's.

Dentre os métodos de segunda ordem existentes, existe um consenso geral da comunidade de análise numérica que a classe de métodos de otimização chamados métodos do gradiente conjugado tratam de problemas de grande escala de maneira efetiva [SIL98].

Os métodos do gradiente conjugado possuem sua estratégia baseada em um modelo geral de otimização, porém com procedimentos mais eficientes que utilizam informações de segunda ordem. Tais métodos são projetados para exigir menos cálculos do que outros métodos de segunda ordem e apresentar taxas de convergência maiores que as obtidas em métodos tradicionais [SIL98].

A maioria dos métodos de segunda ordem utiliza um procedimento de busca para a determinação da taxa de aprendizagem. Tal busca envolve um grande número de avaliações da função ou de sua derivada, tornando o processo bastante custoso do ponto de vista computacional. [MOL93] introduziu uma nova variação no algoritmo de gradiente conjugado, denominado Gradiente Conjugado Escalonado (*Scaled Conjugated Gradient - SCG*), que evita a busca a cada iteração. A aplicação deste método nas equações que fazem a retro-propagação do erro nos algoritmos das redes multi-camadas torna possível o cálculo exato da informação aplicável aos métodos de gradiente conjugado.

### 6.1.3 Software para simulação de RNA

Dentre os softwares existentes para simulação de Redes Neurais Artificiais, se destacam o Joone [JOO10] e o IBM SPSS Statistics [SPS10a]. Estes softwares permitem que seja modelada a rede neural, configurando os nodos, camadas, função de ativação e outros parâmetros que se façam necessários para o treinamento e validação da rede. Tais aplicativos também permitem a leitura de arquivos contendo os dados de treinamento e de validação, facilitando assim o trabalho de forma geral. Foram avaliados ambos os softwares, sendo o SPSS escolhido para o desenvolvimento da RNA proposta. Tal

escolha se deve ao fato do SPSS ter um sistema mais robusto de simulação de Redes Neurais e também por já estar sendo utilizado em outras tarefas deste trabalho.

#### 6.1.4 Normalização dos dados

Segundo [HEA05], as entradas de uma rede neural são sempre números reais. Para processar uma informação não-numérica é necessário um processo que execute a normalização dos dados para a representação numérica. Caso a informação já seja numérica, a normalização dos dados também deve aplicar algum cálculo que deixe o valor de entrada nos limites utilizados pela função de ativação.

De acordo com [SPS10b], a normalização de variáveis nominais (não-numéricas) é realizada mediante a criação de um nodo de entrada para cada valor possível da variável nominal. Por exemplo, a variável material possui dois valores possíveis, desta forma a rede terá dois nodos de entrada para esta variável. O primeiro nodo será ativado quando o material for alvenaria e o segundo nodo será ativado quando o material for vidro. Um nodo ativado recebe o valor 1 como entrada, já um nodo desativado recebe o valor 0.

Para as variáveis do tipo escala, o software SPSS conta com três opções de normalização [SPS10b]:

1. Padronizado: o novo valor de entrada se dá pela equação  $(x - m)/d$ , onde  $x$  é o valor de entrada original,  $m$  é a média dos valores de entrada, e  $d$  é o desvio padrão.
2. Normalizado: os novos valores sempre estarão entre 0 e 1. A fórmula aplicada é:  $(x - min)/(max - min)$ , onde  $x$  é o valor original,  $min$  é o menor valor encontrado para a variável, e  $max$  é o maior valor possível.
3. Normalizado com ajustes: os novos valores resultarão sempre entre -1 e 1. A fórmula é uma nova versão da opção anterior:  $[2*(x - min)/(max - min)] - 1$ .

## 6.2 Cenário móvel

Mesmo sendo desenvolvido um modelo de predição para o cenário móvel utilizando regressão linear (Capítulo 5), também foi desenvolvido um modelo de RNA para este cenário, permitindo assim uma comparação dos resultados de ambos modelos.

Para cada um dos cenários (móvel e estacionário) foi construído um modelo de RNA diferente, pois as características desses sistemas e o objetivo de cada um são distintos. Os modelos foram desenvolvidos utilizando o software SPSS.

### 6.2.1 Modelo

Conforme explicado na Seção 6.1.3, os dados de entrada da rede devem ser classificados em tipos dentro do software SPSS. Os fatores definidos como variáveis dos experimentos (Tabela 4 – Cenário móvel) serviram de entradas para esta rede. As entradas foram classificadas com os seguintes tipos: (i) material: nominal; (ii) sentido: nominal; (iii) distância: escala; (iv) potência: escala; (v) leituras: escala. A normalização dos dados de entrada se deu através da opção “Normalizado” do SPSS, onde o valor proporcional entre 0 e 1 é encontrado para cada valor de entrada original.

Para o cenário da etiqueta em movimento foi desenvolvido um modelo de RNA *feedforward* multi-camadas possuindo seis nodos na camada de entrada da rede neural. O nodo 1 possui a distância em linha reta entre a antena do leitor e a etiqueta, o nodo 2 recebe a potência configurada no leitor, os nodos 3 e 4 representam os dois sentidos possíveis e os nodos 5 e 6 constituem os materiais da parede.

Neste ponto é necessário decidir quantas camadas ocultas o modelo terá e a quantidade de neurônios para cada camada oculta. [HEA05] descreve algumas regras possíveis para se obter o melhor número de camadas e neurônios ocultos. Uma das regras cita que a quantidade de neurônios na camada oculta deve ser menor do que duas vezes o número de neurônios da camada de entrada. Também são especificados métodos de teste para determinar o número de camadas ocultas, sendo aplicada neste trabalho a abordagem de testar o modelo com poucas camadas e ir aumentando progressivamente a cada teste, até a taxa de erros da rede atingir um valor satisfatório.

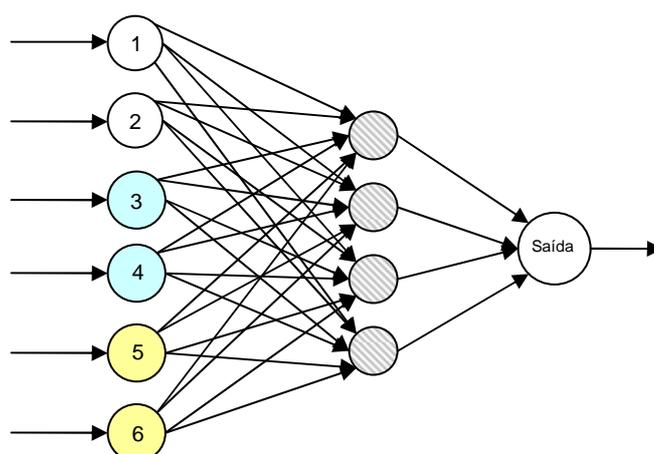


Figura 16 – RNA *feedforward* multi-camadas (cenário móvel).

O modelo deste cenário foi finalizado apresentando seis nodos na camada de entrada, quatro nodos na camada oculta e um nodo na camada de saída. O nodo da

camada de saída é o responsável por apresentar o resultado final, definido nesta abordagem como sendo o número de leituras da etiqueta para a entrada especificada (combinação entre material, sentido, distância e potência). A função de ativação utilizada em todas as camadas foi a sigmóide, desta forma a rede somente resulta valores entre 0 e 1. A Figura 16 ilustra o modelo completo.

Para o processo de aprendizado da RNA é necessário selecionar amostras que possam treinar o algoritmo em conjunto com suas saídas desejadas correspondentes. O treinamento da RNA proposta foi executado utilizando as informações registradas no banco de dados do software de coleta de dados apresentado na Seção 4.2. Para realizar o treinamento da rede, todas as informações registradas nos experimentos foram utilizadas, deixando para uma rodada nova de experimentos a tarefa de validação da RNA. Deste modo, o nodo 1 (distância) foi treinado com as duas distâncias que tiveram medições no experimento (0,40 e 0,80 m), o nodo 2 (potência) recebeu a informação registrada nos experimentos das potências 24,50, 26,50, 28,50, 30,50 e 32,50 dBm, os nodos 4 e 5 representam a ativação dos sentidos “entrando” e “saindo”, respectivamente; e por fim, os nodos 5 e 6 constituem os valores dos materiais, sendo ativado o nodo 5 quando o material é vidro e o nodo 6 quando o material é alvenaria.

