

VINICIUS RUIZ MARTINS

**IDENTIFICAÇÃO AUTOMÁTICA DE SITUAÇÕES DE EMERGÊNCIA
ATRAVÉS DE TÉCNICAS DE FUSÃO DE SINAIS VITAIS E DE
MOVIMENTOS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Engenharia Elétrica, na área de concentração em Automação.

Orientador: Prof. Dr. Teodiano Freire Bastos Filho.

Co-orientador: Prof. Dr. Rodrigo Varejão Andreão.

VITÓRIA
2008

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

M386i Martins, Vinicius Ruiz, 1980-
Identificação automática de situações de emergência através de técnicas de fusão de sinais vitais e de movimentos / Vinicius Ruiz Martins. – 2008.
113 f. : il.

Orientador: Teodiano Freire Bastos Filho.
Co-Orientador: Rodrigo Varejão Andreão.
Dissertação (mestrado) – Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Teoria bayesiana de decisão estatística. 2. Simulação (Computadores). 3. Inteligência artificial. 4. Sistemas de suporte de decisão. I. Bastos Filho, Teodiano Freire. II. Andreão, Rodrigo Varejão. III. Universidade Federal do Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

VINICIUS RUIZ MARTINS

**IDENTIFICAÇÃO AUTOMÁTICA DE SITUAÇÕES DE EMERGÊNCIA
ATRAVÉS DE TÉCNICAS DE FUSÃO DE SINAIS VITAIS E DE
MOVIMENTOS**

Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisição parcial para a obtenção do Grau de Mestre em Engenharia Elétrica - Automação.

Aprovada em 15/04/2008.

COMISSÃO EXAMINADORA

Prof. Dr. Teodiano Freire Bastos Filho
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Rodrigo Varejão Andreão
Universidade Federal do Espírito Santo
Co-orientador

Prof. Dr. Evandro Ottoni Teatini Salles
Universidade Federal do Espírito Santo

Prof. Dr. Jugurta Rosa Montalvão Filho
Universidade Federal de Sergipe

Dedico este trabalho à minha família.

Agradecimentos

Durante a longa jornada que foi este Mestrado, muitos foram os que colaboraram, seja de maneira direta ou indireta, para a elaboração deste trabalho. Assim, não citando nomes (seriam muitos!), àqueles que me apoiaram deixo registrado aqui os meus sinceros agradecimentos.

Mas se faz necessário um agradecimento especial a algumas pessoas:

- à meus pais: estes são realmente especiais. Sempre me apoiando, esta é uma torcida que vale por mil.
- ao meu irmão do coração e minha cunhadinha Kekel. Esses dois são “gente boa”!
- à minha namorada Keilinha, que esteve do meu lado e aturou o meu mau humor em alguns (vários!) momentos. Te amo môa!
- ao professor Teodiano, que desde a minha iniciação científica vem me apoiando em meus projetos, se dedicando na obtenção de bons resultados. Agradeço todo o aprendizado que me foi passado e, claro, a sua amizade.
- ao professor Rodrigo, que me indicou ao estágio na França e que, por várias vezes, se dispôs a sanar minhas inúmeras dúvidas, com muito interesse e competência. Muito obrigado por esta oportunidade inesquecível que me foi concedida.
- aos professores Evandro e Jugurta, pela aceitação do convite para participar da banca de avaliação, e pela valorosa contribuição na melhoria deste trabalho.
- aos professores do Departamento de Engenharia Elétrica da UFES e aos alunos dos laboratórios LAI 1 e LAI 2, sempre dispostos a fornecer ajuda quando necessário.
- aos meus amigos que me viram “sumir” por um bom tempo.

Agradeço também ao FACITEC pela bolsa de estudos concedida no Brasil, e ao *Projet RNTS – TANDEM*, pela bolsa de pesquisa concedida no *Institut National des Télécommunications – INT*, na França.

Meu muito obrigado a todos que fazem parte do grupo de discussão Bayes Net Toolbox, que me forneceu um suporte de grande valor durante o desenvolvimento dos softwares.

Je voudrais remercier à Jérôme Boudy, qui m'a reçu avec beaucoup de zèle et d'hospitalité. Votre aide a été d'extrême importance pour le développement de ce travail.

J'aimerais remercier encore aux Mesdames Bernadette Dorizzi, Marie-Thérèse Courcier et Patricia Fixot pour la réception ; à François Steenkeste et à Jean-Louis Baldinger pour le développement des plates-formes des capteurs fixe et mobile ; à l'Imad Belfeki, à Hamid Medjahed et à Dan Istrate pour l'aide; et à tous ceux que j'ai eu le plaisir de connaître pendant mon séjour en France.

À tous je remercie l'accueil et la patience que vous avez eu pour comprendre mon Français!

Sumário

Lista de Tabelas	9
Lista de Ilustrações.....	10
Simbologia.....	12
Resumo	13
Abstract	14
CAPÍTULO 1 — INTRODUÇÃO	15
1.1 Descrição da Área.....	15
1.2 Motivação	18
1.3 Definição do Problema	20
1.4 Metodologia.....	21
1.5 Estrutura deste Documento.....	23
CAPÍTULO 2 — ESTADO DA ARTE.....	24
2.1 Telemonitoramento de Pacientes em Domicílio.....	24
2.1.1 <i>Trabalhos Existentes</i>	26
2.2 Princípio de Resolução	27
2.3 Doença de Alzheimer e Irregularidades da Frequência Cardíaca.....	29
2.3.1 <i>Doença de Alzheimer</i>	29
2.3.2 <i>Irregularidades da Frequência Cardíaca</i>	30
CAPÍTULO 3 — FUSÃO DE SINAIS USANDO REDES BAYESIANAS.....	33
3.1 Fusão de Sinais	33
3.2 Redes Bayesianas	36
3.2.1 <i>Introdução</i>	36
3.2.2 <i>Por que Utilizar as Redes Bayesianas?</i>	41
3.2.3 <i>Algoritmo Junction Tree</i>	44
3.2.3.1 <i>Algoritmo de Pearl – Message Passing</i>	46
3.2.4 <i>Aprendizagem</i>	49

CAPÍTULO 4 — O PROJETO TELEPAT.....	51
4.1 Introdução.....	51
4.2 O Módulo Portátil RFpat.....	52
4.3 O Sistema GARDIEN.....	54
CAPÍTULO 5 — O PROGRAMA DE FUSÃO PROPOSTO.....	58
5.1 Método de Diagnóstico Baseado em Redes Bayesianas.....	59
5.2 Fusão de Dados – Modo ‘Off-line’.....	61
5.2.1 <i>Primeiras Versões</i>	61
5.2.2 <i>Versão Final</i>	64
5.3 Fusão de Dados – Modo ‘On-line’.....	68
5.4 Etapa de Treinamento.....	69
CAPÍTULO 6 — SIMULAÇÃO DE DADOS.....	73
6.1 Definição do Problema.....	74
6.2 O Programa de Simulação Desenvolvido.....	75
6.2.1 <i>Inclinação</i>	77
6.2.2 <i>Movimentação</i>	77
6.2.3 <i>Frequência Cardíaca</i>	78
6.2.3.1 <i>Modelo de Atividade Cardíaca</i>	79
6.2.3.2 <i>Modelo Cosinor</i>	80
6.2.3.3 <i>Perfil Anormal</i>	82
6.2.4 <i>Características do Programa de Simulação</i>	87
6.3 Resultados da Simulação.....	90
CAPÍTULO 7 — EXPERIMENTOS E RESULTADOS.....	95
7.1 Introdução.....	95
7.2 Resultados Obtidos.....	96
7.3 Discussão.....	101
CAPÍTULO 8 — CONCLUSÕES E PERSPECTIVAS.....	103
REFERÊNCIAS BIBLIOGRÁFICAS.....	107

Lista de Tabelas

Tabela 1-1 – Indicadores Populacionais 1991/2000 – Brasil (IBGE, 2006).	19
Tabela 3-1 – Vantagens comparativas das Redes Bayesianas (NAİM, 2004).	44
Tabela 4-1 – As diferentes medidas fornecidas pelo RFpat.	53
Tabela 5-1 – Estados dos sinais dos sensores.....	65
Tabela 5-2 – Dados de aprendizagem modificados.....	71

Lista de Ilustrações

Figura 1-1 – População brasileira de 80 anos ou mais de idade por sexo – 1980-2050 (IBGE, 2008).....	19
Figura 2-1 – Sistema de informação do telemonitoramento de pacientes em domicílio.....	25
Figura 2-2 – Etapas de resolução de um problema de decisão.....	28
Figura 2-3 – Exemplos de irregularidades na frequência dos batimentos cardíacos.....	31
Figura 3-1 – Esquema conceitual de fusão de dados.....	33
Figura 3-2 – Exemplo de estrutura de uma Rede Bayesiana.....	39
Figura 3-3 – Exemplo de um ciclo fechado.....	39
Figura 3-4 – Moralização de um grafo.....	45
Figura 3-5 – Triangulação e construção de uma <i>junction tree</i>	46
Figura 3-6 – Coleta e distribuição das mensagens – protocolo <i>message passing</i>	46
Figura 3-7 – Processo de inferência.....	48
Figura 4-1 – Sistema TELEPAT.....	51
Figura 4-2 – Implantação do sistema GARDIEN em um quarto de hospital.....	54
Figura 4-3 – Conexão dos sensores por ligação a cabo.....	55
Figura 4-4 – Plano esquemático da sala de testes do INT – sistema GARDIEN.....	56
Figura 4-5 – Área de atuação dos sensores infravermelhos instalados no teto.....	57
Figura 5-1 – Esquema do processo de fusão de dados.....	59
Figura 5-2 – Estrutura da rede - Primeira versão.....	62
Figura 5-3 – Estrutura da rede - Segunda versão.....	64
Figura 5-4 – Estrutura da rede - versão final.....	65
Figura 5-5 – Tabelas de probabilidades para os nós de entrada da rede: (a) Inclinação ('I'), (b) Movimento ('M'), (c) Pulso ('P') e (d) B5.....	66
Figura 5-6 – Tabelas de probabilidade condicional para os nós intermediários da rede: (a) Bradicardia, (b) Taquicardia e (c) Queda.....	66
Figura 5-7 – Tabela de probabilidade condicional para o nó Controle.....	67
Figura 5-8 – Arquivo texto de saída da rede.....	67
Figura 5-9 – Gráfico de probabilidades do nó Controle.....	68

Figura 5-10 – Exemplo de saída da fusão em modo <i>on-line</i> - (a) relatório de classificação, (b) gráfico de probabilidades do nó CONTROLE.	69
Figura 5-11 – Esquema do processo de aprendizagem da rede.	72
Figura 6-1 – Organograma simplificado do programa de simulação.	76
Figura 6-2 - Método Cosinor - análise do ritmo senoidal.	81
Figura 6-3 – Parâmetros de simulação (bpm x horas) – exemplo de bradicardia.	83
Figura 6-4 – Programa de simulação – aquisição de dados.	87
Figura 6-5 – Modelo de arquivo gerado pela simulação.	88
Figura 6-6 – Dados de inclinação de um arquivo normal já registrado.	91
Figura 6-7 – (a) Simulação de agitação com valores <i>default</i> e (b) dados de agitação de um arquivo normal já registrado.	91
Figura 6-8 – Simulação de frequência cardíaca normal utilizando o modelo Cosinor.	92
Figura 6-9 – Simulação de taquicardia utilizando a criação de um novo perfil.	92
Figura 6-10 – Simulação de bradicardia utilizando características do sinal normal tomado como base – (a) modo ‘adicionar’, (b) modo ‘substituir/permanecer’.	93
Figura 6-11 – Exemplo de arquivo de saída e gráficos correspondentes.	94
Figura 6-12 – Resultado da simulação – janela de comando do Matlab.	94
Figura 7-1 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 1.	96
Figura 7-2 – Arquivo texto com resultado da rede - Cenário 1.	97
Figura 7-3 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 2.	97
Figura 7-4 – Arquivo texto com resultado da rede - Cenário 2.	98
Figura 7-5 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 3.	99
Figura 7-6 – Arquivo texto (resumido) com resultado da rede - Cenário 3.	99
Figura 7-7 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 4.	100
Figura 7-8 – Arquivo texto (resumido) com resultado da rede - Cenário 4.	100
Figura 7-9 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 5.	101
Figura 7-10 – Arquivo texto (resumido) com resultado da rede - Cenário 5.	101

Simbologia

bpm	Batimentos por minuto
$P(X)$	Probabilidade da observação X
$P(X Y)$	Probabilidade da observação X dado a verossimilhança Y
$P(Y X)$	Probabilidade <i>a posteriori</i>
$P(X, Y)$	Probabilidade conjunta de X e Y
E	Conjunto de evidências
I	Sensor de inclinação do sistema RFpat
M	Sensor de movimento do sistema RFpat
P	Sensor de frequência cardíaca do sistema RFpat
B e 5	Sensores de inclinação de sistema GARDIEN
$B5$	Variável representativa dos sensores B e 5
f_d	Frequência de início
t_d	Tempo de início
f_a	Frequência de chegada
t_a	Tempo de chegada
t_{du}	Duração da anomalia
t_{an}	Duração da transição anomalia/normal
et	Ruído aditivo gaussiano

Resumo

O objetivo principal do presente trabalho é o desenvolvimento de uma plataforma multimodal para monitoramento de pessoas idosas e/ou com problemas cardíacos, sendo também objeto do trabalho o desenvolvimento de um programa de simulação de dados fisiológicos.

É proposto um sistema para identificação automática de situações de urgência baseado no emprego de redes probabilísticas (com o uso da técnica Rede Bayesiana) para a fusão de sinais vitais e de movimento, fornecidos por um sistema de telemonitoramento de pacientes em domicílio, chamado TELEPAT. São mostradas as vantagens de se utilizar redes probabilísticas, além da metodologia utilizada para se obter a classificação dos sinais dos sensores. Finalmente, o emprego desse sistema como uma importante ferramenta para auxiliar pacientes idosos ou pacientes com problemas cardíacos é demonstrado através de experimentos.

Para a constituição de bases de dados fisiológicos sintéticos foi desenvolvido um programa de simulação de dados, a partir de sinais normais perturbados artificialmente, de acordo com os perfis de situações alarmantes. Os sinais simulados podem ser utilizados nas etapas de treinamento e teste da Rede Bayesiana para fusão de dados, além de servirem para outros programas, com outras finalidades.

Os programas, tanto de simulação quanto de fusão de dados, foram desenvolvidos em ambiente Matlab.

Abstract

The main aim of this thesis is the development of a multimodal system for monitoring aged people and/or people with cardiac diseases, as well as the development of a physiological data simulation software.

A system for automatic identification of urgency situations is proposed, which is based on the application of probabilistic networks (with the use of the Bayesian Network) for the vital signals and movement fusion. These signals are supplied by a telemonitoring system of patients in domicile, called TELEPAT. Besides the methodology used to get the classification of the sensor signals, the advantages of using probabilistic networks are shown. Finally, the application of this system as an important tool to assist aged or patients with cardiac diseases is demonstrated by experiments.

For the formation of synthetic physiological databases, the development of a data simulation software by using normal signals disturbed artificially (in accordance with the profiles of alarming situations) was done. The simulated signals can be used in the training and test stages of the Bayesian Network for data fusion, and also attend a number of software, with other purposes.

The data simulation and data fusion software had been developed in Matlab environment.

CAPÍTULO 1 – INTRODUÇÃO

O trabalho efetuado nesta Dissertação de Mestrado se situa no contexto do telemonitoramento de pacientes em domicílio, que é uma área importante da telemedicina moderna.

Este capítulo inicial tem por objetivo principal situar o contexto da pesquisa, apresentando o estado da arte, a descrição da área, a definição do problema, e a metodologia utilizada para a realização deste trabalho.

1.1 Descrição da Área

Segundo uma resolução do Conselho Federal de Medicina, “telemedicina é o exercício da medicina através da utilização de metodologias interativas de comunicação audiovisual e de dados, com o objetivo de assistência, educação e pesquisa em saúde. Isto engloba ações como videoconferência médica, teleconsulta, telediagnóstico por imagem, educação à distância e formação continuada, estudos de caso e trabalho colaborativo com uso de recursos de rede, entre outros” (AGÊNCIA CT, 2006).

De acordo com a Rede Universitária de Medicina – RUTE, e de acordo com a Organização Mundial de Saúde - OMS, “telemedicina compreende a oferta de serviços ligados aos cuidados com a saúde, nos casos em que a distância é um fator crítico; tais serviços são prestados por profissionais da área da saúde, usando tecnologias de informação e de comunicação para o intercâmbio de informações válidas para diagnósticos, prevenção e tratamento de doenças e a contínua educação de prestadores de serviços em saúde, assim como para fins de pesquisas e avaliações” (RUTE, 2008).

A telemedicina, no contexto deste trabalho, é uma aplicação tecnológica que representa a utilização de novas tecnologias de informação e da comunicação dentro do setor médico. Ela intervém na ação médica interpondo uma ferramenta de comunicação entre os médicos ou entre um médico e seu paciente.

A telemedicina não substituirá jamais o contato imediato médico-paciente, mas vem se juntar às ferramentas do médico para fornecer um melhor atendimento ao paciente, representando um suporte considerável para melhorar os cuidados e as condições de vida de muitas pessoas.

Os primeiros programas de telemedicina foram adotados nos anos 60 a 70 por países de grandes extensões, onde a densidade populacional é pequena, para responder aos problemas de isolamento geográfico de certas populações. Esta iniciativa buscava uma solução à dificuldade de acesso aos centros de saúde especializados. Tomando como exemplo algumas das primeiras experiências desenvolvidas, temos na Austrália o "controle psicoterapêutico à distância", na Escócia o "dermatologia e medicina à distância para as plataformas petroleiras", e o "tele-cuidado" nas zonas rurais dos Estados Unidos (DUCHÈNE, 2004).

A telemedicina tem encontrado atualmente variados campos de aplicação, e se divide em diferentes termos, o que torna difícil determinar uma topologia unânime (ISTRATE, 2003). Serão apresentadas cinco categorias de aplicação em telemedicina:

- Telemonitoramento: registro telemétrico, geralmente em domicílio, de parâmetros fisiológicos ou do ambiente ou do comportamento de um paciente, transmitido em seguida aos profissionais da saúde responsáveis.
- Teleconsulta: exame de um paciente ou análise de dados de interesse sem interação física direta. Podem-se distinguir dois tipos de teleconsulta: (1) ou o paciente consulta, de sua própria iniciativa, um médico por uma rede de comunicações interposta entre as duas partes; (2) ou o médico consultado solicita um aviso de diagnóstico (telediagnóstico) ou terapêutico (tele-avaliação), sendo essa solicitação endereçada a um outro médico situado à distância (tele-imagem, teleradiologia).
- Tele-assistência: ajuda terapêutica levada diretamente ao paciente, à distância, consequência possível da teleconsulta.
- Telecirurgia: chamada também de telemanipulação, é a manipulação de material médico (instrumentos cirúrgicos) controlado remotamente por um profissional da saúde, para a realização de cirurgias à distância em um paciente. Em muitos casos a robótica tem sido uma aliada da telecirurgia. Cada vez mais são utilizados robôs para essa tarefa, sendo esses robôs controlados por um médico à distância.

- Tele-educação ou educação à distância: utilização de uma ferramenta computacional específica para ajudar na formação contínua dos médicos, tendo como exemplos: contatos profissionais via redes, consulta de informações médicas (banco de dados, imagens, acompanhamento de estudos epidemiológicos e de testes clínicos), consulta de cursos de formação e videoconferências nas universidades (tele-aprendizagem) e reuniões.

Em particular, a aplicação denominada comumente telemonitoramento de pacientes em domicílio é fundamental para melhorar a qualidade de vida de pessoas que necessitam de cuidados especiais ou uma atenção particular. Ela visa a utilização, dentro da casa ou local onde se encontra uma pessoa, de um dispositivo que permita a captura de informações sobre seu estado de saúde, a fim de tornar possível a um profissional da saúde um diagnóstico, e assim definir uma ajuda ao paciente à distância.

De acordo com as categorias precedentes, esta ferramenta de mediação entre um médico e seu paciente deve, portanto, levar em consideração vários elementos:

- *Telemonitoramento*: para o registro de um conjunto de parâmetros referentes ao paciente;
- *Teleconsulta*: à iniciativa, ou do paciente ou dos profissionais da saúde, para a análise dos dados telemonitorados;
- *Tele-assistência*: para fornecer à distância, quando possível, uma ajuda direta ao paciente.

Além do compromisso a certas regras técnicas, clínicas, jurídicas, econômicas ou éticas ligadas à telemedicina em geral, a consideração conjunta desses três elementos dentro do contexto do telemonitoramento de pacientes em domicílio impõe, particularmente, definir um compromisso entre a necessidade de minimizar a coação provocada pelo telemonitoramento ao paciente (sensores pouco numerosos, dispositivos não invasivos) e de maximizar as informações disponíveis para o profissional da saúde de acordo com o seu método de diagnóstico e de assistência.

1.2 Motivação

Com o advento de novas tecnologias de informação e de comunicação houve uma expansão das possibilidades em matéria de serviços, sendo que a telemedicina (e a Engenharia Biomédica como um todo) é uma área que tem se beneficiado desse desenvolvimento. O envelhecimento da população, as novas políticas de saúde, e a evolução das atitudes perante as novas tecnologias constituem desafios econômico e social essenciais. O conjunto desses fatores contribui para o desenvolvimento de projetos visando uma melhora na qualidade de vida e também de sistemas de telemonitoramento médico.

Os avanços da medicina permitiram um crescimento mais significativo dos idosos "sêniores". Pesquisa realizada pelo IBGE (Instituto Brasileiro de Geografia e Estatística) mostra que a população brasileira tem envelhecido e que o ritmo de crescimento dos idosos com idade elevada tem sido mais intenso. O segmento de 75 anos ou mais, por exemplo, representava 26,1% da população de idosos brasileiros em 2006. Dez anos antes, eles correspondiam 23,5% (SPITZ, 2007). De acordo com a PNAD (Pesquisa Nacional por Amostra de Domicílios) divulgada em setembro de 2007, pesquisa feita pelo IBGE com o objetivo de investigar diversas características socioeconômicas da população, realizada em todas as regiões do Brasil, a população idosa brasileira correspondia a 19 milhões de pessoas, ou 10% da população brasileira em 2006.

Já de acordo com a pesquisadora Lucia Cunha, da Síntese dos Indicadores Sociais, elaborada pelo IBGE, a população de 80 anos ou mais está crescendo em todos os países. Por exemplo, no Japão a esperança de vida está em quase 80 anos.

A Tabela 1-1 e a Figura 1-1 mostram o resultado obtido do último censo demográfico ocorrido no Brasil, feito pelo IBGE. Nota-se que, além do crescimento da população idosa com os anos, há uma razão de dependência relativamente baixa por parte dos idosos (9,07% no ano 2000). Isso contribui no interesse em fornecer soluções que possam facilitar a manutenção da saúde dessa população idosa como, por exemplo, o atendimento e/ou monitoramento em domicílio. Também, com o aumento do número de idosos, a perspectiva é o aumento do número de doenças e complicações pertinentes a essa faixa etária, e ainda um crescimento no número de atendimentos médicos e atendimentos de urgência.

Levando em consideração o atual quadro médico e a situação dos hospitais no Brasil, que normalmente é precária, principalmente em relação aos estabelecimentos destinados ao atendimento público gratuito, é de grande interesse o desenvolvimento de mecanismos ou ser-

viços que ajude a aliviar a demanda de profissionais e/ou leitos em hospitais, além de dotar de uma melhor qualidade os atendimentos intensivos realizados em hospitais.

Tabela 1-1 – Indicadores Populacionais 1991/2000 – Brasil (IBGE, 2006).

Indicadores	1991	2000	Taxa de Crescimento (%)
População			
0 a 14 anos	50.988.432	50.266.122	-0,16
%	34,73	29,60	
15 a 64 anos	88.751.196	109.597.948	2,39
%	60,45	64,55	
15 a 24 anos	28.582.350	34.081.330	1,99
%	19,47	20,07	
65 anos ou mais	7.085.847	9.935.100	3,86
%	4,83	5,85	
Total	146.825.475	169.799.170	1,64
%	100,00	100,00	
Razão de Dependência (%)			
Total	65,43	54,93	
Jovens	57,45	45,86	
Idosos	7,98	9,07	

Um sistema de monitoramento beneficiaria o acompanhamento de pacientes sob o uso de medicamentos, como antibióticos, antialérgicos, antidepressivos e antiarrítmicos, principalmente se o paciente for acometido por alguma doença que, de alguma maneira, reage de forma danosa ao interagir com algum tipo de medicamento.

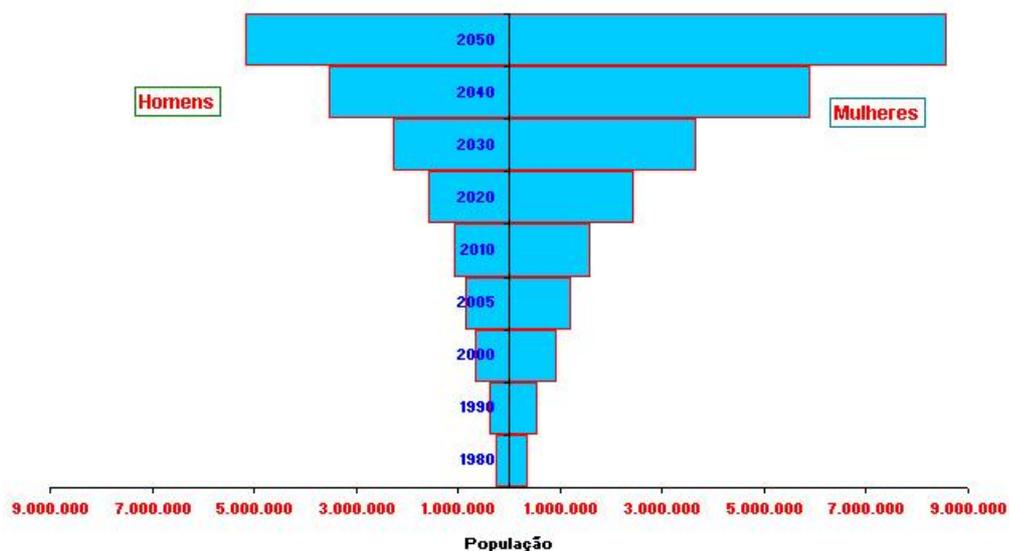


Figura 1-1 – População brasileira de 80 anos ou mais de idade por sexo – 1980-2050 (IBGE, 2008).

Ainda, um sistema desse tipo, além de realizar o monitoramento de situações alarmantes, poderá também ser utilizado para fazer o registro de sinais biológicos do paciente, como a frequência cardíaca por exemplo, para ser utilizado posteriormente na análise do quadro clínico desse paciente por um médico. A medição dos batimentos cardíacos do paciente ao longo do dia poderá servir como método de diagnóstico de problemas relacionados ao coração. Um registro gráfico, de certa maneira, pode ser uma ferramenta muito importante em uma avaliação médica, podendo ajudar na escolha do tratamento do paciente.

Devido à importância da análise dos dados biológicos do paciente, é de grande interesse a criação de um sistema de análise autônomo que auxilie o médico no diagnóstico do paciente e na avaliação do seu histórico de dados biológicos e de movimentos que, por vezes, resulta em uma análise de horas, realizada por um profissional da saúde, o que é um trabalho repetitivo e passível de erros.

1.3 Definição do Problema

Pretende-se desenvolver um programa de fusão de dados biológicos, obtidos a partir de sensores posicionados no corpo do paciente, que apresente as seguintes características:

- 1) Desempenho satisfatório para ser utilizado em um paciente em domicílio;
- 2) Facilidade de uso por uma pessoa não especialista;
- 3) Eficiência de execução;
- 4) Facilidade de manutenção/adaptação;
- 5) Aproveitamento do código fonte do programa para aprimoramentos e/ou inclusão de outras funções;

Além disso, um programa de simulação deve ser implementado para fornecer dados de teste para o programa de fusão de dados, além de servir para outros propósitos que envolvam os mesmos tipos de dados biológicos.

O programa de fusão de dados deverá ser capaz de receber os dados advindos tanto de sensores (no caso sensores piroelétricos) já instalados em um ambiente quanto de sensores biométricos ligados ao paciente, de manipular esses dados e fornecer uma resposta relacionada à atual condição do paciente (se houve uma queda do paciente ou se há alguma irregularidade relacionada aos batimentos cardíacos desse paciente).

As estruturas dos programas de fusão e de simulação de dados desenvolvidos devem levar em consideração a inclusão de outras funções no futuro. Dessa forma, todas as funções devem ser programadas de forma que haja uma modulação do programa final.

É muito importante que as estruturas dos programas sejam de fácil entendimento, na medida do possível. Não do ponto de vista de programação, mas de organização e de encadeamento das idéias, já que futuramente essas estruturas serão utilizadas por outros programadores. Além disso, uma documentação adequada deve ser fornecida.

A respeito da manutenção e/ou adaptação do programa de fusão de dados, como o mesmo será utilizado para o monitoramento de um paciente em estado de observação médica, se for necessário realizar uma manutenção ou adaptação, quanto mais rápido for esse processo, melhor será o atendimento ao paciente. Assim, justifica-se disponibilizar uma documentação de qualidade, de fácil leitura, e o desenvolvimento de um código-fonte de fácil entendimento.

O funcionamento do programa deve ser o mais simples e intuitivo possível, já que será utilizado por uma pessoa não especialista, ou por um paciente em seu domicílio.

1.4 Metodologia

O objetivo principal deste trabalho consiste no estudo e na implementação de uma Rede Bayesiana para estabelecer um processo de decisão no contexto de um sistema de telemonitoramento de paciente em domicílio. A meta é tornar a identificação remota de uma situação de emergência (relativa a um paciente sozinho em sua casa) mais adaptável e robusta para um profissional da saúde em um centro de atendimento.

O trabalho desenvolvido foi dividido nas seguintes etapas:

- Estudo de bibliografias relacionadas às Redes Bayesianas;
- Implementação de uma Rede Bayesiana e de técnicas de processamento dos dados dos sensores relativos ao terminal portátil do paciente;
- Realização de melhorias no processo de decisão por meio dos recursos de flexibilidade e auto-adaptação, que caracteriza a Rede Bayesiana;
- Desenvolvimento e implementação de um simulador de sinais de frequência cardíaca para um banco de dados multimodal.

Foi necessário realizar uma pesquisa bibliográfica, no decorrer do trabalho, relacionada aos sinais biológicos utilizados, como forma de manipular e adaptar a rede de maneira mais real possível (BIBLIOMED, 2005 e KADI, BELFEKI, 2005).

O simulador foi parte integrante, de profunda importância, deste trabalho, pois foram necessários sinais extras para teste da rede, já que não havia disponibilidade de um grande conjunto de dados para essa tarefa. Além disso, o simulador permite o teste de várias condições possíveis, o que faz com que o mesmo seja utilizado em trabalhos futuros que requeiram um banco de dados de sinais de taquicardia e bradicardia.

Os trabalhos efetuados no curso desta Dissertação de Mestrado foram realizados no *Institut National des Télécommunications* – INT, na França, e na Universidade Federal do Espírito Santo – UFES, por meio de uma cooperação UFES/INT. Esses trabalhos se situaram no contexto dos projetos TANDEM e TELEPAT, programas do RNTS (*Réseau National Technologies pour la Santé*), financiado pela *French National Research Agency*, os quais têm como objetivo a realização de um serviço de telemonitoramento em domicílio de pessoas idosas e/ou que sofram de alguma patologia cardíaca. No caso específico do trabalho desta Dissertação de Mestrado, o objeto da pesquisa foi propor e implementar soluções para automação e processamento de sensores para situações que envolvam casos de pré-Alzheimer, ou seja, a idéia inicial foi a de trabalhar na adaptação de um dispositivo de monitoramento remoto que envolva sinais biológicos de pessoas que estejam na fase inicial dessa doença. Essa idéia é viável, já que os meios de tratamento disponíveis para essa doença são limitados e de alto custo.

