



UNIVERSIDADE
ESTADUAL DE LONDRINA

FABRÍCIO FERRACIOLI

AVALIAÇÃO DE USABILIDADE:
INTEGRAÇÃO DE ABORDAGENS CONVENCIONAIS E
AUTOMATIZADAS USANDO TÉCNICAS DE WEB MINING

FABRÍCIO FERRACIOLI

AVALIAÇÃO DE USABILIDADE:
INTEGRAÇÃO DE ABORDAGENS CONVENCIONAIS E
AUTOMATIZADAS USANDO TÉCNICAS DE WEB MINING

Trabalho de Dissertação de Mestrado apresentado a
Universidade Estadual de Londrina como parte dos
requisitos para obtenção do título de Mestre em
Ciência da Computação.

Orientadora: Profa. Dra. Maria Angélica de
Oliveira Camargo Brunetto.

Londrina
2011

Catálogo Elaborado pela Divisão de Processos Técnicos da Biblioteca Central da
Universidade Estadual de Londrina

Dados Internacionais de Catalogação-na-Publicação (CIP)

F368a Ferracioli, Fabrício

Avaliação de usabilidade: integração de abordagens convencionais e automatizadas usando técnicas de Web Mining / Fabrício Ferracioli. - Londrina, 2011. 256 f.: il.

Orientador: Maria Angelica de Olivera Camargo-Brunetto.

Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ciência da Computação, 2011.

Inclui bibliografia.

1. Sites da Web - Desenvolvimento - Teses. 2. Mineração de dados (Computação) - Teses. 3. Interfaces de Usuário (Sistema de computador) - Teses. 4. Redes Neurais (Computação) - Teses. 5. Análises por conglomerados - Teses. I. Camargo-Brunetto, Maria Angelica de Oliveira. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. III. Título

CDU519 68.04

FABRÍCIO FERRACIOLI

AVALIAÇÃO DE USABILIDADE:
INTEGRAÇÃO DE ABORDAGENS CONVENCIONAIS E
AUTOMATIZADAS USANDO TÉCNICAS DE WEB MINING

Trabalho de Dissertação de Mestrado apresentado a Universidade Estadual de Londrina como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

Profa. Dra. Maria Angelica de Olivera Camargo-
Brunetto
Universidade Estadual de Londrina - UEL

Prof. Dr. Pedro Paulo da Silva Ayrosa
Universidade Estadual de Londrina - UEL

Prof. Dr. Rodolfo Miranda de Barros
Universidade Estadual de Londrina - UEL

Prof. Dr. Marcelo Soares Pimenta Universidade
Universidade Federal do Rio Grande do Sul -
UFRGS

Londrina, 09 de Setembro de 2011.

Agradecimentos

Agradeço a Deus pela sabedoria.

Agradeço a meus pais, e seu apoio incondicional.

Agradeço a minha irmã pelo apoio.

Agradeço a minha orientadora por tudo.

Agradeço as empresas ICOMP e a FS Produções Artísticas, seus sócios e funcionários pelo apoio e compreensão durante o desenvolvimento desse trabalho.

Agradeço ao professor Fábio Pitta, Nidia Hernandez, Thais Santanna e a todos os integrantes do laboratório LFIP pela colaboração durante os testes de usabilidade.

Agradeço a Carlos Ferraz e Isabela Gasparini pela colaboração durante a avaliação heurística.

Agradeço a Gabriel Gilini e Lucas Gameiro pela ajuda em São Paulo.

Agradeço a João José pela conversa que resultou na idéia desenvolvida nesse trabalho.

Se fosse fácil achar o caminho das pedras, tantas pedras no caminho não seria ruim.

Humberto Gessinger

FERRACIOLI, Fabrício. **Avaliação de usabilidade**: integração de abordagens convencionais e automatizadas usando técnicas de Web Mining. 2011. 256 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina. 2011.

RESUMO

Cada vez mais a importância da usabilidade vem sendo percebida no desenvolvimento de software. No ambiente Web sua importância é ainda maior, devido a sua característica descentralizada e anônima. Mesmo com o reconhecimento crescente de sua importância, a usabilidade ainda é vista como algo que consome muitos recursos. Este trabalho explora duas alternativas que visam tratar essa barreira, a técnica *Discount Usability* e a automatização. Ambas permitem economizar recursos como tempo, dinheiro e em alguns casos não exigem conhecimento especializado para utilização. No contexto de automatização, uma técnica que pode ser utilizada é o *Web Mining*, mais particularmente através da análise de utilização usando os *logs* de servidor para buscar por padrões de comportamento dos usuários de uma aplicação. O *Web Mining*, por sua vez, faz uso de técnicas de clusterização como parte de suas tarefas. Resultados de pesquisas nessa área mostram que métodos híbridos são melhores em aspectos importantes, como qualidade dos resultados e escalabilidade. A literatura também mostra que métodos automatizados de análise de usabilidade são capazes de auxiliar métodos não automatizados, o que traz vários benefícios. Com base nisso, esse trabalho propõe um método híbrido de avaliação de usabilidade que utiliza a rede neural GHSOM seguida do método *K-means*. Adicionalmente foi desenvolvido um protótipo de *software* que utiliza a abordagem proposta em uma aplicação Web da área de cuidados com saúde de modo a validá-la. O método proposto foi capaz de auxiliar uma avaliação de usabilidade utilizando a abordagem *Discount Usability*, fornecendo dados sobre o comportamento dos usuários em tarefas bem estabelecidas, melhorando o gasto de recursos desse tipo de avaliação. Além disso, o algoritmo de *Web Mining* proposto foi melhor que o já reconhecido SOM + K-means, como mostraram os valores lambda de Wilk medidos para cada tarefa de clusterização. Por fim, uma metodologia de avaliação de usabilidade que combina o algoritmo híbrido proposto e a *Discount Usability*, é proposta como meio de avaliação de usabilidade destinada a grupos sem especialistas de usabilidade.

Palavras-chave: Web-mining. Usabilidade. Discount Usability. Clusterização. SOM. GH – SOM. Métodos híbridos.

FERRACIOLI, Fabrício. **Usability evaluation**: integration of conventional and automated approaches using Web Mining techniques. 2011. 256 f. Dissertation (Master's in Computer Science) – State University of Londrina, Londrina. 2011.

ABSTRACT

The importance of usability has been increasingly perceived in the software development. In the Web environment its importance is even greater due to its decentralized and anonymous characteristics. Even with the growing recognition of its importance, usability is still seen as something that consumes many resources. This paper explores two alternatives that address this barrier, the Discount Usability technique and automation. Both allow you to save resources such as time, money and in some cases do not require specialized knowledge to use. In automation context, a technique that can be used is the Web Mining, particularly through the utilization analysis using the server logs to check for patterns of users behavior of an application. Web Mining, in turn, makes use of clustering techniques as part of their tasks. Results of research in this area show that hybrid methods are better in important aspects such as quality of results and scalability. The literature also shows that automated methods of usability analysis are capable of assisting non-automatic, which brings several benefits. Based on this, this paper proposes a hybrid method that uses neural network method GHSOM and then the K-means method. Additionally we developed a software prototype that uses the approach proposed in a health care web application in order to validate it. The proposed method was able to help a usability evaluation approach using the Discount Usability, providing data on users' behavior on tasks well established, improving the resources spent in this type of evaluation. Moreover, the proposed Web Mining algorithm was better than the already recognized SOM + K-means, as shown by the Wilk's lambda values measured for each clustering task. Finally, an usability evaluation methodology that combines the proposed hybrid algorithm and Discount Usability, is proposed as a means of usability evaluation aimed to groups without usability experts.

Keywords: Web-mining. Usability. Discount usability. Clustering. SOM. GHSOM. Hybrid methods.

Lista de Figuras

Figura 2.1 –	Exemplo notação para meta de nível usabilidade para erros cometidos por hora utilizando o sistema (Adaptado de (NIELSEN, 1998)).....	27
Figura 2.2 –	Curva de custo-benefício de um projeto típico adaptado de (NIELSEN, 1994a).....	41
Figura 2.3 –	Media de problemas encontrados por quantidade de avaliadores em avaliações heurísticas adaptado de (NIELSEN, a).....	43
Figura 2.4 –	Técnicas da Discount Usability agrupadas para utilização em teste de usabilidade.....	47
Figura 2.5 –	Taxonomia de Web Mining adaptado de (SRIVASTAVA; DESIKAN; KUMAR, 2005).....	49
Figura 2.6 –	Arquitetura básica de uma Rede Neural SOM	78
Figura 2.7 –	Inicialização das unidades para uma mapa expandido.....	84
Figura 3.1 –	Modelo da abordagem híbrida proposta.....	86
Figura 3.2 –	Fluxo de informações nos algoritmos pre-clusterização	93
Figura 4.1 –	Representatividade dos clusters da tarefa 1	130
Figura 4.2 –	Representatividade dos clusters da tarefa 2	131
Figura 4.3 –	Representatividade dos clusters da tarefa 3	133
Figura 4.4 –	Representatividade dos clusters da tarefa 4	135
Figura 4.5 –	Representatividade dos clusters da tarefa 5	136
Figura 4.6 –	Representatividade dos clusters da tarefa 7	138
Figura 4.7 –	Representatividade dos clusters da tarefa 8	139
Figura 4.8 –	Representatividade dos clusters da tarefa 10	140

Lista de Tabelas

Tabela 2.1 – Detalhamento do retorno de investimento para cada métrica estudada	29
Tabela 2.2 – Detalhamento do retorno de investimento para cada métrica estudada	75
Tabela 3.1 – Organização do arquivo de log comum do Apache HTTP Server	88
Tabela 3.2 – Definição do arquivo de entrada com configurações	89
Tabela 4.1 – Resultados encontrados durante as avaliações heurísticas	119
Tabela 4.2 – Representatividade de cada heurística para os problemas encontrados.....	119
Tabela 4.3 – Resultados encontrados durante os testes de usabilidade	121
Tabela 4.4 – Quantidade de problemas iguais encontrados em ambas avaliações.....	124
Tabela 4.5 – Quantidade de usuarios para cada tarefa na avaliação automatizada	127
Tabela 4.6 – Medidas de qualidade dos clusters para a tarefa 1	131
Tabela 4.7 – Medidas de qualidade dos clusters para a tarefa 2	132
Tabela 4.8 – Medidas de qualidade dos clusters para a tarefa 3	134
Tabela 4.9 – Medidas de qualidade dos clusters para a tarefa 4	135
Tabela 4.10 – Medidas de qualidade dos clusters para a tarefa 5	137
Tabela 4.11 – Medidas de qualidade dos clusters para a tarefa 7	137
Tabela 4.12 – Medidas de qualidade dos clusters para a tarefa 8	139
Tabela 4.13 – Medidas de qualidade dos clusters para a tarefa 10	141
Tabela A.1 – formato para considerações dos avaliadores da avaliação heurística ...	168
Tabela A.2 – Exemplo do formato para considerações dos avaliadores da avaliação heurística	170
Tabela B.1 – Avaliação de severidade para a tarefa 1	173
Tabela B.2 – Avaliação de severidade para a tarefa 2	175
Tabela B.3 – K Avaliação de severidade para a tarefa 3	176
Tabela B.4 – Avaliação de severidade para a tarefa 4	178
Tabela B.5 – Avaliação de severidade para a tarefa 5	179
Tabela B.6 – Avaliação de severidade para a tarefa 6	180
Tabela B.7 – Avaliação de severidade para a tarefa 7	181
Tabela B.8 – Avaliação de severidade para a tarefa 8	182
Tabela B.9 – Avaliação de severidade para a tarefa 9	183
Tabela B.10 – Avaliação de severidade para a tarefa 10	184
Tabela B.11 – Avaliação de severidade para a tarefa 11	184

Lista de Quadros

Quadro 3.1 –	Formato comum do access log do Apache HTTP Web Server.....	87
Quadro 3.2 –	Exemplo do formato comum do access log do Apache HTTP Web Server	87
Quadro 3.3 –	Exemplo do formato estendido do access log do Apache HTTP Web Server.....	88
Quadro 3.4 –	Formato padrão para o arquivo de configurações do filtro	88
Quadro 3.5 –	Formato padrão para chamada do do Pré-Processamento.....	94
Quadro 3.6 –	Formato padrão do arquivo de opções do GHSOM.....	95
Quadro 3.7 –	Formato do inputFile para o GHSOM	96
Quadro 3.8 –	Formato do descriptionFile para o GHSOM.....	97
Quadro 3.9 –	Arquivo de descrição e estrutura básicas do SOM	99
Quadro 3.10 –	Arquivo com a descrição das unidades do SOM	100
Quadro 3.11 –	Arquivo com vetor de pesos do SOM.....	101
Quadro 3.12 –	Padrão para os valores dos vetores do arquivo de pesos.....	102
Quadro 3.13 –	Formato padrão para execução do GHSOM	103
Quadro 3.14 –	Formato padrão para execução do K-means	104
Quadro F.1 –	Arquivo de entrada para tarefa cadastrar paciente	206
Quadro F.2 –	Arquivo de entrada para a tarefa avaliação inicial	207
Quadro F.3 –	Arquivo de entrada para a tarefa PFP	207
Quadro F.4 –	Arquivo de entrada para a tarefa SGRQ	207
Quadro F.5 –	Arquivo de entrada para a tarefa CRQ.....	208
Quadro F.6 –	Arquivo de entrada para a tarefa Shuttle.....	208
Quadro F.7 –	Arquivo de entrada para a tarefa DP6	208
Quadro F.8 –	Arquivo de entrada para a tarefa Pressões Respiratórias	209
Quadro F.9 –	Arquivo de entrada para a tarefa Evolução do Paciente	209
Quadro F.10 –	Arquivo de entrada para a tarefa Relatórios.....	209

Lista de Siglas e Abreviaturas

APP	Associação dos Profissionais de Propaganda
ART1	<i>Adaptive Resonance Theory 1</i>
ART2	<i>Adaptive Resonance Theory 2</i>
BIRCH	Balanced Iterative Reducing and Clustering using Hierarchies
Borg D	Escala de Borg para Dispneia
Borg F	Escala de Borg para Esforço
CEP	Código de Endereçamento Postal
CF	<i>Clustering Feature</i>
COBWEB	Sistema incremental para clusterização hierarquica conceitual
CPU	<i>Central Processing Unit</i>
CRQ	<i>Chronic Respiratory Disease Questionnaire</i>
CVF	Capacidade Vital Forçada
DATASUS	Departamento de Informatica do Sistema Único de Saúde
DOM	Document Object Model
DP6	Teste de caminhada de 6 minutos
DPOC	Doença Pulmonar Obstrutiva Crônica
FEV1	<i>Forced Expiration Volume in 1 second</i>
FC	Frequência Cardíaca
FS	Algoritmo classico para reconhecimento de caminhos de acesso frequente
GDL	Guideline Definition Language
GG	<i>Growing Grid</i>
GHSOM	<i>Growing Hierarchical Self-Organizing Map</i>
GKA	<i>Genetic K-means Algorithm</i>
HTML	<i>Hyper Text Markup Language</i>
HTTP	<i>Hyper Text Transfer Protocol</i>
HU	Hospital Universitário
IHC	Interação Humano-Computador
IP	<i>Internet Protocol</i>
ISO	<i>International Organization for Standardization</i>
ISP	<i>Internet Service Provider</i>

I/O	<i>Input/Output</i>
LFIP	Laboratório de Pesquisa em Fisioterapia Pulmonar
MQE	Medium Quantization Error
NASA	National Aeronautics and Space Administration
NP-Hard	Non-deterministic Polynomial-time Hard
PA	Pressão Arterial
PBIRCH	Parallel Balanced Iterative Reducing and Clustering using Hierarchies
PFE	Pico de Fluxo Expiratório
PPF	Prova de Função Pulmonar
PHP	PHP Hyper Text Preprocessor
POS BD	Apos Bronco Dilatador
SacarWeb	Software para avaliação cardio-respiratória na Web
SBC	Sociedade Brasileira de Computação
SGRQ	St. George's Respiratory Questionnaire Paul Jones
SOM	<i>Self Organizing Map</i>
SpO ₂	Saturação o de oxigênio
SSB	<i>Sum of Squares Between</i>
SST	<i>Sum of Squares Total</i>
SSW	<i>Sum of Squares Within</i>
TCC	Trabalho de Conclusão de Curso
UDESC	Universidade do Estado de Santa Catarina
UEL	Universidade Estadual de Londrina
UFRGS	Universidade Federal do Rio Grande do Sul
URL	<i>Uniform Resource Locator</i>
VEF1	Volume de Expiração Forçada em 1 segundo
VFC	<i>Vital Forced Capacity</i>

Sumário

1	Introdução	17
1.1	Objetivos	18
1.2	Justificativa	19
1.3	Contribuições	21
1.4	Organização do Trabalho	22
2	Revisão da Literatura	23
2.1	Usabilidade	23
2.1.1	Engenharia de Usabilidade	24
2.1.1.1	Conheça o usuário	26
2.1.1.2	Análise competitiva	26
2.1.1.3	Definindo metas de usabilidade	27
2.1.1.4	Projetos paralelos	30
2.1.1.5	Projeto participativo	30
2.1.1.6	Coordenando o projeto total da interface	31
2.1.1.7	Aplicação de recomendações e análise heurística	31
2.1.1.8	Prototipação	32
2.1.1.9	Avaliação da Interface	33
2.1.1.10	Projeto Interativo	35
2.1.1.11	Coleta de feedback de uso de campo	35
2.1.2	Alternativa aos elevados custos de testes de usabilidade	35
2.1.3	A Discount Usability Engineering	39
2.1.3.1	Avaliação heurística	42
2.1.3.2	Pensando alto simplificado	45
2.1.3.3	Protótipos	46
2.1.3.4	Teste de usabilidade	47
2.2	Web Mining	48
2.2.1	Importância das técnicas de clusterização em Web Mining	50
2.3	Trabalhos Relacionados	51
2.3.1	Trabalhos relacionados a avaliação de usabilidade automatizada	51

2.3.2	Trabalhos relacionados a Web Mining	57
2.3.2.1	Síntese dos métodos apresentados	70
2.4	Revisão das técnicas de clusterização utilizadas	75
2.4.1	K-means	75
2.4.2	Mapas Auto-Organizáveis	77
2.4.3	O Growing Hierarquical Self-Organizing Map	80
3	 Materiais e Métodos	85
3.1	 Método híbrido proposto	86
3.1.1	Algoritmos de pré-clusterização	87
3.1.1.1	Definindo os arquivos de entrada para o pré- processamento	93
3.1.2	Clusterização com GHSOM	94
3.1.2.1	O pacote GHSOM	95
3.1.2.2	Executando a clusterização com GHSOM	102
3.1.3	Aplicação do K-means	103
3.1.3.1	Implementação utilizada do K-means	104
3.1.3.2	Executando a clusterização com K-means	104
3.1.4	Análise humana dos resultados do algoritmo híbrido	105
3.1.5	Possibilidades para transcrição dos problemas de usabilidade em tarefas de Desenvolvimento	106
3.2	 Caracterização da aplicação Web alvo do método híbrido proposto	107
3.3	 Aplicação do Discount Usability	108
3.3.1	Realização da avaliação heurística	110
3.3.2	Realização do teste de usabilidade	112
3.4	 Verificação da consistência das abordagens de usabilidade utilizadas	119
4	 Resultados e Discussão	118
4.1	 Resultados da Discount Usability	118
4.1.1	Resultados da avaliação heurística	118
4.1.2	Resultados dos testes de usabilidade	120
4.1.3	 Comparação da Avaliação Heurística com os Testes de Usabilidade	124

4.2	Resultados da aplicação do método híbrido proposto.....	
4.2.1	Tarefa 1: Cadastrar paciente	129
4.2.2	Tarefa 2: Avaliação inicial	131
4.2.3	Tarefa 3: Exame PFP.....	132
4.2.4	Tarefa 4: Exame SGRQ.....	134
4.2.5	Tarefa 5: Exame CRQ	135
4.2.6	Tarefa 6: Exame Shuttle.....	137
4.2.7	Tarefa 7: Exame DP6.....	137
4.2.8	Tarefa 8: Exame Pressões Respiratórias	138
4.2.9	Tarefa 9: Evolução do paciente.....	139
4.2.10	Tarefa 10: Relatórios.....	140
4.3	Comparação das avaliações com Discount Usability e com o método proposto	141
4.4	Generalização dos resultados obtidos com o estudo de caso	144
4.5	Generalização da metodologia proposta para outras aplicações Web	146
5	Conclusão e trabalhos futuros	147
5.1	Trabalhos Futuros	150
	Referências.....	152
	APÊNDICE A – Guia para avaliação heurística do SacarWeb.....	160
A.1	Cenários de utilização.....	161
A.1.1	Inserir os dados pessoais do paciente.....	162
A.1.2	Preencher a avaliação inicial.....	162
A.1.3	Criar uma nova avaliação	162
A.1.4	Informar os dados dos exames.....	162
A.1.4.1	PFP	163
A.1.4.2	Pressões Respiratórias	163
A.1.4.3	SGRQ	163
A.1.4.4	CRQ.....	164
A.1.4.5	Shuttle	165
A.1.4.6	DP6	166

A.1.5	Verifique a evolução dos pacientes entre as avaliações	166
A.1.6	Obtenha um relatório de dados	167
A.2	Protocolo de execução da avaliação.....	167
APENDICE B – Problemas encontrados durante avaliação heurística.....		171
B.1	Tarefa 1: Cadastrar Paciente.....	171
B.2	Tarefa 2: Preencher a avaliação inicial.....	174
B.3	Tarefa 3: Criar uma nova avaliação	175
B.4	Tarefa 4: Exame PFP.....	176
B.5	Tarefa 5: Exame SGRQ	178
B.6	Tarefa 6: Exame CRQ.....	179
B.7	Tarefa 7: Exame Shuttle.....	180
B.8	Tarefa 8: Exame DP6	181
B.9	Tarefa 9: Exame Pressões Respiratórias	182
B.10	Tarefa 10: Evolução do Paciente	183
B.11	Tarefa 11: Relatórios	184
APÊNDICE C – Guia para teste de usabilidade do SacarWeb.....		185
C.1	Importante.....	185
APÊNDICE D – Tarefas a serem realizadas durante teste de usabilidade do		
	SacarWeb.....	187
D.1	Observações importantes	187
APÊNDICE E – Síntese dos problemas encontrados durante os testes de		
	usabilidade.....	188
E.1	Tarefa 1: cadastrar paciente	188
E.1.1	Boas práticas observadas	190
E.2	Tarefa 2: Avaliação Inicial.....	190
E.2.1	Boas práticas observadas	192
E.3	Tarefa 3: Avaliação.....	192
E.3.1	Boas práticas observadas	194
E.4	Tarefa 4: Shuttle	194
E.4.1	Boas práticas observadas	196
E.5	Tarefa 5: SGRQ	196
E.5.1	Boas praticas observadas	197
E.6	Tarefa 6: DP6	197
E.6.1	Boas práticas observadas	199

E.7	Tarefa 7: CRQ	199
E.7.1	Boas práticas observadas	200
E.8	Tarefa 8: PFP	200
E.8.1	Boas práticas observadas	202
E.9	Tarefa 9: Pressões Respiratórias	202
E.9.1	Boas práticas observadas	202
E.10	Tarefa 10: Evolução do paciente	203
E.10.1	Boas práticas observadas	203
E.11	Tarefa 11: Relatórios	204
E.11.1	Boas práticas observadas	205
APÊNDICE F – Definição das tarefas para a abordagem automatizada		
	proposta	206
F.1	Tarefa 1: Cadastrar Paciente	206
F.2	Tarefa 2: Preencher avaliação inicial	207
F.3	Tarefa 3: Exame PFP	207
F.4	Tarefa 4: Exame SGRQ	207
F.5	Tarefa 5: Exame CRQ	208
F.6	Tarefa 6: Exame Shuttle	208
F.7	Tarefa 7: Exame DP6	208
F.8	Tarefa 8: Exame Pressões Respiratórias	209
F.9	Tarefa 9: Evolução do Paciente	209
F.10	Tarefa 10: Relatórios	209
APÊNDICE G – Resultados da avaliação utilizando a abordagem proposta		
G.1	Tarefa 1: Cadastrar Paciente	210
G.2	Tarefa 2: Preencher avaliação inicial	218
G.3	Tarefa 3: Exame PFP	234
G.4	Tarefa 4: Exame SGRQ	237
G.5	Tarefa 5: Exame CRQ	240
G.6	Tarefa 6: Exame Shuttle	241
G.7	Tarefa 7: Exame DP6	242
G.8	Tarefa 8: Exame Pressoes Respiratorias	244
G.9	Tarefa 9: Evolução do Paciente	246
G.10	Tarefa 10: Relatórios	246

1 Introdução

A *World Wide Web* proporcionou às organizações um meio de quebrar barreiras territoriais para atingir seu público universalmente, de forma barata e ágil. Essa oportunidade foi percebida por diversas organizações, sendo a Web adotada massivamente como plataforma para aplicações com variados fins. Pela sua característica organizacional descentralizada e inserção facilitada, cada vez mais a Web é utilizada (O'REILLY, 2005), sendo uma tarefa crucial para o sucesso e diferenciação da concorrência. Uma das maneiras de se obter essa diferenciação é a satisfação dos usuários (MORANDINI, 2003). Apesar de poder ser caracterizada como subjetiva, ela pode ser medida objetivamente e uma das formas de avaliá-la é através do conceito de usabilidade.

A usabilidade é um fator de qualidade importante em qualquer sistema interativo. Aplicações projetadas com cuidados relacionados a usabilidade e que procuram manter bons níveis dela desde a fase de projeto, possuem maiores chances de sucesso frente as concorrentes, uma vez que os estudos realizados visam atender as necessidades específicas do público alvo do software (NIELSEN, 1998). A usabilidade também é de grande utilidade em processos de melhoria de aplicações já existentes, podendo resolver problemas que estão dificultando o sucesso ou adoção satisfatória de um software (MORANDINI, 2003).

Na Web a usabilidade ganha ainda mais importância, visto que uma das consequências imediatas da falta de usabilidade pode ser a migração do usuário para uma aplicação concorrente, devido a facilidade de encontrar alternativas semelhantes de aplicações Web (NIELSEN; LORANGER, 2006). Além deste fator, na Web também é maior a complexidade de uma avaliação de usabilidade, pois na maioria dos casos o ambiente em que o sistema opera não é controlado, os usuários são totalmente desconhecidos e heterogêneos (MORANDINI, 2003). Ações focadas em usabilidade podem ser o diferencial competitivo frente a concorrentes no caso de aplicações comerciais, ou o diferencial de sucesso na adoção de uma ferramenta, no caso de uma aplicação Web que possui operações complexas e ligadas a objetivos específicos, como operações e processos críticos dentro do ciclo de trabalho (NIELSEN; LORANGER, 2006).

Apesar de estar sendo cada vez mais percebida com importância (FERRE et al., 2001), (NIELSEN, 2008), abordagens tradicionais de estudos de usabilidade são conhecidas por serem custosas com relação a recursos financeiros e tempo (NIELSEN, 1994a) (MORANDINI, 2003). Muitas vezes os especialistas da área tendem a usar configurações complexas de testes, buscando pelos melhores métodos, teorias e laboratórios. Esse tipo de rigor necessita de profissionais mais capacitados, maior gama de conhecimentos de áreas diferentes, maior quantidade de usuários para testes, locais e recursos mais adequados. Todas essas necessidades além de custosas podem intimidar profissionais envolvidos no projeto do software não habituados a terminologia, (NIELSEN, 1998), (AU et al., 2008) e por esse motivo desestimula uma atenção apurada com a usabilidade.

Por outro lado, a falta de usabilidade pode significar o fracasso de um projeto, culminando em resultados com os quais uma organização não pode arcar. Um meio de contornar o problema de modo a atingir essa necessidade são abordagens menos custosas de avaliação de usabilidade, como a técnica de *Discount Usability*¹ ou análises automatizadas, capazes de diminuir custos e apresentar resultados mais rapidamente, mantendo um nível de qualidade tão satisfatório quanto o de uma abordagem tradicional. Com esse fim, análises automatizadas podem ser realizadas em uma aplicação Web através de uma técnica conhecida por *Web mining*². Essa técnica é aplicável para objetivos variados em uma aplicação Web, mas sua utilização para análises automatizadas de usabilidade tem mostrado resultados tão confiáveis quanto os obtidos por abordagens tradicionais (KOSALA; BLOCKEEL, 2000). Por ser uma alternativa mais rápida e menos custosa, o estudo dessas técnicas é de interesse.

1.1 **Objetivos**

O objetivo geral deste trabalho foi analisar formas de integração de mecanismos não tradicionais de auxílio para detecção automatizada de falhas de usabilidade em aplicações Web com mecanismos tradicionais.

Para isso, uma metodologia automatizada de avaliação de usabilidade é proposta, utilizando a técnica de *Web mining* de utilização como base para análise dos dados. Desse modo, foi desenvolvido um método para mineração dos dados, e sendo assim o trabalho possui como objetivos específicos:

1. desenvolvimento de um método clusterização de dados de utilização da Web necessário

¹Será discutida em 2.1.3

²Será discutida em 2.2

na metodologia de avaliação automatizada proposta;

2. realizar uma avaliação tradicional de usabilidade baseada em *Discount Usability* de modo a verificar as possibilidades de integração com a proposta automatizada.

1.2 Justificativa

Dentre as três abordagens possíveis no *Web Mining* (COOLEY; MOBASHER; SRIVASTAVA, 1997) (KOSALA; BLOCKEEL, 2000) - conteúdo, estrutura e utilização - utilização foi selecionada neste trabalho pois apresenta resultados mais expressivos com relação a análise de usabilidade pois trata de dados de interações reais dos usuários com a aplicação Web (SCHWERZ; MORANDINI; SILVA, 2007). Isso é explicado pelo fato da fonte dos dados de mineração serem as interações dos usuários com a aplicação Web encontradas nos *logs* do servidor. Além disso, os resultados alcançados com mineração de estrutura e conteúdo não seriam capazes de fornecer resultados diretamente relacionados as interações dos usuários, pois são recursos relativos a aplicação Web. Desse modo, mineração de utilização possui a maior relevância para esse trabalho, além de atacar a maior fonte de gastos em abordagens tradicionais de teste de usabilidade, o teste com usuários.

Uma vez que os *logs* de servidor possuem tamanho indefinido, e em muitos casos de aplicações reais esse tamanho é muito extenso uma análise de *log* sem métodos computacionais torna-se praticamente inviável, além de ser necessário métodos escaláveis que suportem grandes aplicações do mundo real. Ainda mais importante que diminuir a quantidade de dados para análise, é identificar grupos de usuários com comportamentos similares, característica muito importante para a análise de usabilidade. Algoritmos de clusterização são um recurso computacional que atendem esses requisitos e por isso foram selecionados neste trabalho. Duas abordagens de clusterização clássicas que possuem essas características são a rede neural *Self Organizing Maps* e o método K-means, utilizados na proposta desse trabalho. Essa decisão foi baseada em experimentos analisados no capítulo 2 que mostraram claramente que os métodos atendem essas necessidades.

As melhorias propostas em (KUO; HO; HU, 2002) e (KUO et al., 2006), que também utilizam algoritmos de clusterização, como a rede neural não supervisionada SOM em conjunto com o algoritmo *K-means* e sua variação genética, respectivamente, provaram ser bastante efetivas em segmentação de mercado. Outro trabalho (KUO; LIAO; TU, 2005) com a rede neural não-supervisionada ART2 também em conjunto com o algoritmo *K-means* e sua variação genética confirmou ser efetiva em *Web Mining*. A junção de dois algoritmos forneceu

resultados melhores para clusterização de dados nos trabalhos citados, o que justifica a análise de diferentes combinações de métodos para uma mesma área. Particularmente, a junção de uma rede neural que forneça a solução inicial para um método estatístico como o *K-means* provou ser capaz de atingir bons níveis de clusterização tanto em ambiente controlado, quanto em experimentos práticos, como mostraram os trabalhos analisados.

Em (SORIANO-ASENSI et al., 2008) os Mapas Auto-Organizáveis com Crescimento Hierárquico (GHSOM) são utilizados em *Web Mining*, devido a deficiência da rede neural SOM tradicional em determinar a quantidade de neurônios antes do treinamento. Os resultados alcançados são satisfatórios em termos de clusterização e complexidade computacional, pois são capazes de lidar com grandes quantidades de dados, característica muito comum no ambiente Web. Além disso, a rede GHSOM é capaz de mapear possíveis hierarquias existentes nos dados clusterizados, característica interessante para o campo de aplicação selecionado. Um exemplo comum acontece com *e-commerces*, onde é comum encontrar usuários que chegam até uma página de pedido para verificar o valor final somente para consulta e aqueles que realmente compram. Apesar de ambos realizarem a mesma tarefa, um conclui e outro abandona, o que é uma informação importante para os fornecedores do *e-commerce* e que deve ser capturada pelo algoritmo. Uma vez que a rede neural GHSOM é capaz de superar limitações encontradas na rede SOM tradicional e lidar com características facilmente encontradas no ambiente Web do mundo real, justifica-se sua utilização como método de clusterização de dados.

O *K-means* foi escolhido como segunda etapa do método proposto por ser considerado o algoritmo padrão para clusterização de dados (BRADLEY; FAYYAD, 1998) e por ser capaz de melhorar a saída de outros métodos de clusterização (KANUNGO et al., 2002). Além disso, observa-se na revisão da literatura na seção 2.3 que a maior parte de novos métodos propostos é comparada a resultados alcançados pelo *K-means*, assim como também é comum encontrar métodos híbridos com ele envolvido em uma das etapas.

A escolha da área de cuidados com saúde como alvo do experimento prático é válida pois é comum que aplicações dessa área possuam tarefas bastante específicas, característica crucial para avaliação da ferramenta, e desempenham papel de grande importância na organização de entidades de saúde, muitas vezes fazendo parte de processos organizacionais e sendo de uso obrigatório. Além disso, a usabilidade é tida como uma das características cruciais para a adoção de sistemas nessa área (ROSE et al., 2005) e avaliações automatizadas tendem a possuir uma necessidade crescente dentro do processo de usabilidade de aplicações web (BRINCK; HOFER, 2002).

Os dados clusterizados pela abordagem híbrida proposta foram analisados de maneira humana de maneira a encontrar falhas de usabilidade de a partir dos mapeamentos encontrados pela abordagem automatizada. A falta de usabilidade em sistemas com as características apresentadas pode resultar na não adoção da ferramenta, culminando na falha de processos e até mesmo em consequências mais graves relacionadas a saúde dos que dependem desse tipo de ferramenta. Além disso, a usabilidade nesse tipo de aplicação é tida como barreira para sistemas efetivos (BADENOCH; TOMLIN, 2004), (KUSHNIRUK; PATEL; CIMINO, 1997).

1.3 Contribuições

O método híbrido desenvolvido foi comparado de maneira semelhante a abordagem encontrada em (KUO; HO; HU, 2002), (KUO et al., 2006) e (KUO; LIAO; TU, 2005). Esperava-se que o método proposto fosse melhor que o método encontrado em (KUO; HO; HU, 2002), um método híbrido que utiliza a rede neural SOM tradicional em conjunto com o K-means. Isso porque o GHSOM, utilizado para o desenvolvimento desse trabalho, possui melhorias em relação ao SOM, e por isso a abordagem a desenvolvida tendia a ser superior. O fato foi confirmado na maior parte dos casos testados através do valor lambda de Wilk, medida utilizada também no trabalho (KUO; HO; HU, 2002).

Profissionais interessados em aspectos de experiência do usuário de aplicações Web poderão utilizar a metodologia automatizada aqui descrita para a realização de avaliações de usabilidade em sistemas Web, fomentando processos de criação ou melhoria das aplicações. Assim como afirma (MORANDINI, 2003) os resultados apresentados na avaliação de usabilidade automatizada proposta são capazes de validar os dados encontrados por uma avaliação de usabilidade tradicional, além de fornecer dados adicionais obtidos com maior facilidade através dela. A utilização do método proposto também é vantajosa pois necessita de menos tempo e pode ser realizada por profissionais com baixa experiência na área de usabilidade, como mostrou o estudo de caso ao avaliar a aplicação SacarWeb. (NIELSEN, 1998) mostra que a usabilidade é muito impopular entre desenvolvedores, devido a falta de conhecimento ou até mesmo por acreditarem que gera trabalho adicional (AU et al., 2008). Essa dificuldade também é atacada pela abordagem automatizada, que retira boa parte do trabalho relacionado a usabilidade das mãos de um analista. Além disso, o método proposto pode ser integrado com facilidade a abordagens tradicionais de avaliação de usabilidade, como mostrou nosso estudo de caso com *Discount Usability*.

Além das contribuições relacionadas ao método desenvolvido, houve a con-

tribuição relacionada a aplicação da abordagem *Discount Usability*. O estudo de caso realizado mostrou que mesmo equipes formadas por pessoas com baixa experiência em usabilidade são capazes de realizar avaliações após serem apresentadas aos conceitos, como prevê o trabalho (NIELSEN, 1998).

A comunidade externa à Universidade também é beneficiada através da aplicação da análise de usabilidade automatizada, uma vez que a aplicação Web utilizada no Hospital Universitário da Universidade Estadual de Londrina foi analisada. A aplicação, chamada SacarWeb (CAMARGO-BRUNETTO; ROSSI, 2005), é utilizada por um grupo de pesquisa da área de saúde, focada em reabilitação física de pacientes com doenças respiratórias crônicas. Ainda na área da saúde, é possível realizar melhorias em sistemas fornecidos pelo Datasus³, Departamento de Informática do Sistema Único de Saúde.

Além disso, devido a aplicação satisfatória da metodologia proposta na área de cuidados com saúde, é possível utilizá-la em outras áreas de conhecimento de aplicações Web. Através disso, são abertas possibilidades de utilização em diferentes aplicações Web de modo a melhorar sua usabilidade mais facilmente. Desse modo, possivelmente esse trabalho poderá ajudar na inclusão digital, política incentivada pelo Ministério da Ciência e Tecnologia⁴, auxiliando no acesso dos cidadãos com menos conhecimento da interação com o computador, inclusive em sistemas Web governamentais. (CARVALHO, 2003), (SILVINO; ABRAHÃO, 2003) mostram que a usabilidade possui importância também nesse tipo de processo.

1.4 Organização do trabalho

O restante do trabalho encontra-se organizado da seguinte maneira. No capítulo 2 é realizada uma revisão da bibliografia para os principais conceitos de Usabilidade, *Web Mining* e técnicas de clusterização utilizadas, assim como uma revisão do estado da arte para *Web Mining* e abordagens automatizadas de usabilidade. No capítulo 3 o método híbrido proposto deste trabalho é detalhado, assim como os materiais e métodos utilizados durante as avaliações de usabilidade com *Discount Usability* e validação. No capítulo 4 os resultados do trabalho são exibidos e discutidos, separados por categoria de método utilizado e ao final comparando ambas avaliações. A apresentação do trabalho é concluída no capítulo 5 onde revemos o que o trabalho pretendia, o que foi realizado e sugerimos trabalhos futuros. Os apêndices reúnem documentos utilizados durante as avaliações e seus resultados completos.

³<http://www.datasus.gov.br>

⁴<http://www.mct.gov.br>

2 Revisão da Literatura

Este capítulo apresenta os conceitos empregados neste trabalho e uma revisão do estado da arte para aplicações de análise automatizada de usabilidade e *Web Mining*. O leitor que já possuir conhecimento sobre usabilidade e *Web Mining* poderá se dirigir diretamente para a seção de trabalhos relacionados 2.3, que apresenta o estado da arte em aplicações de análise automatizada de usabilidade e *Web Mining*.

2.1 Usabilidade

A usabilidade é uma característica de sistemas interativos, definida pela ISO 9241-11 do seguinte modo: “Extensão em que um produto pode ser utilizado por usuários específicos para atingir objetivos específicos com efetividade, eficiência e satisfação em um contexto específico de uso”. Dessa forma, a usabilidade é um atributo, e apesar de mensurável estatisticamente, é uma característica subjetiva definida na ocorrência da utilização da ferramenta e dependente de contexto. A ISO 9241-11 mostra benefícios da medição de usabilidade em termos de desempenho e satisfação dos usuários (STANDARDIZATION, 1998).

Em linhas gerais, é interessante medir as três características que fazem parte da definição de usabilidade:

- Efetividade: precisão e perfeição com que usuários atingem objetivos específicos;
- Eficiência: recursos gastos com relação a efetividade e completude com que o objetivo foi atingido;
- Satisfação: conforto e aceitabilidade do sistema de trabalho para seus usuários e outras pessoas afetadas por seu uso.

Os métodos para avaliação de usabilidade podem ter várias finalidades: auxiliar em futuros projetos de interface, identificação e diagnóstico de problemas, avaliação da implementação entre

outros. As avaliações podem ser quantitativas ou qualitativas. Na perspectiva qualitativa, a procura principal deve ser o atendimento das necessidades dos usuários, o que é em boa parte dos casos uma medida muito subjetiva. Já na perspectiva quantitativa temos uma referência aos valores que podem ser obtidos através de medições da interação entre os usuários e os componentes de um web site, em termos da eficiência, da eficácia, da satisfação do usuário, da intuitividade e da facilidade de aprender e lembrar (MORANDINI, 2003). Outra classificação está vinculada aos métodos, que podem ser de inspeção, que não utilizam usuários finais, ou de teste, que utilizam (HOLZINGER, 2005).

Apesar de não existir um consenso sobre qual metodologia ou técnica deva ser utilizada para se obter uma avaliação eficiente (MORANDINI, 2003), algumas abordagens são bem conhecidas na avaliação da usabilidade e melhores resultados são obtidos ao combinar métodos de inspeção e testes (HOLZINGER, 2005). O fato das técnicas de Engenharia de Software não incorporarem elementos da Engenharia de Usabilidade também dificulta a adoção de técnicas de usabilidade por desenvolvedores sem seu conhecimento (NIELSEN, 1998), (SEFFAH; DESMARAIS; GULLIKSEN, 2005). Essa deficiência tem como resultado 80% dos custos de manutenção de software relacionados com problemas de utilização, e que destes, 64% estão relacionados a usabilidade (SEFFAH; DESMARAIS; GULLIKSEN, 2005). Esses valores podem diminuir drasticamente principalmente ao se considerar a usabilidade em estados iniciais do ciclo de vida de um projeto (NIELSEN, 1998).

2.1.1 Engenharia de Usabilidade

As metas de usabilidade a serem alcançadas por uma aplicação são definidas durante a Engenharia de Usabilidade. Ela especifica características de usabilidade quantitativamente no início do projeto, período em que também acontecem atividades significativas, mas é medida em etapas posteriores do ciclo de vida (FERRE et al., 2001), (NIELSEN, 1998). As decisões tomadas no início do projeto além de afetarem todo seu ciclo de vida, terão repercussão em novas versões dele, uma vez que fatores humanos envolvidos em um produto particular tem grande impacto no lançamento de novas versões do mesmo sistema. Um dos grandes benefícios em se definir metas de usabilidade se deve ao fato de que se estas forem claras e articuladas ajudam o projeto de interface a focar seus esforços, simplificando-o e agilizando-o (MAYHEW, 1999).

Segundo (NIELSEN, 1998), o ciclo de vida da engenharia de usabilidade é dividido nas seguintes etapas:

2.1 Usabilidade

1. Conheça o usuário
 - Características individuais
 - Tarefas atuais e desejadas
 - Análise funcional
 - Evolução do usuário e do trabalho
2. Análise competitiva
3. Definindo metas de usabilidade
 - Análise de impacto financeiro
4. Projetos paralelos
5. Projeto Participativo
6. Coordenando o projeto total da interface
7. Aplicação de recomendações e análise heurística
8. Prototipação
9. Avaliação da Interface
10. Projeto interativo
 - Captura de lógica de projeto
11. Coleta de *feedback* de uso de campo

Mesmo não incluindo todas as etapas, o esforço de engenharia de usabilidade pode obter sucesso. O caminho de menor custo é obtido ao iniciar as atividades de usabilidade assim que o projeto for iniciado, o que evita mudanças tardias e de alto custo no projeto.

Antes de realizar qualquer uma das atividades acima é importante utilizar alguns métodos para que elas sejam aplicadas efetivamente. É necessário definir um plano explícito do que se pretende realizar ao utilizar determinado método. Para testes empíricos, por exemplo, é importante definir quantos usuários testar, que tipo, quais serão as tarefas executadas e um esquema de tempo para os estudos. Também é importante que o planejamento seja revisto por uma pessoa que não esteja na mesma equipe e que possa criticar de uma perspectiva neutra. Por fim, é importante realizar um teste piloto com cerca de 10-15% dos recursos pretendidos para os testes. O plano então deverá ser revisto com 85-90% dos recursos para solucionar as dificuldades encontradas durante o piloto.

2.1.1.1 Conheça o usuário

O primeiro passo, e o mais básico no processo, é estudar os potenciais usuários e a utilização do produto. Características individuais dos usuários e variabilidade em tarefas são dois fatores que possuem o maior impacto em usabilidade, e necessitam ser estudados com cautela.

Desse modo é necessário conhecer as pessoas que utilizarão o sistema. Características individuais que auxiliam a definir níveis de complexidade da interface são: experiência com o trabalho, nível educacional, idade, experiência com computadores, habilidade com o idioma e de leitura, entre outros. Uma definição geral é que a interface deve ser o mais simples possível, de modo que os usuários consigam utilizá-la sem treinamento.

A **análise de tarefas** também é essencial para o projeto do sistema. Os objetivos gerais dos usuários devem ser estudados em conjunto com o modo que eles executam determinada tarefa, suas necessidades de informação e como lidar com circunstâncias excepcionais ou emergências. Deve-se buscar situações que possam servir de metáforas para a interface com o usuário, especialmente de usuários alvo, assim como identificar fraquezas nas situações, como locais onde ocorrem muitas falhas, tempo demais é gasto ou algum desconforto é percebido.

Em conjunto com a análise de tarefas, deve-se realizar a **análise de intenção** que deve identificar as razões pela qual uma tarefa é realizada.

Finalmente deve-se prever a **evolução do usuário** conforme ele conhece o *software*, possibilitando a ele novos modos de utilização do sistema. Por esse motivo um sistema não deve ser projetado somente para usuários iniciantes, mas também prevendo necessidades de usuários experientes que coexistem com os iniciantes.

2.1.1.2 Análise competitiva

Produtos existentes e possivelmente concorrentes são normalmente o melhor protótipo que se pode ter para um produto em desenvolvimento. É possível realizar análises heurísticas e testes empíricos com usuários neles, pois uma vez implementado, o nível de facilidade para realização de testes aumenta. As conclusões obtidas a partir desses testes podem servir como fonte de comparação ou definição de metas para um *software* em fase de projeto. Outro fator importante é que uma vez que o produto está disponível ao público, é possível que grande esforço em sua estabilidade tenha se despendido, e por isso ele está menos suscetível a falhas que um protótipo, o que deixa os testes mais realistas.

2.1.1.3 Definindo metas de usabilidade

A usabilidade compreende diversos aspectos algumas vezes conflitantes. Um deles é a opinião dos usuários, considerada importante na concepção de um sistema, apesar dos usuários dificilmente conhecerem suas necessidades reais. A maioria dos projetos não consegue priorizar todos os aspectos com pesos iguais, sendo necessário criar prioridades baseadas na análise dos usuários e suas tarefas. Antes de iniciar o projeto de interface, é importante definir as métricas de usabilidade de interesse do projeto e especificar as metas da interface com o usuário em termos de usabilidade medida a partir das métricas.

As metas de usabilidade devem levar em conta o perfil dos usuários da aplicação e as tarefas por eles executadas dentro de um contexto. A meta de usabilidade requer a especificação do nível mínimo de aceitação para os aspectos de interesse de modo que seja possível disponibilizar o produto. Atenção é necessária ao definir as métricas, pois algumas metas são interessantes para alguns tipos de usuários e tarefas, enquanto não são para outros (MAYHEW, 1999). Um exemplo é a facilidade de aprendizado, que não é tão importante quando os usuários utilizarão o sistema frequentemente e após treinamento intensivo, enquanto sistemas raramente utilizados e com pouco ou nenhum treinamento devem priorizar esse objetivo.

Um nível maior de detalhamento para as metas pode incluir o nível planejado e o nível atual. Esses valores podem ser obtidos dos níveis de usabilidades obtidos nos testes com softwares existentes e também é possível inserir o melhor valor teoricamente possível, mesmo que ele não possa ser atingido. A figura 2.1 exibe uma possível notação para definição dos níveis de usabilidade a serem atingidos por um sistema. Neste exemplo, a métrica utilizada é de erros por hora cometidos no sistema.