### 6.2.2 Resultados

O treinamento da rede foi executado utilizando o recurso de Redes Neurais Multicamadas do software SPSS. Foram realizados vários testes de tentativa e erro para verificar quais seriam os parâmetros (conjunto de dados, algoritmo de aprendizado, modo de treinamento por lote ou seqüencial, critério de parada) que obteriam a menor taxa de erros. Por fim, os parâmetros de execução selecionados para o treinamento foram: (i) Conjunto de dados de treinamento: 100% dos dados dos experimentos. (ii) Modo de treinamento: lote (*batch*); (iii) Algoritmo de aprendizado: Gradiente Conjugado Escalonado; (iv) Critério de parada: parar quando a alteração do erro relativo atingir 0,0001.

Em uma primeira etapa de treinamento foram utilizados os dados de cada repetição de experimento, afim da RNA encontrar algum padrão para o número de leituras dentro das repetições realizadas. No entanto, o treinamento da rede desta maneira teve uma taxa de erros (RMSE, Seção 6.1.2.1) relativamente alta, ficando em 1,631. O fator de maior impacto neste resultado insatisfatório é a grande variância das leituras entre as repetições realizadas de uma mesma situação, não permitindo que o ajuste de pesos da rede minimizasse a taxa de erros.

Partindo do pressuposto que uma RNA também pode ser treinada utilizando somente uma amostra para cada situação distinta de entrada, decidiu-se utilizar como entrada de treinamento uma base de dados contendo a média das leituras para cada situação. Sendo assim, o número de amostras para o treinamento foi reduzido de 193 para 40, que é o número de situações distintas deste cenário. O treinamento foi realizado novamente com a mesma modelagem e parâmetros utilizados anteriormente. Ao final desse treinamento a taxa de erros (RMSE) da rede foi de 0,039. A melhora na taxa de erros mostra que a rede conseguiu realizar a aproximação dos valores inferidos com os valores reais, onde o ajuste dos pesos pelo método do Gradiente Conjugado Escalonado foi mais satisfatório do que o executado anteriormente.

O gráfico da Figura 17 exibe a relação entre as leituras dos experimentos (eixo X) e as leituras inferidas pela RNA (eixo Y). Por exemplo, nos dados de treinamento onde havia resultado real de 10 leituras, o valor inferido variou entre 9 e 13, aproximadamente. Um valor próximo da linha diagonal representa um resultado mais satisfatório, onde a diferença entre o valor do experimento e o valor inferido é menor.

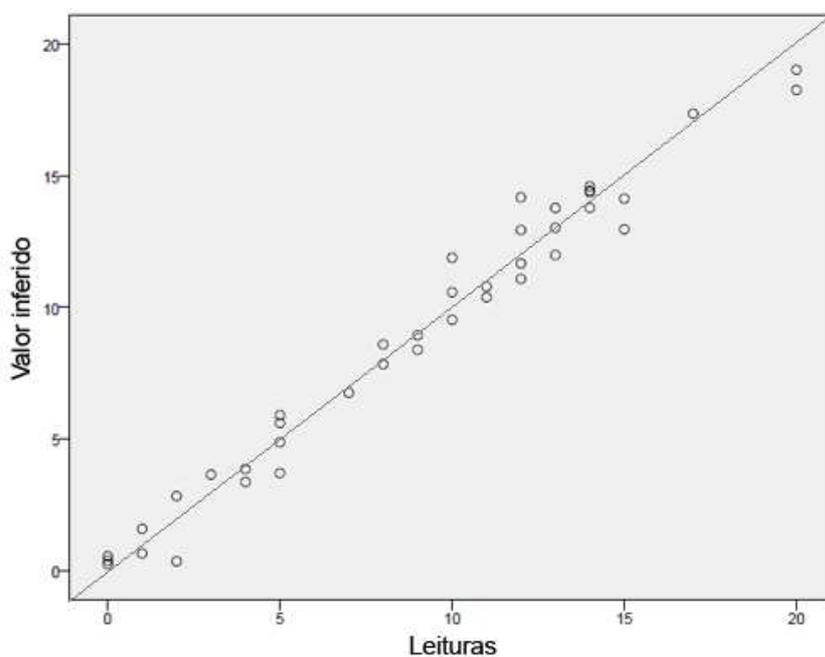


Figura 17 – Relação entre leituras reais e leituras inferidas pela RNA ao final do treinamento (cenário móvel).

Com o intuito de apresentar os resultados completos da intensidade de leituras no ambiente, foi realizada a predição de leituras para o cenário móvel com uma resolução de distância de 0,1 m e escala de potência de 1 dBm. A distância foi limitada entre 0,1 e 1m

e a potência entre 24,5 e 32,5 dBm. Os gráficos da Figura 18 exibem os resultados da cobertura do número de leituras inferido pela RNA desenvolvida.

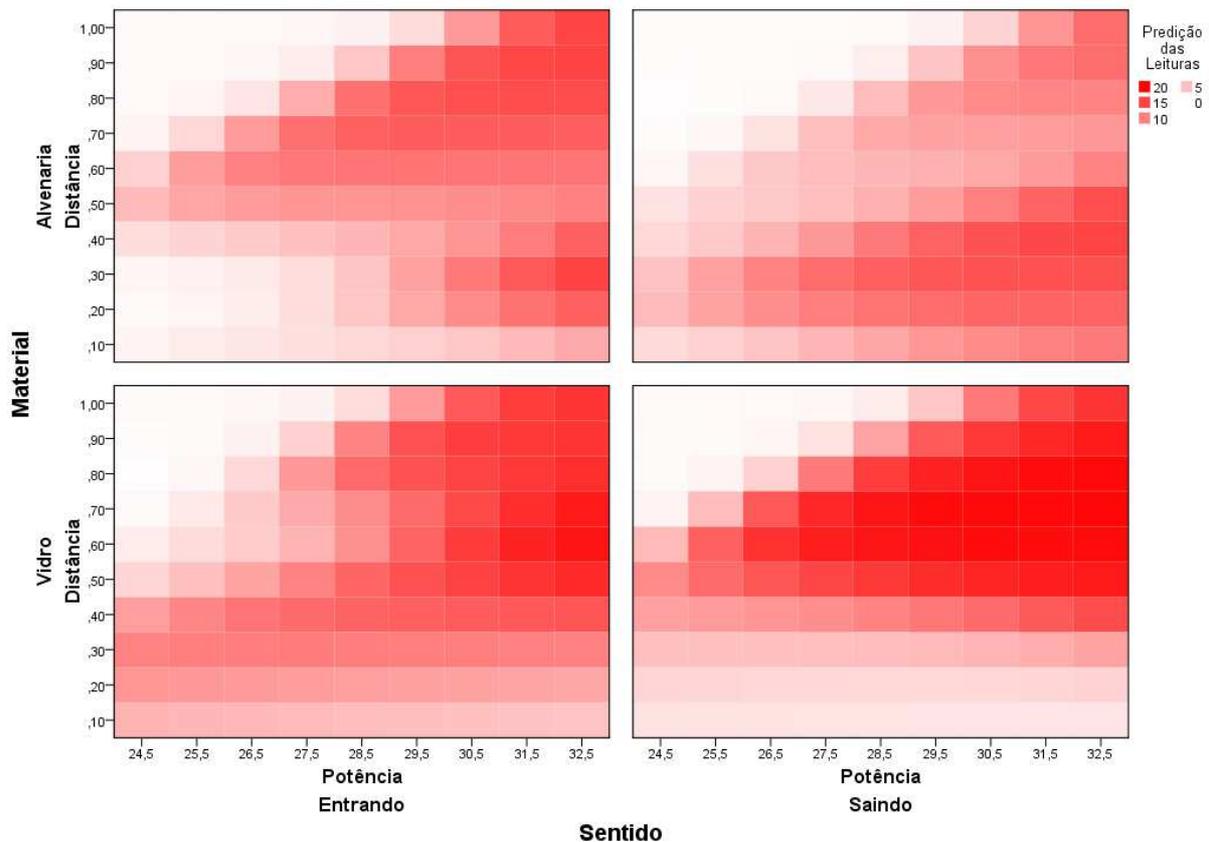


Figura 18 – Resultados da RNA para o cenário móvel.