Os parceiros desse projeto são: o INT, líder do projeto e responsável pelos desenvolvimentos eletrônicos e de processamento dos dados precedentes à transmissão ao centro de telemonitoramento, o ESIEE (*Ecole Supérieure d'Ingénieurs en Electroniques et Electrotechniques*), que é o responsável pela parte de informática e transmissão sobre IP, o ISTM (*Institut Supérieur de Technologie et de Management*), incumbido da resolução dos problemas de adaptação em relação com o SAMU 92 (Serviço de ajuda médica de urgência francês), o INSERM (*Institut National de la Santé et de la Recherche Médicale*) encarregado do estudo de implantação de um espaço de medição em um ambiente médico, e o UTT (*Université de Technologie de Troyes*) responsável pelo estudo de questões relacionadas à qualidade de serviço.

Além do desenvolvimento do simulador de sinais de frequência cardíaca para o presente projeto, foi realizada também uma parceria com o ESIGETEL (*École Supérieure d'Ingénieurs en Informatique et Genie des Télécommunications*) para o desenvolvimento de uma ferramen-

ta para criação de um banco de dados biológicos relativos à frequência cardíaca, além de uma parceria com o hospital Broca, de Paris, para os projetos relacionados a problemas que envolvam mal de Alzheimer.

1.5 Estrutura deste Documento

Este capítulo introduziu a descrição da área que será abordada, apresentando o conceito de telemedicina, apresentou os problemas envolvidos e as motivações que levaram à idealização desse trabalho.

O Capítulo 2 trata da definição de telemonitoramento de pacientes em domicílio, o princípio de resolução de um sistema desse tipo e apresenta diversos trabalhos existentes na área. Nesse capítulo são definidas também as irregularidades da frequência cardíaca abordadas neste trabalho.

No Capítulo 3 são discutidos os conceitos de fusão de sinais e Redes Bayesianas, para situar o leitor no contexto do trabalho. São descritas as vantagens de se utilizar as Redes Bayesianas, além de expor detalhes construtivos importantes que definem uma Rede Bayesiana.

O Capítulo 4 apresenta o projeto TELEPAT, responsável pela realização do serviço de telemonitoramento proposto por essa Dissertação. São detalhados os subsistemas que o compõe: o terminal portátil RFpat e o sistema fixo GARDIEN.

No Capítulo 5 está detalhado o programa de fusão de dados desenvolvido, o qual utiliza a técnica Rede Bayesiana descrita no Capítulo 3. Este programa realiza a fusão dos sinais de saída dos sensores dos sistemas RFpat e GARDIEN, discutidos no Capítulo 4.

O programa de simulação de dados desenvolvido é apresentado no Capítulo 6, juntamente com a definição do problema e as vantagens de se realizar um sistema desse tipo. São apresentadas as características desse programa, com a descrição dos sinais de saída utilizados, além de detalhar as opções disponíveis pelo programa para cada um desses sinais. Ao final são exibidos os resultados obtidos por este programa.

No Capítulo 7 são apresentados os resultados obtidos pelos experimentos realizados com o programa de fusão de dados.

Finalmente, o Capítulo 8 conclui o trabalho, apresentando suas vantagens e sugerindo trabalhos futuros.

CAPÍTULO 2 – ESTADO DA ARTE

2.1 Telemonitoramento de Pacientes em Domicílio

O trabalho de pesquisa efetuado no curso desta Dissertação de Mestrado se situa dentro do contexto do telemonitoramento de pacientes em domicílio, que representa uma das dimensões da telemedicina. Esta aplicação leva em consideração elementos de telemonitoramento, de teleconsulta e de tele-assistência.

A idéia principal desses sistemas é permitir às pessoas viverem em suas casas o maior tempo possível, com maior independência, em um ambiente de conforto e com segurança, o que tende a diminuir a ansiedade do paciente. Esses sistemas têm por finalidade detectar e prevenir a ocorrência de situações críticas no domicílio ou uma degradação do estado de saúde de uma pessoa (CASTELLI, 2001; VIRONE, 2003). Esses sistemas representam, assim, uma alternativa momentânea ou durável à hospitalização ou aos recursos nos estabelecimentos para acomodação de longa duração (casas de repouso, asilos ou centros especializados). O paciente não será, portanto, forçado a abandonar seu lar e sua vida social. Ele terá uma grande autonomia em seu ambiente social e privativo, todos se beneficiando dos serviços preventivos de saúde. Em particular, esses sistemas têm uma relação de interesse maior com os idosos, mas, em geral, com pessoas que apresentam riscos de enfermidades motoras (que possam sofrer quedas, por exemplo) ou cognitivas (depressão, demência senil, etc.), ou que necessitam de cuidados ou uma atenção excepcional (diabéticos, asmáticos, etc.).

Outra contribuição desse tipo de sistema de telemonitoramento é um atendimento mais rápido e eficaz no caso de uma urgência real, graças ao fornecimento contínuo dos parâmetros fisiológicos do paciente, à distância (BOUDY, BALDINGER, 2007). A identificação confiável de uma situação de urgência limita os falsos alarmes que contribuem na saturação dos centros de chamadas de urgência médica (SAMU – Serviço de Atendimento Móvel de Urgência).

O telemonitoramento de um paciente em domicílio se baseia em um sistema de informação global que abrange os seguintes elementos (Figura 2-1):

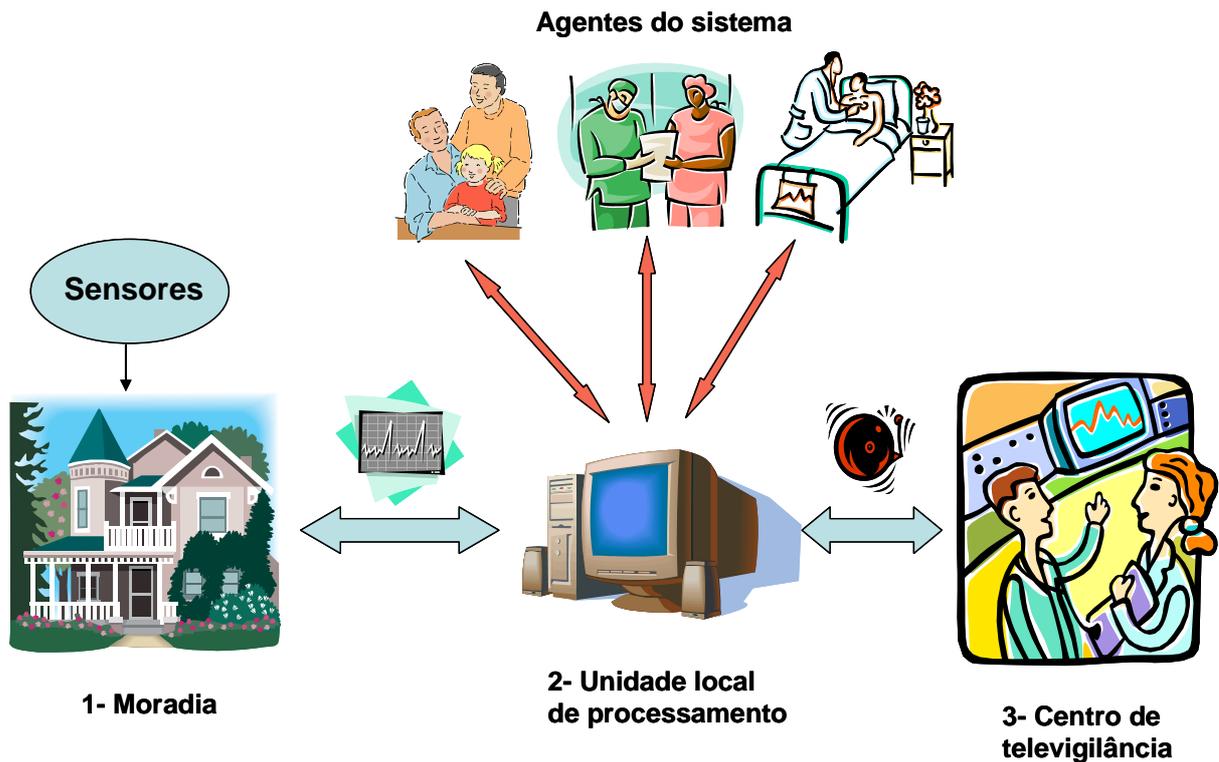


Figura 2-1 – Sistema de informação do telemonitoramento de pacientes em domicílio.

1- Um conjunto de sensores de diferentes tipos (fisiológicos, no ambiente, de movimento, etc.) instalados dentro da moradia ou transportado pela pessoa, conectado em rede para a coleta dos dados em tempo real, e de equipamentos automáticos, para adaptar o habitat da pessoa às suas capacidades pessoais, motrizes e cognitivas;

2- Uma unidade local de processamento, situada na moradia do paciente, responsável pelo armazenamento e processamento dos sinais recebidos pelos sensores, pelo gerenciamento de uma base de conhecimento relativa à pessoa telemonitorada, e pela emissão de mensagens de alarme.

3- Um centro de televigilância para o tratamento das mensagens e alarmes recebidos das áreas telemonitoradas (onde ocorre a interpretação dos dados e tomadas de decisão).

Um conjunto de agentes (grupo médico, pessoa telemonitorada e membros de sua família) pode então ter acesso a todo momento, após autenticação e conforme seus privilégios, aos dados do sistema, no nível da unidade local de tratamento.

2.1.1 Trabalhos Existentes

Existem diversos trabalhos, desenvolvidos e em desenvolvimentos pelo mundo, que são relacionados a sistemas de monitoramento médico de pacientes. Resumidamente, a seguir, serão mostrados alguns trabalhos desenvolvidos.

- HTCMS: O *Home Telecare Management System* é um protótipo desenvolvido para testar a possibilidade de uso de um sistema de monitoramento em um ambiente real. A ferramenta de classificação utiliza a técnica Rede Neural (RODRÍGUEZ, et al., 1994).
- CarerNet: referente a uma arquitetura auxiliar que estabelece as necessidades e requerimentos dos sistemas de telemonitoramento, promove a integração de sistemas para identificar a natureza dos dispositivos de monitoramento e a inteligência requerida para o sistema. (WILLIAMS, et al., 1998).
- SmartBo: Analisa a concepção de casas adaptadas às exigências do telemonitoramento médico. São criados cenários simulando moradias de indivíduos fictícios com diferentes problemas. As complicações encontradas são listadas e analisadas (ELGER, FURUGREN, 1998).
- Monitoring Health Status: Propõe um monitoramento em domicílio sem perturbar as atividades diárias do usuário. Envolve, por exemplo, o monitoramento durante o sono e no banheiro, por meio de sensores (OGAWA, TOGAWA, 2000).
- Monitoring Behavior in Home: Desenvolve um habitat inteligente, relacionado à saúde do usuário que realiza um monitoramento remoto do comportamento diário do paciente em domicílio por meio de sensores no ambiente e por um dispositivo detector de queda (NOURY, et al., 2000).
- TISSAD: Tem por objetivo principal a definição de uma arquitetura genérica, modular e aberta para os sistemas de telemonitoramento, adaptável a diversas patologias tratadas em domicílio (THOMESSE, et al., 2001).
- AID HOUSE: Analisa a concepção de casas adaptadas às exigências do telemonitoramento médico. Explora a possibilidade do uso de alguns dispositivos para controle automático de portas, janelas, cortinas, luzes, etc., e monitora a movimentação dos ocupantes da casa para identificar potenciais situações de emergência (BONNER, 1998).
- Monitoring of Daily Physical Activity: Neste trabalho é proposto um sensor cinemático portátil para o monitoramento da atividade física de uma pessoa pela detecção das posturas do corpo e pelos períodos de caminhada (NAJAFI, et al., 2003).

- HAT: O *Home Asthma Telemonitoring* é um sistema destinado especificamente aos pacientes que sofrem de asma. O sistema estabelece uma troca de informações constantes entre o paciente e a central médica, para permitir uma decisão mais eficiente (FINKELSTEIN, et al., 2001).
- Home Health Care: Desenvolve uma aplicação baseada na tecnologia Java para um sistema de monitoramento de diabéticos (LIND, et al., 2001).
- Distress Sound Extraction System: É implementado um sistema capaz de detectar e identificar eventos sonoros em condições ruidosas, sendo essa informação utilizada conjuntamente com outros sensores (fisiológicos, de posição, etc.) para realizar um monitoramento médico à distância (ISTRATE, et al., 2006).
- E-Vanguard for Emergency: É um sistema que ajuda as operações de resgate e nos serviços de primeiros socorros. Consiste em dotar os médicos e responsáveis pelo resgate com PDAs e laptops com GPS, que se comunicam com o subsistema (composto por sensores fisiológicos) transportado pela pessoa ferida (CHEN, et al., 2003).

2.2 Princípio de Resolução

O conteúdo deste capítulo é baseado na Tese de Doutorado de Duchène, que trata de um sistema de telemonitoramento médico de pessoas a domicílio (DUCHÈNE, 2004).

Em toda problemática de decisão, é indispensável especificar bem os objetivos e o contexto da tomada de decisão, para assegurar uma boa adequação dos sistemas implementados à resolução do problema proposto. Esta etapa é particularmente crítica dentro do contexto de problemáticas complexas, havendo vários níveis de detalhes entre os conhecimentos disponíveis, os dados registrados e os objetivos da decisão. É o caso do presente projeto, pois são disponíveis dados de “baixo nível” (os sinais dos sensores), para uma tomada de decisão de “alto nível” (a identificação de situações de queda e condicionamento cardíaco de uma pessoa). Faz-se necessário, portanto, avaliar adequadamente as necessidades, exigências e restrições do sistema, a fim de que as decisões correspondam da melhor maneira aos objetivos.

As etapas de um processo de decisão, que se inserem dentro de um esquema recursivo, são apresentadas a seguir (ver Figura 2-2):

- Definição do contexto e dos objetivos da decisão: o objetivo é o de especificar os tipos de dados, os conhecimentos e as informações necessárias a considerar em relação aos objetivos do processo de decisão. Algumas questões importantes que podem ajudar nesta etapa são, por exemplo: Quais são as observações pertinentes necessárias dentro do contexto experimental? Qual o nível de detalhe a ser considerado para as observações? Qual o desempenho considerado para o sistema de decisão?
- Registro ou geração de dados apropriados ao contexto da decisão: o desenvolvimento de um processo experimental para a coleta de dados adaptados ao problema de decisão é conduzido pelos objetivos e pelo contexto da decisão, definidos na etapa anterior, de resolução de problemas. Porém, a aquisição de conjuntos de dados suficientemente completos e representativos para um sistema de decisão pode ser difícil para certas aplicações. Para fazer frente à falta de dados experimentais, o desenvolvimento de um processo de simulação se torna útil como uma primeira etapa de resolução do problema. Variando-se os parâmetros da simulação podem-se obter dados representativos para um conjunto de situações possíveis, que não são simples de observar na realidade, em um processo experimental. O desempenho do sistema de decisão poderá ser avaliado de forma mais completa graças à simulação.
- Teste dos métodos adaptados para a resolução do problema de tomada de decisão: os dados adquiridos ou gerados pela simulação, para a experimentação, são utilizados como entradas do processo de decisão implementado para a resolução do problema. A sensibilidade e a especificidade dos algoritmos experimentados devem corresponder às exigências de resolução do problema definido durante a primeira etapa.

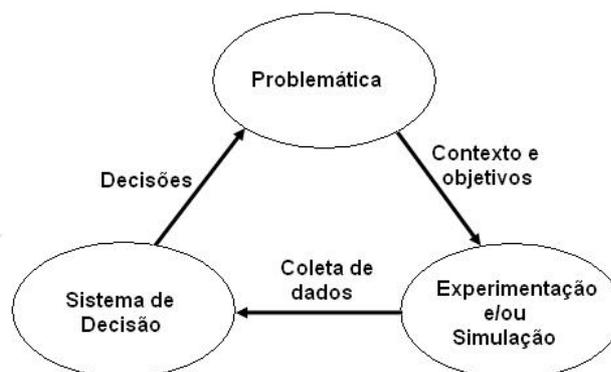


Figura 2-2 – Etapas de resolução de um problema de decisão.

Uma vez que o algoritmo de decisão tenha sido implementado e testado, a adequação entre o desempenho da decisão e os objetivos de resolução permite determinar se uma próxima etapa de refinamento do processo de decisão será necessária ou não. A solução do problema de decisão deve também levar em consideração as etapas de validação das escolhas efetuadas e dos resultados obtidos em cada fase do processo, a fim de assegurar constantemente a sua adequação ao contexto e aos objetivos de decisão.

2.3 Doença de Alzheimer e Irregularidades da Frequência Cardíaca

Uma das aplicações da plataforma de telemonitoramento em domicílio é o monitoramento de pacientes que estejam no estágio que precede a doença de Alzheimer, chamado de pré-Alzheimer. Além disso, o sistema pode ser empregado no monitoramento de pessoas idosas e/ou que sofram de alguma patologia que cause variações anormais nos batimentos cardíacos, como a bradicardia e a taquicardia.

A doença de Alzheimer e as irregularidades da frequência cardíaca são detalhadas nas seções seguintes.

2.3.1 Doença de Alzheimer

A doença (ou mal) de Alzheimer, conhecida como demência senil do tipo Alzheimer e chamada também de esclerose ou caduquice, é uma doença degenerativa do cérebro, caracterizada por uma perda das faculdades cognitivas superiores, manifestando-se inicialmente por alterações da memória episódica. Estas faltas amnésicas agravam-se com a progressão da doença, e são posteriormente acompanhados por faltas visuo-espaciais e de linguagem. O início da doença pode muitas vezes dar-se com simples alterações de personalidade, com ideação paranóide (que é um transtorno de personalidade em que ocorre desconfiança e suspeita quanto aos outros, de modo que seus motivos são interpretados como malévolos). Os indivíduos com o transtorno supõem que as outras pessoas os exploram, prejudicam ou enganam, ainda que não exista qualquer evidência apoiando esta idéia. Eles suspeitam, com base em poucas

ou nenhuma evidências, que os outros estão conspirando contra eles e que poderão atacá-los subitamente, a qualquer momento e sem qualquer razão (BALLONE, 2007).

Este tipo de demência apresenta uma complexa etiologia, ou seja, as causas que contribuem para a ocorrência dessa doença podem ter várias origens, onde fatores genéticos e ambientais influenciam na patogênese (ou origem da doença). A idade, o sexo, a poliproteína E4, a avançada idade paterna, a doença cerebrovascular, o trauma craniano, o infarto do miocárdio e os efeitos imunológicos também são citados na literatura como sendo fatores de risco para a doença de Alzheimer. Alterações cromossômicas, fatores ambientais, agentes infecciosos, toxinas, grau de escolaridade, e ainda a ocupação podem também estar relacionados. Esta doença degenerativa do sistema nervoso central é caracterizada por perda gradual e progressiva da memória e outras atividades cognitivas, presença de distúrbios psicóticos como depressão, alucinações e até mesmo comportamentos agressivos. Geralmente há deterioração dos movimentos motores, desorientação (o que pode ocasionar quedas), mudança de personalidade, depressão e comportamento bizarro (SILVA, GALLO, et al., 2007). Existem outras doenças que estão associadas ao mal de Alzheimer, como a ansiedade, osteoporose (também responsável por quedas), diabetes e hipertensão arterial (ALZHEIMER MED, 2007).

2.3.2 Irregularidades da Frequência Cardíaca

Outro ponto abordado neste trabalho, de uma forma mais ampla, é um método de monitorar a frequência cardíaca de um paciente em domicílio. Para isso, faz-se necessária uma análise sobre as irregularidades da pulsação cardíaca (ou frequência cardíaca), sendo tratados aqui dois tipos de perturbações: a bradicardia e a taquicardia. O monitoramento da frequência cardíaca se faz necessário, inclusive, durante um tratamento com medicamentos que acarretem na mudança da frequência cardíaca. Além disso, a medição da frequência cardíaca pode ser utilizada como um mecanismo simplificado, na ausência de ECG, para a identificação de insuficiência cardíaca (HIDEAKI, 2005).

Arritmia cardíaca (ou disritmia cardíaca) é o nome genérico de diversas perturbações que alteram a frequência e/ou o ritmo dos batimentos cardíacos. A arritmia é uma alteração do ritmo cardíaco em relação ao normal. As arritmias podem levar à morte e constituir, por isso, um caso de emergência médica. A maior parte delas é, no entanto, inofensiva. O nódulo sinusal, na aurícula direita do coração, é um grupo de células que regula esses batimentos através de impulsos elétricos que estimulam a contração do músculo cardíaco ou miocárdio. Quando

esses impulsos elétricos são emitidos de forma irregular ou conduzidos de forma deficiente, pode ocorrer arritmia cardíaca. Na Figura 2-3 tem-se um exemplo de irregularidades na frequência dos batimentos cardíacos. Esta figura mostra registros reais em papel milimetrado, onde o eixo das ordenas representa o nível de tensão do sinal (padrão onde 10 mm representa 1 mV), e cada divisão do eixo do tempo (abscissas) equivale a 0,2 segundos.

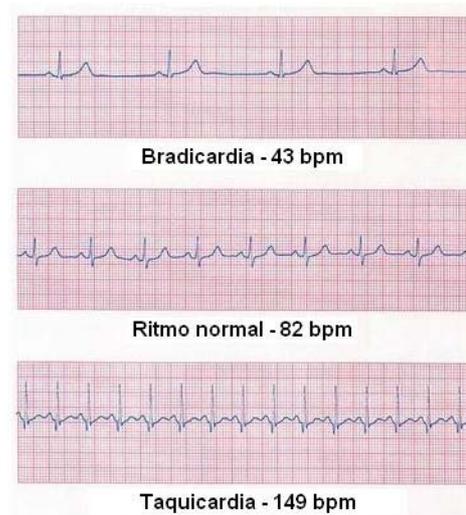


Figura 2-3 – Exemplos de irregularidades na frequência dos batimentos cardíacos.

Bradicardia é um termo utilizado na Medicina para designar uma arritmia em que haja diminuição na frequência cardíaca. Convencionou-se como normal, em um ser humano adulto, uma frequência cardíaca (ou batimentos cardíacos) entre 60 e 100 bpm (batimentos por minuto). Portanto, as frequências cardíacas abaixo de 60 bpm constituem a bradicardia (OLSON, 1998). Em algumas pessoas pode ser uma ocorrência normal, como em atletas. São conhecidos vários tipos de bradicardia, cada um com suas características peculiares. Os marca-passos cardíacos são utilizados no tratamento desse tipo de arritmia (BIBLIOMED, 2005).

A bradicardia pode ocorrer devido a variações normais do funcionamento do organismo (neste caso chamada de bradicardia fisiológica), ou devido a alguma doença (neste caso bradicardia patológica). A bradicardia pode causar baixa pressão, devido à diminuição do esforço do músculo cardíaco, e pode causar também danos ao organismo devido a menor circulação do sangue, levando menos oxigênio às células. Podem ocorrer tonturas e desmaios, que são os principais sintomas da bradicardia.

Já a taquicardia é um termo médico utilizado para designar uma arritmia em que haja um aumento da frequência cardíaca. Existem vários tipos, algumas extremamente graves. A

partir de 100 bpm, inclusive, considera-se que há taquicardia no ser humano. Assim como para a bradicardia, a taquicardia pode ocorrer devido a variações normais do funcionamento do organismo (neste caso chamada de taquicardia fisiológica, que ocorre normalmente durante atividades físicas e situações de estresse emocional), ou devido a alguma doença (neste caso taquicardia patológica, que pode acontecer, por exemplo, na presença de anemia). No caso da taquicardia durante a realização de esforços, como por exemplo os exercícios físicos, decorrido alguns minutos o batimento cardíaco deve voltar a uma frequência normal. Na taquicardia, o principal sintoma é a palpitação. Note-se que taquicardia não é sinônimo de ataque cardíaco (EYE, 2007).

Muitas arritmias são completamente assintomáticas, e a grande parte da população apresenta alguns episódios arrítmicos durante o dia e nem se dá conta disso. A verdade é que existem alguns tipos de arritmias que estão mais associados à ocorrência de sintomas do que outros, podendo depender também de vários outros fatores: frequência dos episódios, frequência cardíaca atingida durante a arritmia, presença de doença cardíaca prévia, entre outros. O sintoma mais comum é a palpitação (também conhecido como "batedeira"), que pode ocorrer tanto nas bradicardias quanto nas taquicardias. Algumas pessoas são bastante sensíveis e apresentam grande desconforto na presença desse sintoma. Outro sintoma comum é a síncope (ou desmaio) caracterizada pela recuperação imediata e espontânea. O indivíduo pode sentir também falta de ar, mal-estar, e outros sintomas que vão depender da presença ou não de outras doenças. Algumas vezes, quando a arritmia é mais grave, o paciente pode apresentar confusão mental, fraqueza, hipotensão (pressão baixa), dor no peito (angina), caracterizando uma emergência médica e levando à aplicação de tratamento imediato para evitar a morte do paciente.

Entretanto, a arritmia é uma condição patológica. Uma arritmia pode se expressar sob a forma de uma taquicardia ou bradicardia. Assim, nem sempre quando ocorre uma bradicardia ou taquicardia deve-se, necessariamente, haver uma arritmia.

Na abordagem de fusão de sinais deste trabalho estão envolvidos tanto sinais de bradicardia como de taquicardia, não importando a origem dessas perturbações.

CAPÍTULO 3 – FUSÃO DE SINAIS USANDO REDES BAYESIANAS

Neste capítulo é abordado o conceito de fusão de sinais, utilizado no processamento das informações obtidas do ambiente de vida do paciente. Os sinais que serão fusionados são fornecidos pelos sensores do sistema TELEPAT, e o processamento é realizado no servidor local.

Foi utilizada a técnica Rede Bayesiana para realizar a fusão de sinais, devido às suas características favoráveis ao problema proposto, as quais são descritas na seção 3.2.2.

3.1 Fusão de Sinais

Fusão de dados é geralmente definida como o uso de técnicas que combinam informações obtidas de uma ou várias fontes de naturezas diferentes para se chegar a uma interpretação de uma situação, ou seja, combina múltiplos dados para produzir uma informação coerente e mais completa ao usuário. Esta última é muito mais confiável e precisa do que obter a informação desejada a partir de apenas um tipo de informação fornecida por somente um sensor isolado. A Figura 3-1 ilustra um processo de fusão de dados.

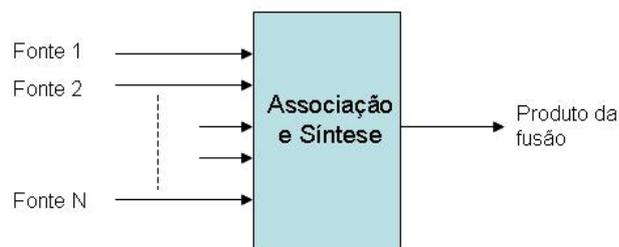


Figura 3-1 – Esquema conceitual de fusão de dados.

Associação é a correlação do assunto tratado por todas as fontes de informação de entrada, e a síntese é a combinação dessa informação em um único produto da fusão (KESSEL, 2006).

O processo de fusão de dados é caracterizado por contínuos refinamentos de suas estimativas e avaliação da necessidade de fontes adicionais, ou modificação do processo propriamente dito, para alcançar um melhor resultado.

Existem muitas vantagens em se utilizar uma técnica de fusão de dados (WALD, 2002):

- Robustez e confiança: o sistema continua operacional mesmo se uma ou várias fontes de informação são perdidas ou apresentam mau funcionamento;
- Extensão da cobertura no tempo e espaço: as informações são coletadas em um intervalo de tempo menor devido o uso de várias fontes de sinais;
- Aumento da dimensão do espaço de dados: para aumentar a qualidade da informação deduzida, além de reduzir a vulnerabilidade do sistema;
- Ambigüidade reduzida: uma informação mais completa fornece uma melhor discriminação entre as hipóteses disponíveis;
- Provisão de soluções para a grande quantidade de informações que são disponíveis atualmente.

Se vários sensores são usados, a combinação das observações fornecerá um aumento da estimativa do resultado. Por exemplo, se sensores diferentes são utilizados para monitorar a posição do paciente, a combinação dos sinais de saída dos mesmos irá fornecer um resultado mais preciso sobre uma possível queda do paciente. Uma vantagem estatística é conseguida pela adição de N observações independentes, assumindo que os dados são combinados de uma maneira ótima. Esse mesmo resultado pode também ser obtido pela combinação de N observações de um único sensor. Ainda, alguns sinais podem, em parte ou mesmo totalmente, ser corrompidos, por exemplo, por uma interferência de sinais de radiofrequência ou pela deterioração do sistema. Neste caso, é indispensável que o sistema de recepção e fusão de dados possa efetuar uma combinação com outras modalidades de telemonitoramento, tais como sensores de infravermelho (para determinação da posição da pessoa no ambiente – sistema GARDIEN). Isto é então objeto de uma fusão entre estas diversas modalidades sensoriais, de forma que o sistema seja capaz de tomar uma decisão eficiente e robusta ante situações de distúrbios do ambiente ou de limites materiais.

Outra vantagem envolve o uso de posições relativas ou deslocamento de múltiplos sensores para aumentar o processo de observação. Por exemplo, dois sensores que medem a posi-

ção de um objeto podem ser ajustados para determinar a posição de um objeto por triangulação (HALL, LLINAS, 2001).

A técnica de fusão de sinais é uma aplicação bastante difundida e utilizada em vários domínios como, por exemplo, nas áreas militares, comerciais e, particularmente, por um grande número de sistemas biológicos.

Freqüentemente, em problemas de fusão de dados de sensores, cada sensor possui resoluções temporal e espacial diferentes, e cada um deve resolver o problema da correspondência, ou seja, decidir quais eventos de um sensor correspondem aos mesmos eventos relatados pelos outros sensores. Nesse caso, mesmo com a falta de dados de um dos sensores, o uso de um sistema baseado na fusão de dados, por exemplo, pode permitir a tomada de decisão fazendo inferência sobre os dados desconhecidos (o que é conhecido como *missing data*). Além disso, visto que cada sensor tem somente uma possibilidade limitada de dar uma interpretação correta ao problema, a combinação das possibilidades de todos os sensores tende a aumentar a probabilidade de uma interpretação válida.

O programa de fusão de dados pode ter dois tipos diferentes de construção da arquitetura de *hardware* e *software*: do tipo centralizado e do tipo descentralizado. Na arquitetura do tipo centralizada, uma unidade de processamento central única fica encarregada de executar todo o processo de fusão de dados necessário. Este tipo de arquitetura agrega um melhor custo benefício por utilizar um único processador ao invés de vários, em comparação com a arquitetura do tipo descentralizada. Entretanto, o modelo de arquitetura centralizado oferece uma confiabilidade menor que o modelo de arquitetura descentralizado, uma vez que uma falha na unidade de processamento central compromete todo o sistema. A arquitetura do tipo descentralizada (ou distribuída) é baseada em sistemas sensoriais e sistemas controladores distribuídos, interligados em rede, na qual nenhum sistema individual está em nível hierárquico superior a qualquer outro. Desta forma, o sistema se torna mais robusto, confiável e mais tolerante a falhas, em comparação com a abordagem centralizada. Nesse último caso, a falha de um subsistema não implica, necessariamente, na falha do sistema global.

O modelo de arquitetura utilizado no sistema desenvolvido é do tipo misto, centralizado de uma parte e descentralizado de outra parte, pois cada sistema (RFpat e GARDIEN) tem sua unidade de processamento (que realiza algum tratamento dos sinais antes que eles sejam transmitidos à unidade de processamento central), mas a fusão de dados é executada em uma única unidade.

Existem diversas técnicas para se realizar a fusão de dados, destacando-se: Filtros de Kalman (WELCH, 2006 e HARRIS, 2002), Redes Neurais (FREEMAN, 1991 e BELFEKI,

2006), Modelos Ocultos de Markov - HMM (FÉRON, 2003 e MÜLLER, 2006), K-Médias e K-Vizinhos-Mais-Próximos (DUDA, 2001 e BELFEKI, 2006), Árvores de Decisão (FONSECA, 1994 e PAPASTAVROU, 1994), e Redes Bayesianas (JENSEN, 2001 e PESSETE, 2002).