Figura 2.1: Exemplo notação para meta de nível usabilidade para erros cometidos por hora utilizando o sistema (Adaptado de (NIELSEN, 1998))

Algumas métricas relacionadas a desempenho ou eficiência dizem respeito a quantidade de esforço necessária aos usuários para atingir seus objetivos. Algumas delas são (MAYHEW, 1999), (NIELSEN, 1998):

- Tempo para completar uma tarefa na primeira tentativa

2.1 Usabilidade

- Tempo para completar uma tarefa após um tempo específico sem utilizar o software
- Tempo para completar um conjunto de instruções particular
- Tempo para executar uma tarefa comparado ao de um usuário experiente
- Tempo para aprender um critério
- Tempo para atingir desempenho de usuário experiente
- Tempo gasto corrigindo erros
- Quantidade de ícones lembrados após completar a tarefa
- Quantidade de botões acionados para completar a tarefa
- Quantidade e tipo de erro por tarefa
- Quantidade de acessos a ajuda ou documentação
- Quantidade de usuários que comete um erro em particular
- Quantidade de usuários completando uma tarefa com sucesso

Métricas relacionadas a efetividade estão relacionadas a completude e precisão com que os usuários atingem seus objetivos. Algumas delas são (NIELSEN, 1998):

- Porcentagem de funções relevantes utilizadas
- Porcentagem de tarefas completadas com sucesso na primeira tentativa
- Porcentagem de usuários capazes de completar uma tarefa com sucesso
- Porcentagem de usuários que realizam tarefas chave sem ler o manual
- Quantidade de erros persistentes
- Quantidade de erros por unidade de tempo
- Quantidade de erros cometidos ao realizar tarefas específicas
- Quantidade de pedidos de auxílio para completar uma tarefa

Medidas de satisfação ou emocionais referem-se a como os usuários se sentem sobre o sistema. Alguns exemplos são (NIELSEN, 1998):

2.1 Usabilidade

- Razão de adjetivos positivos por negativos para descrever o produto
- Porcentagem de clientes que avaliam melhor o produto atual que o anterior
- Porcentagem de clientes que se sentem dominando o produto
- Porcentagem de clientes que recomendariam o produto a um amigo após duas horas de uso
- Porcentagem de clientes que acham o produto mais fácil de utilizar que um concorrente

É possível utilizar outros tipos de métrica, como taxas de conversão (quantidade de vendas ou geração de interesse em um produto) e valores de tráfego (estatísticas relacionadas a *page views*) e utilização de uma função alvo (quantidade de cliques num *link* para uma informação crucial) (NIELSEN, 2008).

Paralelamente é possível mensurar o impacto financeiro da usabilidade para o sistema, que pode ser utilizado como uma métrica. Para isso deve-se estimar a quantidade de pessoas que utilizará o sistema, seus custos e o tempo aproximado que o utilizarão.

Tendo em mãos os níveis de usabilidade e custos de utilização antes do projeto do sistema é possível medir a melhoria após o projeto e o retorno de investimento. Em (NIELSEN, 2003) os ganhos de usabilidade variaram de 100% a 202%, com uma média de 135% para um investimento de 10% do valor do projeto. Os valores específicos obtidos para as métricas utilizadas são visualizados na tabela 2.1:

Métrica	Melhoria média em projetos Web
Vendas / taxa de conversão	100%
Tráfego / número de visitantes	150%
Desempenho do usuário / produtividade	161%
Uso de características específicas (alvo)	202%

Tabela 2.1: Detalhamento do retorno de investimento para cada métrica estudada

Em alguns projetos essa é uma medida de fácil obtenção, como em **e-commerces**, onde os valores das vendas podem ser utilizados como métrica, ou em intranets, onde os ganhos em produtividade são encontrados multiplicando o tempo economizado pelo custo hora. Para a maioria das aplicações, essa medida depende dos objetivos da companhia que mantém a aplicação. Tipicamente, quanto mais pessoas utilizam um projeto, melhor é seu retorno de investimento, uma vez que os benefícios surgem do valor agregado que a facilidade de uso traz a cada usuário.

Em um estudo posterior (NIELSEN, 2008), os ganhos médios diminuíram para 83%, apesar dos custos permanecerem quase constantes. Isso é explicado por fatores como o crescimento da importância de interfaces bem projetadas, que na época do primeiro estudo não eram tão comuns. O fato dos investimentos em usabilidade também não terem aumentado pode ser um fator determinante para esse resultado. Apesar desses resultados, a taxa de conversão (razão entre visitantes casuais que se convertem em visitantes que realizam ações desejadas em um software, como uma compra ou o envio de uma pesquisa de opinião) dobrou, o que é muito positivo pois é considerada a métrica mais importante de usabilidade para Web.

2.1.1.4 Projetos paralelos

Normalmente é interessante ter projetos paralelos de diferentes projetistas para projetos preliminares. A ideia é explorar diferentes alternativas antes de selecionar uma abordagem que seja desenvolvida em detalhes e com atividades de usabilidade mais detalhadas. É importante que os diferentes projetos sejam desenvolvidos separadamente, uma vez que diversidade é o objetivo dessa etapa. Com os diversos projetos em mãos é possível aproveitar as melhores ideias de cada um. Também é possível alocar diferentes projetistas para partes diferentes do projeto. Apesar de parecer um método ineficiente em termos de custos, projetos paralelos são muito baratos, pois a maioria das ideias não será implementada do modo que seriam se não fossem julgadas em uma etapa de projetos paralelos. Como várias abordagens de projeto são avaliadas simultaneamente, o tempo de desenvolvimento é reduzido e um produto pode ir para o mercado mais rapidamente (NIELSEN, 1998).

2.1.1.5 Projeto Participativo

Nessa etapa os projetistas devem ter acesso a uma quantidade de usuários representativos do sistema que realmente o utilizarão. Usuários tipicamente levantam questões que a equipe de desenvolvimento dificilmente imaginaria sozinha, especialmente com relação a potenciais inadequações entre o modelo de tarefa dos usuários e o imaginado pelos desenvolvedores. Mas para que os usuários sejam capazes de ajudar nessa etapa é importante apresentar as ideias de projeto de modo que eles possam compreender, como protótipos.

Como é comum que os usuários não saibam o que precisam ou desejam, não se deve limitar um projeto a realizar somente algumas questões. É importante mostrar a eles as possibilidades atuais, verificar sua utilidade e então perguntar aos usuários sua opinião. Em projetos muito longos, também pode ser interessante trocar o grupo de usuários periodicamente para não se correr o risco da diminuição da representatividade do grupo atual.

2.1.1.6 Coordenando o projeto total da interface

Consistência é uma das características de usabilidade mais importantes e deve ser aplicada nos meios que formam a interface total, incluindo a documentação, sistema de ajuda, tutoriais e treinamentos. A consistência é uma característica que é medida entre as diversas versões de um *software* e é fundamental que seja aplicada inclusive em famílias de produtos.

Por isso é importante que o controle do desenvolvimento do projeto seja centralizado, incentivando a utilização de padrões de interface para atingir a consistência. Em termos tecnológicos, a consistência pode ser atingida com código reutilizável e compartilhado.

2.1.1.7 Aplicação de recomendações e análise heurística

Recomendações listam princípios bem conhecidos de projeto de interfaces que devem ser seguidos no desenvolvimento do projeto. Em qualquer projeto, diversos níveis de recomendações podem ser utilizados:

- Recomendações gerais, aplicáveis a qualquer interface;
- Recomendações específicas de uma categoria, para um tipo de sistema em desenvolvimento;
- Recomendações específicas de um produto, para um produto individual.

Todas essas Recomendações podem ser utilizadas como base para avaliações heurísticas. A diferença entre padrões e recomendações é que um padrão especifica como a interface deve aparecer para o usuário, enquanto as recomendações fornecem características de usabilidade da interface. Normalmente, padrões seguem recomendações de usabilidade tradicionais, sendo assim seguir um padrão pode deixar a interface mais usável.

Consideradas um método informal, as avaliações heurísticas são baseadas em padrões de qualidade e construídas por meio de experiência prática e conhecimento teórico, sendo capazes de revelar diversos e variados problemas. Envolve especialistas julgando quando cada elemento do diálogo segue princípios de usabilidade estabelecidos ou heurísticas (NIELSEN, 1994b). Basicamente, observa-se a interface para se formular uma opinião sobre o que é bom e o que é ruim sobre ela, opinião que deve ser orientada por um documento guia (NIELSEN; MOLICH, 1990).

2.1.1.8 Prototipação

Não se deve iniciar a implementação de um projeto com base somente em projetos de interface primitivos. Ao invés disso, avaliação de usabilidade antecipada pode ser realizada baseada em protótipos do sistema final que poderá ser desenvolvido mais rapidamente, com custos menores e ser alterado quantas vezes forem necessárias até que se tenha um bom entendimento da interface com o usuário (NIELSEN, 1998). Além disso, recomenda-se que em algum ponto do projeto de interface um estágio seja dedicado a prototipação (COYETTE; KIEFFER; VANDERDONCKT, 2007).

A ideia da prototipação é diminuir custo e tempo de desenvolvimento e desenvolver algo que possa ser testado com usuários reais. Isso é atingido diminuindo pontos de implementação, o que pode ser realizado de dois modos. Diminuir características é chamado de prototipação vertical, que inclui funcionalidades completas, mas apenas algumas selecionadas de todo o sistema. A segunda é chamada de prototipação horizontal, que inclui uma camada com toda a interface, mas sem funções completas (NIELSEN, 1998).

Outra maneira de classificar o protótipo está relacionada a sua similaridade com o sistema final que será desenvolvido, chamada de classificação por fidelidade. A fidelidade é dita ser alta caso a representação do protótipo seja a mais próxima possível da interface final, em termos gráficos, de navegação e comunicação. Em contrapartida, a fidelidade é dita baixa quando o protótipo representa apenas parcialmente a interface final. Independente do tipo e material utilizado, todas as fidelidades são capazes de encontrar problemas de usabilidade (COYETTE; KIEFFER; VANDERDONCKT, 2007).

É possível reduzir tanto a quantidade de características quanto a de funcionalidades para chegar a um cenário onde só é possível simular a interface em um nível e caminho previamente planejado. Algumas práticas que ajudam a produzir protótipos mais rapidamente são:

- Direcionar menos esforços a eficiência da implementação;
- Aceitar código de menor qualidade;
- Utilizar algoritmos simplificados que não tratam todos os casos especiais;
- Utilizar um humano com experiência na manipulação de sistemas para superar algumas limitações
- Utilizar um computador diferente do alvo da plataforma;

- Utilizar interfaces de baixa fidelidade;
- Utilizar dados falsos e outros conteúdos;
- Utilizar protótipos de papel ao invés de sistemas funcionando;
- Utilizar uma interface totalmente imaginária onde o experimentador descreve a possível interface oralmente.

Protótipos podem ser utilizados em uma forma participativa de projeto chamada prototipação interativa, onde o protótipo é feito e alterado enquanto é utilizado, baseado em comentários dos usuários.

2.1.1.9 Avaliação da Interface

São duas as dimensões de classificação para os testes de interface: os que utilizam ou não usuários reais e os em que a interface real está implementada ou não (NIELSEN, 1998). Métodos de prototipação conseguem ser efetivos o bastante para influenciar um projeto e alterar seu direcionamento em seus estágios iniciais e heurísticas permitem avaliar a usabilidade sem originar gastos com um teste com usuários.

As avaliações são capazes de comprovar diagnósticos produzidos por avaliações heurísticas ou *checklists*. Elas são baseadas na observação de interações reais e são muito efetivas na identificação de problemas que se manifestam durante a interação dos usuários com o sistema. Através delas é possível obter métricas quantitativas de usabilidade, analisando a efetividade, eficácia, entre outras características (MORANDINI, 2003).

Independente do método de avaliação utilizado, os resultados esperados são uma lista de problemas de usabilidade com a interface assim como sugestões de como dar suporte as estratégias dos usuários. É comum não ser possível resolver todos os problemas, então eles devem ser priorizados com base em seu impacto no desempenho do usuário. Esse impacto é medido a partir de avaliações de especialistas de usabilidade isolados, o que nem sempre é uma medida confiável. A avaliação dos especialistas deve seguir um parâmetro único, que pode ser baseado na gravidade do problema ou na combinação de quantos usuários espera-se que venham a enfrentá-lo e o quanto esses usuários são prejudicados pelo problema. É possível chegar a esses valores observando quantos usuários enfrentam determinado problema com a interface, a quantidade de tempo necessária para contorná-lo e se possível comparar os resultados com um projeto paralelo que não apresente a mesma deficiência na interface sendo observada.

Também é possível avaliar uma interface através de *checklists*, uma técnica que compara características de usabilidade de um sistema com padrões qualitativos, onde profissionais com conhecimentos diferentes podem de forma satisfatória tomar parte sem necessidade de serem especialistas (MORANDINI, 2003). Os *checklists* são semelhantes a questionários de auto-avaliação usados pelos avaliadores, e por esse motivo são de fácil aplicação, possuem metodologia sistematizada e fornecem uma classificação final da interface, e são mais completos e detalhados que heurísticas (NIELSEN; MOLICH, 1990). Eles possuem nove critérios de avaliação (JOHNSON, 1996):

- Clareza visual
- Consistência
- Compatibilidade
- *Feedback* informativo
- Explicitude
- Funcionamento apropriado
- Flexibilidade e controle
- Prevenção e correção de erros
- Guia e suporte ao usuário

Apesar de não exigirem conhecimentos avançados para aplicação, devido a quantidade de perguntas realizadas sua aplicação é muitas vezes deixada de lado. Por ser um método superficial, é melhor utilizado como base de avaliações de usabilidade (MORANDINI, 2003).

Alguns aspectos de usabilidade são mais facilmente estudados simplesmente perguntando ao usuário (NIELSEN, 1998). Normalmente essas características são relativas a satisfação subjetiva do usuário e possíveis ansiedades, o que é difícil de mensurar objetivamente. Apesar disso, essas características são importantes em muitos sistemas, como os voltados para entretenimento, onde a diversão proporcionada é muito mais importante que a velocidade que uma ação é executada, por exemplo.

A avaliação de interface é um dos processos mais custosos da Engenharia de Usabilidade. Entretanto, é reconhecido que seus resultados são essenciais e não devem ser deixados de lado. Mesmo que não seja possível realizar os métodos com grande rigor e quantidade

de recursos, os benefícios de se utilizar métodos racionais de engenharia de usabilidade para avaliar a interface ao invés de apenas disponibilizar a aplicação sem avaliação são muito maiores que os benefícios incrementais de utilizar os métodos rigorosos para determinado projeto (NIELSEN, 1998).

2.1.1.10 Projeto Interativo

Baseado nos problemas de usabilidade descobertos e oportunidades trazidas pelo teste empírico, uma nova versão da interface pode ser produzida. Apesar disso, atenção é necessária ao propor novas soluções pois estas podem originar novos problemas de usabilidade. Também é possível que uma mudança para solucionar um problema relacionado a um parâmetro de usabilidade impacte negativamente em outro.

Durante o período de aceitação de um novo componente para interface do sistema é melhor realizar testes através de heurísticas, consultores e especialistas em usabilidade e até mesmo usuários especializados. Essa medida evita que os usuários realizem testes com componentes muito limitados, que devem ser testados integrados a todo o sistema.

As razões para uma decisão de projeto de interface devem ser armazenadas de modo que seja possível acessá-las posteriormente. Essas informações são importantes durante o projeto e desenvolvimento. Como mudanças ocorrerão em algum ponto é importante saber as razões da decisão inicial para aquele ponto para que essas não sejam sacrificadas por um objetivo de menor importância. Além disso essa documentação é capaz de auxiliar os escritores técnicos com a documentação e durante a tradução e é importante para se manter a consistência.

2.1.1.11 Coleta de feedback de uso de campo

O maior objetivo do trabalho de usabilidade após a publicação do produto é obter dados de usabilidade para as próximas versões e novos produtos. Estudos do uso natural do produto em seu campo mostram como usuários reais o utilizam trabalhando no ambiente real, o que pode fornecer muito mais dados do que os disponíveis em laboratório.

2.1.2 Alternativa aos elevados custos de testes de usabilidade

Os resultados das avaliações de usabilidade são de grande interesse pois são idealmente aplicados em usuários típicos da aplicação, num ambiente de utilização típico e

com tarefas típicas. Por outro lado, os testes de usabilidade são comumente tidos como uma técnica onerosa devido a (MORANDINI, 2003), (AU et al., 2008):

- dificuldade de obtenção de usuários;
- conotação negativa do teste para os usuários;
- custo (tempo e recursos).

Devido a sua efetividade, recomenda-se que o teste de usabilidade nunca seja deixado de lado, sendo fundamental pois mostra informações diretas de como os usuários utilizam o sistema além de ser importante para diversos profissionais de usabilidade (NIELSEN, 1998). Cada método auxilia na identificação de uma gama específica de problemas, mas não existe um método capaz de identificar todos os problemas possíveis, apesar da análise de interação do usuário ser um dos mais completos e considerado referência (WINCKLER; PIMENTA, 2002). Já (FERRE et al., 2001) acredita não ser possível prever a usabilidade de um sistema sem testes com usuários reais. Técnicas que priorizam algumas atividades e dispensam outras, tem como característica marcante a redução da necessidade de recursos e tempo. Uma delas é a *Discount Usability*¹ que reduz custos de usabilidade em até 50% e obtêm resultados tão expressivos quanto técnicas mais custosas (NIELSEN, 1998), (NIELSEN, 1994a). Apesar desses resultados, ela mantém uma das principais origens de gastos, o teste com usuários (MARTY; TWIDALE, 2005).

A literatura vem mostrando que é possível agilizar e diminuir os custos desse tipo de avaliação através de sua automatização (BRINCK; HOFER, 2002). Além disso a obtenção e análise preliminar dos dados é padronizada e mais ágil, além de não estar condicionada a visão de um avaliador humano (IVORY; HEARST, 2001). Para isso utiliza-se os dados produzidos pela interação dos usuários reais com o *website*, que ficam armazenados em seu servidor. Esse assunto será abordado na seção 2.2. As vantagens de avaliações de usabilidade automatizadas comparadas as não-automatizadas são (IVORY; HEARST, 2001), (BRINCK; HOFER, 2002):

- redução dos custos da avaliação através da automatização de etapas críticas, pois métodos que automatizam a captura de dados, análise e crítica diminuem o tempo necessário para a avaliação, e conseqüentemente seus custos;
- crescente consistência dos erros não cobertos, pois em alguns casos é possível definir um modelo de conclusão da tarefa que um software pode utilizar para buscar por desvios, sendo também possível detectar padrões que sugerem possíveis erros;

¹Ver seção 2.1.3

- predição de custos de tempo e erros em todo o projeto, pois avaliações automatizadas podem ser aplicadas em qualquer estágio do projeto, cobrindo todas características;
- redução da necessidade de avaliadores, pois um avaliador que não tem experiência em determinada característica da avaliação coberta pelo software não precisará ser auxiliado por outro avaliador;
- crescente cobertura de características atendidas, pois em avaliações não-automatizadas nem sempre é possível avaliar todos aspectos da interface devido a restrições de tempo, custo e recursos, o que não ocorre com avaliações automatizadas;
- permite comparações entre projetos alternativos, o que não ocorre sempre em avaliações não-automatizadas devido a restrições de custos.

Boa parte das ferramentas de análise automatizada de usabilidade não necessitam que seu usuário seja um especialista para identificar problemas de usabilidade (WINCKLER; PIMENTA, 2002). Esse cenário é verdadeiro mesmo quando é necessário analisar os resultados apresentados pela ferramenta de análise em busca de problemas mais complexos ou não explícitos.

As ferramentas para avaliação automatizada de usabilidade podem ser classificadas segundo as seguintes características (IVORY; HEARST, 2001), (BRINCK; HOFER, 2002):

- Classe do método: tipo e nível da avaliação conduzida;
- Tipo do método: como a avaliação é conduzida dentro da classe do método;
- Tipo de automação: descreve o aspecto da avaliação que é automatizado;
- Nível de esforço: descreve o tipo de esforço necessário para executar o método.

Os métodos são divididos em cinco classes:

- Baseados em teste: um avaliador observa usuários interagindo com a interface para determinar problemas de usabilidade;
- Inspeção: um avaliador não especialista utiliza um conjunto de critérios ou heurísticas para identificar problemas de usabilidade em potencial;
- Questionamento: usuários fornecem *feedback* em um interface através de entrevistas e pesquisas;

2.1 Usabilidade

- Modelagem analítica: um avaliador aplica modelos de usuário e interface para gerar previsões de usabilidade;
- Simulação: um avaliador aplica modelos de usuário e interface para imitar um usuário interagindo com a aplicação e reportar os resultados dessa interação;

Teste, inspeção e questionamento são abordagens apropriadas para propósitos formativos e somativos e possuem grande influência da engenharia de software. Já modelagem analítica e simulação são abordagens de engenharia que permitem ao avaliador prever níveis de usabilidade de um sistema.

Os tipos de automação são:

- Nenhum: nenhuma automação é realizada;
- Captura: software grava automaticamente dados de usabilidade;
- Análise: problemas de usabilidade em potencial são identificados automaticamente;
- Crítica: análise automatizada e sugestões de melhorias.

O esforço humano necessário para utilização de uma técnica automatizada de análise de usabilidade é classificado em:

- Esforço mínimo: não necessita de modelagem ou utilização de interface;
- Desenvolvimento de modelo: o avaliador precisa desenvolver um modelo de interface de usuário e/ou de usuário de modo a aplicar o método;
- Utilização informal: é necessário completar tarefas escolhidas abertamente;
- Utilização formal: é necessário completar tarefas selecionadas especificamente.

A partir das diferentes saídas possíveis de ferramentas automatizadas, é possível realizar os seguintes tipos de análises (BRINCK; HOFER, 2002):

- Assistência em projeto: ferramentas de usabilidade que podem ser utilizadas durante o início do projeto para auxiliar um projetista a criar candidatos que atendam as especificações do site e necessidades dos usuários;

- Coleta de dados: com avaliação direcionada a sintaxe, nenhum uso do site pode ser necessário para coletar informações sobre usabilidade, uma vez que sua utilização pode ser simulada através de modelos. Outras abordagens podem necessitar de envolvimento significativo dos usuários como o teste de usuários ou questionários;
- Visualização de dados: ferramentas de visualização ajudam um avaliador humano a compreender dados coletados e realizar inferências;
- Avaliação descritiva e preditiva: dados de usabilidade de sites podem ser apresentados em formato estatístico. Análises descritivas básicas ajudam o avaliador entender a estrutura do site e podem ajudar na avaliação manual. Análises preditivas podem indicar o quanto um site é usável;
- Identificação de problemas: técnicas de diagnóstico que identificam uma característica específica do site que poderá causar problemas de usabilidade;
- Recomendações: ferramentas capazes de avaliar soluções alternativas e recomendar uma que seja ótima;
- Recuperando o projeto: ferramenta automatizada que ajuda o desenvolvedor a reparar e a adaptar o código HTML.

2.1.3 A Discount Usability Engineering

Como nem sempre é possível realizar todas as etapas que compreendem a Engenharia de Usabilidade, é necessário selecionar as mais relevantes para determinado projeto. Uma abordagem conhecida, limitada somente a alguns aspectos da engenharia de usabilidade, é a *Discount Usability Engineering*. Ela procura se adaptar a limitações orçamentárias e de tempo, duas características tidas como chave na escolha de uma abordagem de ações relacionadas a usabilidade, e atinge bons níveis de resultados.

Outro ponto importante é o conhecimento necessário para utilizar a técnica. Por utilizar métodos mais simples que não exigem conhecimento aprofundado de usabilidade nem grande infra-estrutura, a técnica é de fácil aplicação e ultrapassa a barreira intimidante que muitas vezes impede que seja realizada qualquer atividade de usabilidade em um projeto (NIELSEN, 1998).

A *discount usability* se baseia em quatro técnicas:

- Observação do usuário e da tarefa;

2.1 Usabilidade

- Cenários;
- Pensando alto simplificado;
- Avaliação heurística.

O princípio básico de foco nos usuários é mantido e sempre deve ser seguido. Ele pode ser atingido de diversos modos, como visitas ao local em que realizam as tarefas, onde eles seriam observados desempenhando-as. É importante que o avaliador permaneça em silêncio e deixe os usuários trabalhar sem interferir.

Cenários são um modo de obter *feedback* rápido e constante dos usuários, pois simulam apenas uma parte bastante limitada de um sistema e determinam o caminho que o usuário precisa seguir. Por serem pequenos e limitados, é possível alterá-los frequentemente e em conjunto com testes pequenos de pensando alto, é possível testar todas as versões.

O método pensando alto simplificado envolve um usuário, que utiliza o sistema por um tempo determinado, para um dado número de tarefas enquanto é convidado a “pensar alto”. Ao verbalizar seus pensamentos, o usuário permite ao observador determinar não apenas o que ele está realizando com a interface, mas porque ele está fazendo isso. Essas informações podem ajudar a apontar elementos na interface que causam confusão. A análise dos dados é realizada com base em notas feitas pelo observador.

Já guias de usabilidade possuem muitas de regras a serem seguidas, o que é intimidante. A técnica de heurísticas (NIELSEN, 1998) reduz a complexidade para um conjunto pequeno de 10 regras que explicam em proporções largas problemas observados em projetos de interface. Pessoas com pouca experiência são capazes de encontrar problemas seguindo as heurísticas, e os problemas restantes podem ser revelados com o pensando alto simplificado. As 10 regras são apresentadas a seguir:

- Diálogo simples e natural, que não contenha informações irrelevantes ou raramente necessárias;
- Fale do mesmo modo que os usuários, expressando diálogos de forma clara e com conceitos familiares a eles;
- Minimize a carga de memória do usuário, não exija que ele tenha que se lembrar de uma informação de uma parte da interface em outro ponto;
- Consistência, pois é difícil para um usuário perceber que diferentes palavras, situações ou ações significam a mesma coisa;

2.1 Usabilidade

- *Feedback* para os usuários de forma a mantê-los informados do que acontece no sistema;
- Saídas bem marcadas, pois é comum que os usuários escolham ações por engano e queiram uma saída de emergência ao estado indesejado da aplicação;
- Atalhos podem acelerar a interação para o usuário avançado, fazendo com que o sistema atenda tanto usuários com e sem experiência;
- Boas mensagens de erro, que devem ser indicadas na linguagem dos usuários, sem códigos de linguagens de programação, indicando claramente o problema e sugerindo uma solução;
- Previna erros com um projeto cuidadoso;
- Ajuda e documentação devem ser fornecidos, com conteúdo fácil de buscar, focado na tarefa do usuário, listando passos concretos para se tomar e não ser muito grande.

Por ser uma abordagem que busca otimizar a utilização de recursos, a *discount usability* além de ser de fácil aplicação pode gerar resultados com custo benefício maior que de técnicas mais elaboradas (NIELSEN, 1994a), como mostra a figura 2.2.

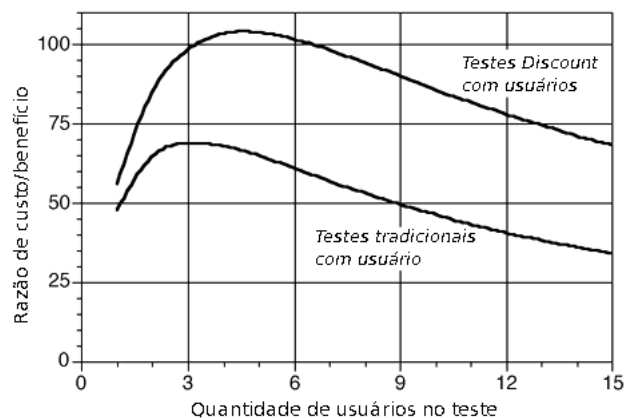


Figura 2.2: Curva de custo-benefício de um projeto típico adaptado de (NIELSEN, 1994a)

Resultados de estudos mostram que a melhoria alcançada mesmo com pessoas com pouco conhecimento de usabilidade pode variar entre 50 e 76% (NIELSEN, 1994a). Também é comum que os resultados obtidos ao se aplicar a *Discount Usability* sejam melhores que ao aplicar técnicas mais elaboradas, pois seus métodos possuem ênfase em iterações antecipadas e rápidas com frequente entrada de usabilidade (NIELSEN, 2009).

Um estudo do custo benefício da técnica de *Discount Usability* (NIELSEN, 1994a) com um protótipo de interface para um sistema interno de uma companhia de telefone

onde é necessário conhecimento específico do domínio para operação foi conduzido. Nele foram encontrados 44 problemas com 11 avaliadores, sendo que 40 destes eram chave, através da avaliação heurística. Aos avaliadores também foi solicitada uma estimativa de melhorias após solução dos problemas encontrados. As estimativas foram que a curva de aprendizado seria reduzida em 0.5 dias e o usuário avançado teria um ganho de 10%. Desse modo, 14.000 dias de usuário no primeiro ano de utilização para 2000 usuários realizando 3.3% mais rapidamente seriam economizados. Em termos financeiros, se cada usuário custar \$100, somente metade dos problemas puderem ser resolvidos, 20% do valor da melhoria na curva de aprendizado for descartado e 30% da melhoria na utilização for ignorada, os ganhos em um ano são de \$540.000. Assumindo que foram necessárias 400 horas para aplicar as melhorias no software com custo de \$100, os ganhos são reduzidos para \$500.000. O custo do projeto de usabilidade foi de \$10.500, o que resulta em um benefício 48 vezes maior que o custo.

Após a análise heurística, um estudo com 4 usuários reais foi realizado, confirmando 17 dos problemas encontrados anteriormente e 4 novos. O custo total desses testes foi de \$2.250. Uma análise de gravidade dos problemas mostrou que a gravidade dos problemas encontrados na avaliação heurística foram de 98.41 pontos. Desse modo, os 17 problemas encontrados em ambas avaliações tiveram uma gravidade de 41.56. Logo, uma estimativa da gravidade dos 4 novos problemas ($4/17 \times 41.56/98.41$) é de 0.99. Assumindo que as melhorias ao resolver esses problemas são as mesmas de resolver os anteriores, os ganhos chegam a $\$500.000 \times 0.99 = \49.500 , o que resulta em um benefício 22 vezes maior que os custos.

2.1.3.1 Avaliação Heurística

Uma avaliação heurística é um método de inspeção sistemática pela usabilidade de uma interface de usuário cujo objetivo é encontrar problemas em seu projeto (NIELSEN, 1992). Uma avaliação envolve um conjunto pequeno de avaliadores examinando a interface e julgando o cumprimento de princípios de usabilidade reconhecidos, ou heurísticas. Ela é realizada verificando a interface e tentando formular uma opinião sobre o que é bom e ruim sobre ela. É ideal que essas avaliações sejam conduzidas sob um conjunto de regras, como as encontradas nos típicos documentos guia.

Apesar de em geral recomendar-se a utilização de avaliadores experientes, as heurísticas são um método de avaliação que permitem pessoas com poucos conhecimentos sobre usabilidade possam executá-la, apesar desses normalmente não serem capazes de perceber muitos problemas de usabilidade (NIELSEN, 1992). Como diferentes avaliadores encontram

2.1 Usabilidade

problemas diferentes, recomenda-se utilizar mais de um avaliador quando esses não tiverem muita experiência. Um número recomendado está entre três e cinco avaliadores. A figura 2.3 mostra uma relação entre a quantidade média de problemas encontrados pela de avaliadores (NIELSEN, a).

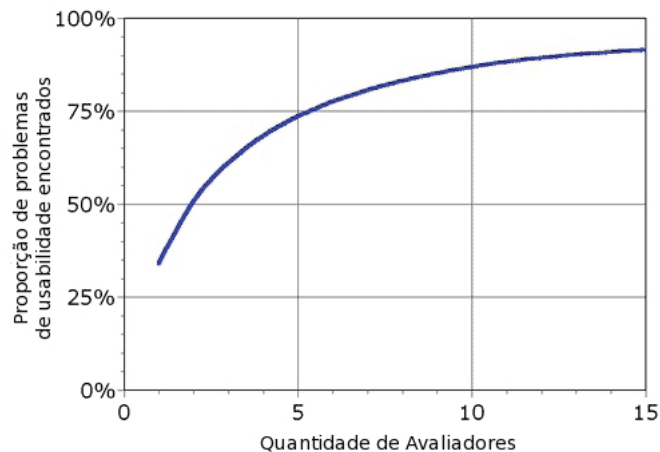


Figura 2.3: Média de problemas encontrados por quantidade de avaliadores em avaliações heurísticas adaptado de (NIELSEN, a)

Os avaliadores devem analisar a interface individualmente e somente após todos eles completarem a avaliação poderão se comunicar e trocar suas opiniões. Essa medida previne que suas opiniões sejam influenciadas. Os resultados encontrados podem ser escritos pelo próprio avaliador ou por um observador presente durante a avaliação, que transcreve as opiniões expressas pelo avaliador. O observador também pode assistir o avaliador a operar a interface no caso de problemas, mas não deve interpretar as ações realizadas por ele na interface, somente seus comentários sobre ela, uma vez que o responsável pela avaliação é o avaliador. Em avaliações de aplicações de domínios específicos deve-se responder as questões do avaliador, quando estes não são especialistas naquele domínio, o que provavelmente irá aumentar a efetividade da avaliação pois permite a ele interpretar a interface com respeito as características daquele domínio. Apesar de ser interessante fornecer dicas aos avaliadores, o momento correto de realizar é quando eles realmente estiverem com problemas e já tiverem comentado o problema de usabilidade em questão.

Em sistemas de utilização generalizada normalmente não é necessário fornecer auxílio aos avaliadores. No caso de sistemas para domínios específicos, em que os avaliadores não possuem conhecimento sobre ele, é preciso auxílio durante a avaliação. Uma abordagem de sucesso é fornecer cenários típicos de utilização aos avaliadores, listando os passos que um usuário deve tomar para realizar algumas tarefas reais.

Durante a avaliação, o avaliador interage com a interface várias vezes inspecionando os elementos de diálogo e comparando-os com uma lista de princípios de usabilidade conhecidos. Essas heurísticas são regras gerais que descrevem propriedades comuns de interfaces usáveis e são uma das grandes vantagens da avaliação heurística, pois são especificações escritas e bem definidas (NIELSEN, 1992). Os avaliadores podem decidir como preferem avaliar a interface, mas recomenda-se que eles passem por ela pelo menos duas vezes. A primeira para ter noção do fluxo da interação e escopo geral do sistema. A segunda para focar em elementos específicos de diálogo.

Uma avaliação heurística fornece como resultado uma lista de problemas de usabilidade, com referências aos princípios violados pelo projeto em cada caso, na opinião do avaliador. Mesmo não oferecendo sugestões de correção, por oferecer referências de usabilidade estabelecidas aos problemas, é simples gerar uma interface revisada que esteja de acordo com as referências violadas (NIELSEN, 1998).

A severidade das falhas encontradas na interface pode ser avaliada após a união de todas as falhas encontradas pelos diferentes avaliadores. Uma vez que cada um encontrará um conjunto diferente dos demais, e fornecerá uma descrição diferente para o problema, eles deverão ser compilados em um documento único enviado a todos os avaliadores. Igualmente a fase de avaliação, a severidade das falhas deve ser dada individualmente e sem considerar as realizadas por outros avaliadores. Um ponto importante é, caso seja possível, fornecer acesso ao sistema para os avaliadores, fazendo com que eles não dependam somente de suas memórias e da descrição do problema para lembrar da interface.

A seguinte métrica pode ser utilizada para indicar a severidade de um problema:

- 0 Não concordo, não é um problema de usabilidade;
- 1 Apenas um problema cosmético, precisa ser corrigido apenas se houver tempo extra no projeto;
- 2 Pequeno problema de usabilidade, baixa prioridade para correção;
- 3 Grande problema de usabilidade, importante corrigir e deve ser dada alta prioridade;
- 4 Catástrofe de usabilidade, imperativo corrigir antes do produto ser publicado.

Uma média satisfatória é obtida com quantidade semelhante a de avaliadores para a avaliação heurística, sendo três um número aceitável (NIELSEN, b).

A avaliação heurística é um método que faz parte da técnica de *Discount Usability*, apesar de pesquisas independentes confirmarem que é um método eficiente de avaliação de usabilidade isolado (NIELSEN, 1992), (NIELSEN, 1998).

2.1.3.2 Pensando alto simplificado

Pensando alto simplificado é um método de avaliação de usabilidade tido como um dos testes mais valiosos dentro da engenharia de usabilidade (HOLZINGER, 2005). O usuário, através da verbalização dos seus pensamentos enquanto utiliza a interface, permite ao avaliador compreender as estratégias do usuário e os conceitos mal interpretados pelos mesmos. Uma das grandes vantagens do método é que ele permite conhecer o porquê um usuário realiza uma ação. O método pensando alto tradicional é conduzido com psicólogos, especialistas em interface, possui seu áudio e vídeo gravados e realiza uma análise detalhada de protocolo. Uma das maiores diferenças entre o método tradicional e o simplificado é que este não necessita de filmagens, podendo as análises ser realizadas somente com base em notas feitas pelo avaliador (NIELSEN, 1998).

Sua principal desvantagem é que ele não é um método bem adequado a análises de desempenho, pelo contrário, sua vantagem é a riqueza dos dados qualitativos coletados de um número pequeno de usuários.

Uma das questões em torno do método é a importância dada a opinião dos usuários. Não se deve levar em conta todas as observações feitas por eles, durante o teste, o avaliador deve tomar nota do que os usuários estão fazendo durante a etapa do experimento em que encontrarem dificuldades. Dados de onde os usuários olhavam tem maior validade do que pedidos dos usuários por determinadas características da interface. Além disso é responsabilidade do avaliador interpretar os comentários do usuário, e não apenas aceitá-los.

Como pensar alto também não é um comportamento natural para maioria das pessoas, alguns usuários tem muita dificuldade durante o teste. Devido a isso ele pode ser mais difícil de conduzir e ter impacto nos resultados. O avaliador deve continuamente perguntar ao usuário o que ele pensa com questões como: “O que você está pensando agora?” ou “O que você acha que essa mensagem significa?”, mas somente após verificar que o usuário percebeu a mensagem e está pensando sobre ela. Se o usuário perguntar se pode realizar determinada ação no sistema, o avaliador não deve responder, mas manter o usuário falando com uma pergunta como “O que você acha que vai acontecer se fizer isso?”. Se o usuário parecer surpreso ao realizar uma ação, o avaliador pode perguntar “Era isso que você esperava que acontecesse?”. Por não ser comum ao usuário pensar alto, uma prática que ajuda nos

testes é uma prévia de como deverão se comportar antes de iniciar o testes, feita pelo avaliador (NIELSEN, 1998).

2.1.3.3 Protótipos

Protótipo é um modelo funcional criado para testar ideias de projeto. Podem ser utilizados para examinar conteúdo, estética e técnicas de interação da perspectiva de projetistas, clientes e usuários.

Protótipos muito similares ao produto final são chamados de protótipos de alta fidelidade, enquanto os menos similares são chamados de protótipos de baixa fidelidade. É comum que sejam empregados os mesmos métodos, técnicas de interação e aparência do produto final em um protótipo de alta fidelidade, e por isso ele necessita de mais recursos que um protótipo de baixa fidelidade. Recomenda-se utilizar protótipos de alta fidelidade caso os clientes avaliem os de baixa fidelidade como não profissionais (WALKER; TAKAYAMA; LANDAY, 2002) e também para refinar as ideias entre múltiplas interações (NIELSEN, 2003).

Vantagens de se utilizar protótipos de alta fidelidade são proporcionar interações mais realistas e melhor condução das possibilidades de projeto. Por outro lado, ela pode fazer com que os projetistas tornem-se relutantes a explorar e alterar todo o espaço do projeto (WALKER; TAKAYAMA; LANDAY, 2002). Ambos níveis de fidelidade são capazes de encontrar problemas (COYETTE; KIEFFER; VANDERDONCKT, 2007), não havendo diferenças entre a quantidade de problemas encontradas pelos dois tipos de protótipos, concluindo-se que a estética não influencia nas percepções de usabilidade (WALKER; TAKAYAMA; LANDAY, 2002).

Cenários são um protótipo minimalista, simples e de baixo custo de produção em que se descreve apenas uma interação de sessão única sem qualquer flexibilidade para o usuário e por isso não são realistas. Eles combinam as duas principais características de protótipos, reduzindo a quantidade de funcionalidades e características, desse modo não é possível interagir com dados reais nem se movimentar pelo sistema com total liberdade (NIELSEN, 1998). Já (CARROLL, 2000) define cenários como histórias sobre pessoas e suas atividades que destacam objetivos sugeridos pela aparência e comportamento do sistema, o que as pessoas tentam fazer no sistema, quais procedimentos são adotados, não adotados e executados com ou sem sucesso e por fim, quais as interpretações que as pessoas fazem do que acontece com elas. Cenários são o tipo de protótipo utilizados na abordagem *Discount Usability*.

Um cenário é uma descrição encapsulada de um usuário individual, utilizando um conjunto específico de funções, para atingir uma saída específica, sobre certas circunstân-

cias e em um intervalo de tempo determinado. Eles podem ser utilizados em dois pontos do projeto. Durante o início de forma a expressar e compreender como os usuários irão interagir com o sistema futuro. Também podem ser utilizados como forma de avaliar precocemente a interface para ter *feedback* dos usuários sem que seja necessário construir um protótipo (NIELSEN, 1998).

2.1.3.4 Teste de usabilidade

Teste com usuários reais é o método de avaliação de usabilidade mais fundamental e insubstituível, pois fornece informações sobre como as pessoas utilizam o software e seus problemas exatos enquanto utilizam a interface, é importante para diversos profissionais de usabilidade (NIELSEN, 1998). Também é considerado um dos métodos mais completos e serve de referência para outros (WINCKLER; PIMENTA, 2002). Por essas características acredita-se que somente com testes de usabilidade seja possível prever a usabilidade de um sistema (FERRE et al., 2001).

Uma das características mais marcantes do teste de usabilidade é seu rigor metodológico, de modo que características como confiabilidade, que indica se é possível obter os mesmos resultados caso seja necessário repetir o teste, e validade, que indica se os resultados refletem as questões de usabilidade a serem testadas, não comprometam seus resultados.

O teste de usabilidade pode ser realizado utilizando diversas técnicas, inclusive pensando alto simplificado e protótipos, o que é recomendado na abordagem *Discount Usability*. A figura 2.4 mostra como a aplicação das técnicas é agrupada quando se decide seguir esse padrão:

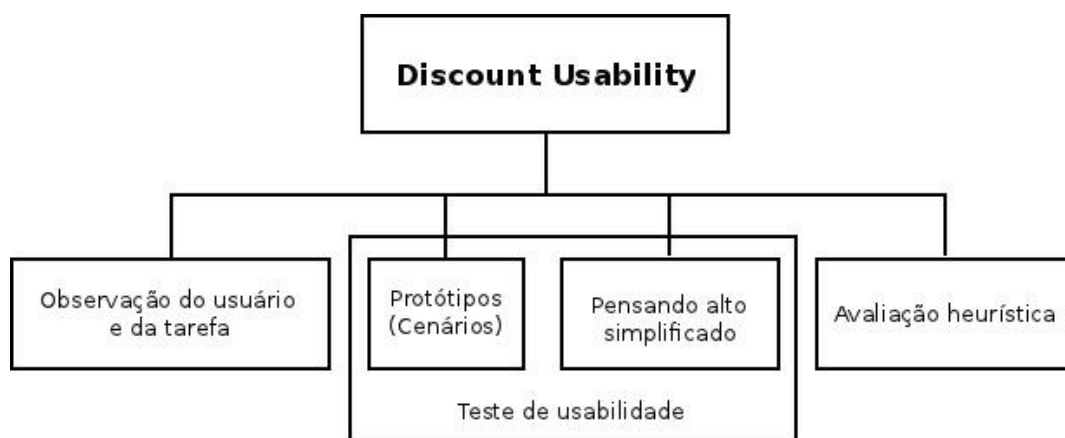


Figura 2.4: Técnicas da *Discount Usability* agrupadas para utilização em teste de usabilidade

2.2 Web Mining

Web mining é a aplicação de técnicas de *data mining* em amplos repositórios de informação na Web (MOBASHER et al., 1996). Ela é uma tarefa que pode ser dividida em quatro etapas:

1. Encontrar os recursos: basicamente buscar pelos recursos que fomentarão a mineração de conhecimento;
2. Seleção da informação e pré-processamento: eliminação de dados e itens irrelevantes que não possuem nenhuma importância para a análise de dados da Web. As únicas informações importantes são aquelas que representam acessos dos usuários ao Web site (COOLEY; MOBASHER; SRIVASTAVA, 1997);
3. Descoberta de conhecimento: nessa etapa diversos tipos de mineradores de padrões de acesso podem ser empregados dependendo das necessidades do analista, como padrões sequenciais, análise de caminhos, descoberta de regras de associação, clusterização e classificação (COOLEY; MOBASHER; SRIVASTAVA, 1997);
4. Análise: conjunto de técnicas e ferramentas que permitem ao analista utilizar os padrões encontrados na etapa anterior. Entre elas temos estatística, gráficos e visualizações e análise de usabilidade (COOLEY; MOBASHER; SRIVASTAVA, 1997).

Em cada uma das etapas podem ser utilizadas diferentes técnicas e com objetivos variados.

Web mining pode ser focada sob três diferentes abordagens: **mineração de conteúdo**, **mineração estrutural** e **mineração de utilização** (COOLEY; MOBASHER; SRIVASTAVA, 1997) (KOSALA; BLOCKEEL, 2000), mas não são abordagens com fronteiras bem definidas e é possível que uma abordagem utilize recursos de outras. Segundo (SRIVASTAVA; DESIKAN; KUMAR, 2005), essas delimitações possuem taxonomia como observado na figura 2.5.

Mineração de conteúdo (TIEDTKE; MÄRTIN; GERTH, 2002) (SPILIOPOULOU, 2000) descobre informações úteis de documentos, dados e recursos da Web. Possui grande utilidade no desenvolvimento de ferramentas inteligentes para recuperação de informação, mas pode sofrer com o problema do conteúdo englobar um amplo conjunto de dados. Utiliza como base principalmente técnicas de *text mining*, a descoberta de conhecimento em texto. Na Web essa abordagem pode ter dois focos: melhoria ou assistência para encontrar

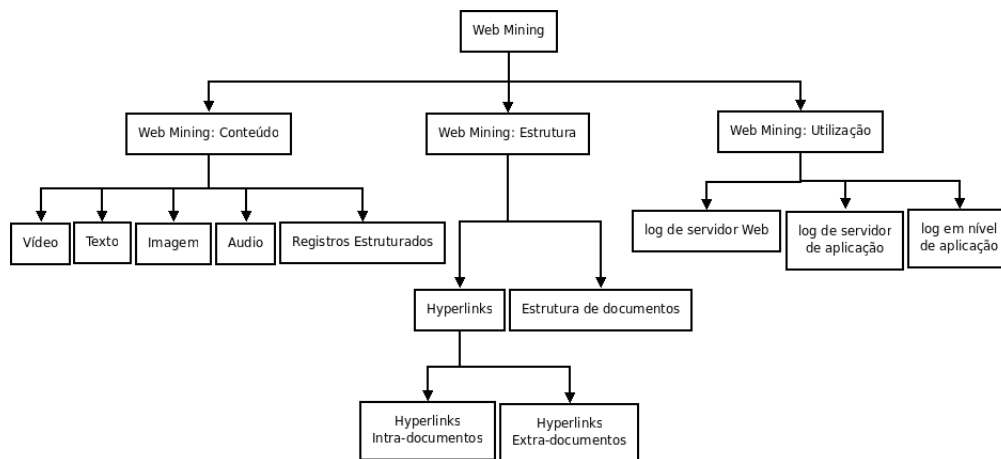


Figura 2.5: Taxonomia de *Web Mining* adaptado de (SRIVASTAVA; DESIKAN; KUMAR, 2005)

ou filtrar informação aos usuários com base em perfis de usuários informados ou inferidos pelo sistema de acordo com o comportamento; ou modelar os dados da Web e integrá-los de modo que buscas mais sofisticadas do que somente com palavras-chave possam ser realizadas (KOSALA; BLOCKEEL, 2000).

A mineração de estrutura (LI; KIT, 2005) realiza análises automatizadas relacionadas à estrutura da aplicação Web e aos dados gerados durante a utilização dos usuários da aplicação. As análises estruturais encontram erros relacionados ao funcionamento da aplicação e a maneira como ela está estruturada pelos arquitetos de informação, além de categorizar páginas Web e gerar informações como a similaridade e relacionamento entre diferentes Web sites. Normalmente é embasada na topologia dos *links* existentes na aplicação (KOSALA; BLOCKEEL, 2000).

Já a **mineração de utilização** (SCHWERZ; MORANDINI; SILVA, 2007) baseia-se em interações reais dos usuários com o *Website* comumente gravadas em arquivos de *log* presentes em várias camadas da arquitetura Web. São considerados dados secundários pois são derivados das interações dos usuários com a aplicação (KOSALA; BLOCKEEL, 2000). Uma vez que as interações que alimentam a análise são reais, os problemas encontrados refletem exatamente as deficiências da aplicação para o período analisado, algo que pode sofrer desvios em abordagens tradicionais. Além disso, arquivos de *log* com interações do usuário ajudam a mostrar onde o usuário paralisou sua tarefa, onde ele perdeu tempo e quais erros são encontrados mais frequentemente (NIELSEN, 1998). Em linhas gerais, observa-se que os dados dos arquivos de *log* do servidor armazenam dados que servem como fonte para análises de usabilidade tradicionais, onde esses dados são obtidos através de outros meios como os *checklists*, análises heurísticas, observações de interações dos usuários por especialistas, entre outras. As análises, normalmente são realizadas em laboratórios, com ambientes preparados para o

teste. Entretanto, os usuários devem estar totalmente dispostos a colaborar e expressar suas opiniões com relação as interações com o sistema testado, o que nem sempre ocorre, gerando resultados que não refletem a realidade. Como a coleta de dados para análise automatizada é realizada de forma imperceptível durante a interação do usuário com a aplicação, esse tipo de comportamento não ocorre (MORANDINI, 2003).