Na Figura 18, os dois gráficos superiores exibem os resultados para o material “alvenaria” e os gráficos inferiores para o material “vidro”. Analisando a gradiente de cores obtida nos gráficos, para cada distância nota-se uma boa convergência do número de leituras em relação à potência (eixo X do gráfico), sendo que em todos os casos o aumento da potência causou um incremento no número de leituras. Já na análise da distância, alguns poucos casos tiveram comportamento fora do esperado, como no material “alvenaria”, sentido “entrando”; onde houveram leituras altas em distâncias próximas, decréscimo de leituras nas distâncias intermediárias (entre 0,4 e 0,7 m), seguidos de um acréscimo nas distâncias mais longas.

Analisando o comportamento do sistema em relação aos materiais da parede, a principal diferença visível nos gráficos é a de que o material “vidro” obteve, em várias potências e distâncias, um número de leituras maior em relação ao material “alvenaria”. Pode-se inferir também que o sentido “saindo” oferece uma melhor cobertura de leituras, tanto no que diz respeito à distância quanto à potência, isto ocorre possivelmente porque

o indivíduo portador do crachá permanece mais tempo do lado de dentro da sala com a etiqueta tendo visada livre para a antena do leitor.

### 6.2.3 Validação

Assim como a validação do modelo que utilizou regressão linear múltipla (Seção 5.4), a RNA desenvolvida foi validada a partir dos dados obtidos em novos experimentos. Desta forma a rede foi testada com valores nunca antes conhecidos por ela, ou seja, dados diferentes dos que foram utilizados no treinamento da rede.

Os dados dos experimentos de validação foram obtidos e analisados da mesma forma como descrito na Seção 5.4 (Validação do modelo utilizando regressão linear). Executando a RNA para as situações dos experimentos de validação, foi obtida a inferência de leituras para cada uma das amostras, conforme pode ser visualizado na Tabela 10.

Tabela 10 – Amostras da validação para o modelo de RNA do cenário móvel

<b>Material</b>	<b>Sentido</b>	<b>Distância (m)</b>	<b>Potência (dBm)</b>	<b>Leituras (real)</b>	<b>Leituras (inferência)</b>
Alvenaria	Entrando	0,4	24,5	4	3
Alvenaria	Entrando	0,4	26,5	5	4
Alvenaria	Entrando	0,4	28,5	6	6
Alvenaria	Entrando	0,4	30,5	6	8
Alvenaria	Entrando	0,4	32,5	6	12
Alvenaria	Entrando	0,6	25,5	6	8
Alvenaria	Entrando	0,6	27,5	5	11
Alvenaria	Entrando	0,6	29,5	12	11
Alvenaria	Entrando	0,6	31,5	15	11
Alvenaria	Entrando	0,8	24,5	1	0
Alvenaria	Entrando	0,8	26,5	5	2
Alvenaria	Entrando	0,8	28,5	8	11
Alvenaria	Entrando	0,8	30,5	11	14
Alvenaria	Entrando	0,8	32,5	12	14
...	...	...	...	...	...

Na Figura 19 pode ser visualizado um gráfico onde é realizada a comparação entre todos os resultados. Neste gráfico o valor medido no experimento de validação é comparado com o valor inferido. Quanto mais distante da linha diagonal estiver a elipse, maior é a diferença entre os valores. O erro absoluto, calculado da mesma forma que na Seção 5.4, é representado no gráfico através do tamanho de cada elipse (quanto maior a elipse, maior é o erro absoluto).

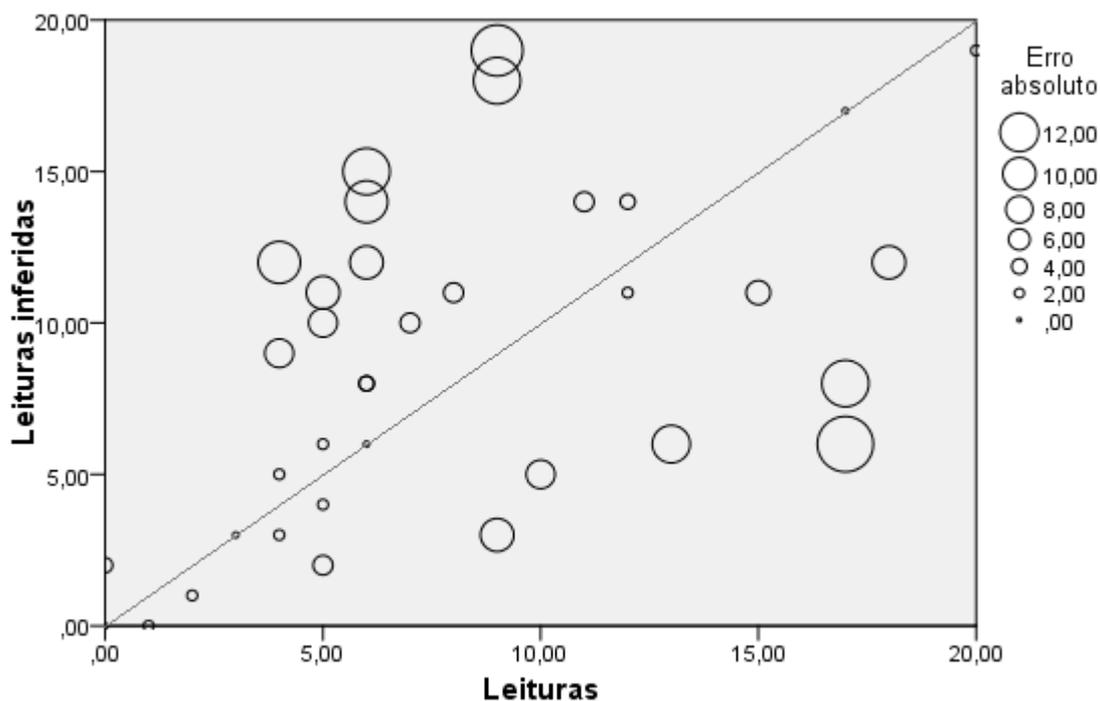


Figura 19 – Relação entre o número de leituras medido e inferido na validação da RNA (cenário móvel)

Conforme explicado na validação do modelo linear, uma comparação direta entre os valores dos experimentos e os valores inferidos não levaria a um resultado conciso. Desta forma, optou-se por utilizar o mesmo método de classificação desenvolvido na validação do modelo de regressão linear, descrito no final da Seção 5.4. A Tabela 11 mostra os resultados de média e desvio padrão dos erros para a validação do modelo desenvolvido no cenário móvel.

Tabela 11 – Diferença entre o valor inferido e o valor medido na validação do modelo utilizando RNA (cenário móvel)

Material/Sentido	Entrando		Saindo	
	Média	Desvio padrão	Média	Desvio padrão
Alvenaria	2,5	1,8	4,0	3,3
Vidro	5,8	4,5	7,0	4,1
<b>Média geral</b>	3,9			
<b>Desvio padrão geral</b>	3,33			

Conforme já explicado no desenvolvimento do modelo utilizando regressão linear, não é de conhecimento um erro mínimo desejado para a inferência do número de leituras em um sistema RFID. Assim, optou-se por realizar a interpretação do desvio padrão geral dos erros absolutos. Levando em consideração que a faixa de leituras obtida nos experimentos iniciais deste cenário ficou entre 0 e 25 (Seção 4.3), o desvio padrão geral obtido (3,33) representa 13% do valor máximo de leituras. Considera-se que este percentual não representa uma alteração tão grande no comportamento do sistema. O fato de alguns pontos inferidos terem ficado distantes do valor medido mostra o quão inconstante é este sistema, onde somente uma base estatística que abranja mais variáveis e tenha muito mais repetições poderia trazer resultados mais satisfatórios.

Em comparação com o modelo obtido usando regressão linear múltipla, tanto a média quanto o desvio padrão geral deste modelo tiveram resultados inferiores ao modelo apresentado no Capítulo 5.

### 6.3 Cenário estacionário

As seções seguintes descrevem o modelo de RNA desenvolvido para o cenário estacionário, bem como os resultados e a validação do modelo.

#### 6.3.1 Modelo

O desenvolvimento do modelo para o cenário estacionário teve poucas diferenças em relação à modelagem do cenário móvel, visto que também se desenvolveu uma RNA *feedforward* multi-camadas. Uma das diferenças está na camada de entrada, que foi definida com sete nodos.