3.2 Redes Bayesianas

3.2.1 Introdução

Antes de começar a descrever uma Rede Bayesiana, faz-se necessário colocar aqui alguns conceitos importantes (SAMPAIO, 2007). Uma variável aleatória é dita discreta quando assume uma distribuição de probabilidades discreta, ou seja, uma variável cujos valores possíveis são limitados ou ilimitados, mas numeráveis. Portanto, se uma variável aleatória é discreta então o conjunto de todos os possíveis valores que ela pode assumir é finito ou contavelmente infinito. Uma distribuição de probabilidade discreta é caracterizada pela sua função massa de probabilidade, ou seja, uma variável X é dita discreta se

$$\sum_u P(X = u) = 1 \quad \text{Eq. 3-1}$$

com u percorrendo todo o conjunto de possíveis valores da variável aleatória discreta X . Uma observação a ser feita é que variáveis contínuas podem ser discretizadas para se adequarem às Redes Bayesianas.

Sendo assim, podemos definir uma variável discreta X com estados finitos x_1, x_2, \dots, x_{N1} ($N1$ estados), e uma variável discreta Y com estados finitos y_1, y_2, \dots, y_{N2} ($N2$ estados).

Uma Probabilidade Condicional é representada como $P(x_i | y_k)$, ou seja, é a probabilidade de x_i dada a ocorrência de y_k . Uma Probabilidade Conjunta (também conhecida como *joint*) é representada como $P(x_i, y_k)$, ou seja, é a probabilidade de ocorrência de x_i e de y_k conjuntamente.

Algumas regras fundamentais são a Marginalização e Regra de Bayes (ou Lei de Bayes):

$$P(y_k) = \sum_{i=1}^{N1} P(x_i, y_k) = \sum_{i=1}^{N1} P(y_k | x_i) \cdot P(x_i) \quad (\text{Marginalização}) \quad \text{Eq. 3-2}$$

$$P(x_i, y_k) = P(x_i | y_k) \cdot P(y_k) = P(y_k | x_i) \cdot P(x_i) \quad \text{Eq. 3-3}$$

$$P(x_i | y_k) = \frac{P(x_i, y_k)}{P(y_k)} = \frac{P(y_k | x_i) \cdot P(x_i)}{P(y_k)} \quad (\text{Lei de Bayes})$$

De uma outra maneira, podemos escrever a Lei de Bayes da forma mostrada na Equação 3-4, que é um Modelo Causal (considera-se as causas gerando efeitos).

$$P(\text{causa} | \text{efeito}) = \frac{P(\text{efeito}, \text{causa})}{P(\text{efeito})} = \frac{P(\text{efeito} | \text{causa}) \cdot P(\text{causa})}{P(\text{efeito})} \quad \text{Eq. 3-4}$$

Para a Lei de Bayes descrita na Equação 3-3, o conjunto de variáveis x_i é considerado como as variáveis aleatórias de distribuição *a priori* $P(x_i)$. A utilização desta distribuição *a priori* permite a introdução de restrições no conjunto de parâmetros da rede, a fim de melhorar a estimação (notadamente no caso onde o conjunto de treinamento é limitado). Utilizando a Lei de Bayes, esta distribuição *a priori* é utilizada para construir a distribuição *a posteriori* $P(x_i | y_k)$, sobre as quais serão baseadas as etapas de treinamento e classificação (BOITE, 2000). $P(y_k | x_i)$ é dita como a verossimilhança da função.

As tabelas com os valores de probabilidade criadas durante a análise de uma Rede Bayesiana seguem os modelos:

- $P(Y)$ é uma tabela com valores $P(Y=y_1), P(Y=y_2), \dots, P(Y=y_s)$, cuja soma desses valores dá o valor 1;
- $P(X | Y)$ é uma tabela com valores $P(X=x_i | Y=y_k)$, com $1 \leq i \leq N_1, 1 \leq k \leq N_2$;
- $P(X, Y)$ é uma tabela com valores $P(X=x_i, Y=y_k)$.

As relações de probabilidade fundamentais para as tabelas são:

$$P(Y) = \sum_X P(X, Y) = \sum_X P(Y | X) \cdot P(X) \quad \text{Eq. 3-5}$$

$$P(X | Y) = \frac{P(Y | X) \cdot P(X)}{P(Y)} = \alpha \cdot P(Y | X) \cdot P(X) \quad \text{Eq. 3-6}$$

onde α é um fator de normalização.

As Redes Bayesianas são um tipo particular de modelos gráficos probabilísticos, resultado do casamento entre a teoria dos grafos e a teoria das probabilidades, utilizado no apoio à decisão em aplicações que envolva inteligência artificial. Podem ser chamadas também de Redes de Crença (*Belief Networks*) ou Redes Probabilísticas (*Probabilistic Networks*). Houve um grande aumento de interesse em se utilizar essa ferramenta, nos últimos anos, devido a fornecer um formalismo eficiente para lidar com as incertezas e as complexidades de um problema prático (HATON, 2006).

Os modelos gráficos probabilísticos são grafos em que os nós representam as variáveis aleatórias, e a falta de um arco entre dois nós supõe uma independência condicional¹ entre as duas variáveis correspondentes. Quando as transições não são orientadas, o modelo gráfico é um *Campo de Markov*, e dois conjuntos de nós A e B são condicionalmente independentes em relação a um terceiro conjunto C somente se todas as transições entre A e B passam por C. Quando os arcos são orientados e o gráfico é acíclico, o modelo gráfico é uma Rede Bayesiana (CASTILLO, et al., 1998, DODIER, 1999 e CHARNIAK, 1991). A vantagem de não permitir ciclos é que se torna possível atualizar rapidamente os algoritmos, visto que não haverá caminho para que o processo de inferência probabilística fique em um circuito fechado indefinidamente.

Os parâmetros de uma Rede Bayesiana são (JENSEN, 2000 e JENSEN, 2001):

- i. A estrutura do modelo, isto é, os nós (representando as variáveis aleatórias) e os arcos (representando as dependências probabilísticas entre as variáveis associadas);
- ii. A distribuição de probabilidade condicional a cada nó (CPD, do inglês *Conditional Probability Distribution*, ou Distribuição de Probabilidade Condicional).

A CPD define a Tabela de Probabilidade Condicional, ou seja, define $P(x(i) | x(Pa(i)))$, sendo $x(i)$ o $i^{\text{ésimo}}$ nó, e $x(Pa(i))$ são os pais do nó i . A CPD quantifica as relações probabilísticas causais entre os nós da estrutura, fornecendo as relações de independência condicionais, listando as probabilidades do nó filho para cada combinação de valores dos seus nós pais. Algumas CPDs serão exemplificadas na seção 5.2.2.

A topologia do grafo codifica as relações de independência entre as variáveis aleatórias. A Figura 3-2 exemplifica uma Rede Bayesiana composta por quatro nós (quatro variáveis aleatórias). A estrutura acíclica de uma Rede Bayesiana é formada por arcos orientados (modelo

¹ Dado um evento C, os eventos A e B são condicionalmente independentes se

$$P(A \cap B | C) = P(A | C) \cdot P(B | C)$$

direto), de maneira que se pode dizer que um nó S_i não depende de todos os nós anteriores, mas apenas de seus nós pais (que são aqueles diretamente ligados ao nó S_i), ou seja:

$$P(S_i | S_1, \dots, S_{i-1}) = P(S_i | Pa(S_i)) \quad \text{Eq. 3-7}$$

sendo $Pa(S_i)$ a representação dos nós pais de S_i (OLIVIER, 2006). De acordo com a Figura 3-2, os nós S_2 e S_3 possuem S_1 como nó pai, o nó S_4 possui S_2 e S_3 como nós pais, e S_1 não possui pais. Nota-se que os arcos são orientados e formam uma estrutura que não possui ciclos fechados.

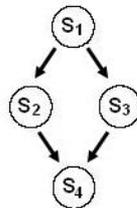


Figura 3-2 – Exemplo de estrutura de uma Rede Bayesiana.

Para cada variável S_i com pais $Pa(S_i) = \{S_1, \dots, S_{i-1}\}$ está associada a tabela $P(S_i | Pa(S_i))$. Esta relação estabelece a independência local entre variáveis da rede. Um arco do nó S_1 para o nó S_2 significa que S_1 é “causa” de S_2 , ou seja, diz-se que S_1 é pai de S_2 .

Se houver algum ciclo direcionado (ver Figura 3-3), não se pode dizer que a estrutura representa uma Rede Bayesiana. No caso de modelos indiretos, a estrutura é formada por arcos não orientados, o que não especifica a representação da variável: se é uma causa ou se é um efeito.

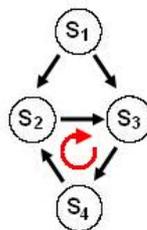


Figura 3-3 – Exemplo de um ciclo fechado.

Uma rede bayesiana pode seguir uma regra Causal e uma regra de Diagnóstico. Quando se fala em uma rede Causal, o sentido dos arcos é das causas para os efeitos (o que leva em consideração que há circunstâncias em que a causa não gera um efeito). Já em uma rede de

Diagnóstico, o sentido dos arcos é dos efeitos para as causas (o que leva em consideração que a causa para um efeito pode ser outra).

Uma Rede Bayesiana pode ser composta por variáveis aleatórias discretas (por exemplo, uma distribuição multinomial), por variáveis aleatórias contínuas (por exemplo, distribuições gaussianas ou exponencial) ou pela mistura das duas. Cada variável que compõe a estrutura tem um conjunto finito de estados mutuamente exclusivos, ou seja, todas as variáveis admitem estar em apenas um estado.

Após a definição da estrutura da rede, é necessário estabelecer as Tabelas de Probabilidades *a priori*, as Tabelas de Probabilidade Conjunta e as Tabelas de Probabilidade Condicional. A Tabela de Probabilidade *a priori* é o conjunto das probabilidades de uma dada variável para cada estado possível, cuja soma das probabilidades deve ser 100%. A Tabela de Probabilidade Conjunta é o conjunto das probabilidades conjuntas de duas ou mais variáveis para cada configuração possível de estados, cuja soma deve ser 100%. A Tabela de Probabilidade Condicional é o conjunto das probabilidades condicionais de uma variável com respeito a uma ou mais variáveis, para cada configuração possível de estados. A soma não necessariamente deve ser 100%.

A evidência sobre uma variável é uma declaração das certezas dos seus estados. Faz-se a evidência de uma variável quando se vai testar a rede de acordo com uma entrada específica. Por exemplo, para a Rede Bayesiana criada para a fusão de dados, os dados dos sensores do dispositivo portátil (que serão as entradas da rede) são evidenciados, já que seus valores são conhecidos.

A inferência realizada em uma Rede Bayesiana é a atualização das probabilidades *a posteriori* das variáveis da rede, baseadas nas evidências disponíveis. É uma função essencial para a finalidade de uma Rede Bayesiana, sendo muito importante para o bom desempenho da rede, e existem vários algoritmos para realizar essa operação.

O aprendizado é uma etapa onde é possível “ensinar” os parâmetros e/ou a estrutura para uma Rede Bayesiana, dado um banco de dados para aprendizado. Entretanto, aprender uma estrutura é uma tarefa muito difícil, devido à sua complexidade. Além disso, também é difícil realizar um aprendizado quando há informação incompleta. Quando a estrutura da rede é conhecida, e a observabilidade é parcial, o método para o aprendizado se baseia no algoritmo EM (do inglês *Expectation Maximization*, ou esperança e maximização).

Resumidamente, uma Rede Bayesiana é um modelo matemático baseado em nós e arcos que representam, respectivamente, as variáveis de um universo $U=(A_1, A_2, \dots, A_n)$ e as dependências entre elas, realizadas com o uso de métodos probabilísticos. Ela nada mais é que uma

representação de uma distribuição de probabilidades, que é representada a partir de conhecimentos explícitos (incorporação do conhecimento do usuário) ou que é aproximada a partir de uma base de dados, e é a partir da distribuição aproximada que são efetuadas as inferências (ou deduções dos valores). Toda previsão resultada de uma Rede Bayesiana é, portanto, por definição, uma probabilidade.

3.2.2 Por que Utilizar as Redes Bayesianas?

De acordo com o tipo de aplicação, a utilização prática de uma Rede Bayesiana pode ser considerada da mesma forma que outros modelos: redes neuronais, sistemas especialistas, árvores de decisão, modelos de análise de dados (regressão linear), modelos lógicos, etc. Naturalmente, para a escolha do método são levados em consideração diferentes critérios, como a facilidade de uso, o custo e o atraso na conclusão de uma solução. Fora todas as considerações teóricas, os seguintes aspectos das Redes Bayesianas lhes rendem, para numerosos casos, vantagens e a preferência a outros modelos (NAÏM, 2004):

- Capacidade de reunir e fundir conhecimentos de naturezas diversas num mesmo modelo (aquisição de conhecimentos): pode-se utilizar dados históricos ou empíricos, conhecimentos de especialista (expresso sobre a forma de regras lógicas, de equações, de estatísticas ou de probabilidades subjetivas), e observações. Por exemplo, dentro do mundo industrial, cada uma das suas fontes de informação, individualmente, é incapaz de fornecer uma representação precisa e realista do sistema analisado;
- Representação de conhecimentos: a representação gráfica de uma Rede Bayesiana é explícita, intuitiva e compreensível por uma pessoa não-especialista, o que facilita a validação do modelo, suas evoluções eventuais e acima de tudo a sua utilização. Ela permite representar as assertivas de independência de forma visual e de fácil percepção. Para ilustrar, uma pessoa não-especialista estará mais encorajada a utilizar um modelo em que ela pode compreender seu funcionamento do que em utilizar um modelo do tipo “caixa preta”, no qual os processos entre as entradas e as saídas ficam desconhecidos pelo usuário;
- Utilização dos conhecimentos: uma Rede Bayesiana é polivalente (ou multifuncional), no sentido da sua utilização. Pode-se servir de um mesmo modelo para avaliar, estimar, prognosticar, diagnosticar, ou otimizar as decisões, o que torna rendoso os esfor-

ços de construção de uma Rede Bayesiana. As Redes Bayesianas podem controlar ou prever o comportamento de um sistema;

- Qualidade na oferta em matéria de programas: existe hoje uma grande gama de programas para iniciar e tratar as Redes Bayesianas. Essas ferramentas apresentam funcionalidades mais ou menos evoluídas: aprendizagem das probabilidades dos nós, aprendizagem da estrutura da rede, possibilidade de assimilar variáveis contínuas, variáveis de utilidade e de decisão (que são tipos de variáveis utilizadas em Redes Bayesianas), etc.

As duas propriedades fundamentais das Redes Bayesianas são, primeiramente, de serem compostas por grafos orientados, isto é, representam causalidades e não apenas simples correlações, e, além disso, as Redes Bayesianas garantem uma correspondência entre a distribuição de probabilidade subjacente e o grafo associado.

As Redes Bayesianas permitem considerar, dentro de um mesmo formalismo, a representação de modelos de causalidades e as estatísticas multivariadas. É dessa maneira que Redes Bayesianas, assim como as técnicas Árvores de Decisão e Redes Neurais, são umas das técnicas mais utilizadas para *data-mining*. *Data-mining* (conhecido também como Extração de Conhecimento ou Análise de Dados/Padrões) é um processo de mineração de dados, ou seja, realiza a extração de informações interessantes (não-trivial, implícita, previamente desconhecida e potencialmente útil) de dados armazenados em grandes massas de dados (SILVA, 2003).

A decisão em utilizar as Redes Bayesianas neste trabalho, para a análise e fusão dos dados do sistema, se deve a ser uma técnica que apresenta qualidades inerentes a uma aplicação de domínio médico: permite análise de grandes quantidades de informações, possui capacidade de juntar e fusionar conhecimentos de naturezas diversas num mesmo modelo e possibilidade de trabalhar as incertezas e complexidades de um problema prático. Além disso, como comentado anteriormente, um ponto importante a se levar em consideração é a sua representação gráfica intuitiva, que o torna compreensível a uma pessoa não especialista e facilita a sua utilização (HIDEAKI, 2005; LUNA, 2004; PINTO, 2005).

Uma propriedade bastante interessante das Redes Bayesianas é que as probabilidades não precisam ser exatas para serem úteis. Probabilidades aproximadas, ainda que sejam supostos valores subjetivos, fornecem bons resultados. Redes Bayesianas são geralmente robustas ao ‘conhecimento imperfeito’.

Outros pontos importantes, que favorecem o uso das Redes Bayesianas são:

- Torna o processo de inferência computacionalmente eficaz;
- Permite a construção de um sistema baseado na opinião de especialistas;
- Possui modelo direto (em comparação com um modelo indireto – utilizado em Redes de Markov):
 - Possui facilidades para construção da estrutura (arco de ligação "A causa B");
 - A aprendizagem e os ajustes dos dados são mais simples, pois existe um condicionamento conhecido entre as variáveis (relações determinísticas);
- Permite analisar a rede do efeito para a causa (*bottom-up*) e da causa para o efeito (*top-down*).

A vantagem do modelo direto (estrutura acíclica formada por arcos orientados) sobre um modelo indireto (estrutura formada por arcos não orientados) está na facilidade e simplicidade na construção da estrutura, o que torna mais simples o aprendizado dessa rede (relacionamentos determinísticos), e torna o processo de inferência eficiente computacionalmente, pois não é necessário calcular a função de distribuição de probabilidades conjunta global, e sim apenas realizar cálculos locais entre um nó e seus pais (PESSETE, 2002). E no caso de aplicações em sistemas *on-line*, como o deste trabalho em questão, a velocidade de processamento é um fator importante.

Do ponto de vista das aplicações, as vantagens e inconvenientes das Redes Bayesianas, de acordo com algumas das técnicas concorrentes, podem se resumir na Tabela 3-1. Estão agrupadas algumas vantagens e inconvenientes de acordo com três tópicos: a aquisição, a representação e a utilização de conhecimentos. Cada linha da Tabela 3-1 corresponde a uma característica, que pode ser uma vantagem ou um problema específico. Caso a técnica considerada leve em consideração o problema (ou vantagem) apresentado, um sinal de * estará adicionado dentro da caixa correspondente. Um sinal de + é adicionado quando a técnica é avaliada como a melhor dentre as listadas na Tabela, do ponto de vista da característica considerada (segundo a avaliação de NAÏM, 2004).

A generalidade do formalismo das Redes Bayesianas, tanto em termos de representação como de utilização, a torna difícil de manipular a partir de certo tamanho da rede. A complexidade das Redes Bayesianas não se traduz somente em termos de compreensão pelos utilizadores. Todos os problemas subjacentes são, praticamente, de complexidade não polinomial, e conduzem ao desenvolvimento de algoritmos aproximados, onde o comportamento não é garantido em problemas que envolvam redes de grandes dimensões.

3.2.3 Algoritmo Junction Tree

Quando a rede bayesiana apresenta ciclos indiretos, tomando como exemplo o esquema mostrado na Figura 3-2, o algoritmo utilizado para a passagem de parâmetros dos nós será aplicado duas vezes, pois as informações de S2 e S3 que chegam a S4 não são independentes, já que necessitam de uma causa comum, S1. Uma aproximação comumente utilizada na passagem de parâmetros dos nós da rede é converter a Rede Bayesiana em uma estrutura de árvore, transformando os nós em *clusters* para formar uma *junction tree*, definir as CPDs para os *clusters* e então utilizar o algoritmo de passagem de parâmetros localmente nessa árvore.

Tabela 3-1 – Vantagens comparativas das Redes Bayesianas (NAÏM, 2004).

Habilidades	Análise de Dados	Redes Neurais	Árvores de Decisão	Sistemas Expertos	Redes Bayesianas
AQUISIÇÃO					
Somente especialista				+	
Somente dados	*	+	*		*
Misto	*	*	*		+
Incremental		*			+
Generalização	*	+	*		*
Dados incompletos		*			+
REPRESENTAÇÃO					
Incerteza				*	+
Legibilidade	*		*	*	+
Facilidade		*	+		
Homogeneidade					+
UTILIZAÇÃO					
Função elaborada	*			*	+
Utilidade econômica	*	*			+
Desempenho	*	+			

Junction tree (que pode também ser chamado de *Clique Trees*, *Tree Decomposition* ou *Join Trees*) é um conceito vindo da teoria dos grafos, que representa um importante papel em problemas que envolvam inferência probabilística. Esse método é usado em *machine learning* para realizar uma inferência (marginalização) exata em gráficos em geral, obtendo uma sequência ótima de eliminação de variáveis e cria uma estrutura para propagar as multiplicações e marginalizações das tabelas.

O problema a ser resolvido na etapa de inferenciação é calcular a probabilidade condicional de um nó, ou de um conjunto de nós, dados os valores observados de outro conjunto de nós. A idéia do algoritmo *junction tree* é a de decompor o cálculo global das probabilidades

conjuntas em um conjunto interligado de cálculos locais. O procedimento desse algoritmo está resumidamente descrito a seguir (PASKIN, 2003; XU, ZHU, 2004 e HUANG, DARWICHE, 1994):

1. Moralização: converte um grafo direto em um grafo indireto. O processo consiste em conectar os nós que têm filhos em comum e tornar os arcos não-direcionados (como mostra a Figura 3-4).
2. Triangulação e construção da *junction tree*: esse processo determina a ordem de escolha das variáveis do gráfico para a construção da *junction tree*. De acordo com a Figura 3-5-a, a ordem de eliminação adotada, para definir um gráfico de *clusters*, foi *f, d, e, c, b* (a linha pontilhada entre os nós C e B indica a o processo de triangulação realizado entre três nós – B, C e E.). O gráfico completo (Figura 3-5-b) é obtido após encontrar o peso dos *clusters*, que é definido pelo número de separadores (representado pelos retângulos) em cada interligação.
3. Propagação das probabilidades: se baseia no protocolo *Message-Passing* (ver seção 3.2.3.1), que realiza a propagação das probabilidades (inferência). De acordo com esse protocolo, um *cluster* B envia uma mensagem para seu vizinho C somente depois que ele tenha recebido mensagens de todos os vizinhos exceto C. Uma forma de realizar essa tarefa é, após escolhido um *cluster* R para ser a origem, executar os procedimentos Coletar(R) e Distribuir (R) (como mostrado na Figura 3.6):
 - Coletar: para cada filho B de R, chamar iterativamente o procedimento Coletar(B) e então passar uma mensagem de B para R;
 - Distribuir: para cada filho B de R, passar uma mensagem para B e então iterativamente chamar Distribuir(B).

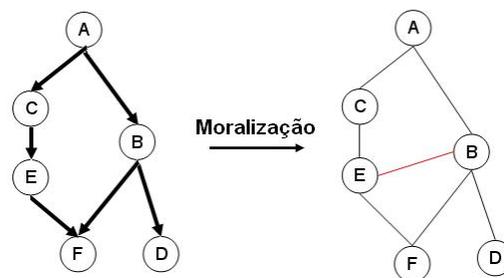


Figura 3-4 – Moralização de um grafo.

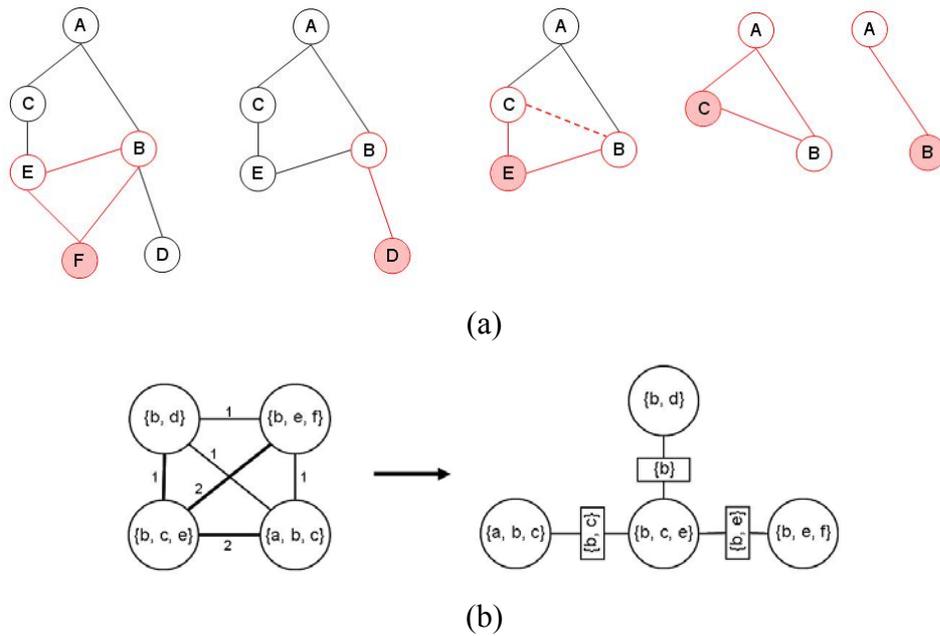


Figura 3-5 – Triangulação e construção de uma *junction tree*.

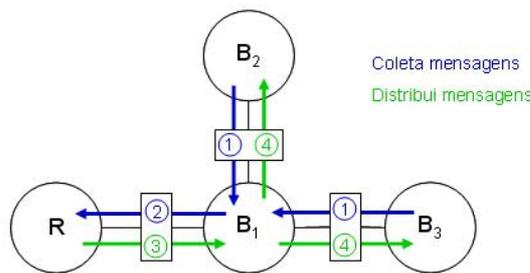


Figura 3-6 – Coleta e distribuição das mensagens – protocolo *message passing*.

3.2.3.1 Algoritmo de Pearl – Message Passing

A seguir será mostrado um resumo do funcionamento do algoritmo de Pearl, que realiza a passagem dos parâmetros da rede.

O objetivo da inferência é calcular $P(X | E)$ para um conjunto de variáveis X e E , com E representando um conjunto de observações (evidências) e X representando as variáveis não-observadas. Adotando E_i^- como a evidência acessível através dos filhos de X_i , e E_i^+ como a evidência acessível através dos pais de X_i , $P(X | E)$ pode ser escrito como (SINGLIAR, 2003):

$$\begin{aligned}
 P(X | E) &= P(X | E_i^+, E_i^-) = \alpha \cdot P(E_i^+, E_i^- | X) \cdot P(X) = \\
 &= \alpha \cdot P(E_i^- | X) \cdot P(E_i^+ | X) \cdot P(X) = \alpha \cdot P(E_i^- | X) \cdot P(X | E_i^+) \\
 P(X | E) &= \alpha \cdot \lambda(X) \cdot \pi(X)
 \end{aligned}
 \tag{Eq. 3-8}$$

sendo α uma constante de normalização. Os parâmetros de passagem do algoritmo são:

$$\lambda(X_i) = P(E_i^- | X) \quad \text{Eq. 3-9}$$

$$\pi(X_i) = P(X | E_i^+) \quad \text{Eq. 3-10}$$

O algoritmo de inferência utiliza o princípio de decomposição executado sobre a variável X_i para estimar a probabilidade de uma variável dada a observação. O algoritmo atua da seguinte maneira (HATON, et al., 2006):

- Calcula os $\lambda(X_i)$ de uma maneira recursiva para todos os X_i começando pelos filhos;
- Transmite estes valores a partir dos filhos;
- Calcula $\pi(\text{filhos})$ em função dos λ ;
- Propaga $\pi(\text{filhos})$ para os outros nós e calcula a probabilidade da variável correspondente em função da observação.

A Figura 3-7 ilustra esse processo. Para o cálculo de uma mensagem- π , a informação necessária é a CPD correspondente ao nó pai. Tomando como exemplo dois nós X e Z , sendo Z um nó pai de X , para se obter π basta ter a CPD de Z . Assumindo dois nós X e Y , sendo Y um nó filho de X , com Y observado, calculando a lei de Bayes:

$$P(X | Y) = \frac{P(X) \cdot P(Y | X)}{P(Y)} = \alpha \cdot P(X) \cdot P(Y | X) \quad \text{Eq. 3-11}$$

a informação não disponível $P(Y | X)$ é passada pela mensagem- λ . Adotando $\lambda_{XY}(x)$ como a mensagem- λ enviada entre X e Y , e $\pi_{XY}(x)$ como a mensagem- π enviada entre X e Y , com y_1, \dots, y_c representando os nós filhos de X e z_1, \dots, z_p representando os nós pais de X , as equações seguintes são utilizadas para calcular λ e π :

$$\lambda(x) = \prod_{j=1}^c \lambda_{y_j, X}(x) \quad \text{Eq. 3-12}$$

$$\pi(x) = \sum_{z_1, \dots, z_p} P(x | Z) \prod_{j=1}^p \pi_{z_j, X}(z_j) \quad \text{Eq. 3-13}$$

As mensagens π transmitidas, adotando X como nó pai e Y como nó filho, são dadas por:

$$\pi_{XY_j}(x) = \alpha \pi_X(x) \prod_{k \neq j} \lambda_{Y_k X}(x) \quad \text{Eq. 3-14}$$

As mensagens λ transmitidas, adotando X como nó pai e Y como nó filho, e levando em consideração outros pais V de Y , são dadas por:

$$\lambda_{Y_j X}(x) = \sum_{y_j} \lambda_{Y_j}(y_j) \sum_{v_1, \dots, v_q} p(y_j | v_1, \dots, v_q) \prod_{k=1}^q \pi_{V_k Y_j}(v_k) \quad \text{Eq. 3-15}$$

Na inicialização, para os nós sem pais $\pi(x_i) = P(x_i)$ (probabilidade *a priori*), e para os nós sem filhos $\lambda(x_i) = I$.

Resumindo, as iterações ocorrem, até que nenhuma alteração aconteça, da seguinte maneira (para cada nó X) (NEAPOLITAN, 2004):

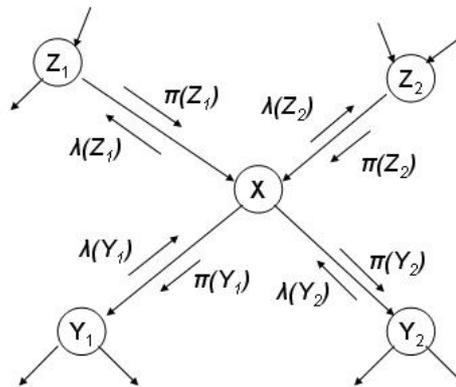


Figura 3-7 – Processo de inferência.

- Se X recebeu todas as mensagens- π de seus pais, calcula $\pi(x)$. O nó Z que enviará uma mensagem- π ao seu filho X precisa das mensagens- π de todos os seus pais e das mensagens- λ de todos os seus filhos exceto X ;
- Se X recebeu todas as mensagens- λ de seus filhos, calcula $\lambda(x)$. O nó Y que enviará uma mensagem- λ ao seu pai X precisa das mensagens- λ de todos os seus filhos e das mensagens- π de todos os seus pais exceto X ;
- Se $\pi(x)$ foi calculado e X recebeu todas as mensagens- λ de todos os seus filhos (exceto Y), calcula $\pi_{XY}(x)$ e o envia a Y ;
- Se $\lambda(x)$ foi calculado e X recebeu todas as mensagens- π de todos os seus pais (exceto Z), calcula $\lambda_{XZ}(x)$ e o envia a Z ;
- Se X recebeu todas as mensagens de seus vizinhos, calcula $P(X | E) = \alpha \lambda(X) \pi(X)$.

3.2.4 Aprendizagem

O treinamento (ou aprendizado) em Redes Bayesianas pode ter como objetivo a obtenção (ou atualização) das probabilidades da rede, da estrutura da rede, ou de ambos, com o intuito de aumentar o potencial da rede. Os algoritmos de aprendizagem permitem projetar um aprendizado de maneira bem completa:

- Na falta total de conhecimentos pode-se utilizar uma estrutura de rede mais adaptada, isto é, uma estrutura que realize a quantificação das relações de dependência e de independência entre as diferentes variáveis e os parâmetros (ou probabilidades).
- Se estiverem disponíveis conhecimentos *a priori* sobre a estrutura, e uma base de exemplos representativos, a determinação das matrizes de probabilidades condicionais (que são os parâmetros da rede) pode ser efetuada por simples cálculo de frequências, por determinação do máximo de verossimilhança, ou por métodos bayesianos.

Esses métodos podem ser estendidos para o contexto de bases de dados incompletos.

Levando-se em consideração a ótica de estabelecer um compromisso entre aprendizagem e generalização, é igualmente possível efetuar a aprendizagem fazendo o controle (ou restringindo) da estrutura da rede.

O princípio geral da aprendizagem em uma Rede Bayesiana é descrito pela fórmula geral:

$$A \text{ Posteriori} = \text{Verossimilhança} \times A \text{ Priori} \quad \text{Eq. 3-16}$$

A princípio, a definição para conhecimento *a priori* é de um tipo de conhecimento adquirido unicamente através do pensamento dedutivo. O conhecimento *a priori* costuma ser contrastado com o conhecimento *a posteriori*, que é aquele tipo de conhecimento que requer certa experiência.