Além disso, os arquivos de *log* já são armazenados por padrão na maioria dos servidores Web, estando disponíveis sem que seja necessária nenhuma alteração em configurações do sistema ou da aplicação Web. Por esse motivo, não será necessário buscar por usuários aptos a realizar o teste.

Os arquivos de *log* são analisados em busca de padrões comportamentais dos usuários dentro da aplicação Web. Os padrões são buscados através dos caminhos de navegação dos usuários dentro do *website*. Esses padrões serão organizados em *clusters* com características similares e comparados com comportamentos padrão definidos pela equipe que administra a aplicação que está passando pela análise. Através da análise e comparação dos padrões encontrados com os padrões esperados definidos pelo administrador da aplicação, espera-se obter informações sobre o estado da usabilidade da aplicação (CATLEDGE; PITKOW, 1995).

2.2.1 Importância das técnicas de clusterização em Web Mining

Mesmo que o administrador de uma aplicação conheça muito bem os dados gerados por ela e seus usuários, em muitos casos a quantidade de dados disponível é muito extensa para ser analisada manualmente. Separar um conjunto finito de dados não rotulados em um conjunto finito e discreto de estruturas naturais e escondidas é uma abordagem que diminui a quantidade de dados a serem analisados, facilitando assim a análise. Essa tarefa é chamada de classificação não-supervisionada ou clusterização (XU; WUNSCH, 2005).

Clusterização é muito útil quando se deseja explorar naturezas desconhecidas dos dados que estão integradas com poucas ou nenhuma informação (XU; WUNSCH; SOCIETY, 2009). A técnica é amplamente empregada em pesquisa científica, e em ciência da computação ela é muito difundida no campo de *Web Mining* (XU; WUNSCH, 2005), (COOLEY; MOBASHER; SRIVASTAVA, 1997). Neste trabalho, os algoritmos de clusterização aplicam-se a dados sequenciais, entretanto, os algoritmos de clusterização podem ser aplicados nas três categorias de *Web Mining*. Alguns exemplos de algoritmos de clusterização encontrados na literatura são: Método de Ward, BIRCH, *K-means*, *K-means* genético, métodos baseados em *Fuzzy*, Redes Neurais e Algoritmos Genéticos (KOSALA; BLOCKEEL, 2000), (KOLARI; JOSHI,

2004), (XU; WUNSCH, 2005). Os algoritmos de clusterização também podem ser aplicados a grandes conjuntos de dados de alta densidade, situação comum na área de *Web Mining*. Pesquisas relacionadas a escalabilidade estão gerando novos métodos propostos para lidar com essa questão (XU; WUNSCH, 2005).

Clusterização é de grande importância para *Web Mining* devido a natureza dos dados disponíveis para mineração. Além das questões relacionadas a dimensão e densidade, é comum que eles sejam conhecidos superficialmente. No caso de análise de dados de navegação, apesar de conhecer o conteúdo de um arquivo de *log*, não se sabe exatamente o conhecimento possível de ser adquirido a partir dele. Neste trabalho, a partir dos dados contidos nos *logs* é possível classificar usuários através de seu comportamento de navegação. Outra possibilidade é a predição de requisições de usuários enquanto ele utiliza a aplicação (KOSALA; BLOCKEEL, 2000). Já em mineração de conteúdo, apesar de conhecer o conteúdo de páginas únicas, pode ser interessante agrupá-las para oferecer páginas semelhantes a usuários de uma aplicação Web (KOLARI; JOSHI, 2004).

Sendo assim, por se adaptarem a área de estudo, inclusive com flexibilidade para seus subcampos, atenderem a necessidades oriundas da natureza dos dados analisados e por oferecer resultados com possibilidades de aplicação diversas, algoritmos de clusterização são uma ferramenta importante para o campo de *Web Mining*.

2.3 Trabalhos Relacionados

Nesta seção serão apresentados os principais trabalhos que tratam da área de automatização da avaliação de usabilidade assim como trabalhos relacionados ao uso de diferentes técnicas de clusterização utilizadas em *Web Mining*. O detalhamento das técnicas utilizadas neste trabalho será apresentado na seção 2.4.

2.3.1 Trabalhos relacionados a avaliação de usabilidade automatizada

As abordagens de avaliação automatizada de usabilidade podem ser várias, como programas que verificam a sintaxe, assessores de projeto de páginas e analisadores de usabilidade (IVORY; HEARST, 2001) (TIEDTKE; MÄRTIN; GERTH, 2002). Além disso, avaliação automatizada permite melhorar a cobertura da usabilidade, cada vez mais necessária, diminuindo custos em termos de tempo e recursos humanos (IVORY; HEARST, 2001). O trabalho de (IVORY; HEARST, 2001) realiza uma extensa revisão de ferramentas já existentes além de forne-

cer informações genéricas como classificação dos algoritmos por automação, etapas realizadas, entre outras já abordadas neste trabalho.

O AWUSA (TIEDTKE; MÄRTIN; GERTH, 2002) é um sistema de análise de usabilidade que busca cobrir o aspecto da usabilidade em conjunto com o aspecto de marketing, utilizando uma abordagem baseada em tarefas e integrando questões de usabilidade com tópicos de arquitetura de informação do *website*. Para a análise dos dados gerados pelos usuários, *data mining* é aplicado em duas fases: análise estática, em tempo de projeto e definição das páginas e análise dinâmica, realizada durante a própria análise dos dados dos usuários. O trabalho utiliza *Web Mining* de utilização como recurso de análise dos dados. Desse modo, o AWUSA combina em um sistema análise de arquivo de *log*, avaliação de usabilidade automatizada e *web mining*, utilizando XML e Java como plataformas de desenvolvimento do sistema. Os objetivos da ferramenta são: encontrar e visualizar os caminhos dos usuários pelo *website*, encontrar e visualizar desvios entre as tarefas pretendidas e o uso atual, encontrar e visualizar localização e eventos onde as tarefas são abortadas, encontrar e visualizar áreas com usabilidade baixa, fornecer causas plausíveis para problemas de usabilidade detectados, encontrar novos objetivos e tarefas dos usuários enquanto estão no *website* e classificação de diferentes grupos de usuários com mapeamento em vários objetivos e grupos. Durante o projeto, as páginas estáticas são analisadas, o que gera métricas para cada uma delas. Nessa etapa, o *Document Object Model* e a representação visual das páginas são as principais áreas de análise, utilizando *Web Mining* de conteúdo e estrutura. Nessa etapa é detectada a conformidade com padrões de codificação, erros de sintaxe e compatibilidade com navegadores, contagem de elementos de conteúdo e análise de posicionamento de dados, cores e tamanho. Depois o arquivo de *log* é pré-processado, então sessões e seus caminhos são gerados. Em seguida os dados de utilização são comparados com o modelo definido da tarefa e as técnicas de *data mining* são aplicadas. Na análise dinâmica a mineração de utilização é aplicada para mineração de caminhos, geração de *clusters* e detecção de regras de associação. Uma das características mais interessantes do AWUSA é trabalhar com diferentes fontes de recurso, que vão desde os arquivos do *website* no servidor, passam pela estrutura dos *links* na navegação, pela definição das tarefas ou objetivos dos usuários dentro da aplicação analisada e pelos diferentes grupos alvo de determinada tarefa. Além de realizar a detecção de problemas de usabilidade, o AWUSA é capaz de classificá-los em cinco categorias: interação, informação, arquitetura, projeto e técnico. Os problemas são detectados de duas maneiras, analisando as páginas diretamente ou analisando o comportamento dos usuários e comparando com padrões de utilização conhecidos por seus problemas de usabilidade inerentes. O AWUSA integra os dados de todas as minerações realizadas para estabelecer relações entre a estrutura estática e

conteúdo do *website* com sua utilização, além de realizar medidas estatísticas sobre os dados extraídos. Problemas de usabilidade são encontrados através da análise de atributos que podem ser traduzidos em possíveis problemas ou comparando padrões detectados com sintomas considerados problemas de usabilidade. Por fim o AWUSA analisa as tarefas finalizadas com sucesso ou não e clusterização de usuários. Todos os resultados da aplicação são armazenados em formato XML que ficarão disponíveis para visualização baseada em camada de modelos que integra todos os dados contidos no XML. Por fim, os autores concluem que a melhor abordagem de usabilidade é a integração de técnicas tradicionais com automatizadas, pois beneficia-se com as vantagens de ambas. Com relação ao trabalho é possível concluir que o AWUSA é um *framework* de análise automatizada de usabilidade bastante abrangente e pela quantidade de conhecimento agregado é uma ferramenta robusta.

Já o trabalho (LI; KIT, 2005) observa que a mineração estrutural não foi bem explorada para análise automatizada de usabilidade, e apresentam um novo uso do modelo de grafo de *hyperlinks* para descrever um projeto de um algoritmo de mineração de conjuntos de itens espaciais frequentes (NS) de um *website*. É possível modelar a estrutura de *links* de um *website* através de grafos. Os autores sugerem um novo algoritmo chamado *Adaptive Window Algorithm* para descobrir a estrutura navegacional de um *website*, de onde é formada uma matriz esparsa. Uma das vantagens do algoritmo é que ele é capaz de descobrir estruturas de repetição entre as páginas, como menus, diminuindo a sua complexidade com relação a outros algoritmos em *websites* muito densos. Um experimento mostrou que a proposta dos autores é melhor que o algoritmo Eclat, que possui a mesma função. A análise de usabilidade é realizada com base em tarefas definidas pelo administrador do *website* através da definição de páginas alvo ou localizações. Essas localizações determinam se o usuário chegou à página alvo antes de finalizar a tarefa. As atividades dos usuários são interceptadas por um servidor *proxy*, filtradas e armazenadas em banco de dados. Os resultados experimentais exibiram resultados como quantidade de cliques, tempo médio e taxa de erros em três diferentes aplicações. Paralelamente foi solicitado aos usuários que respondessem a um questionário qualitativo, cujos resultados coincidiram com os resultados da análise automatizada.

O trabalho de (KORT; POOT, 2005) aborda rapidamente a integração da análise de *log* de eventos com avaliações qualitativas. Visto a crescente complexidade das aplicações os autores acreditam que *log* de eventos é capaz de avaliar aplicações desse tipo. Os modelos existentes de *log* de eventos focam na avaliação de usabilidade baseada em regras, como *checklist*, ou modelos. Uma das conclusões que os autores chegam é que somente o *log* de eventos não é capaz de fornecer dados melhores que os de um teste de usuário, mas é capaz de ajudar na sua compreensão. Por isso, eles sugerem que as abordagens sejam utilizadas em

conjunto. Um piloto de aplicação complexa contendo as seguintes características foi avaliado: utilizada em múltiplos contextos e em diferentes dispositivos móveis, alvo de múltiplos grupos e com interface dinâmica. Além do *log* de eventos, foi utilizado um *log* de alterações na interface, combinado com dados qualitativos obtidos através de questionários e entrevistas. Os resultados da avaliação com as abordagens combinadas do piloto do sistema permitiu que o projeto completo da aplicação fosse avaliado sem utilizar muito tempo e recursos.

O trabalho de (BEIREKDAR et al., 2005) realiza uma avaliação automatizada da usabilidade e acessibilidade de *websites* verificando se o código HTML atende recomendações formalmente expressadas através de uma linguagem baseada em XML chamada *Guideline Definition Language* (GDL). O fato da linguagem que expressa as recomendações ser separada da lógica de avaliação permite manter diferentes recomendações sem que o código do programa seja afetado, o que é considerado uma deficiência de outras abordagens automatizadas. O programa implementado deve ser capaz de aplicar qualquer tipo de guia de recomendação desejado. Apesar da existência de importantes guias de recomendações sobre acessibilidade e usabilidade, a maior parte deles acaba não sendo utilizada em projetos, pois necessita de grande esforço para ser aplicada. Nesse aspecto, as ferramentas automatizadas são grandes aliadas pois economizam recursos e principalmente tempo. Um dos objetivos dos autores é mostrar que os resultados desse tipo de ferramenta também podem ser parametrizados para produzir relatórios destinados a diferentes tipos de usuários, um genérico e um projetista. Assim como em outras abordagens, a automação da avaliação é total, variando conforme as recomendações sendo avaliadas. Uma das vantagens do programa desenvolvido é que ele é capaz de mostrar a página avaliada nos resultados, facilitando a interpretação dos resultados.

O trabalho de (AU et al., 2008) trata da avaliação automatizada de sistemas para *handhelds*, utilizados como solução móvel em companhias que desejam melhorar o modo de trabalho de seus funcionários e sua produtividade. Por essa característica, a usabilidade desses sistemas é um fator crucial. Mesmo com sua importância, o teste de usabilidade acaba sendo negligenciado e deixado de lado devido a fatores como recursos e tempo. Isso também acontecia em testes funcionais, mas com o desenvolvimento de ferramentas automatizadas, passaram a ser adotados com mais frequência em projetos. Observando essa característica, os autores acreditam que a usabilidade pode também se beneficiar da automação para ser mais constantemente aplicada em projetos. A própria natureza dos programas para *handheld* faz com que eles sejam suscetíveis a baixa usabilidade. Os autores realizam três estudos de caso com falhas comuns em sistemas para esse tipo de dispositivo, um de uma tela de cadastro, uma de realização de um pedido e uma página de uma loja. Ao todo, 17 problemas são apresentados e classificados pelo autor como apenas um subconjunto dos vários que os

sistemas para *handheld* apresentam em geral. Após analisar características de mercado da área de usabilidade, os autores colocam como objetivo a automação destinada a desenvolvedores sendo possível realizar testes de regressão durante o projeto. Os objetivos da ferramenta de avaliação são:

- teste de usabilidade efetivo, onde a ferramenta deve detectar tanto quanto ou mais problemas que uma metodologia convencional;
- maior velocidade, pois a ferramenta deve analisar mais rapidamente que um método tradicional;
- maior eficiência, uma vez que partes do processo serão assumidas por ela, os desenvolvedores estarão livres para realizar outras tarefas;
- maior eficácia e precisão, pois os resultados produzidos estarão sempre funcionalmente corretos e erros limitados a entradas incorretas;
- menor demanda de recursos, a automação deve ser capaz de reduzir tempo, recursos humanos, equipamento e custo;
- maior flexibilidade, vários aspectos de usabilidade devem ser testados, permitindo customização e facilitando teste em vários estágios do projeto;
- consistência, a ferramenta deve ter um meio de manter os padrões testados durante o projeto, para que eles possam ser utilizados por diferentes pessoas;
- promover a usabilidade, a ferramenta deve encorajar os *stakeholders* a estarem envolvidos no processo de usabilidade.

Os autores visualizam como capacidades desejáveis da ferramenta a coleta e interpretação de vários aspectos de aplicações, possibilitando a análise de aspectos como a disposição de itens de formulário, quantidade de cliques, rolagem e entrada de usuário para completar uma tarefa e taxa de erros. A ferramenta também deve oferecer padrões de projeto para telas comuns para guiar desenvolvedores dando um ponto de início com boas soluções, provendo também a consistência. Além disso, a ferramenta deve ser capaz de verificar se um padrão de projeto foi implementado corretamente. Outra característica desejada é a gravação da interação do usuário com o software com possibilidade de destacar determinadas ações tomadas pelo usuário durante o teste. Também é de interesse que a ferramenta dê suporte a interpretação de emoções do usuário durante o teste, pois comumente representam informações sobre a usabilidade. Os resultados da aplicação devem ser apresentados em diversos níveis de detalhe

e com múltiplas perspectivas, apesar de ser considerado mais importante que seja rápida e fácil de interpretar. O software implementado realiza a gravação das ações do usuário através de ações realizadas em formulário em formato XML, mas apenas um subconjunto de ações foi possível. Os padrões de entrada da ferramenta também são codificados no formato XML, com uma opção embutida na própria ferramenta para criação ou entrada manual. A ferramenta também é capaz de comparar o comportamento esperado pelo desenvolvedor contra o comportamento real dos usuários. Conformidade com métricas de usabilidade também é realizada, baseada em valores aceitáveis definidos automaticamente ou pelos desenvolvedores. A quantidade de determinadas ações do usuário tem sua quantidade verificada para assegurar que estão dentro de um padrão aceitável. A ferramenta também é capaz de fornecer um resumo com medidas estatísticas sobre os testes e a frequência de uso por componente. Apesar da grande quantidade de métricas cobertas, a ferramenta limita-se a análise de formulários e não implementou várias funcionalidades propostas. Apesar disso, nenhuma ferramenta de teste de usabilidade automatizada analisada foi capaz de atender todos os requisitos e a proposta dos autores teve um dos melhores resultados.

O trabalho (SCHWERZ; MORANDINI; SILVA, 2007) apresenta uma proposta de evolução do ErgoMonitor de (MORANDINI, 2003) para reduzir o trabalho de identificação de tarefas do especialista da aplicação alvo da avaliação de usabilidade. O ErgoMonitor realiza a captura de dados de arquivos de *log* de servidores Web para monitoração de dados de usabilidade. Através desses dados e modelos de comportamento esperados a ferramenta calcula métricas que quantificam a usabilidade do site. No ErgoMonitor é necessário realizar as seguintes etapas para que uma análise de usabilidade seja realizada:

- O especialista da aplicação seleciona quais tarefas devem compor o modelo de tarefas monitoradas e comportamento;
- um filtro é realizado nos arquivos de *log* de onde resulta o modelo de tarefas e comportamentos verificados, de onde serão determinadas as métricas;
- são determinadas as métricas que quantificam a usabilidade da interface;
- os dados são armazenados em um banco de dados histórico com os valores obtidos pela interface, permitindo a comparação de diversas etapas de um mesmo projeto;
- geração de relatórios com os dados obtidos da avaliação de usabilidade.

Apesar de ter se mostrado capaz de realizar inspeções de usabilidade, em sua primeira versão o ErgoMonitor necessitava da especificação manual de cada tarefa a ser assistida pelo *software* e

o mecanismo de limpeza do arquivo de *log* não havia sido implementado. O fato do especialista da aplicação ter que definir as tarefas está sujeito a erros e pode ser uma tarefa que leva grande quantidade de tempo ao se considerar uma aplicação muito grande. O ErgoMonitor considera que uma tarefa é formada pelo conjunto de páginas que um usuário precisa atingir em uma tarefa, a primeira página, a última página que identifica o sucesso e uma função de transição que indica que o usuário trocou de uma página para outra. O foco do trabalho (SCHWERZ; MORANDINI; SILVA, 2007) é o desenvolvimento de um mecanismo em que o especialista da aplicação manipule o modelo de tarefas, registrando o site a ser avaliado e as URLs de suas páginas. Então ele poderá registrar as tarefas que deseja avaliar. Após realizar o filtro do arquivo de *log*, com base nas tarefas definidas para análise, o ErgoMonitor busca identificar o comportamento dos usuários com uma das seguintes classificações: acesso a tarefa, tarefa com sucesso, tarefa sem sucesso, tarefa com página de erro ou tarefa com página de ajuda. Um experimento prático foi realizado com os novos módulos do ErgoMonitor. O arquivo de *log* do site da Construtora Cidade Verde, com 3 meses de tráfego, 22.66 Mb e mais de 113000 acessos foi analisado. Foram analisadas 64 tarefas, que geraram um arquivo após o filtro de 638 Kb e 3442 acessos. Esses dados mostram que a identificação manual das tarefas é praticamente impossível devido a grande quantidade de acessos de tarefas. O experimento foi capaz de identificar acessos, sucesso e falhas nas tarefas, mas não erros e acessos a ajuda uma vez que essas páginas não estavam disponíveis. Devido a esses resultados os autores concluem que o ErgoMonitor é capaz de realizar análises de usabilidade satisfatoriamente e que somente após o desenvolvimento do modelo de tarefas e comportamentos é que foi possível analisar grandes *websites*.

2.3.2 Trabalhos relacionados a Web Mining

Em geral, técnicas de *Data Mining* típicas podem ser aplicadas para mineração de dados após o pré-processamento de dados, entretanto existem variações dessas técnicas empregadas com o mesmo objetivo em *Web Mining* (KOSALA; BLOCKEEL, 2000). Algumas das técnicas aqui abordadas, como conjuntos *fuzzy* e redes neurais e algoritmos genéticos, constituem as principais ferramentas utilizadas na área (PAL; TALWAR; MITRA, 2002).

Iniciamos com o algoritmo *Balanced Iterative Reducing and Clustering using Hierarchies* (**BIRCH**), proposto por (ZHANG; RAMAKRISHNAN; LIVNY, 1996) como alternativa as abordagens de aprendizado de máquina e estatísticas que não se preocupavam com o fato de trabalharem com recursos limitados, mais particularmente não considerando o fato do conjunto de dados ser tão largo que não sendo possível armazená-lo em memória. Segundo os autores,

a maior vantagem obtida foi a possibilidade de trabalhar com grandes quantidades de dados, além das seguintes:

- ser local, o que possibilita que cada decisão de clusterização seja realizada sem necessidade de releitura de todos os dados ou *clusters* existentes;
- explorar a observação de que o espaço de dados é usualmente não uniforme, e nem todo ponto de dados é igualmente importante para clusterização;
- realizar uso total da memória disponível para derivar os *subclusters* mais finos possíveis, diminuindo custos de I/O;
- o método que não requer todo o conjunto de dados ao omitir passos opcionais.

No trabalho de (FU; SANDHU; SHIH, 1999) os padrões de utilização dos usuários são clusterizados a partir dos arquivos de *log* do servidor utilizando o BIRCH combinado a indução orientada a atributos pois os autores consideram estudos anteriores de mineração de dados de utilização da *Web* deficientes, por falharem na profundidade de compreensão dos padrões de acesso dos usuários. Outra motivação do trabalho foi dada pelo fato de que em um servidor existe a possibilidade da existência de uma grande quantidade de páginas, o que leva a uma representação de usuário com vetores de grandes dimensões, dificultando o processo de encontrar usuários com acessos comuns as páginas, diversidade que leva a *clusters* com poucos usuários ou poucas semelhanças, sendo clusterização uma boa solução.

A generalização das sessões de usuário, inicialmente utiliza o processo de indução orientado por atributos das sessões, que necessita de uma hierarquia de páginas. Essa hierarquia de páginas é construída de modo a generalizar cada URL, onde o nó folha, ou página simples, é apontado por um nó pai, ou página geral. Após isso a indução orientada por atributos da sessão é aplicada, o que também é chamado de subida de árvore. A quantidade de níveis a subir é decidida pelo usuário ou por um limiar, especificando o número de páginas nos resultados.

Na sequência, o algoritmo BIRCH é aplicado sobre as sessões. Ele constrói uma árvore chamada *Clustering Feature* (CF) como resultado da clusterização, através da inserção de objetos, representados como vetores, de modo incremental na CF.

Os autores realizaram experimentos em diferentes conjuntos de dados, variando entre 50.000 e 500.000 entradas no *log*, com tempo de inatividade definido em 30 minutos. O número de sessões variou de 18.184 até 178.655, com um nível de generalização

definido em 2, que é o nível abaixo da raiz. Isso possibilitou uma redução no conjunto de dados variando de 2/3 a 9/10. Também verificou-se que o tempo de execução cresce linearmente de acordo com a entrada.

Os resultados mostraram que o método proposto é escalável, mas alguns problemas como generalização excessiva fazem com que os *clusters* percam sua validade no caso de generalização muito alta das sessões, o que causa perda de detalhes e coloca a validade dos *clusters* em dúvida. Mesmo com grande redução na quantidade de sessões, os autores consideram a generalização das páginas para organizá-las de modo hierárquico antes de aplicar a generalização das sessões, de modo a diminuir a dimensão dos dados de sessão. Também é proposta a utilização dos dados dos usuários contidos em cadastros no site como parte dos dados utilizados na clusterização (FU; SANDHU; SHIH, 1999).

Em (GARG et al., 2006) o BIRCH é melhorado através de uma implementação paralela, denominada PBIRCH. Essa melhoria foi proposta devido a característica do BIRCH utilizar toda a memória disponível e em alguns casos necessitar de duas leituras dos dados de entrada. Os principais resultados alcançados foram a escalabilidade e melhora de velocidade quase linear com a quantidade de processadores. Acredita-se que a qualidade dos *clusters* obtidos também é melhor, uma vez que a versão paralela apresenta maior quantidade de CFs e também devido a soma de Rádios dos *clusters* da versão paralela ser menor do que no tradicional.

A aplicação da técnica de **modelos de Markov** para predição de futuras explorações da Web é apresentada por (HAFRI et al., 2003). Os autores identificam duas necessidades para garantir a usabilidade de *Websites* que se propõem a oferecer alternativas interessantes de personalização de navegação baseadas em preferências de usuários detectadas automaticamente. São elas: (1) obter representações da Web exaustivas e comprimidas, e (2) prover diferentes estratégias de exploração adaptadas as necessidades dos usuários, foco dos autores que investigam a utilização da Web baseada em explorações históricas e perfis de usuários. Utilizando modelos de Markov eles conseguem endereçar a questão de manter os dados do esquema de inferência obtidos durante o tempo, pois a informação é mutável e o grupo de usuários amplamente desconhecido e pode se alterar durante o processo. Outra motivação foram deficiências encontradas em abordagens que utilizavam cadeias de Markov.

O princípio da probabilidade máxima é utilizado sem perda de generalidade para estimar a matriz que representa troca de um estado para outro e a distribuição de probabilidade inicial dos estados correspondentes ao espaço de estado. A distribuição de probabilidades é criada a partir do histórico dos usuários, tidos como boas predições de suas próximas explorações. Variações do processo de Markov acomodam peso para mais de um estado histórico,

sendo possível então utilizar cada exploração anterior para predição das futuras.

A matriz de transição de Markov é clusterizada de modo a diminuir a sua dimensão. Estados similares são agrupados alcançando reduções de $\log N$ através de uma variação do método *k-medoids*, que substitui o método de amostragem por heurísticas, melhorando o custo do método de troca e podendo aumentar a confiabilidade. Esse método é um algoritmo de clusterização relacionado ao *K-means* mas que utiliza medoids, ponto com distância média mínima para todos os outros objetos em um mesmo *cluster*, como protótipos do *cluster* (XU; WUNSCH; SOCIETY, 2009).

Páginas acessadas com grande frequência ou raramente acessadas não são representações ideais de comportamento dos usuários e foram descartadas. A abordagem foi testada em um conjunto de 90785 sessões atingindo resultados como a possibilidade de predição probabilística, análise de caminhos e clusterização. Nos testes foram considerados diferentes quantidades de classes, iterações e número máximo de vizinhos. Foi suposto existir quatro comportamentos típicos de usuário e cinco iterações para o algoritmo de clusterização. Dos experimentos concluiu-se que o tempo de execução é proporcional a quantidade de vizinhos, que variou de 10 a 50, apesar do tempo de execução crescer em menor proporção que a quantidade de vizinhos. A quantidade de vizinhos também influenciou na qualidade da distorção. Outros pontos importantes foram detectados, como a seleção de atributos do conjunto de dados que inicialmente possuía 200 sendo utilizados efetivamente somente 19; a quantidade de sessões também foi reduzida de 230000 para 90785; pequenos centros de gravidade nos *clusters*, associados a poucas páginas que representam sessões típicas; modelos de Markov com altas dimensões mas esparsos, devido a grande quantidade de páginas, mas muitas sem utilização, necessitando de muitos recursos computacionais; e por fim ficou a questão em aberto de como os administradores de *Web sites* podem utilizar os comportamentos representativos encontrados.

Redes Neurais Artificiais possuem sua estrutura baseada na compreensão atual do sistema nervoso humano. Elas exploram muitas hipóteses simultaneamente concorrentes utilizando redes paralelas massivamente conectadas compostas de muitos elementos computacionais com conexões de pesos variáveis. Modelos de redes neurais são especificados através da topologia da rede, características dos nós e regras de treinamento (LIPPMANN, 2003). As redes neurais podem ser utilizadas em uma ampla gama de problemas, mas possuem grande potencial para reconhecimento de padrões (WANG, 2002), área de aplicação fundamental para *Web mining* (SRIVASTAVA et al., 2000).

Estudos têm demonstrado o sucesso da capacidade de aprendizado das redes

neurais, sendo mais utilizadas, no caso de *Web Mining*, as redes com aprendizado não supervisionado pela sua boa capacidade de clusterização. Entre as técnicas utilizadas encontram-se os **Mapas Auto-Organizáveis** (SOM) e a Teoria da Ressonância Adaptativa (ART) 1 e 2.

Duas variações da rede neural SOM são apresentadas em (SMITH; NG, 2003), o WEBSOM que agrupa páginas Web com base em seu conteúdo, e o método proposto pelos autores, o LOGSOM, que organiza as páginas de acordo com os padrões de navegação dos usuários. Os interesses dos usuários expressos durante sua navegação são uma fonte de dados mais útil para oferecer serviços aos próprios usuários, como a personalização de navegação, sugestão de páginas, entre outros. Observando isso, os autores propõem o LOGSOM o que possibilita atualizações regulares refletindo os interesses dos usuários e resultados possivelmente melhores que do WEBSOM. A rede neural SOM foi selecionada devido a sua capacidade de representar graficamente os relacionamentos entre os *clusters*, ser flexível o bastante para ser adotada em uma abordagem de clusterização hierárquica para apresentar grandes quantidades de pontos de dados e pela sua viabilidade baseada na organização de páginas Web já ter sido provada. Os arquivos de *log* utilizados foram retirados do servidor da *Monash University, School of Business Systems*, que possuía 170515 entradas. Esse servidor é basicamente acessado por alunos em laboratórios, o que facilitou tarefas de pré-processamento. Com os usuários identificados e sessões agrupadas, uma tabela é gerada no formato transações x URLs, que no experimento possuía 8054 transações e 235 URLs.

Essa quantidade de transações torna a aplicação do SOM impraticável, pois a quantidade de entradas para o algoritmo deve ser equivalente ao número de transações. Além disso, em sistemas reais, como uma ferramenta de busca, a quantidade de dados é infinitamente maior, o que torna essa tarefa ainda mais crítica, tanto em termos de processamento quanto em aplicabilidade da técnica. Esse problema foi solucionado combinando particionamento de hipergrafo e regras de associação, de modo a clusterizar os documentos Web com a mesma abordagem pretendida inicialmente, baseada no comportamento de navegação dos usuários. Ainda para reduzir a dimensão, foi utilizado o algoritmo *K-means* que agrupou as transações em 9 grupos, valor escolhido arbitrariamente. Dentro de cada grupo de transação processado foi utilizado o total da coluna como atividade do grupo de transação, os grupos de transações são utilizados como características ao invés de transações. Com esse agrupamento, as dimensões foram reduzidas. A magnitude foi reduzida de 8054 para 9.

A arquitetura utilizada e saída desejada da rede SOM é um mapa bidimensional 16x16. Os resultados obtidos foram satisfatórios, entretanto é difícil medir quantitativamente a eficiência do SOM, mesmo ao se comparar com outros métodos. Apesar disso, os

resultados experimentais mostraram que as páginas foram clusterizadas adequadamente, inclusive com a organização hierárquica das páginas sendo semelhante as preferências de navegação dos usuários. Ao analisar a vizinhança também foi observado um comportamento semelhante, o que validou a aplicação do LOGSOM. A qualidade dos resultados também foi estudada pela alteração de parâmetros do SOM, α que representa a quantidade de trocas de pesos dos nós no processo de aprendizagem e Ω , que determina a quantidade de repetições do processo de aprendizagem dentro de um ciclo. A principal deficiência detectada na pesquisa foi a utilização de *logs* de exemplo, que não necessariamente caracterizam aplicações ou dados do mundo real, o que torna o maior desafio do LOGSOM o tamanho dos dados de uma aplicação Web real. Outra questão importante é a determinação do valor de *clusters* do *K-means* durante o pré-processamento, que deve ser cuidadosamente selecionado. Também não é claro como a topologia da rede SOM é definida, sendo apresentadas somente questões relativas a variação de parâmetros internos ao SOM.

Em (SORIANO-ASENSI et al., 2008) a questão de personalização de grandes portais governamentais para o cidadão é discutida. Muitas vezes, a análise da interação dos usuários com a aplicação não é possível devido a grande quantidade de acessos. No trabalho é proposta a utilização de uma variação do algoritmo SOM para analisar um portal de cidadãos na Web, o Infoville XXI do governo de Valência, pois o algoritmo clássico possui a deficiência da necessidade de se conhecer o tamanho adequado da camada de saída antes da aplicação do algoritmo. A variação proposta, o *Growing Hierarchical Self-Organizing Map (GHSOM)*, tenta resolver esse problema com um tempo de treinamento baixo comparado a outras abordagens para o mesmo problema, como o algoritmo *Growing Grid (GG)*, o que é importante para grandes quantidades de dados, como em portais. Apesar dessas vantagens, o autor considera que o GHSOM possui a deficiência de não apresentar um mapa único como resultado do treinamento, mas uma estrutura hierárquica de mapas. O trabalho propõe uma nova abordagem que junta toda a informação dispersa pela estrutura hierárquica. Os dados utilizados para classificar os usuários passaram por um filtro que eliminava os seguintes tipos de dados:

- relativos à usuários anômalos, como acessos de administradores com propósito de testes e usuários que apresentam quantidade de acessos excessivamente elevada, fugindo do padrão;
- descritores de serviços do portal com acessos excessivamente baixos ou altos, com desvios padrão maiores do que três vezes o valor médio, o que resultou em 16 descritores utilizados de 22 disponíveis.

Em uma análise inicial com os 16 descritores, verificou-se que 83% dos usuários focam atenção em apenas 5 descritores, o que levou os autores a uma nova redução no conjunto de dados, agora formada por 1676 padrões representando as frequências de acesso aos cinco descritores. Utilizar o conjunto total não iria trazer contribuições relevantes além de fazer mais difícil a extração de conhecimento.

Um dos problemas do algoritmo GG é que devido a sua abordagem de crescimento do mapa gerado pelo SOM ele deve repetir o treinamento até encontrar o tamanho adequado do mapa, tornando o processo de treinamento muito lento. O GHSOM, ao invés de inserir linhas ou colunas de neurônios a um mapa já existente, como no GG, insere uma nova estrutura que mapeia em detalhes padrões representados por um neurônio com grande erro de quantização. Apesar dessa estrutura hierárquica melhorar o tempo de aprendizagem, é difícil analisá-la entre os diversos mapas. Uma vez que áreas do mapa precisam de mais detalhes que outras, elas podem estar representadas em diferentes níveis. Entretanto, a partir das localizações dos neurônios, é possível representar os dados hierárquicos em apenas um mapa preservando os relacionamentos topológicos. Para isso, os autores utilizaram um novo SOM de duas dimensões que recebe as localizações dos neurônios na estrutura hierárquica como padrões de treinamento. Desse modo os pesos dos neurônios após esse último treinamento estão diretamente relacionados a localização do neurônio na estrutura hierárquica. Deve-se assegurar que cada neurônio nesse mapa não representa mais do que um neurônio na estrutura hierárquica, por isso a quantidade de neurônios do segundo mapa deve ser igual ou ligeiramente maior que da estrutura hierárquica. Neurônios adicionais são usados para interpolar áreas com baixa densidade de dados. O experimento com o portal Infoville XXI conseguiu determinar os tópicos de maior interesse dos cidadãos assim como algumas particularidades em um mapa de 9 linhas e colunas, ou 81 neurônios, utilizando 1000 padrões de entrada e 676 como teste.

Alternativas de visualização dos *clusters* após a fase de aprendizado são oferecidas pelos autores. Essa é uma característica importante e deficitária no SOM (KUO et al., 2006). Uma delas é o histograma de dados, que representa a quantidade de padrões associados aos *clusters*. Após analisar os padrões de distribuição no mapa, deve-se estudar o significado de áreas com alta concentração de padrões. Para isso foi utilizada uma representação baseada em componentes planos. Também foi utilizada uma terceira representação útil para ter informações sobre a quantidade de padrões e suas características representadas por cada neurônio. As representações utilizadas tornam mais simples análises quantitativas e qualitativas.

Utilizando a rede neural ART1 para clusterização, (RAJU et al., 2007) também realiza análise de utilização da Web de usuários que representam a minoria, que muitas

vezes não são encontrados por outros algoritmos considerados eficientes. Buscar por padrões nesse tipo de usuário inicialmente pode parecer irrelevante, entretanto padrões como atividades de “hacking” foram encontradas em até 0.04% das requisições. A abordagem utilizada no artigo assume que com um algoritmo de clusterização apropriado, sessões similares correspondentes aos usuários podem ser agrupadas no mesmo *cluster*. Os *logs* do servidor foram pré-processados com heurísticas que realizavam a filtragem de dados e ao fim do processo armazenados em um banco de dados. Desse banco de dados é extraído um vetor de padrões binários com as URLs acessadas por um grupo de usuários organizacionalmente relacionados.

O **ART1** é uma rede neural não-supervisionada que possui a vantagem de se adaptar as alterações dos padrões dos usuários ao longo do tempo sem perder informação sobre seus padrões anteriores. Comparado ao SOM e ao *K-means*, o ART1 é capaz de produzir *clusters* relativamente estáveis e com qualidade. Nos experimentos realizados para teste, foram utilizados arquivos de *log* disponibilizados pela NASA. Ao se comparar a abordagem utilizada com outras da área de mineração de padrões sequenciais, verificou-se que com valores de suporte entre 0.02 e 0.06% a abordagem utilizada é muito superior as utilizadas como comparação, que com valores de 0.06% já se tornam exponenciais, o que só ocorre com a abordagem testada em valores abaixo de 0.025%. Ainda existe a necessidade de realizar testes dessa abordagem em iterações com Web sites reais.

Dificuldades semelhantes também são mencionadas em (RANGARAJAN et al., 2004) para agrupar usuários Web com base em seus interesses extraídos de *logs*. Uma delas é justamente a alteração desses durante o tempo, como também observado em (RAJU et al., 2007) e a potencial infinidade e diversidade de usuários. Em seu trabalho, (RANGARAJAN et al., 2004) também utiliza a rede neural ART1 com acessos representados como um vetor binário a partir dos arquivos de *log* fornecidos pela NASA, comparando o desempenho obtido com a do *K-means*. O trabalho tentava realizar a previsão de URLs a serem requisitadas por um usuário, o que chegou a níveis de até 97.78%, com média de 92.3%. A técnica utilizada previa requisições baseadas em grupos de usuários clusterizados através do ART1, o que é bom em termos de desempenho pois algumas abordagens realizam previsões baseadas em acessos de apenas um usuário. A comparação com o *K-means* foi realizada pela alteração do parâmetro de vigilância, entre 0.3 e 0.5, e as medidas de qualidade são a distância entre os *clusters* (distância *intercluster*) e a distância média entre membros do mesmo *cluster* (distância *intracluster*). A quantidade de *clusters* informada como entrada ao *K-means* era a apresentada nos resultados do ART1. Chegou-se a conclusão que o algoritmo ART1 obtém *clusters* mais homogêneos comparado ao *K-means*, o que indica estabilidade na clusterização, atributo associado a *clusters* de alta qualidade. Os resultados de previsão de requisições foram muito superiores aos de outras

pesquisas, que chegaram a no máximo 75% de efetividade. Apesar dos bons resultados, redes neurais como o ART1 conseguem apenas capturar, mas não conseguem utilizar as propriedades inerentes de auto-similaridades da *World Wide Web*.

Em (LIANGUN; BIN; YUANYUAN, 2009) uma melhoria na rede neural **ART2** é proposta e aplicada a área de *Data Mining*. No subsistema de atenção cada neurônio aumenta um neurônio de acompanhamento, que armazena a informação de escopo do modelo de classe correspondente. No subsistema de ajuste um sistema de *reset* B é adicionado, que realiza uma medição de distância caso não seja enviado um sinal de *reset* do sistema original. Caso o valor obtido seja maior que um limiar deve ser criada uma nova classe para armazenar o padrão, operação que utiliza os valores dos neurônios de acompanhamento. Um experimento com dados gerados por uma função que produz pontos espaciais divididos em nove categorias foi realizado comparando o ART2 clássico com as melhorias propostas pelos autores. Segundo os autores, as melhorias propostas aumentam o escopo da informação no processo de clusterização, o que melhora efetivamente sua eficiência, reduz chamadas ao parâmetro de alerta e forma resultados de clusterização de estrutura hierárquica.

Melhorias significativas relacionadas ao comportamento das interações com a Web no mundo real são alcançadas com a utilização de **teoria dos conjuntos Fuzzy** (AROTARITEI; MITRA, 2004). A abordagem não é limitada por fontes de dados quebradiças, sobrepostas ou por dados incompletos, situações muito comuns em arquivos de *log* com dados reais e também é capaz de classificar um usuário em diferentes grupos de interesse, algo muito comum na Web (KOLARI; JOSHI, 2004). Em (SONG; SHEPPERD, 2006) um algoritmo é proposto pelos autores e a clusterização de usuários e páginas baseados na teoria dos conjuntos *Fuzzy* é utilizada como comparação. Algumas definições são apresentadas pelos autores para possibilitar a utilização dos conceitos relacionados a *Fuzzy*. São elas:

- conjunto de usuários;
- conjunto de páginas;
- subconjunto *fuzzy* que reflete o comportamento de navegação de um usuário;
- subconjunto *fuzzy* que reflete o comportamento de todos os usuários em uma página;
- medida de similaridade entre subconjuntos *fuzzy* utilizada para classificação.

Ao identificar a deficiência da necessidade de identificação das sessões de usuário em métodos de clusterização, os autores decidem desenvolver um método baseado

em análise de vetores, visto que nessa abordagem é desnecessário realizar essa tarefa. Outra vantagem da abordagem proposta é ser não-sequencial, evitando dificuldades de desempenho e escalabilidade que podem surgir durante a mineração de padrões sequenciais. Para desenvolvimento do algoritmo, os autores adaptaram conceitos conhecidos de teoria dos grafos: adjacência entre páginas, caminho em um *Website*, *hits* de usuário em uma página, *Website* e estado de uma URL. São analisados três aspectos de padrões de navegação: *clusters* de usuários, páginas e caminhos de acesso frequente. A mineração de caminhos de acesso frequente é baseada nos *clusters* de usuários e páginas, e a descoberta de *clusters* de usuários e páginas no comportamento navegacional dos usuários sem assumir grupos predefinidos. Para o método proposto pelos autores supõe-se que usuários com interesses similares devem ter padrões de navegação similares, páginas Web associadas devem ser navegadas por usuários com interesses similares, e os padrões de navegação gerais não se alteram durante um dado período de tempo para um dado usuário, enquanto diferentes padrões de navegação de usuários podem diferir durante um período específico de tempo.

O algoritmo proposto foi testado com um conjunto de dados reais do *log* do *Website* da Universidade Xi'an Jiaotong, dividido em cinco conjuntos de diferentes tamanhos no intervalo de 200 a 1000 KB. Para cada um desses conjuntos foi utilizado o método de validação cruzada de 3 dobras como estratégia de validação. Comparado ao *K-means*, ao método baseado no modelo incremental COBWEB (FISHER, 1987) e ao algoritmo clássico de reconhecimento de caminhos de acesso frequente FS (CHEN; PARK; YU, 1998) a precisão do algoritmo proposto foi melhor. Para uma determinação mais formal da significância dos resultados, os autores utilizaram o teste de *Mann-Whitney* (MANN; WHITNEY, 1947) - alternativa ao teste t, é um teste de hipótese estatística não paramétrico para determinar se duas amostras independentes possuem valores diferentes de médias - para comparar as medianas dos valores de precisão dos métodos. Com ele encontrou-se que o método de clusterização de páginas e usuários é melhor que o COBWEB e *K-means*. Além disso, os tempos de CPU medidos dos métodos propostos foram melhores, tendo sua complexidade a partir do tamanho da entrada crescendo quase linearmente e sendo menor que o clássico FS, mostrando que o método proposto é escalável.

Por outro lado, ao ser comparado com os resultados dos métodos baseados na teoria dos conjuntos *Fuzzy*, o algoritmo proposto teve uma precisão menor. Novamente, o teste de *Mann-Whitney* foi utilizado para comparar as medianas. Para clusterização de usuários e páginas verificou-se que os métodos *Fuzzy* são melhores que o *K-means* e o COBWEB. Os resultados ainda mostraram que eles são capazes de pequenas melhorias em relação aos outros métodos desenvolvidos. Outra importante conclusão está relacionada ao tempo de execução,

menor em relação aos métodos utilizados para comparação e com crescimento quase linear em relação ao crescimento da entrada de dados (SONG; SHEPPERD, 2006).

Originalmente proposto por (MACQUEEN, 1967), o método estatístico de clusterização de dados **K-means** é um dos mais populares e eficientes métodos utilizados para essa finalidade. Nele, objetos são selecionados aleatoriamente e utilizados como centróides de *clusters*. Os objetos são associados a diferentes *clusters* baseados em sua distância dos centróides. *Clusters* recém formados são utilizados para determinar novos centróides. Esse processo continua até que os *clusters* se estabilizem (HARTIGAN; WONG, 1979).

Algumas variações da proposta original vem sendo utilizadas em *Web mining*, como a que agrega conceitos de algoritmos genéticos (KRISHNA; MURTY, 2002) ou de conjuntos irregulares (LINGRAS; WEST, 2004).

O trabalho de clusterização em *Web mining* possui desafios adicionais em relação a outras áreas (LINGRAS; WEST, 2004). Os *clusters* tendem a ser vagos ou com fronteiras imprecisas, a composição de um objeto em um *cluster* pode não ser precisamente definida, existe uma probabilidade de um objeto ser candidato a mais de um *cluster* e a presença de dados incompletos e sobrepostos. Devido a essas características, houve a necessidade de adaptar o *K-means* a teoria de conjuntos irregulares (PAWLAK, 1997), utilizada para encontrar intervalos de usuários Web em (LINGRAS; WEST, 2004). Ao incorporar a teoria de conjuntos irregulares ao *K-means*, é necessário adicionar o conceito de limites superior e inferior, o que modifica a equação que calcula os centróides. Também é necessário determinar o critério para quando um objeto pertence ao limite superior ou inferior de um *cluster*.

O Website analisado é utilizado por alunos do primeiro ano de ciência da computação da Universidade Saint Mary's, que possuem comportamentos de navegação bastante diferenciados, com acessos semanais variando entre 30000 e 360000. São analisados conjuntos de visitas de um dado IP ao invés de acessos únicos que deverão ser classificados em três categorias. Os *logs* do servidor foram pré-processados, e alguns dados foram selecionados para representação dos usuários. Apesar de acreditarem que os resultados obtidos na clusterização dos alunos foram satisfatórios, os autores mencionam que é provável que o método não possa ser estendido para grandes períodos de análise, no caso maior que seis semanas, ou 361906 entradas com 22996 visitas. Também são discutidos os valores para os parâmetros de aprendizagem do algoritmo, sendo fornecidos os valores em que foram obtidos os melhores resultados.

Nos resultados da clusterização houve grande quantidade de limites superiores sobrepostos, o que mostra que o algoritmo atende a necessidade de limites de *clusters*

sobrepostos. Comparações da abordagem proposta com o *K-means* clássico e o *Fuzzy C-means* (DUNN, 1973), (BEZDEK, 1981), uma alteração do *K-means* em que é possível que um ponto esteja em mais de um *cluster* medido através de um grau de pertinência, mostraram que tornou-se mais fácil a identificação de similaridades e diferenças entre a utilização de *Websites* do que o *K-means* clássico, e que também pode ser um complemento a abordagem do *Fuzzy C-means*, entretanto o trabalho condiciona a explicação desse ponto a futuras publicações.

Estudos realizados em (KUO; HO; HU, 2002), (KUO et al., 2006) e (KUO; LIAO; TU, 2005) demonstraram que **métodos híbridos** utilizando as redes neurais como uma de suas partes conseguem melhorar resultados de abordagens simples ou as que utilizam apenas um tipo de algoritmo. Um método híbrido é proposto em (KUO; HO; HU, 2002) com dois estágios, combinando a rede neural SOM com o método *K-means*. Nessa abordagem a rede neural fica responsável pela primeira etapa do processo, descobrir a quantidade de *clusters* e pontos de partida para então aplicar o *K-means* que encontra a solução final. Esse trabalho utilizou segmentação de mercado como área de estudo, e seus resultados foram comparados a outro método de dois estágios: método da variância mínima de Ward, um método de clusterização hierárquica cujo objetivo é minimizar o aumento da soma quadrática dos erros internos as classes causada pela junção de dois *clusters* (XU; WUNSCH; SOCIETY, 2009), seguido do *K-means*. A medida de comparação utilizada foi a taxa de erros de classificação. Adicionalmente, a rede neural SOM teve seus resultados, sendo aplicada isoladamente, comparados com os métodos híbridos. Em geral, encontrou-se que a taxa de erros de classificação da rede neural SOM isoladamente foi a maior, e que ambos os métodos de dois estágios podem oferecer melhor desempenho. Apesar da taxa média de erros de classificação do método de dois estágios proposto ser menor que o método de dois estágios tradicional, o teste *t* de amostra indicou que essa diferença não é significativa. Entretanto, através de uma segmentação de mercado com dados reais o método proposto atingiu melhores resultados de acordo com o valor Lambda de Wilk - razão entre a variância interna do grupo e a total, onde valores altos indicam que não há diferenças entre médias internas de grupos, e valores próximos a zero indicam que a origem da variância total está entre os grupos - e análise discriminante - método que encontra uma combinação linear de características que agrupam ou separam duas ou mais classes de objetos.