As entradas foram classificadas com os seguintes tipos de dados: (i) material: nominal; (ii) sentido: nominal; (iii) distância X: escala; (iii) distância Y: escala; (iv) potência: escala; (v) leituras: escala. A normalização dos dados de entrada foi realizada através da opção “Normalizada com ajustes” do SPSS, onde o valor proporcional entre -1 e 1 é encontrado para cada valor de entrada original. Ao longo dos testes de treinamento esta opção foi a que resultou a menor taxa de erros.

Os nodos 1 e 2 representam as distâncias entre a antena do leitor e a etiqueta, sendo que o nodo 1 recebe a distância horizontal no plano 2D e o nodo 2 recebe a distância vertical (eixos X e Y da Figura 7, respectivamente). O nodo 3 considera a potência configurada no leitor, os nodos 4 e 5 representam os dois sentidos possíveis e os nodos 6 e 7 constituem os materiais da parede.

Para decidir o número de camadas ocultas deste modelo foi realizada a mesma abordagem utilizada no modelo do cenário móvel. Este modelo apresentou uma melhor

taxa de erros quando foi testado utilizando duas camadas ocultas, diferente do modelo anterior que utilizou somente uma camada oculta. O modelo deste cenário foi finalizado apresentando sete nodos na camada de entrada, quatro nodos na primeira camada oculta, três nodos na segunda camada oculta e apenas um nodo na camada de saída. O nodo da camada de saída é o responsável por apresentar o resultado final, definido como sendo o número de leituras da etiqueta para a entrada especificada (combinação entre material, sentido, distâncias e potência). A função de ativação utilizada em todas as camadas foi a sigmóide. A Figura 20 ilustra o modelo completo.

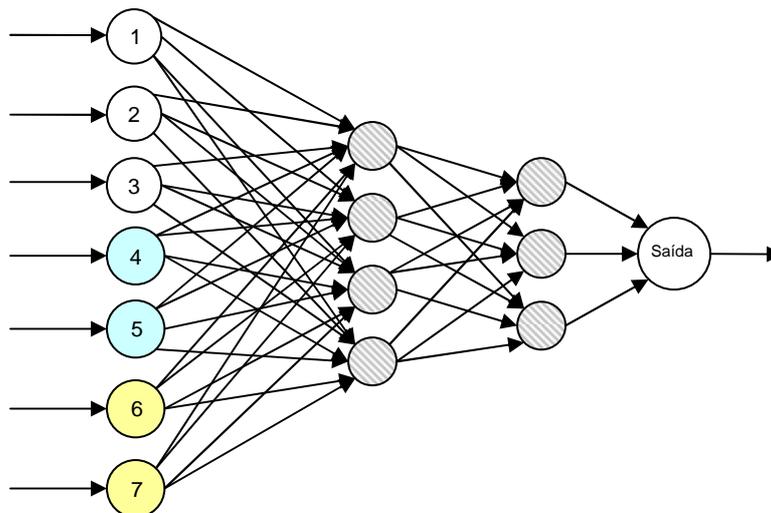


Figura 20 – RNA *feedforward* multi-camadas (cenário estacionário).

Da mesma forma que no cenário móvel, o treino da RNA foi executado utilizando todas as informações registradas nos experimentos iniciais, deixando para uma rodada nova de experimentos a tarefa de validação da RNA. O nodo 1 (distância no eixo X) foi treinado com as cinco distâncias medidas nos experimentos (-0,80, -0,40, 0, 0,40 e 0,80 m). O nodo 2 (distância no eixo Y) recebeu as duas distâncias que tiveram medições nos experimentos (0,40 e 0,80 m), o nodo 3 (potência) recebeu a informação registrada nos experimentos das potências 24,50, 26,50, 28,50, 30,50 e 32,50 dBm. Os nodos 5 e 6 representam a ativação dos sentidos “entrando” e “saindo”, respectivamente; e por fim, os nodos 7 e 8 constituem os valores dos materiais, sendo ativado o nodo 7 quando o material é vidro e o nodo 8 quando o material é alvenaria.

### 6.3.2 Resultados

O treinamento da RNA executado no software SPSS contou com os mesmos parâmetros utilizados no modelo do cenário móvel (Seção 6.2.2). Assim como no modelo

anterior, o resultado da rede treinada utilizando cada repetição de experimento também se mostrou insatisfatório, tendo um RMSE de 30,174. A partir disto se optou por utilizar a média das leituras para cada situação, tendo assim um total de 360 amostras. Ao final desse treinamento a taxa de erros (RMSE) da rede foi de 0,733. A relação entre as leituras dos experimentos reais com os valores inferidos pela RNA pode ser visualizada na Figura 21.

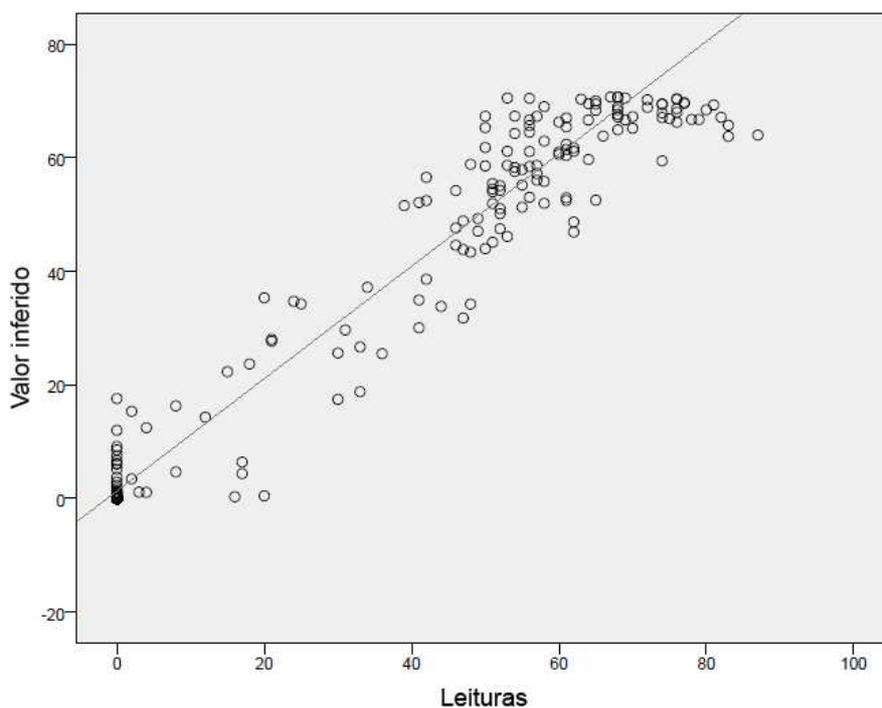


Figura 21 – Relação entre leituras reais e leituras inferidas pela RNA (cenário estacionário).

Após o treinamento da RNA, criou-se uma base de dados para inferência do número de leituras utilizando entradas ainda não conhecidas pela rede. A base de dados possui informações de ambos os materiais e ambos os sentidos de locomoção. Utilizou-se uma resolução de distâncias de 0,1 m e escala de potência de 0,5 dBm. O perímetro de cobertura foi limitado em 2 m no eixo X e 1 m no eixo Y, de acordo com o plano 2D do ambiente. A inferência realizada pela RNA resultou em uma base de dados com informações de número de leituras em todos os pontos do ambiente para todas as situações de material, sentido e potência. Os gráficos da Figura 22 exibem a cobertura da leitura de etiquetas inferida pela RNA para a potência 30 dBm. As dez elipses internas de cada gráfico representam as posições onde foram realizados os experimentos reais.