A fórmula mostrada na Equação 3-16 condiciona a modificação do conhecimento contido dentro da rede pela aquisição de novos exemplos. Ela mostra que o conhecimento dito *a priori*, ou em um instante qualquer, dentro da rede, é transformado em conhecimento *a posteriori* em função da verossimilhança de observação de exemplos examinados segundo o conhecimento inicial. Em outras palavras, quanto mais os dados dos exemplos observados se

desviam dos dados do conhecimento da rede, mais modificações eles farão nesses conhecimentos.

Teoricamente, esta fórmula (que é a fórmula de Bayes aplicada ao conhecimento) é válida tanto para a aprendizagem de parâmetros da rede quanto para a aprendizagem da estrutura da rede.

De fato, a função da aprendizagem em uma Rede Bayesiana é examinar, no meio de um conjunto de distribuições, a distribuição mais próxima possível da distribuição representada pelos dados. A estimação das probabilidades é realizada a partir de dados, chamados dados de treinamento (ou bases de treinamento), que podem ser rotulados ou não em função das classes presentes. Se a base de treinamento é rotulada em termos de classes, é dito que há um *treinamento supervisionado*, caso contrário, é dito que há um *treinamento não supervisionado*. O treinamento utilizado foi o supervisionado, que consiste em determinar os parâmetros da rede dado um conjunto de vetores associados a cada uma das classes, o seja, os exemplos já estão classificados.

Em relação à *superaprendizagem*, que vem a ser a cópia exata da distribuição representada pelos dados de exemplo, para evitá-la deve-se limitar o número de parâmetros livres do modelo.

Como regra geral, se a pertinência dos vetores nas diferentes classes não é conhecida, não é possível ou simplesmente não é necessária, os parâmetros da rede podem ser determinados iterativamente pelo *algoritmo EM*, que realiza a maximização da função de verossimilhança $P(y_k | x_i)$.

CAPÍTULO 4 – O PROJETO TELEPAT

4.1 Introdução

O projeto TELEPAT tem como objetivo a realização de um serviço de telemonitoramento confiável de pessoas idosas e/ou com problemas cardíacos em domicílio, baseado na utilização original de duas modalidades complementares de televigilância. Ou seja, o projeto TELEPAT tem por objetivo principal possibilitar uma medicação adequada ao paciente de acordo com as suas necessidades, baseado no telediagnóstico por meio da utilização de um ambiente inteligente e de dados vitais do paciente. Esse sistema processa as informações, obtidas pelas medições, em uma unidade de processamento (servidor local), para fornecer um diagnóstico. Essas informações são capturadas com o uso de variados tipos de sensores para captação de sinais de fontes diversas.

O sistema TELEPAT, conforme mostra a Figura 4-1, é composto por um Sistema Móvel ou terminal móvel portátil, de tamanho reduzido, transportado pelo paciente (BALDINGER, BOUDY, et al., 2004), um Sistema Fixo ou dispositivo fixo de detecção de presença, com sensores piroelétricos (STEENKESTE, BANERJEE, et al., 2005), integrado ao ambiente onde se encontra o paciente, uma Base Doméstica, e um Servidor Remoto (BOUDY, 2006).

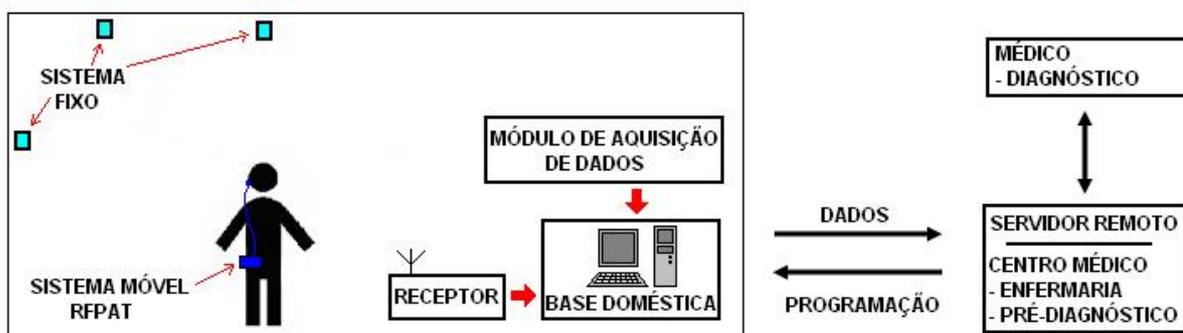


Figura 4-1 – Sistema TELEPAT.

A Base Doméstica (servidor local), também chamada de dispositivo de recepção fixo, é integrado ao ambiente do paciente (em seu domicílio ou quarto do hospital, por exemplo), compondo uma rede de comunicação doméstica. O conjunto composto pelo Sistema Móvel e pelo Sistema Fixo é conectado ao Servidor Local, sendo que o primeiro sistema utiliza ondas de rádio (receptor/transmissor VHF, com o receptor ligado ao Servidor Local por meio de conexão serial RS232 ou USB) e o segundo sistema utiliza comunicação serial RS232 da Plataforma de Pré-processamento à Base Doméstica, mas com a possibilidade de se utilizar dois transceptores (um ligado ao Módulo de Aquisição de Dados e outro ligado ao Servidor Local), possibilitando uma conexão sem-fio. Essa Plataforma de Pré-processamento (ou Módulo de Aquisição de Dados) é responsável por receber os dados dos sensores piroelétricos (sensores infravermelhos otimizados para detectar a radiação humana) e convertê-los a um sinal válido para ser utilizado na Base Doméstica.

O Servidor Remoto (ou Servidor Centralizado) recebe os sinais emitidos pelas modalidades de telemonitoramento desde o Servidor Local por meio de uma rede privada virtual (VPN) ou pela Internet, e permite ao centro de televigilância médica acompanhar continuamente as informações biomédicas recebidas e prover um atendimento de emergência caso haja algum alarme. Pode-se estender a aplicação do telemonitoramento fazendo uso de PDAs ou de *smartphones* para o acesso do médico à base doméstica ou ao servidor remoto.

4.2 O Módulo Portátil RFpat

Esse sistema de telemonitoramento comporta:

- Um terminal portátil, concebido e desenvolvido por BALDINGER, et al., 2004, transportado pelo paciente (fixado em sua cintura) e equipado por três sensores de movimento e um de medição de frequência cardíaca. Ele é dotado de um emissor RF (sinais de radiofrequência), com capacidade de transmissão em torno de dez metros de distância.
- Uma base doméstica, do tipo PC, para a recepção dos dados dos sensores enviados pelo terminal portátil, e para pós-tratamento desses dados. Ela é conectada à Internet para a conexão com o Servidor Remoto de telemonitoramento.

O circuito responsável pelo envio dos dados à base doméstica, por meio do emissor RF, realiza a transmissão com uma taxa de envio de 30 em 30 segundos.

O dispositivo móvel é composto por sensores, os quais proporcionam as seguintes informações:

- O nível de atividade ou agitação do paciente, medido com a ajuda de um sensor de movimento colocado dentro do próprio terminal;
- A situação do paciente (postura: em pé/sentado ou deitado), medidos com a ajuda de sensores de movimento também instalados no interior do terminal;
- A frequência cardíaca, medida durante 30 segundos por um sensor de pulsos (infravermelho) do tipo pinça, aplicada na orelha ou no dedo;
- O dispositivo móvel possui também um botão de chamada para emitir os dados mais recentes a um servidor remoto, no caso, por exemplo, de o paciente sentir a necessidade de um atendimento médico urgente.

O condicionamento dos sinais captados pelo terminal envolve um tratamento de eliminação de artefatos nesses sinais, além de eliminar interferências do sinal da rede elétrica. As saídas disponibilizadas pelo terminal, sobre a forma de estados ou medidas contínuas, estão representadas na Tabela 4-1.

Tabela 4-1 – As diferentes medidas fornecidas pelo RFpat.

Sensor	Representação
Inclinação	binária (0 ou 1)
Movimentação (ou agitação)	4 bits (valores de 0 a 15)
Frequência cardíaca (ou pulsação)	8 bits (valores de 0 a 255)

Para a saída Inclinação, o bit '0' indica que o usuário está na posição 'de pé', e o bit '1' indica que o usuário está na posição 'deitado'. Para a saída Movimentação, o valor '0' é indicativo de ausência de movimento, e o valor '15' é indicativo de um estado de agitação intensa, sendo a variação entre os dois extremos realizada de forma linear. Para a saída Frequência Cardíaca, os valores representam os batimentos por minuto do usuário, cuja faixa de variação (limite dos sensores) vai de 0 bpm a 200 bpm (que é um limite mais que suficiente, tendo em vista os níveis fisiológicos relacionados à frequência cardíaca do ser humano).

O terminal móvel é fixado na cintura por meio de uma presilha, que é colocada sobre a calça. No caso de sistemas de uso contínuo e prolongado, é importante levar em consideração a facilidade de uso, já que o paciente deverá utilizar esse aparelho durante todo o dia. Tamanho e peso foram itens de importância fundamental no desenvolvimento do protótipo RFpat, além do uso de componentes eletrônicos de baixo consumo de energia, para aumentar a autonomia das baterias.

4.3 O Sistema GARDIEN

O Sistema Fixo GARDIEN é um dispositivo de monitoramento do ambiente (STEENKESTE, 2005). Ele é instalado no ambiente de vida da pessoa monitorada, e permite controlar a pessoa à distância para prevenir acidentes e agravações de doenças. Esse sistema é constituído de uma rede de nove sensores piroelétricos de detecção de presença (como mostra a Figura 4-2, na implantação dentro de um quarto de hospital). Essa rede de sensores é conectada a um Módulo de Aquisição de Dados (Plataforma de Pré-processamento) que se comunica com um computador distante (Servidor Remoto), seja por ligação a cabo, seja por intermédio de ligação sem-fio (Figura 4-3). A taxa de envio de dados do sistema GARDIEN ao Servidor Remoto é de 0,5 segundos.

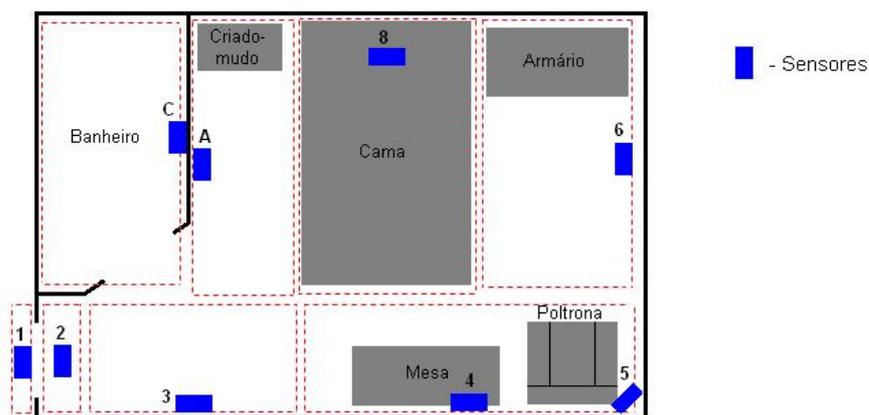


Figura 4-2 – Implantação do sistema GARDIEN em um quarto de hospital.

Em sua primeira versão (a versão original), instalado em um quarto de hospital, o sistema GARDIEN comporta uma rede de nove sensores: oito deles são instalados ao nível do teto

de um quarto de um paciente e orientados verticalmente, de cima para baixo, para detectar todos os movimentos realizados dentro de uma zona de dois metros quadrados aproximadamente (para cada sensor); o nono sensor é fixado a 1 metro do solo, aproximadamente, e orientado horizontalmente (sensor 5, da Figura 4-2). Dentro do grupo dos oito sensores, dois são designados ao monitoramento da porta do quarto do paciente (sensores 1 e 2), um outro monitora a utilização do banheiro (sensor C), outro detecta a agitação do paciente em sua cama (sensor 8) e quatro são determinados a monitorar a movimentação do paciente em seu quarto (sensores 3, 4, 6 e A). Enfim, o sensor horizontal 5 tem como função confirmar ou não uma situação de queda eventual.



Figura 4-3 – Conexão dos sensores por ligação a cabo.

Para permitir a realização do objetivo de telemonitoramento do sistema TELEPAT, a saber, a fusão dos dados de saída dos sensores infravermelhos do GARDIEN de uma parte, e dos sensores embarcados do RFpat de outra parte, um sistema GARDIEN experimental foi desenvolvido dentro das instalações do INT, mais precisamente em uma sala de informática do *Département Electronique et Physique*. Este sistema, instalado no INT, apresenta algumas diferenças quando comparado com o sistema original (tendo em conta o espaço disponível), a começar pela disposição dos sensores. A Figura 4-4 representa a planta baixa e o corte AB da sala de testes do INT onde está instalado o sistema GARDIEN. Essa sala foi dividida em duas partes com o uso de um cortinado feito de material resistente e opaco (simulando uma parede), e cada área foi nomeada Área 1 e Área 2 (BELFEKI, 2006).

Os sensores piroelétricos de cor azul (3, 4, 6, 8, A e C) estão instalados no teto, a uma distância de aproximadamente 2,30 metros do solo (entre 2,10 e 2,50 metros, com possibilidade de regulagem de altura). Cada um desses sensores tem a finalidade de detectar a movimentação de uma pessoa no ambiente (detecção vertical), cobrindo, cada um, uma área retangular de aproximadamente dois metros quadrados ao nível do solo. Os sensores piroelétricos

de cor vermelha (5 e B) estão instalados a aproximadamente 1,0 metro do solo, sobre bancadas (um em cada Área – Área 1 para sensor B e Área 2 para sensor 5), de maneira que a zona de detecção seja horizontal e cubra toda a área de atuação na qual ele esteja posicionado. Diferentemente dos outros sensores de detecção vertical, os sensores 5 e B possuem uma detecção orientada (e não formando um cone), no caso horizontalmente. Foram utilizados dois sensores horizontais (e não um, como na versão original) por uma questão de distância de detecção, pois a sala de experimentação do INT é menor em superfície, mas mais larga que um quarto de hospital. Cada um desses dois sensores tem a finalidade de detectar se houve queda ou não do paciente.

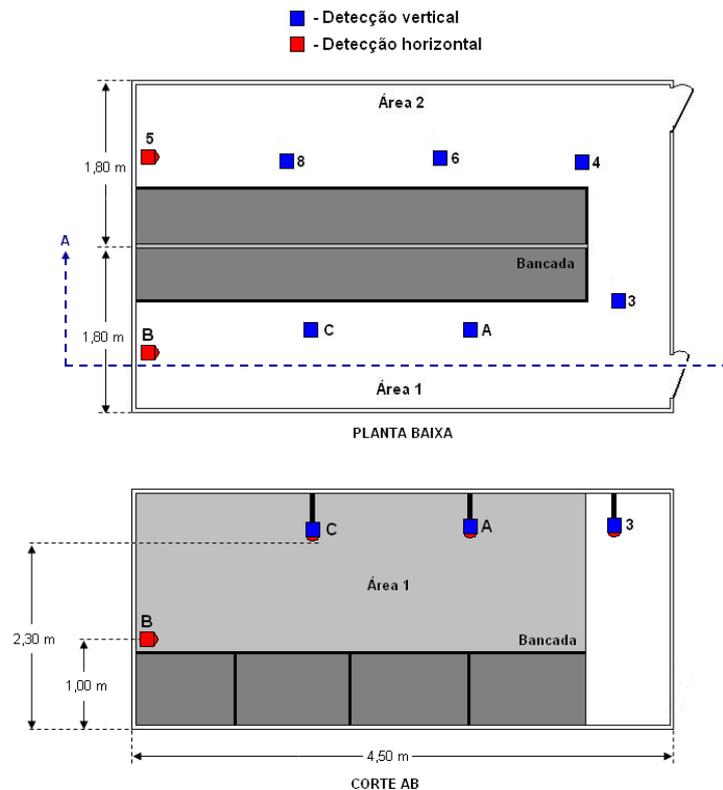


Figura 4-4 – Plano esquemático da sala de testes do INT – sistema GARDIEN.

Os sensores infravermelhos utilizados possuem uma regulagem do campo de detecção, permitindo o aumento ou diminuição do campo de atuação dos mesmos. Essa regulagem se torna necessária, para que o campo de atuação de um sensor não ultrapasse o campo de atuação de outro sensor (independência condicional dos sensores), o que interferiria no correto funcionamento do algoritmo do programa de detecção da posição do paciente no ambiente. De acordo com a Figura 4-5, a Situação 1 não é desejada, já que o campo de atuação dos dois

sensores estão se sobrepondo. Nesse caso, os sensores são corrigidos até se conseguir uma condição como mostrada na Situação 2.

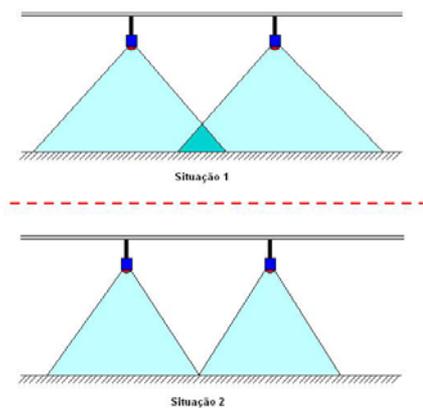


Figura 4-5 – Área de atuação dos sensores infravermelhos instalados no teto.

CAPÍTULO 5 – O PROGRAMA DE FUSÃO PROPOSTO

Como foi mencionado no Capítulo 1 sobre a definição do problema, um dos pontos principais deste trabalho é o desenvolvimento de um programa que realize a fusão dos dados de saída dos sistemas RFpat e GARDIEN.

O objetivo principal do programa de fusão é oferecer uma medicação apropriada a pessoas idosas e/ou que sofram de patologias cardíacas, sendo que os benefícios atendem também ao interesse de avaliação da solução tecnológica e das suas evoluções potenciais, como por exemplo a evolução orientada ao telediagnóstico (ou seja, o diagnóstico realizado à distância) ou a aplicação baseada no princípio do habitat inteligente (como a comunicação com dispositivos de uma casa, realizada de forma automática).

O telemonitoramento utilizando esse programa de fusão se baseia na idéia de fornecer uma saída que indique uma condição alarmante, empregando três situações relacionadas com o paciente: queda, bradicardia e taquicardia (ou a combinação entre eles). As condições alarmantes de saída são especificadas em probabilidades, já que para se realizar o trabalho de monitoramento foi utilizada Rede Bayesiana. O esquema da fusão de dados está exemplificado na Figura 5-1. Nele estão identificados alguns exemplos de saída da rede, como, por exemplo, a visualização dos resultados em um monitor, a geração de um relatório para armazenamento em um banco de dados, e/ou um sinal de alerta. Nesta mesma Figura, ‘I’, ‘M’ e ‘P’ significam Inclinação, Movimentação e Pulsação, respectivamente. C_1 , C_2 , etc., simbolizam os sensores piroelétricos do ambiente, que podem ser utilizados pelo sistema GARDIEN.

Para a realização da Rede Bayesiana aplicada ao sistema, foi utilizada a plataforma Matlab 6.5, com o uso de várias funções do *toolbox* BNT – *Bayes Net Toolbox* (MURPHY, 1998), dedicado à avaliação de Redes Bayesianas. O BNT é uma biblioteca de código aberto com funções Matlab para auxiliar na criação, inferência e aprendizagem de modelos gráficos. Esse projeto conta atualmente com a participação de inúmeros pesquisadores no desenvolvimento e atualização das ferramentas disponíveis por essa biblioteca.

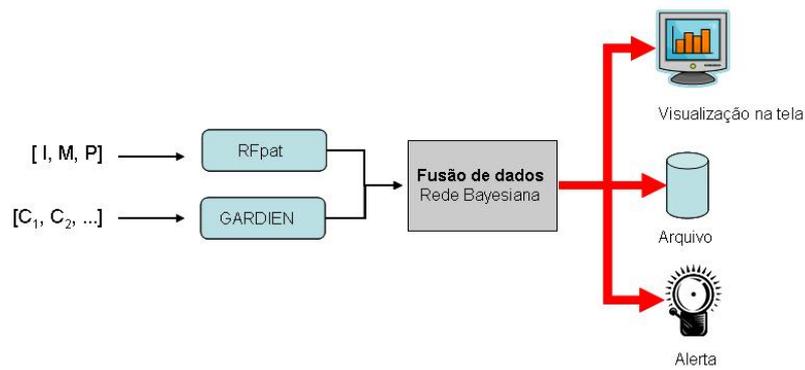


Figura 5-1 – Esquema do processo de fusão de dados.

A escolha desse toolbox se deve ao fato de possuir vantagens apreciáveis em comparação com outros programas similares, tais como:

- Código fonte aberto;
- Possibilidade de uso de API (Matlab);
- Aprendizagem tanto de parâmetros quanto da estrutura da rede;
- Suporte a diagramas de influência;
- Uso tanto de modelos diretos quanto de modelos indiretos;
- Grande quantidade de métodos de inferência.

Com o propósito de testar a rede, além do uso do *toolbox* BNT, foi utilizado também o programa Netica (NORSYS, 2007), específico para aplicações com Redes Bayesianas e diagramas de influência.

5.1 Método de Diagnóstico Baseado em Redes Bayesianas

Neste trabalho foram analisados três sinais gerados pelo terminal portátil RFpat (‘Inclinação’, ‘Movimentação’ e ‘Pulsção’ do paciente) e dois sinais gerados pelo sistema GARDIEN (sensores piroelétricos B e 5 – ver Figura 4-4). Esses sinais se apresentam na forma de eventos e são representados por nós da Rede Bayesiana (variáveis aleatórias discretas). À medida que os eventos vão ocorrendo, a probabilidade condicional associada aos mesmos se propaga ao longo da rede, tendo em vista, a princípio, a identificação das seguintes situações alarmantes:

- Queda do paciente;
- Freqüência cardíaca baixa (bradicardia);
- Freqüência cardíaca alta (taquicardia).

Em conseqüência, um sinal de alarme é gerado como resultado da análise da rede. Num primeiro instante, após o sistema acusar o recebimento dos eventos gerados pelos sensores que indiquem alguma anormalidade com o paciente (que é verificada de acordo com a probabilidade de ocorrência de uma das três situações alarmantes descritas acima), é gerado um aviso de ‘Atenção’, indicando a data, hora e a probabilidade do evento. Caso no próximo conjunto de dados de entrada recebido persista a situação de ‘Atenção’, um aviso de ‘Alerta’ é exibido.

Além dos nós que representam os eventos gerados pelos sensores, têm-se ainda os nós intermediários, que representam os estados do paciente. Para estabelecer os parâmetros das tabelas de probabilidades que relacionam os nós da rede (as CPDs) foram adotados alguns critérios para definir os estados do paciente, ou seja, foram estabelecidos os valores dos sinais inclinação, movimentação, pulsação e sensor 5 ou B que, combinados, dão os estados normal, taquicardia, bradicardia e queda do paciente (SOBIECKI, BOURKEB, 2006). Como exemplo de possíveis cenários, pulsação maior que 120 batimentos por segundos juntamente com nível de agitação do paciente maior que 70% podem representar um estado de taquicardia; e pulsação menor que 40 batimentos por segundo juntamente com nível de agitação menor que 70% podem representar um cenário para um estado de bradicardia. O valor de porcentagem tem como referência a faixa de valores para a movimentação, que vai de 0 a 15.

Foram desenvolvidas duas interfaces Matlab responsáveis por mostrar os resultados do telemonitoramento de acordo com a necessidade de uso: uma para o modo *off-line* e outra para o modo *on-line*, que serão explicados nas seções 6.2 e 6.3, respectivamente.

O desenvolvimento dessas topologias de rede foi possível após a aquisição de dados para testes. Esses dados foram adquiridos por meio de diferentes registros de dados normais (de uma pessoa saudável) adquiridos com o uso dos sistemas RFpat e GARDIEN, e de dados simulados para bradicardia e taquicardia, por meio do programa de simulação.

5.2 Fusão de Dados – Modo ‘Off-line’

A modalidade de fusão em modo *off-line*, como o próprio nome diz, leva em consideração a determinação dos sinais de alerta quando os sistemas RFpat e GARDIEN não estão enviando dados para o computador, ou seja, os dados são obtidos dessas unidades de monitoramento e gravados no computador para, posteriormente, serem disponibilizados para o programa de fusão. Essa modalidade serve para analisar conjuntos de dados registrados pelos sensores.

Foram testadas várias topologias de rede para a fusão de dados no modo *off-line*. Nas próximas seções será abordada a evolução dos modelos desenvolvidos, apresentando as topologias que foram consideradas como teste e versão final.

5.2.1 Primeiras Versões

A princípio, foram testadas várias topologias de rede, mas serão colocadas aqui apenas duas versões consideradas mais importantes como comparação. Nessas duas versões serão mostrados os nós que foram substituídos ou mantidos devido à observação de seu comportamento na rede.

A primeira topologia está representada na Figura 5-2, onde estão mostradas as relações de independência condicional entre os nós. Essa rede possui apenas variáveis discretas, ou seja, os nós admitem apenas um número finito de estados. Para essa estrutura estão sendo analisados apenas os sinais provenientes do módulo RFpat. Foi utilizado esse método de estudo como forma de testar, por parte, o funcionamento do sistema, sendo a princípio apenas os dados do terminal portátil RFpat incluídos na rede e, posteriormente, realizada a inclusão dos dados dos sensores B e 5 do sistema GARDIEN.

Na estrutura mostrada na Figura 5-2 estão representados oito nós discretos, onde três deles, os nós Inclinação, Movimento e Pulso representam a entrada da rede (que são os dados de saída do terminal RFpat). Os estados que cada nó pode receber são:

- Inclinação:
 - vertical: indica que o usuário está na posição vertical;
 - horizontal: indica que o usuário está na posição horizontal.
- Movimento:
 - menos 70: indica que o usuário está com nível de movimentação abaixo de 70 %;

- mais 70: indica que o usuário está com nível de movimentação acima de 70 %;
- Pulso:
 - normal: indica que o usuário possui a sua frequência cardíaca em níveis normais;
 - alto baixo: indica que o usuário possui a sua frequência cardíaca em níveis anormais (ou mais alto ou mais baixo que o normal);

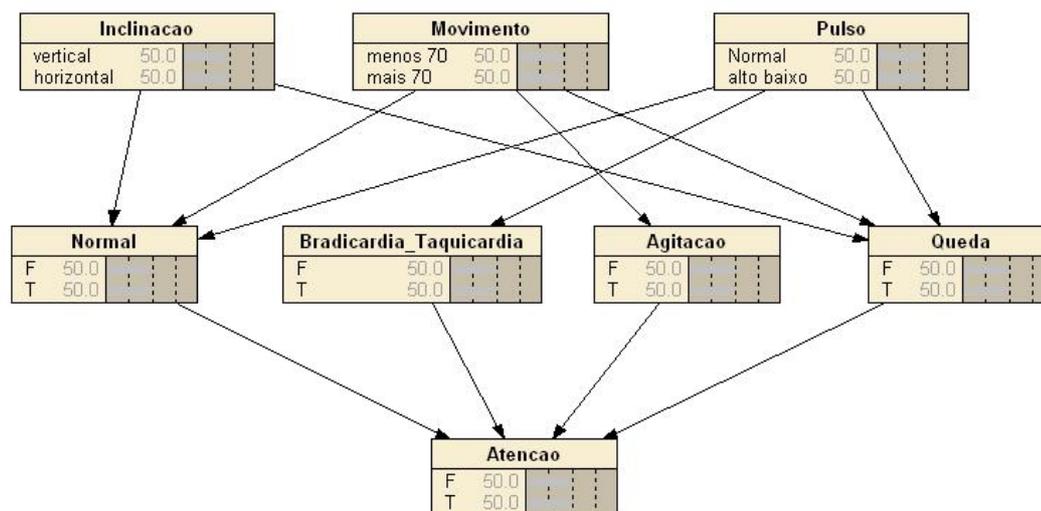


Figura 5-2 – Estrutura da rede - Primeira versão.

Os sinais de entrada relativos à movimentação do usuário (nó Movimento) são gerados dentro de uma faixa de valores que vai de 0 a 15 (4 bits). Desta forma, um nível de movimentação acima de 70% (condição ‘mais 70’) indica que o sinal de entrada possui valor maior que 10. Esse valor específico foi escolhido com base em dados médicos referente à condição de um idoso que sofre uma queda, e também tendo como referência a sensibilidade do sensor de movimento do terminal portátil RFpat. Para a condição ‘menos 70’ o sinal de entrada deve possuir valor igual ou menor que 10, o que representa movimentação abaixo dos 70%. Deve-se deixar claro que a tabela de probabilidades para o nó ‘Queda’ leva em consideração não apenas movimento acima de 70%, como indicativo de uma situação de queda, mas também situações em que pode-se ter pouco movimento e queda, ou também muito movimento e não ter queda. Esses cenários dependem da situação analisada.

Para os sinais de frequência cardíaca (nó ‘Pulso’), a condição ‘normal’ corresponde a um sinal de entrada entre 40 e 120 batimentos por minuto. A condição ‘alto baixo’ indica um sinal de entrada tanto maior que 120 bpm quanto menor que 40 bpm, denotando uma taquicardia ou bradicardia, respectivamente.

Nota-se a presença de um nó intermediário ‘Normal’, que tem a função de identificar, a partir da análise dos dados de entrada, uma condição de normalidade com o usuário. Esse nó se tornou uma forma redundante de verificar essa condição, cuja retirada não acarretou em mudanças significativas no funcionamento da rede, sendo assim descartado nas versões seguintes.

O nó intermediário ‘Bradycardia_Taquicardia’ modela o sinal de frequência cardíaca permitindo verificar se há uma anormalidade com o batimento cardíaco do usuário, e o nó intermediário ‘Agitação’ analisa os dados do nó ‘Movimento’ para verificar o nível de agitação do usuário. Da mesma forma que o nó ‘Normal’, o nó ‘Agitação’ foi descartado por não causar mudanças expressivas no desempenho da rede, utilizando-se do empirismo para testar e observar o funcionamento desses nós.

O nó mediador ‘Queda’, como o próprio nome diz, tem a função de receber os dados de entrada da rede e analisar se há uma possível situação de queda do paciente.

Um sinal de alerta é gerado a partir da resposta do nó ‘Atenção’. Esse sinal tem dois níveis, sendo especificado a partir da resposta fornecida por esse nó, que será uma probabilidade indicativa de uma situação alarmante com o paciente. De acordo com essa probabilidade, pode-se ter uma situação de alerta normal (que é determinada em função do paciente analisado, sendo, por exemplo, uma probabilidade de 50%) e de alerta máximo (também determinada em função do paciente, podendo ser, por exemplo, uma probabilidade de 70%).

A segunda versão da topologia da rede de teste está representada na Figura 5-3. Para essa estrutura estão sendo analisados, além dos sinais provenientes do módulo RFpat, os sinais advindos dos sensores 5 e B do sistema GARDIEN. Dessa forma, uma situação de queda será melhor determinada. O nó de entrada B5 representa os sinais provenientes dos sensores 5 ou B, e possui os estados F (do inglês *false*) e T (do inglês *true*), onde ‘T’ indica que o sensor está detectando a presença de uma pessoa no ambiente (paciente teoricamente de pé), e ‘F’ indica que o paciente está, teoricamente, deitado, pois o sensor não está detectando a presença de movimento no ambiente. Vale ressaltar que os estados dos nós de entrada ‘Inclinação’, ‘Movimento’ e ‘Pulso’, para essa estrutura, são os mesmos da estrutura anterior. Além disso, os nove nós dessa estrutura são do tipo discreto.

Os nós intermediários ‘Bradycardia_Taquicardia’ e ‘Queda’ possuem a mesma função que a mostrada para a estrutura da rede da Figura 5-2.

O sinal de saída da rede, ainda em dois níveis, conta agora com dois nós. Em primeira instância existe o nó ‘Atenção’, que fornece um sinal de atenção, e em segunda instância existe o nó ‘ALERTA’ que exhibe um sinal de alerta (de maior relevância, em comparação ao sinal

de atenção). Este último leva em consideração os estados dos nós ‘Atenção’ e do nó ‘Tempo’. O nó ‘Tempo’ indica apenas se houve um sinal de atenção no ciclo anterior, ou seja, seu estado é verdadeiro (*true*) quando contabiliza um sinal de atenção por mais de um ciclo (por mais de 30 segundos).