Algoritmos genéticos tem se mostrado populares e aplicados satisfatoriamente em problemas de otimização. São baseados em princípios genéticos, onde uma população inicial de indivíduos representa uma possível solução. Essa população passa por uma série de operações genéticas, conhecidas como *crossover* e mutação. Os indivíduos se esforçam pela sobrevivência baseada num esquema de seleção que possui um *bias* para seleção dos indivíduos que melhor representam uma solução. Os indivíduos selecionados formam a próxima geração

e o processo continua. Após um número de gerações o programa converge para a solução ótima, representada pelo melhor indivíduo (CRONIN, 2003).

A generalização alcançada por algoritmos genéticos é melhor que a de algoritmos estatísticos como o *K-means* (KUO; LIAO; TU, 2005). Adicionalmente, eles também são capazes de melhorar a saída deste mesmo algoritmo e até mesmo ser combinado com ele (XU; WUNSCH; SOCIETY, 2009), abordagem que gerou o Algoritmo Genético *K-means* (GKA), considerado alternativa melhor que os algoritmos separados. Além disso, os resultados desse algoritmo podem ser melhorados através da aplicação de uma rede neural como solução inicial e entrada ao *K-means* genético, como mostram (KUO et al., 2006) e (KUO; LIAO; TU, 2005).

Em (KUO; LIAO; TU, 2005) os autores observaram que os algoritmos genéticos vinham sendo aplicados satisfatoriamente em problemas de busca, sendo em algumas condições melhores que algoritmos de clusterização tradicionais. O algoritmo *K-means* genético é proposto em (KRISHNA; MURTY, 2002), sendo provado que ele é capaz de convergir para um ótimo global. Nos trabalhos (KUO et al., 2006) e (KUO; LIAO; TU, 2005), um método híbrido de duas etapas é proposto utilizando uma rede neural e o algoritmo *K-means* genético em seguida. A solução da rede neural na primeira etapa dos algoritmos desenvolvidos é utilizada como entrada do algoritmo *K-means* genético, ligeiramente modificado em relação ao trabalho de (KRISHNA; MURTY, 2002) para receber os *clusters* iniciais da rede neural, não utilizar reprodução, utilizar mutação baseada em distância para escapar dos mínimos locais, e utilizar um operador *K-means* de convergência rápida. A abordagem proposta provou não ser limitada pela necessidade da determinação prévia da quantidade de *clusters*, além de superar os resultados apresentados pelo algoritmo original.

Dois métodos são comparados em (KUO; LIAO; TU, 2005): a rede ART2 seguida de *K-means* e a rede ART2 seguida do *K-means* genético. A mesma linha de raciocínio é seguida em (KUO et al., 2006), que utiliza a rede SOM seguida do *K-means* genético comparada aos resultados encontrados em (KUO; HO; HU, 2002).

Conjuntos de dados simulados pelo método de Monte Carlo, método utilizado para analisar propriedades de algoritmos de clusterização, foram utilizados para testar os métodos, que foram avaliados segundo as taxas de erros de classificação e variação interna do *cluster*. Para o SOM + *K-means* genético, após 35 gerações a variação interna do *cluster* não decresce significativamente, sua performance é a melhor entre os métodos com os quais ele é comparado, e a taxa de erros de classificação entre os métodos não difere (KUO et al., 2006). Já o método ART2 + *K-means* genético possui variações internas de *cluster* menores que de seus concorrentes, mesmo com o ART2 + *K-means* sendo melhor em altas densidades

em termos de erros de classificação. Segundo resultados do método de comparação múltipla de Scheff, não há diferença significativa entre ART2 + *K-means* e ART2 + *K-means* genético nem entre o SOM + *K-means* e SOM + *K-means* genético (KUO; LIAO; TU, 2005).

Testes práticos com dados reais foram realizados de modo a verificar a eficácia dos algoritmos e validar os dados encontrados a partir dos dados simulados. O algoritmo SOM + *K-means* genético (KUO et al., 2006) foi aplicado a área de segmentação de mercado. A variação interna de *cluster* obtida por ele foi melhor em comparação aos métodos *K-means*, SOM + *K-means*. O ART2 + *K-means* genético (KUO; LIAO; TU, 2005) foi aplicado a área de *Web mining*, utilizando dados do site da empresa Web PDA e gerando uma aplicação de recomendação para os clientes. Esse sistema coleta os caminhos dos usuários e implementa a mineração com o algoritmo proposto, clusterizando os dados. Por fim, o agente de recomendação analisa o escopo do usuário e recomenda páginas a ele. Os resultados obtidos na comparação entre os métodos, houve um ganho de 9,25% na variação interna do *cluster* a favor do ART2 + *K-means* genético em relação ao ART2 + *K-means*. A precisão na clusterização também foi maior para o algoritmo ART2 + *K-means* genético. Outra melhora alcançada pelo mesmo algoritmo foi uma melhor percepção da quantidade de *clusters* obtidos comparado ao SOM + *K-means* genético.

Desse modo, algoritmos híbridos, possuindo algoritmos genéticos ou redes neurais, podem ser uma boa alternativa para tarefa de clusterização de dados, e como visto, alcançam bons resultados na área de *Web mining*. Os autores ainda sugerem como trabalho futuro a substituição das redes neurais por um sistema de colônia de formigas para determinação da quantidade de *clusters* para o *K-means* genético (KUO; LIAO; TU, 2005).

2.3.2.1 Síntese dos métodos apresentados

Entre os vários métodos apresentados é possível identificar pontos positivos e negativos em relação a suas características, relacionadas a complexidade computacional, aplicação prática na área de *Web mining*, complexidade de utilização, integração com outros métodos e capacidade de suprir necessidades encontradas no ambiente Web.

A maior parte das técnicas apresentadas foi aplicada ao campo de *Web Mining*, e é possível utilizá-las em aplicações reais. As técnicas apresentadas não utilizadas em *Web Mining* foram relevantes para esse trabalho pois foram empregadas em campos similares ou comparadas a trabalhos cujo foco era *Web Mining*. A tabela 2.2 contém uma síntese dos métodos apresentados:

2.3 Trabalhos Relacionados

Referência	Problema	Método Proposto	Métodos confrontados	Testes realizados	Resultados
(FU; SANDHU; SHIH, 1999)	Recursos limitados, como memória. Falta compreensão dos padrões de acesso dos usuários	Indução orientada por atributos + BIRCH		log do servidor da University of Missouri-Rolla	Trabalhar com grandes quantidades de dados e escalabilidade
(GARG et al., 2006)	BIRCH utiliza toda memória e pode necessitar de duas leituras dos dados de entrada	PBIRCH	BIRCH		Escalabilidade, paralelismo, crescimento quase linear, maior qualidade dos clusters
(HAFRI et al., 2003)	Deficiências das cadeias de Markov, como a necessidade de categorias predefinidas antes de aplicar o método	Modelos de Markov		log do site KDD Cup	Predição probabilística; análise de caminhos; clusterização

2.3 Trabalhos Relacionados

(SMITH; NG, 2003)	Dados de navegação do usuário são melhores para agrupar páginas do que dados sobre o conteúdo das páginas	Logsom	Websom	log da Monash University School of Bussiness System	Clusters adequados com organização hierárquica e vizinhança semelhante
(SORIANO-ASENSI et al., 2008)	Análise de dados de grandes portais e necessidade de conhecer a dimensão da saída desejada	GHSOM	SOM, Growing Grid	log do Infoville XXI de Valência	Mais rápido que o Growing Grid; mapeou os dados segundo a necessidade
(RAJU et al., 2007)		ART1		logs fornecidos pela NASA	Análise de minorias e detecção de atividades de hacking
(RANGARAJAN et al., 2004)	Previsão de URLs a serem visitadas pelos usuários	ART1	K-means	logs fornecidos pela NASA	Análise de minorias; atividades de hacking

2.3 Trabalhos Relacionados

(LIANGUN; BIN; YU-ANYUAN, 2009)	Definir o parâmetro de alerta antes do treinamento; os autores consideram a estratégia <i>winner takes all</i> problemática; incapacidade de manifestar todas as relações de entrada; aplicações limitadas somente ao escopo atingido pelo ART2	ART2 melhorado	Simulação com função que produz pontos espaciais divididos em 9 categorias	Data Mining	Melhor eficiência devido ao aumento de escopo da informação durante clusterização e resultados com estrutura hierárquica
(LINGRAS; WEST, 2004)	clusters vagos; fronteiras imprecisas; um objeto pode estar presente em mais de um cluster; dados incompletos e sobrepostos	K-means + teoria dos conjuntos irregulares	K-means; Fuzzy C-means	logs do Website destinado aos alunos do primeiro ano da Universidade Saint Mary's	Atende a necessidade de clusters sobrepostos, melhor que K-means para interpretação dos resultados

2.3 Trabalhos Relacionados

(SONG; SHEPPERD, 2006)	Fontes de dados quebradiças, sobrepostas, incompletas e possibilidade de usuários simultaneamente em mais de um grupo	Teoria dos conjuntos Fuzzy	K-means; FS; Cobweb	log da universidade de Xi'an Jiaotong	Grande precisão; melhor que o K-means e Cobweb para classificar usuários; tempo de execução melhor e crescimento quase linear com relação a entrada
(KUO; HO; HU, 2002)	Bom para a área de aplicação e estudar a possibilidade de integração de Redes Neurais e análise multivariada	SOM + K-means	Variância Mínima de Ward + K-means; SOM	Segmentação de Mercado	Melhor desempenho e melhor taxa média de erros
(KUO; LIAO; TU, 2005)	Melhorar e facilitar a utilização dos algoritmos	ART2 + K-means; ART2 + GKA	comparados entre eles e com (KUO et al., 2006); Simulação do método de Monte Carlo	logs da empresa Web PDA	Melhor percepção da quantidade de clusters, precisão e variação interna do cluster

(KUO et al., 2006)	Melhorar e facilitar a utilização dos algoritmos	SOM + GKA	Comparada (KUO; HO; HU, 2002)	Segmentação de mercado	Melhor performance, desempenho e variação interna
--------------------	--	-----------	-------------------------------	------------------------	---

Tabela 2.2: Detalhamento do retorno de investimento para cada métrica estudada

2.4 Revisão das técnicas de clusterização utilizadas

Esta seção detalha especificamente as técnicas de clusterização utilizadas neste trabalho.

2.4.1 K-means

O termo *K-means* foi utilizado pela primeira vez por (MACQUEEN, 1967), apesar da ideia ter sido apresentada em (STEINHAUS, 1957). Já o algoritmo padrão foi proposto por (LLOYD, 1982) como uma técnica de modulação de código de pulso, e por isso o método *K-means* também é conhecido como algoritmo de Lloyd, que originalmente foi concebido para dados escalares (KANUNGO et al., 2002). Apesar de ser considerada uma boa heurística para a clusterização *K-means*, o algoritmo de Lloyd não constitui uma solução ótima, pois essa ainda não foi encontrada. Em alguns casos, o problema é considerado NP-Hard.

O objetivo do algoritmo é dividir M pontos de dimensão N em k *clusters* de modo que a soma dos quadrados internos do *cluster* sejam minimizados. A solução buscada deve preferencialmente não ser um ótimo local, ou seja, ao movimentar um ponto da solução atual de um *cluster* para outro as somas dos quadrados internos dos *clusters* não serão reduzidas. O algoritmo requer como entrada uma matriz $M \times N$, onde M é a quantidade de pontos e N a dimensão dos pontos, e uma matriz de k centros iniciais de dimensão N . Em geral, o procedimento busca por uma partição k com ótimo local de soma dos quadrados internos do *cluster* pela movimentação de pontos de um *cluster* para o outro até que seja atingido um critério de conversão.

Mais detalhadamente, a clusterização *K-means* inicia escolhendo os centros iniciais $m^{(k)}$ de algum modo, por exemplo aleatoriamente. Então um algoritmo iterativo de dois passos é iniciado. No passo de atribuição cada ponto n é associado ao meio mais próximo

através de uma função que calcula a distância entre o ponto e o centro, normalmente é utilizada a distância Euclideana 2.1.

$$d(x, y) = \frac{1}{2} \sum_i (x_i - y_i)^2 \quad (2.1)$$

Sendo assim, cada ponto n é associado ao centro mais próximo $k^{(n)}$ 2.2.

$$k^{(n)} = \arg \min_k d(m^{(k)}, x^{(n)}) \quad (2.2)$$

Essa associação também pode ser realizada de maneira a indicar responsabilidades $r_k^{(n)}$, como mostra a equação 2.3.

$$r_k^{(n)} = \begin{cases} 1, & \text{se } k^{(n)} = k \\ 0, & \text{se } k^{(n)} \neq k \end{cases} \quad (2.3)$$

No passo de atualização, os centros são ajustados de modo a condizer com as médias dos pontos de dados pelos quais eles são responsáveis. O ajuste é realizado como mostra a equação 2.4.

$$m^{(k)} = \frac{\sum_n r_k^{(n)} x^{(n)}}{R^{(n)}} \quad (2.4)$$

onde $R^{(n)}$ é a responsabilidade total do centro k 2.5,

$$R^{(k)} = \sum_n r_k^{(n)} \quad (2.5)$$

Quando $R^{(k)} = 0$, ou seja, não possui responsabilidades, o meio $m^{(k)}$ permanece no local em que está.

Esses dois passos são repetidos até que não seja necessário realizar mais nenhuma associação (MACKAY, 2003).

Geralmente para pontos que não são equidistantes de dois centros, o algoritmo converge para mínimos locais, que não necessariamente representam um mínimo global. A determinação cuidadosa de *clusters* iniciais pode reduzir drasticamente o tempo de execução do algoritmo (HARTIGAN; WONG, 1979), (KANUNGO et al., 2002), além de evitar o caso em que centros iniciais vazios permaneçam vazios após a execução do algoritmo (BRADLEY; FAYYAD, 1998).

De modo a obter respostas satisfatórias do algoritmo, é possível realizar testes com entradas aleatórias para a quantidade de centros e para os centros iniciais. Mesmo

com o *K-means* sendo um algoritmo de rápida execução, realizar esses testes em grandes quantidades de dados é uma tarefa árdua. Outras alternativas para inicialização de centros são apresentadas em (BRADLEY; FAYYAD, 1998).

2.4.2 Mapas Auto-Organizáveis

O algoritmo conhecido por Mapas Auto-Organizáveis (SOM) é baseado em aprendizado competitivo. Nele, os neurônios de saída de uma grade competem entre si para serem disparados, resultando em apenas um neurônio de saída ativo em um determinado instante de tempo. Em uma rede neural SOM os neurônios estão dispostos como nós em uma grade normalmente uni ou bidimensional topologicamente ordenada. Os neurônios se tornam seletivamente sintonizados a vários estímulos de entrada no decorrer do processo de aprendizado. As localizações dos neurônios vencedores se tornam ordenadas entre si de forma que um sistema de coordenadas significativo para diferentes características de entrada é criado sobre a grade. Desse modo um mapa auto-organizável é um mapa topográfico dos padrões de entrada no qual as coordenadas dos neurônios são indicativos das características estatísticas intrínsecas contidas nos padrões de entrada (HAYKIN; ENGEL, 2001). Já (KOHONEN, 1982) não restringe os resultados apenas a mapas topográficos, mas a mapas de padrões relacionados a uma característica arbitrária ou espaço de atributos, e em qualquer nível de abstração.

O desenvolvimento do SOM como modelo neural foi motivado pela característica distintiva do cérebro humano, que organiza entradas sensoriais diferentes em locais diferentes, como se fossem representadas por mapas computacionais ordenados topologicamente.

Cada neurônio da grade SOM está totalmente conectado com os nós da camada de entrada, como mostra a figura 2.6. Cada padrão de entrada apresentado a grade consiste de um foco de atividade contra um fundo em repouso, sendo que a localização e a natureza desse foco variam entre os padrões de entrada. Por isso todos os neurônios da grade devem ser expostos a um número suficiente de diferentes realizações do padrão de entrada, assegurando que o processo de auto-organização amadureça apropriadamente (HAYKIN; ENGEL, 2001).

O algoritmo de treinamento primeiramente inicializa os pesos sinápticos da grade, o que pode ser feito através da atribuição de pequenos valores aleatórios, fazendo que nenhuma organização prévia seja imposta ao mapa de características. Após a inicialização, há três processos essenciais envolvidos na formação do mapa (HAYKIN; ENGEL, 2001):

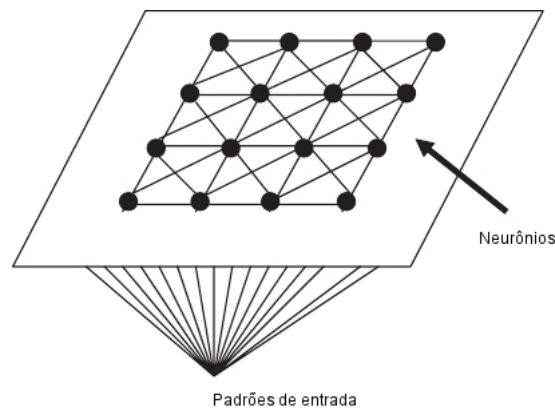


Figura 2.6: Arquitetura básica de uma Rede Neural SOM

1. Competição: para cada padrão de entrada os neurônios da grade calculam seus respectivos valores de uma função discriminante, que fornece a base para competição entre os neurônios. O que tiver maior valor resultante dessa função é declarado vencedor;
2. Cooperação: o neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo assim a base para a cooperação entre os neurônios vizinhos;
3. Adaptação sináptica: mecanismo que permite que os neurônios excitados aumentem seus valores individuais da função discriminante relacionada ao padrão de entrada aplicando ajustes a seus pesos sinápticos. Esses ajustes permitem que em uma posterior apresentação de um padrão de entrada semelhante o neurônio obtenha uma resposta melhorada.

A localização espacial de um neurônio de saída em um mapa topográfico corresponde a um domínio ou característica particular do dado retirado do espaço de entrada (HAYKIN; ENGEL, 2001).

Áreas com alta probabilidade de ocorrência de padrões de amostra serão representados por áreas maiores no mapa. Após o processo de aprendizagem um conjunto de vetores peso é atualizado dentro da vizinhança do neurônio vencedor. A vizinhança é determinada de modo topológico e decresce com o tempo. Dado o neurônio vencedor J sobre a apresentação de um padrão de entrada x , sua vizinhança de atualização Ω_j inicia com um amplo campo e gradualmente encolhe com o tempo até que não haja mais neurônios, ou seja $\Omega_j = \emptyset$. Correspondentemente, o paradigma de aprendizado transita de aprendizado por competição suave, que atualiza a vizinhança de neurônios, para aprendizado por competição forte, que atualiza apenas o vencedor. A equação de atualização do neurônio j na iteração t

é escrita como 2.6;

$$w_j(t+1) = \begin{cases} w_j(t) + \eta(t)(x - w_j(t)), & \text{se } j \in \Omega_j(t), \\ w_j(t), & \text{se } j \notin \Omega_j(t) \end{cases} \quad (2.6)$$

onde $\eta(t)$ é a taxa de aprendizado decrescente que pode ser substituída pela função de vizinhança $h_{Jj}(t)$, que resulta na seguinte equação 2.7:

$$w_j(t+1) = w_j(t) + h_{Jj}(t)(x - w_j(t)) \quad (2.7)$$

É comum que a função de vizinhança seja Gaussiana o que é apropriado para representar a interação biológica lateral na forma de uma curva de sino

$$h_{Jj} = \eta(t) \exp\left(\frac{-\|r_J - r_j\|^2}{2\sigma^2(t)}\right) \quad (2.8)$$

onde r_J e r_j representam as posições dos neurônios correspondentes na grade e $\sigma(t)$ é a função decrescente de tamanho de núcleo. Uma possível escolha dessa função é

$$\sigma(t) = \sigma_0 \exp(-t/\tau) \quad (2.9)$$

onde σ_0 é o valor inicial e τ é a constante de tempo.

Os passos básicos do SOM são:

1. Determine a topologia e inicialize o vetor de pesos $w_j(0)$ para $j = 1, \dots, K$ aleatoriamente;
2. Apresente um padrão de entrada x a rede, escolha o nó vencedor J que tenha a menor distância para x , ou seja

$$J = \arg_j \min(\|x - w_j\|) \quad (2.10)$$

3. Calcule a taxa de aprendizado atual e o tamanho da vizinhança;
4. Atualize o vetor de pesos de todos os neurônios na vizinhança de J utilizando a equação 2.7
5. Repita os passos de 2 até 4 até que a mudança da posição do neurônio esteja abaixo de um número positivo pequeno predeterminado.

A técnica possui como vantagens aproximação da densidade do espaço de entrada e independência da ordem de padrões de entrada em lote. Já a necessidade de

determinar as dimensões da grade de entrada *a priori* e a ineficiência de manipular ruído pode limitar sua aplicação em situações reais. Um SOM treinado pode sofrer má representação da densidade do espaço de entrada, onde áreas de baixa densidade podem estar excessivamente representadas e áreas de grande densidade inferiormente. Adicionalmente, quando a dimensão dos padrões de entrada são maiores que a dimensão da saída da rede a topologia não é preservada perfeitamente (XU; WUNSCH; SOCIETY, 2009).

2.4.3 O Growing Hierarquical Self-Organizing Map

O *Growing Hierarquical Self-Organizing Map* (GHSOM) é uma modificação da tradicional rede SOM (KOHONEN, 1982) que busca solucionar duas deficiências encontradas nela. A primeira delas é a utilização de uma arquitetura de rede fixa em termos de número e organização dos elementos de processamento neural, que deve ser definido antes do treinamento. Em casos onde a quantidade de dados é muito ampla e desconhecida, essa tarefa torna-se muito difícil e os resultados da rede dependem diretamente dessa etapa. A segunda é que relações hierárquicas entre os dados de entrada não são refletidas satisfatoriamente por estarem inseridas no mesmo espaço representacional, dificultando sua identificação.

Desse modo, é interessante que uma rede seja capaz de determinar a quantidade e organização das unidades durante o treinamento e que seja capaz de refletir relações hierárquicas entre os dados, uma vez que essas relações podem ser observadas em uma ampla gama de aplicações. A rede GHSOM foi proposta em (DITTENBACH; MERKL; RAUBER, 2000) com a intenção de endereçar ambas deficiências.

O GHSOM possui uma estrutura hierárquica de múltiplas camadas, onde cada uma delas consiste de um número independente de SOMs. Iniciando do nível ao topo, cada mapa cresce em tamanho para representar uma coleção de dados em um nível específico de detalhes. Após melhorias relacionadas a granularidade da representação dos dados ser atingida, as unidades são analisadas, buscando-se verificar se um nível específico mínimo de granularidade foi atingido. As unidades que representam dados muito diversos são expandidas para formar um novo mapa em uma camada subsequente, onde os dados serão representados com maior nível de detalhamento. Esses mapas continuam crescendo em tamanho até que uma melhoria na qualidade de representação dos dados seja atingida. As unidades que já representam dados de modo homogêneo não necessitam de expansão em novas camadas. Desse modo, o GHSOM resultante será totalmente adaptado para refletir através de sua arquitetura a estrutura hierárquica inerente aos dados (RAUBER; MERKL; DITTENBACH, 2002).

O algoritmo inicia com uma camada virtual 0, que possui apenas uma uni-

dade. O vetor de pesos dessa unidade é inicializado como a média de todos os dados de entrada e o Erro Médio de Quantização (MQE) da unidade, ou mqe_0 , também é calculado. A expressão para o cálculo do mqe_i é dada em 2.11 como a distância Euclideana média entre o vetor modelo m_i e os n_c vetores de entrada x_j que são elementos do conjunto de vetores de entrada C_j , mapeados nessa unidade i :

$$mqe_i = \frac{1}{n_c} \cdot \sum_{x_j \in C_i} \|m_i - x_j\|, n_c = |C_i|, C_i \neq \emptyset \quad (2.11)$$

O ponto de partida do aprendizado se dá com o cálculo do mqe_0 da unidade na camada 0, como pode ser visto em 2.12. n_x é o número de todos os vetores de entrada x do vetor de entrada X e m_0 denota a média dos dados de entrada:

$$mqe_0 = \frac{1}{n_x} \cdot \sum_{x_i \in X} \|m_0 - x_i\|, n_x = |X| \quad (2.12)$$

O mqe mede a dissimilaridade entre as entradas mapeadas em uma unidade particular e é utilizado para controlar o processo de crescimento da rede neural, que no GHSOM é definido como uma fração τ_2 do mqe_0 . Esse cálculo é chamado de critério global de parada, como visto na expressão 2.13:

$$mqe_i < \tau_2 \cdot mqe_0 \quad (2.13)$$

As unidades que não satisfazem essa equação devem ter unidades adicionadas de forma a fornecer mais espaço no mapa para representação. O erro de quantização qe , como visto em 2.14, pode ser utilizado para ao invés do mqe como critério global de parada, guiando o processo de treinamento. Seu cálculo é realizado como a soma das distâncias entre o vetor de pesos de uma unidade i e os vetores de entrada mapeados nessa unidade, como visto em 2.15:

$$qe_i = \sum_{x_j \in C_i} \|m_i - x_j\| \quad (2.14)$$

$$qe_i < \tau_2 \cdot qe_0 \quad (2.15)$$

A utilização do qe , ao invés do mqe , como medida de qualidade global do GHSOM é uma escolha que segue melhor o princípio característico do SOM de fornecer mais espaço no mapa para regiões mais densamente populadas. Desse modo, utilizar a expressão 2.15 ao invés da 2.13 resulta em uma distribuição de dados mais fiel que mapeia diferenças mais finas em *clusters* densamente populados (RAUBER; MERKL; DITTENBACH, 2002).

O treinamento inicia com um mapa de dimensões 2×2 na camada 1, que

segue as regras do algoritmo SOM padrão. Um padrão de entrada é selecionado aleatoriamente e apresentado a rede neural. Cada unidade determina sua ativação de acordo com a distância Euclideana entre seu vetor de pesos e o vetor de entrada. A unidade vencedora, ou que apresenta menor distância, e sua vizinhança são adaptados, ou seja, as diferenças entre os componentes do vetor são diminuídas, sendo o vencedor mais semelhante ao padrão de entrada após a adaptação. Esse processo de treinamento é repetido uma quantidade λ de vezes, ou iterações. Após essas repetições a unidade com maior desvio entre seu vetor de pesos e o padrão de entrada é selecionada como a unidade de erro (e), seguindo o cálculo apresentado em 2.14. Entre a unidade de erro e seu vizinho mais diferente (d), em termos do espaço de entrada, uma nova linha ou coluna de unidades é inserida.

Esse processo de crescimento da rede pode ser descrito formalmente da seguinte maneira. Seja C_i o subconjunto de vetores x_j dos dados de entrada que são mapeados na unidade i , como $C_i \subseteq X$, e m_i o vetor modelo da unidade i . Então, a unidade de erro e é determinada como a unidade com o máximo qe como visto na equação 2.16:

$$e = \arg \max_i \left(\sum_{x_j \in C_i} \|m_i - x_j\| \right), n_c = |C_i|, C_i \neq \emptyset \quad (2.16)$$

Já a unidade vizinha d mais dissimilar a unidade de erro e é determinada como na expressão 2.17 onde N_e é o conjunto de unidades vizinhas da unidade e :

$$d = \arg \max_i (\|m_e - m_i\|), m_i \in N_e \quad (2.17)$$

Após a inserção da nova linha ou coluna os vetores de peso dessas novas unidades são inicializados com a média de seus vizinhos. Após a inserção a taxa de aprendizado e de vizinhança são reiniciadas com seus valores originais. Essa estratégia é muito semelhante a utilizada em outra adaptação da rede neural SOM que também aumenta de tamanho inserindo linhas ou colunas durante o aprendizado, o *Growing Grid* (FRITZKE, 1995). A diferença é que no GHSOM a taxa de aprendizado e a de vizinhança são decrescentes, ao invés de valores fixos, além do *Growing Grid* produzir apenas um mapa SOM (DITTENBACH; MERKL; RAUBER, 2000), (RAUBER; MERKL; DITTENBACH, 2002).

Conforme unidades vão sendo adicionadas ao GHSOM, seus qe s diminuem e quanto menor o valor do qe , melhor o mapa está treinado. Um mapa cresce até que todas as unidades satisfaçam o critério global de parada. De modo a explicitar a estrutura hierárquica dos dados, cada mapa explica apenas uma porção de similaridade dos dados. O processo de crescimento continua até que o erro médio de quantização do mapa, ou MQE seja reduzido

a um valor τ_1 de qe_u da unidade u na camada predecessora da hierarquia. Portanto, o mapa representa os dados mapeados na unidade u da camada acima em mais detalhes (DITTENBACH; MERKL; RAUBER, 2000). O MQE de um mapa é calculado como a média de todos os erros de quantização das unidades do subconjunto U de mapas em que as unidades estão mapeadas 2.18.

$$MQE_m = \frac{1}{n_u} \cdot \sum_{i \in U} qe_i, n_u = |U| \quad (2.18)$$

Em termos gerais, o critério de parada para o crescimento de um mapa único m é definido como 2.19, sendo qe_u o erro de quantização da unidade u correspondente na camada acima:

$$MQE_m < \tau_1 \cdot qe_u \quad (2.19)$$

Observa-se que quanto menor o parâmetro τ_1 escolhido maior será o mapa, explicando seus dados em uma alta granularidade. Para um τ_1 alto, uma representação de dados mais detalhada será delegada a mapas adicionais em hierarquias mais baixas. O parâmetro τ_1 serve como medida de controle para a profundidade da arquitetura resultante do GHSOM (RAUBER; MERKL; DITTENBACH, 2002).

Quando o treinamento de um mapa é finalizado de acordo com a expressão 2.19, cada unidade deve ser verificada pelo atendimento do critério global de parada 2.15. Caso os dados representados sejam muito diversos, eles são expandidos em uma nova camada, formando um novo mapa. Nessa nova unidade, é gerado novamente um novo mapa de 2×2 unidades, de maneira similar ao mapa da camada 1 originado da camada 0. A inicialização pode ser feita aleatoriamente, apesar disso usualmente distorcer a topologia geral de mapas vizinhos. De modo a fornecer uma orientação geral dos mapas individuais nas várias camadas da hierarquia, sua orientação deve estar conforme a orientação dos dados em seu mapa pai, o que pode ser obtido com uma inicialização coerente das unidades em um novo mapa. Considerando que a unidade p será expandida em um novo mapa 2×2 na próxima camada da hierarquia. Os quatro vetores e peso desse mapa, s_1 a s_4 serão inicializados de modo a espelhar a orientação das unidades vizinhas de p , o que é obtido através da adição de uma fração dos vetores peso na vizinhança de p (DITTENBACH; RAUBER; MERKL, 2001). Um exemplo da inicialização de novos mapas pode ser visto na figura 2.7.

No exemplo, um mapa 3×3 possui vetores peso a até i . Um exemplo de

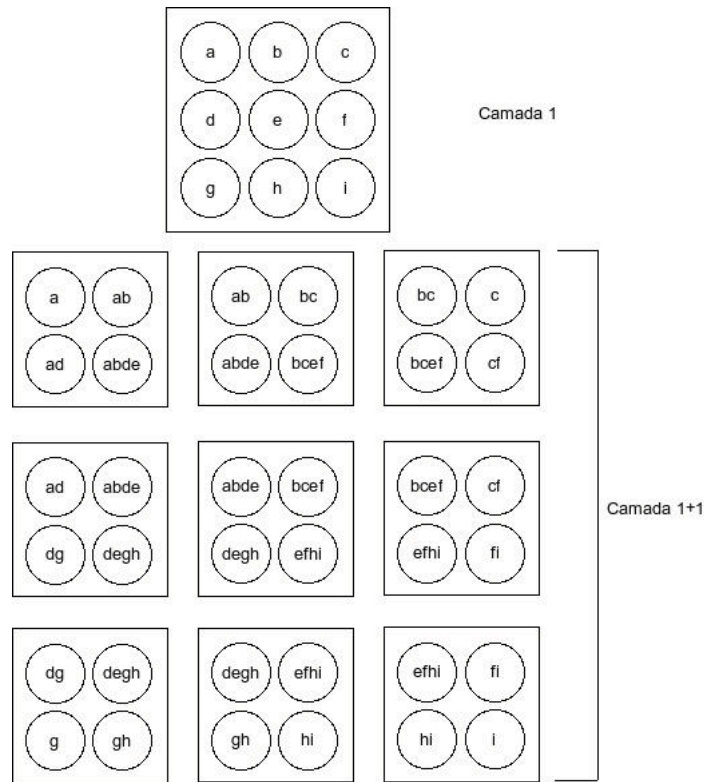


Figura 2.7: Inicialização das unidades para uma mapa expandido

inicialização para as unidades a e e pode ser visto em 2.20:

$$\begin{array}{l|l}
 a_1 = a & e_1 = e + \frac{(a-e)+(b-e)+(d-e)}{3} \\
 a_2 = a + \frac{(b-a)}{2} & e_2 = e + \frac{(b-e)+(c-e)+(f-e)}{3} \\
 a_3 = a + \frac{(d-a)}{2} & e_3 = e + \frac{(d-e)+(g-e)+(h-e)}{3} \\
 a_4 = a + \frac{(b-a)+(d-a)+(e-a)}{3} & e_4 = e + \frac{(f-e)+(h-e)+(i-e)}{3}
 \end{array} \quad (2.20)$$

O novo mapa se auto-organizará até que seu qe seja reduzido ao critério 2.19 (DITTENBACH; MERKL; RAUBER, 2000). Logo, o critério de parada do algoritmo é atingido quando não é mais necessário adicionar mapas na hierarquia, ou seja, os mapas atuais apresentam um nível de granularidade de representação de dados satisfatório (RAUBER; MERKL; DITTENBACH, 2002).

Experimentos mostram que além das vantagens previstas em projeto, outros resultados interessantes podem ser obtidos com a utilização do GHSOM, quando comparado a rede SOM tradicional. O tamanho total do mapa será geralmente reduzido, pois apenas quantidade necessária de unidades é criada em cada nível de representação. Além disso, os ramos não crescem na mesma profundidade em todos os nós da hierarquia. Com mapas menores, também é possível executar mais rapidamente o cálculo de ativação e avaliação de unidade vencedora (RAUBER; MERKL; DITTENBACH, 2002), resultando em um processo de treinamento mais rápido (SORIANO-ASENSI et al., 2008).

3 Materiais e Métodos

Para atingir os objetivos desse trabalho foram realizadas as seguintes etapas, detalhadas abaixo:

- Desenvolvimento da metodologia proposta para abordagem automatizada
 - Desenvolvimento de algoritmos de pré-clusterização;
 - Desenvolvimento do algoritmo de clusterização híbrida proposta;
 - Desenvolvimento de um protótipo de software que auxilie na análise de usabilidade;
- Seleção de uma aplicação alvo para avaliação de usabilidade;
- Validação do protótipo em ambiente real de uma aplicação Web da área da saúde;
 - Interpretação humana dos mapeamentos gerados pela clusterização;
- Realização de avaliação de usabilidade utilizando a metodologia *Discount Usability*;
- Verificação da consistência dos resultados de ambas as abordagens de usabilidade utilizadas para avaliação;
- Mensurar melhorias que o sistema alvo poderá obter ao aplicar as correções aos erros de usabilidade encontrados.

Para executar ambas avaliações de usabilidade foi necessário ter uma aplicação a ser analisada. Uma vez que o foco desse trabalho é em aplicações da área da saúde, a aplicação selecionada para análise opera nessa área e é chamada SacarWeb. Sua caracterização pode ser encontrada na seção 3.2.

A aplicação SacarWeb passou por duas análises de usabilidades com diferentes abordagens, a primeira utilizando as técnicas propostas pela *Discount Usability*, e a segunda utilizando a técnica de análise automatizada proposta por esse trabalho. O cenário de ambas as abordagens é detalhado nas seções 3.3 e 3.1 respectivamente.

3.1 Método híbrido proposto

A abordagem proposta neste trabalho possui a seguinte visão geral de seu modelo, como mostra a figura 3.1.

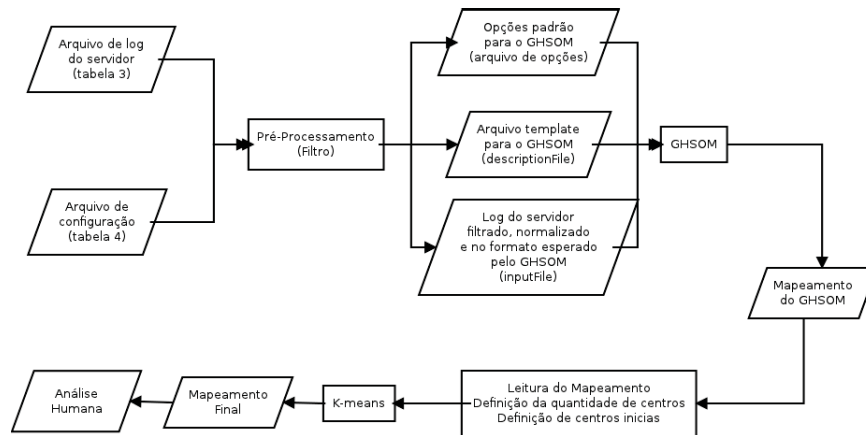


Figura 3.1: Modelo da abordagem híbrida proposta

Uma vez que foram utilizadas técnicas de *Web mining*, tarefas relacionadas ao pré-processamento dos dados dos arquivos de *log* do servidor foram necessárias, pois a literatura mostra que não é possível utilizar o arquivo sem nenhum filtro anterior ao processo de clusterização. Além disso, somente tarefas bem determinadas dentro da aplicação foram analisadas, e como todas as requisições de usuários são armazenadas nos arquivos de *log*, é necessário que boa parte delas seja descartada.

A proposta de um método híbrido se baseou na pesquisa apresentada em (KUO; HO; HU, 2002), (KUO et al., 2006) e (KUO; LIAO; TU, 2005). A rede neural GHSOM possui base nos trabalhos (DITTENBACH; MERKL; RAUBER, 2000) e (RAUBER; MERKL; DITTENBACH, 2002). Para o *K-means* utilizou-se o método clássico (LLOYD, 1982) com definição de centros iniciais. Adicionalmente, o método proposto nesse trabalho, baseado em análise de arquivo de *log*, está classificado com relação a avaliação de usabilidade automatizada como um método de teste, capaz de capturar e analisar dados com esforço necessário em níveis formais, informais e desenvolvimento de modelos, variando de acordo com o método empregado especificamente (IVORY; HEARST, 2001).

De modo semelhante a abordagem encontrada nos trabalhos (KUO; HO; HU, 2002), (KUO et al., 2006) e (KUO; LIAO; TU, 2005), o método proposto teve seus resultados comparados aos de outro algoritmo híbrido. Neste caso, a abordagem que utiliza GHSOM e *K-means* é comparada a outra que utiliza SOM e *K-means*. Estratégia semelhante foi realizada no trabalho (KUO et al., 2006) que utilizou métodos semelhantes em sua proposta híbrida. A

comparação entre os métodos também teve base em (KUO; HO; HU, 2002), que utilizou o valor lambda de Wilk para calcular a diferença entre os métodos.

3.1.1 Algoritmos de pré-clusterização

Apesar de não executar nenhuma tarefa que compreenda a construção de conhecimento, os algoritmos de pré-clusterização são de grande importância para que sejam atingidos resultados satisfatórios no processo de clusterização (KOSALA; BLOCKEEL, 2000), (SPILIOPOULOU, 2000), (SRIVASTAVA; DESIKAN; KUMAR, 2005). Recebendo os arquivos de *log* do servidor Web como entrada, os dados de navegação disponíveis serão abstraídos e gerarão um arquivo de saída. Esse arquivo será a entrada da abordagem híbrida proposta, devendo estar num formato em que seja possível adquirir conhecimento a partir dos dados contidos nele (SRIVASTAVA et al., 2000). Neste trabalho o arquivo que possui esse formato serve como entrada da rede GHSOM.

Os arquivos de *log* do servidor Web armazenam as requisições realizadas pelos usuários do servidor, armazenando informações como seu IP, recurso requisitado, horário do pedido, resposta do servidor, entre outras. Neste trabalho utiliza-se o formato de *log* do Apache HTTP Server¹, selecionado por ter uma licença flexível e por ser o servidor Web mais utilizado atualmente no mercado². Mais especificamente, utiliza-se o *access log* do Apache³, que armazena todas as requisições processadas pelo servidor. Cada requisição armazenada no arquivo geralmente possui o chamado formato comum com visto em 3.1:

```
%h %l %u %t "%r" %>s %b
```

Quadro 3.1: Formato comum do access log do Apache HTTP Web Server

onde:

Esse formato de *log* é semelhante ao observado em 3.2, onde a identidade da máquina cliente e o id do usuário não estão disponíveis:

```
127.0.0.1 -- [10/Oct/2000:13:55:36 -0700] "GET /apache_pb.gif
HTTP/1.0" 200 2326
```

Quadro 3.2: Exemplo do formato comum do access log do Apache HTTP Web Server

¹<http://httpd.apache.org>

²http://news.netcraft.com/archives/web_server_survey.html

³<http://httpd.apache.org/docs/2.2/en/logs.html#accesslog>

3.1 Método híbrido proposto

%h	Endereço de IP do cliente que fez a requisição ao servidor
%l	Identidade da máquina cliente que fez a requisição, normalmente não disponível
%u	Id de usuário que está realizando a requisição, normalmente indisponível
%t	Tempo em que a informação foi requisitada, no formato [dia/mês/ano:hora:minuto:segundo zona]
%r	Entre aspas duplas essa informação é a solicitação do cliente, que contém o método utilizado, o recurso solicitado e o protocolo utilizado
%>s	Código de status da resposta do servidor à requisição do cliente
%b	O tamanho do objeto retornado ao cliente

Tabela 3.1: Organização do arquivo de log comum do Apache HTTP Server

Também está disponível o formato combinado de *log* que possui informações adicionais sobre a requisição do usuário, como a versão do aplicativo (agente de usuário) e sistema operacional utilizado para acessar um recurso e, caso seja pertinente, a URL de origem da requisição atual (*Referer*). O formato de *log* estendido é semelhante ao exemplo em 3.3:

```
127.0.0.1 -- [10/Oct/2000:13:55:36 -0700] "GET /apache_pb.gif
HTTP/1.0" 200 2326 "http://www.example.com/start.html" "
Mozilla/4.08 [en] (Win98; I ;Nav)"
```

Quadro 3.3: Exemplo do formato estendido do access log do Apache HTTP Web Server

É possível configurar o servidor Web Apache para produzir qualquer um dos tipos de *log* disponíveis. Buscando uma maior abstração da entrada de dados, os algoritmos de pré-clusterização foram desenvolvidos para serem capazes de filtrar ambos formatos.

Utilizou-se a linguagem de programação Python⁴ para desenvolver os algoritmos de pré-clusterização. Sua principal entrada, como já mencionado, é o caminho para o *access log*. Uma segunda entrada é necessária, que é um arquivo com as configurações do programa que possui uma linha para cada tarefa que se deseja analisar, onde cada linha do arquivo deve estar no formato em 3.4:

```
ui uf mr id
```

Quadro 3.4: Formato padrão para o arquivo de configurações do filtro

onde:

A partir dessas duas entradas é iniciado o processo de filtragem, que é dividido nas seguintes etapas:

⁴<http://www.python.org/>

ui	URL inicial da tarefa
uf	URL final da tarefa
mr	Quantidade máxima de requisições entre as URLs inicial e final
id	String identificadora da tarefa

Tabela 3.2: Definição do arquivo de entrada com configurações

1. Eliminação de URLs irrelevantes para o objetivo da análise;
2. Identificação de sessões de usuários;
3. Criação de arquivos de entrada para a rede neural GHSOM.

Em cada etapa, foram observados pontos relevantes encontrados na literatura. Para eliminação de acessos irrelevantes foi preciso encontrar URLs que não representariam interações do usuário com a aplicação. Entre elas estão arquivos de imagens, folhas de estilo e *scripts* com processamento no lado do cliente. É comum eliminar esse tipo de recurso durante essa etapa, mantendo somente o arquivo HTML requisitado. Para essa tarefa pode ser utilizada a busca por sufixos na URLs contidas no *log*, abordagem utilizada nesse trabalho. Apesar desses recursos serem utilizados pelos usuários, não são acessados diretamente, mas referenciados pelas páginas acessadas. Sendo assim, os acessos a eles contidos no *log* não representam o comportamento de navegação de um usuário. Também pode ser utilizada outra técnica que é a eliminação ou preservação de URLs contidas em uma listagem fornecida ao programa (COOLEY; MOBASHER; SRIVASTAVA, 1999). Essa técnica é parcialmente utilizada neste trabalho. Isso pois todas entradas do *log* são eliminadas, exceto as que estão entre as *mr* requisições encontradas entre *ui* e *uf*.

Outro aspecto a ser observado é a determinação da existência de acessos importantes não localizados no arquivo de *log*, que podem estar em cache em servidores *proxy* ou localmente nos agentes dos usuários. Pode-se superar parcialmente esse tipo de problema utilizando *cookies*, registro explícito de usuário e *cache busting*, que é a prática de prevenir que os agentes de usuário utilizem versões armazenadas localmente de uma página, forçando um novo *download* do servidor a cada visualização. Nenhuma dessas soluções possui resultados expressivos (COOLEY; MOBASHER; SRIVASTAVA, 1997), (SPILIOPOULOU, 2000), (SRIVASTAVA et al., 2000) pois dependem do lado do cliente e podem ser desabilitadas pelos usuários. *Cookies*, por exemplo, podem ser eliminados pelos usuários, e *cache busting* elimina a vantagem pela qual o cache foi criado, além de ser possível desabilitá-lo no cliente. O registro de usuário seria a melhor solução, pois com ele é possível utilizar dados adicionais dos usuários e identificar com facilidade suas sessões. Entretanto, a maioria dos usuários não navega em aplicações

3.1 Método híbrido proposto

Web que exigem registro ou fornecem informações falsas (COOLEY; MOBASHER; SRIVASTAVA, 1999).

Servidores *proxy* também originam problemas com identificação de usuários. Apesar de não ser necessário identificá-los, é necessário distingui-los (SPILIOPOULOU, 2000). A maior parte dos formatos de arquivo de *log* armazena como únicas informações dos usuários o nome da máquina e o agente de usuário utilizado. Devido a isso, os seguintes casos podem ocorrer (SRIVASTAVA et al., 2000):

- Endereço de IP único - Múltiplas Sessões: provedores de Internet (ISPs) tipicamente possuem vários servidores *proxy* pelos quais os usuários acessam a Web. Um único endereço *proxy* pode ter diversos usuários acessando uma aplicação Web, potencialmente no mesmo período de tempo;
- Múltiplos endereços de IP - Sessão única: alguns ISPs ou ferramentas de privacidade associam aleatoriamente a cada requisição do usuário um novo IP, sendo assim uma única sessão de usuário pode ter vários IPs diferentes;
- Múltiplos endereços de IP - Usuário único: um usuário que acessa a Web de diferentes locais terá diferentes endereços IP de uma sessão para outra, o que torna difícil a detecção de visitas repetidas de um mesmo usuário;
- Múltiplos agentes de usuário - Usuário único: novamente um usuário que utiliza mais de um agente, mesmo que em um só local, é detectado como mais de um usuário.

É possível solucionar esse tipo de dificuldade utilizando a estrutura hierárquica da aplicação e heurísticas baseadas em agente do usuário, nome da máquina, endereço de IP e informações temporais (COOLEY; MOBASHER; SRIVASTAVA, 1997). Também é possível utilizar informações adicionais do arquivo de *log*, como a URL de referência, caso elas estejam disponíveis (SPILIOPOULOU, 2000). Uma das heurísticas, baseada em agente de usuário, assume que caso duas requisições de um mesmo IP venham de agentes de usuário diferentes logo elas são realizadas por usuários diferentes. Outra heurística utiliza o *log* de acesso estendido, buscando pelo *referrer*, e a topologia da aplicação Web. Caso uma página requisitada não seja atingível de nenhuma das páginas visitadas pelo usuário, então assume-se que é um segundo usuário navegando com o mesmo IP (COOLEY; MOBASHER; SRIVASTAVA, 1999). Mesmo sendo efetivas para alguns casos, essas heurísticas falham para comportamentos de navegação facilmente encontrados nos usuários.