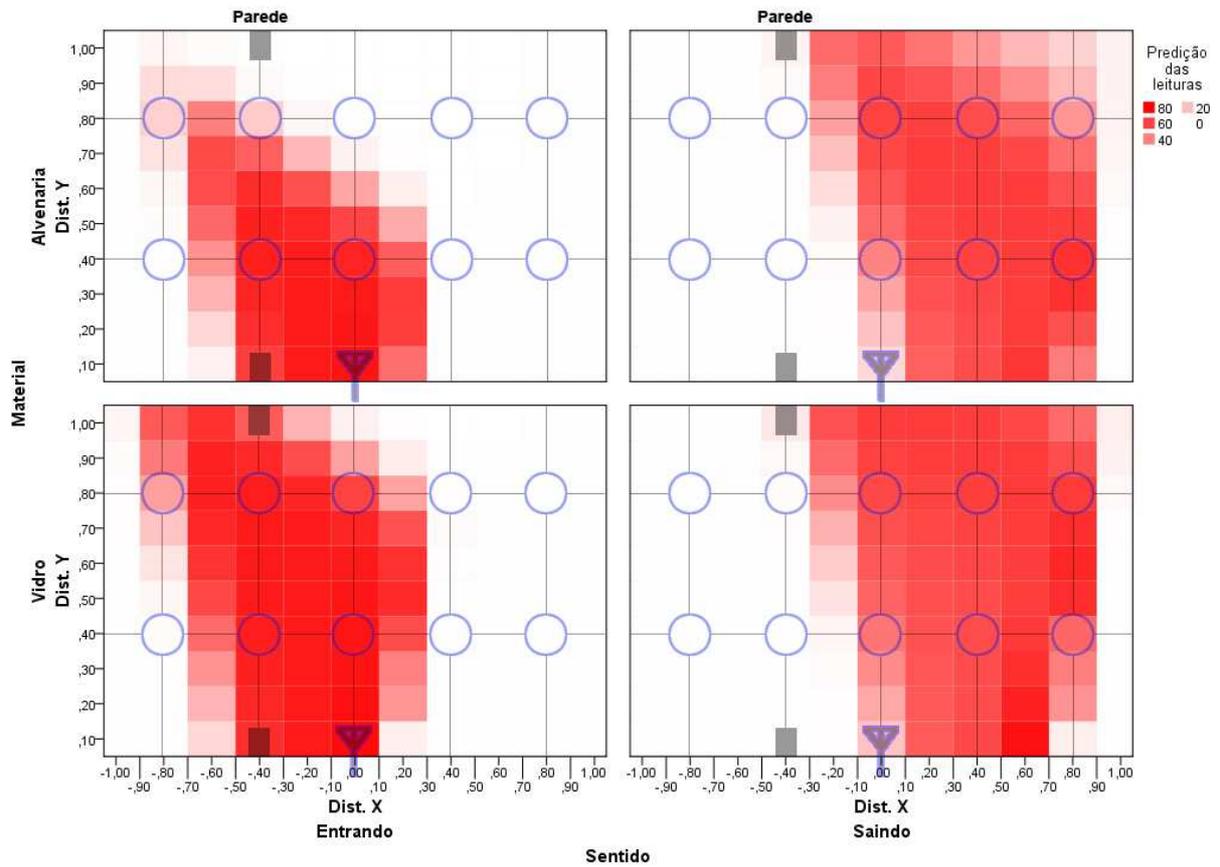


Figura 22 – Cobertura de leituras inferida no cenário estacionário (potência: 30 dBm).

Na Figura 22, os gráficos exibem a cobertura completa do número de leituras no lado externo e interno da sala. Analisando os resultados para o sentido de locomoção “entrando”, entende-se que a cobertura de leituras aparenta ter uma vazão em direção ao lado de fora da sala, o que é razoável considerando o fato do indivíduo estar vindo de lá com a etiqueta em frente ao seu corpo. Pode-se notar também que no sentido “entrando” não existem leituras em distâncias superiores a 0,2 m (eixo X), devido ao fato que o indivíduo portador do crachá está de costas para a antena do leitor, encobrindo a etiqueta.

O sentido “saindo” mostra uma cobertura de leituras onde não existe comunicação entre leitor e etiqueta no lado externo da sala, local onde o indivíduo já está com a etiqueta encoberta, permanecendo de costas para a antena. Analisando o interior da sala, a cobertura de leituras do sentido “saindo” é praticamente completa, sendo notável que as distâncias mais longas (tanto do eixo X quanto do eixo Y) têm um número de leituras inferior às distâncias mais próximas ao leitor.

Ainda analisando os resultados da Figura 22, na comparação dos resultados entre os materiais “alvenaria” e “vidro”, percebe-se que não existem muitas diferenças na cobertura das leituras. A única situação mais sobressalente acontece no sentido

“entrando”, onde o material “alvenaria” possui uma cobertura de leituras menor do que o material “vidro”.

Uma análise da predição de leituras em relação à potência do leitor também é necessária. Os gráficos da Figura 23 exibem o número de leituras inferido para cada potência, sendo que nestes dados a etiqueta está sempre posicionada em frente ao leitor (eixo X do plano 2D em 0 m). Na maioria das combinações de material e sentido, quanto maior a potência, maior é o número de leituras, independente da distância. A única exceção se dá no material “vidro”, sentido “saindo”; onde o número de leituras entre as potências 25 e 27 dBm está maior do que os valores obtidos nas potências superiores.

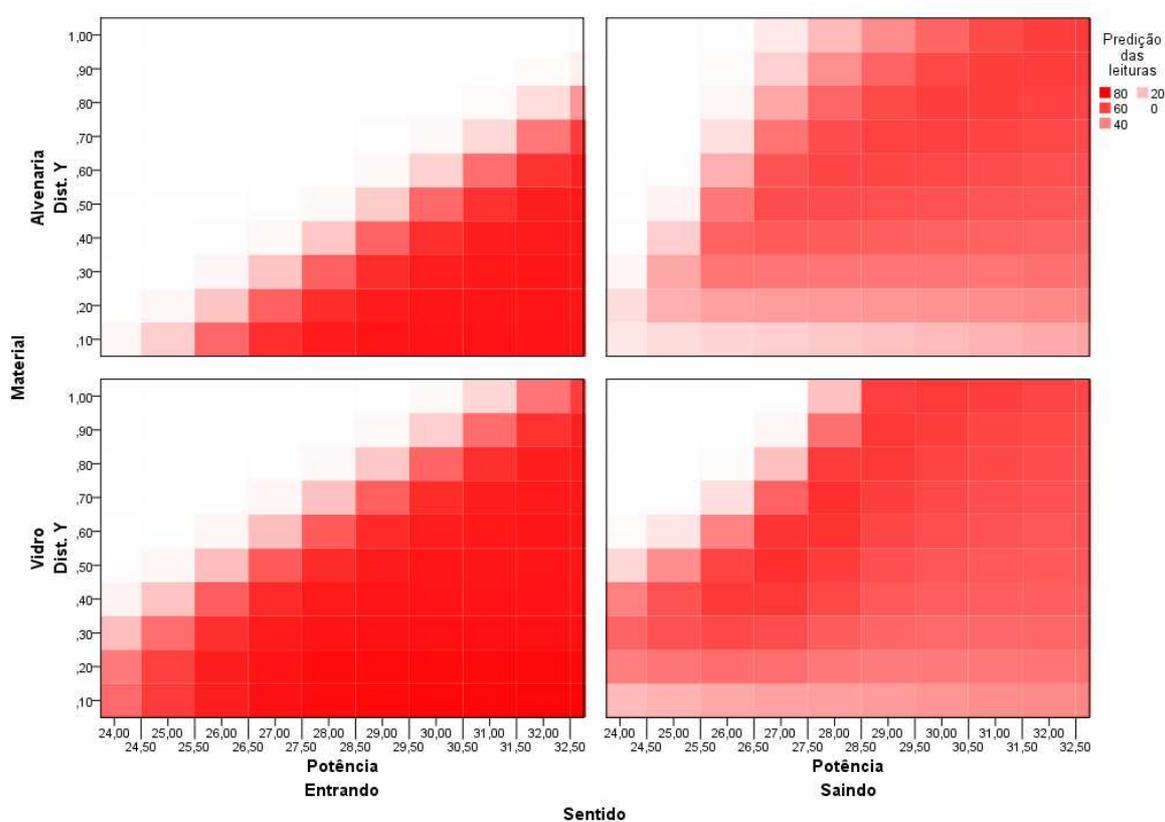


Figura 23 – Relação da inferência entre material, sentido, distância (Y) e potência no cenário estacionário.

Na análise da distância em linha reta entre o leitor e a etiqueta (eixo Y), pode-se perceber que na inferência do sentido “saindo” as distâncias curtas tiveram um desempenho menor do que o número de leituras para distâncias mais longas, ao contrário do que acontece no sentido “entrando”.

### 6.3.3 Validação

Da mesma forma como descrito nas validações dos modelos anteriores (Seções 5.4 e 6.2.3), o cenário estacionário também utilizou duas fases de experimentos a fim de coletar dados para realizar a validação do modelo desenvolvido. Os novos experimentos realizados tiveram o objetivo de comparar o resultado das inferências da RNA com as leituras obtidas em testes reais.