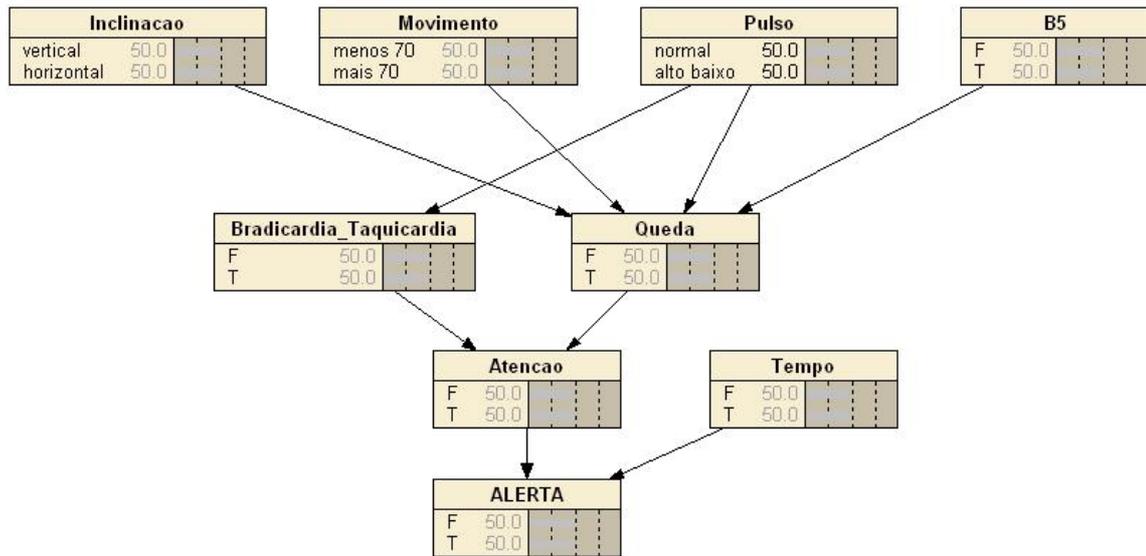


Figura 5-3 – Estrutura da rede - Segunda versão.

5.2.2 Versão Final

A estrutura final da Rede Bayesiana para a fusão dos dados está mostrada na Figura 5-4. Os seus oito nós são discretos, sendo que o nó ‘Pulso’ agora tem três estados: baixo, normal e alto. Em relação às versões anteriores, a condição ‘alto baixo’ foi separada em ‘alto’ (para sinais referentes à taquicardia – maiores que 120 bpm) e ‘baixo’ (para sinais referentes à bradicardia – menores que 40 bpm).

Para os nós de entrada da rede foram estabelecidos os estados dos sinais de acordo com o tipo de sensor, como mostra a Tabela 5-1.

Esta topologia apresenta, separadamente, um nó para o estado de bradicardia e um nó para o estado de taquicardia. O resultado da fusão é fornecido pelo nó ‘CONTROLE’.

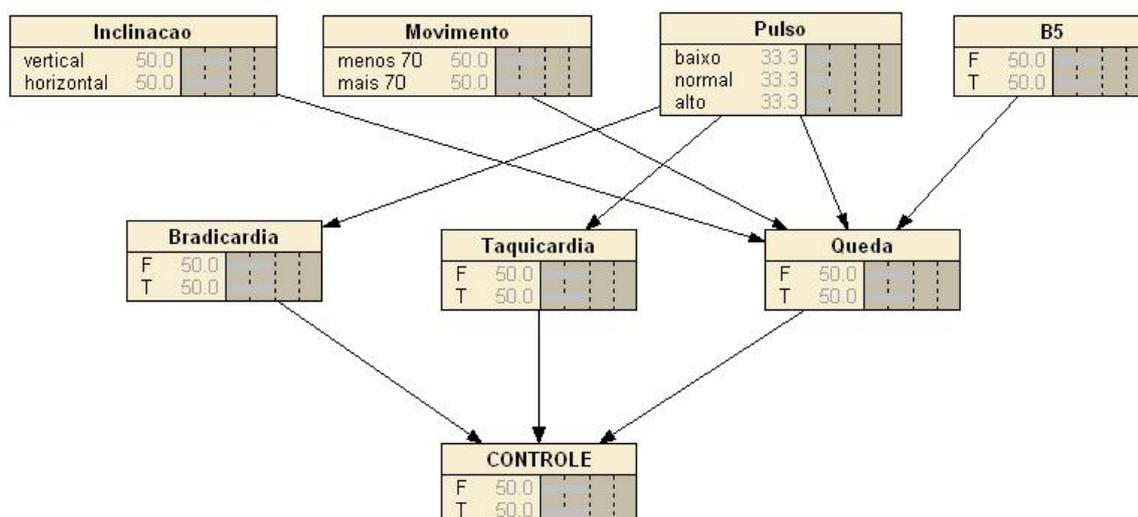


Figura 5-4 – Estrutura da rede - versão final.

Tabela 5-1 – Estados dos sinais dos sensores.

Sensor	Estado
Inclinação	Horizontal ou vertical (deitado ou em pé)
Movimentação	Porcentagem de atividade
Frequência cardíaca (FC)	FC < 40 bat/min (Bradicardia) 40 < FC < 120 bat/min (Normal) FC > 120 bat/min (Taquicardia)
B5	Horizontal ou vertical (deitado ou em pé)

Após a criação da estrutura da Rede Bayesiana, deve-se especificar as probabilidades *a priori*, ou seja, a quantificação inicial da Rede Bayesiana. A aproximação indireta para calcular as probabilidades condicionais, dadas pela Lei de Bayes da equação 3-16, possui esse inconveniente, pois as probabilidades *a priori* são geralmente difíceis de estimar de maneira confiável. Para a determinação das probabilidades condicionais foi realizada uma etapa de aprendizado, utilizando para isso dados obtidos em simulações de cenários pré-determinados, com o uso do equipamento de telemonitoramento em laboratório. Essa etapa será melhor explicada na seção 6.4.

Para preencher as tabelas de probabilidades *a priori* foi empregado um modo empírico, a partir de dados bibliográficos e informações de profissionais da área médica. Tal método permitiu testar a robustez da rede, já que as probabilidades iniciais utilizadas não condiziam, tão precisamente (teoricamente), com uma situação real. Além disso, tal método possibilitou a

obtenção das probabilidades de forma rápida. Durante as avaliações realizadas com a rede, puderam-se efetuar correções nos valores de algumas probabilidades que se mostraram incoerentes. As Tabelas de probabilidades *a priori* estão mostradas nas Figuras 5-5, 5-6 e 5-7. Os valores ‘F’ e ‘T’ representam ‘falso’ e ‘verdadeiro’, respectivamente.

<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px;">P(I = vertical)</td> <td style="padding: 2px;">P(I = horizontal)</td> </tr> <tr> <td style="padding: 2px;">0,5</td> <td style="padding: 2px;">0,5</td> </tr> </table>		P(I = vertical)	P(I = horizontal)	0,5	0,5	<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px;">P(M = menos 70)</td> <td style="padding: 2px;">P(M = mais 70)</td> </tr> <tr> <td style="padding: 2px;">0,5</td> <td style="padding: 2px;">0,5</td> </tr> </table>		P(M = menos 70)	P(M = mais 70)	0,5	0,5		
P(I = vertical)	P(I = horizontal)												
0,5	0,5												
P(M = menos 70)	P(M = mais 70)												
0,5	0,5												
(a)		(b)											
<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px;">P(P = baixo)</td> <td style="padding: 2px;">P(P = normal)</td> <td style="padding: 2px;">P(P = alto)</td> </tr> <tr> <td style="padding: 2px;">0,33</td> <td style="padding: 2px;">0,34</td> <td style="padding: 2px;">0,33</td> </tr> </table>			P(P = baixo)	P(P = normal)	P(P = alto)	0,33	0,34	0,33	<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px;">P(B5 = F)</td> <td style="padding: 2px;">P(B5 = T)</td> </tr> <tr> <td style="padding: 2px;">0,5</td> <td style="padding: 2px;">0,5</td> </tr> </table>	P(B5 = F)	P(B5 = T)	0,5	0,5
P(P = baixo)	P(P = normal)	P(P = alto)											
0,33	0,34	0,33											
P(B5 = F)	P(B5 = T)												
0,5	0,5												
(c)			(d)										

Figura 5-5 – Tabelas de probabilidades para os nós de entrada da rede: (a) Inclinação (‘I’), (b) Movimento (‘M’), (c) Pulso (‘P’) e (d) B5.

<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="padding: 2px;">Pulso</th> <th style="padding: 2px;">P(Bradicardia = F)</th> <th style="padding: 2px;">P(Bradicardia = T)</th> </tr> <tr> <td style="padding: 2px;">baixo</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> </tr> <tr> <td style="padding: 2px;">normal</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td style="padding: 2px;">alto</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> </table> <p style="text-align: center;">(a)</p>	Pulso	P(Bradicardia = F)	P(Bradicardia = T)	baixo	0	1	normal	1	0	alto	1	0	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="padding: 2px;">Inclinação</th> <th style="padding: 2px;">Movimento</th> <th style="padding: 2px;">Pulso</th> <th style="padding: 2px;">B5</th> <th style="padding: 2px;">P(Queda = F)</th> <th style="padding: 2px;">P(Queda = T)</th> </tr> </thead> <tbody> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0</td><td style="padding: 2px;">1</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0</td><td style="padding: 2px;">1</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0,5</td><td style="padding: 2px;">0,5</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">baixo</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0,3</td><td style="padding: 2px;">0,7</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0</td><td style="padding: 2px;">1</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">F</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">F</td><td style="padding: 2px;">0</td><td style="padding: 2px;">1</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">normal</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">menos 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">vertical</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> <tr><td style="padding: 2px;">horizontal</td><td style="padding: 2px;">mais 70</td><td style="padding: 2px;">alto</td><td style="padding: 2px;">T</td><td style="padding: 2px;">1</td><td style="padding: 2px;">0</td></tr> </tbody> </table> <p style="text-align: center;">(c)</p>	Inclinação	Movimento	Pulso	B5	P(Queda = F)	P(Queda = T)	vertical	menos 70	baixo	F	1	0	horizontal	menos 70	baixo	F	0	1	vertical	mais 70	baixo	F	1	0	horizontal	mais 70	baixo	F	0	1	vertical	menos 70	baixo	F	1	0	horizontal	menos 70	baixo	F	0,5	0,5	vertical	mais 70	baixo	F	1	0	horizontal	mais 70	baixo	F	0,3	0,7	vertical	menos 70	normal	F	1	0	horizontal	menos 70	normal	F	0	1	vertical	mais 70	normal	F	1	0	horizontal	mais 70	normal	F	0	1	vertical	menos 70	normal	T	1	0	horizontal	menos 70	normal	T	1	0	vertical	mais 70	normal	T	1	0	horizontal	mais 70	normal	T	1	0	vertical	menos 70	alto	T	1	0	horizontal	menos 70	alto	T	1	0	vertical	mais 70	alto	T	1	0	horizontal	mais 70	alto	T	1	0	vertical	menos 70	alto	T	1	0	horizontal	menos 70	alto	T	1	0	vertical	mais 70	alto	T	1	0	horizontal	mais 70	alto	T	1	0
Pulso	P(Bradicardia = F)	P(Bradicardia = T)																																																																																																																																																																	
baixo	0	1																																																																																																																																																																	
normal	1	0																																																																																																																																																																	
alto	1	0																																																																																																																																																																	
Inclinação	Movimento	Pulso	B5	P(Queda = F)	P(Queda = T)																																																																																																																																																														
vertical	menos 70	baixo	F	1	0																																																																																																																																																														
horizontal	menos 70	baixo	F	0	1																																																																																																																																																														
vertical	mais 70	baixo	F	1	0																																																																																																																																																														
horizontal	mais 70	baixo	F	0	1																																																																																																																																																														
vertical	menos 70	baixo	F	1	0																																																																																																																																																														
horizontal	menos 70	baixo	F	0,5	0,5																																																																																																																																																														
vertical	mais 70	baixo	F	1	0																																																																																																																																																														
horizontal	mais 70	baixo	F	0,3	0,7																																																																																																																																																														
vertical	menos 70	normal	F	1	0																																																																																																																																																														
horizontal	menos 70	normal	F	0	1																																																																																																																																																														
vertical	mais 70	normal	F	1	0																																																																																																																																																														
horizontal	mais 70	normal	F	0	1																																																																																																																																																														
vertical	menos 70	normal	T	1	0																																																																																																																																																														
horizontal	menos 70	normal	T	1	0																																																																																																																																																														
vertical	mais 70	normal	T	1	0																																																																																																																																																														
horizontal	mais 70	normal	T	1	0																																																																																																																																																														
vertical	menos 70	alto	T	1	0																																																																																																																																																														
horizontal	menos 70	alto	T	1	0																																																																																																																																																														
vertical	mais 70	alto	T	1	0																																																																																																																																																														
horizontal	mais 70	alto	T	1	0																																																																																																																																																														
vertical	menos 70	alto	T	1	0																																																																																																																																																														
horizontal	menos 70	alto	T	1	0																																																																																																																																																														
vertical	mais 70	alto	T	1	0																																																																																																																																																														
horizontal	mais 70	alto	T	1	0																																																																																																																																																														
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="padding: 2px;">Pulso</th> <th style="padding: 2px;">P(Taquicardia = F)</th> <th style="padding: 2px;">P(Taquicardia = T)</th> </tr> <tr> <td style="padding: 2px;">baixo</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td style="padding: 2px;">normal</td> <td style="padding: 2px;">1</td> <td style="padding: 2px;">0</td> </tr> <tr> <td style="padding: 2px;">alto</td> <td style="padding: 2px;">0</td> <td style="padding: 2px;">1</td> </tr> </table> <p style="text-align: center;">(b)</p>	Pulso	P(Taquicardia = F)	P(Taquicardia = T)	baixo	1	0	normal	1	0	alto	0	1																																																																																																																																																							
Pulso	P(Taquicardia = F)	P(Taquicardia = T)																																																																																																																																																																	
baixo	1	0																																																																																																																																																																	
normal	1	0																																																																																																																																																																	
alto	0	1																																																																																																																																																																	

Figura 5-6 – Tabelas de probabilidade condicional para os nós intermediários da rede: (a) Bradicardia, (b) Taquicardia e (c) Queda.

Bradicardia	Taquicardia	Queda	P(CONTROLE = F)	P(CONTROLE = T)
F	F	F	1	0
T	F	F	0	1
F	T	F	0	1
T	T	F	0,2	0,8
F	F	T	0	1
T	F	T	0	1
F	T	T	0	1
T	T	T	0,2	0,8

Figura 5-7 – Tabela de probabilidade condicional para o nó Controle.

Dada as evidências, ou seja, dados os valores das variáveis dos nós observáveis (que são aqueles cujos valores são obtidos diretamente do experimento), pode-se realizar a inferência probabilística da rede, que é o processo de estimação da nova distribuição de probabilidade condicional, fornecendo a saída da rede. A estratégia de inferência adotada foi o uso do algoritmo *Junction Tree* (ver seção 3.2.3). Os nós observáveis são os nós representantes da entrada da rede: Inclinação, Movimento, Pulso e B5.

A saída da rede, com o resultado da fusão, é apresentada na forma de um arquivo texto com as indicações de data, horário, sinais de entrada (inclinação, movimentação, pulsação e B5) e as probabilidades dos nós intermediários e do nó CONTROLE, além de também ser disponibilizado um gráfico contendo a probabilidade do nó CONTROLE. As Figuras 5-8 e 5-9 mostram um exemplo de um arquivo texto de saída e de um gráfico de probabilidades, respectivamente. Para a Figura 5-8: *Jour, Heure, Inclin, Mouv, Pouls, B5, Brady, Tachy, Chute e Control* representam respectivamente Data, Hora, Inclinação, Movimentação, Freq. Cardíaca, Sensor B5, Bradicardia, Taquicardia, Queda e Controle.

```

Resultats du Reseau Bayesian

  Jour      Heure      Inclin  Mouv  Pouls  B5 | Brady  Tachy  Chute | Control
-----
15-01-2008  15:31:40    0    08   072    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:32:10    0    06   056    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:32:40    0    11   062    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:33:10    0    12   056    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:33:40    0    13   056    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:34:10    0    12   054    0 | 0 %   0 %   0 % | 0 %
15-01-2008  15:34:40    0    04   032    0 | 100 % 0 %   0 % | 100 %
15-01-2008  15:35:10    0    03   036    0 | 100 % 0 %   0 % | 100 %

--> CONTROL : Bradycardie
--> **ALARME** : Bradycardie

```

Figura 5-8 – Arquivo texto de saída da rede.

Nota-se na Figura 5-8 que na interface *off-line* os dados do sistema TELEPAT (RFpat e GARDIEN) são inseridos no arquivo texto de saída da rede com uma defasagem de 30 segundos na indicação do horário de medição, o que obedece a taxa de transferência dos dados que

ocorre nas transmissões realizadas pelo terminal RFpat, que é maior que o período de envio do sistema GARDIEN (0,5 segundos).

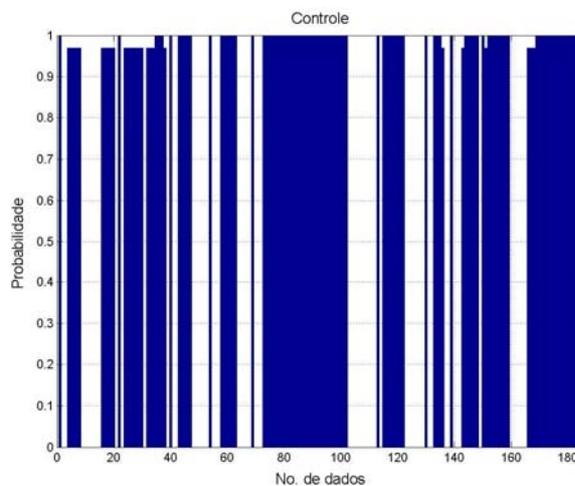


Figura 5-9 – Gráfico de probabilidades do nó Controle.

5.3 Fusão de Dados – Modo ‘On-line’

A modalidade de fusão em modo *on-line* leva em consideração a determinação dos sinais de alerta de forma *on-line*, ou seja, no mesmo instante que os dados são obtidos pelas unidades de monitoramento RFpat e GARDIEN (obedecendo o período de amostragem dos mesmos), os dados são simultaneamente enviados ao programa de fusão de dados. O programa, por sua vez, irá tratar os dados recebidos por meio da Rede Bayesiana integrada, e fornecerá a situação atual do paciente na tela, podendo ou não apresentar um sinal de alerta.

Os parâmetros da Rede Bayesiana para essa modalidade são os mesmos da versão *off-line* final. Depois de realizado o aprendizado da rede, esta estará pronta para receber os dados dos sensores e realizar a classificação.

Na interface *on-line* o processamento é realizado em ciclos de 30 segundos, que é o período de envio dos dados do sistema RFpat. A cada classificação no modo *on-line*, ou seja, a cada ciclo, é gerado um relatório e um gráfico contendo ou não as informações de situações alarmantes, como se pode ver na Figura 5-10.

A Figura 5-10-(a) exibe um relatório com a classificação de quatro ciclos de envio/classificação dos dados, sendo que no primeiro ciclo não havia problemas com o paciente,

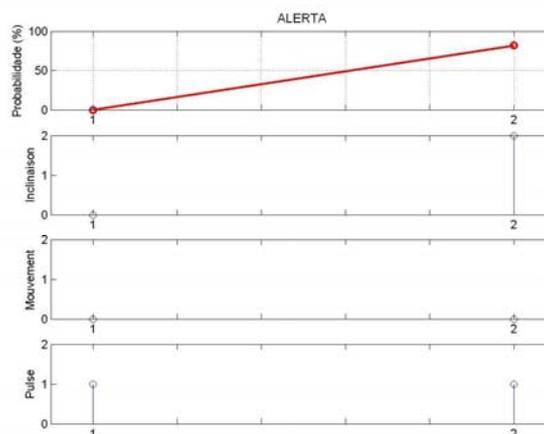
no segundo ciclo já surge o estado ‘Atenção!’ (com probabilidade de 64,8% de ocorrência de uma situação de risco), e nos dois últimos ciclos já aparece o estado ‘Alerta!’ (devido à permanência do estado ‘Atenção!’).

```

*** Para parar o programa escreva "stop(t)" ***,
Atencao! --> 64.8 % - 7/6/2006, 11:14:17
*** Para parar o programa escreva "stop(t)" ***,
Atencao! --> 64.8 % - 7/6/2006, 11:14:17
Atencao! --> 64.8 % - 7/6/2006, 11:14:47
*** ALERTA! --> 64.8 % - 7/6/2006, 11:14:47 ***
*** Para parar o programa escreva "stop(t)" ***,
Atencao! --> 64.8 % - 7/6/2006, 11:14:47
Atencao! --> 64.8 % - 7/6/2006, 11:15:17
*** ALERTA! --> 64.8 % - 7/6/2006, 11:15:17 ***
*** Para parar o programa escreva "stop(t)" ***,
>> stop(t)
>> |

```

(a)



(b)

Figura 5-10 – Exemplo de saída da fusão em modo *on-line* - (a) relatório de classificação, (b) gráfico de probabilidades do nó CONTROLE.

A Figura 5-10-(b) exibe os dois últimos valores recebidos pelos sensores (pontos ‘1’ e ‘2’ do eixo das abscissas), e a respectiva saída da rede em probabilidade. Neste gráfico, os campos Inclinação, Movimento e Pulsação indicam os estados para as entradas da rede (neste exemplo sem o sensor B5), de acordo com o valor indicado no eixo das ordenadas. Para Inclinação, ‘0’ indica posição vertical e ‘2’ indica posição horizontal. Para Movimento, ‘0’ indica movimento abaixo de 70% e ‘2’ indica movimento acima de 70%. Para Pulsação, ‘0’ indica bradicardia, ‘1’ indica pulsação normal e ‘2’ indica taquicardia.

5.4 Etapa de Treinamento

Para a etapa de treinamento são necessários um algoritmo de aprendizado e uma base de dados, onde esta última corresponde ao domínio do problema modelado pela rede. No caso

específico deste projeto, a base de dados contém sinais correspondentes a situações de queda e de problemas cardíacos, obtidos de casos registrados anteriormente. O objetivo final da etapa de treinamento é obter as probabilidades que melhor correspondam à realidade do problema modelado pela Rede Bayesiana.

Em Redes Bayesianas para auxílio à decisão médica, o principal obstáculo em utilizar o aprendizado é a dificuldade na obtenção da base de dados, pois a coleta dos dados que possam auxiliar a construção de sistemas especialistas não é uma prática comum no atendimento médico, e quando existente é de difícil acesso.

Para se obter os dados de treinamento, foram realizadas simulações de cenários para situações de queda utilizando o sistema de monitoramento. Para realizar as simulações, uma pessoa transporta consigo o terminal portátil RFpat na sala de teste onde está instalado o sistema GARDIEN e simula situações e movimentos pré-determinados, por instantes de tempo fixo. Entretanto, para simular taquicardia e/ou bradicardia seria preciso o apoio de voluntários que sofram desses problemas. Dessa maneira, foi utilizado o programa de simulação para obter os dados referentes à bradicardia e taquicardia. Como o processo de aprendizado é realizado também com dados reais, o perfil comportamental do usuário, caracterizado dentro de suas condições habituais de vida, é levado em consideração.

Foram gerados quatro tipos de dados de treinamento: inclinação, movimentação, frequência cardíaca e queda¹. Os dados são coletados e comparados com o esquema traçado para o cenário, projetando assim uma tabela com os valores planejados (os rótulos) e os valores medidos. Essa tabela é então manipulada e formatada para se usar na aprendizagem da Rede Bayesiana.

Durante os primeiros testes de montagem do protótipo RFpat foram gerados dados inconsistentes devido a algumas falhas nas medições (hardware). Dessa forma, foram necessárias algumas mudanças em certas situações nitidamente anormais. Como exemplo, um trecho da tabela com os dados referentes às situações de queda originais e os equivalentes modificados está mostrado na Tabela 5-2.

A estimação das variáveis individualmente não modela adequadamente as relações de dependência entre as variáveis da rede (relações de causa e efeito). Portanto, foi implementada uma etapa de aprendizagem onde todos os parâmetros da rede são ajustados a partir de um conjunto de amostras para treinamento. A estratégia de treinamento adotada considera que

¹ Para a inclinação, em todos os dados de treinamento foi utilizado o sensor de inclinação do RFpat, e em algumas situações foram utilizados os sensores B e 5 do GARDIEN.

apenas alguns nós são observáveis (conseqüentemente, os nós restantes são considerados não-observáveis), dado o conjunto de amostras disponíveis para treinamento. Os valores dos nós não-observáveis são estimados pelo método de treinamento (via inferência), sendo os valores mais apropriados computados com respeito aos eventos observados.

Tabela 5-2 – Dados de aprendizagem modificados.

Dados de Aprendizagem				
Inclinação	Movimentação	Pulsação	Queda original	Queda modificada
vertical	mais 70%	normal	T	F
vertical	mais 70%	normal	T	F
horizontal	menos 70%	normal	T	T
horizontal	menos 70%	normal	T	T
horizontal	menos 70%	normal	T	T
horizontal	menos 70%	normal	T	T
horizontal	menos 70%	normal	F	T
horizontal	menos 70%	normal	F	T

Em virtude de termos nós não-observáveis (falta de informação), foi necessário utilizar algoritmo EM (*Expectation-Maximization*)¹ para treinar a rede, baseando-se no conjunto de dados de treinamento. O algoritmo EM realiza o treinamento de maneira iterativa, onde a cada iteração ele realiza dois passos: o primeiro consiste em inferir o valor sobre os nós não-observáveis a partir das observações obtidas do experimento, enquanto que no segundo passo ele re-estima (maximiza) os valores das probabilidades.

O número de iterações necessárias para a convergência do algoritmo de treinamento do *toolbox* BNT, que estabelece o critério de parada, é definido pelo limiar estipulado, de acordo com a relação mostrada na Equação 5-1.

$$\frac{|f(t) - f(t-1)|}{TM} < \text{limiar} \quad , \quad \text{Eq. 5-1}$$

sendo $f(t)$ o logaritmo da verossimilhança e $f(t-1)$ o logaritmo da verossimilhança do instante anterior. TM é o termo médio, dado por:

$$TM = \frac{(|f(t) + f(t-1)|)}{2} \quad , \quad \text{Eq. 5-2}$$

¹ A função de aprendizado utilizada do *toolbox* BNT (MURPHY, 2003) foi a '*learn_params_em*', empregada geralmente em situações que haja perdas de dados ou variáveis escondidas (observação parcial). Esta função utiliza o algoritmo EM para realizar a estimativa do ponto, ou seja, utiliza os métodos *Maximum Likelihood* ou *Maximum a Posteriori*.

O esquema do processo de aprendizagem da rede está apresentado na Figura 5-11.

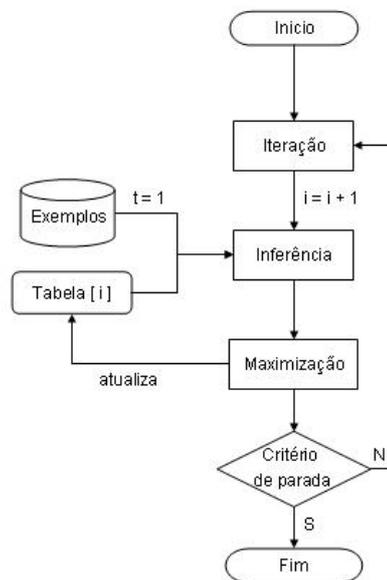


Figura 5-11 – Esquema do processo de aprendizagem da rede.

CAPÍTULO 6 – SIMULAÇÃO DE DADOS

Neste trabalho, foi utilizado um processo de simulação para a geração de um grande número de seqüências temporais representativas de um comportamento habitual de uma pessoa. Esse processo se situa como uma parte integrante do esquema de decisão (ver Figura 2-2). A utilização da simulação é condicionada pelo contexto e pelos objetivos do problema, a fim de produzir dados apropriados à experimentação dos algoritmos de decisão. A implementação e a análise de todo o sistema de decisão necessita de conjuntos de dados realistas e adaptados à problemática estudada. Esses conjuntos de dados podem ser constituídos a partir de experimentações realizadas em campo e/ou geradas por simulação.

No domínio do telemonitoramento de pacientes em domicílio, as experimentações em um ambiente realista são, no início, bastante difíceis. Não se pode dispor de conjuntos de dados suficientemente realistas e completos para constituir a base de dados de sistemas de decisão sobre os hábitos de vida de uma pessoa em domicílio. Além do mais, um estudo completo e confiável necessita levar em conta vários perfis de pessoas, em face de vários tipos de situações. Isto permite analisar o sistema de decisão, testando de maneira ideal todos os casos que o sistema deverá ser capaz de tratar. De qualquer forma, a constituição por experimentação destes conjuntos de teste suficientemente completos e representativos se mostra, dentro do contexto do trabalho, uma tarefa difícil: observar, dentro das condições “normais” de vida em domicílio, um grande número de pessoas, tendo em vista todo um conjunto de situações. É por esta razão que se utiliza a simulação, como um meio de contornar as dificuldades de constituição de numerosos conjuntos de dados experimentais completos e representativos. A utilização de dados experimentais registrados em um ambiente realista se torna, contudo, absolutamente indispensável como uma segunda etapa de validação dos sistemas de decisão. Assim, a utilização da simulação constitui uma primeira etapa de testes. A validação do sistema de simulação foi realizada, em um primeiro momento, por meio da comparação com sinais reais.

As vantagens da simulação de dados se situam em vários níveis:

- Permite a produção de grandes conjuntos de dados para uma experimentação de algoritmos de decisão;

- Gera dados representativos de um máximo de situações para tratamento pelo sistema de decisão: vários perfis de pessoas e de situações em um domicílio, dentro do contexto do trabalho;
- Possibilita a construção de um processo confiável e facilmente compreensível por qualquer agente do sistema;
- Proporciona o aprimoramento do conhecimento das probabilidades condicionais (*a posteriori*) dos parâmetros estudados, com a análise das tendências e variações conjuntas de diferentes parâmetros observados simultaneamente dentro do contexto do telemonitoramento de pacientes em domicílio;
- Facilita o teste da eficácia e da robustez dos algoritmos de decisão, fazendo variar os parâmetros da simulação para a geração de dados, cobrindo um máximo de casos possível.

6.1 Definição do Problema

A falta de dados experimentais dentro do nosso contexto de pesquisa de uma parte, e as vantagens do uso de dados simulados de outra parte, motivou a utilização de um processo de simulação. O objetivo é, portanto, a construção de um modelo em que a modificação de alguns parâmetros permita simular diferentes perfis comportamentais, isto é, gerar dados associados ao monitoramento de diferentes pessoas. Modificações progressivas dos parâmetros permitem também simular a evolução de cada pessoa para diferentes situações.

Para se desenvolver um algoritmo de simulação são levadas em conta informações relacionadas a diferentes domínios. No caso particular deste trabalho é considerada a área médica, tanto para a construção do modelo como também para sua validação. O estudo inicial necessita de conhecimentos variados com respeito, por exemplo, aos hábitos da vida cotidiana de uma pessoa, as características dos parâmetros fisiológicos que poderão ser gerados em domicílio, ou ainda o método de análise estatística de dados experimentais.

Devido à falta de um profissional da área médica para a assistência na escolha dos parâmetros de simulação, e à necessidade urgente de um simulador para testes, foi desenvolvido um algoritmo com base em dados médicos obtidos de trabalhos na área, informações de professores com conhecimento do assunto e com base no bom senso.

O ponto principal do programa de simulação é a geração de dados anormais para a frequência cardíaca, ou seja, bradicardia e taquicardia, com a construção de um arquivo patológico realista a partir de dados normais perturbados artificialmente por diversos perfis de situações alarmantes. Como comentado anteriormente, a bradicardia vem a ser uma redução do número de batimentos cardíacos por minuto, enquanto a taquicardia representa um aumento do número de batimentos cardíacos por minuto. Particularmente nesse quesito, o procedimento de simulação se torna extremamente necessário pela dificuldade de obter dados reais para teste. Além da simulação da frequência cardíaca, o programa simula ainda o nível de inclinação e a taxa de movimentação da pessoa monitorada, de acordo com parâmetros estipulados pelo hardware do sistema (faixa de precisão, nível de sensibilidade, etc.).