Nesse trabalho o problema com *proxy* não é tratado, uma vez que a aplicação Web a ser analisada é acessada em rede interna sem necessidade de acesso a um *proxy*.

Outra tarefa realizada durante o pré-processamento é a identificação de transações ou sessões. Antes de realizar qualquer classificação dos dados é necessário agrupá-los em unidades lógicas como sequências de URLs acessadas por um usuário durante uma visita única a uma aplicação Web. Esse problema é semelhante ao problema de identificação de usuários únicos, sendo que a transação difere da sessão em tamanho, o que depende do critério adotado para identificá-las (COOLEY; MOBASHER; SRIVASTAVA, 1997), (SRIVASTAVA et al., 2000). Devido ao uso da ferramenta para análise de usabilidade, somente tarefas bem determinadas serão extraídas do *log*. Logo, as transações são bem definidas e específicas, e possuem tamanho determinado no arquivo de entrada com as configurações para o filtro através do parâmetro *mr*. Nesse trabalho uma transação é tratada como uma sessão. A sessão teve seu tempo máximo de ociosidade limitado em trinta minutos, como utilizado em (SRIVASTAVA et al., 2000), (FU; SANDHU; SHIH, 1999).

Outro ponto que pode ser endereçado nessa etapa é a filtragem de acessos de *spiders* automatizados. Acessos originados de aplicações como essas podem prejudicar a clusterização pois são capazes de gerar ruídos entre os acessos de usuários comuns. As principais técnicas de filtragem de acessos automatizados tentam identificar um comportamento não humano durante o acesso aos recursos da aplicação (SPILIOPOULOU, 2000). Esse tipo de acesso não é tratado pelo sistema, visto que a aplicação analisada possui acesso controlado cedido a grupos de pesquisa interessados na área de ciências da saúde.

Para boa parte das questões pertinentes durante a fase de pré-processamento deve-se considerar a privacidade dos usuários. Qualquer ação que identifique os usuários, seus acessos e sessões, que sempre são dados privados, deve ser evitada. Algumas alternativas de ações encontradas para resolver deficiências relacionadas a identificação de sessões não levam em conta essa medida, o que é algo prejudicial quando não se possui autorização dos usuários para utilização de dados identificados e pertencentes a eles.

A remoção de itens indesejados e a seleção dos pertinentes é feita com a seguinte política: verifica-se no arquivo filtrado onde se encontram URLs iniciais (*ui*) das tarefas definidas no arquivo de configuração. Definidos esses pontos deve-se procurar pela URL final (*uf*) definida para aquela tarefa dentro de um limite máximo de requisições (*mr*) permitido para tarefa. Ao se encontrar uma URL final ou ser atingida a quantidade máxima de requisições para uma tarefa, todas as URLs acessadas dentro do intervalo recém determinado fazem parte do caminho percorrido pelo usuário durante a execução da tarefa.

Os dados selecionados do arquivo de *log* são:

- IP de usuário;
- URL requisitada;
- Data da requisição;
- *Status* de resposta do servidor;
- Tamanho da resposta do servidor.

Todas informações são armazenadas em um banco de dados *sqlite*⁵ e posteriormente acessadas para construção das sessões de usuário para cada tarefa. As sessões são tratadas separadamente para cada tarefa levando em conta as definições do arquivo de configuração do filtro. Nessa aplicação uma sessão de usuário é o caminho percorrido por ele enquanto executa uma tarefa bem determinada. Esse caminho sempre possui início em uma URL inicial presente nas configurações fornecidas e pode terminar com sucesso na URL final prevista, ou em outra qualquer, disponível ou não na aplicação analisada, dentro do limite máximo de requisições para a tarefa em análise.

Outro fator que deve ser levado em conta durante a construção das sessões de usuário é o tempo máximo ocioso entre duas requisições. Caso esse tempo seja ultrapassado, a sessão atual em análise deve ser finalizada por inatividade (FU; SANDHU; SHIH, 1999). Se uma nova página requisitada posteriormente ao final da sessão inativa estiver presente no conjunto de URLs iniciais, uma nova sessão independente da anterior é iniciada.

Na saída da etapa de pré-processamento são gerados dois arquivos necessários para entrada do GHSOM desenvolvido como projeto da Universidade Tecnológica de Viena (DITTENBACH; MERKL; RAUBER, 2000). Esse sistema foi selecionado por ser uma solução bem documentada e por já ter apresentado boas soluções em outros experimentos com resultados encontrados na literatura. Mais detalhes sobre o pacote utilizado podem ser encontrados em 3.1.2.1.

Como já mencionado todo o módulo de filtragem foi desenvolvido utilizando a linguagem de programação Python. Inicialmente, suas partes constituintes foram desenvolvidas de modo procedural como programas independentes. Após um determinado ponto

⁵<http://www.sqlite.org/>

3.1 Método híbrido proposto

de desenvolvimento do código, sua complexidade aumentou, sendo necessário refatorar o programa alterando seu paradigma de programação para o Orientado a Objeto, de modo a facilitar o gerenciamento do código.

O objetivo esperado com essa medida foi atingido, uma vez que o código tornou-se mais homogêneo. A utilização do filtro também foi melhorada, pois os programas passaram a se comunicar melhor encapsulados em classes, sendo necessário realizar somente uma chamada ao método principal para que todo o filtro funcione, ao contrário do que ocorria na versão procedural onde cada programa deveria ser invocado separadamente. Um dos pontos que dificultava esse aspecto era a comunicação entre os programas, que deveria ser realizada manualmente através de parâmetros. Já na versão atual, métodos são invocados um após o outro automaticamente, sem a necessidade de se preocupar com os parâmetros que eles devem receber, como é possível visualizar na figura 3.2.

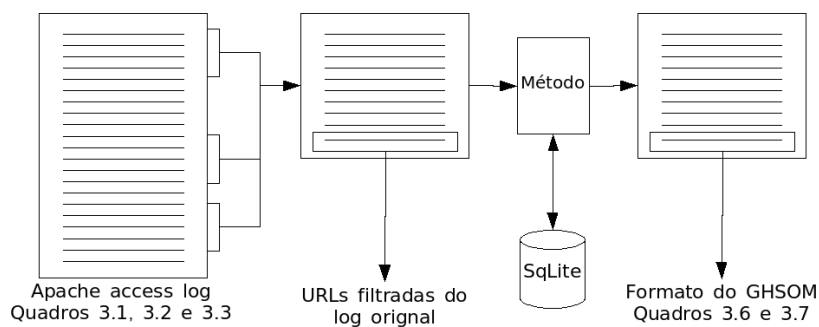


Figura 3.2: Fluxo de informações nos algoritmos pré-clusterização

Outro ponto notável foi a reutilização de código. Alguns pontos da implementação comuns entre as funções que constituem a abordagem automatizada estavam duplicados entre elas. Na reescrita para o padrão Orientado a Objeto esses trechos de código foram encapsulados em métodos, tornando a implementação mais generalizada.

3.1.1.1 Definindo os arquivos de entrada para o pré-processamento

Como já mencionado, a método híbrido proposto recebe como entrada o arquivo de *log* do servidor e um arquivo de opções. Somente o arquivo denominado *access log* do servidor Web Apache é aceito, independente se está no formato padrão ou estendido. Esse é o arquivo que registra todas as requisições processadas pelo servidor, e sua localização é dependente de sistema, podendo ser encontrada na documentação do servidor⁶. Ao encontrar o *access log* recomenda-se utilizar uma cópia somente com o período desejado para seleção como entrada, evitando conflito de utilização do arquivo com o servidor Web.

⁶<http://httpd.apache.org/docs/2.0/logs.html#accesslog>

3.1 Método híbrido proposto

O segundo arquivo de entrada, denominado arquivo de opções ou configurações, define as tarefas desejadas para análise. As tarefas devem ser definidas preferencialmente por uma pessoa com bom conhecimento da aplicação alvo, uma vez que conhecimento sobre o fluxo de dados e comportamento esperado do usuário é necessário para resultados mais fiéis a aplicação. Cada linha do arquivo constitui uma tarefa para análise, devendo seguir o seguinte formato apresentado no quadro 3.4. É possível visualizar exemplos para definição do arquivo de entrada no apêndice F. Ao definir esse arquivo será necessário informar a URL inicial ou onde a tarefa é considerada iniciada, URL final ou onde considera-se que a tarefa foi finalizada, a quantidade de URLs possíveis entre a URL inicial e final, levando em conta a URL final. A definição desse último parâmetro deve levar em conta outros recursos disponíveis na aplicação para o usuário, como ajuda e documentação, requisições realizadas de maneira assíncrona, ou recursos específicos da aplicação analisada. O último parâmetro é o nome que identificará a tarefa durante a análise automatizada.

Após definir ambos arquivos, o pré-processamento pode ser executado através da chamada semelhante ao quadro 3.5.

```
python filter.py -i access_log_path -c config_file_path [-f true  
-S]
```

Quadro 3.5: Formato padrão para chamada do do Pré-Processamento

É necessário ter o interpretador da linguagem python instalado para execução do programa. O parâmetro `-i` define o caminho do *access log*, `-c` define o caminho do arquivo com as tarefas a serem analisadas, ambos parâmetros são obrigatórios. Como parâmetros opcionais o programa fornece `-f`, que força a remoção de arquivos de cache de banco de dados de análises realizadas anteriormente e `-S` realiza a clusterização utilizando o SOM tradicional, e não o GHSOM. Ao final dessa etapa serão fornecidos três arquivos de entrada para o GHSOM, com nomes nos seguintes formatos:

- nome_da_tarefa.in.som
- nome_da_tarefa.prop
- nome_da_tarefa.t.som

3.1.2 Clusterização com GHSOM

Após a etapa de pré-processamento e com os arquivos que contém os dados de onde será extraído conhecimento, representando acessos retirados do *log* do servidor relativo

3.1 Método híbrido proposto

a tarefas específicas, pode ser iniciada a etapa de clusterização. Como já mencionado, a estratégia adotada é dividida em duas etapas, sendo a primeira compreendida pela rede neural não supervisionada *Growing Hierarchical Self-Organizing Map* (GHSOM).

3.1.2.1 O pacote GHSOM

Os autores do GHSOM fazem parte do Departamento de Tecnologia em Software da Universidade Tecnológica de Viena, e possuem um projeto direcionado ao desenvolvimento do GHSOM. Atualmente o projeto possui duas versões, uma que utiliza a linguagem de programação C++⁷, estável mas descontinuada, e outro implementado em Java⁸ batizado de SOMToolbox.

Uma vez que existe uma implementação confiável, pois utilizada em trabalhos com experimentos práticos de sucesso (DITTENBACH; MERKL; RAUBER, 2000), (DITTENBACH; RAUBER; MERKL, 2001) e (RAUBER; MERKL; DITTENBACH, 2002), é fornecida pelos próprios autores, com uma licença flexível e com fins de pesquisa, decidiu-se utilizar a implementação do GHSOM dos autores nesse trabalho. Apesar de descontinuada, essa versão oferece uma solução completa para o GHSOM, o que não acontece com o SOMToolbox. Além disso a versão disponibilizada é estável, ao contrário da versão disponível para o SOMToolbox, que está em constante desenvolvimento. Outro fator importante na escolha do pacote em C++ foi a apresentação de bons resultados encontrados na literatura como em (DITTENBACH; MERKL; RAUBER, 2000), (RAUBER; MERKL; DITTENBACH, 2002) e outros⁹.

Uma vez que a implementação do GHSOM de seus autores será utilizada, deve-se atender seus pré-requisitos para se obter resultados satisfatórios. Uma dessas necessidades é justamente a entrada do algoritmo, que nesse caso é gerada pelos filtros já abordados nesse trabalho. Seu arquivo de entrada define propriedades do algoritmo GHSOM, além de informar os arquivos que contém as entradas da rede neural. O arquivo de opções contém dados formatados como texto simples e deve seguir o padrão em 3.6:

```
propriedade1=valor1
propriedade2=valor2
propriedade3=valor3
...
```

Quadro 3.6: Formato padrão do arquivo de opções do GHSOM

⁷<http://www.ifs.tuwien.ac.at/andi/ghsom/index.html>

⁸<http://www.ifs.tuwien.ac.at/dm/somtoolbox/>

⁹<http://www.ifs.tuwien.ac.at/dm/publications.html>

3.1 Método híbrido proposto

Algumas das propriedades definidas no arquivo de opções são:

- Valor to τ_1 ;
- Valor do τ_2 ;
- Taxa de aprendizado inicial;
- Valores iniciais para X e Y , que devem ser iguais a 2 para o GHSOM;
- `inputFile` e `descriptionFile`, descritos a seguir.

As propriedades `inputFile` e `descriptionFile` devem possuir um caminho para um arquivo de texto plano. No caso de `inputFile`, o arquivo deve conter os vetores de entrada para a rede neural definidos nos campos entre “<>”, além de respeitar o formato em 3.7:

```
$TYPE <nomedovetor>
$XDIM <# de vetores de entrada>
$YDIM 1
$VECDIM <# de características>
<carac. 1> <carac. 2> ... <nome do vetor 1>
<carac. 1> <carac. 2> ... <nome do vetor 2>
.
.
.
```

Quadro 3.7: Formato do `inputFile` para o GHSOM

A primeira linha TYPE contém uma palavra única com o nome do vetor. XDIM é um inteiro com a quantidade de vetores no arquivo. YDIM possui sempre o valor 1. VECDIM é um inteiro com a dimensão do espaço característico. Desse modo, há uma linha para cada vetor, listando os valores dos atributos como números reais separados por espaço, seguidos por um *label*, que pode ser textual. A matriz formada será $VECDIM+1 \times XDIM$

Já o `descriptionFile` descreve as dimensões do espaço característico, e pode ser usado para associar etiquetas aos *clusters*. O formato deve seguir o padrão em 3.8:

```
$TYPE <nomedotemplate>
$XDIM 7
$YDIM <# de vetores de entrada>
$VECDIM <# de características>
```

3.1 Método híbrido proposto

```

0 <nome da carac. 1> <df> <tf> <min_tf> <max_tf> <mean_tf>
1 <nome da carac. 2> <df> <tf> <min_tf> <max_tf> <mean_tf>
2 <nome da carac. 3> <df> <tf> <min_tf> <max_tf> <mean_tf>
.
.
.
# carac. - 1 <nome da carac. n> <df> <tf> <min_tf> <max_tf> <
  mean_tf>

```

Quadro 3.8: Formato do descriptionFile para o GHSOM

Novamente TYPE é um identificador textual para o arquivo. XDIM sempre será 7, pois os vetores possuem 7 colunas de informação. YDIM é definido com a quantidade de vetores desejados para análise. VECDIM novamente possui a dimensão do espaço característico. As linhas após o cabeçalho de cada linha descrevem uma dimensão, com as colunas representando respectivamente, o número da dimensão iniciando em 0, o nome daquela dimensão como palavra única, a frequência do termo do atributo na coleção de dados, a frequência do documento do atributo, e a frequência mínima, máxima e média do termo. Caso o GHSOM não seja utilizado para análise textual, esses valores podem ser definidos para 1, pois não serão utilizados durante o treinamento.

Os dois arquivos, *inputFile* e *descriptionFile*, são gerados com os valores obtidos do *log* do servidor. Nesses arquivos estão inseridos somente as URLs acessadas durante uma sessão, mas é possível adicionar ainda o tempo de permanência na URL, resposta enviada pelo servidor além do tamanho da resposta. A utilização desses dados pode ser interessante para gerar *clusters* mais específicos, mas essa possibilidade não será explorada nesse trabalho. Já o arquivo de propriedades é gerado com valores padrão, que devem ser revisados durante a execução do processo para calibragem da rede neural GHSOM. Todos os vetores do arquivo *inputFile* possuem o mesmo tamanho, gerando uma matriz homogênea. Em casos onde a quantidade de acessos em uma sessão não é suficiente para completar uma linha, são adicionados valores -1, pois não interferem no treinamento do GHSOM.

Com os arquivos de entrada da rede GHSOM finalizados e com o arquivo de propriedades preenchido com os valores corretos, é possível executar o primeiro passo da abordagem proposta. O pacote GHSOM deve ser executado manualmente, informando o arquivo de propriedades como parâmetro. Ao final do processamento, o pacote gera arquivos de saída no diretório definido pela variável *savePath* no arquivo de propriedades.

3.1 Método híbrido proposto

Os arquivos de saída do GHSOM são divididos em no mínimo três tipos, com diferentes características do mapa em cada um deles. Adicionalmente poderá ser gerado um quarto tipo de arquivo codificado em HTML contendo uma visualização do mapa final. Esse tipo de arquivo é gerado ao se definir a variável de entrada *saveAsHTML* como verdadeira no arquivo de entrada com as propriedades. Os outros arquivos são:

- Descrição do mapa
- Descrição da unidade
- Vetor de pesos

Esses tipos de arquivos são gerados para cada mapa na hierarquia do GHSOM. O arquivo com a descrição do mapa possui a estrutura básica do SOM, fornecendo todos os parâmetros utilizados para o treinamento, sendo inicialmente escrito como resultado do processo de treinamento do SOM, mas informações adicionais podem ser encontradas por outros programas. Esse arquivo possui o formato em 3.9:

```
#SOM Map Description File
#created by ghsom 1.4 (Growing Hierarchical Self-Organizing Map)
#Michael Dittenbach
#
$TYPE rect
$XDIM 6
$YDIM 5
$VEC_DIM 481
$STORAGE_DATE Thu Apr 5 16:19:14 2001
$TRAINING_TIME 3.519999981
$LEARNRATE_TYPE ghsom-specific (sawtooth)
$LEARNRATE_INIT 0.5
$NEIGHBORHOOD_TYPE ghsom-specific (sawtooth)
$NEIGHBORHOOD_INIT 3
$RAND_INIT 17
$ITERATIONS_TOTAL 5100
$NR_TRAINVEC_TOTAL 51
$VEC_NORMALIZED 1
$URL_TRAINING_VEC vectors/demo_1.tfxidf
$URL_WEIGHT_VEC output/som_demo1_a_1_1_0_0.wgt
```

3.1 Método híbrido proposto

```

$URL_MAPPED_INPUT_VEC  vectors/demo_1.tfxidf
$URL_NODE_DESCR  output/som_demo1_a_1_1_0_0.node
$DESCRIPTION  NONE
$GHSOM_EXPAND_CYCLES  100
$GHSOM_MAX_CYCLES  0
$GHSOM_TAU_1  1
$GHSOM_TAU_2  1

```

Quadro 3.9: Arquivo de descrição e estrutura básica do SOM

A descrição do mapa é iniciada com um bloco de comentários precedidos pelo caractere `#`. `TYPE` descreve a topologia do mapa, assim como `TOPOLOGY`. `XDIM` é um número inteiro com a quantidade de unidades na direção `x`, semelhante a `YDIM` que possui o valor da direção `y`. `VECDIM` é um inteiro que representa a dimensionalidade do vetor de pesos do mapa. `STORAGE_DATE` e `STORAGE_TIME` são campos opcionais com valores de tempo para data de armazenamento do mapa treinado e tempo de treinamento, respectivamente. `LEARNRATE_TYPE` é uma *string* com o tipo da taxa de aprendizado, também um campo opcional. `LEARNRATE_INIT` é um campo opcional com o valor inicial da taxa de aprendizado. `NEIGHBORHOOD_TYPE` é uma *string* opcional com o tipo de região de vizinhança e `NEIGHBORHOOD_INIT` também é um campo opcional de valor real com a abrangência inicial da vizinhança. `RAND_INIT` é um inteiro utilizado como valor inicial do gerador de números aleatórios. `ITERATIONS_TOTAL` é a quantidade de iterações do processo de treinamento. `ITERATIONS_BUFFERED` é a quantidade de iterações de um ciclo do processo de treinamento quando utilizada a leitura por *buffer*. `NR_TRAINVEC_TOTAL` define a quantidade de vetores de entrada utilizados para treinamento. `NR_TRAINVEC_BUFFERED` possui a quantidade de vetores de entrada utilizados para treinamento no ciclo quando utilizada a leitura de vetores de entrada com *buffer*. `VEC_NORMALIZED` indica se os vetores de entrada foram normalizados antes do processo de treinamento. `QUANTERROR_MAP` é o valor do erro de quantização do mapa. `QUANTERROR_VEC` possui o valor médio do erro de quantização dos vetores de entrada do mapa ou $QUANTERROR_MAP / NR_TRAINVEC_TOTAL$. `URL_TRAINING_VEC` é o caminho dos vetores de entrada utilizados para treinamento. `URL_TRAINING_VEC_DESCR` possui o caminho do diretório com a descrição dos vetores de entrada utilizados no treinamento. `URL_WEIGHT_VEC` é o caminho do vetor de pesos do mapa treinado. `URL_QUANTERR_MAP` contém o caminho para o arquivo contendo o erro de quantização do mapa treinado. `URL_MAPPED_INPUT_VEC` são várias strings com os caminhos dos arquivos contendo vetores mapeados no mapa treinado. `URL_MAPPED_INPUT_VEC_DESCR` contém os caminhos para os arquivos

3.1 Método híbrido proposto

com as descrições dos vetores de entrada mapeados no mapa treinado. URL_UNIT_DESCR possui a URL do arquivo que contém a descrição das unidades do mapa treinado.

Já o tipo de arquivo com descrição das unidades do SOM consiste de dois blocos. O primeiro descreve a estrutura geral do SOM e o segundo fornece uma descrição específica para cada unidade. Um pequeno trecho de exemplo desse tipo de arquivo segue o exemplo em 3.10:

```
#SOM Node Description File
#created by ghsom1.4 (Growing Hierarchical Self-Organizing Map)
#Michael Dittenbach
#
$TYPE rect
$XDIM 6
$YDIM 5
$POS_X 0
$POS_Y 0
$NODE_ID 1_1_0/0_0/0
$QUANTERROR_NODE 0.9115136862
$NR_VEC_MAPPED 3
$MAPPED_VECS prsr/hst/mer_fns97.html.idv prsr/hst/mer_wirn97.
    html.idv
$NR_SOMS_MAPPED 0
```

Quadro 3.10: Arquivo com a descrição das unidades do SOM

Os valores que descrevem as unidades podem se repetir várias vezes, no caso do exemplo podem existir até 30 repetições desses blocos. As informações possíveis de serem encontradas nesse arquivo são no cabeçalho são: TYPE, com a topologia do mapa, XDIM, com a quantidade de unidades na direção x e YDIM com a mesma informação para direção Y. O bloco de informações pode conter as seguintes informações sobre cada unidade: POS_X, inteiro com a coordenada x da unidade e POS_Y, outro inteiro com o valor da posição y da unidade. UNIT_ID é uma *string* com uma etiqueta para unidade. QUANTERROR_UNIT armazena o erro de quantização da unidade. QUANTERROR_UNIT_AVG possui o erro de quantização médio do vetor de entrada da unidade. AC_POS_X contém a coordenada x da unidade na visualização AC do SOM, assim como AC_POS_Y possui a coordenada y. UMAT_UNIT possui a distância média para representação U-Matriz da unidade. UMAT_RIGHT armazena a distância do vizinho a direita para representação em U-Matriz, UMAT_DOWN armazena a distância

3.1 Método híbrido proposto

média do vizinho de baixo para representação de U-Matriz no caso de mapas hexagonais ou a distância do arranjo para o vizinho de baixo para representação da U-Matriz no caso de mapas retangulares. UMAT_DOWN_LEFT e UMAT_DOWN_RIGHT armazenam os mesmos valores para os cantos inferior esquerdo e direito, respectivamente. NR_VEC_MAPPED possui o número de vetores de entrada mapeados na unidade. MAPPED_VECS é uma listagem de VEC_IDs de vetores de entrada mapeados na unidade. MAPPED_VECS_DIST possui as distâncias pelas quais os vetores são mapeados na unidade. NR_SOMS_MAPPED armazena a quantidade de SOMs mapeados na unidade. URL_MAPPED_SOMS é uma listagem com os caminhos para os arquivos de descrição dos mapas. MAPPED_SOM_DIST é uma lista com as distâncias pelas quais os vetores SOM são mapeados na unidade, no caso do GH-SOM o erro médio de quantização da unidade. NR_UNIT_LABELS é o número de etiquetas para unidade. UNIT_LABELS é uma listagem de etiquetas, UNIT_LABELS_QE possui o erro de quantização para elas, UNIT_LABELS_WGT possui seus pesos, UNIT_LABELS_LEFT, UNIT_LABELS_RIGHT, UNIT_LABELS_UP e UNIT_LABELS_DOWN possuem uma listagem de labels da unidade e UNIT_LABELS_LEFT_DIFF, UNIT_LABELS_RIGHT_DIFF, UNIT_LABELS_UP_DIFF e UNIT_LABELS_DOWN_DIFF possuem a diferença das etiquetas para o vizinho esquerdo, direito, superior e inferior, respectivamente. URL_RELATED_UNITS possui uma lista dos caminhos das unidades relacionadas. Por fim, DESCRIPTION pode ser uma descrição textual da unidade.

Finalmente, o arquivo com o vetor de pesos descreve os pesos dos vetores treinados pelo SOM. O arquivo possui dois blocos, sendo o primeiro com a estrutura geral do SOM e o segundo com os vetores de pesos. Um exemplo do formato desse tipo de arquivo segue em 3.11:

```
#SOM Weight Vector File
#created by ghsom 1.4 (Growing Hierarchical Self-Organizing Map)
#Michael Dittenbach
#
$TYPE rect
$XDIM 6
$YDIM 5
$VEC_DIM 3
0.01748991385 1.82042394e-24 9.309764873e-14 1_1_0/0_0/0
0.06499654055 2.154417392e-24 7.153842102e-11 1_1_0/0_1/0
```

Quadro 3.11: Arquivo com vetor de pesos do SOM

Os valores do cabeçalho são: TYPE, com a topologia do mapa, XDIM com a quantidade de unidades na direção x, similar a YDIM com a quantidade de unidades na direção y e VEC_DIM com a dimensão dos vetores de peso do mapa. Os valores do conteúdo do arquivo seguem o padrão em 3.12:

```
<x_1_1> ... <x_1_n> <label_1>
...
...
...
<x_m_1> ... <x_m_n> <label_m>
```

Quadro 3.12: Padrão para os valores dos vetores do arquivo de pesos

Esse padrão é uma lista de n elementos de vetor para m vetores, onde $m = XDIM \cdot YDIM$, com valores reais seguidos de uma etiqueta para o vetor de pesos.

Com base nesses arquivos gerados na saída do GHSOM deverá ser possível utilizar o algoritmo *K-means*, que compreende a segunda etapa da estratégia proposta para clusterização dos dados.

3.1.2.2 Executando a clusterização com GHSOM

A segunda parte da abordagem é constituída pela execução do programa responsável pela clusterização com GHSOM. Caso tenha sido escolhida a clusterização com SOM na primeira etapa, os passos a serem executados permanecem os mesmos, pois os arquivos de entrada contém as informações necessárias para execução dos diferentes algoritmos. Os dados de entrada do algoritmo estão contidos nos arquivos resultantes do pré-processamento. Apesar de três arquivos serem gerados, recomenda-se alterar somente um deles, pois os outros estão em formato esperado pelo programa e não possuem informações de calibração da rede neural. O único arquivo que recomenda-se alteração possui nome no formato nome_da_tarefa.prop. Esse arquivo define os valores dos parâmetros de entrada para o GHSOM. Por padrão, ele está preenchido com valores recomendados pelos autores do algoritmo, e pode ser utilizado sem nenhuma alteração. O formato do arquivo segue o padrão do quadro 3.6 e os possíveis parâmetros são encontrados na documentação oficial do pacote ¹⁰. Por padrão, os dados de saída do ghsom são armazenados em um diretório chamado *output* dentro do diretório de trabalho. Esse parâmetro pode ser alterado, mas o diretório deve existir. Recomenda-se remover arquivos do diretório de saída de execuções realizadas anteriormente, pois estes podem

¹⁰http://www.ifs.tuwien.ac.at/andi/ghsom/download/ghsom_guide-1.6.html

interferir nos resultados. Para execução do GHSOM execute o comando seguindo o formato do quadro 3.13.

```
./ghsom nome_da_tarefa .prop
```

Quadro 3.13: Formato padrão para execução do GHSOM

O parâmetro `nome_da_tarefa.prop` deve ser substituído pelo nome definido no arquivo de configuração do pré-processamento. Ao final dessa etapa os arquivos necessários para execução do *K-means* estarão no diretório de saída definido no arquivo de configuração.

3.1.3 Aplicação do K-means

Como já mencionado, o algoritmo de clusterização *K-means* compreende a segunda etapa da abordagem híbrida proposta para a tarefa de *Web mining*. A partir do mapeamento resultante do GHSOM a entrada do *K-means* é definida. Particularmente, dado qualquer algoritmo de clusterização, o *K-means* pode ser aplicado como etapa de pós-processamento de modo a melhorar a distorção final (KANUNGO et al., 2002). Historicamente, o método possui duas deficiências que interferem diretamente no resultado apresentado ao final do processamento:

- Definição da quantidade de centros antes da execução do método;
- Definição de centros iniciais bons o bastante de modo a evitar mínimos locais.

Variações do método original vem sendo propostas de modo a superar essas deficiências. É possível que devido as várias estratégias que surgiram, o *K-means* seja um dos algoritmos mais amplamente estudados e utilizados (KANUNGO et al., 2002), sendo considerado inclusive o algoritmo padrão para clusterização (BRADLEY; FAYYAD, 1998). Neste trabalho é utilizada uma abordagem semelhante a (KUO; HO; HU, 2002) e (KUO; LIAO; TU, 2005), que utilizam uma rede neural antes do *K-means*, de modo a descobrir essas duas informações de maneira rápida, automatizada e confiável. Ao final da execução espera-se obter *clusters* estáveis e relevantes o suficiente para utilização em situações reais, como a proposta de experimentação desse trabalho, a detecção automática de falhas de usabilidade em aplicações Web da área da saúde.

3.1.3.1 Implementação utilizada do K-means

Durante a fase de filtragem de dados, a linguagem de programação Python foi selecionada como meio para desenvolvimento dos algoritmos. Nessa etapa ela também foi utilizada, mais especificamente devido as funções fornecidas pela biblioteca SciPy¹¹. Trata-se de uma biblioteca para computação científica, que oferece maior precisão no tratamento de ponto flutuante, representação numérica, entre outras características. Além disso, possui diversos métodos de diferentes áreas já implementados, como é o caso do *K-means*. A SciPy depende de outra biblioteca chamada NumPy¹², que é o pacote fundamental responsável por adicionar as características de maior precisão a biblioteca padrão da linguagem Python.

3.1.3.2 Executando a clusterização com K-means

Ao final da clusterização com GHSOM a primeira parte da abordagem híbrida está finalizada, e os arquivos necessários para executar a segunda parte, constituída pelo *K-means* foram todos gerados. Novamente será necessário utilizar o módulo de interpretação da linguagem Python, além dos módulos da linguagem Scipy e Numpy. Para execução dessa etapa, a linha de comando no formato do quadro 3.14 deve ser informada.

```
python pre_kmeans.py -c nome_da_tarefa.prop [-n -S]
```

Quadro 3.14: Formato padrão para execução do K-means

O parâmetro `-c` informa o arquivo de configuração da tarefa, com o nome composto pelo valor definido para nome da tarefa no arquivo de configuração do pré-processamento. Todos os outros parâmetros são opcionais, `-n` indica que o arquivo de entrada no *K-means* não deve ter seus valores normalizados. Recomenda-se utilizar essa opção somente quando o algoritmo falhar. O parâmetro `-S` indica que a etapa de clusterização anterior foi realizada com o SOM, e não com o GHSOM.

Ao final dessa etapa, serão gerados os mapeamentos de usuários para a tarefa além dos gráficos com a distribuição dos *clusters*. Os arquivos com os mapeamentos possuem nome no formato `nome_da_tarefa.prop.clusters` para o GHSOM e `nome_da_tarefa.prop.som.clusters` para SOM. Os gráficos não são salvos automaticamente, apenas gerados e exibidos para o usuário.

¹¹<http://www.scipy.org/>

¹²<http://numpy.scipy.org/>

3.1.4 Análise humana dos resultados do algoritmo híbrido

É necessário analisar os dados após a finalização da execução do programa. Apesar de algumas medidas serem calculadas automaticamente, é necessário que elas sejam contextualizadas para cada aplicação, uma vez que cada uma possui uma modelagem diferente. Devido a isso, recomenda-se que esse passo seja executado por uma pessoa com conhecimento da aplicação alvo da avaliação, preferencialmente a que definiu as tarefas analisadas. É possível que uma pessoa sem conhecimentos da aplicação realize a análise, mas o significado dos dados pode não ser aparente para ela.

Em termos gerais, recomendamos as seguintes etapas para análise dos resultados do método automatizado:

- Verificar os padrões apresentados nos *clusters* mais e menos representativos;
- Verificar se os *clusters* selecionados apresentam padrões esperados;
- Verificar desvios não previstos nos caminhos das tarefas;
- Verificar tempos de execução da tarefa não esperados;
- Verificar a representatividade dos padrões de acesso que coincidem com o esperado;
- Verificar padrões que possuam acessos a ajuda e documentação;
- Realizar outras verificações específicas para a aplicação em análise.

As etapas de análise humana recomendadas podem ser realizadas em qualquer ordem, entretanto a execução de algumas preliminarmente pode facilitar outras. Um caso é a verificação dos padrões mais e menos representativos. Os padrões com maior porcentagem dentro da análise representam o comportamento mais executado no sistema, e em caso de falha é um ponto a ser tratado no sistema. Já os com menor representatividade podem ser preocupantes caso representem padrões de sucesso, uma vez que os usuários não estão conseguindo realizar a tarefa da maneira prevista.

Desvios não previstos podem evidenciar um comportamento de adaptação dos usuários frente a um problema da aplicação alvo de teste ou uma maneira aparentemente mais confortável para o usuário realizar uma tarefa. Nesses casos deve-se verificar se o projeto da aplicação está de acordo com o público ou se o modelo para análise da tarefa deve ser refeito.

Os tempos de finalização de uma tarefa podem dizer muito sobre a tarefa. Tempos muito acima do esperado podem representar uma tarefa de difícil conclusão. Tempos muito baixos e repetidos podem ser a evidência de uma tarefa com erro conceitual, pois o usuário imagina que deve ser realizada de uma maneira e faz isso rapidamente várias vezes, remetendo a URLs de erro ou reiniciando a tarefa várias vezes.

Padrões com grande quantidade de acessos com tempo muito baixo podem representar desistência ou falta de compreensão da tela, uma vez que os usuários deixam a tarefa rapidamente após acessá-la.

Deve-se atentar para padrões que apresentam acessos de ajuda e documentação, pois além de mostrar que a tarefa pode estar difícil de interpretar, também mostra se a ajuda/documentação fornecida atende seu propósito. Isso pode ser verificado através do padrão de finalização da tarefa, em caso de sucesso é provável que a ajuda/documentação esteja atingindo seu objetivo, e em caso de falhas ela e a tarefa devem passar por revisões.

Devido as características únicas de cada aplicação, é possível que os dados após a clusterização apresentem ainda mais informações sobre o comportamento dos usuários e usabilidade da aplicação. Por esse motivo, etapas adicionais de análise dos dados podem ser executadas de modo a encontrar informações específicas relacionadas a aplicação. Essa característica evidencia ainda mais a importância do conhecimento da aplicação alvo da avaliação por parte do avaliador.

Adicionalmente, caso seja realizada uma avaliação tradicional de usabilidade, é possível verificar se há convergência entre esta e a avaliação com a metodologia proposta, de maneira semelhante a realizada por esse trabalho. Essa atividade pode validar ambas as avaliações e inclusive aumentar a quantidade de dados disponíveis para remodelagem da aplicação. A avaliação realizada por esse trabalho pode ser visualizada na seção 4.3.

3.1.5 Possibilidades para transcrição dos problemas de usabilidade em tarefas de desenvolvimento

Apesar de serem reconhecidamente um bom ponto de partida para correções de problemas em um software e gerar tarefas de desenvolvimento, problemas de usabilidade não são facilmente reconhecidos em uma aplicação por desenvolvedores (HORNBAEK; FRØK-JAER, 2005). Mesmo que sendo baseados em literatura e métricas bem conhecidas, é possível que os desenvolvedores que irão tratar o problema não tenham conhecimento para interpretar um problema de determinada interface. Por esse motivo a transcrição dos problemas em ta-

refas de desenvolvimento explícitas se faz necessária. Uma abordagem para essa atividade é encontrada no trabalho (JURISTO; MORENO; SEGURA, 2007), onde os autores associam 9 características comuns de usabilidade baseadas em literatura com a tarefas de desenvolvimento. Entretanto, o nível de informação proposto ainda não é suficiente e encontrado na maior parte das heurísticas e guias de usabilidade. Normalmente o problema pode ser solucionado aumentando a especificidade da informação através da adição de mais informações. Também surge o problema da transcrição dessas informações, normalmente recomenda-se que o grupo de desenvolvedores se reúna com profissionais de interação humano-computador para definir as tarefas. Entretanto, essa abordagem é normalmente custosa e não viável em empresas de pequenas dimensões. Devido a toda problemática, os autores sugerem um guia que incentiva os usuários a obter as necessidades funcionais dos problemas de usabilidade sem depender de especialistas de usabilidade. Primeiro foram extraídas e categorizadas as informações sobre características funcionais de usabilidade de diferentes autores de IHC, que resultou na lista de 9 características já mencionada. Informações mais específicas foram adicionadas as características de usabilidade, assim como o objetivo do problema sendo avaliado. Essas informações permitem que os diferentes *stakeholders* se comuniquem sobre o problema. Através de um experimento prático mostrou-se que esses guias ajudam os desenvolvedores a determinar se e quando um uma característica de usabilidade se aplica a um sistema em particular. Apesar dos resultados positivos, as características de usabilidade cobertas no trabalho não foram suficientes para fazer o *software* plenamente usável, e para solucionar essa questão os autores pretendem adicionar mais recomendações de usabilidade em trabalhos futuros.

O trabalho (HORNBAEK; FRØKJAER, 2005) cita que as propostas de novos projetos podem ser mais simples de entender, integradas diretamente ao projeto e mais estimulantes para desenvolvedores, apesar de não saber se as informações se concretizam na prática. O trabalho explora as diferenças entre a descrição de problemas de usabilidade e as propostas de novos projetos. O experimento prático realizado pelos autores demonstrou que propostas de novos projetos podem ser melhores que a descrição de problemas de usabilidade em termos práticos por serem mais concretas, mais fáceis de compreender para o desenvolvedor e mais úteis durante o desenvolvimento.

3.2 Caracterização da aplicação Web alvo do método híbrido proposto

O SacarWeb (Software para Avaliação Cardio-Respiratória na Web) é um software utilizado por um grupo de pesquisa em fisioterapia na Universidade Estadual de

Londrina (UEL) responsável pela reabilitação de pacientes com doenças crônicas respiratórias (CAMARGO-BRUNETTO; ROSSI, 2005). Essas doenças, em geral, precisam ser monitoradas continuamente para controle de ocorrência de crises. Com o objetivo de ajudar os profissionais a monitorar seus pacientes individualmente e estudar grupos de pacientes de acordo com características específicas o SacarWeb foi criado. Utilizando o software, estudantes, pesquisadores e professores podem ensinar e aprender conceitos, mantendo um banco de dados com a identidade dos pacientes, resultados de exames clínicos, e utilizá-los para observar a evolução dos pacientes assim como extrair grupos de pacientes com características de interesse.

Através do módulo de cuidados com o paciente, que possui as funcionalidades mais importantes do sistema, é possível coletar e analisar dados dos exames de pacientes. Cada avaliação consiste de um conjunto de exames. Os exames disponíveis são agrupados por categoria:

- Questionários para avaliar a qualidade de vida: *Chronic Respiratory Disease Questionnaire* (CRQ) e *St. George's Respiratory Questionnaire Paul Jones* (SGRQ) (JONES; QUIRK; BAVEYSTOCK, 1991);
- Testes de função pulmonar, para os quais o sistema fornece opções de diferentes equações de normalidade, incluindo o padrão brasileiro (TISIOLOGIA, 1996), americano de Knudson (KNUDSON et al., 1976) e o europeu da *European Respiratory Society* (MILLER et al., 2005);
- Exames para testar a capacidade de exercício: *Shuttle Walk Test* (PAYNE; SKEHAN, 1996), e o Teste dos 6 minutos de caminhada (TROOSTERS; GOSELINK; DECRAMER, 1999).

3.3 Aplicação do Discount Usability

Um dos objetivos é que tanto os testes empíricos quanto o automatizado proposto por esse trabalho auxiliem a aplicação a atingir os níveis de usabilidade definidos. Em linhas gerais, espera-se que a aprendizagem seja melhorada, a taxa de erros diminua e se possível que a taxa de memorização aumente, apesar desse item não ser avaliado.

Discount usability foi selecionada por ser uma abordagem que reduz custos e não necessitar de vasta experiência para ser aplicada. Além disso é um bom ponto de partida para equipes não habituadas a realizar análises de usabilidade em seu ciclo de desenvolvimento, podendo evoluir, ser refinada a necessidades específicas entrar definitivamente no ciclo de vida

3.3 Aplicação do *Discount Usability*

(NIELSEN, 1994a). Pelo fato da aplicação já estar em funcionamento, a única técnica não utilizada da *Discount Usability* será a de cenários, substituída por protótipo de alta fidelidade, no caso o próprio *software*. Essa decisão foi tomada pois o objetivo é medir a usabilidade do sistema atual, e não testar um projeto alternativo, caso em que seria mais interessante a utilização de cenários. Ao utilizar cenários para medir o sistema atual seriam necessários mais gastos, o que vai contra o objetivo desse trabalho. Após a aplicação das técnicas da *Discount Usability*, será realizado o teste automatizado com base nos dados de *log*.

Este trabalho realizou a primeira abordagem de usabilidade aplicada ao projeto SacarWeb. É de interesse do laboratório de desenvolvimento do projeto que a usabilidade seja levada em conta em projetos futuros. Os resultados encontrados pela *Discount Usability* foram cruzados com os resultados da análise automatizada proposta. Com isso pretende-se validar a proposta, além de verificar se ela complementa técnicas tradicionais ou guia/valida o que deve ser/for encontrado por elas.

Os protótipos do teste foram limitados a funções bem determinadas no sistema, para que fosse possível comparar os resultados dos testes empíricos com os testes da análise de *log*. Entretanto, o sistema testado já está em fase de produção e as funcionalidades e características do sistema não estarão limitadas, somente as tarefas a serem executadas pelo usuário.

As tarefas analisadas foram:

- Inserção de um paciente no sistema;
- Inserção da avaliação inicial de um paciente;
- Inserção de um dos seguintes exames: PFP (Prova de Função Pulmonar), Pressões Respiratórias, SGRQ (*St. George's Respiratory Questionary*), CRQ (*Chronic Respiratory Questionnaire*), Shuttle e DP6 (Teste de Caminhada de 6 minutos);
- Verificar a evolução do paciente;
- Realizar a aplicação de um filtro nos dados de modo que seja possível utilizar os resultados como base em uma pesquisa (relatório).

Os participantes do projeto LFIP, usuários da aplicação SacarWeb, são organizados de modo específico para as necessidades do projeto. Um usuário nunca é responsável por todos os tipos de exames executados, mas somente por um grupo de exames, como por exemplo os questionários de qualidade de vida. A experiência dos usuários com relação ao

conhecimento específico necessário para realizar os exames também é variável, uma vez que o projeto inclui alunos desde o segundo ano do curso de fisioterapia até doutorandos. É comum que os usuários com menos experiência utilizem o sistema com mais frequência, realizando a transcrição dos dados entre os formulários de exame impressos para o formato do SacarWeb.

Os recursos utilizados durante todas as avaliações também foram padronizados do seguinte modo:

- Utilização de dispositivo compatível com mouse e teclado;
- Utilização do navegador Mozilla Firefox versão 3.6.x com javascript e cookies habilitados;
- Resolução de tela mínima de 1024x768.

Durante as avaliações era recomendado evitar interferências externas, minimizando a presença de pessoas fora do escopo da avaliação no local do teste, evitando telefones ligados e utilizando somente o navegador, sem nenhum outro programa em execução.

3.3.1 Realização da avaliação heurística

A avaliação heurística foi realizada por três profissionais:

- Maria Angélica de Oliveira Camargo-Brunetto
- Carlos Henrique Ferraz
- Isabela Gasparini

A avaliadora Maria Angélica de Oliveira Camargo-Brunetto (AV_1) é graduada em Ciência da Computação pela Unicamp, possui mestrado e Doutorado em Ciência da Computação na área de Matemática Computacional pela Universidade Federal do Rio Grande do Sul (UFRGS). Atua como docente e pesquisadora na Universidade Estadual de Londrina (UEL), onde desenvolve e orienta projetos nas áreas de Informática aplicada em Saúde, aplicações Web adaptativas, técnicas de aprendizagem de máquina, entre outras. Foi selecionada para a avaliação por fazer parte do projeto do SacarWeb e ter o melhor conhecimento sobre a aplicação e seu domínio de conhecimento, o que é um grande diferencial uma vez que a aplicação é bastante específica. Outro fator que influencia na escolha é o fato de não ter experiência em avaliações de usabilidade, o que está relacionado aos objetivos desse trabalho.

3.3 Aplicação do Discount Usability

Carlos Henrique Ferraz (AV_2) é graduado em Artes Visuais - Multimídia; especialista em Comunicação Visual em Mídias Interativas; atuante no mercado de internet desde 2000, trabalhou na Agência Digital ICOMP por quase 4 anos, como Web Designer e Arquiteto de informação, criando e gerenciando a atual equipe de trabalho da agência formada por designers, desenvolvedores, programadores e redatores, hoje é responsável pelo desenvolvimento de projetos na FS Produções Artísticas, empresa que atua no setor de *show business*. Focado no desenvolvimento de portais e sites, destacam-se os trabalhos realizados para os principais artistas sertanejos, como Fernando e Sorocaba e Luan Santana, onde foram aplicados conceitos de Arquitetura de Informação e Usabilidade para facilitar o relacionamento entre os artistas e fãs, tendo seu trabalho reconhecido nacionalmente em prêmios e veículos de mídia online. Como principais projetos relacionados a usabilidade, foram aplicados a reestruturação de portais como Gramática Online, APP Londrina e Londrina *Convention Visitours Bureau*. Foi selecionado para avaliação por fazer parte da área de projeto de interfaces e ter experiência de 10 anos na área em projetos comerciais. Também não possui experiência na execução de testes de usabilidade, o que é interessante para esse trabalho, mas não faz parte da equipe do SacarWeb.

Isabela Gasparini (AV_3) atualmente é doutoranda em Ciência da Computação pela UFRGS na área de IHC (Interação Humano Computador), com qualificação sobre avaliação de usabilidade. Tem mestrado na mesma instituição e área. Professora da UDESC onde atua na área de IHC e usabilidade desde 2004, tanto no ensino (disciplinas), pesquisa (projetos, iniciação científica, TCC) e extensão (avaliação com usuários). Em 2006, juntamente com seus alunos de graduação, foi uma das equipes finalistas da “Competição de Avaliação” (nível graduação), parte integrante do Simpósio De Fatores Humanos em Sistemas Computacionais (IHC), patrocinado pela SBC. Neste trabalho a equipe utilizou 3 técnicas de avaliação, sendo uma delas a avaliação heurística¹³. Já realizou diversas avaliações de usabilidade na universidade, em especial, utilizando a avaliação heurística de Nielsen. Atualmente também desenvolve o Portal da Usabilidade¹⁴, onde pretende fazer dele um espaço para a comunidade de IHC trocar ideias, ser um canal de comunicação e inovação, além de questionar o que é usabilidade nos dias atuais. Foi selecionada para essa avaliação devido a sua experiência na área, sendo que os resultados de sua avaliação podem ser parâmetro de validade para as avaliações dos outros avaliadores.

Todas avaliações foram realizadas sob supervisão do mantenedor do projeto, que era responsável por fornecer assistência ao avaliador no caso de dúvidas relacionadas ao

¹³Detalhes em http://www.ufmg.br/swib/arquivos/chamada_da_competicao.pdf

¹⁴<http://portaldausabilidade.com.br/>

domínio da aplicação ou como operar o sistema. As informações eram fornecidas somente após o avaliador expor o problema de usabilidade em questão.

Para uma melhor compreensão das tarefas a serem avaliadas, foi fornecida aos avaliadores uma grande quantidade de dados para serem utilizados durante a avaliação. Cada avaliador deveria inserir dados de dois pacientes com três avaliações completas no sistema. Desse modo, cada tarefa seria realizada pelo menos duas vezes. Os dados fornecidos estavam em fichas de pacientes reais, com dados reais, fornecidos pelo laboratório LFIP.

Para realizar a avaliação os avaliadores seguiram as instruções contidas em um documento preparado pelo mantenedor do projeto, com informações sobre a avaliação e domínio de conhecimento da aplicação. Este documento pode ser visualizado no apêndice A. Um segundo documento (NIELSEN, 1992) apresentando as heurísticas e com um estudo prático sobre sua aplicação também foi fornecido de modo a familiarizar os usuários com a atividade.