Assim como no cenário móvel, a primeira fase da validação utilizou um ambiente com características semelhantes ao local dos experimentos originais, sendo o objetivo desta fase comparar os valores para as mesmas situações que foram treinadas na rede. Na segunda fase da validação foram realizados experimentos com distâncias e potências nunca antes registrados. A nova distância vertical (eixo Y) testada na validação foi de 0,6 m. No eixo horizontal (X), foram validadas as distâncias intermediárias -0,6, -0,2, 0,2, 0,4 m. Neste processo de validação também se utilizou potências intermediárias em relação aos experimentos iniciais: 25, 26, 27, 28, 29, 30, 31 e 32 dBm.

Foi coletado um total de 308 amostras de validação, sendo 180 na primeira fase e 128 na fase dois. A partir dos dados coletados realizou-se a comparação das novas leituras com os valores inferidos pela rede. A comparação de todos os valores inferidos com o número de leituras obtido em experimentos reais pode ser visualizado no gráfico da Figura 24. Alguns destes resultados estão descritos de forma textual na Tabela 12.

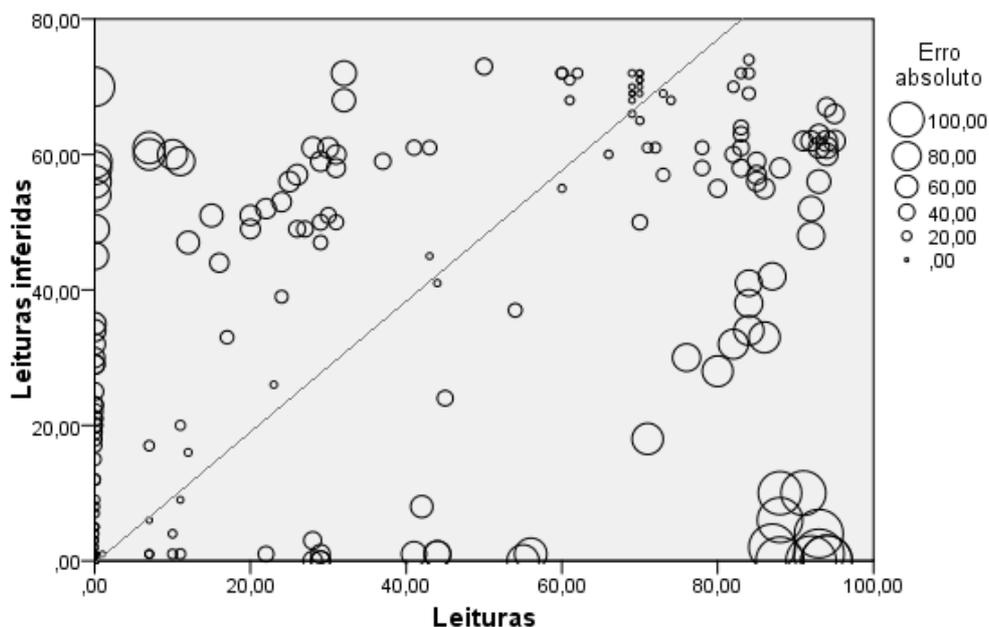


Figura 24 – Relação entre o número de leituras medido e inferido na validação da RNA (cenário estacionário)

Tabela 12 – Amostras da validação para o modelo de RNA do cenário estacionário

<b>Material</b>	<b>Sentido</b>	<b>Dist. X (m)</b>	<b>Dist. Y (m)</b>	<b>Potência (dBm)</b>	<b>Leituras (real)</b>	<b>Leituras (inferência)</b>
Vidro	Saindo	0,2	0,6	25	0	3
Vidro	Saindo	0,2	0,6	26	16	44
Vidro	Saindo	0,2	0,6	27	95	66
Vidro	Saindo	0,2	0,6	28	71	61
Vidro	Saindo	0,2	0,6	29	88	58
Vidro	Saindo	0,2	0,6	30	85	57
Vidro	Saindo	0,2	0,6	31	85	56
Vidro	Saindo	0,2	0,6	32	86	55
Vidro	Saindo	0,6	0,6	25	0	3
Vidro	Saindo	0,6	0,6	26	15	51
Vidro	Saindo	0,6	0,6	27	92	62
Vidro	Saindo	0,6	0,6	28	83	64
Vidro	Saindo	0,6	0,6	29	93	63
Vidro	Saindo	0,6	0,6	30	91	62
Vidro	Saindo	0,6	0,6	31	94	60
Vidro	Saindo	0,6	0,6	32	85	59
...	...	...	...	...	...	...

A análise dos resultados da validação do cenário estacionário não difere da abordagem utilizada para o cenário móvel, onde uma comparação direta entre os valores dos experimentos de validação e os valores inferidos não levaria a um resultado conciso, visto que o cenário que produziu tal modelo também tem uma variância grande nos resultados entre uma repetição e outra. Desta forma, decidiu-se utilizar o mesmo método de classificação desenvolvido na validação dos modelos anteriores (Seções 5.4 e 6.2.3). A Tabela 13 mostra os resultados de média e desvio padrão dos erros para a validação do modelo desenvolvido no cenário estacionário.

Tabela 13 – Diferença entre o valor inferido e o valor medido na validação do modelo utilizando RNA (cenário estacionário)

<b>Material/Sentido</b>	<b>Entrando</b>		<b>Saindo</b>	
	<b>Média</b>	<b>Desvio padrão</b>	<b>Média</b>	<b>Desvio padrão</b>
<b>Alvenaria</b>	14,1	25,1	15,0	16,4
<b>Vidro</b>	14,7	23,7	20,3	16,0
<b>Média geral</b>	15,2			
<b>Desvio padrão geral</b>	21,0			

Os resultados da validação deste modelo apresentaram uma média e desvio padrão acima do esperado, tais valores têm uma significância razoável se verificarmos

que a faixa de leituras dos experimentos iniciais deste cenário se manteve entre 0 e 96 (Seção 4.3). Levando em consideração essa faixa de leituras, o desvio padrão geral (21) dos erros representa 22% do valor máximo de leituras. Considera-se que este é um percentual relativamente alto para o modelo, podendo inferir um comportamento errôneo para o sistema em determinados casos. A distribuição dos pontos do gráfico da Figura 24 mostra muitos casos onde um número de leituras inferido está distante do valor obtido no experimento real. A sensibilidade alta dos dispositivos e a influência de variáveis externas podem ser os principais motivadores destas diferenças que, da mesma forma que ocorreram durante os experimentos iniciais, também vieram a acontecer nos experimentos de validação.

## 7 CONCLUSÃO

O trabalho apresentou três modelos estatísticos de inferência para um cenário específico de uso da tecnologia RFID. O primeiro modelo abordou o método chamado Regressão Linear Múltipla, sendo aplicado em um cenário com a etiqueta em movimento. Os dois outros modelos utilizaram Redes Neurais Artificiais, sendo aplicados tanto em um cenário móvel quanto em um cenário estacionário.

Por se tratarem de modelos empíricos, uma base histórica de amostras necessitou ser coletada a partir de experimentos em campo. Para realizar a tarefa de experimentação foi desenvolvido um software capaz de automatizar a coleta dos dados. Durante o projeto dos experimentos, toda a configuração dos dispositivos foi previamente analisada, sendo detectados os principais fatores que influenciavam na detecção das etiquetas RFID. Alguns fatores foram definidos como fixos na configuração do ambiente, já os fatores mais importantes foram escolhidos como variáveis para os modelos desenvolvidos, tendo seus valores alternados durante a execução dos experimentos. O valor inferido pelos modelos é o número de leituras de uma etiqueta RFID, tal informação é capaz de demonstrar quais são as configurações do sistema e características do ambiente que proporcionam uma melhor comunicação no sistema RFID.

Durante a execução dos experimentos foi detectada uma alta variância no número de leituras entre as repetições dos testes em uma mesma configuração de ambiente. Apesar disto, os modelos desenvolvidos foram capazes de inferir as configurações que proporcionam o melhor e o pior número de leituras, ou seja, a melhor e a pior comunicação entre os dispositivos. Os resultados de inferência dos modelos apontaram que o número de leituras aumenta à medida que a potência do equipamento leitor é incrementada. Não foi detectada uma diferença relevante de comportamento entre os materiais da parede do ambiente onde foram realizados os ensaios.