6.2 O Programa de Simulação Desenvolvido

A partir das necessidades relacionadas à etapa de testes do protótipo de fusão de dados dos sensores do Sistema TELEPAT (utilizando Redes Bayesianas), foi desenvolvido o programa de simulação de dados. Deve-se deixar claro que esse programa de simulação foi criado tendo em vista as necessidades do projeto de pesquisa, com suas características e particularidades, apesar de que se pode utilizá-lo para outros fins, já que os sinais gerados correspondem com a realidade.

O objetivo principal é o de simular os sinais fornecidos pelo sistema RFpat de acordo com uma situação particular (determinada pelo usuário do programa), ou seja, simular sinais de inclinação, movimentação e frequência cardíaca, com o intuito de reproduzir um cenário em que o paciente esteja utilizando o dispositivo portátil.

O programa, desenvolvido inteiramente por meio da plataforma Matlab versão 6.5, é dividido em etapas, e para cada uma delas são mostradas na tela do computador as opções de escolha disponíveis ao usuário, sendo que a alternativa escolhida terá influência direta na geração dos sinais. A representação gráfica da organização do programa de simulação, com a indicação dos arranjos e as inter-relações de suas unidades constitutivas (e o limite das atribuições de cada uma delas), está retratada na Figura 6-1.

A descrição dos sinais de saída do programa de simulação, além das opções disponíveis pelo programa para cada um desses sinais, está comentada nas seções a seguir.

6.2.1 *Inclinação*

Os dados de saída para esse tipo de sinal são do tipo binário, como os fornecidos pelo terminal RFpat: o nível lógico ‘1’ representa uma pessoa na posição “deitada”, e o nível lógico ‘0’ representa uma pessoa na posição “de pé”. As opções de escolha para a simulação de inclinação são:

- Utilizar dados já registrados: extrai os valores relacionados à inclinação do indivíduo de um arquivo escolhido de um banco de dados específico (ver Figura 6-4), sendo esse banco de dados composto por arquivos com dados biométricos reais ou simulados;
- Criar um novo perfil: neste caso, o usuário será levado a escolher um valor ‘0’ ou ‘1’ para a geração do sinal. Esse modo vem a ser útil em um teste forçado de queda do paciente, por exemplo, adotando o valor ‘1’ para inclinação para fazer a análise dos sinais de movimentação e frequência cardíaca.

O bloco que indica ‘Tempo total’, que pode ser visto em vários locais do organograma, é a representação de uma função que solicita ao usuário o período total desejado para a simulação. Esse período é então convertido em número de dados de simulação, sendo empregada uma taxa de amostragem de 30 segundos (que é o período do maior ciclo de envio, utilizado pelo terminal RFfpat – como mostrado na seção 4-2).

6.2.2 *Movimentação*

Assim como no terminal RFpat, os dados gerados para a movimentação variam em uma escala que vai de 0 a 15, dependendo do nível de agitação. As opções de escolha para a simulação da agitação são:

- Utilizar dados já registrados: extrai os valores relacionados à agitação do indivíduo de um arquivo escolhido de um banco de dados específico, conforme o procedimento adotado na etapa de escolha do sinal de inclinação.

- Utilizar dados simulados com valores padrão: é gerado um sinal de agitação com valores de média e desvio padrão pré-definidos. O nível médio de agitação e o desvio padrão da agitação fixados foram 7,5 e 9, respectivamente.
- Criar um novo perfil: é gerado um sinal de agitação, sendo que os valores de média e desvio padrão são determinados pelo usuário.

Os valores de média e desvio padrão utilizados como valores padrão foram obtidos extraíndo a média aritmética das médias e dos desvios padrão de sinais normais (obtidos com o terminal RFpat).

Para a simulação utilizando tanto a criação de um novo perfil quanto os valores padrão, é gerado um conjunto de dados aleatórios normalizados, com distribuição normal (com média 0 e desvio padrão 1). A Equação 6-1 mostra a aplicação da média e desvio padrão no conjunto de dados normalizados,

$$M = \mu + \sigma \cdot A \quad , \quad \text{Eq. 6-1}$$

sendo M o novo conjunto de dados referentes à Movimentação, μ a média desejada, σ o desvio padrão desejado e A é uma variável aleatória com distribuição normal (média zero e variância um). Nota-se que os dados gerados para a Movimentação não dependem da variável inclinação, ficando a critério do usuário a combinação destas duas variáveis.

Assim como o dispositivo RFpat, que limita os valores de saída em 0 e 15 para sinais referentes à agitação do indivíduo, os sinais simulados adotam esses mesmos parâmetros. Dessa forma, após a simulação da agitação, todos os valores acima de 15 são fixados em 15, e todos os valores abaixo de 0 são fixados em 0.

6.2.3 *Frequência Cardíaca*

As opções de escolha para a simulação da frequência cardíaca são:

- Utilizar dados já registrados: da mesma forma que na escolha dos sinais de inclinação e movimentação, os valores relacionados à frequência cardíaca são extraídos de um arquivo escolhido de um banco de dados específico.

- Criar um novo perfil "Normal": é realizada a simulação com o objetivo de se obter valores normais de frequência cardíaca, ou seja, são gerados sinais relativos a uma pessoa saudável.
- Criar um novo perfil "Anormal": é realizada a simulação com o objetivo de se obter valores de frequência cardíaca relacionados a um dos dois tipos de alterações dos batimentos cardíacos: bradicardia ou taquicardia.

No modo ‘Perfil Normal’, existem duas opções de escolha: ou utilizar um ‘Modelo de Atividade Cardíaca’ ou um modelo chamado ‘Cosinor’ (DUCHÈNE, 2004).

6.2.3.1 Modelo de Atividade Cardíaca

O método utilizado para a simulação dos valores de frequência cardíaca é determinado, de certa forma, também pelo empirismo, além de conhecimentos técnicos da área. Para uma simulação mais próxima possível do real é necessário considerar uma distribuição específica de valores de frequência cardíaca em função da postura e do nível de atividade (sendo que o valor médio da frequência observada dentro de um contexto dado é uma função do nível de atividade, e esta é igualmente específica à postura). Dessa forma, a construção de um modelo mais realista levaria em conta a sensibilidade da frequência cardíaca à postura da pessoa e à intensidade das atividades físicas realizadas. Além das variações intra-individuais (que são as variações que ocorrem para um mesmo indivíduo), existe uma grande variabilidade inter-individual (que são as variações que ocorrem entre um indivíduo e outro), em particular as variações da frequência cardíaca média no estado de repouso. Nesse caso é que entra o Modelo de Atividade Cardíaca. Assim, a simulação das variações da frequência cardíaca compreenderia:

- Variação de Repouso ($F_{Crep}(t)$), que seriam as seqüências temporais correspondentes às variações da frequência cardíaca da pessoa em repouso;
- Custo Cardíaco da Atividade ($\Delta FC(t)$), que seriam variações introduzidas aos valores da frequência cardíaca de repouso, em função da postura e do nível de atividade observados no mesmo momento.

Um valor de Custo Cardíaco de Atividade pode ser determinado a cada instante a partir de uma distribuição apropriada em função do trabalho realizado, estimado pelo nível de atividade para cada tipo de postura.

A Variação da Frequência Cardíaca, $Fc(t)$, seria então obtida reunindo as Variações de Repouso às variações do Custo Cardíaco da Atividade:

$$Fc(t) = Fc_{rep}(t) + \Delta Fc(t) \quad \text{Eq. 6-2}$$

A função para realizar uma simulação utilizando o Modelo da Atividade Cardíaca não foi implementada nesse programa de simulação. Entretanto, a infra-estrutura para suportar tal função já está desenvolvida.

6.2.3.2 Modelo Cosinor

O método mais utilizado pelos cronobiologistas (pessoas que estudam a variação dos fenômenos biológicos com o tempo, em especial aqueles que são recorrentes, como a frequência cardíaca) para a análise de ritmos circadianos (ou seja, durante um dia) em dados experimentais coletados em um período de 24 horas é a técnica do Cosinor. Essa técnica utiliza, para dados coletados durante um período de 24 horas, uma aproximação senoidal para a representação da função. Esse método é aplicado porque nem sempre se necessita de uma grande precisão na estimação dos parâmetros do ritmo cardíaco observado. Além disso, a estimação de um sinal que represente a condição de repouso para uma pessoa deve ser registrada no curso de 24 horas, com um baixo nível de atividade, o que dificulta um ensaio prático.

O método Cosinor foi utilizado para a estimação do ritmo circadiano da frequência cardíaca de repouso para cada pessoa. Este método, para análise de ritmos biológicos, é utilizado na identificação de ritmicidade no comportamento de um sinal, e consiste no ajuste da curva cosseno aos dados, o que possibilita descrever as séries temporais de indivíduos por meio de parâmetros rítmicos estimados pelo método dos mínimos quadrados, por exemplo, identificando parâmetros como acrofase e MESOR (AFONSO, 2006). O Método dos Mínimos Quadrados é uma técnica de otimização matemática que procura encontrar o melhor ajuste para um conjunto de dados, tentando minimizar a soma dos quadrados das diferenças entre a curva ajustada (no caso uma cossenóide) e os dados (no caso as medições da frequência car-

díaca). O método Cosinor também serve para verificar se a série temporal apresenta algum tipo de recorrência periódica.

As características e variáveis desse modelo são as seguintes (ver Figura 6-2):

1. MESOR (*Midline Estimating Statistics Of Rythm*): representa o nível médio em torno do qual os valores de freqüência oscilam;
2. Amplitude: medida do máximo valor das variações possíveis em um sinal completo;
3. Acrofase: momento do dia em que a aproximação senoidal do ritmo cardíaco alcança o seu valor máximo.

A variação temporal dos valores de freqüência cardíaca para o repouso, $F_{c_{rep}}(t)$, é então exprimida em função desses três parâmetros, MESOR, amplitude e acrofase:

$$F_{c_{rep}}(t) = ME + A \cdot \text{sen}\left(\frac{2\pi}{24} \cdot t + \phi\right), \quad \text{Eq. 6-3}$$

onde ME representa o MESOR, A representa a amplitude, t é o tempo exprimido em horas e ϕ representa a fase das variações, em unidades trigonométricas. A Equação 6-3 é específica para um período de 24 horas.

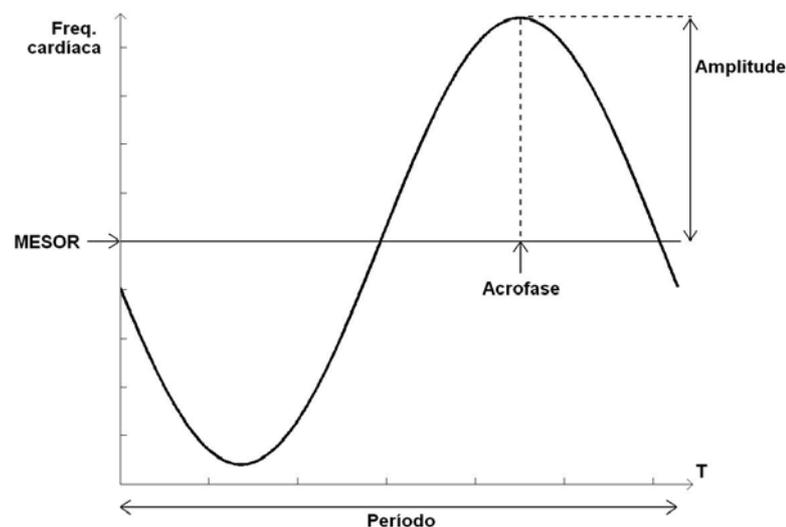


Figura 6-2 - Método Cosinor - análise do ritmo senoidal.

De acordo com as análises realizadas a respeito do método cosinor, os valores para os diferentes parâmetros são de aproximadamente:

- MESOR \cong 65 bpm;

- Amplitude entre 0,7 e 13,9 bpm;
- Acrofase entre 10h40min e 19h30min, para o período médio de ocorrência do valor mais forte em um dia.

6.2.3.3 Perfil Anormal

No modo Perfil Anormal, existem duas opções de escolha: ou utilizar a opção *taquicardia* (para simular situações relacionadas a um aumento anormal da frequência cardíaca) ou a opção *bradicardia* (para simular situações relacionadas a uma diminuição anormal da frequência cardíaca). No esquema mostrado na Figura 6-1 está representado apenas o fluxograma referente à opção ‘bradicardia’, já que o fluxograma, e conseqüentemente as funções do programa, para a opção de simulação de taquicardia, são similares.

Assim, analisando as opções ‘bradicardia’ e ‘taquicardia’ conjuntamente, pode-se gerar um novo perfil ou utilizar como base de simulação um arquivo normal já registrado anteriormente (escolhido de um banco de dados específico).

A Figura 6-3 apresenta como exemplo uma simulação de frequência cardíaca com dois eventos de bradicardia, sendo que para cada um desses eventos há as frequências e tempos correspondentes, ou seja, índice 1 para episódio 1 e índice 2 para episódio 2 (f_{d1} e t_{d1} para episódio 1, f_{d2} e t_{d2} para episódio 2, etc.). Os sinais simulados, tanto para a opção de geração de um novo perfil quanto para a opção de utilizar um arquivo normal, apresentam cinco partes distintas (ver Figura 6-3), que são definidas da seguinte forma:

1. Início do sinal (estado normal (trecho A)): pode ser um novo sinal do tipo normal (quando a opção é a de criar um novo perfil) ou um trecho inicial de um arquivo existente (quando a opção é a de usar um arquivo normal já registrado);
2. Transição normal/anomalia (trecho B): representa a passagem do estado normal para o estado anormal (bradicardia ou taquicardia);
3. Duração da anormalidade (trecho C): período em que ocorre a anomalia, ou seja, ou a fase em que a frequência está mais baixa, no caso de bradicardia, ou a fase em que a frequência está mais alta, no caso de taquicardia;
4. Transição anomalia/normal (trecho D): representa a passagem do estado anormal para o estado normal;

5. Fim do sinal (trecho E): pode ser um novo sinal do tipo normal (quando a opção é a de criar um novo perfil) ou o término de um arquivo existente (quando a opção é a de usar um arquivo normal já registrado);

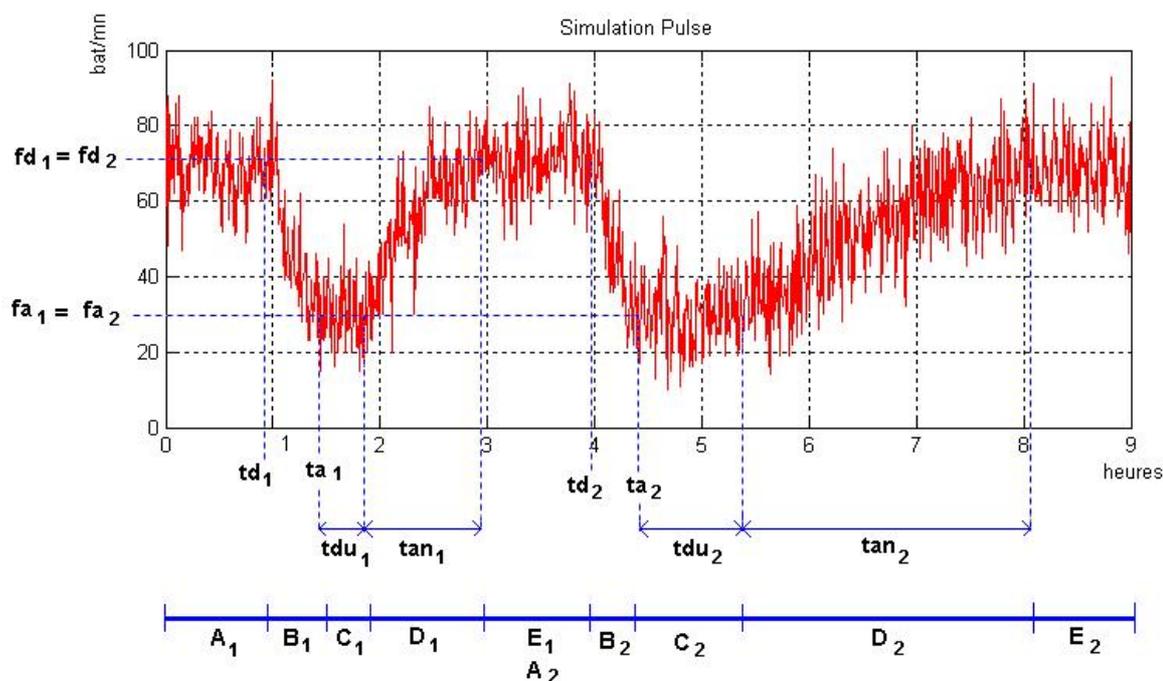


Figura 6-3 – Parâmetros de simulação (bpm x horas) – exemplo de bradicardia.

De acordo com a opção escolhida para a simulação de dados de frequência cardíaca, o usuário será levado a determinar alguns (ou a totalidade) desses parâmetros (ver Figura 6-3, para um exemplo de bradicardia):

- A frequência de início (f_d), que significa a frequência de início da transição normal/anomalia;
- O tempo de início (t_d), que significa o tempo de início da transição normal/anomalia;
- A frequência de chegada (f_a), que significa a frequência do término da transição normal/anomalia, ou seja, a frequência da anomalia propriamente dita;
- O tempo de chegada (t_a), que significa o tempo do término da transição normal/anomalia;
- A duração da anomalia (t_{du}), que significa o período de tempo que ocorre a anomalia, ou seja, é o tempo total em que a frequência fica no valor da frequência de chegada f_a ;

- A duração da transição anomalia/normal (t_{an}), que é o período de tempo que ocorre na transição do estado de anomalia para o estado normal.

Além disso, deve-se determinar o tipo da transição anomalia/normal, sendo que a passagem do estado de anomalia para o estado normal deve obedecer a uma função do tipo senoidal ou a uma função do tipo quadrática, e determinar o ruído aditivo gaussiano (et), que é o valor do desvio padrão aplicado ao sinal. A escolha das funções senoidal e quadrática se deve ao fato de que se desejava uma alteração dos batimentos cardíacos o mais próximo do real.

Na opção de geração de um novo perfil o usuário deverá estabelecer a duração total do sinal de saída, escolher quantos eventos da anomalia (sendo que anomalia, nesse caso, significa ou bradicardia ou taquicardia) ele deseja incluir nesse sinal, e definir os parâmetros do sinal (como os descritos acima) para cada evento da anomalia.

Neste modo de simulação, são apresentados ao usuário valores de frequência e desvio padrão que exercem a função de auxílio, como referência. Os valores fornecidos como referência para a simulação de bradicardia são:

- Frequência normal (início da transição normal/bradicardia) = 70 bpm;
- Frequência de bradicardia = 20 bpm;
- Valor do ruído gaussiano (desvio padrão) = 9.

Para a simulação de taquicardia, os valores fornecidos como referência são:

- Frequência normal (início da transição normal/taquicardia) = 70 bpm;
- Frequência de taquicardia = 140 bpm;
- Valor do ruído gaussiano (desvio padrão) = 9.

Estes parâmetros foram utilizados por serem característicos em situações de bradicardia e taquicardia, de acordo com as referências pesquisadas para esta dissertação. O valor para a frequência na simulação de uma bradicardia é limitada a 10 bpm (valor mínimo), assim como o valor para a frequência na simulação de uma taquicardia é limitada a 160 bpm.

A função do tipo senoidal, utilizada na transição anomalia/normal, obedece a uma equação do tipo:

$$F(x) = M + A \cdot \text{sen}(x) \quad , \quad \text{Eq. 6-4}$$

sendo M a média e A a amplitude do sinal desejado.

A função quadrática não foi implementada no programa de simulação, já que a função senoidal atendeu às necessidades de simulação. Entretanto, a infra-estrutura para suportar a função quadrática (ou outra função desejada) no programa de simulação já está desenvolvida.

No caso da opção de simulação tomando como base um arquivo normal já registrado, será adicionado um sinal de taquicardia ou bradicardia no mesmo, obedecendo às opções ‘adicionar’ ou ‘substituir’, que serão detalhadas a seguir:

- ‘Adicionar’: Nessa opção, os intervalos ‘transição normal/anomalia’, ‘duração da anormalidade’ e ‘transição anomalia/normal’, do sinal simulado, são acrescentadas ao sinal normal já registrado, num certo instante, sem modificá-lo por completo. As alterações que poderão ocorrer no sinal normal registrado são: o valor da frequência do sinal antes da anomalia e o valor da frequência do sinal após a anomalia. Assim, pode haver um deslocamento da média da frequência do sinal real, tanto para mais quanto para menos, dependendo do valor de entrada escolhido pelo usuário para a ‘frequência de início’.
- ‘Substituir’: Nessa opção, os intervalos ‘transição normal/anomalia’, ‘duração da anormalidade’ e ‘transição anomalia/normal’, do sinal simulado, substituem os valores do sinal normal já registrado, sendo que a dimensão dessa substituição (ou seja, o intervalo de tempo do sinal normal que será substituído) é dependente do tamanho do sinal simulado e do ponto de inclusão desse sinal no arquivo normal. Da mesma forma que a opção ‘adicionar’, pode haver um deslocamento da média da frequência do sinal real, tanto para mais quanto para menos, dependendo do valor de entrada escolhido pelo usuário para as ‘frequência de início’ e para a frequência após o intervalo t_{du} da anomalia, sendo esta última a frequência utilizada apenas nesta opção de simulação. Caso o período total da simulação (definido inicialmente pelo usuário – função ‘Tempo total’) seja maior que o período efetivo de simulação (que vem a ser a duração da parte utilizada do arquivo normal mais a duração da anomalia simulada), há a necessidade de continuar a simulação de alguma forma, seja em um estado normal, seja em um estado anormal. Para isso, existem duas alternativas:
 - ‘Retornar’: Nessa opção, após a duração da anomalia, o sinal retorna a um estado normal (ou seja, é utilizada uma transição do tipo ‘anomalia/normal’ no sinal). O valor de frequência para retorno é estipulado pelo usuário.

- ‘Permanecer’: Nessa opção, a média do sinal após o intervalo ‘duração da anormalia’ permanece a mesma, ou seja, a frequência é inalterada. Dessa forma, a anormalidade continua até que se termine o período total da simulação.

O ruído aplicado nas simulações de frequência cardíaca (chamado de ruído branco gaussiano) obedece ao método (ou distribuição) de Box-Muller (DONADIO, 2007, e WEISSTEIN, 2007), que é uma forma rápida e prática de gerar uma distribuição normal a partir de uma distribuição uniforme. Sejam x_1 e x_2 valores independentes gerados por uma distribuição uniforme entre 0 e 1. Então z_1 e z_2 (como definido pelas Equações 6-5 e 6-6, que representam a distribuição de Box-Muller) têm distribuição normal com média $\mu = 0$ e variância $\sigma^2 = 1$.

$$z_1 = \sqrt{-2 \cdot \ln(x_1)} \cdot \cos(2 \cdot \pi \cdot x_2) \quad \text{Eq. 6-5}$$

$$z_2 = \sqrt{-2 \cdot \ln(x_1)} \cdot \text{sen}(2 \cdot \pi \cdot x_2) \quad \text{Eq. 6-6}$$

Os valores z_1 e z_2 podem ser facilmente modificados para diferentes médias e variâncias, como mostra as Equações 6-7 e 6-8:

$$z_1' = \mu + \sqrt{\sigma^2} \cdot z_1 \quad \text{Eq. 6-7}$$

$$z_2' = \mu + \sqrt{\sigma^2} \cdot z_2, \quad \text{Eq. 6-8}$$

sendo μ a média e σ^2 a variância desejadas.

Para algumas opções de simulação de frequência cardíaca que não seja a criação de um novo perfil, os valores de média e variância do sinal simulado são obtidos pela extração da média e variância do sinal real normal já registrado. Esses valores são utilizados para, por exemplo, adicionar ruído nos valores simulados, tornando-os os mais reais possíveis.

Porém, pode ser que o sinal real já registrado apresente algumas irregularidades (que pode ter ocorrido durante a execução do registro), o que faria com que o ruído do sinal simulado não se assemelhe ao ruído do sinal real. Uma funcionalidade disponível para a simulação de frequência cardíaca é a possibilidade de atenuar o nível de ruído do sinal gerado como saída. Como exemplo, na simulação em que se adiciona uma bradicardia em um sinal normal

real, utilizando como base para o ruído da bradicardia o desvio padrão do sinal real, se esse sinal real variar com muita intensidade (o que geralmente ocorre no início da medição, ocasionando um aumento do desvio padrão) o ruído do sinal de bradicardia simulado será maior que o desejado, pois o desvio padrão do sinal real será maior. Dessa forma, pode-se diminuir o ruído até um nível desejado, tendo como auxílio um gráfico com o resultado da alteração.

6.2.4 Características do Programa de Simulação

A Figura 6-4 exibe o esquema da aquisição de dados pelo programa de simulação, mostrando que o processo de simulação pode utilizar tanto sinais normais como sinais simulados.

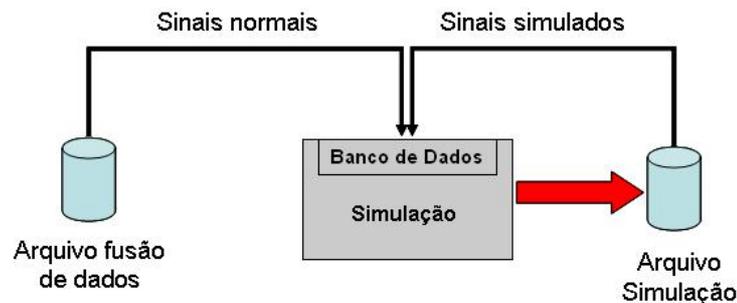


Figura 6-4 – Programa de simulação – aquisição de dados.

Após a simulação dos três dados biológicos desejados (inclinação, movimentação e frequência cardíaca) é disponibilizada para o usuário (que deseja utilizar esses dados no programa de fusão de dados biológicos ou utilizá-los em outra aplicação) a opção de criação de um arquivo. O modelo de um arquivo gerado pela simulação está mostrado na Figura 6-5. Esse arquivo é gerado e armazenado em um banco de dados de simulação ('Arquivo Simulação', na Figura 6-4) com extensão "day", e contém dados referentes à data (formato 'dia-mês-ano') e horário (formato 'hora:minuto:segundo') de registro dos dados, além dos valores de inclinação, movimentação e pulsação relacionados ao horário de medição. O modelo do arquivo de saída está vinculado às necessidades de outros programas que fazem parte do sistema TELEPAT, assim como também está vinculado aos requisitos de entrada do programa de fusão de dados biológicos (Rede Bayesiana). O restante dos dados (as outras colunas do arquivo de saída, com valores zero) não especificados aqui, não interfere no funcionamento do programa de fusão de dados, e estão relacionados a outros programas que irão utilizar esse mesmo arquivo de dados. O horário indicado em cada linha do arquivo de dados tem um atraso de

30 segundos em relação ao horário da linha anterior, o que representaria uma medição com um ciclo de geração de 30 segundos (que é a mesma taxa de transmissão adotada pelo terminal RFpat).

Data	Horário	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5	Coluna 6	Coluna 7	Coluna 8
07-06-2006	11:04:18	0	0	0	0	15	042		
07-06-2006	11:04:48	0	0	0	0	15	041		
07-06-2006	11:05:00	0	0	0	0	15	061		
07-06-2006	11:05:30	0	0	0	0	15	082		
07-06-2006	11:06:00	0	0	0	0	15	081		
07-06-2006	11:06:30	0	0	0	0	15	077		
07-06-2006	11:07:00	0	0	0	0	15	077		
07-06-2006	11:07:30	0	0	0	0	15	082		
07-06-2006	11:08:00	0	0	0	0	14	082		
07-06-2006	11:08:30	0	0	0	0	15	083		
07-06-2006	11:09:00	0	0	0	0	15	082		
07-06-2006	11:09:30	0	0	0	0	15	080		
07-06-2006	11:10:00	0	0	0	0	15	076		
07-06-2006	11:10:30	0	0	0	0	15	078		
07-06-2006	11:11:00	0	0	0	0	15	081		
07-06-2006	11:11:30	0	0	0	0	15	082		
07-06-2006	11:12:00	0	0	0	0	15	081		
07-06-2006	11:12:30	0	0	0	0	15	081		
07-06-2006	11:13:00	0	0	0	0	15	080		
07-06-2006	11:13:30	0	0	0	0	14	081		
07-06-2006	11:14:00	0	0	1	0	04	078		
07-06-2006	11:14:30	0	0	1	0	00	071		

Figura 6-5 – Modelo de arquivo gerado pela simulação.

Para a criação do arquivo de saída, os vetores gerados pelo programa de simulação, que correspondem aos dados de inclinação, movimentação e frequência cardíaca, devem ter o mesmo comprimento. Como o tamanho de cada vetor depende da opção de simulação selecionada pelo usuário, há a necessidade de se utilizar uma função que faça a correção da dimensão dos vetores. A prioridade é sempre adotar o tamanho do maior vetor para que nenhum dado simulado seja perdido. Desta forma, os vetores menores deverão se adaptar ao vetor maior, sendo necessário incrementar esses vetores com dados que não prejudiquem a real intenção de simulação dos mesmos. Assim, foram desenvolvidos dois tipos de correção da dimensão dos vetores:

1. Repetição de todos os valores: A correção dos vetores relacionados à inclinação, à movimentação e à frequência cardíaca normal do tipo Cosinor se dá com a repetição do sinal desde o seu início. Por exemplo, se for necessário adicionar 35 minutos de simulação para corrigir o vetor, haverá uma repetição dos primeiros 35 minutos de simulação desse sinal. Para a correção dos vetores relacionados à frequência cardíaca anormal (taquicardia e bradicardia), o último valor de frequência

desses vetores é tomado como referência para a criação dos dados complementares, ou seja, ele será adotado como a média constante do sinal restante. O ruído gaussiano é adicionado a esse novo sinal com o mesmo desvio padrão adotado anteriormente à correção.

2. Repetição dos últimos valores: Nesse caso, a correção dos vetores relacionados à inclinação e à movimentação se dá com a repetição do último valor dos mesmos. Por exemplo, se durante a criação do sinal o último valor de inclinação foi '1', para o restante do sinal (correspondente à correção) o valor será '1'. Para a correção dos vetores relacionados à frequência cardíaca, assim como no modo de correção com repetição de todos os valores, o último valor de frequência desses vetores é tomado como referência para a criação dos dados complementares. O ruído gaussiano é adicionado a esse novo sinal também com o mesmo desvio padrão adotado anteriormente à correção.

Os arquivos que compõem o banco de dados do qual se utilizam os dados para as simulações são obtidos com o uso do dispositivo móvel RFpat e pelos próprios arquivos simulados (ou seja, para se gerar novos arquivos pode-se fazer uso de modificações de arquivos gerados anteriormente). Cada vez que é realizada uma simulação, ao final da mesma há uma opção para o usuário de salvar ou não os dados gerados em um arquivo 'day'. Caso a opção escolhida seja a de salvar, esses dados ficarão em uma pasta comum, e estarão disponíveis para serem utilizados em uma nova simulação.

Nesta etapa, os arquivos obtidos com o uso do sistema RFpat são gerados com o objetivo de testar o equipamento e de obter dados reais para utilizar no programa de simulação. Para isso, uma pessoa saudável transporta consigo o equipamento por um período de tempo pré-determinado, para que posteriormente se possam coletar os dados considerados normais.

É importante mencionar que, dependendo do tipo de simulação realizada, caso sejam utilizados arquivos já registrados no banco de dados, os campos 'data' (informando dia, mês e ano) e 'horário' (informando horas, minutos e segundos) não serão alterados. Por exemplo, no caso de uma simulação de bradicardia com adoção do modo de substituição de dados do arquivo original para a inclusão do sinal simulado, serão levados em consideração a data e o horário inicial do arquivo original. O sinal simulado que substituirá o sinal original adotará a data e o horário do sinal substituído, não alterando o tempo total do sinal original (a não ser que haja uma correção final da dimensão dos vetores).