As considerações dos avaliadores foram inseridas em um documento padronizado. Nesse documento todas as tarefas e as heurísticas eram fornecidas na ordem de execução das tarefas no sistema, de modo a facilitar o preenchimento e recordar os avaliadores das heurísticas e tarefas. Sendo assim, o avaliador deveria somente identificar o local correto para informar o problema encontrado. Esse documento pode ser visto no apêndice A na tabela A.1.

Os relatórios com as considerações dos avaliadores foram posteriormente analisados e condensados em um único relatório de falhas de usabilidade pelo mantenedor. Esse relatório foi novamente revisto pelos avaliadores que forneceram notas de gravidade aos problemas relatados nele, como mostrado em 2.1.3.1. Durante esse processo os avaliadores podiam utilizar o sistema com os mesmos dados de acesso fornecidos para avaliação, de modo a facilitar a avaliação da severidade do problema pois o avaliador poderia não se recordar da tela.

3.3.2 Realização do teste de usabilidade

Os testes de usabilidade foram conduzidos com os usuários alvo da aplicação, selecionados entre os participantes do projeto LFIP. Para o teste foi definido que os usuários não deveriam ter nenhuma experiência anterior com o SacarWeb e pouca experiência com o domínio da aplicação. Por esse motivo, somente alunos do segundo ano foram selecionados para participar dos testes. O avaliador utilizou o método *Simplified Thinking Aloud*, e por esse motivo o áudio do ambiente foi gravado durante a avaliação. Além disso, o avaliador produziu notas com base em suas percepções sobre a utilização do sistema. O avaliador também foi

3.3 Aplicação do Discount Usability

orientado a realizar anotações com base em sua experiência sobre o que pode ser bom ou ruim durante a interação dos usuários. Os testes de usabilidade foram executados no mesmo local em que os usuários utilizam o sistema, de modo a deixá-los mais à vontade com o ambiente.

Nessa avaliação os usuários foram convidados a inserir duas avaliações de um paciente no sistema, sendo que os exames inseridos eram somente os de especialidade do usuário. Novamente os dados eram reais e fornecidos pelo laboratório LFIP.

Antes de realizar os testes de usabilidade com os usuários alvo, foram realizadas duas avaliações piloto de modo a calibrar as avaliações. Os usuários que participaram dessas avaliações não faziam parte do grupo alvo, e não possuíam conhecimento específico da área. Por esse motivo, os conceitos mais importantes do sistema foram apresentados antes da avaliação, de modo a familiarizar os usuários com conceitos específicos e importantes para avançar pelo sistema.

Os resultados da avaliação piloto demonstraram que na segunda interação dos usuários havia um ganho expressivo de desempenho, nesse caso inserção de exames, diminuindo inclusive a quantidade de erros cometidos. Devido a isso, decidiu-se alterar a relevância dos dados considerados para as medições, visto que os da primeira interação refletem melhor o comportamento dos usuários para os objetivos da avaliação, principalmente aprendizagem. Esperava-se que o mesmo comportamento se repetisse com os usuários alvo, o que iria interferir nos resultados da avaliação. A intenção inicial era avaliar os dois conjuntos de dados para avaliar aprendizado e desempenho, mas devido ao comportamento apresentado na avaliação piloto, decidiu-se descartar a segunda interação do usuário para essas métricas.

Para o teste de usabilidade com usuários alvo foram selecionadas quatro pessoas as características desejadas, sendo a divisão das especialidades dentro do SacarWeb as seguintes:

1. Shuttle
2. PFP e Pressões Respiratórias
3. DP6
4. SGRQ e CRQ

Além de realizarem as tarefas, foi solicitado aos usuários que respondessem dois questionários, um antes da avaliação¹⁵, com o objetivo de identificá-los e obter dados

¹⁵<http://goo.gl/NzbUh>

sobre sua experiência anterior com computadores e conceitos da aplicação para elaboração de um perfil, e outro após¹⁶, cujo objetivo era obter dados relativos a satisfação dos usuários com relação ao SacarWeb.

Com base nos dados produzidos durante a avaliação, o avaliador deveria produzir um documento com suas considerações após a cada avaliação. Recomendou-se que esses documentos fossem produzidos em um intervalo pequeno após o término das avaliações, pois as informações ainda poderiam ser facilmente recordadas pelo avaliador. Novamente, esses documentos seriam revisados e compilados em apenas um documento reunindo todos os problemas encontrados durante os testes de usabilidade pelo mantenedor.

3.4 Verificação da consistência das abordagens de usabilidade utilizadas

Para validação do método desenvolvido realizou-se análises de *log* com dados reais de uma aplicação Web da área de cuidados com saúde chamada SacarWeb. A partir da análise dos *clusters* resultantes da abordagem híbrida foi possível realizar constatações relacionadas a usabilidade da aplicação, ou uma avaliação automatizada de usabilidade, de maneira semelhante a (SCHWERZ; MORANDINI; SILVA, 2007) e (MORANDINI, 2003). Uma avaliação de usabilidade baseada em *Discount Usability* foi realizada. O objetivo dessa avaliação é comparar os resultados alcançados pela abordagem automatizada, verificando se ambos convergem e se os resultados do software são capazes de auxiliar uma verificação de usabilidade, que valida o protótipo implementado. Adicionalmente, espera-se que com o desenvolvimento da ferramenta e estudo das técnicas utilizadas na *Discount Usability* a equipe do SacarWeb possa integrar atividades relacionadas a usabilidade no ciclo de vida de desenvolvimento do *software*.

O protótipo confronta padrões encontrados durante a clusterização com padrões esperados que deverão ser fornecidos pelos administradores da aplicação Web, o que é chamado de análise de arquivo de *log* baseada em tarefas (IVORY; HEARST, 2001), muitas vezes utilizada em conjunto com correspondência de padrões. Em (MORANDINI, 2003) mostrou-se que é possível encontrar dados relacionados a:

- tempo de execução de uma tarefa;
- caminhos comuns para realização de uma tarefa;
- quantidade de desvios em caminhos esperados;

¹⁶<http://goo.gl/g8xqG>

3.4 Verificação da consistência das abordagens de usabilidade utilizadas

- quantidade de erros e pedidos de ajuda na execução de uma tarefa;
- desistência na realização de tarefas;

Essas medidas estão relacionadas tanto a eficácia quanto a eficiência da aplicação, medidas importantes para determinação de uma boa usabilidade, mas não são limitadas a somente essas medidas.

Os resultados são apresentados pela aplicação em forma de gráfico, que contém os *clusters* e o percentual de representatividade com relação a amostra, e em formato textual com os caminhos de cada *cluster*, seu tempo médio e medidas de qualidade do *clusters* ao final.

A análise dos resultados é realizada através do tempo de realização da tarefa dentro do *cluster* e caminhos percorridos no *cluster*. Deve-se levar em conta as definições dos arquivos de entrada de cada tarefa, que definem a URL inicial e final da tarefa, além da quantidade esperada para que seja completada. Essa tarefa deve ser deixada para uma pessoa que tenha conhecimento profundo da aplicação sendo analisada, pois ela terá conhecimento suficiente para definir as tarefas satisfatoriamente. Na definição das tarefas é necessário atenção para o parâmetro que define a quantidade de URLs máxima para realização de uma tarefa, pois ela deve levar em conta possíveis consultas a ajuda, documentação ou páginas de erro intermediárias.

Os documentos de entrada produzidos para as tarefas analisadas da aplicação SacarWeb podem ser vistos no apêndice F. Uma maneira de definir esses documentos é verificar a URL da tarefa em sua primeira e última tela, verificando se ela passa por telas intermediárias e se existem caminhos alternativos - como consulta a ajuda e documentação, ou passos intermediários não obrigatórios - que podem ser executados. Durante a análise dos resultados é importante que o avaliador conheça cada tarefa, de modo a identificar URLs corretas ou não nos caminhos, tempos de finalização de tarefa esperados e quantidade de URLs ótima para finalizar uma tarefa. Essas características que guiam o avaliador para definir se um *cluster* representa sucesso, fracasso ou outro estado intermediário.

Os mesmos dados serão utilizados para comparação dos resultados entre a avaliação com *Discount Usability* e automatizada. Padrões de erro ou dificuldade identificados na avaliação tradicional podem ser relacionados a padrões com tempo alto de conclusão da tarefa, caminhos inesperados ou consultas a ajuda.

Além da validação através da análise de usabilidade, a estratégia híbrida desenvolvida será comparada a estratégia proposta em (KUO; HO; HU, 2002), o algoritmo

SOM + *K-means*, utilizando a mesma medida proposta nesse trabalho, o valor lambda de Wilk (λW). Essa medida é a razão entre a variância interna dos grupos (SSW) e a variância total (SST), sendo a sua fórmula definida como mostra a equação 3.1:

$$\lambda W = \frac{SSW}{SST} \quad (3.1)$$

Um valor alto para 3.1 implica que não há diferença nas médias internas dos grupos, enquanto um valor próximo de zero que a origem total da variância está entre os grupos. Desse modo, quando menor o valor da equação, melhores são os *clusters*, uma vez que homogeneidade interna e heterogeneidade externa são características significantes da clusterização (KUO; HO; HU, 2002).

Para chegar aos valores de λW cálculos intermediários são necessários. Primeiro a soma dos X , que é a soma das observações de cada grupo. Em seguida, é possível calcular a média simples de cada grupo, ou $Xbar$ e a média simples ao quadrado $Xbar^2$. Outro valor calculado é a soma dos X^2 , onde as observações de um grupo são somadas elevadas a segunda potência. Finalmente a soma dos quadrados é soma de todos $(X - Xbar)^2$.

Com esses valores é possível calcular a variância. SSW, ou variância interna aos grupos é calculada com todas as somas dos quadrados dos grupos, o grau de liberdade dos grupos - definido como a soma de $(n - 1)$ entre todos os grupos, que pode ser calculado também como o total de observações subtraído da quantidade de grupos. Com esses dois dados, SSW é definida como a soma dos quadrados interna dos grupos dividida pelo grau de liberdade interno dos grupos.

A variância entre grupos SSB é calculada com a variância da média, $S_{xbar}^2 = (\sum (Xbar - grande_media)^2) / (quantidade_grupos - 1)$. A grande média é calculada como a média das médias. Esse valor é transformado na variância entre grupos (SSB) ao se multiplicar o valor pela quantidade de itens em cada grupo.

A relação entre as variância pode ser calculada através do valor F, que é a razão entre SSB/SSW. Por fim, a variância total SST é calculada como $(X - grande_media)^2$. O desenvolvimento mais detalhado das equações pode ser encontrado em ^{17, 18, 19}.

Uma vez que o algoritmo SOM necessita da quantidade de neurônios antes do treinamento, definiremos esse valor como a quantidade de neurônios utilizada ou um valor

¹⁷<http://goo.gl/XBkKf>

¹⁸<http://goo.gl/yo1Fl>

¹⁹<http://goo.gl/MnEft>

3.4 Verificação da consistência das abordagens de usabilidade utilizadas

um pouco maior no treinamento do GHSOM, prática utilizada em (SORIANO-ASENSI et al., 2008). Os valores padrão para o arquivo de entrada que gera mapas SOM foram definidos conforme documentação encontrada em ²⁰

²⁰http://www.ifs.tuwien.ac.at/andi/somlib/download/somlib_guide.html

4 Resultados e Discussão

Este capítulo detalha os resultados obtidos pela avaliação realizada com a abordagem automatizada proposta pelo trabalho e a realizada com *Discount Usability*. Inicialmente os resultados são apresentados isoladamente, então as avaliações tradicionais realizadas são comparadas e por fim, verifica-se a convergência dos dois tipos de avaliação realizados no SacarWeb.

4.1 Resultados da Discount Usability

Essa seção apresenta os resultados alcançados através das avaliações realizadas pela abordagem proposta pela *Discount Usability*.

4.1.1 Resultados da avaliação heurística

Todos os problemas encontrados pelos avaliadores foram considerados no documento que os compila. Problemas comuns foram agrupados, mas as heurísticas indicadas pelos avaliadores mantidas. Os resultados da avaliação podem ser encontrados no apêndice B. Além dos problemas, suas avaliações de severidade são fornecidas. Alguns resultados da severidade são indicados por valores reais, os quais recomendamos utilizar a convenção de considerar o inteiro mais próximo como valor de severidade.

Uma análise mais detalhada dessas avaliações pode ser realizada através das tabelas 4.1 e 4.2.

A tabela 4.1 mostra que a avaliadora AV_1 foi responsável por 55.10% dos problemas encontrados. Acreditamos que esse valor é devido ao grande conhecimento do domínio da aplicação que a avaliadora possui, visto que é uma pessoa envolvida no projeto. Os outros dois avaliadores tiveram resultados bastante próximos, 21.77% e 23.13% para o avaliador AV_2 e AV_3 respectivamente. A quantidade de problemas comuns encontrados pelos três avaliadores representou 38.68%, um valor considerável que mostra que vários problemas

4.1 Resultados da Discount Usability

Tarefa	Qtd. problemas			$F = \sum_{n=1}^3 AV_n$	$C = \bigcap_{p \in F} X$	$F - C$
	AV_1	AV_2	AV_3	Final	Comuns	Compilada
1 - Cadastrar Paciente	13	5	5	23	1	21
2 - Avaliação Inicial	8	3	3	13	2	11
3 - Avaliação	8	5	5	18	6	12
4 - Shuttle	5	2	2	9	3	6
5 - SGRQ	10	3	9	22	11	12
6 - DP6	7	4	2	13	4	9
7 - CRQ	8	2	3	13	6	7
8 - PFP	11	3	2	16	4	12
9 - Pressões Respiratórias	2	2	1	5	1	4
10 - Evolução do Paciente	5	2	2	9	2	7
11 - Relatórios	4	1	1	6	1	5
Total	81	32	34	147	41	106

Tabela 4.1: Resultados encontrados durante as avaliações heurísticas

se mostraram de forma clara para todos os avaliadores. Uma vez que a quantidade total de problemas encontrados por todos avaliadores era 147 e desses 41 eram semelhantes, 106 problemas distintos foram encontrados na aplicação, o que representa 72.11% da quantidade original.

Heurística	Quantidade	Porcentagem
Diálogo natural e simples	23	17.69%
Fale a língua dos usuários	7	5.38%
Minimize a carga de memória do usuário	6	4.62%
Consistência	25	19.23%
Feedback	3	2.31%
Saídas claramente marcadas	13	10%
Atalhos	6	4.62%
Boas mensagens de erro	14	10.77%
Previna erros	22	16.92%
Ajuda e documentação	11	8.46%

Tabela 4.2: Representatividade de cada heurística para os problemas encontrados

Já a tabela 4.2 mostra que a maior parte dos problemas encontrados estão relacionados a **consistência** da aplicação. Esse tipo de constatação é preocupante, pois observamos que itens comuns entre as telas não se comportam da mesma maneira em todas elas, nem sempre possuem mesmo posicionamento e apresentação visual. Esse tipo de comportamento da aplicação não incentiva a confiança por parte dos usuários, dificultando a interação.

Observamos também que **Diálogo natural e simples** e **Prevenção de erros** são duas áreas tão críticas quanto consistência. Possivelmente devido a diálogos que poderiam ser mais simples para os usuários que acabam induzindo ao erro. Esse tipo de comportamento também foi percebido devido a falta de consistência de partes da aplicação.

Por fim, a avaliação da severidade dos problemas indicou 3 (2.83%) de problemas graves avaliados como graves por todos os avaliadores, 8 (7.55%) em que o valor final era um número real onde o valor 4 era o inteiro mais próximo. Desse modo, a quantidade total de catástrofes de usabilidade é de 11, ou 10.38%.

4.1.2 Resultados dos testes de usabilidade

Os resultados individuais de cada usuário foram avaliados independentemente, e depois condensados em somente um relatório, com problemas comuns agrupados em somente uma ocorrência. A pessoa responsável por essas tarefas foi a mesma que realizou os testes, ou o mantenedor da aplicação. O documento final levou em conta os problemas encontrados em todos os testes de usabilidade realizados, piloto e com usuários reais. Entretanto, os dados do teste piloto tiveram um peso menor ao se considerar o problema. Um problema ocorrido somente na avaliação piloto ganha uma observação com relação a isso, e deve ser considerado um problema somente caso também tenha ocorrido na avaliação com usuários da aplicação alvo.

Com base no formulário respondido pelos usuários antes da avaliação, o seguinte perfil de usuário alvo foi utilizado durante a avaliação: usuários predominantemente do sexo feminino, em média 19.75 anos, com conhecimento de computação regular e média de 8.67 anos usando o computador. Atualmente sua utilização é diária, em média 1:52 horas para estudo, pesquisa e entretenimento, em alguns casos trabalho. Todos os usuários fazem parte do projeto LFIP em média por 11.5 meses e estão no terceiro ano do curso de fisioterapia. Ao considerar também os usuários piloto, temos usuários de ambos os sexos, ainda com pessoas do sexo feminino predominantes, com média de idade alterada para 20.34 anos, conhecimento de computação variando de regular a bom e utilizando o computador a 8.8 anos. Atualmente o utilizam diariamente, em média 3:35 horas para estudo pesquisa e entretenimento, em alguns casos trabalho. Observa-se que os usuários de piloto realmente não são representativos comparados aos usuários reais, pois além de não terem ligação com a área de conhecimento específico do SacarWeb, possuem mais experiência com computadores e o utilizam por mais tempo para tarefas semelhantes. Essa conclusão é confirmada pelo aumento no valor das médias dos dados coletados.

4.1 Resultados da Discount Usability

Problemas que representaram grande dificuldades para os usuários foram classificados como problemas graves, e devem ter alta prioridade em tarefas de correção. Esses problemas foram identificados quando uma das situações ocorria:

- usuários necessitavam de ajuda do avaliador para continuar a tarefa;
- usuários não conseguiam cumprir a tarefa devido a um ponto específico;
- grande perda de desempenho devido a um ponto específico;
- usuários reais indicavam um erro de conceito, terminologia ou algo que pudesse prejudicar a interpretação dos resultados do exame.

Esses problemas são indicados através de *itálico*.

Adicionalmente foram listadas as boas práticas encontradas em cada tarefa, que podem ser utilizadas em todas as telas do programa de modo a padronizar a interação com base em boas experiências da própria aplicação. O documento pode ser visto no apêndice E.

Essa avaliação encontrou os valores contidos na tabela 4.3:

Tarefa	Qtd. de problemas	Qtd. problemas graves
1 - Cadastrar paciente	10	3
2 - Avaliação inicial	11	1
3 - Avaliação	12	4
4 - Shuttle	7	0
5 - SGRQ	6	2
6 - DP6	12	2
7 - CRQ	8	2
8 - PFP	13	3
9 - Pressões respiratórias	5	1
10 - Evolução do paciente	7	0
11 - Relatórios	16	9
Total	107	27

Tabela 4.3: Resultados encontrados durante os testes de usabilidade

Como é possível observar foram encontrados 107 problemas de usabilidade através da avaliação com os usuários, e destes 27 são considerados graves, o que representa 25.23% da quantidade total. Considerando que problemas graves podem diminuir o desempenho dos usuários, limitar sua taxa de sucesso, impedir que tarefas sejam finalizadas e mostrar erros conceituais da aplicação, essa taxa pode ser preocupante.

4.1 Resultados da Discount Usability

A tarefa 11 teve grande quantidade de problemas graves (56.25%), o que é muito alto. Essa taxa representa bem a realidade da tarefa, que não foi completada com sucesso por nenhum usuário sem nenhum auxílio, e só completada por um usuário real (17%) após quatro tentativas. Essa tarefa claramente necessita de ações relacionadas a usabilidade. As sugestões desse trabalho são:

- montagem de um guia, documentação ou ajuda para os usuários na tela, explicando seu objetivo, como utilizar e exibindo exemplos de resultados bem formados;
- remodelagem total baseada nos dados coletados por essa avaliação e nos de novos testes realizados durante o projeto.

Além dessas, outras conclusões podem ser tiradas desses resultados. Vários problemas se mostraram redundantes em diferentes telas do sistema relacionados a diferentes aspectos, conceitos, terminologia e elementos de tela. Um desses problemas que se mostrou de forma bastante clara foi a falta de consistência em elementos comuns do sistema. Um problema desse tipo ocorreu com os botões de ação “gravar”, “cancelar” e “finalizar”, que não possuem uma boa terminologia para as ações executadas, não possuem posicionamento padronizado entre as telas e nem comportamento. Esse problema pode ser visto em todas telas cuja interação dos usuários com formulários é necessária, principalmente na área de exames. Outro problema comum entre diversas telas estava relacionado a data do exame, que ficava próximo a áreas com dados somente de leitura. Como esses dados se encontram em campos de formulário sem nenhuma diferença para o campo data, ficava difícil de visualizar que havia um dado a ser preenchido naquela área. Além disso, a validade do valor do campo nem sempre era verificada. O fato dos campos somente leitura não terem um padrão diferenciado dos campos de entrada de dados também atrapalha a interação dos usuários. Os poucos campos somente leitura diferenciados não são padronizados entre as telas. Outro ponto que é necessário verificar está relacionado a campos obrigatórios para submissão de um formulário, raramente informados pelo sistema. Como em algumas telas o comportamento do sistema era retornar um formulário com dados inválidos em branco, mostrar quais são os campos obrigatórios e quais valores eles esperam receber poderia ajudar nesse tipo de interação.

Devido a comportamento ou respostas inesperadas do sistema os usuários criaram um hábito que apesar de ajudá-los, era nocivo: preencher todos os campos de formulário, mesmo quando não era necessário ou quando o dado não estava disponível. Em algum momento de todas as avaliações os usuários tiveram que informar dados de um mesmo formulário mais de uma vez, pois estavam fora do padrão esperado pela aplicação. Ao submeter os dados com erros, o formulário era retornado com os campos vazios, o que era muito

4.1 Resultados da Discount Usability

frustrante para os usuários. De modo a evitar esse tipo de problema, também foi comum eles decidirem preencher os formulários completamente. Na maior parte do sistema esse tipo de comportamento não era necessário, mas os usuários se sentiam mais confortáveis fazendo isso, apesar de não ser o comportamento padrão em uma aplicação e deixar a execução da tarefa mais lenta.

A diferença de performance e aprendizado do sistema foi bastante clara em relação aos usuários do teste piloto e do teste com usuários alvo. Sugerimos que isso ocorre devido ao conhecimento do domínio dos usuários reais, que tiveram desempenho muito elevado principalmente na área de exames, onde seu conhecimento específico era aplicado. Além disso, alguns problemas relacionados a terminologias e conceitos foram facilmente contornados por esses usuários, enquanto os usuários do teste piloto necessitavam de explicações adicionais. O aprendizado para esses usuários também era mais fácil pois os exames eram comuns a eles, o que ajudava a reconhecer a interface, seus elementos e como interagir com eles. Apesar dessas diferenças a avaliação piloto foi útil para encontrar problemas com usuários sem ou com pouca experiência do domínio de conhecimento e para calibrar a avaliação com usuários reais. A calibração foi importante, pois os ganhos de desempenho esperados na segunda interação com as telas ocorreram também na avaliação real, conforme previsto através do piloto.

Apesar desses e outros problemas, o formulário aplicado após o teste de usabilidade mostrou que em geral os usuários se sentiram confortáveis no sistema e acreditam que ele pode ser útil em suas atividades diárias, acelerando e facilitando processos realizados manualmente. Dos usuários reais, 50% utilizaria o sistema novamente, e o restante utilizaria mas acredita que ele precisa de pequenas correções. Todos usuários reais acreditam terem alcançado nível iniciante na utilização do sistema e somente um deles não acredita que seu conhecimento específico da área foi importante para utilização do sistema. Os usuários acreditam que explicação sobre as siglas utilizadas, um manual ou ajuda na aplicação pode melhorar sua interação. Devido ao pedido de um manual ou ajuda na aplicação, é possível constatar que os usuários não notaram a sua existência em todos os exames, possivelmente por não estar apresentada através da palavra “ajuda”, mas por um ícone remetendo a ela. Por fim, apesar dos problemas durante as avaliações, os usuários acreditam que o sistema é de fácil aprendizado e que o software atende seu objetivo. Os usuários do teste piloto tiveram respostas semelhantes, e algumas delas esperadas, como a dificuldade em aprender e utilizar partes muito específicas do sistema e interpretar seus resultados por não serem da área. Mesmo não sendo da área utilizariam o SacarWeb novamente e sugerem melhorias para interação, interpretação dos dados e gráficas.

4.1.3 Comparação da avaliação heurística com os testes de usabilidade

A medida de similaridade dos problemas encontrados é importante, pois dá maior representatividade, validade e embasamento para 13.62% (29) dos problemas. Apesar disso, ambos testes são importantes, uma vez que problemas considerados graves foram observados somente em uma das avaliações. A tabela 4.4 mostra em detalhes a quantidade de problemas iguais entre as avaliações encontrados por tarefa. Na tabela, a coluna “Problemas Comuns (U)” considera os valores de problemas iguais com base nos testes de usabilidade, enquanto “Problemas Comuns (H)” com base nas heurísticas.

Tarefa	Problemas comuns (U)	Problemas comuns (H)
1 - Cadastrar paciente	4	6
2 - Avaliação inicial	3	3
3 - Avaliação	3	3
4 - Shuttle	3	3
5 - SGRQ	3	3
6 - DP6	2	3
7 - CRQ	3	3
8 - PFP	3	3
9 - Pressões respiratórias	0	0
10 - Evolução do paciente	3	2
11 - Relatórios	2	2
Total	29	31

Tabela 4.4: Quantidade de problemas iguais encontrados em ambas avaliações

Os problemas podem ser vistos em detalhes nos apêndices B e E. Ao observar a tabela 4.4, é possível notar quantidades diferentes apontadas para problemas comuns entre eles para as mesmas tarefas. Isso ocorre pois um problema observado através de uma heurística pode estar associado a mais de um problema em um teste de usabilidade com usuários e um problema apresentado com os usuários pode estar associado a mais de uma heurística. Neste trabalho consideramos a hipótese com menor valor, apesar da pequena diferença entre a de maior valor, onde 31 problemas (14.55%) são semelhantes entre as avaliações.

Dessa maneira, a quantidade total de problemas distintos é 184, considerando a hipótese com menor quantidade de problemas comuns, e 182 na hipótese que considera maior quantidade de problemas comuns. Esses totais são formados pelos 107 problemas encontrados durante os teste de usabilidade somados aos 106 problemas distintos da avaliação heurística menos a quantidade de problemas semelhantes entre elas considerados.

Através dos resultados encontrados com as avaliações propostas pela *Discount Usability* sugerimos que o aprendizado pode ser melhorado consideravelmente para os usuários após a implantação de um novo projeto de interface que considere os principais problemas encontrados. Uma vez que esses problemas estão relacionados com conceitos chave no sistema para os usuários, concluímos que atualmente ele é difícil de entender visto a gravidade dos problemas, apesar dos usuários acreditarem no contrário. Com esse tipo de medida, acreditamos que a quantidade de erros cometidos por usuários em fase de adaptação ao sistema diminuirá, uma vez que o *software* será mais fácil de entender e utilizar. Acreditamos que os problemas relacionados a heurística consistência devem ser revistos com cuidado, pois melhorar esses pontos no sistema pode melhorar a experiência em toda a aplicação.

Deve-se levar em consideração que o grupo de desenvolvimento pode utilizar os resultados dessa avaliação para guiar as próximas iterações de desenvolvimento, principalmente porque ela possui pontos claros com os principais problemas a serem solucionados, dados oriundos da prioridade dada aos problemas da avaliação heurística e observação dos usuários nos testes de usabilidade.

Com o conhecimento adquirido do usuários através dos testes também acreditamos que o mantenedor da aplicação pode ser de grande ajuda durante a correção dos problemas.

Pelos resultados alcançados por esse trabalho acreditamos ser possível integrar a equipe de desenvolvimento do SacarWeb uma estratégia de usabilidade baseada em *Discount Usability*. Também acreditamos ser possível em equipes semelhantes, que possuem somente usuários sem conhecimento específico de Usabilidade e façam parte de equipes com pessoas predominantemente da Computação e área de Saúde.

Um estudo relacionado mostrou que avaliadores, com conhecimento similar de Ciência da Computação e Usabilidade são capazes de diferenciar um projeto bom de um ruim após aprender alguns conceitos de usabilidade (NIELSEN, 1994a). Como esperado por (NIELSEN, 1998), pessoas com pouco conhecimento de usabilidade são capazes de encontrar problemas com heurísticas e conduzir inspeções de usabilidade. Nesse trabalho, avaliadores com mais conhecimento - sobre o sistema ou métodos de inspeção - obtiveram melhores resultados.

4.2 Resultados da aplicação do método híbrido proposto

As tarefas analisadas durante os testes automatizados foram as mesmas observadas através da *Discount Usability*. Por outro lado, a seleção dos usuários utilizou critérios diferentes, e foi realizada da seguinte maneira:

- Os usuários não deveriam ter nenhum contato prévio com o SacarWeb;
- Os usuários poderiam estar em qualquer período do curso de Fisioterapia e qualquer período de participação no projeto LFIP;
- Pelo menos quatro usuários por especialidade deveriam participar dos testes.

Com base nessas necessidades e disponibilidade de usuários, foram selecionados:

- 5 usuários para os exames de Pressões Respiratórias e PFP;
- 5 usuários para o exame DP6
- 6 usuários para os exames SGRQ e CRQ
- 5 usuários para o exame Shuttle

Os usuários deveriam realizar a inserção de duas avaliações com os exames de sua especialidade para um paciente de dentro do período de 1 semana. Os usuários foram orientados a realizar as tarefas sem nenhum tipo de auxílio, e preferencialmente no laboratório do LFIP. Também foi explicado que a aplicação estaria sendo avaliada, e não os usuários. As instruções gerais da avaliação foram fornecidas a todos os usuários através de um documento disponível online. Esse documento pode ser visto no apêndice C. Além do documento com as orientações gerais para a realização das tarefas alvo da avaliação, foi fornecido um segundo documento descrevendo as tarefas que deveriam ser realizadas pelos usuários, que pode ser visto no apêndice D.

De modo a também obter dados qualitativos, foi solicitado aos usuários o preenchimento de dois formulários, um antes e outro após a realização das tarefas da avaliação. O primeiro possui o objetivo de identificar os usuários e verificar seu perfil¹. Já o formulário preenchido ao final da avaliação tinha como objetivo a busca de informações sobre a satisfação

¹<http://goo.gl/NzbUh>

4.2 Resultados da aplicação do método híbrido proposto

dos usuários com relação ao SacarWeb, com base em sua experiência de uso recente ². Esses foram os mesmos formulários aplicados aos usuários durante os testes de usabilidade.

Com base nas respostas dos formulários, preenchidas por dezessete pessoas, os usuários em média tem a idade de 22.21 anos, com mulheres predominantes e conhecimento variando de baixo a bom. O período de experiência médio informado por eles foi de 5.11 anos com frequência de uso diária de aproximadamente 3.29 horas. A maior parte dos usuários utiliza o sistema para estudo, entretenimento, pesquisa e trabalho. A experiência dos participantes na área vai de alunos do terceiro ano até o doutorado, fazendo parte do projeto por 2.43 anos em média do projeto LFIP. Os participantes possuíam múltiplas especialidades, e puderam preencher mais de um exame. A quantidade de usuários por exame, com base no formulário aplicado antes da tarefa pode ser vista na tabela 4.5:

Tarefa	Quantidade de avaliadores
Exame DP6	4
Exame CRQ	3
Exame SGRQ	4
Exame Pressões Respiratórias	4
Exame PFP	5
Exame Shuttle	3

Tabela 4.5: Quantidade de usuários para cada tarefa na avaliação automatizada

Entretanto, a quantidade de respostas do formulário aplicado após a avaliação foi menor, somente 13 usuários responderam, mesmo que o protocolo da avaliação tenha deixado claro que era um passo obrigatório. Em geral os usuários acharam o sistema útil para armazenar dados e calcular valores automaticamente, mas acreditam que é difícil em alguns pontos. Somente um usuário apresentou *feedback* negativo. Dos usuários que responderam, 53.84% utilizariam o sistema novamente, 38.46% condicionaram a utilização posterior a alguma correção ou melhoria do sistema e um usuário, 7.6% não utilizaria o sistema novamente. O usuário que não utilizaria o sistema parece ter apresentado muitas dificuldades ao utilizar o sistema, pois seu *feedback* foi negativo. Em geral os usuários acreditaram que o sistema era fácil de aprender, mas expressaram que uma melhor apresentação facilitaria o uso. A maior parte dos usuários acredita ter atingido nível de utilização iniciante na aplicação, com 23% acreditando terem atingido nível intermediário. 92.3% dos usuários acreditam que seus conhecimentos foram importantes para utilização do sistema, para interpretação das telas do sistema, saber onde inserir os dados no sistema e interpretar seus resultados. 23% dos usuários acreditam que o sistema falha ao atingir seus objetivos, sendo que um deles acredita que o

²<http://goo.gl/g8xqG>

sistema que utiliza, planilhas de Excel, é mais prático. Já o restante dos usuários acredita que o sistema é prático e importante, pois se adapta a rotina dos usuários, ajuda a armazenar os dados corretamente e no gerenciamento das informações.

Na área para comentários e melhorias, dois usuários reportaram dificuldades com a data base, nos questionários um usuário reportou o problema com os questionários, onde campos não obrigatórios são marcados como obrigatórios, outro usuário acredita que se os dados pudessem ser exportados para planilhas do Excel para realização de análises estatísticas em programas específicos posteriormente poderia ser interessante, apesar desse tipo de funcionalidade poder ser implementada no sistema. Finalmente, um usuário reportou problemas ao encontrar áreas do sistema e falta de flexibilidade em formulários com dados obrigatórios incompletos. As sugestões dos usuários estão de acordo com dificuldades percebidas na avaliação heurística e teste de usabilidade, como os problemas com a data base da avaliação, campos não obrigatórios do exame SGRQ e dificuldade de encontrar áreas no sistema. Isso pode indicar que a opinião dos usuários em aplicações de áreas de conhecimento específicas pode ser de grande valor e tão bem fundamentado quanto os dados adquiridos em avaliações de usabilidade. Além disso mostram como é importante testar o sistema em etapas preliminares do projeto, pois o *feedback* dos usuários pode guiar e melhorar o desenvolvimento do sistema, o que confirma a importância da atividade **conheça o usuário** da *Discount Usability*.

O protótipo de software desenvolvido utilizou o arquivo de log produzido durante o período de utilização dos usuários selecionados, que foi estendido a pedido de alguns usuários até o dia 07 de Abril de 2011. Sem nenhum tipo de filtro, o período selecionado possuía 3638 entradas. Cada tarefa teve uma diferente quantidade de URLs filtradas, uma vez que suas definições são específicas, como mostra o apêndice F. Durante os testes detectamos uma deficiência no projeto da aplicação, o que impossibilitou a análise das tarefas **Shuttle** e **Evolução do paciente**. O problema está relacionado a URLs inicial e final da tarefa, que no caso dessas tarefas são iguais. Uma vez que são iguais, não é possível definir quando o usuário inicia e termina a tarefa, uma vez que dois cenários podem ser visualizados:

- quando o usuário acessa a URL pela primeira vez, a tarefa está sendo iniciada e pela segunda vez, finalizada;
- todo acesso do usuário a URL indica um novo início da tarefa.

Certamente o segundo cenário não pode ser utilizado em termos práticos, mas a utilização do primeiro não garante a validade dos dados, uma vez que existe a possibilidade do usuário ter reiniciado a tarefa.

As tarefas que puderam ser analisadas são detalhadas nas próximas seções. Por simplicidade apresentamos somente os dados gráficos. Os resultados textuais detalhados apresentados pela aplicação podem ser visualizados no apêndice G. É possível obter em todas as tarefas a ocorrência de desistência de tarefas, caracterizadas pelo início de duas tarefas, uma logo após a outra, sendo considerada a desistência na primeira tarefa. Por esse motivo o tempo de realização da tarefa é definido para zero, pois uma vez que houve uma desistência não há uma URL para calcular o tempo que se permaneceu na primeira visita.

Ao final do processo, uma avaliação da qualidade dos *clusters* é realizada, utilizando o mesmo método de (KUO; HO; HU, 2002), o valor Lambda de Wilk. Uma vez que chegar nesse valor necessita de passos intermediários, foi possível calcular mais valores de qualidade, como visto abaixo:

- Variância interna dos grupos (SSW);
- Variância entre grupos (SSB);
- Variância total (SST);
- Valor-F (F);
- Valor Lambda de Wilk (λW).

Apresentaremos os valores de ambos métodos utilizados, GHSOM + *K-means* e SOM + *K-means*, entretanto, a análise de usabilidade considerada será somente a apresentada pelo algoritmo GHSOM + *K-means*, proposta deste trabalho. O valor lambda de Wilk é utilizado como comparação entre os algoritmos, como proposto em (KUO; HO; HU, 2002), e quanto menor esse valor, melhor a clusterização, como mostra o desenvolvimento da equação 3.1.

4.2.1 Tarefa 1: Cadastrar paciente

Cada usuário realizou essa etapa somente uma vez, visto que somente um paciente deveria ser utilizado. A figura 4.1 mostra como estão distribuídos os *clusters* finais.

Os *clusters* mais representativos, 2 e 6, representam comportamentos predominantemente errados. No *cluster* 2 é notável a presença de várias ocorrências da URL de busca de pacientes (`/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons`), o que pode evidenciar que o usuário não estaria realmente cadastrando um usuário ou estaria confuso com o fluxo de dados da aplicação.

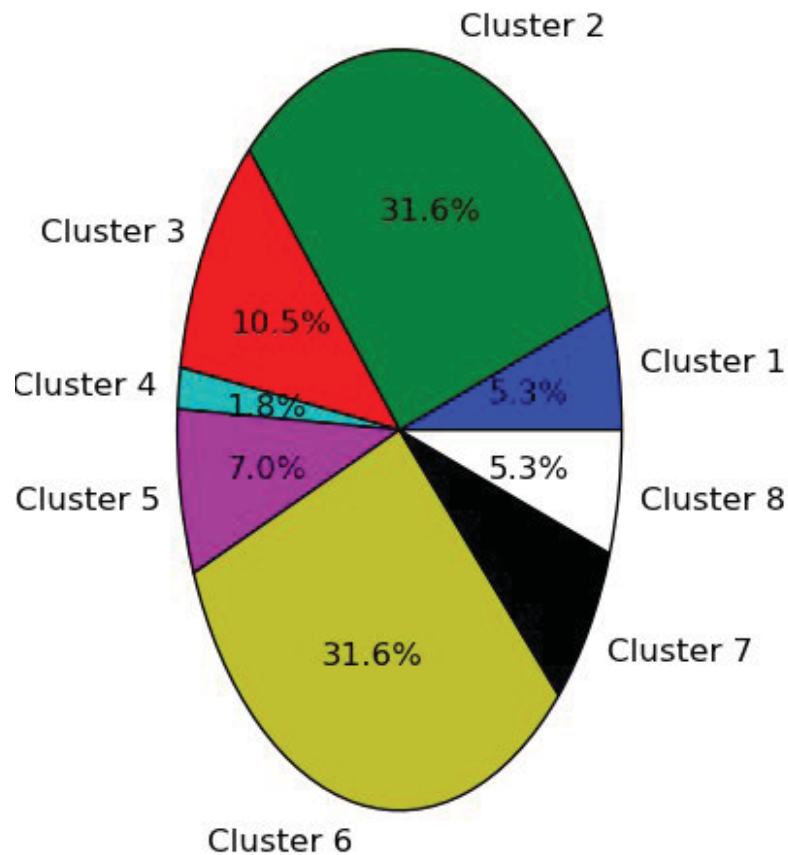


Figura 4.1: Representatividade dos clusters da Tarefa 1

Já no *cluster 6* é notável a grande presença desistência de tarefas. Curiosamente padrões de sucesso podem ser notados nesse *cluster*, que provavelmente foram inseridos nesse padrão devido ao tempo de finalização da tarefa, que foi bem rápido nesse caso. É provável que esses usuários estivessem retornando de uma inserção com erros, mais rápida que o primeiro contato com tela.

Padrões de sucesso são encontrados nos *clusters 1, 3, 7 e 8*, predominantemente, além daqueles presentes no *cluster 6*. A diferença básica entre eles é o tempo de conclusão da tarefa, que está de acordo com o observado durante a avaliação com usuários. Acreditamos que os tempos menores são obtidos na segunda interação com a tela, provavelmente após cometer erros de preenchimento no primeiro contato. Os tempos mais altos são atribuídos ao primeiro contato dos usuários com a tela.

O fato da maior parte dos *clusters* indicar acessos com algum tipo de erro pode ser preocupante nessa tarefa, que deve ser investigada com maior profundidade para detecção pontual dos problemas e suas causas, o que já foi realizado na avaliação através de

Discount Usability.

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.6:

Equação de qualidade	SOM	GHSOM
F	1.55138688998	1.30278807492
λW	3.33419979364	0.241064961748

Tabela 4.6: Medidas de qualidade dos clusters para a tarefa 1

Para esse caso, como mostra o valor λW da tabela 4.6, o algoritmo proposto é melhor que SOM + *K-means*(KUO; HO; HU, 2002).

4.2.2 Tarefa 2: Avaliação inicial

Essa tarefa é realizada somente uma vez por usuário, pois está diretamente ligada ao paciente e não a avaliação. Os *clusters* foram distribuídos como mostra a figura 4.2.

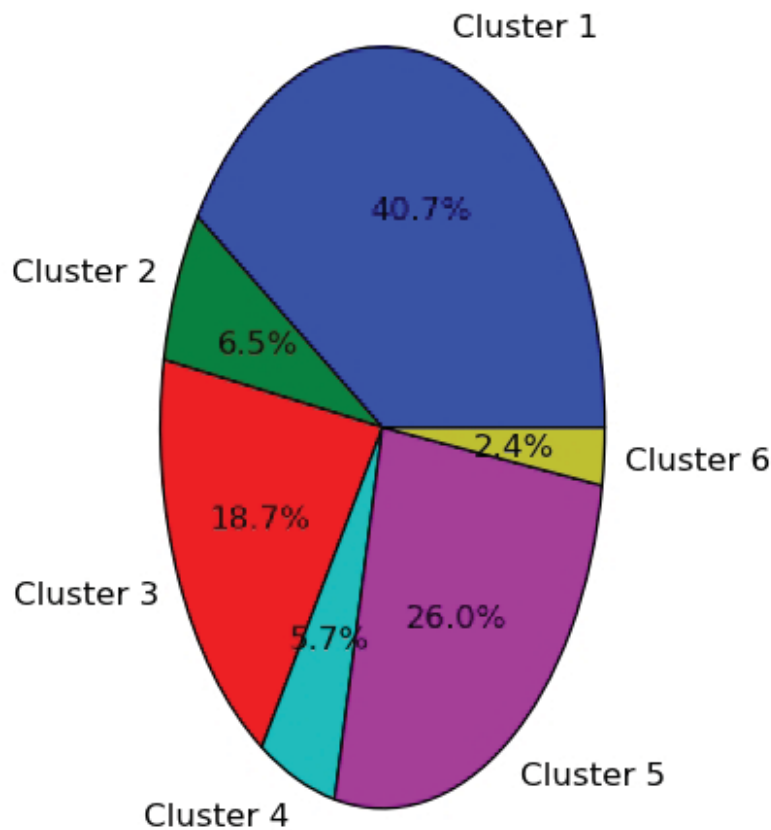


Figura 4.2: Representatividade dos clusters da Tarefa 2

Claramente o *cluster* 1 é o mais representativo nesse caso. Aparentemente ele reúne padrões de acesso com tempo baixo de finalização e problemas ou desistência da tarefa. É possível que os usuários estivessem buscando por alguma outra função, uma vez que não há um padrão bem definido para suas URLs. Curiosamente, dois padrões de acesso de sucesso foram classificados no nele.

O *cluster* 5 mostra comportamento semelhante, mas também mostra alguns erros da aplicação, que atinge sua URL de erro padrão, possivelmente devido a falhas nessa rotina disparados por entradas do usuário, o que também ocorre no *cluster* 3. O que valida o pensamento de que o usuário está navegando procurando por algo é o tempo de finalização da tarefa, ou atingimento da quantidade máxima de URLs, que é muito baixo.

Os *clusters* restantes, 2, 4 e 6, representam caminhos de sucesso, variando no tempo de finalização da tarefa. O *cluster* mais diferenciado é o 4, que possui tempos bem menores que dos outros. Isso pode ser devido as opções padrão contidas no formulário, que ajudam no preenchimento rápido, ou devido a erros dos usuários no preenchimento que em sua segunda passagem pela tela obtém melhor desempenho.

Novamente, a maior parte dos *clusters* representa algum erro durante a execução da tarefa, e devido a isso a tarefa também precisa ser avaliada de maneira mais aprofundada.

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.7:

Equação de qualidade	SOM	GHSOM
F	1.93681105071	0.570678073258
λW	2.89495926707	0.127714786374

Tabela 4.7: Medidas de qualidade dos clusters para a tarefa 2

Para essa tarefa, a tabela 4.7 mostra que o algoritmo proposto teve melhor desempenho que a proposta de (KUO; HO; HU, 2002).

4.2.3 Tarefa 3: Exame PFP

O exame PFP deveria ser realizado pelo menos duas vezes por usuário. Levando em consideração a quantidade de usuários que afirmou ser especialista em PFP, cinco usuários participaram dessa avaliação. Selecionamos a quantidade de três URLs após a inicial uma vez que é possível trocar o padrão de normalidade do exame e consultar a ajuda/documentação. Entretanto, a tarefa pode ser concluída de forma ótima em apenas duas interações.

4.2 Resultados da aplicação do método híbrido proposto

Após a análise pelo método proposto, os *clusters* finais para a tarefa ficaram configurados como mostra a figura 4.3.

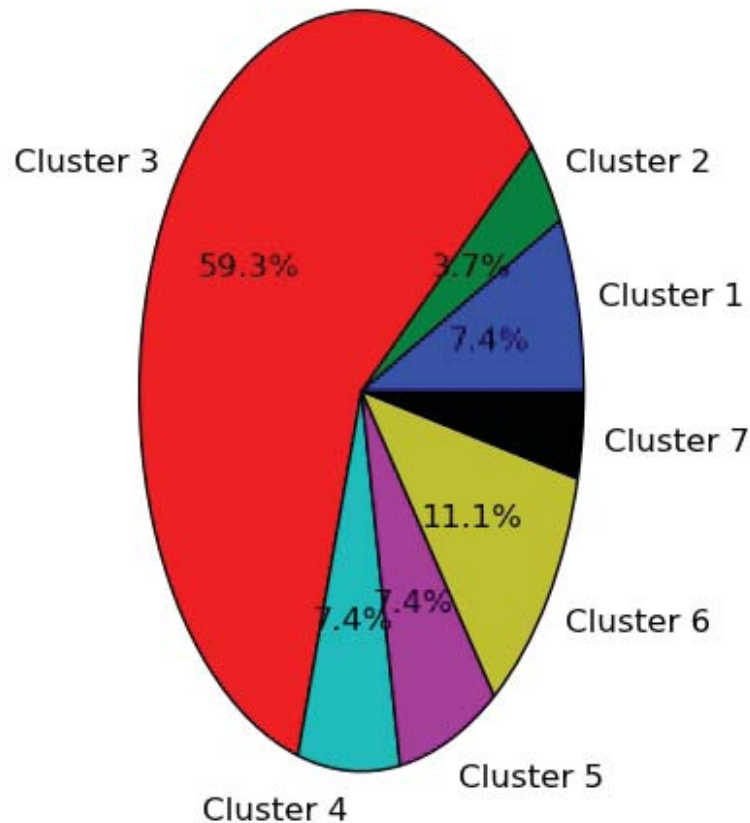


Figura 4.3: Representatividade dos clusters da Tarefa 3

Os *clusters* 1, 2, 6 e 7 representam acessos com sucesso, sendo diferentes no tempo de finalização da tarefa. O *cluster* 7 apresenta um tempo bastante alto para conclusão da tarefa, o que pode ser explicado pela presença de campos no formulário digital que não estão presentes no formulário impresso, o que pode confundir os usuários. Já entre os valores do *cluster* 6, os tempos são muito baixos, o que pode ser um acesso em que o usuário preencheu somente campos obrigatórios ou está na segunda interação com exame. Os *clusters* 2 e 1 apresentam um tempo de finalização da tarefa mais próximo do esperado.

Já o *cluster* 4 sugere que os usuários nele classificados estavam navegando pela aplicação buscando algum recurso ou procurando aprender sobre ela, uma vez que não existe um padrão lógico de URLs. O *cluster* 3, o mais representativo nesse caso, reúne grande quantidade de acessos únicos a URL de início da tarefa, o que pode representar desistência da tarefa.

Nessa tarefa a representatividade dos *clusters* de sucesso é suficiente para acreditar que a aplicação não está errada, porém a quantidade de acessos com padrões de erro é preocupante. Devido a isso, acreditamos que a tarefa necessita de atenção relacionada a usabilidade pois possui problemas que apesar de não impedirem o usuário de finalizá-la, pode deixar a operação mais difícil ou lenta.