A posição da etiqueta dentro do ambiente também é um fator de impacto no número de leituras, ocasionando uma comunicação diferenciada em determinados pontos do ambiente. A diferença no número de leituras em virtude do posicionamento dos componentes é motivada pelos efeitos dos inúmeros valores máximos ou mínimos de sinal distribuídos ao longo de toda região de cobertura. Tais efeitos decorrem da interferência construtiva ou destrutiva das diversas ondas refletidas no cenário de operação, fato que não foi considerado no presente trabalho e deve ser objeto de pesquisa em trabalhos futuros.

A validação dos modelos se deu através de novos experimentos em campo, onde os modelos obtiveram um erro relativamente baixo no cenário móvel e um erro alto em determinados casos do cenário estacionário. Na maioria dos casos, tais erros são aceitáveis se for considerada a variância existente durante os próprios experimentos de coleta dos dados, onde fatores como a alta sensibilidade dos dispositivos e variáveis externas não consideradas podem impactar nos valores das amostras obtidas.

Portanto, essa Dissertação documenta o processo de criação de modelos estatísticos para a inferência do funcionamento de um sistema RFID. Tal inferência se dá a partir de um cenário típico da tecnologia, onde a predição pode informar o número de leituras que uma etiqueta obtém para determinada configuração do ambiente. A partir dos modelos desenvolvidos é possível simular um ambiente que utiliza tal tecnologia, trazendo benefícios para o projeto de implantação desses sistemas.

## 7.1 Trabalhos futuros

Durante a realização dos experimentos necessários para a geração dos modelos estatísticos, pôde ser confirmada a influência de outros fatores do ambiente que não foram considerados no projeto dos ensaios iniciais. Tais fatores podem causar impacto significativo no funcionamento de sistemas RFID. Um dos fatores presenciados durante os ensaios pode ter relação com o nível de luminosidade incidente no ambiente. Um trabalho minucioso, onde todas as outras variáveis do ambiente sejam isoladas, pode verificar o impacto da variância desta característica sobre o funcionamento do sistema RFID.

A metodologia documentada para geração dos modelos não se restringe ao cenário dos crachás inteligentes, também podendo ser aplicada a outros cenários típicos de uso de sistemas RFID, tais como esteiras por onde trafegam produtos contendo etiquetas RFID afixadas, ou então *pallets* carregando caixas que possuem etiquetas RFID.

Uma variável que não está incluída no trabalho proposto diz respeito ao número de etiquetas existentes no ambiente. Um sistema RFID atuando com mais de uma etiqueta pode ter um comportamento diferenciado na camada física quando comparado a um sistema onde existe somente uma etiqueta. Em um sistema com várias etiquetas atuando simultaneamente também é necessário tratar o problema de colisão das mensagens da camada lógica.

Finalmente, é interessante a implantação dos modelos desenvolvidos neste trabalho em um software simulador que permita o desenho e projeto visual de ambientes RFID. Já existe no grupo de pesquisa um trabalho cujo objetivo é o desenvolvimento de

um software deste gênero, sendo que tal aplicação também irá fornecer uma interface para integração de modelos de predição externos.



## REFERÊNCIAS BIBLIOGRÁFICAS

- [ALI10] Alien Technology. "Alien Technology - Products - RFID Tags". Capturado em: <http://www.alientechnology.com/tags>, Dezembro de 2010.
- [AZA08a] Azambuja, M. C.; Marcon, C.; Hessel, F. P. "Survey of Standardized ISO 18000-6 RFID Anti-Collision Protocols". In: The Second International Conference on Sensor Technologies and Applications, 2008, pp. 468-473.
- [AZA08b] Azambuja, M. C.; Marcon, C.; Hessel, F. P. "A Communication Protocol and Physical Characteristics Simulator for an RFID Sensor Environment". In: IEEE IWCMC'08 International Wireless Communications and Mobile Computing Conference, 2008, pp. 1093-1098.
- [AZA10] Azambuja, M. C.; Jung, C. F.; Caten, C. S. T.; Hessel, F. P. "*RFID-Env Methods and Software Simulation for RFID Environments*", Business Process Management Journal, vol. 16-6, Nov 2010, pp. 1014-1038.
- [BAR97] Barreto, J. M. "Introdução às Redes Neurais Artificiais". In: V Escola Regional de Informática da SBC Regional Sul, 1997, pp.41-71.
- [CHW07] Chwif, L.; Medina, C. A. "Modelagem e simulação de eventos discretos: teoria e aplicações". São Paulo: Ed. do Autor, 2007, 2ª edição, 254p.
- [EPC10] EPC Global. "Regulatory status for using RFID in the UHF spectrum 13 December 2010". Capturado em: [http://www.gs1.org/docs/epcglobal/UHF\\_Regulations.pdf](http://www.gs1.org/docs/epcglobal/UHF_Regulations.pdf), Dezembro de 2010.
- [FIN03] Finkenzeller, K. "RFID Handbook: Fundamentals and Applications in Contact Less Smart Cards and Identification". New York: Wiley & Sons, 2003, 2ª edição, 446p.
- [FLO09] Floerkemeier, C.; Sarma, S. "*RFIDSim - A Physical and Logical Layer Simulation Engine for Passive RFID*", IEEE Transactions on Automation Science and Engineering, vol. 6-1, Jan 2009, pp. 33-43.
- [GAR09] Garcia, C. "Modelagem e Simulação de Processos Industriais e de Sistemas Eletromecânicos". São Paulo: USP, 2009, 2ª edição, 678p.
- [GLO06] Glover, B.; Bhatt, H. "RFID Essentials". Sebastopol: O'reilly, 2006, 1ª edição, 288p.

- [HAY01] Haykin, S. "Redes Neurais: Princípios e prática". Porto Alegre: Bookman, 2001, 2ª edição, 900p.
- [HEA05] Heaton, J. "Introduction to Neural Networks with Java". Heaton Research Inc, 2005, 1ª edição, 380p.
- [JAI10] Jain, A.; Agarwal, S.; Jagannatha, S. K.; Narayanaswamy, G.; Engels, D. W. "Performance analysis of alien squiggle tag in human presence". In: Antennas and Propagation Society International Symposium (APSURSI), 2010, pp. 1-4.
- [JEC06] Jechlitschek, C. "A Survey Paper on Radio Frequency IDentification (RFID) Trends". Capturado em: <http://www1.cse.wustl.edu/~jain/cse574-06/ftp/rfid/index.html>, Novembro de 2010.
- [JO07] Jo, M.; Lim, C.; Zimmers, E. W. "*RFID Tag Detection on a Water Content Using a Back-propagation Learning Machine*", KSII Transactions on Internet and Information Systems, vol. 1-1, Dez 2007, pp. 19-32.
- [JO09a] Jo, M.; Youn, H. Y.; Cha, S.; Choo, H. "*Mobile RFID Tag Detection Influence Factors and Prediction of Tag Detectability*", IEEE Sensors Journal, vol. 9-2, Fev 2009, pp. 112-119.
- [JO09b] Jo, M.; Youn, H. Y.; Chen, H. "*Intelligent RFID Tag Detection Using Support Vector Machine*", IEEE Transactions on Wireless Communications, vol. 8-10, Out 2009, pp. 5050-5059.
- [JON07] Jones, E. C.; Chung, C. A. "RFID in logistics: a practical introduction". CRC Press, 2007, 520p.
- [JOO10] Joone. "an Object Oriented Neural Engine". Capturado em: <http://sourceforge.net/projects/joone/>, Novembro 2010.
- [KIM08] Kim, D. "Interference analysis of UHF RFID systems". In: Progress In Electromagnetics Research B, 2008, pp. 115-126.
- [KNO96] Knoblauch, R.; Pietrucha, M.; Nitzburg, M. "*Field Studies of Pedestrian Walking Speed and Start-up Time*", Transportation Research Record, vol. 1538, 1996, pp. 27-38.
- [KOV06] Kovacs, Z. L. "Redes Neurais Artificiais: Fundamentos e Aplicações". São Paulo: Ed. Livraria da Física, 2006, 4ª edição, 174p.