Em todas as etapas do programa de simulação que necessitam da entrada de dados pelo usuário, com o uso do teclado, existem funções que informam ao usuário um possível erro de digitação, sendo necessária então uma nova entrada. Isso é um recurso interessante do ponto de vista do uso do programa, caso se cometa algum erro durante a digitação dos dados exigidos.

Deve-se deixar claro que todas as funcionalidades desse programa de simulação foram desenvolvidas para atender, além do programa de fusão de dados, os requisitos do projeto TELEPAT, pois vários outros pesquisadores irão utilizar essa plataforma de simulação para outros testes. Por exemplo, a opção de simulação de frequência cardíaca anormal com substituição do arquivo normal (com as sub-opções ‘retornar’ e ‘permanecer’) foi desenvolvida para atender às necessidades específicas de um subprojeto do projeto TELEPAT.

6.3 Resultados da Simulação

São apresentados, a seguir, exemplos de resultados obtidos com o programa de simulação de dados biológicos para cada opção de simulação.

- Inclinação (ver Figura 6-6): exemplo de uma simulação utilizando um arquivo com dados normais já registrados, durante um período de 12 horas. Nota-se que o número de dados simulados para inclinação foi menor que o número de dados simulados para movimentação e/ou frequência cardíaca, pois houve correção da dimensão do vetor de dados de inclinação (neste caso houve repetição de todos os valores de inclinação).
- Movimentação (ver Figura 6-7): exemplo de uma simulação com os valores *default* (média 7,5 e desvio padrão 9,7) e de uma simulação utilizando um arquivo com dados normais já registrados (média 7,93 e desvio padrão 5,76)

Apesar de se ter utilizado uma gaussiana para modelar sinais que representem a movimentação (agitação) de uma pessoa, pode-se utilizar outras funções que representem, de outra maneira, o sinal real.

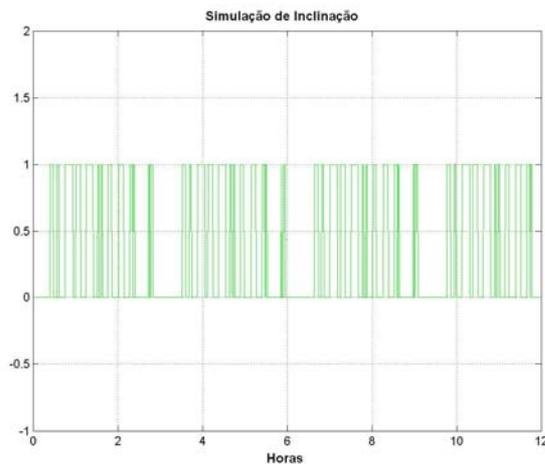


Figura 6-6 – Dados de inclinação de um arquivo normal já registrado.

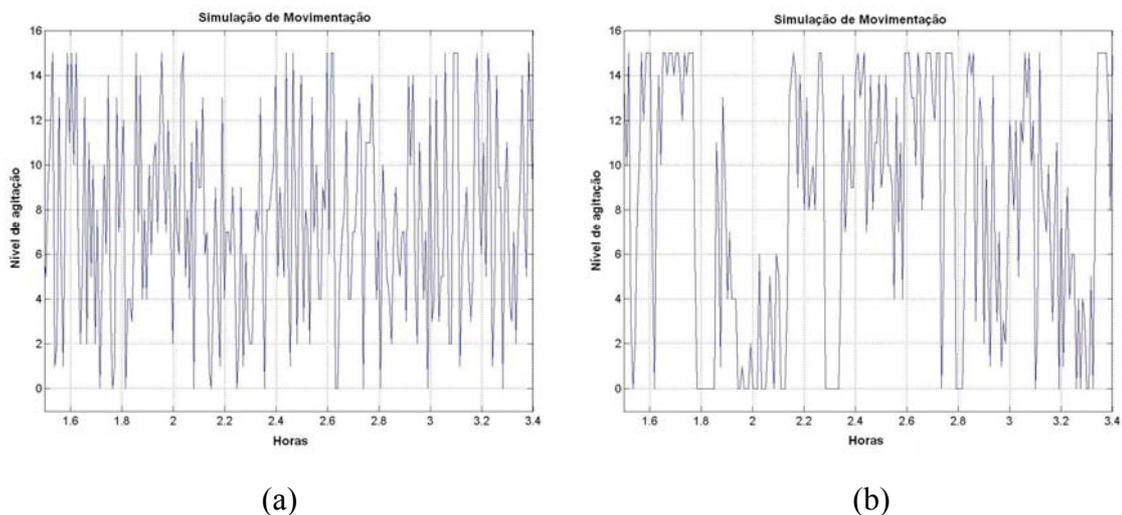


Figura 6-7 – (a) Simulação de agitação com valores *default* e (b) dados de agitação de um arquivo normal já registrado.

- Frequência cardíaca:
 - Modelo Cosinor (ver Figura 6-8): exemplo de uma simulação utilizando o modelo Cosinor, com valor de MESOR igual a 65 bpm, amplitude de 7 bpm e acrofase em 16 horas.
 - Taquicardia (ver Figura 6-9): exemplo utilizando a criação de um novo perfil de taquicardia. Nesse exemplo foram criadas duas ocorrências de taquicardia em um período de 10 horas. O valor médio escolhido para o sinal normal (assim como o valor para a frequência de início da transição normal/taquicardia para as duas ocorrências) é de 70 bpm, e a frequência de taquicardia para as duas ocorrências é de 140 bpm. Todos os intervalos de início e fim das etapas normal/taquicardia e taquicardia/normal, para as duas ocorrências, foram definidos pelo usuário. As transições taquicardia/normal

são do tipo senoidal. A duração da primeira taquicardia é de 30 minutos e a duração da segunda taquicardia é de 45 minutos. Para as duas ocorrências o desvio padrão é de 9 bpm.

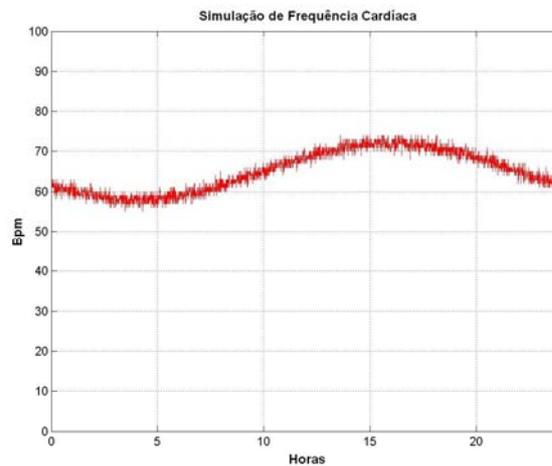


Figura 6-8 – Simulação de frequência cardíaca normal utilizando o modelo Cosinor.

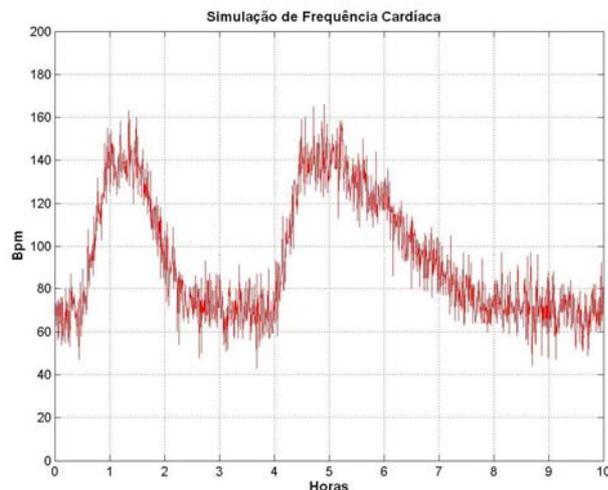


Figura 6-9 – Simulação de taquicardia utilizando a criação de um novo perfil.

- Bradycardia (ver Figura 6-10): exemplos utilizando os valores de média e variância do sinal normal tomado como base. No exemplo (a) a bradicardia é adicionada ao sinal normal (não substitui o sinal original), a frequência de início da transição normal/bradicardia é adotada como a frequência do sinal real no ponto onde começará a simulação, assim como a frequência após a transição bradicardia/normal é adotada como a frequência do sinal real do ponto onde finaliza a simulação. A duração total do sinal é de 4 horas e 18 minutos, a transição normal/bradicardia tem duração de 30 minutos, a frequência de bradicardia é de 20 bpm, a sua duração é de 40 minutos e a

transição bradicardia/normal ocorre em 20 minutos (função senoidal). No exemplo (b) a bradicardia substitui o sinal normal. A frequência de início da transição normal/bradicardia é adotada como a frequência do sinal real no ponto onde começará a simulação e, após o intervalo ‘duração da anormalidade’, a média da frequência permanece a mesma até finalizar a simulação. A duração total do sinal é de 3 horas, a transição normal/bradicardia tem duração de 30 minutos, a frequência de bradicardia é de 20 bpm e a sua duração é de aproximadamente 1 hora (até finalizar a simulação). O desvio padrão, para os dois casos, é de 8,93 bpm.

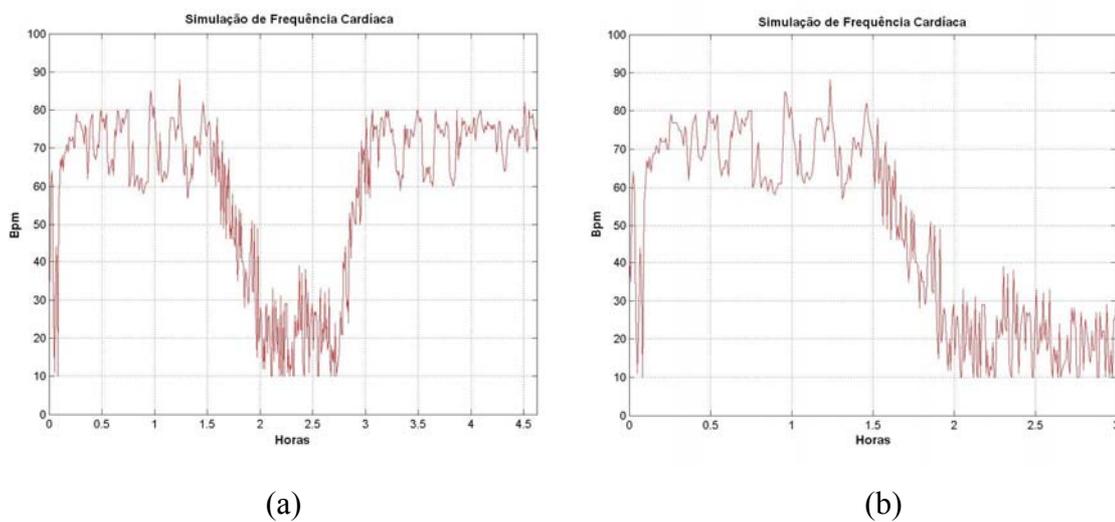


Figura 6-10 – Simulação de bradicardia utilizando características do sinal normal tomado como base – (a) modo ‘adicionar’, (b) modo ‘substituir/permanecer’.

Um exemplo de arquivo de saída (expondo apenas os 31 primeiros valores), assim como os gráficos correspondentes aos valores de inclinação, movimentação e frequência cardíaca contidos no arquivo, mostrados ao fim da simulação, estão representados na Figura 6-11. Percebe-se que a simulação ocorre a cada 30 segundos (assim como no sistema RFpat).

A Figura 6-12 mostra um exemplo da janela de comando do Matlab com a apresentação do resultado da geração dos dados referentes a cada sinal de saída, contendo informações do tipo: tempo total de simulação, valor médio, valor do desvio padrão, duração da anormalidade, etc. As opções de simulação exemplificadas na Figura 6-12 foram: inclinação com valor 1, criação de um novo perfil de agitação com média 7,4 e desvio padrão 8, e para a simulação de frequência cardíaca um sinal com duas taquicardias (com frequência de 140 bpm e desvio padrão 9), sendo que a duração para a primeira taquicardia foi de 20 minutos e para a segunda taquicardia foi de 40 minutos e 10 segundos.

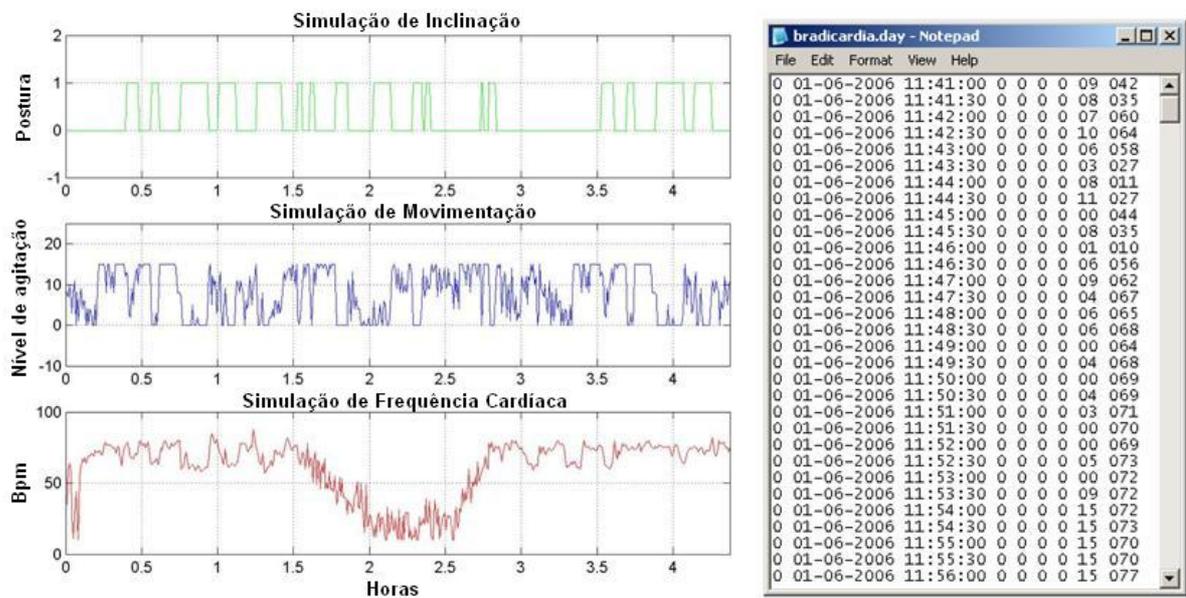


Figura 6-11 – Exemplo de arquivo de saída e gráficos correspondentes.

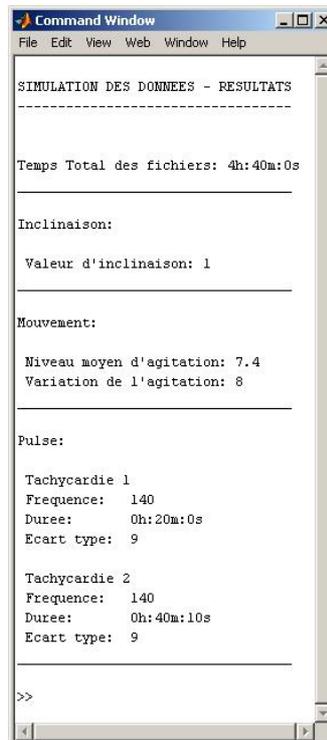


Figura 6-12 – Resultado da simulação – janela de comando do Matlab.

CAPÍTULO 7 – EXPERIMENTOS E RESULTADOS

7.1 Introdução

Nesse capítulo serão abordados os resultados obtidos por alguns dos experimentos realizados com o programa de fusão de dados biológicos.

Para a realização dos testes da rede (incluindo as etapas de treinamento e classificação) utilizaram-se dados reais (adquiridos por meio do uso dos sistemas RFpat e GARDIEN, por uma pessoa saudável), e dados simulados por meio do programa de simulação desenvolvido.

Foram adotados dois critérios para avaliação de resultados de Redes Bayesianas: sensibilidade (S) e de valor preditivo positivo (predição P) (OLIVEIRA, et al., 2008).

Fazendo a análise para uma situação de queda, por exemplo, o critério da sensibilidade S determina a porcentagem da fração de situações de queda corretamente detectadas:

$$S = \frac{VP}{FN + VP} \quad (\text{Sensibilidade}) \quad \text{Eq. 7-1}$$

VP (Verdadeiro Positivo) é o número de quedas corretamente detectado, e FN (Falso Negativo) é o número de quedas que não foi detectado pelo programa de fusão. FN + VP representa o número de situações de quedas descritas nos arquivos de teste (base de teste).

O critério do valor preditivo positivo corresponde à capacidade do sistema de detectar quedas verdadeiras.

$$P = \frac{VP}{FP + VP} \quad (\text{Predição}) \quad \text{Eq. 7-2}$$

FP (Falso Positivo) é o número de quedas detectado pelo programa de fusão que não estão de acordo com as situações de queda da base de teste, ou seja, o programa detectou quedas que, na verdade, não existem. FP + VP representa o número total de situações de quedas detectadas (tanto as detecções corretas como as detecções incorretas).

Os testes do programa de fusão foram realizados na modalidade *off-line*, pois a topologia do tipo *on-line* tem o mesmo funcionamento, já que utiliza um algoritmo baseado em Rede Bayesiana semelhante.

7.2 Resultados Obtidos

Foram definidos cinco cenários de teste, contendo situações de queda, bradicardia e taquicardia no ambiente domiciliar mostrado na Figura 4-4. Foram realizados alguns testes com dados incompletos, devido aos mesmos não utilizarem as saídas dos sensores B e 5.

- **Cenário 1:** uma pessoa entra na sala de testes (área 1), guarda as suas chaves, senta para ler um jornal em torno de 2 minutos, sem movimentos bruscos e depois se levanta e vai para a área 2, onde essa pessoa cai, permanecendo no chão por 2 minutos e tenta se levantar, mas não consegue.

A representação dos dados dos sensores (entrada da rede) e o resultado da fusão (nó CONTROLE) são mostrados na Figura 7-1. Para a representação do sensor B5 no gráfico, nível '0' indica que a pessoa está levantada, e nível '1' indica que a pessoa está deitada. O arquivo de texto gerado com as probabilidades dos nós é mostrado na Figura 7-2.

Para a Figura 7-2, assim como para as Figuras relacionadas aos arquivos de texto gerados nos próximos cenários, *Jour*, *Heure*, *Inclin*, *Mouv*, *Pouls*, *B5*, *Brady*, *Tachy*, *Chute* e *Control* representam respectivamente Data, Hora, Inclinação, Movimentação, Freq. Cardíaca, Sensor B5, Bradicardia, Taquicardia, Queda e Controle.

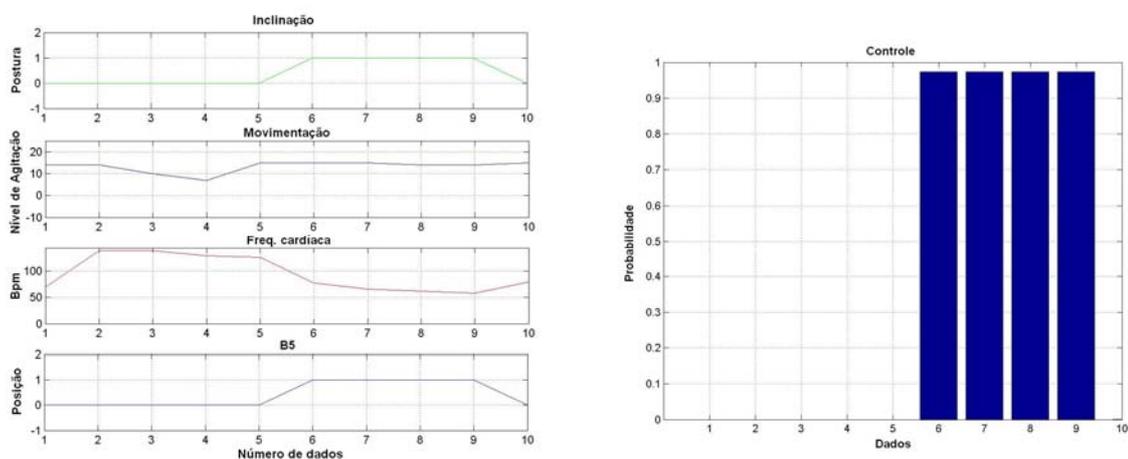


Figura 7-1 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 1.

Jour	Heure	Inclin	Mouv	Pouls	B5	Brady	Tachy	Chute	Control
20-12-2007	15:58:02	0	14	070	0	0 %	0 %	0 %	0 %
20-12-2007	15:58:32	0	14	140	0	0 %	0 %	0 %	0 %
20-12-2007	15:59:02	0	10	140	0	0 %	0 %	0 %	0 %
20-12-2007	15:59:32	0	07	130	0	0 %	0 %	0 %	0 %
20-12-2007	15:60:02	0	15	127	0	0 %	0 %	0 %	0 %
20-12-2007	16:00:32	1	15	078	1	0 %	0 %	97.37 %	97.37 %
20-12-2007	16:01:02	1	15	066	1	0 %	0 %	97.37 %	97.37 %
20-12-2007	16:01:32	1	14	062	1	0 %	0 %	97.37 %	97.37 %
20-12-2007	16:02:02	1	14	058	1	0 %	0 %	97.37 %	97.37 %
20-12-2007	16:02:32	0	15	080	0	0 %	0 %	0 %	0 %

---> CONTROL : Chute
 ---> **ALARME** : Chute
 ---> **ALARME** : Chute
 ---> **ALARME** : Chute

Figura 7-2 – Arquivo texto com resultado da rede - Cenário 1.

Nota-se que a classificação, no arquivo de texto, foi identificada como “queda” (“*chute*”), com 97,37% de probabilidade de ocorrência. A avaliação baseada na identificação de situações de queda do paciente forneceu como resultados:

- Sensibilidade: 100 %.
 - Predição: 75 %.
- Cenário 2: mesmas situações descritas para o Cenário 1, porém sem apresentar ocorrências de queda por parte da pessoa monitorada.

A Figura 7-3 apresenta os dados dos sensores e as probabilidades do nó CONTROLE. A Figura 7-4 mostra o arquivo de texto gerado pelo programa de fusão.

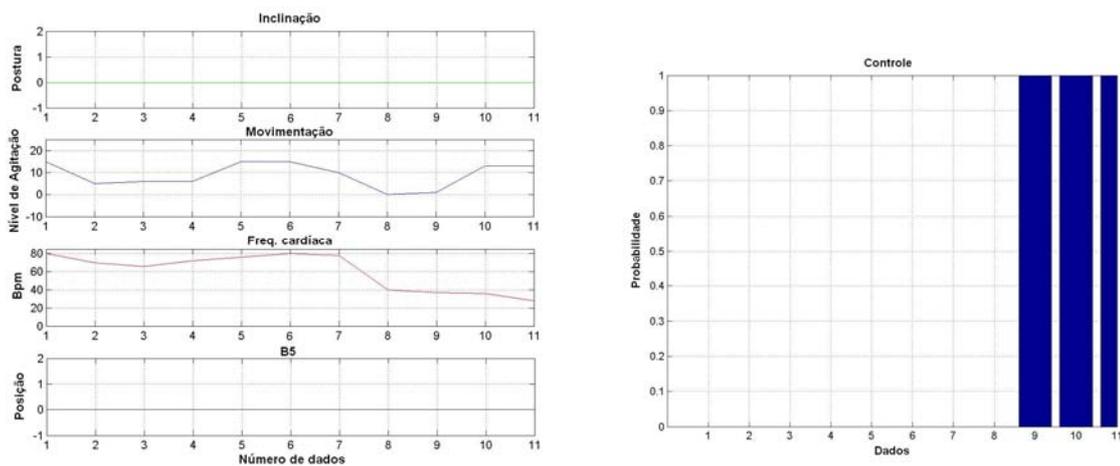


Figura 7-3 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 2.

Foram identificadas três situações de bradicardia que não estavam previstas. Uma hipótese pode ser mau contato do sensor de pulsação, devido a algum movimento brusco por parte da pessoa que utiliza o sensor, ou devido ao começo do processo de retirada desse sensor, já

que esses eventos se apresentam no fim da medição. Quando não há um bom contato eletrodo/pele, a pulsação se apresenta em níveis bem abaixo do normal.

Jour	Heure	Inclin	Mouv	Pouls	B5	Brady	Tachy	Chute	Control
20-12-2007	16:10:03	0	15	080	0	0 %	0 %	0 %	0 %
20-12-2007	16:10:33	0	05	070	0	0 %	0 %	0 %	0 %
20-12-2007	16:11:03	0	06	066	0	0 %	0 %	0 %	0 %
20-12-2007	16:11:33	0	06	072	0	0 %	0 %	0 %	0 %
20-12-2007	16:12:03	0	15	076	0	0 %	0 %	0 %	0 %
20-12-2007	16:12:33	0	15	080	0	0 %	0 %	0 %	0 %
20-12-2007	16:13:03	0	10	078	0	0 %	0 %	0 %	0 %
20-12-2007	16:13:33	0	00	040	0	0 %	0 %	0 %	0 %
20-12-2007	16:14:03	0	01	037	0	100 %	0 %	0 %	100 %
20-12-2007	16:14:33	0	13	036	0	100 %	0 %	0 %	100 %
20-12-2007	16:14:33	0	13	028	0	100 %	0 %	0 %	100 %

--> CONTROL : Bradycardie
 --> **ALARME** : Bradycardie
 --> **ALARME** : Bradycardie

Figura 7-4 – Arquivo texto com resultado da rede - Cenário 2.

Como não existem ocorrências de quedas nesse cenário, o cálculo de sensibilidade e preditivo positivo foi realizado para a classe normal (neste caso, as medidas levam o nome de especificidade e preditivo negativo). Os resultados obtidos foram:

- Especificidade: 100 %.
- Predição: 100 %.

Levando-se em conta que a análise foi realizada tomando-se como referência situações de queda, pode-se concluir que os resultados estão de acordo, já que o sistema acertou em detectar apenas situações normais com o paciente, ou seja, não houve quedas.

- Cenário 3: uma pessoa permanece sentada, trabalhando em um computador na área 1. Ela sai por 6 minutos, retornando em seguida por 44 minutos. Ao final do cenário, ocorre uma queda.

A Figura 7-5 apresenta os dados dos sensores e as probabilidades do nó CONTROLE. A Figura 7-6 mostra o arquivo de texto gerado pelo programa de fusão, com algumas medições ocultadas, devido ao texto do arquivo ser extenso.

A classificação das medições finais indica uma situação de queda de 1 minuto de duração. A avaliação baseada na identificação de situações de queda do paciente forneceu como resultados:

- Sensibilidade: 100 %.
- Predição: 100 %.

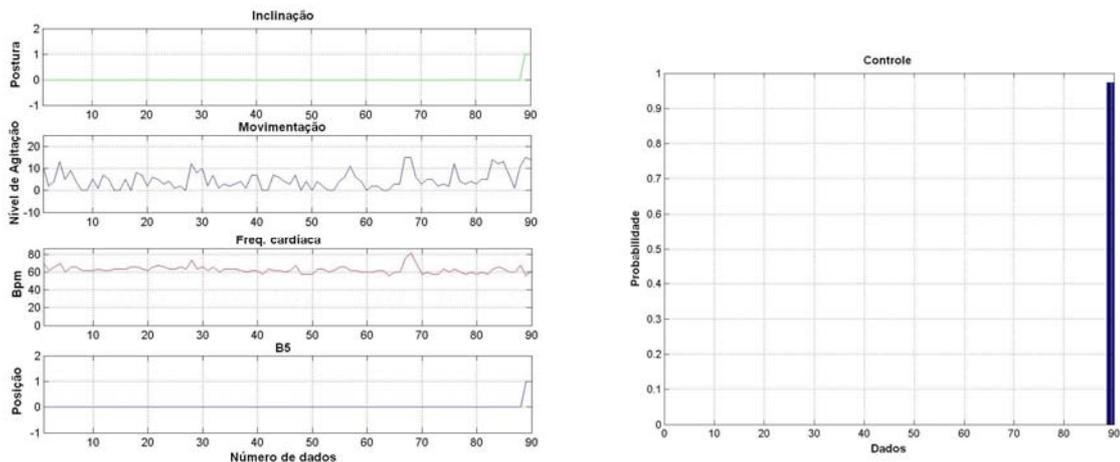


Figura 7-5 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 3.

Jour	Heure	Inclín	Mouv	Pouls	B5	Brady	Tachy	Chute	Control
09-01-2008	12:05:16	0	11	070	0	0 %	0 %	0 %	0 %
09-01-2008	12:05:46	0	02	062	0	0 %	0 %	0 %	0 %
09-01-2008	12:06:16	0	04	066	0	0 %	0 %	0 %	0 %
09-01-2008	12:06:46	0	13	070	0	0 %	0 %	0 %	0 %
.
.
09-01-2008	12:57:16	0	01	060	0	0 %	0 %	0 %	0 %
09-01-2008	12:57:46	0	11	068	0	0 %	0 %	0 %	0 %
09-01-2008	12:58:16	1	15	056	1	0 %	0 %	97.37 %	97.37 %
09-01-2008	12:58:46	1	14	062	1	0 %	0 %	97.37 %	97.37 %

--> CONTROL : chute
--> **ALARME** : chute

Figura 7-6 – Arquivo texto (resumido) com resultado da rede - Cenário 3.

- **Cenário 4:** Utilizou-se um sinal simulado de bradicardia. A medição dura 12 horas, e é composta por sinais de inclinação com valor 0 (posição vertical) e de movimentação com valores padrão (7,5 para nível médio e 9 para desvio padrão). Os sensores B e 5 não foram utilizados neste teste devido a simulação dos mesmos não ser disponível.

A Figura 7-7 apresenta os dados dos sensores e as probabilidades do nó CONTROLE. A Figura 7-8 mostra o arquivo de texto gerado pelo programa de fusão (também resumido, devido ao arquivo ser extenso).

Nota-se a detecção de uma bradicardia isolada às 14h17m30s. Essa detecção não é correta. O motivo principal deste erro se deve ao fato da Rede Bayesiana proposta não levar em conta o estado anterior do paciente. A informação anterior é uma informação adicional, para aumentar a precisão do sistema. Uma maneira de se resolver isso seria com o emprego de uma Rede Bayesiana Dinâmica (mais detalhes no Capítulo 8).

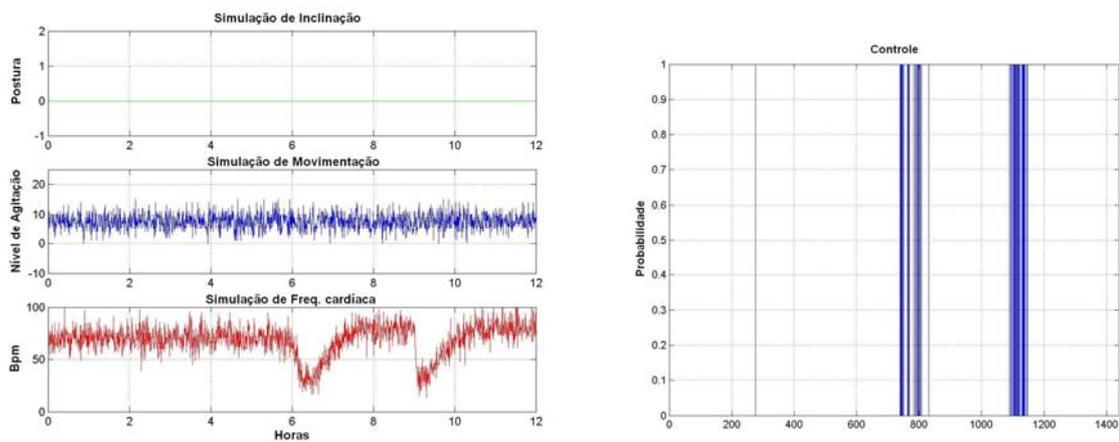


Figura 7-7 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 4.