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.8:

Equação de qualidade	SOM	GHSOM
F	2.42877304277	2.89865442752
λW	0.102856000022	0.08284989382

Tabela 4.8: Medidas de qualidade dos clusters para a tarefa 3

Para essa tarefa a tabela 4.8 mostra que a proposta desse trabalho teve melhor desempenho que o algoritmo SOM + *K-means* (KUU; HO; HU, 2002).

4.2.4 Tarefa 4: Exame SGRQ

Das tarefas analisadas, o exame SGRQ é o mais representativo para o método proposto, uma vez que possui três passos intermediários, a tela de retorno e a presença de ajuda. Logo, possui um caminho ideal bem definido e longo para ser analisado. A configuração de quantidade de URLs após a inicial foi definida para seis e segundo os questionários pré-avaliação, quatro usuários participaram dela. Após a utilizar o protótipo desenvolvido, os *clusters* ficaram como mostra a figura 4.4:

Os *clusters* 3 e 4 representam sucesso com tempo diferente entre eles, mas compatível com o desempenho esperado. Acreditamos que o *cluster* 4 representa acessos da segunda interação com o sistema, pois o tempo de conclusão da tarefa é menor. O *cluster* 2 também apresenta uma ocorrência de sucesso, apesar de ser caracterizado por padrões de acesso de usuários navegando pelo sistema procurando uma função ou buscando aprender mais sobre ele, uma vez que várias telas com finalidades diferentes são acessadas nesse *cluster*, não existindo homogeneidade entre os acessos. Por fim, o *cluster* 1 parece representar desistências na execução da tarefa.

Acreditamos que essa tarefa não possui nenhum problema expressivo de usabilidade, uma vez que a maior parte dos padrões apresentados são de sucesso. Também acreditamos que os erros detectados podem ter baixa representatividade, pois aparentemente são padrões de usuários realizando acessos para encontrar algum ponto do sistema.

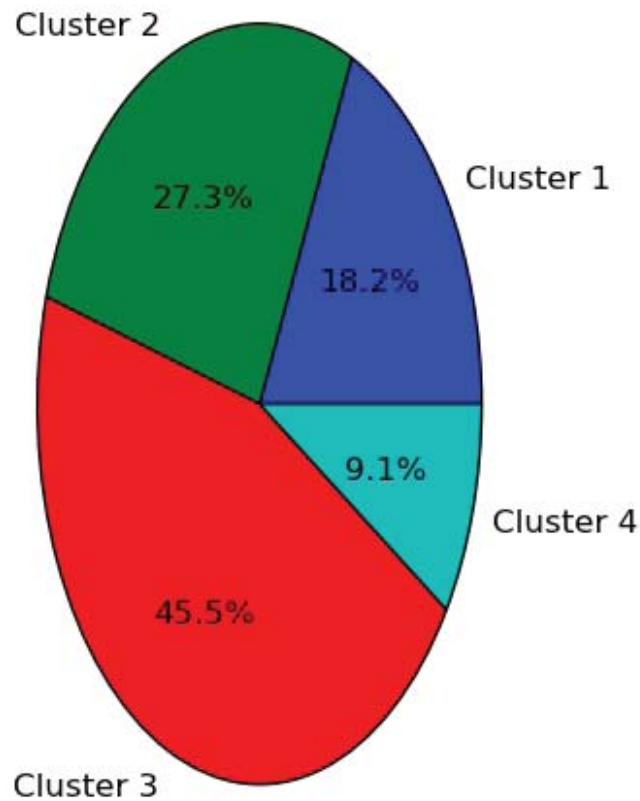


Figura 4.4: Representatividade dos clusters da Tarefa 4

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.9:

Equação de qualidade	SOM	GHSOM
F	3.52200649491	5.39552225607
λW	5.08679792769	3.80218570621

Tabela 4.9: Medidas de qualidade dos clusters para a tarefa 4

A tabela 4.9 mostra que para tarefa o algoritmo proposto é melhor que o SOM + *K-means* (KUO; HO; HU, 2002).

4.2.5 Tarefa 5: Exame CRQ

O exame CRQ possui uma etapa intermediária, que apresenta os resultados do exame, e então possibilita a operação gravar. Além disso, também é disponibilizada ajuda/documentação. Desse modo, a quantidade de URLs após a inicial foi definida para quatro.

Com essas configurações, os *clusters* para a tarefa ficaram como mostra a figura 4.5:

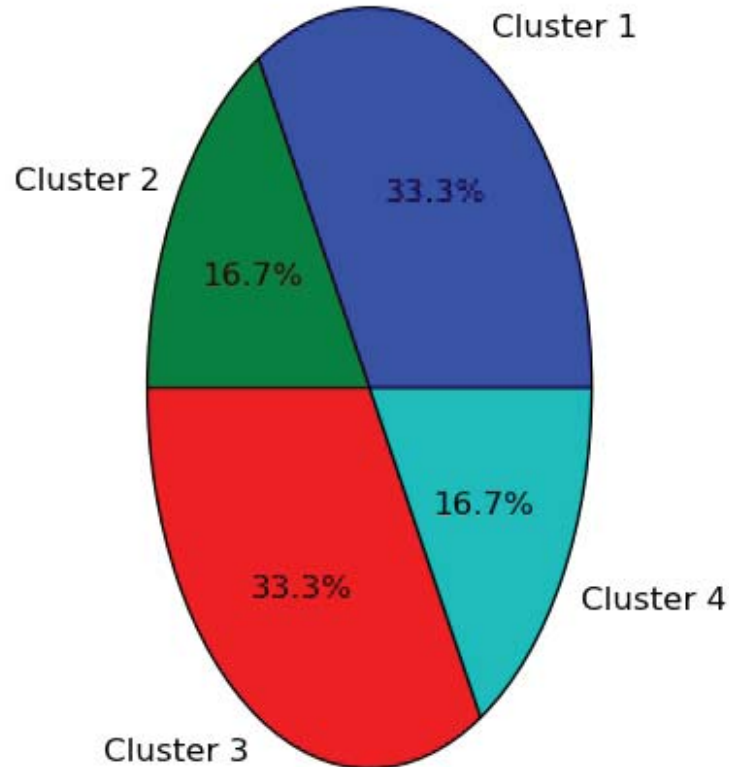


Figura 4.5: Representatividade dos clusters da Tarefa 5

Nesse caso somente o *cluster* 3 representa um caso de erro na realização da tarefa, todos os outros representam sucesso sem nenhuma passagem por outras URLs fora das contidas no menor caminho da tarefa. A diferença entre os *clusters* 1, 2 e 4 é o tempo, que para uma tarefa relativamente simples é alto, que mesmo no melhor tempo ainda pode ser considerado alto. Desse modo, apesar dos usuários conseguirem finalizar a tarefa, deve-se atentar para o modo de realizá-la, melhorando o desempenho, o que pode ser devido ao modo como as informações são apresentadas aos usuários.

Já os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.10:

Essa tarefa apresentou um resultado exatamente igual entre os algoritmos, como pode ser observado na tabela 4.10. Atribuímos esse comportamento a quantidade de dados disponíveis para essa tarefa, que foram bastante reduzidas, não sendo suficientes para gerar diferenças entre as propostas.

Equação de qualidade	SOM	GHSOM
F	0.364391820678	0.364391820678
λW	1.01397528087	1.01397528087

Tabela 4.10: Medidas de qualidade dos clusters para a tarefa 5

4.2.6 Tarefa 6: Exame Shuttle

O exame shuttle foi uma das tarefas em que não foi possível avaliar a abordagem automatizada proposta uma vez que suas URL inicial e final são iguais.

4.2.7 Tarefa 7: Exame DP6

Exame DP6 possui somente a URL inicial e final, com a possibilidade de acesso a ajuda/documentação em qualquer etapa do processo. Desse modo, a quantidade de URLs após a inicial foi configurada para três. Os *clusters* foram organizados conforme mostra a figura 4.6.

Para essa tarefa os *clusters* 2 e 3 representam sucesso. O tempo para finalização da tarefa também parece estar correto, uma vez que alguns conceitos da tela são difíceis de interpretar. Os outros *clusters*, que podem representar erros, parecem classificar usuários que estavam buscando alguma função no sistema ou aprendendo sobre ele e aparentemente não representam erros graves com a tela do exame. Entretanto, um modo mais ágil de inserção dos exames poderia ser estudado em um projeto alternativo.

Por fim, os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.11:

Equação de qualidade	SOM	GHSOM
F	2.71544310966	1.68587333092
λW	2.97779864456	4.81854470682

Tabela 4.11: Medidas de qualidade dos clusters para a tarefa 7

Para essa tarefa, a proposta de (KUO; HO; HU, 2002) obteve melhor desempenho que a proposta desse trabalho, como mostra a tabela 4.11.

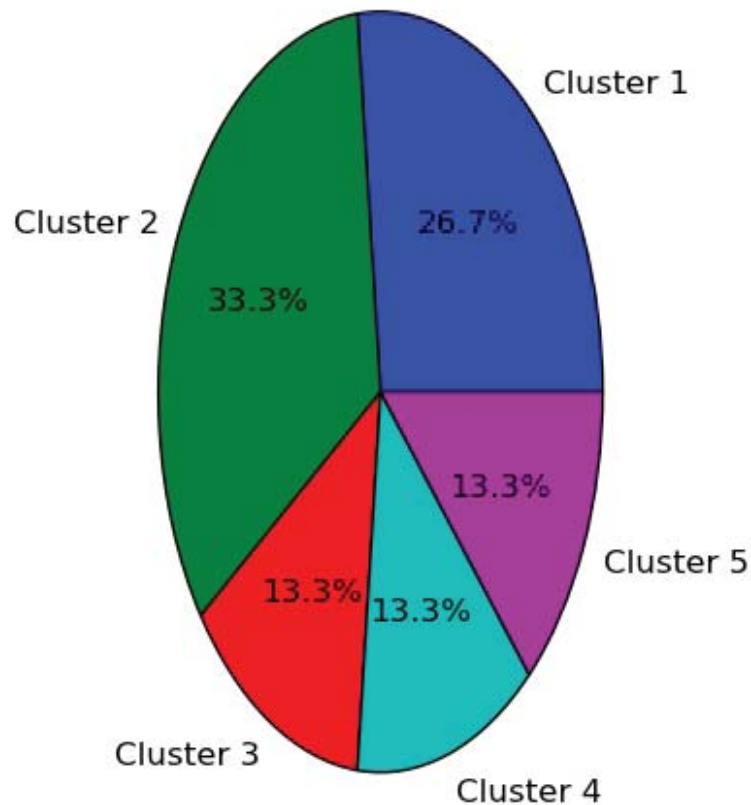


Figura 4.6: Representatividade dos clusters da Tarefa 7

4.2.8 Tarefa 8: Exame Pressões Respiratórias

Certamente essa é a tarefa mais simples do sistema, pois consiste em inserir três valores em um formulário e então gravá-lo. Apesar da simplicidade, a ajuda/documentação é encontrada no local sistema. Por esse motivo foi definida a quantidade de três URLs após a inicial. Os *clusters* apresentados podem ser visualizados na figura 4.7:

Os *clusters* 1, 2, 4 e 5 representam acessos com sucesso na finalização da tarefa. O *cluster* 1 apresenta um tempo de finalização muito alto. Os outros *clusters* de sucesso apresentam tempo compatível com o esperado e nesse caso, acreditamos que a diferença entre a primeira e segunda avaliação não é percebida, devido a facilidade da tarefa. O *cluster* 3 representa acessos a tarefa e desistência da mesma. Por ser uma tarefa simples e não aparentemente não ter ocorrido erros graves, acreditamos que apenas uma pesquisa de opinião sobre a tarefa completa a avaliação com o método proposto.

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.12:

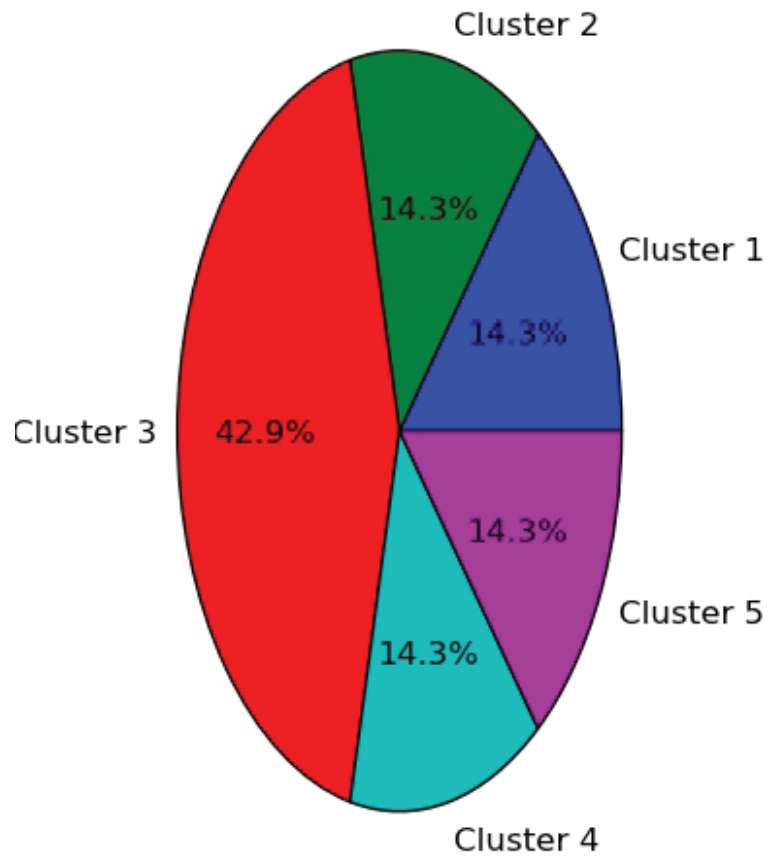


Figura 4.7: Representatividade dos clusters da Tarefa 8

Equação de qualidade	SOM	GHSOM
F	0.714160472651	0.530568014208
λW	7.63959329273	2.99502797048

Tabela 4.12: Medidas de qualidade dos clusters para a tarefa 8

Para essa tarefa o trabalho de (KUO; HO; HU, 2002) foi inferior a proposta deste trabalho, como mostra a tabela 4.12.

4.2.9 Tarefa 9: Evolução do paciente

A evolução do paciente foi uma das tarefas em que não foi possível avaliar a abordagem automatizada proposta uma vez que suas URL inicial e final são iguais.

4.2.10 Tarefa 10: Relatórios

Todos os usuários deveriam passar pela tarefa Relatórios pelo menos uma vez. Em sua tela é possível definir filtros dinâmicos com base nos exames, e a cada adição de filtro uma nova requisição é feita. Uma vez que cada usuário estaria realizando a análise com base em seu exame de especialidade, acreditamos que no máximo dois filtros seriam definidos. Nenhuma tela intermediária é apresentada, apesar de requisições serem realizadas entre a tela inicial e final. Desse modo, a quantidade de URLs entre após a inicial foi definida para quatro. Os *clusters* ficaram configurados como mostra a figura 4.8:

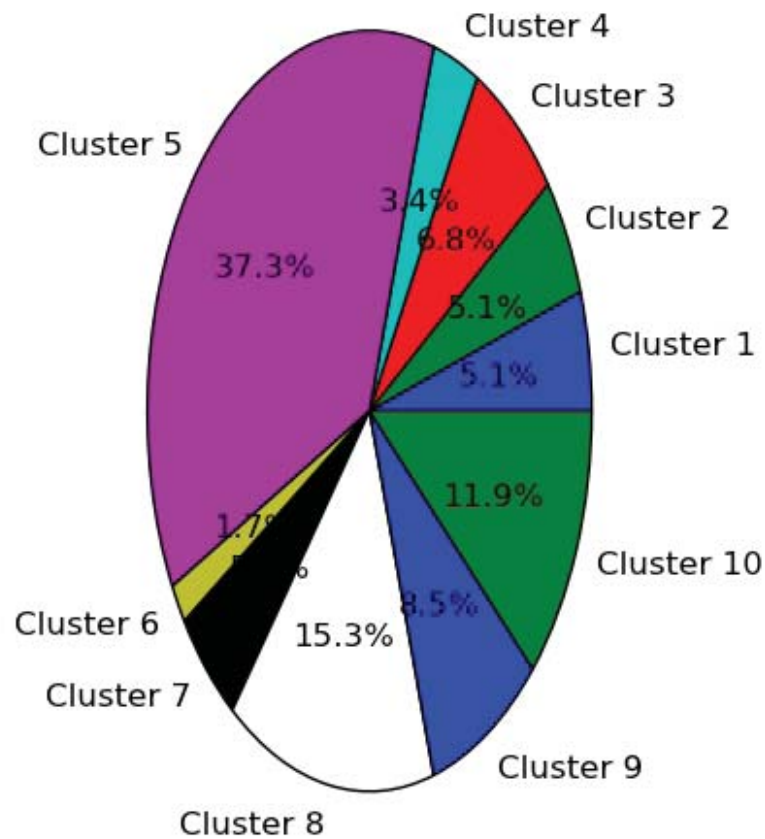


Figura 4.8: Representatividade dos clusters da Tarefa 10

A maior parte dos padrões identificadas foram de erros. Alguns deles levando para a tela de erro padrão da aplicação, responsável por uma tela mais amigável quando um erro inesperado ocorre. Esse comportamento evidencia um problema inesperado, provavelmente a partir de entrada do usuário, algo grave e que deve ser tratado.

Padrões de erro ou navegação buscando aprender mais sobre a aplicação estão representados pelos *clusters* 2, 4, 5, 7, 8, 9 e 10. Os acessos vieram desde usuários que

querem finalizar rapidamente a tarefa, com tempos de acesso curto, a usuários que parecem desejar aprender sobre ela, com longos tempos de acesso.

Padrões de acesso que podem representar sucesso estão nos *clusters* 1, 3 e 6. O *cluster* 3 devido aos pequenos tempos de acesso, parece representar o caso em que o usuário inicia a tarefa e preenche alguns campos rapidamente com o objetivo de terminar rapidamente. Já os *clusters* 1 e 6 parecem ter mais validade em termos de tempo de acesso, mas não é possível saber se os resultados obtidos pelos usuários foram satisfatórios, apesar de terem atingido a URL de sucesso.

Devido a grande quantidade de padrões de acesso representando erros, acreditamos que um cuidado aprimorado com essa tarefa em relação a usabilidade é necessário.

Os valores de qualidade dos *clusters* para essa tarefa estão na tabela 4.13:

Equação de qualidade	SOM	GHSOM
F	1.17392963873	1.01794617049
λW	82.5518024917	6.96553359463

Tabela 4.13: Medidas de qualidade dos clusters para a tarefa 10

Para essa tarefa, a proposta do SOM + *K-means* (KUO; HO; HU, 2002) obteve desempenho inferior em relação a abordagem deste trabalho.

4.3 Comparação das avaliações com *Discount Usability* e com o método proposto

Neste capítulo comparamos os resultados de ambas avaliações realizadas no SacarWeb, a avaliação baseada em *Discount Usability* e no método automatizado proposto por esse trabalho. As comparações analisam os *clusters* apresentados pelo protótipo desenvolvido, principalmente os caminhos comuns e o tempo para finalizar a tarefa. Essa análise é realizada pelo mantenedor da aplicação, que possui conhecimento sobre os padrões esperados a partir da arquitetura da aplicação. O mesmo foi responsável pelos arquivos de entrada para o protótipo, que indicam os padrões das tarefas, pelo teste de usabilidade e confecção do arquivo único da avaliação heurística. Acreditamos que essa prática é a mais recomendada, uma vez que os resultados serão interpretados mais corretamente por alguém com conhecimento da aplicação alvo.

Para a **tarefa 1: Cadastrar paciente** levando em consideração os testes

de usabilidade, o baixo tempo para conclusão na tarefa pode ter ocorrido em uma segunda tentativa de cadastro, após uma primeira tentativa com erro, ou em um cadastro de um segundo usuário, comportamento observado nos testes piloto e repetido nos testes com usuários reais. O fato do *cluster 2* aparentemente apresentar um padrão de busca pela função dentro do SacarWeb se relaciona com o fato observado tanto nos testes de usabilidade, quanto na avaliação heurística, da área não ser facilmente encontrada. O problema grave encontrado pela avaliação heurística e teste de usabilidade relacionado ao código do paciente pode ser o motivo dos diferentes tempos na finalização da tarefa, que também pode ter ocorrido com o campo CEP. Outro fato que pode contribuir para tempos mais altos na finalização da tarefa é o fato dos formulários impresso e digital estarem diferentes. Acreditamos que usuários com maior tempo médio de acesso, poderiam estar em seu primeiro contato com o SacarWeb e buscavam dados em formulários impressos diferentes do digital.

Um problema que ficou bastante claro durante os testes de usabilidade foi a confusão dos usuários com o termo avaliação. Na tarefa **tarefa 2: Avaliação inicial** ao considerar esse problema é possível que o *cluster 2* represente a confusão dos usuários para compreender o conceito de avaliação. A maior parte dos usuários não sabia diferenciar entre uma avaliação inicial e a avaliação, acreditando que se tratavam do mesmo conceito. Outra dificuldade está relacionada com o acesso a essa tarefa, que não fica imediatamente após o cadastro do paciente, e deveria ser o caminho natural do usuário. Esse fato incentiva o usuário a navegar pelo sistema procurando pela tela realizando leitura rápida por busca de informações relevantes. Nesse processo é possível que ele chegue a página mais de uma vez sem identificar que está na tela correta. Também é possível observar tempos rápidos na finalização da tarefa, o que é associado ao fato de serem fornecidos valores padrão para as respostas, foco automático em respostas que precisam de detalhes e semelhança entre os formulários impresso e digital. Além disso, opiniões de usuários indicavam esses fatores como bons facilitadores do preenchimento.

A **tarefa 3: Exame PFP** apresenta semelhanças entre o tempo esperado dos usuários, após observação durante os testes de usabilidade, e o tempo atingido nos *clusters 1 e 2*. O tempo do *cluster 7* é muito alto para usuários especialistas no exame. Esse comportamento é creditado ao fato dos formulários serem muito diferentes entre si e da nomenclatura utilizada entre eles também, o que confundiu até mesmo um usuário especialista nesse exame durante os testes de usabilidade e também foi observado na avaliação heurística. Tempos muito pequenos são observados no *cluster 6*, o que pode ser a segunda interação do usuário com o exame ou uma entrada de testes com erros. Outro comportamento que pode ter sido influenciado pela diferença entre os formulários pode ser observado no *cluster 3*, ou

também por comportamento observado durante avaliação heurística e testes de usabilidade, pois os usuários conhecem o exame pelo termo Espirometria, enquanto o SacarWeb utiliza PFP, problema que foi considerado grave.

Na **tarefa 4: Exame SGRQ** observamos que o tempo esperado dos usuários após testes de usabilidade ocorre no *cluster* 3. O *cluster* 4 aparentemente possui um tempo de realização da tarefa durante a segunda interação do usuário com o exame. O *cluster* 2 apresenta um padrão de sucesso com o usuário retornando para conferir dados de etapas anteriores do exame com tempo de finalização ainda dentro do esperado, apesar de estar classificado em um *cluster* com padrão de usuários navegando pelo sistema. A falta de grandes padrões de erro nos *clusters*, como observado em outras tarefas, pode ser reflexo da boa prática observada nos testes de usabilidade, a igualdade entre os formulários impresso e digital.

Conforme observado durante avaliação heurística e testes de usabilidade para a **tarefa 5: Exame CRQ**, a diferença de tempo entre os padrões de sucesso contidos nos *clusters* 1, 2 e 4 pode ser a falta de ordenação entre os campos de falta de ar no início do exame, principalmente ao observar o *cluster* 1. Essa foi uma das principais características de dificuldade no exame, que além de não ter a mesma ordenação possuía diferenças entre o texto das opções entre os formulários impresso e digital. Apesar desse ser um dos principais problemas da aplicação, o exame não apresentou grandes quantidades de padrões com erro, assim como observado durante as avaliações com *Discount Usability*. Neste exame é curioso o fato que a grande diferença entre os formulários não dificultou o preenchimento, mas parece ter tido efeito contrário e facilitado, uma vez que o formulário impresso é confuso de interpretar.

A **tarefa 6: Exame Shuttle** não pode ser analisada através da abordagem proposta por esse trabalho, devido a limitação encontrada em tarefas com URLs inicial e final iguais. Desse modo deixamos como referência a comparação entre a avaliação heurística e testes de usabilidade apresentada no capítulo 4.1.3.

Para a **tarefa 7: Exame DP6** o *cluster* 2 apresenta um tempo bastante semelhante ao apresentado pelos usuários durante os testes de usabilidade, uma vez que existem vários campos a serem preenchidos. Acreditamos que esse tempo pode ser melhorado, pois existe uma diferença entre os formulários impresso e digital, que necessita de atenção redobrada dos usuários para inserir dados nos locais corretos. O *cluster* 3 apresenta um padrão de acesso de sucesso com tempo menor, o que pode ser a segunda interação do usuário com o sistema.

A **tarefa 8: Exame pressões respiratórias**, possui uma das telas mais simples do sistema. Os tempos de realização da tarefa nos *clusters* de sucesso 2, 4 e 5 foram

bastante semelhantes ao tempo esperado dos usuários após os testes de usabilidade. O alto tempo apresentado no *cluster 1* não é considerado comum, entretanto um usuário alertou para o fato do protocolo do exame estar implementado de modo ligeiramente diferente do realmente utilizado. Acreditamos que o usuário ao reparar com esse fato deve ter buscado ajuda no guia da avaliação, fornecido aos usuários do teste automatizado. O fato de poucos padrões de erro terem sido apresentados pelo método proposto também reflete comportamento das avaliações realizadas com *Discount Usability*.

A **tarefa 9: Evolução do Paciente** também não pode ser analisada através da abordagem proposta por esse trabalho, devido a limitação encontrada em tarefas com URLs inicial e final iguais. Novamente deixamos como referência a comparação entre a avaliação heurística e testes de usabilidade apresentada no capítulo 4.1.3.

Finalmente, a **tarefa 10: Relatórios** foi a mais problemática nos testes de usabilidade e heurística. De maneira semelhante, os *clusters* apresentados pela abordagem automatizada em sua maioria foram compostos por padrões de erro na aplicação. Conforme avaliação com *Discount Usability*, esse comportamento pode estar ligado a vários problemas de usabilidade identificados na tela, principalmente ao fato da falta de compreensão da maior parte dos temas pelos usuários alvo e a falta de ajuda. Outro padrão semelhante entre as avaliações pode ser observado através do tempo de acesso, inicialmente rápido com os usuários intencionados a terminar rapidamente, visto que se trata da última tarefa, e após acessos com erro, procurando entender melhor a tarefa e conseqüentemente aumentando os tempos de acesso. A presença da tela padrão de erros do sistema, devido a entradas inesperadas ou fora do padrão também foi comum entre as avaliações.

4.4 Generalização dos resultados obtidos com o estudo de caso

O estudo de caso mostrou que a metodologia proposta está de acordo com os resultados encontrados na literatura, em termos de análise de usabilidade automatizada, de *Web Mining* e metodologias híbridas de análise de usabilidade. A metodologia proposta é semelhante aos trabalhos (TIEDTKE; MÄRTIN; GERTH, 2002), (MORANDINI, 2003), (LI; KIT, 2005), (BEIREKDAR et al., 2005), (SCHWERZ; MORANDINI; SILVA, 2007) e (AU et al., 2008). Após a análise da literatura, ficou claro que o trabalho mais completo analisado foi (AU et al., 2008), apesar de ter atingido parcialmente os objetivos propostos. Entretanto, todas as áreas propostas para análise podem formar a ferramenta automatizada de análise mais completa.

Foi possível notar a semelhança entre os trabalhos em termos de métricas de usabilidade cobertas, recursos utilizados para análise, modo de definição das tarefas, conclusões quanto as melhorias de métodos automatizados e integração com métodos tradicionais. Por outro lado a abordagem proposta se mostrou mais flexível em alguns pontos como na definição das tarefas, que na maior parte dos casos precisa ter todo seu caminho definido. Ao definir somente as URLs inicial e final, o avaliador ganha flexibilidade para descobrir o que está acontecendo após uma tarefa específica ser iniciada, independentemente das URLs percorridas entre o início e fim da tarefa. O fato do avaliador definir as tarefas para análise foi utilizado em outras abordagens, e também apresenta bons resultados, assim como destinar a avaliação automatizada para grupos com pouca experiência em usabilidade.

Também é importante notar que o algoritmo de clusterização proposto para a abordagem automatizada, o GHSOM + *K-means*, se mostrou efetivo e melhor que o SOM + *K-means*. Devido a isso, é possível utilizá-lo em outras aplicações de clusterização. Sua vantagem é facilidade de uso, uma vez que os parâmetros iniciais necessários no SOM não são necessários no GHSOM. Por esse motivo, pela arquitetura do GHSOM e por resultados de estudos anteriores, acreditamos que o algoritmo proposto pode ser de execução mais rápida, uma vez que não é necessário definir parâmetros iniciais diversas vezes até que se obtenha uma clusterização satisfatória.

Diferentes estudos de caso foram realizados nos trabalhos estudados, com áreas de conhecimento diferentes da observada neste trabalho. Por esse motivo é possível que a metodologia proposta possa ser aplicada em diferentes áreas de conhecimento de aplicações Web, uma vez que ela é semelhante a outras aplicações que já obtiveram sucesso em outras áreas.

Também foi possível observar o sucesso na utilização da metodologia proposta, integrando a abordagem automatizada com *Discount Usability*. Essa comportamento também é observado na literatura, e mostrou de forma clara neste trabalho que esse tipo de estratégia para análise de usabilidade traz benefícios. Além da metodologia proposta, é possível que outras abordagens não automatizadas de usabilidade se integrem a abordagem proposta, o que pode aumentar os benefícios de diferentes métodos de análise de usabilidade.

4.5 Generalização da metodologia proposta para outras aplicações Web

O estudo de caso mostrou que a metodologia proposta é aplicável em aplicações Web da área de saúde, com usuários possuindo alto nível de conhecimento da área da aplicação. Visto a semelhança e melhorias frente a outros trabalhos analisados, pode ser possível aplicar a metodologia proposta em outras áreas de conhecimento, tão específicas quanto a deste trabalho ou mais genéricas.

Para que isso seja possível os passos para utilização da abordagem automatizada devem ser seguidos, dando a devida atenção as etapas onde o conhecimento específico da aplicação alvo de análise é necessário. Indicações gerais como a utilização de um especialista na aplicação alvo ou domínio da aplicação devem ser seguidas, pois garantem maior fidelidade dos resultados.

Para a integração do metodologia automatizado com a avaliação baseada em *Discount Usability* recomendamos que a literatura básica sobre *Discount Usability* seja de conhecimento da equipe responsável pela avaliação. Também recomendamos que a avaliação automatizada seja realizada por último, pois um maior conhecimento da aplicação, seus usuários e problemas será adquirido ao se realizar a avaliação com *Discount Usability* inicialmente.

5 Conclusão e trabalhos futuros

Uma questão que buscou-se responder nesse trabalho foi: é possível que ferramentas automatizadas auxiliem na análise de usabilidade? Trabalhos analisados no capítulo 2.3 mostram que sim. Mais que auxiliar ganhos de desempenho, confiabilidade, menor uso de recursos e facilidade de aplicação são pontos de grande vantagem de análises de usabilidade automatizadas.

O objetivo desse trabalho foi fornecer uma metodologia automatizado para busca de falhas de usabilidade, facilitando, agilizando e reduzindo custos do processo. Nossa proposta de metodologia automatizada para análise de usabilidade chegou a conclusões semelhantes as já encontradas na literatura. Além disso, foi possível identificar melhorias na metodologia proposta frente aos trabalhos analisados e no método híbrido computacional proposto (GHSOM + *K-means*).

Também foram objetivos desse trabalho mostrar que é possível utilizar técnicas de usabilidade mais simples que técnicas tradicionais com grande rigor metodológico em equipes sem especialistas de usabilidade e verificar a possibilidade de integração da abordagem de avaliação de usabilidade *Discount Usability* com a metodologia automatizada proposta. Um estudo de caso foi realizado com uma aplicação Web de modo a validar a metodologia proposta. Através desse estudo de caso foi possível verificar que a *Discount Usability* pode ser utilizada por pessoas não especialistas em usabilidade com treinamento básico na abordagem. O estudo de caso também verificou que foi possível integrar a abordagem automatizada a técnica de *Discount Usability*, o que traz melhores resultados para aplicação analisada e maior validade aos dados encontrados em ambas as avaliações. Assim como outros trabalhos indicam, concluímos que métodos híbridos de avaliação de usabilidade podem ser bastante úteis em análises de usabilidade.

Apesar da importância da usabilidade ser reconhecida e comprovada, é comum que seja deixada de lado ou tenha poucos recursos investidos, o que pode causar muitas iterações de manutenção na aplicação em busca de correções de erros. Um dos motivos que

causa isso é a dificuldade da maioria das equipes em integrar esse tipo prática em seus ciclos de vida, devido a falta de conhecimento. Nosso estudo de caso mostrou que a metodologia proposta pode ser utilizada por equipes com essa característica que desejam realizar análises de usabilidade em suas aplicações.

A técnica de avaliação automatizada se baseou em *Web Mining* de utilização, utilizando os *logs* de acesso do servidor *Web* da aplicação alvo da avaliação. A técnica de *Web Mining* selecionada foi a clusterização, utilizando um algoritmo híbrido com a rede neural GHSOM e o método K-means.

Para validação das técnicas utilizadas e da estratégia proposta nesse trabalho, utilizamos a aplicação SacarWeb como alvo do estudo de caso, uma aplicação Web da área da saúde. Foram realizadas duas avaliações, uma baseada na metodologia automatizada proposta e outra em *Discount Usability*. Usuários reais da aplicação foram selecionados para participar em todas as avaliações, seguindo critérios bem estabelecidos para estudo principalmente do aprendizado e taxa de erros da aplicação.

Na avaliação baseada em *Discount Usability*, foram realizadas avaliações heurísticas com três avaliadores, sendo dois com baixa experiência na área de usabilidade, e testes de usabilidade com os usuários do SacarWeb, com o mantenedor da aplicação sendo responsável pelos testes.

Os resultados da avaliação heurística e do teste de usabilidade apresentaram grande quantidade de problemas graves, o que é preocupante. Problemas de consistência e conceitos básicos da aplicação interpretados de modo errado pelos usuários são os principais problemas da aplicação, o que dificulta bastante a aprendizagem e interação dos usuários nas diversas telas do sistema. Outro ponto notável foi a grande quantidade de problemas relacionados a prevenção de erros e simplicidade dos diálogos. Durante as avaliações foi notável a necessidade do conhecimento específico dos usuários para a interação com o sistema, diferença que ficou bastante clara entre os testes de usabilidade piloto e o realizado com usuários alvo. Ambas as avaliações foram úteis e bem executadas pelos avaliadores, uma vez que há grande quantidade de problemas comuns entre elas. Também concluímos que a equipe do SacarWeb poderá utilizar os resultados satisfatoriamente em iterações de desenvolvimento futuras, uma vez que os problemas identificados facilmente serão transformados em tarefas de desenvolvimento (FERRACIOLI; BRUNETTO, 2011). Por possuírem boa fundamentação, as soluções também serão facilmente encontradas, pois os problemas estão bem documentados na literatura. Entretanto, como mostrou o capítulo 3.1.5, é comum que seja necessário transcrever os problemas de usabilidade de modo mais claro em forma de tarefas para a equipe

de desenvolvimento. Essa é uma etapa adicional deixada como sugestão para exploração em trabalhos futuros.

A abordagem automatizada utilizou usuários reais, que não puderam ser os mesmos que haviam participado da avaliação com *Discount Usability*. Buscou-se por padrões em tarefas definidas pelo mantenedor da aplicação e informadas ao protótipo. Essas informações são consideradas os padrões esperados do comportamento do usuário, e foram utilizadas como base dos filtros por caminhos percorridos. Posteriormente elas foram clusterizadas, sendo que os resultados apresentados pelo protótipo de software são os *clusters* de usuários agrupados por características de navegação comuns. Esses resultados tinham dois formatos, uma imagem com o gráfico dos *clusters* gerados e sua representatividade e os *clusters* com formato textual, com os caminhos e tempo de execução de cada representante do *cluster*. Esses resultados foram úteis para identificar padrões de erros nas tarefas, acessos a áreas do sistema não relacionados a tarefa sendo analisada, tempo de realização das tarefas e padrões de sucesso.

Ao relacionar os resultados da avaliação automatizada com as avaliações do *Discount Usability* foram encontradas várias semelhanças e explicações para resultados de ambas avaliações. Uma vez que principalmente os resultados dos testes de usabilidade são qualitativos, os dados quantitativos da avaliação automatizada foram bastante úteis para validar o que foi encontrado, com semelhanças também com a avaliação heurística, método bem fundamentado em boas práticas reconhecidas na literatura. Esse é um importante resultado, pois caso fosse necessário avaliar o desempenho dos usuários no sistema através da *Discount Usability*, mais um teste de usabilidade seria necessário. Outra vantagem é falta de desvios que ocorrem nos testes tradicionais, uma vez que além de ser incomum, não é confortável para o usuário utilizar o computador sendo observado por outra pessoa e dizendo o que está pensando.

Ainda em termos de validação da proposta, realizamos a validação dos algoritmos utilizados durante a clusterização. O algoritmo SOM + *K-means* já provou ser uma boa escolha de clusterização (KUO; HO; HU, 2002). Entretanto, a dificuldade em determinar a quantidade de neurônios na saída do algoritmo SOM permanece como uma dificuldade na proposta, portanto a proposta do algoritmo GHSOM + *K-means* é válida em termos de utilização, devido a característica do GHSOM em definir sua topologia durante o aprendizado. Além disso, utilizamos a mesma medida de confiança utilizada no desenvolvimento da proposta SOM + *K-means*, o valor lambda de Wilk. Segundo os experimentos realizados nesse trabalho, o algoritmo proposto foi melhor em 6 (75%) tarefas, obteve desempenho igual em

1 (12.5%) e foi pior somente em 1 (12.5%). Baseado no desempenho apresentado, o método proposto é válido para nosso estudo de caso pois obteve melhor desempenho de clusterização em relação ao SOM + *K-means*.

Finalmente concluímos que a proposta desse trabalho foi efetiva em definir uma metodologia de análise de usabilidade combinando o método automatizado proposto e a *Discount Usability*. Reconhecemos que os testes com usuários não devem ser eliminados de abordagens de usabilidade, mas podem ser melhorados em termos de velocidade e dados coletados com ferramentas automáticas, como a proposta por esse trabalho. E por fim, deixamos como sugestão de método de clusterização nossa proposta do GHSOM + *K-means*, que obteve melhor desempenho na maior parte dos casos de análise frente ao SOM + *K-means*.

5.1 Trabalhos Futuros

- É necessário testar o algoritmo proposto com uma quantidade de dados maior, pois como pode ser visto no capítulo 2.3, a maior parte dos algoritmos possuía grande quantidade de dados para análise;
- Apesar de nossa estratégia utilizar dados de uma aplicação real, sua utilização é limitada a um grupo muito pequeno, que não gera grandes quantidades de dados. Além de testar com uma maior quantidade de dados, o algoritmo proposto pode confrontado com propostas testadas também com o trabalho (KUO; HO; HU, 2002): são elas (KUO; LIAO; TU, 2005) e (KUO et al., 2006);
- Implementação alternativa do método de clusterização utilizando teoria dos conjuntos *fuzzy* em alguma das etapas devido ao seus bons resultados apresentados em (SONG; SHEPPERD, 2006);
- Implementação de uma estratégia para contornar o problema de tarefas com as URLs inicial e final iguais, o que impossibilitou a avaliação da usabilidade de duas tarefas através do método proposto. É possível que uma estratégia *look ahead*, onde a segunda URL poderia ser configurada como entrada opcional ao algoritmo, assim como o *Web Mining* de estrutura poderia auxiliar a contornar o mesmo problema;
- Adicionar mais parâmetros ao arquivo de configuração de modo a melhorar a automação da avaliação. Dois parâmetros representativos seriam a URL de ajuda e a de erros, caso ambas existam. Dessa maneira é mais simples rastrear a quantidade de vezes

que a ajuda foi acessada em um *cluster* assim como a quantidade de erros média para tarefa;

- Algumas abordagens de análise automatizada observadas no capítulo 2.3 buscavam separar a lógica da aplicação das métricas de usabilidade monitoradas, facilitando a definição de novas métricas ou medidas específicas por aplicação. Apesar de parcialmente implementado nesse trabalho, uma metodologia que aproveite todo o potencial dessa abordagem pode obter melhores resultados;
- Utilizar as outras abordagens de *Web Mining* pode ajudar a melhorar os resultados do algoritmo, como abordagem estrutural, que poderia ajudar a validar os caminhos percorridos pelos usuários e verificar se realmente o melhor caminho está sendo o mais comumente utilizado;
- Em aplicações genéricas a mineração de conteúdo também poderia ajudar a determinar, por exemplo, o interesse dos usuários na aplicação sendo avaliada. Resultados expressivos puderam ser observados em avaliações automatizadas de usabilidade que utilizaram mais de uma abordagem de *Web Mining* no capítulo 2.3;
- Projetos alternativos para as tarefas com os principais problemas do SacarWeb, que podem ser testadas através de protótipos através da *Discount Usability*, que pode ser integrada a equipe de desenvolvimento, em conjunto com a ferramenta desenvolvida nesse trabalho;
- Um estudo comparativo dos resultados alcançados por esse trabalho e novas métricas poderiam ser utilizadas, como a taxa de memorização dos usuários;
- Proposta de uma etapa adicional para a metodologia automatizada que transcreva os problemas de usabilidade tarefas para a equipe de desenvolvimento.

Referências

- AROTARITEI, D.; MITRA, S. Web mining: a survey in the fuzzy framework. *Fuzzy Sets and Systems*, v. 148, n. 1, p. 5–19, November 2004. ISSN 01650114.
- AU, F. T. W. et al. Automated usability testing framework. In: *Proceedings of the ninth conference on Australasian user interface - Volume 76*. Darlinghurst, Australia, Australia: Australian Computer Society, Inc., 2008. (AUIC '08), p. 55–64. ISBN 978-1-920682-57-6. Disponível em: <<http://portal.acm.org/citation.cfm?id=1378349>>.
- BADENOCH, D.; TOMLIN, A. How electronic communication is changing health care: Usability is main barrier to effective electronic information systems. *BMJ*, v. 328, n. 7455, p. 1564+, June 2004.
- BEIREKDAR, A. et al. Flexible reporting for automated usability and accessibility evaluation of web sites Human-Computer interaction - INTERACT 2005. In: COSTABILE, M.; PATERNÒ, F. (Ed.). *Human-Computer Interaction - INTERACT 2005*. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2005, (Lecture Notes in Computer Science, v. 3585). cap. 25, p. 281–294. ISBN 978-3-540-28943-2. Disponível em: <http://dx.doi.org/10.1007/11555261_25>.
- BEZDEK, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms (Advanced Applications in Pattern Recognition)*. [S.l.]: Springer, 1981. Hardcover. ISBN 0306406713.
- BLACK, L. F.; HYATT, R. E. Maximal respiratory pressures: normal values and relationship to age and sex. *The American review of respiratory disease*, v. 99, n. 5, p. 696–702, May 1969. ISSN 0003-0805.
- BRADLEY, P. S.; FAYYAD, U. M. Refining initial points for k-means clustering. In: *Proceedings of the 15th International Conference on Machine Learning (ICML98)*. [S.l.: s.n.], 1998. p. 91–99.
- BRINCK, T.; HOFER, E. Automatically evaluating the usability of web sites. In: *CHI '02 extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2002. (CHI EA '02), p. 906–907. ISBN 1-58113-454-1. Disponível em: <<http://dx.doi.org/10.1145/506443.506652>>.
- CAMARGO-BRUNETTO, M. A. de O.; ROSSI, A. R. Uma aplicação web para análise de exames cardiorrespiratórios. In: *Simpósio de Qualidade de Software - Workshop de Informática Médica: Anais do Workshop de Informática Médica*. [S.l.: s.n.], 2005.
- CAMELIER, A. et al. Avaliação da qualidade de vida pelo questionário do hospital saint george na doença respiratória em portadores de doença pulmonar obstrutiva crônica: validação de uma nova versão para o brasil. *Jornal Brasileiro de Pneumologia*, v. 32, n. 2, April 2006. ISSN 1806-3713.

- CARROLL, J. M. Making use : scenario-based design of human-computer interactions. In: _____. 1st. ed. [S.l.]: MIT Press, 2000. ISBN 0262032791.
- CARVALHO, J. O. F. de. O papel da interação humano-computador na inclusão digital. *Transinformação*, v. 15, n. 3, p. 75–89, 2003. ISSN 0103-3786.
- CATLEDGE, L. D.; PITKOW, J. E. Characterizing browsing strategies in the world-wide web. *Computer Networks and ISDN Systems*, v. 27, n. 6, p. 1065–1073, April 1995. ISSN 01697552.
- CHEN, M.-S.; PARK, J. S.; YU, P. S. Efficient data mining for path traversal patterns. *IEEE Transactions on Knowledge and Data Engineering*, v. 10, n. 2, p. 209–221, March 1998. ISSN 10414347. Disponível em: <<http://dx.doi.org/10.1109/69.683753>>.
- COOLEY, R.; MOBASHER, B.; SRIVASTAVA, J. Web mining: Information and pattern discovery on the world wide web. *IEEE International Conference Tools with Artificial Intelligence*, IEEE Computer Society, Los Alamitos, CA, USA, p. 0558–567, 1997.
- COOLEY, R.; MOBASHER, B.; SRIVASTAVA, J. Data preparation for mining world wide web browsing patterns. *Knowledge and Information Systems*, v. 1, p. 5–32, 1999.
- COYETTE, A.; KIEFFER, S.; VANDERDONCKT, J. Multi-fidelity prototyping of user interfaces. In: BARANAUSKAS, C. et al. (Ed.). *Human-Computer Interaction – INTERACT 2007*. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2007, (Lecture Notes in Computer Science, v. 4662). cap. 16, p. 150–164. ISBN 978-3-540-74794-9.
- CRONIN, B. Annual review of information science and technology. In: _____. [S.l.]: Information Today Inc., 2003. cap. Web Mining: Machine Learning for Web Applications, p. 289–330.
- DITTENBACH, M.; MERKL, D.; RAUBER, A. The growing hierarchical self-organizing map. In: *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. [S.l.]: IEEE Comput. Soc, 2000. p. 15–19. ISBN 0-7695-0619-4.
- DITTENBACH, M.; RAUBER, A.; MERKL, D. Recent advances with the growing hierarchical self-organizing map. In: ALLINSON, N. et al. (Ed.). *Proceedings of the 3rd Workshop on Self-Organizing Maps*. Lincoln, England: Springer, 2001. (Advances in Self-Organizing Maps), p. 140–145.
- DUNN, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, Taylor & Francis, v. 3, n. 3, p. 32–57, 1973. Disponível em: <<http://dx.doi.org/10.1080/01969727308546046>>.
- FERRACIOLI, F.; BRUNETTO, M. A. de O.-C. Using and integrating discount usability engineering in the life cycle of a health care web application. In: VERDIER, C.; THOM, L. H.; RONCANCIO, C. (Ed.). *Proceedings of the Health Ambient Information Systems Workshop, HamIS 2011*. Gramado, Brazil: CEUR-WS.org, 2011. ISSN 1613-0073.
- FERRE, X. et al. Usability basics for software developers. *IEEE Software*, v. 18, n. 1, p. 22–29, Jan 2001. ISSN 07407459.

- FISHER, D. H. Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, Springer Netherlands, v. 2, n. 2, p. 139–172, sep 1987. ISSN 0885-6125. Disponível em: <<http://dx.doi.org/10.1007/BF00114265>>.
- FRITZKE, B. Growing grid - a self-organizing network with constant neighborhood range and adaptation strength. *Neural Processing Letters*, v. 2, n. 5, p. 9–13, September 1995. ISSN 1370-4621.
- FU, Y.; SANDHU, K.; SHIH, M. yi. Clustering of web users based on access patterns. In: . [S.l.: s.n.], 1999.
- PBIRCH: A Scalable Parallel Clustering algorithm for Incremental Data*. 315–316 p.
- GUYATT, G. H. et al. A measure of quality of life for clinical trials in chronic lung disease. *Thorax*, v. 42, n. 10, p. 773–778, October 1987. ISSN 0040-6376.
- HAFRI, Y. et al. A markovian approach for web user profiling and clustering. In: . [S.l.: s.n.], 2003. p. 563.
- HARTIGAN, J. A.; WONG, M. A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Blackwell Publishing for the Royal Statistical Society, v. 28, n. 1, p. 100–108, 1979. ISSN 00359254.
- HAYKIN, S.; ENGEL, P. M. *Redes neurais : princípios e prática*. Bookman, 2001. ISBN 8573077182. Disponível em: <<http://www.worldcat.org/isbn/8573077182>>.
- HOLZINGER, A. Usability engineering methods for software developers. *Commun. ACM*, ACM, New York, NY, USA, v. 48, n. 1, p. 71–74, January 2005. ISSN 0001-0782.
- HORNBAEK, K.; FRØKJAER, E. Comparing usability problems and redesign proposals as input to practical systems development. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 2005. (CHI '05), p. 391–400. ISBN 1-58113-998-5. Disponível em: <<http://dx.doi.org/10.1145/1054972.1055027>>.
- IVORY, M. Y.; HEARST, M. A. The state of the art in automating usability evaluation of user interfaces. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 33, n. 4, p. 470–516, December 2001. ISSN 0360-0300.
- JOHNSON, G. I. Usability evaluation in industry. In: _____. [S.l.]: Taylor & Francis Ltd., 1996. cap. The usability checklist approach revisited, p. 179–188.
- JONES, P.; QUIRK, F.; BAVEYSTOCK, C. The st george's respiratory questionnaire. *Respiratory Medicine*, v. 85, p. 25–31, September 1991. ISSN 09546111.
- JURISTO, N.; MORENO, A.; SEGURA, M. I. S. Guidelines for eliciting usability functionalities. *IEEE Transactions on Software Engineering*, IEEE Computer Society, Los Alamitos, CA, USA, v. 33, n. 11, p. 744–758, 2007. ISSN 0098-5589. Disponível em: <<http://dx.doi.org/10.1109/TSE.2007.70741>>.
- KANUNGO, T. et al. An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Washington, DC, USA, v. 24, n. 7, p. 881–892, July 2002. ISSN 0162-8828.