- [LAR10] Larson, R.; Farber, B. "Estatística aplicada". São Paulo: Pearson Prentice Hall, 2010, 4ª edição, 637p.
- [MIC10] Microsoft Office Excel. "Perform a regression analysis". Capturado em: <http://office.microsoft.com/en-us/excel-help/perform-a-regression-analysis-HA001111963.aspx>, Dezembro de 2010.
- [MOL93] Moller, M. F. "*A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning*", Neural Networks, vol. 6-4, 1993, pp. 525-533.
- [MTI10] MTI Wireless Edge Ltd. "Antenna Data Sheet MT-262010/TRH/A". Capturado em: <http://www.mtiwe.com/uploads/product/436.pdf>, Dezembro de 2010.
- [NAR10a] Narciso, M. "Regressão Linear Múltipla". Capturado em: <http://condicaoinicial.com/2010/06/regressao-linear-multipla.html/all/1>, Novembro de 2010.
- [NAR10b] Narciso, M. "Regressão Linear Múltipla - Eliminação de Dados Espúrios". Capturado em: <http://condicaoinicial.com/2010/09/regressao-linear-multipla-eliminacao-de-dados-espurios.html/all/1>, Novembro de 2010.
- [NEŠ00] Nešković, A.; Nešković, N.; Paunović, Đ. "*Indoor Electric Field Level Prediction Model Based on the Artificial Neural Networks*", IEEE Communications Letters, vol. 4-6, Jun 2000, pp. 190-192.
- [NIK08] Nikitin, P.; Rao, K. "Antennas and Propagation in UHF RFID Systems". In: IEEE International Conference on RFID, 2008, pp. 277-288.
- [POL10] Polar Electronic Industries Pty Ltd. "Cables & Accessories: Model RG58". Capturado em: [http://www.polarelectronicindustries.com/model.php?model\\_id=1173](http://www.polarelectronicindustries.com/model.php?model_id=1173), Dezembro de 2010.
- [QTR10] QTrac. "QTrac - RFID Solutions". Capturado em: [http://www.qtrac.com/Solutions\\_Hotspot.asp](http://www.qtrac.com/Solutions_Hotspot.asp), Dezembro de 2010.
- [RAM03] Ramakrishnan, K. N. M. "Performance Benchmarks for Passive UHF RFID Tags", Dissertação de Mestrado, College of Engineering, Anna University at India, 2003, 100p.
- [RAP94] Rappaport, T. S.; Sandhu, S. "*Radio-wave propagation for emerging wireless personal-communication systems*", IEEE Antennas and Propagation Magazine, vol. 36-5, Out 1994, pp. 14-24.

- [RFI09] RFIDBr. "Saint Paul lança crachá RFID de alta frequência". Capturado em: <http://www.rfidbr.com.br/index.php/saint-paul-lanca-cracha-rfid-de-alta-frequencia.html>, Julho de 2010.
- [RID10] Rida, A.; Yang, L.; Tentzeris, M. "RFID-Enabled Sensor Design and Applications". Norwood: Artech House Publishers, 2010, 210p.
- [ROS04] Rosa, H. N. S. "Redes Neurais na Previsão de Séries Temporais". In: I Workcomp-Sul, 2004, 12p.
- [RUS95] Russell, S; Norvig, P. "Artificial intelligence: a modern approach". New Jersey: Prentice-Hall, 1995, 1ª edição, 932p.
- [SCH05] Schafer, T. M.; Maurer, J.; Hagen, J.; Wiesbeck, W. "*Experimental characterization of radio wave propagation in hospitals*", IEEE Transactions on Electromagnetic Compatibility, vol. 47-2, Mai 2005, pp. 304-311.
- [SIL98] Silva, L. N. C. "Análise e Síntese de Estratégias de Aprendizado para Redes Neurais Artificiais", Dissertação de Mestrado, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 1998, 248p.
- [SIL10] Silva, M. "boxplot - Usando R em Auditoria". Capturado em: <http://sites.google.com/site/marcosfs2006/boxplot>, Novembro de 2010.
- [SPI09] Spiegel, M. R.; Stephens, L. J. "Estatística". Porto Alegre: Bookman, 2009, 4ª edição, 597p.
- [SPS10a] SPSS Inc. "IBM SPSS Statistics". Capturado em: <http://www.spss.com/software/statistics/>, Novembro de 2010.
- [SPS10b] SPSS Inc. "SPSS Neural Networks 17.0". Capturado em: <https://www.washington.edu/uware/spss/docs/SPSS%20Neural%20Network%2017.0.pdf>, Dezembro de 2010.
- [XAV09] Xavier, A. P.; Coimbra, A. R.; Oliveira, S. S. T.; Rodrigues, V. J. S.; Vaz, T. R.; Freitas, L. A. "Ferramentas e técnicas para reduzir o custo de implantação de um sistema RFID". In: IV Simpósio RFID, 2009, 8p.
- [ZHA03] Zhang, J. T.; Huang, Y. "Investigation of building dielectric properties and the effects on indoor radio channel". In: Twelfth International Conference on Antennas & Propagation (ICAP), 2003, pp. 207-211.

## ANEXO A – Diagrama de radiação da antena MT-262010/TRH/A

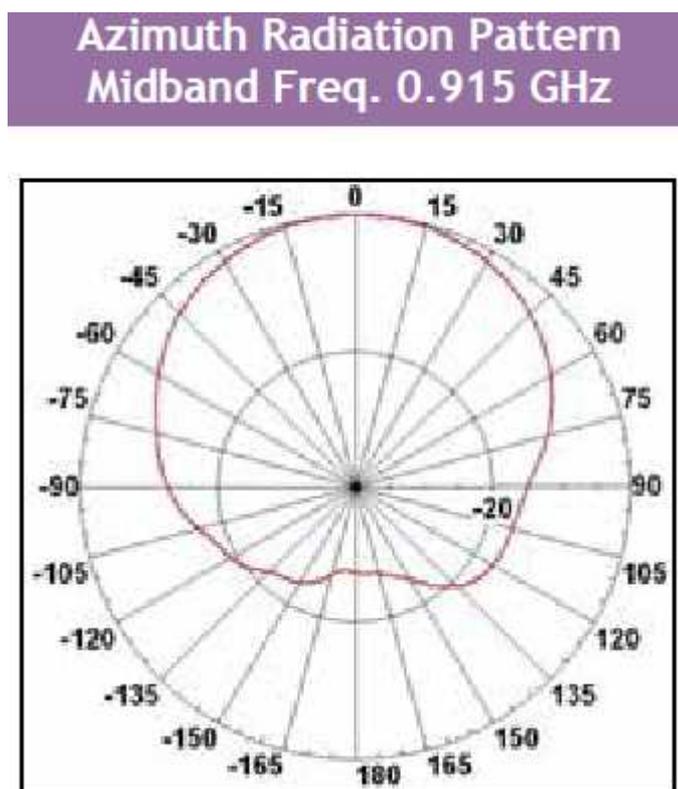


Figura 25 – Diagrama de radiação no plano H (horizontal) da antena MT-262010/TRH/A [MT110].

O diagrama de radiação de uma antena pode ser definido como a representação de uma função matemática que descreve determinadas características deste sistema físico real. As propriedades de radiação nestes diagramas são representadas em função de coordenadas espaciais (cartesianas ou esféricas) e podem mostrar o comportamento do campo, intensidade, ganho, distribuição da radiação, etc.

Na Figura 25, o diagrama mostra o lóbulo de radiação da antena MT-262010/TRH/A. O lóbulo de radiação, peculiar a cada tipo de antena, quando representado em um diagrama mostra as direções em que uma antena pode transmitir ou receber sinais eletromagnéticos e correspondentes características. Um lóbulo é delimitado por regiões de relativas intensidades de radiação e são classificados como principal, laterais ou traseiro.

Ao se considerar este diagrama como o de uma Antena denominada de A, o lóbulo de radiação principal (em direção a  $0^\circ$ ) mostra que a energia radiada para frente, e entre  $-90^\circ$  a  $90^\circ$  é maior do que a radiação alcançada na parte de trás a partir de  $-91^\circ$  e  $91^\circ$ .

Pode-se dizer que, ao se colocar uma segunda Antena B no ponto  $0^\circ$  do diagrama os sinais recebidos nesta direção serão máximos pela Antena B. Quando a Antena B for deslocada em torno da Antena A os sinais recebidos irão diminuir à medida que aumentar a distância do centro do lóbulo principal, ou seja, a energia radiada diretamente para frente pela Antena A é maior do que a energia radiada pelos seus lóbulos laterais e traseiro. Neste caso, pode-se dizer que a antena MT-262010/TRH/A é uma antena que radia a energia eletromagnética de forma direcional.