Jour	Heure	Inclin	Mouv	Pouls	Brady	Tachy	Chute	Control
01-05-2007	12:00:00	0	10	061	0 %	0 %	0 %	0 %
01-05-2007	12:00:30	0	09	071	0 %	0 %	0 %	0 %
01-05-2007	12:01:00	0	11	070	0 %	0 %	0 %	0 %
.
.
01-05-2007	14:17:00	0	08	076	0 %	0 %	0 %	0 %
01-05-2007	14:17:30	0	10	039	100 %	0 %	0 %	100 %
01-05-2007	14:18:00	0	05	065	0 %	0 %	0 %	0 %
.
.
01-05-2007	18:11:30	0	04	050	0 %	0 %	0 %	0 %
01-05-2007	18:12:00	0	10	051	0 %	0 %	0 %	0 %
01-05-2007	18:12:30	0	08	029	100 %	0 %	0 %	100 %
01-05-2007	18:13:00	0	08	029	100 %	0 %	0 %	100 %
01-05-2007	18:13:30	0	07	035	100 %	0 %	0 %	100 %
01-05-2007	18:14:00	0	02	046	0 %	0 %	0 %	0 %
01-05-2007	18:14:30	0	06	019	100 %	0 %	0 %	100 %
01-05-2007	18:15:00	0	05	025	100 %	0 %	0 %	100 %
01-05-2007	18:15:30	0	09	033	100 %	0 %	0 %	100 %
01-05-2007	18:16:00	0	09	031	100 %	0 %	0 %	100 %
.
.

Figura 7-8 – Arquivo texto (resumido) com resultado da rede - Cenário 4.

A avaliação baseada na identificação de situações de bradicardia do paciente forneceu como resultados:

- Sensibilidade: 100 %.
- Predição: 100 %.

• **Cenário 5:** Empregou-se um sinal simulado de taquicardia. A medição dura 12 horas, e é composta por sinais de inclinação com valor 0 (posição vertical) e de movimentação com valores padrão (7,5 para nível médio e 9 para desvio padrão). Os sensores B e 5 não foram utilizados neste teste devido a simulação dos mesmos não ser disponível.

A Figura 7-9 apresenta os dados dos sensores e as probabilidades do nó CONTROLE. A Figura 7-10 mostra o arquivo de texto gerado pelo programa de fusão (também resumido, devido ao arquivo ser extenso).

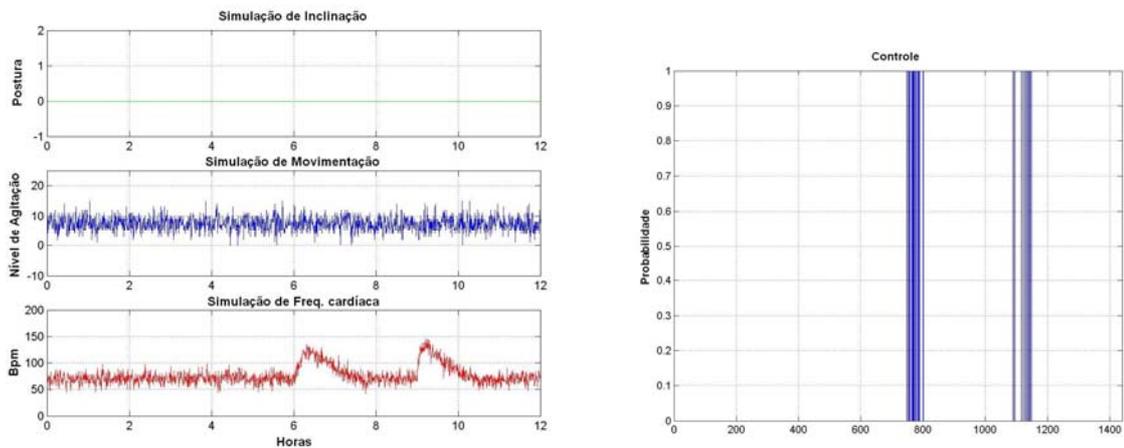


Figura 7-9 – Dados de entrada da rede e resultado do nó CONTROLE - Cenário 5.

Jour	Heure	Inclin	Mouv	Pouls	Brady	Tachy	Chute	Control
01-05-2007	12:00:00	0	08	082	0 %	0 %	0 %	0 %
01-05-2007	12:00:30	0	10	067	0 %	0 %	0 %	0 %
.
01-05-2007	18:12:30	0	04	117	0 %	0 %	0 %	0 %
01-05-2007	18:13:00	0	11	120	0 %	0 %	0 %	0 %
01-05-2007	18:13:30	0	09	124	0 %	100 %	0 %	100 % --> CONTROL : Tachycardie
01-05-2007	18:14:00	0	06	125	0 %	100 %	0 %	100 % --> **ALARME** : Tachycardie
01-05-2007	18:14:30	0	13	118	0 %	0 %	0 %	0 %
01-05-2007	18:15:00	0	05	116	0 %	0 %	0 %	0 %
01-05-2007	18:15:30	0	12	120	0 %	0 %	0 %	0 %
01-05-2007	18:16:00	0	08	114	0 %	0 %	0 %	0 %
01-05-2007	18:16:30	0	10	113	0 %	0 %	0 %	0 %
01-05-2007	18:17:00	0	10	121	0 %	100 %	0 %	100 % --> CONTROL : Tachycardie
01-05-2007	18:17:30	0	07	107	0 %	0 %	0 %	0 %
01-05-2007	18:18:00	0	09	110	0 %	0 %	0 %	0 %
01-05-2007	18:18:30	0	08	135	0 %	100 %	0 %	100 % --> CONTROL : Tachycardie
01-05-2007	18:19:00	0	05	129	0 %	100 %	0 %	100 % --> **ALARME** : Tachycardie
01-05-2007	18:19:30	0	04	128	0 %	100 %	0 %	100 % --> **ALARME** : Tachycardie
01-05-2007	18:20:00	0	08	125	0 %	100 %	0 %	100 % --> **ALARME** : Tachycardie
01-05-2007	18:20:30	0	06	132	0 %	100 %	0 %	100 % --> **ALARME** : Tachycardie
01-05-2007	18:21:00	0	09	105	0 %	0 %	0 %	0 %
.
.
.

Figura 7-10 – Arquivo texto (resumido) com resultado da rede - Cenário 5.

A avaliação baseada na identificação de situações de taquicardia do paciente forneceu como resultados.

- Sensibilidade: 100 %.
- Predição: 100 %.

7.3 Discussão

De acordo com os resultados obtidos, nota-se que a Rede Bayesiana está realizando a classificação das situações de queda, bradicardia e taquicardia, entretanto, fazem-se necessários ajustes nas tabelas a priori e também a obtenção de um número maior de dados confiáveis

de treinamento, para que a rede possa realizar tanto a classificação com maior precisão quanto a fusão de dados mais complexos.

Após a adoção de uma etapa prévia de treinamento para estimação das probabilidades dos nós, o desempenho do sistema melhorou, mostrando que a estratégia a ser adotada para a estimação da contribuição de cada variável requer algum conhecimento sobre as outras variáveis (o que é conseguido pela etapa de treinamento). Vale ressaltar que os primeiros testes, que não são mostrados aqui, foram realizados sem estágio de treinamento. A estimação dos parâmetros de cada variável, realizada apenas por meio de cada variável isoladamente, fornece resultados com qualidade inferior.

Uma estratégia a ser utilizada futuramente para a avaliação da rede e utilizar médicos para analisar os resultados fornecidos pelo sistema. Dos testes da rede podem surgir mudanças no algoritmo, a partir de opiniões de especialistas médicos.

As Redes Bayesianas constituem uma poderosa ferramenta para apoio à decisão e fusão de dados, abrangendo aí os sistemas dedicados ao diagnóstico médico. Sua flexibilidade possibilita utilizar um mesmo sistema com diferentes pacientes (sendo necessária uma adaptação inicial para melhorar o desempenho, respeitando a especificidade de cada um). No entanto, para isso é necessária uma quantidade adequada de dados para a etapa de treinamento.

CAPÍTULO 8 – CONCLUSÕES E PERSPECTIVAS

Nesta Dissertação de Mestrado foi apresentada uma proposta de utilização de Redes Bayesianas para a fusão de sinais vitais e de movimento de pacientes, com geração de alarmes, no contexto de monitoramento de pacientes em domicílio.

Foi desenvolvido um programa de fusão de dados reunindo características essenciais para a realização do telemonitoramento de um paciente em domicílio, o qual possui as seguintes características:

- Facilidade de uso por uma pessoa não especialista;
- Desempenho satisfatório para a função desejada;
- Facilidade de manutenção;
- Reuso de código (modularização).

Do mesmo modo, o programa de simulação criado supriu a necessidade de formação de um banco de dados para teste da rede, além de ser uma ferramenta disponibilizada para geração de dados de taquicardia e bradicardia. O programa de simulação desenvolvido possui as seguintes características:

- Facilidade de uso;
- Representação adequada para uso em testes de algoritmos (que requer o máximo de situações para tratamento pelo sistema de decisão);
- Reuso de código (modularização);
- Desempenho satisfatório para a função desejada;
- Facilidade de manutenção.

Com este trabalho foi possível modelar um sistema de decisão do estado do paciente em domicílio, através do uso de Redes Bayesianas, permitindo determinar o estado do paciente a cada 30 segundos em função dos parâmetros: inclinação, movimentação e frequência cardíaca. As quatro classes definidas foram: taquicardia, bradicardia, queda e normal. Foi alcançado um desempenho superior a 90%, provando que o sistema proposto baseado nas Redes Bayesi-

anas é uma boa escolha para o emprego em aplicações no campo do telemonitoramento. Os resultados obtidos mostram a capacidade de uma Rede Bayesiana em tomar decisões mesmo se as informações de algum evento estejam incompletas e/ou se não há dados reais para o desenvolvimento do algoritmo da rede.

A Rede Bayesiana exibe uma representação gráfica compreensiva, trata das relações de incertezas através de sua representação probabilística, possibilita trabalhar com dados incompletos por meio de seu algoritmo de inferência e permite redundância de sensores.

A saída do sistema com o valor de probabilidade possibilita uma análise mais completa da situação do paciente, pois é levada em consideração a incerteza da resposta. Como exemplo, podemos dizer que um sinal de alarme com probabilidade de 80% tem maior chance de realmente representar uma situação de alerta que um sinal de alarme com probabilidade de 60%. Assim, a pessoa responsável pelo atendimento do paciente poderá fazer um melhor julgamento, com base no quadro médico específico desse paciente.

Contudo, pode-se melhorar o desempenho da classificação com a adição de outros parâmetros de entrada da rede, em particular a atividade do paciente, a variável tempo, o controle do nível da bateria do dispositivo portátil RFpat e o sensor do tipo pinça (para medir a pulsação).

A atividade do paciente seria uma tarefa mais complexa, pois compreenderia determinar a atividade atual do paciente de acordo com a sua condição física, batimentos cardíacos, postura, etc., sendo essa atividade um item importante para se determinar os outros parâmetros. Durante o monitoramento de uma pessoa em seu domicílio, realizado através de seus deslocamentos e posturas, das variações do nível de atividade e de sua frequência cardíaca, é sabido que:

- Os deslocamentos dentro do domicílio seguem o ritmo das atividades cotidianas;
- As posturas sucessivas dependem da atividade realizada e, por consequência, do recinto ocupado no domicílio e do momento do dia;
- O nível de atividade observado é igualmente função da atividade realizada e, portanto, dos deslocamentos observados e das posturas;
- As principais variações da frequência cardíaca são relacionadas à postura e ao nível de atividade.

A informação temporal na entrada da rede consiste no uso de informações de instantes anteriores, compreendendo também o uso das saídas dos nós da classificação precedente. Isso poderia ser conseguido com o uso de uma Rede Bayesiana Dinâmica - RBD (MURPHY,

2002), que é uma extensão de uma Rede Bayesiana tradicional, aumentando assim o potencial de classificação desta última. Para utilizar a variável tempo as saídas da rede seriam comparadas às saídas precedentes, podendo assim ser confirmada uma situação de alerta ou verificar o deslocamento e a posição do paciente no ambiente. Ou seja, além de modelar a probabilidade $P(x(t))$, em um dado instante 't' (procedimento de uma Rede Bayesiana padrão), a RBD também modela a probabilidade $P(x(t) | x(t-1))$ (probabilidade no tempo 't' dada a probabilidade no tempo 't-1'), entre instantes de tempo distintos. Isto vale para problemas onde há uma evolução no tempo do processo analisado.

O monitoramento do nível da bateria do terminal RFpat seria interessante para informar ao usuário e ao médico responsável (via servidor remoto), que a carga da bateria está chegando a um nível baixo. Quando isso ocorre, devido à alteração do nível de tensão de alimentação dos sensores, os resultados dos mesmos podem ser afetados, fornecendo um valor incorreto.

Já o sensor do tipo pinça, que serve para medir a frequência cardíaca, poderia ser trocado por outro de maior confiabilidade (do tipo adesivo, por exemplo), pois em um movimento brusco esse sensor pode sair do lugar e, sem o paciente notar, pode interferir na medição (como aconteceu no experimento do Cenário 2). Além disso, esse sensor pode ser incômodo para o uso constante, ao longo do dia.

Em relação à imprecisão e incerteza referente às variáveis discretas de entrada da rede (saídas dos sensores), pode-se melhorar na decisão por meio do uso de Lógica Fuzzy. A possibilidade de representação dos dados utilizando conjuntos Fuzzy permite maior flexibilidade na definição dos limiares de decisão (tomada de decisão).

Apesar de se utilizar poucas variáveis discretas neste sistema (o que leva a pensar em abordagens alternativas às redes bayesianas) deve-se levar em conta o fato de que, futuramente, serão incluídas novas variáveis obtidas tanto pelo sistema GARDIEN quanto por novos sensores no sistema RFpat.

Para o aperfeiçoamento do sistema, pode ser realizada a fusão dos dados do dispositivo RFpat e do sistema GARDIEN completo, com todos os sensores piroelétricos do ambiente. Dessa forma, juntamente com o monitoramento dos deslocamentos e níveis de atividades do paciente, a precisão da classificação seria incrementada significativamente. A identificação da localização do paciente no ambiente tem influência direta nas probabilidades iniciais da rede. Por exemplo, dependendo do horário de ocorrência de uma situação de queda, caso essa ocorra no quarto, pode ser que o paciente tenha se deitado na cama. Não se pode dizer o mesmo caso essa queda tenha ocorrido no banheiro. Além disso, caso aparente ter ocorrido uma situ-

ação de queda, deve-se verificar se o indicador dos sensores piroelétricos está inerte, ou seja, indicando que o paciente não está se movimentando.

Vale ressaltar que já estão em andamento os testes com um acelerômetro tridimensional instalado no terminal RFpat e o uso de microfones para a detecção de sons do ambiente. O acelerômetro será responsável por informar com precisão se houve uma queda do paciente (por meio da medição do impacto causado pela queda), dotando a Rede Bayesiana de maior exatidão em suas classificações. A análise de sons do ambiente possibilita identificar e extrair sons que possam ser úteis, como por exemplo, detectar um objeto de vidro se quebrando ou um grito (MEDJAHED, et al., 2008).

As pesquisas e soluções desenvolvidas nessa Dissertação poderão ser aproveitadas no projeto Europeu ICADORS - *Integrated Cognitive Assistive Domotic & Robotic Smart Systems*, do qual o INT participa como um dos parceiros. Esse projeto visa prover assistência a idosos debilitados (pré-Alzheimer) em domicílio, por meio do monitoramento de seu estado de saúde. Para isso, serão desenvolvidos dois sistemas complementares: um robô interativo para prover assistência, memorização de eventos, entre outras funções, e um ambiente composto por uma rede de sensores que possam adquirir informações vitais do paciente e gerar alarmes.

Finalmente, é importante destacar que os trabalhos desenvolvidos nesta Dissertação de Mestrado resultaram nas publicações dos artigos (MARTINS, BOUDY, et al., 2007) e (MEDJAHED, ISTRATE, et al., 2008).

REFERÊNCIAS BIBLIOGRÁFICAS

AFONSO, L. S.; SANTOS, J. F. B.; et al. Frequência cardíaca máxima em esteira ergométrica em diferentes horários. *Revista Brasileira de Medicina do Esporte*, v. 12, n. 6. Niterói. Novembro/Dezembro/2006.

AGÊNCIA CT – Ministério da Ciência e Tecnologia. Telemedicina é tema de debate. Novembro/2006. Disponível em <<http://agenciact.mct.gov.br/index.php/content/view/42146.html>>. Acesso em: ago/2008.

ALZHEIMER MED. Informação e Solidariedade. Doenças Associadas. Disponível em: <http://www.alzheimermed.com.br/m3.asp?cod_pagina=935>. Acesso em: dez/2007.

BALDINGER, J. L.; BOUDY, J.; DORIZZI, B.; LEVREY, J.-P.; ANDREÃO, R.; PERPÈRE, C.; DELAVAUULT, F.; ROCARIES, F.; DIETRICH, C.; LACOMBE, A. Tele-surveillance System for Patient at Home: the MEDIVILLE system. *Congrès ICCHP 2004*, Paris. Julho/2004.

BALLONE, G. J. Transtorno de Personalidade Paranóide. *PsiquWeb - psiquiatria geral*. Disponível em: <http://virtualpsy.locaweb.com.br/dsm_janela.php?cod=166>. Acesso em: dez/2007.

BELFEKI, I. Contribution à la Fusion Multimodale appliquée à la Combinaison de deux Systèmes de Télévigilance du Patient à Domicile – RFpat et GARDIEN. *Rapport de Master 2 – TRIED*. INT, France. Setembro/2006.

BIBLIOMED, INC. Arritmia Cardíaca. Setembro 2005. Disponível em: <<http://boasaude.uol.com.br/lib/ShowDoc.cfm?LibDocID=4825&ReturnCatID=357>>. Acesso em: dez/2007.

BOITE, R.; BOURLARD, H.; DUTOIT, T.; HANCQ, J.; LEICH, H. Traitement de la parole. *Collection Électricité. Presses Polytechniques et Universitaires Romandes*. Lausanne, France, 2000.

BONNER, S. G. Assisted Interactive Dwelling House. *3rd TIDE Congress - Technology for Inclusive Design and Equality*. Helsinki, Finland. 1998.

BOUDY, J. Projet TELEPAT – Télésurveillance de Patients à domicile. *Rapport Scientifique Final – Version Finale 2.0*. GET/INT, França. Novembro/2006.

BOUDY, J.; BALDINGER, J. L. Télésurveillance de paramètres vitaux du Patient à Domicile : Télévigilance. *HIT 2007*. Groupe des Ecoles des Télécommunications / INT dept. EPH. Maio/2007.

CASTELLI, E.; ISTRATE, D. Everyday Life Sounds and Speech Analysis for a Medical Telemonitoring System. *7th European Conference on Speech Communication and Technology*. Dinamarca. 2001.

CASTILLO, E.; GUTIÉRREZ, J. M.; HADI, A. S. Sistemas Expertos y Modelos de Redes Probabilísticas. *Monografias da Academia Espanhola de Engenharia*. Madrid. 1998.

CHARNIAK, E. Bayesian Networks without Tears. *American Association for Artificial Intelligence*. 1991.

CHEN, S.-C.; CHEN, H.-W.; LEE, A.; CHAO, K.-H.; HUANG, Y.-C.; LAI, F. E-Vanguard for Emergency - a wireless system for rescue and healthcare. *5th International Workshop on Enterprise Networking and Computing in Healthcare Industry*. Healthcom 2003, p.29 – 35. Junho/2003.

DODIER, R. H. Unified Prediction and Diagnosis in Engineering Systems by Means of Distributed Belief Networks. *Tese de doutorado*. Department of Civil, Environmental and Architectural Engineering. University of Colorado. 1999.

DONADIO, M. How to Generate White Gaussian Noise. *dspGuru – Digital Signal Processing Central*. Disponível em: <<http://www.dspguru.com/howto/tech/wgn.htm>>. Acesso em: mar/2007.

DUCHÈNE, F. Fusion de données multicateurs pour un système de télésurveillance médicale de personne à domicile. *Thèse de doctorat*. INPG Grenoble, França. Outubro/2004.

DUDA, R. O.; HART, P. E.; STORK, D. G. Pattern Classification. Second Edition. *Wiley-Interscience Publication*. 2001.

ELGER, G.; FURUGREN, B. "SmartBo" – An ICT and computer-based demonstration home for disabled people. *3rd TIDE Congress - Technology for Inclusive Design and Equality*. Helsinki, Finland. 1998.

EYE, G. V. Arritmias. *ABC da Saúde e Prevenção Ltda*. 2007. Disponível em: <<http://www.abcdasaude.com.br/artigo.php?32>>. Acesso em: dez/2007.

FÉRON, O.; MOHAMMAD-DJAFARI, A. A Hidden Markov model for Bayesian data fusion of multivariate signals. *Fifth Int. Triennial Calcutta Symposium on Probability and Statistics*. Calcutta University, India. 2003.

FINKELSTEIN, J.; O'CONNOR, G.; FRIEDMAN, R. H. Development and implementation of the Home Asthma Telemonitoring (HAT) System to Facilitate Asthma Self-Care. *10th World Congress on Medical Informatics (MEDINFO)*, p. 810–814. Londres, UK. 2001.

FONSECA, J. M. M. R. da. Indução de Árvores de Decisão. *HistClass – Proposta de um algoritmo não paramétrico*. Universidade Nova de Lisboa. Faculdade de Ciências e Tecnologia. Lisboa. 1994.

FREEMAN, J. A.; SKAPURA, D. M. Neural Networks – Algorithms, Applications, and Programming Techniques. *Addison-Wesley*. 1991

HALL, D. L.; LLINAS, J. Handbook of Multisensor Data Fusion. The Electrical Engineering and Applied Signal Processing Series. *CRC Press*. 2001.

HARRIS, C. ; GAN, Q. Multisensor Data Fusion Using Kalman Filters Based on Neuro-Fuzzy Linearisation. Multisensor Fusion (NATO ASI Series Book), pp. 435-456. *Kluwer Academic Publishers*. 2002.

HATON, J.-P.; CERISARA, C.; FOHR, D.; LAPRIE, Y.; SMAÏLI, K. Reconnaissance automatique de la parole – Du signal à son interprétation. *Dunod*. Paris. 2006.

HIDEAKI, A. S. Construção de uma rede bayesiana aplicada ao diagnóstico de doenças. *Dissertação de Mestrado*. Universidade de São Paulo, Brasil. 2005.

HUANG, C.; DARWICHE, A. Inference in Belief Networks: A Procedural Guide. *International Journal of Approximate Reasoning*, 11:1-158. *Elsevier Science Inc*. New York. 1994.

IBGE - Instituto Brasileiro de Geografia e Estatística. Indicadores Sociodemográficos. Prospectivos para o Brasil 1991-2030. *Arbeit Editora e Comunicação*. Rio de Janeiro. Outubro/2006.

IBGE - Instituto Brasileiro de Geografia e Estatística. Tábuas Completas de Mortalidade. Disponível em:

<http://www.ibge.gov.br/home/estatistica/populacao/projecao_da_populacao/piramide/piramide.shtm>. Acesso em: jan/2008.

ISTRATE, D. M. Détection et reconnaissance des sons pour la surveillance médicale. *Tèse de doutorado*. Institut National Polytechnique de Grenoble. França. Dezembro/2003.

ISTRATE, D.; VACHER, M.; SERIGNAT, J. F. Generic Implementation of a Distress Sound Extraction System for Elder Care. *8th IEEE EMBS Annual International Conference*, p. 3309-3312. New York City, USA. Agosto/2006.

JENSEN, F. V. An Introduction to Bayesian Network. *UCL Press*. Aalborg University, Denmark. 2000.

JENSEN, F. V. Bayesian Network and Decision Graphs. *Springer*. New York. 2001.

KADI, A. BELFEKI, I. Utilisation d'un MLP pour la télésurveillance du patient à domicile. *Projet long*. Université de Versailles. St. Quentin En-Yvelines. Master 2 TRIED. 2004/2005.

KESSEL, R. T. The Dynamics of Information Fusion: Synthetic Versus Misassociation. *Reprint Series. NURC-PR-2006-003*. NATO Undersea Research Centre. La Spezia, Itália. Agosto/2006.

LIND, L.; SUNDVALL, E.; ÅHLFELDT, H. Experiences from Development of Home Health Care Applications based on Emerging Java Technology. 10th World Congress on Medical Informatics (MEDINFO), p. 830–834. London, UK. 2001.

LUNA, J. E. O. Algoritmos EM para Aprendizagem de Redes Bayesianas a partir de Dados Incompletos. *Dissertação de Mestrado*. Departamento de Computação e Estatística. Universidade Federal de Mato Grosso do Sul, Brasil. Julho/2004.

MARTINS, V. R.; BOUDY, J.; ANDREÃO, R. V.; BASTOS-FILHO, T. F. Uma Abordagem de Fusão de Sinais Vitais Baseada em Redes Bayesianas. CLAIB 2007, IFMBE Proceedings 18. *Springer-Verlag*, pág. 178-182. 2007.

MEDJAHED, H.; ISTRATE, D.; BOUDY, J.; BALDINGER, J-L; DORIZZI, B.; BELFEKI, I.; MARTINS, V.; STEENKESTE, F.; ANDREÃO, R. A Multimodal Platform for Database Recording and Elderly People Monitoring. *Biosignals 2008*. Funchal, Portugal. Janeiro/2008.

MÜLLER, S. M. T. Adaptação dos Modelos de Markov para um Sistema de Segmentação e Classificação de Sinais de Eletrocardiograma. *Dissertação de mestrado*. PPGEE - Universidade Federal do Espírito Santo, Vitoria. 2006.

MURPHY, K. P. A Brief Introduction to Graphical Models and Bayesian Network. 1998. Disponível em: <<http://bnt.sourceforge.net/>>. Acesso em: fev/2007.

MURPHY, K. An introduction to Bayesian Networks and the Bayes Net Toolbox for Matlab. *MIT AI Lab*. Maio/2003.

MURPHY, K. P. Dynamic Bayesian Networks: Representation, Inference and Learning. *Tese de Doutorado*. Universidade da Califórnia, Berkeley. 2002.

NAÏM, P.; WUILLEMIN, P.-H.; LERAY, P.; POURRET, O.; BECKER, A. Réseaux bayésiens. *Eyrolles*. Paris. 2004.

NAJAFI, B.; AMINIAN, K.; et al. Ambulatory System for Human Motion Analysis Using a Kinematic Sensor: Monitoring of Daily Physical Activity in the Elderly. *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, p. 711-723. Junho/2003.

NEAPOLITAN, R. E. Learning Bayesian Network. Pearson Prentice Hall. *Prentice Hall Series In Artificial Intelligence*. 2004.

NORSYS Software Corp. *Netica Tutorial*. 2007. Disponível em: <<http://www.norsys.com/netica.html>>. Acesso em: fev/2007.

NOURY, N.; HERVÉ, T.; RIALLE, V.; VIRONE, G.; MERCIER, E. Monitoring behavior in home using smart fall sensor and position sensors. *1st Annual International IEEE-EMBS on Microtechnologies in Medicine and Biology*, p. 607-610. Lyon, França. Outubro/2000.

OGAWA, M.; TOGAWA, T. Attempts at Monitoring Health status in the Home. *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine & Biology*, p. 552-556. Lyon, França. Outubro/2000.

OLIVEIRA, L. S. C.; ANDREÃO, R. V.; SARCINELLI-FILHO, M. Classification of premature ventricular beat using bayesian networks. *Healthinf 2008*. Funchal, Portugal. Janeiro/2008.

OLIVIER, F. De l'identification de structure de réseaux bayésiens à la reconnaissance de formes à partir d'informations complètes ou incomplètes. *Thèse de doctorat*. Institut National de Sciences Appliquées de Rouen. França. 2006.

OLSON, K. R. Bradicardia e Bloqueio Atrioventricular. *IPCS INTOX Databank*. University of Califórnia, San Francisco. 1998. Disponível em: <http://www.intox.org/databank/documents/treat/treatp/trt10_p.htm>. Acesso em: dez/2007.

PAPASTAVROU, J. D. Fusion de données. An Efficient Approximation Scheme for Distributed Multiple Hypothesis Testing. *Traitement du Signal* - Volume 11 – nº 6. Maio/1994.

PASKIN, M. A. A Short Course on Graphical Models. The Junction Tree Algorithms. 2003. Disponível em: <<http://ai.stanford.edu/~paskin/gm-short-course/>>. Acesso em: jan/2008.

PESSETE, R. S.; VIEIRA, K. M. M. Redes Bayesianas no Diagnóstico Médico. *Ciências da Computação quinta fase, INE*, Universidade Federal de Santa Catarina – UFSC, Brasil. 2002.

PINTO, C. M. S. Algoritmos Incrementais para Aprendizagem Bayesiana. *Dissertação de Mestrado*. Departamento de Ciência de Computadores. Faculdade de Economia da Universidade do Porto. Portugal. 2005.

RODRÍGUEZ, M.J.; ARREDONDO, M.T.; POZO, F. del; GÓMEZ, E.J.; MARTÍNEZ, A; DOPICO, A. A Home Telecare Management System. *Computers in Cardiology*, p. 433–436. Setembro/1994.

RUTE – Rede Universitária de Medicina. O que é telemedicina. Disponível em: <<http://rute.rnp.br/sobre/telemedicina/>>. Acesso em: ago/2008.

SAMPAIO, R. M. Redes Bayesianas. *Material referente à Dissertação de Mestrado*. Disponível em: <<http://www.dcc.ufla.br/~rudini/>>. Acesso em: dez/2007.

SILVA, E. M. M.; GALLO, A. K. G.; SANTOS, D. M.; BARÃO, V. A. R.; FREITAS, A. C. Enfermidades do Paciente Idoso. *Pesquisa Brasileira em Odontopediatria e Clínica Integrada*. Vol. 7, número 001, p. 83-88. Universidade Federal da Paraíba, João Pessoa. Janeiro-Abril/2007.

SILVA, G. DA S. Estudo de Técnicas e Utilização de Mineração de Dados em uma Base de Dados da Saúde Pública. *Trabalho de Conclusão*. Universidade Luterana do Brasil. Rio Grande do Sul. Junho/2003.

SINGLIAR, T. Pearl's algorithm. Message passing algorithm for exact inference in polytree BBNs. *Course CS3750*. University of Pittsburgh. Setembro/2003. Disponível em: <<http://www.cs.pitt.edu/~tomas/cs3750/>>. Acesso em: fev/2008.

SOBIECKI, B.; BOURKEB, Y. Application d'un Réseau de Neurones à la Detection de Troubles de Rythme Cardiaque. *Projet de Fin d'Etudes*. Telecom INT, França. 2006.

SPITZ, C. Brasil vê crescimento de idosos “seniores”, afirma IBGE. *Folha Online*. Setembro/2007. Disponível em: <<http://www1.folha.uol.com.br/folha/cotidiano/ult95u332056.shtml>>. Acesso em: jan/2008.

STEENKESTE, F.; BANERJEE, S.; COUTURIER, P. Telesurveillance of Geriatric Patients in a Hospital Using Passive Infra-red Sensors. *The Journal on Information Technology in Healthcare - JITH*. Volume 3, Issue 2. Oxford. 2005.

THOMESSE, J.-P.; BELLOT, D.; BOYER, A.; CAMPO, E.; CHAN, M.; CHARPILLET, F.; FAYN, J.; LESCHI, C.; NOURY, N.; et al. Integrated Information Technologies for Patients Remote Follow-Up and Homecare. *HealthCom 2001*. 2001.

VIRONE, G.; ISTRATE, D.; et al. First Step in Data Fusion between a Multichannel Audio Acquisition and an Information System for Home Healthcare. *EMBC 2003 – IEEE*, p. 1364-1367. 2003.

WALD, L. Data Fusion: Definitions and Architectures - Fusion of Images of Different Spatial Resolutions. *Presses de l'Ecole des Mines de Paris*. 2002.

WEISSTEIN, E. W. Box-Muller Transformation. *MathWorld – A Wolfram Web Resource*. Disponível em: <<http://mathworld.wolfram.com/Box-MullerTransformation.html>>. Acesso em: mar/2007.

WELCH, G.; BISHOP, G. An Introduction to the Kalman Filter. TR 95-041. *Department of Computer Science*. University of North Carolina at Chapel Hill. 2006.

WILLIAMS, G.; DOUGHTY, K.; BRADLEY, D.A. A system approach to achieving CarerNet – an integrated and intelligent telecare system. *IEEE Transactions on Information Technology in Biomedicine*, vol. 2, no. 1. Março/1998.

XU, W.; ZHU, Q. The Junction Tree Algorithm. *Curso CS281A/Stat241A: Statistical Learning Theory*. Lecturer: Michael I. Jordan. Novembro/2004.