- KNUDSON, R. J. et al. The maximal expiratory flow-volume curve. normal standards, variability, and effects of age. *The American review of respiratory disease*, v. 113, n. 5, p. 587–600, may 1976. ISSN 0003-0805. Disponível em: <<http://view.ncbi.nlm.nih.gov/pubmed/1267262>>.
- KOHONEN, T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, v. 43, n. 1, p. 59–69, January 1982. ISSN 0340-1200.
- KOLARI, P.; JOSHI, A. Web mining: research and practice. *Computing in Science & Engineering*, v. 06, n. 4, p. 49–53, June 2004.
- KORT, J.; POOT, H. de. Usage analysis: combining logging and qualitative methods. In: *CHI '05 extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2005. (CHI EA '05), p. 2121–2122. ISBN 1-59593-002-7. Disponível em: <<http://dx.doi.org/10.1145/1056808.1057117>>.
- KOSALA, R.; BLOCKEEL, H. Web mining research: a survey. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 2, n. 1, p. 1–15, June 2000. ISSN 1931-0145.
- KRISHNA, K.; MURTY, M. N. Genetic k-means algorithm. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, v. 29, n. 3, p. 433–439, August 2002.
- KUO, R. J. et al. Integration of self-organizing feature maps neural network and genetic k-means algorithm for market segmentation. *Expert Systems with Applications*, v. 30, n. 2, p. 313–324, February 2006. ISSN 09574174.
- KUO, R. J.; HO, L. M.; HU, C. M. Integration of self-organizing feature map and k-means algorithm for market segmentation. *Computers & Operations Research*, v. 29, n. 11, p. 1475–1493, September 2002. ISSN 03050548.
- KUO, R. J.; LIAO, J. L.; TU, C. Integration of art2 neural network and genetic k-means algorithm for analyzing web browsing paths in electronic commerce. *Decision Support Systems*, v. 40, n. 2, p. 355–374, August 2005. ISSN 01679236.
- KUSHNIRUK, A. W.; PATEL, V. L.; CIMINO, J. J. Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. *Proceedings : a conference of the American Medical Informatics Association / ... AMIA Annual Fall Symposium. AMIA Fall Symposium*, p. 218–222, 1997. ISSN 1091-8280.
- LI, C. hung; KIT, C. chun. Web structure mining for usability analysis. In: *WI '05: Proceedings of the The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*. Washington, DC, USA: IEEE Computer Society, 2005. p. 309–312. ISBN 076952415X.
- LIANGUN, L.; BIN, Z.; YUANYUAN, C. The improved algorithm of art2 in data mining. In: *2009 First International Workshop on Database Technology and Applications*. [S.l.: s.n.], 2009. p. 177–180.
- LINGRAS, P.; WEST, C. Interval set clustering of web users with rough k-means. *Journal of Intelligent Information Systems*, Kluwer Academic Publishers, v. 23, n. 1, p. 5–16, July 2004. ISSN 0925-9902.
- LIPPMANN, R. P. An introduction to computing with neural nets. *IEEE ASSP Magazine*, v. 4, n. 2, p. 4–22, April 2003.

- LLOYD, S. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–137, March 1982. ISSN 0018-9448.
- MACKAY, D. J. C. Information theory, inference, and learning algorithms. In: _____. 1st. ed. [S.l.]: Cambridge University Press, 2003. cap. An Example Inference Task: Clustering, p. 284–292. ISBN 0521642981.
- MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: CAM, L. M. L.; NEYMAN, J. (Ed.). [S.l.: s.n.], 1967. v. 1, p. 281–297.
- MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 18, n. 1, p. 50–60, 1947. ISSN 00034851.
- MARTY, P. F.; TWIDALE, M. B. *Extreme Discount Usability Engineering*. [S.l.], 2005. Disponível em: <<http://www.isrl.uiuc.edu/~twidale/pubs/ExtremeDisc>>.
- MAYHEW, D. J. *The usability engineering lifecycle : a practitioner's handbook for user interface design*. [S.l.]: Morgan Kaufmann Publishers, 1999. Paperback. ISBN 1558605614.
- MILLER, M. R. et al. Standardisation of spirometry. *European Respiratory Journal*, v. 25, n. 2, p. 319–338, August 2005. ISSN 1399-3003.
- MOBASHER, B. et al. Web mining: Pattern discovery from world wide web transactions. 1996. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.57.4087>>.
- MORANDINI, M. *Ergo-Monitor: Monitoramento da usabilidade em ambiente Web por meio de análise de arquivos de Log*. Tese (Doutorado) — Universidade Federal de Santa Catarina, November 2003.
- MOREIRA, G. L. et al. Versão em português do chronic respiratory questionnaire: estudo da validade e reprodutibilidade. *Jornal Brasileiro de Pneumologia*, v. 35, n. 8, August 2009. ISSN 1806-3713.
- NEDER, J. A. et al. Reference values for lung function tests: li. maximal respiratory pressures and voluntary ventilation. *Brazilian Journal of Medical and Biological Research*, v. 32, n. 6, June 1999. ISSN 0100-879X.
- NIELSEN, J. *How to Conduct a Heuristic Evaluation*. [S.l.]. Disponível em: <http://www.useit.com/papers/heuristic/heuristic_evaluation.html>.
- NIELSEN, J. *Severity Ratings for Usability Problems*. [S.l.]. Disponível em: <<http://www.useit.com/papers/heuristic/severityrating.html>>.
- NIELSEN, J. Finding usability problems through heuristic evaluation. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 1992. (CHI '92), p. 373–380. ISBN 0-89791-513-5.
- NIELSEN, J. *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. [S.l.], 1994. Disponível em: <<http://www.useit.com/papers/guerrilahci.html>>.
- NIELSEN, J. Usability inspection methods. In: *CHI '94: Conference companion on Human factors in computing systems*. [S.l.]: ACM, 1994. p. 413–414. ISBN 0-89791-651-4.

- NIELSEN, J. *Usability Engineering*. AP Professional, 1998. ISBN 1125184069. Disponível em: <<http://www.worldcat.org/isbn/1125184069>>.
- NIELSEN, J. *Return on Investment for Usability*. [S.l.], 2003. Disponível em: <<http://www.useit.com/alertbox/roi-first-study.html>>.
- NIELSEN, J. *Usability ROI Declining, But Still Strong*. [S.l.], 2008. Disponível em: <<http://www.useit.com/alertbox/roi.html>>.
- NIELSEN, J. *Discount Usability: 20 Years*. [S.l.], 2009. Disponível em: <<http://www.useit.com/alertbox/discount-usability.html>>.
- NIELSEN, J.; LORANGER, H. *Prioritizing Web usability*. 1. ed. [S.l.]: New Riders, 2006. Paperback. ISBN 0321350316.
- NIELSEN, J.; MOLICH, R. Heuristic evaluation of user interfaces. In: *CHI '90: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 1990. p. 249–256. ISBN 0-201-50932-6.
- O'REILLY, T. *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. [S.l.], 2005. Disponível em: <<http://www.oreillynet.com/pub/a-oreilly/tim/news/2005/09/30/what-is-web-20.html>>.
- PAL, S. K.; TALWAR, V.; MITRA, P. Web mining in soft computing framework: relevance, state of the art and future directions. *IEEE Transactions on Neural Networks*, v. 13, n. 5, p. 1163–1177, November 2002.
- PAWLAK, Z. Rough set approach to knowledge-based decision support. *European Journal of Operational Research*, v. 99, n. 1, p. 48–57, May 1997. ISSN 03772217.
- PAYNE, G. E.; SKEHAN, J. D. Shuttle walking test: a new approach for evaluating patients with pacemakers. *Heart*, v. 75, n. 4, p. 414–418, apr 1996. Disponível em: <<http://dx.doi.org/10.1136/hrt.75.4.414>>.
- PEREIRA, C. A.; SATO, T.; RODRIGUES, S. C. Novos valores de referência para espirometria forçada em brasileiros adultos de raça branca. *Jornal Brasileiro de Pneumologia*, v. 33, n. 4, August 2007. ISSN 1806-3713.
- RAJU, G. T. et al. Knowledge discovery from web usage data: Extraction of sequential patterns through art1 neural network based clustering algorithm. *International Conference on Computational Intelligence and Multimedia Applications*, IEEE Computer Society, Los Alamitos, CA, USA, v. 2, p. 88–92, 2007.
- RANGARAJAN, S. K. et al. Adaptive neural network clustering of web users. *Computer*, v. 37, n. 4, p. 34–40, August 2004.
- RAUBER, A.; MERKL, D.; DITTENBACH, M. The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. *IEEE Transactions on Neural Networks*, v. 13, n. 6, p. 1331–1341, November 2002. ISSN 1045-9227.
- ROSE, A. F. et al. Using qualitative studies to improve the usability of an EMR. *Journal of Biomedical Informatics*, v. 38, n. 1, p. 51–60, feb 2005. ISSN 15320464. Disponível em: <<http://dx.doi.org/10.1016/j.jbi.2004.11.006>>.

- SCHWERZ, A.; MORANDINI, M.; SILVA, S. da. A task model proposal for web sites usability evaluation for the ergomonitor environment. In: . [S.l.: s.n.], 2007. p. 1188–1197.
- SEFFAH, A.; DESMARAIS, M. C.; GULLIKSEN, J. *Human-Centered Software Engineering Integrating Usability in the Software Development Lifecycle*. Springer, 2005. ISBN 9781402040276. Disponível em: <<http://www.worldcat.org/isbn/9781402040276>>.
- SILVINO, A. M.; ABRAHÃO, J. I. Navegabilidade e inclusão digital: usabilidade e competência. *RAE eletrônica*, v. 2, n. 2, December 2003. ISSN 1676-5648.
- SINGH, S. J. et al. Comparison of oxygen uptake during a conventional treadmill test and the shuttle walking test in chronic airflow limitation. *European Respiratory Journal*, v. 7, n. 11, p. 2016–2020, November 1994.
- SINGH, S. J. et al. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax*, v. 47, n. 12, p. 1019–1024, December 1992.
- SMITH, K. A.; NG, A. Web page clustering using a self-organizing map of user navigation patterns. *Decision Support Systems*, v. 35, n. 2, p. 245–256, May 2003. ISSN 01679236.
- SOCIETY, A. T. Ats statement: Guidelines for the six-minute walk test. *American Journal of Respiratory and Critical Care Medicine*, v. 166, n. 1, p. 111–117, 2002.
- SONG, Q.; SHEPPERD, M. Mining web browsing patterns for e-commerce. *Computers in Industry*, v. 57, n. 7, p. 622–630, September 2006. ISSN 01663615.
- SORIANO-ASENSI, A. et al. Web mining based on growing hierarchical self-organizing maps: Analysis of a real citizen web portal. *Expert Systems with Applications*, v. 34, n. 4, p. 2988–2994, May 2008. ISSN 09574174.
- SOUSA, T. C.; JARDIM, J. R.; JONES, P. Validação do questionário do hospital saint george na doença respiratória (sgrq) em pacientes portadores de doença pulmonar obstrutiva crônica no brasil. *Jornal de Pneumologia*, v. 26, n. 3, June 2000. ISSN 0102-3586.
- SPILIOPOULOU, M. Web usage mining for web site evaluation. *Communications of the ACM*, ACM Press, New York, NY, USA, v. 43, n. 8, p. 127–134, August 2000. ISSN 00010782.
- SRIVASTAVA, J. et al. Web usage mining: discovery and applications of usage patterns from web data. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 1, n. 2, p. 12–23, January 2000. ISSN 1931-0145.
- SRIVASTAVA, J.; DESIKAN, P.; KUMAR, V. Web mining - concepts, applications and research directions. In: . [S.l.: s.n.], 2005. p. 275–307.
- STANDARDIZATION, I. O. for. *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*. 1. ed. [S.l.], 1998.
- STEINHAUS, H. Sur la division des corps materiels en parties. *Bull. Acad. Pol. Sci., Cl. III*, v. 4, p. 801–804, 1957.
- TIEDTKE, T.; MÄRTIN, C.; GERTH, N. Awusa - a tool for automated website usability analysis. In: *PreProceedings of the 9 th International Workshop on the Design, Specification and Verification of Interactive Systems*. [S.l.: s.n.], 2002.

- TISIOLOGIA, S. B. de Pneumologia e. *I Consenso Brasileiro sobre Espirometria*. 3. ed. [S.l.], 1996.
- TROOSTERS, T.; GOSSELINK, R.; DECRAMER, M. Six minute walking distance in healthy elderly subjects. *European Respiratory Journal*, European Respiratory Journal, v. 14, n. 2, p. 270–274, 1999.
- WALKER, M.; TAKAYAMA, L.; LANDAY, J. A. High-fidelity or low-fidelity, paper or computer choosing attributes when testing web prototypes. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, Human Factors and Ergonomics Society, p. 661–665, 2002. ISSN 1071-1813. Disponível em: <<http://www.ingentaconnect.com/content/hfes/hfproc/2002/00000046/00000005/art00013>>.
- WANG, D. Pattern recognition: neural networks in perspective. *IEEE Expert*, v. 8, n. 4, p. 52–60, August 2002.
- WINCKLER, M.; PIMENTA, M. S. Avaliação de usabilidade de sites web. In: *Livro Texto da X Escola de Informática da SBC-Sul*. [S.l.]: Sociedade Brasileira de Computação, 2002. p. 85–137.
- XU, R.; WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, v. 16, n. 3, p. 645–678, May 2005. ISSN 1045-9227.
- XU, R.; WUNSCH, D. C.; SOCIETY, I. C. I. *Clustering*. Wiley, 2009. ISBN 9780470276808. Disponível em: <<http://www.worldcat.org/isbn/9780470276808>>.
- ZHANG, T.; RAMAKRISHNAN, R.; LIVNY, M. Birch: an efficient data clustering method for very large databases. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 25, n. 2, p. 103–114, 1996. ISSN 0163-5808.

APÊNDICE A – Guia para avaliação heurística do SacarWeb

Este documento é um guia para a avaliação heurística do SacarWeb (Software para Avaliação Cardio-Respiratória na Web). Pretende-se avaliar principalmente o nível de aprendizagem em cenários bem definidos, com base nas heurísticas definidas em (NIELSEN, 1998), que também são descritas nesse trabalho em 2.1.3.1. Adicionalmente é interessante levar em conta o nível de memorização e erros. Essas características são consideradas importantes devido a alta rotatividade de usuários do sistema, pois ao início de cada ano letivo, novos usuários ingressam no sistema. Além disso, fora do período letivo o sistema é minimamente utilizado, e ao retorno do período letivo antigos usuários deverão lembrar como utilizar o sistema facilmente.

Os principais usuários do sistema são graduandos de Fisioterapia da Universidade Estadual de Londrina que participam de pesquisas desenvolvidas no Laboratório de Fisioterapia Pulmonar (LFIP). Alguns usuários são alunos de residência em Fisioterapia Respiratória, mestrado e doutorado. A principal função do SacarWeb é manter um banco de dados para pesquisas baseadas nas avaliações clínicas executadas em pacientes atendidos pelo ambulatório de fisioterapia do Hospital Universitário. Elas servem como base de estudo no laboratório, o que pode ser feito através do acompanhamento de pacientes individualmente, ou grupos de pacientes. Durante a realização das avaliações clínicas os pesquisadores preenchem os dados em formulários de papel, que posteriormente são inseridos no SacarWeb pelos mesmos usuários.

Por se tratar de uma aplicação Web espera-se que seu funcionamento seja independente de ambiente, mas por questões de padronização das condições de uso entre os avaliadores do software e aproximação do ambiente em que os usuários reais da aplicação se encontram, para essa avaliação será necessário utilizar o seguinte ambiente:

- Dispositivo compatível com mouse e teclado

- Navegador Mozilla Firefox (Versão Mínima 3.6), com Javascript e *cookies* ativados
- Resolução de tela mínima 1024x768

Recomenda-se eliminar possibilidades de interferência externa durante a avaliação. Por isso deve-se evitar a entrada repentina de pessoas no local onde a avaliação será executada, desligar telefones, celulares e outros dispositivos que possam dispersar o avaliador. Também é recomendado executar somente o navegador com o SacarWeb durante a avaliação, evitando que alertas de outros programas desviem a atenção durante a avaliação.

A.1 Cenários de utilização

Dentro do SacarWeb os avaliadores do software deverão se portar como um examinador clínico interessado em obter dados amostrais para uma pesquisa científica. Os dados dos pacientes e suas avaliações, que estarão em fichas de papel, serão transcritos para os locais apropriados dentro do software.

Para utilizar o sistema serão fornecidos os seguintes dados:

- URL: <http://www.protem.dc.uel.br/sacarweb-alt>
- Login: avaliador_ux
- Senha: heuristica

Uma vez no sistema, deverão ser executadas as seguintes ações em ordem, para cada paciente:

1. Inserir os dados pessoais do paciente;
2. Preencher a avaliação inicial;
3. Criar uma nova avaliação;
4. Informar os dados dos exames: PFP, Pressões Respiratórias, SGRQ, CRQ, Shuttle e DP6;
5. Verificar a evolução dos pacientes através das diversas avaliações;
6. Obter um relatório de dados que possa ser utilizado como amostragem de pesquisa para os exames executados.

A.1.1 Inserir os dados pessoais do paciente

Neste ponto um registro eletrônico do paciente é criado no SacarWeb. Após esse cadastro, todas as informações referentes ao paciente poderão ser acessadas através desse registro. Esse é um cadastro básico, e somente informações gerais são inseridas, mas note para a inserção dos valores corretos pois vários campos são utilizados durante as avaliações. Durante a execução da tarefa o campo de código do paciente deverá ter um valor sequencial, baseado no último valor encontrado na listagem de paciente. Também poderão ser ignorados os campos e-mail, raça e tamanho do passo.

A.1.2 Preencher a avaliação inicial

A avaliação inicial é a primeira avaliação realizada com um paciente. Nela questões sobre a histórico de doenças, saúde atual e hábitos do paciente são analisados. Essas informações podem ter influência direta sobre os resultados de exames, assim como em sua interpretação. Além disso podem constituir dados relevantes para análise durante pesquisas. Durante o cadastro da avaliação inicial, caso algum dos campos de peso não esteja informado, o valor encontrado em uma das fichas de avaliação, preferencialmente na imediatamente anterior, deverá ser repetido.

A.1.3 Criar uma nova avaliação

No SacarWeb uma avaliação é composta por uma bateria de exames realizados a partir de uma data base. Uma vez que os exames que constituem uma avaliação são realizados em datas próximas, as variáveis dependentes de data estarão relacionadas a data base da avaliação, e não a data do exame. Ela é a data na qual foi feita o primeiro exame dentre o conjunto de exames. Essa data normalmente está especificada na ficha de dados pessoais do paciente, mas caso não esteja disponível deverá ser utilizada a menor data entre os exames que constituem a avaliação. As variáveis peso e altura também possuem os valores fixados de acordo com o que foi informado em cada avaliação.

A.1.4 Informar os dados dos exames

É permitida somente uma instância de cada exame em uma avaliação. Cada exame possui um objetivo distinto e analisa um determinado ponto de interesse para a pesquisa, e uma determinada característica da enfermidade do paciente. Todos são fundamentados em

literatura científica e possuem seus métodos, cálculos e diagnósticos bem documentados. É possível que em alguns exames não esteja disponível a data de sua realização, nesses casos a data base deverá ser informada como data de realização do exame.

A.1.4.1 PFP

Espirometria, ou PFP, é um teste fisiológico que avalia a função pulmonar, mensurando os volumes e capacidades pulmonares que o indivíduo inspira e expira, em função do tempo. Somente os aspectos mais importantes da espirometria são de interesse para pesquisa, são eles (MILLER et al., 2005), (PEREIRA; SATO; RODRIGUES, 2007):

- Capacidade Vital Forçada (CVF)
- Volume de Expiração forçada em um segundo (VEF1)
- VEF1/CVF

Há uma diferença de nomenclatura entre o formulário impresso e o formulário digital, pois o impresso utiliza nomenclatura em inglês, enquanto o SacarWeb em português. Os campos de interesse serão encontrados como:

- CVF = VFC
- VEF1 = FEV1

Somente duas colunas são transcritas do formulário impresso para o sistema: PRE e POS.

A.1.4.2 Pressões Respiratórias

A medida das Pressões Respiratórias Estáticas Máximas é um exame que mede indiretamente a força dos músculos respiratórios a partir das pressões geradas, ao nível da boca, durante uma inspiração e expiração máximas. (BLACK; HYATT, 1969), (NEDER et al., 1999) Estarão disponíveis diversas medidas de pressões inspiratória e expiratória. O valor selecionado deve ser o maior, desde que haja uma diferença máxima de 10% entre o maior valor e o valor mais próximo menor que ele.

A.1.4.3 SGRQ

O Questionário do Hospital Saint George na Doença Respiratória (SGRQ) foi escrito originalmente em inglês britânico e validado para o Brasil em 2000 (SOUSA; JARDIM;

JONES, 2000). Ele é específico para avaliar a qualidade de vida em portadores de doenças pulmonares obstrutivas crônicas. Entretanto, também tem sido utilizado em portadores de doenças pulmonares restritivas.

O SGRQ aborda os aspectos relacionados a três domínios: sintomas, atividades e impactos psicossociais que a doença respiratória causa ao paciente, divididos em 76 itens.

Cada domínio tem uma pontuação máxima possível (sintomas, 662,5 pontos; atividades, 1.209,1 pontos; e impactos, 2.117,8 pontos). Com isso, os pontos de cada resposta são somados, e o total é referido como um percentual desse máximo (0-100%). Além das pontuações de cada domínio, um score total também é calculado, baseado nos resultados dos três domínios (0-3.989,4 pontos). Uma pontuação maior significa pior qualidade de vida. (JONES; QUIRK; BAVEYSTOCK, 1991), (CAMELIER et al., 2006)

As sentenças das questões do formulário impresso podem estar ligeiramente diferentes das encontradas no SacarWeb, entretanto possuem o mesmo significado. No formulário impresso, a questão 6 possui somente 4 respostas possíveis, apesar de ter 5 campo possíveis para preenchimento. Devido a isso, tenha atenção ao transcrever a resposta dessa questão no sistema. Caso não haja resposta, selecione o campo "1 ou 2 dias".

A.1.4.4 CRQ

O Chronic Respiratory Questionnaire (CRQ, Questionário sobre Doenças Respiratórias Crônicas) é um questionário de qualidade de vida doença-específico respiratório. Foi o primeiro instrumento desenvolvido para mensurar a qualidade de vida em pacientes com Doença Pulmonar Obstrutiva Crônica (DPOC) (GUYATT et al., 1987).

O CRQ contém 20 questões, as quais estão divididas em quatro domínios: dispnéia (5 questões), fadiga (4 questões), função emocional (7 questões) e autocontrole (4 questões). O domínio dispnéia é o único individualizado, ou seja, cada paciente seleciona, em uma lista de 26 itens, as atividades que provocaram dispneia nas últimas duas semanas, podendo relatar também outras atividades que não se encontram na lista. Subsequentemente, dentre as atividades relatadas e selecionadas, o indivíduo escolhe cinco atividades que considera mais importantes e, por meio de uma escala de 7 pontos, gradua a sua dispneia para cada uma dessas atividades. Nessa escala, a pontuação vai de 1 (máximo comprometimento) a 7 (nenhum comprometimento).

Para os demais domínios (fadiga, função emocional e autocontrole), as ques-

tões são padronizadas e o paciente responde a cada questão utilizando a escala de 7 pontos. Os resultados são expressos em média da pontuação de cada domínio. Dessa forma, quanto maior a pontuação obtida, melhor a qualidade de vida do indivíduo (MOREIRA et al., 2009).

A ficha de avaliação do CRQ possui cinco campos que especificam as atividades que causam falta de ar no paciente. No sistema, deverá ser selecionada a atividade que mais se aproximar da resposta contida na folha, pois nem sempre são exatamente coincidentes. Para atividades em branco, a opção “nenhuma falta de ar” deverá ser selecionada. Ao selecionar essa opção, seu valor é preenchido automaticamente. Os valores obtidos para cada ítem estarão disponíveis no quadro “Pontuação das questões”, onde as cinco primeiras respostas são relacionadas as questões de falta de ar.

A.1.4.5 Shuttle

O *Shuttle Walking Test* constitui um teste padronizado de campo para caminhada incremental que provoca um sintoma de desempenho máximo limitado em pacientes com DPOC. Através dele é possível mensurar objetivamente a deficiência e comparar diretamente a capacidade de cada paciente (SINGH et al., 1992). Por ser possível mensurar a capacidade funcional ele também permite medir a capacidade cardiorespiratória de um paciente (SINGH et al., 1994). É recomendado para avaliação objetiva de pacientes com grande variação de deficiência respiratória.

Duas fichas de execução do exame estarão disponíveis, ao informar no sistema utilize a que possui melhor resultado, o que é verificado através do maior estágio atingido. Neste exame, FC é uma abreviação para Frequência Cardíaca. Já PA, Pressão Arterial, que no impresso é informada no formato xx/yy, onde:

- xx denota pressão sistólica
- yy denota pressão diastólica

Outras equivalências entre a versão impressa e do sistema são:

- SpO_2 - Saturação de Pulso de Oxigênio
- Borg D - Escala de Borg de Dispneia
- Borg F - Escala de Borg de Esforço

As informações do impresso contidas na coluna depois da primeira tabela do exame devem ser informadas no sistema no estágio alcançado pelo paciente. Já as informações de recuperação devem ser colocadas no campo de recuperação após o primeiro minuto. Valores fracionários na Escala de Borg devem ser arredondados para o inteiro mais próximo.

A.1.4.6 DP6

O Teste de seis minutos de caminhada (DP6) é um teste comumente utilizado para estimar a capacidade funcional de exercício em pacientes com doenças crônicas, como a DPOC. É um teste indicado para medir a resposta a intervenções médicas em pacientes com doenças do coração ou pulmonares moderadas ou severas. Também é utilizado como medida do estado funcional dos pacientes, assim como prediz a morbidade e mortalidade (TROOSTERS; GOSELINK; DECRAMER, 1999), (SOCIETY, 2002).

Duas fichas de execução do exame estarão disponíveis, ao transcrever os valores para o sistema considere os que possuem melhor resultado. Essa seleção deve ser feita com base na maior distância percorrida. Esse exame é preenchido da seguinte maneira: Cada linha da primeira coluna da tabela impressa possui valores que representam a distância percorrida em metros, iniciados em [30, 60, ..., 780], com incrementos de 30. Seis delas serão marcadas manualmente, uma a cada minuto da caminhada do paciente, com um ponto ou um valor no formato +xx. No primeiro caso, o paciente atingiu exatamente o valor da linha, já no segundo, ele atingiu o valor da linha e ultrapassou xx adicionais daquele valor. A segunda coluna representa o valor da Saturação de Pulso de Oxigênio (SpO_2), e a terceira a Frequência Cardíaca (FC). Novamente Borg D é equivalente a Escala de Borg de Dispneia e Borg F a Escala de Borg de Esforço, e devem ser arredondados para inteiro mais próximo no caso de valores fracionários. As informações do campo "Depois" são informadas na linha que representa o sexto minuto, já as do campo "Recuperação" são inseridas na linha recuperação no primeiro minuto. Em alguns casos os valores de recuperação podem estar divididos em duas colunas, neste caso a primeira coluna representa a recuperação no primeiro minuto, enquanto a segunda coluna representa a recuperação no segundo minuto.

A.1.5 Verifique a evolução dos pacientes entre as avaliações

Após um paciente ter sido avaliado mais de uma vez, é possível acompanhar sua evolução através de gráficos para cada exame. Os gráficos contém os dados informados nos formulários de exames, além de comparações analíticas e valores médios exibidos em tabelas. A partir desses recursos espera-se facilitar o acompanhamento de um paciente e a obtenção

de dados específicos.

A.1.6 Obtenha um relatório de dados

Pesquisas com dados selecionados anonimamente podem ser realizadas no SacarWeb. Essa é uma funcionalidade que pretende disponibilizar informações específicas e confiáveis através de filtros. Essas características são possíveis pois os dados são retirados de dados reais de pacientes. Pretende-se que essas informações possam ser utilizadas como fonte segura para dados de pesquisa.

A.2 Protocolo de execução da avaliação

Todos os passos descritos em A.1 deverão ser realizados duas vezes, exceto o passo 6 que será realizado somente uma. Sendo assim, cada avaliador possuirá dois pacientes, cujas fichas serão compostas por três avaliações distintas. Desse modo, os passos 3 e 4 deverão ser repetidos 3 vezes para cada paciente, pois somente assim será possível ter uma amostragem mínima para realização dos passos 5 e 6.

Com os dados dos exames de todos os pacientes inseridos será possível obter melhores resultados nos passos 5 e 6. É comum que essas telas sejam consultadas somente após a inserção de todos os dados, pois apresentam de modo objetivo informações de várias avaliações. No passo 5, é possível acompanhar a evolução dos valores obtidos por um paciente em um exame determinado através de várias avaliações. Essa tarefa será executada ao final das repetições dos passos anteriores.

Já no passo 6 é possível filtrar dados de toda a base de dados da instituição a qual o usuário pertence, dados que podem ser utilizados como fonte de pesquisa. Uma vez que é possível que os resultados desejados para um filtro sejam vazios, durante a realização da avaliação heurística deve-se atentar para utilização de valores semelhantes aos encontrados durante a inserção das avaliações dos pacientes. É importante notar que os valores apresentados como resultado dependem da quantidade de dados disponíveis para o filtro selecionado. Essa tarefa será executada após a inserção de todos os dados disponíveis ao avaliador do software, o que garante uma maior quantidade de dados para aplicação dos filtros.

Uma vez que as características analisadas pelas heurísticas são abrangentes, será necessário executar alguns passos mais de uma vez para observar se determinada interface atende ou não uma heurística específica. Um exemplo comum são as inserções, que deverão

A.2 Protocolo de execução da avaliação

ser executadas com dados inválidos e válidos, de modo a verificar o comportamento do sistema nas duas ocasiões.

Os avaliadores terão a sua disposição o atual mantenedor do SacarWeb, que conduzirá a avaliação e auxiliará os avaliadores no caso de dúvidas que os impeça de executar uma tarefa.

O avaliador deverá produzir um documento com as suas considerações de usabilidade, baseadas na heurística utilizada, para cada tarefa executada após a avaliação. Recomenda-se utilizar o seguinte formato apresentado na tabela A.1:

Tarefa executada	Heurística Avaliada	Problema Encontrado
tarefa 1	Diálogo natural e simples	
	Fale a língua dos usuários	
	Minimize a carga de memória do usuário	
	Consistência	
	<i>Feedback</i>	
	Saídas claramente marcadas	
	Atalhos	
	Boas mensagens de erro	
	Previna erros	
	Ajuda e documentação	
tarefa 2		
tarefa n		

Tabela A.1: formato para considerações dos avaliadores da avaliação heurística

Um exemplo do formato da tabela A.1 pode ser visualizado na tabela A.2:

Tarefa executada	Heurística Avaliada	Problema Encontrado
Inserção de um paciente no sistema	Diálogo natural e simples	Elementos textuais não necessários nesse momento são apresentados, influenciando os usuários a avaliar uma informação inútil
	Fale a língua dos usuários	1) Ítens textuais podem ser substituídos por ícones facilitando a compreensão e memorização

A.2 Protocolo de execução da avaliação

		2) Ítens textuais exigem nível de conhecimento muito elevado para sua interpretação
	Minimize a carga de memória do usuário	Formato de inserção da data é confuso e não especificado
	Consistência	Mensagens de erro são apresentadas em formato comum dentro da interface, enquanto um padrão bem aceito e reconhecido pelos usuários é utilizar a cor vermelha
	<i>Feedback</i>	Erros de preenchimento são informados somente após submeter o formulário exigindo passos adicionais para finalizar a tarefa
	Saídas claramente marcadas	Não é possível editar ou excluir um cadastro deficiente logo após a inserção
	Atalhos	O formulário de cadastro não é submetido ao utilizar a tecla "Enter" dentro do formulário, exigindo que o usuário utilize o <i>mouse</i> para pressionar o botão que executa a operação
	Boas mensagens de erro	As mensagens de erro apresentadas ao usuário não estão em linguagem clara, apresentam códigos de sistema, o que não auxilia o usuário a resolver o problema

A.2 Protocolo de execução da avaliação

	Previna erros	Por não especificar o formato padrão de data, nem fornecer um exemplo, o usuário insere no formato que achar mais adequado o que pode ser um fonte de erros
	Ajuda e documentação	Nenhum exemplo de entrada válida é fornecido, nem em documentação ou tópicos de ajuda
...

Tabela A.2: Exemplo do formato para considerações dos avaliadores da avaliação heurística

No exemplo, todas as heurísticas a serem avaliadas estão listadas. Recomenda-se listar todos os problemas encontrados em uma tarefa, para só então listar o de uma segunda tarefa. Outra recomendação importante é para vários problemas de uma mesma tarefa relacionados a uma heurística. Como mostrado no exemplo, eles devem ser listados separadamente, para um acompanhamento facilitado de sua resolução.

Os documentos gerados pelos avaliadores serão compilados pelo mantenedor da aplicação em somente um documento. A esse documento será adicionada a coluna severidade do problema de usabilidade, que será informada novamente pelos três avaliadores e retornada ao mantenedor. Com os valores para a gravidade dos problemas em mãos, o mantenedor fará uma média e finalizará a avaliação heurística com os problemas e suas gravidades. Durante a atribuição dos valores de gravidade para os problemas de usabilidade, os avaliadores terão a sua disposição acesso ao SacarWeb com o mesmo nome de usuário e senha utilizado anteriormente, de modo a facilitar a atribuição dos valores.

Recomenda-se que os avaliadores não troquem informações sobre as avaliações realizadas até que ela esteja finalizada, não influenciando assim opiniões individuais sobre os problemas e suas severidades.

APÊNDICE B – Problemas encontrados durante avaliação heurística

Este documento reúne os problemas encontrados durante todas as avaliações heurísticas realizadas por três avaliadores. Problemas semelhantes foram agrupados mas mantém as heurísticas violadas indicadas pelos avaliadores.

Problemas também observados durante os testes de usabilidade são indicados pelo sinal^[tu]; A avaliação da severidade dos problemas também se encontra nesse documento, ao final dos apontamentos para cada tarefa.

B.1 Tarefa 1: Cadastrar Paciente

1. Alguns dados do formulário digital não constam no formulário impresso, como e-mail, raça e tamanho do passo^[tu]. Heurística: Diálogo natural e simples;
2. Não é possível identificar o código do paciente, trocar a terminologia para identificador do paciente pode ser mais fácil^[tu]. Heurísticas: Fale a língua dos usuários/Diálogo natural e simples;
3. Após cadastrar um paciente, ao realizar a consulta, quando se clica no ícone de consulta os dados do paciente já deveria remeter direto para os dados do paciente e não só mostrar as opções que se pode selecionar para aquele paciente. Heurística: Minimize a carga de memória do usuário;
4. Após cadastrar um paciente, realizando uma consulta subsequente, o link cadastrar exames iniciais remete a uma página diferente de avaliação inicial. Heurística: Consistência;
5. As mensagens de erro/alerta para campos obrigatórios não preenchidos são diferentes, para o campo sexo, parece uma janela pop-up, para peso também, mas para altura não. Heurística: Consistência;

B.1 Tarefa 1: Cadastrar Paciente

6. Após gravar o cadastro de um paciente solicitou-se alteração e a mensagem de retorno foi cadastro realizado com sucesso, o correto seria alteração realizada com sucesso. Heurística: Feedback;
7. O sistema não deu suporte suficiente para operação de “redo”. Na ocorrência de alguma inconsistência na gravação dos dados, o sistema retorna a tela de cadastro do paciente em branco. O ideal seria retornar com os dados do paciente, só marcando os campos com inconsistências^[tu]. Heurísticas: Saídas claramente marcadas;
8. Ao se tentar cadastrar um paciente com código já existente, após a mensagem de erro não há nenhum botão de voltar para o tela anterior. Heurísticas: Saídas claramente marcadas;
9. Na consulta de paciente, poderia ter ao lado do ícone de consultar dados, o ícone da lixeira, para o caso de agilizar a remoção. Heurística: Atalhos;
10. Ao deixar de preencher campos obrigatórios, a mensagem que aparece após tentar gravar é muito vaga - “preencher todos os campos marcados como obrigatórios.” - e o retorno é para um cadastro em branco, perdendo-se tudo que foi informado^[tu]. Heurística: Boas mensagens de erro;
11. Após cadastrar, fazer uma alteração, seleciona-se remover, e não aparece os dados do registro que se deseja remover. O usuário corre o risco de remover o paciente errado. Heurística: Previna erros;
12. O sistema deve exibir um alerta para todos os campos obrigatórios deixados branco, e não somente para alguns. Também é interessante destacar o campo com erro de forma gráfica, como alteração de cor. Heurísticas: Previna erros/Boas mensagens de erro;
13. Não há tópico de ajuda para esta atividade. Heurística: Ajuda e documentação;
14. Na tela seleção de paciente, seria interessante a opção de clique no nome/código do paciente. Heurística: Diálogo natural e simples;
15. O Campo de telefone não possui cadastro diferenciado de telefone fixo e celular, também não possui um campo de Pessoa para Contato, como ocorre no formulário impresso. Heurística: Consistência;
16. Apesar do campo raça ser não obrigatório, não possui um valor nulo^[tu]. Heurística: Consistência;

B.1 Tarefa 1: Cadastrar Paciente

- 17.O campo de endereço tem limites de caracteres baixo. Heurística: Consistência;
- 18.Existem dois padrões de cores para cadastro de dados, dois tons de azul. Heurística: Consistência;
- 19.Quando seleciono cadastro, ainda aparece uma outra tela para eu selecionar a opção, poderia minimizar esforço do usuário, abrindo já a tela dos dados do paciente. Heurística: Atalhos;
- 20.Alguns campos deveriam auxiliar o usuário antes de seu preenchimento para evitar erro (telefone, CEP, etc) - como é feito na altura e peso^[tu]. Heurística: Previna erros;
- 21.Depois do feminino, tem dois pontos, confundindo usuário quanto ao preenchimento. Heurística: Previna erros;

A severidade dos problemas pode ser encontrada na tabela B.1.

Tarefa.Problema	AV₁	AV₂	AV₃	Média
1.01	2	3	1	2
1.02	3	3	2	2.667
1.03	3	2	3	2.667
1.04	4	3	4	3.667
1.05	4	2	3	3
1.06	3	2	3	2.667
1.07	4	3	2	3
1.08	3	3	3	3
1.09	2	2	1	1.667
1.10	4	3	4	3.667
1.11	4	3	4	3.667
1.12	2	3	3	2.667
1.13	3	2	2	2.333
1.14	3	2	2	2.333
1.15	3	3	2	2.667
1.16	2	4	2	2.667
1.17	2	3	2	2.333
1.18	1	1	2	1.333
1.19	2	2	3	2.333
1.20	2	2	2	2
1.21	2	2	0	1.333

Tabela B.1: Avaliação de severidade para a tarefa 1

B.2 Tarefa 2: Preencher a avaliação inicial

- 1.A ordem das opções entre o formulário impresso e digital está diferente, enquanto um apresenta o padrão “sim” e “não”, o outro apresenta o padrão “não” e “sim”. Heurísticas: Fale a língua dos usuários/Minimize a carga de memória dos usuários/Previna erros;
- 2.As mensagens de erro quando se deixa o campo peso em branco são diferentes entre a primeira vez que se está preenchendo os dados e a alteração dos dados. Heurística: Consistência;
- 3.Na primeira vez que se preenche os dados o campo “livre de falta de ar” é relatado como obrigatório, mas ao alterar os dados preenchidos deixando este campo em branco e os demais preenchidos, uma mensagem de sucesso de atualização é exibida. Heurística: Consistência;
- 4.O formulário só possibilita a ação “enviar”, enquanto o padrão do sistema utiliza “gravar” e “cancelar”^[tu]. Heurística: Saídas claramente marcadas;
- 5.Ao deixar um campo obrigatório sem preencher, a mensagem de erro é muito vaga. Heurística: Boas mensagens de erro;
- 6.Após consultar uma avaliação inicial e fazer alterações, deixando os campos peso atual e peso anterior em branco a mensagem “Erro ao tentar salvar o exame no banco de dados.” foi exibida. Esta mensagem não é expressiva. E não possui botão de voltar logo após a exibição desta mensagem. Heurística: Boas mensagens de erro;
- 7.Os erros não são prevenidos, pois os alertas são exibidos somente após os campos preenchidos já terem sido reiniciados^[tu]. Heurística: Previna erros;
- 8.Não há ajuda ou documentação para esta atividade. Heurística: Ajuda e documentação;
- 9.A ordenação dos campos de doença não está padronizada entre o formulário impresso e digital. Heurística: Diálogo natural e simples;
- 10.Os formulários estão organizados 2 campos por linha, dificultado o cadastro, seria ideal que fosse logo abaixo do outro campo^[tu]. Heurística: Diálogo natural e simples;
- 11.Somente em algumas questões o espaço para observações detalhadas é disponibilizado, diferentemente do formulário impresso. Heurística: Consistência.

Os valores de severidade são encontrados na tabela B.2.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
2.01	3	3	2	2.667
2.02	3	2	2	2.333
2.03	3	2	3	2.667
2.04	3	2	4	3
2.05	3	3	3	3
2.06	4	3	2	3
2.07	3	2	3	2.667
2.08	3	0	2	1.667
2.09	2	2	2	2
2.10	2	2	1	1.667
2.11	3	3	3	3

Tabela B.2: Avaliação de severidade para a tarefa 2

B.3 Tarefa 3: Criar uma nova avaliação

- 1.O link “detalhes” está perdido e necessita que o usuário conheça a utilidade do botão para utilizá-lo. Talvez seja melhor alterar para “observações”. Heurística: Diálogo natural e simples;
- 2.Os hints (*title*) dos exames deveriam ser mais expressivos (CRQ - Chronic Respiratory Questionnaire, DP6 - Teste da Distância Percorrida em 6 minutos e assim por diante). Heurística: Fale a língua dos usuários;
- 3.Ao cadastrar uma segunda avaliação, o sistema não direciona o usuário automaticamente para ela, pois utiliza a data mais antiga por padrão, campo que diferencia as avaliações. Desse modo o usuário é obrigado a selecionar a data recém criada para que as opções (novos exames) de uma nova avaliação sejam disponibilizadas^[tu]. Heurísticas: Minimizar a carga de memória do usuário/Previna erros/Diálogo natural e simples;
- 4.Exames que ainda não foram realizados (que estão listados no lado direito) não podem estar ofuscados, pois embora o exame já tenha sido acessado, foi para outra data base. Heurística: Consistência;
- 5.Se uma avaliação for criada com data errada não há possibilidade de alterar ou excluir. Heurística: Saídas claramente marcadas;
- 6.Erros de formato de entrada remetem a mensagens de ausência de conteúdo informado, por exemplo, ao digitar altura utilizando vírgula como separador decimal, o erro exibido é: “Favor informar a altura”. Heurística: Boas mensagens de erro;

B.4 Tarefa 4: Exame PFP

- 7.O tamanho da fonte que explica o conceito de avaliação é muito pequeno. Heurística: Ajuda e documentação;
- 8.A mensagem de cadastro com sucesso, está fora dos padrões do sistema, sendo avisado por um alert ao invés de uma página de retorno padrão^[tu]. Heurística: Feedback;
- 9.Como o cadastro não possui nenhum tipo de ajuda ou máscara nos campos de data, o entendimento da entrada dos dados é dificultado, o que pode causar erros^[tu]. Heurísticas: Previna erros/Documentação e ajuda;
- 10.Diferentes cores aparecem destacadas. Heurística: Consistência;
- 11.A ajuda abre uma nova página tirando o foco do usuário. Heurística: Ajuda e documentação;
- 12.Existe uma confusão e falta de padronização com os botões “voltar”, “gravar” e “finalizar”. Além do comportamento não ser o esperado, após gravar uma avaliação a tela apresenta somente o botão finalizar, que não traz nada. Poderia retornar para tela dos exames. Heurísticas: Ajuda e documentação/Saídas claramente marcadas.

A gravidade dos problemas encontrados pode ser visualizada na tabela B.3.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
3.01	2	2	1	1.667
3.02	2	2	1	1.667
3.03	3	2	4	3
3.04	2	3	2	2.333
3.05	3	4	4	3.667
3.06	3	3	3	3
3.07	2	2	2	2
3.08	2	3	1	2
3.09	3	3	3	3
3.10	2	2	2	2
3.11	3	2	2	2.333
3.12	3	3	4	3.333

Tabela B.3: Avaliação de severidade para a tarefa 3

B.4 Tarefa 4: Exame PFP

- 1.A informação “Especialidade: Questionários” está irrelevante. Heurísticas: Diálogo natural e simples;

B.4 Tarefa 4: Exame PFP

- 2.O nome do exame está como PFP, mas ao buscar por informação o conteúdo remete diretamente a espirometria, que é um dos testes da PFP^[tu]. Heurística: Fale a língua dos usuários;
- 3.Na dica (*title*) do campo VFC aparecem valores válidos de 0.1 a 7 e nas informações de 0 a 100. Heurística: Consistência;
- 4.Na dica (*title*) do campo VEF1 aparecem valores válidos de 0.1 a 6 e nas informações de 0 a 100. Heurística: Consistência;
- 5.Os valores em porcentagem que aparecem no laudo estão incoerentes, não esta explícito porcentagem de que, deve ser obtido sobre previsto. Os valores que aparecem na tela após a introdução dos dados parece ser mais coerente com o que esta ficha impressa, calculada pelo equipamento. Heurística: Consistência;
- 6.As mensagens de erro de consistência ou ausência de algum campo preenchido deveriam ser exibidas assim que detectadas e não somente quando se clica em gravar. Heurísticas: Boas mensagens de erro;
- 7.Quando não se informa a data do exame, apesar da mensagem está clara, o cursor deveria ser posicionado no campo data. Heurística: Boas mensagens de erro;
- 8.Ao selecionar o padrão do exame, no campo padrão americano, se transforma em array. Heurística: Consistência;
- 9.A tabulação está incorreta. Heurística: Atalhos;
- 10.Os campos do conjunto Ventilometria pode ter uma mensagem explicando que são campos opcionais^[tu]. Heurística: Previna erros;
- 11.Siglas sem explicações e descrição, causam dificuldades para quem não é da área da saúde ou para usuários iniciantes que não necessariamente conhecem a nomenclatura. Além disso as siglas estão diferentes entre os formulários impresso e digital, dificultando a associação entre os campos dos formulários^[tu]. Heurística: Fale a língua dos usuários;
- 12.O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” não está padronizado com o restante da aplicação^[tu]. Heurística: Saídas claramente marcadas/- Previna erros.

A severidade dos problemas para essa tarefa podem ser visualizadas na tabela

B.4.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
4.01	2	3	1	2
4.02	2	3	2	2.333
4.03	3	3	3	3
4.04	3	3	3	3
4.05	3	3	3	3
4.06	3	3	2	2.667
4.07	2	3	1	2
4.08	3	3	2	2.667
4.09	3	2	3	2.667
4.10	3	2	2	2.333
4.11	3	3	3	3
4.12	3	2	4	3

Tabela B.4: Avaliação de severidade para a tarefa 4

B.5 Tarefa 5: Exame SGRQ

- 1.A informação “Especialidade: Questionários” está irrelevante. Heurística: Diálogo natural e simples;
- 2.O texto do botão “Visualizar dados” é muito vago, melhor seria “Visualizar resultados”. Heurística: Minimize a carga de memória do usuário;
- 3.A questão 6 pode ficar em branco. Falta colocar a observação “Passe para a pergunta 7 se não teve crises graves”. Heurística: Consistência;
- 4.O item B da seção 1, parte 2 pode ficar em em branco, e o sistema não aceita isto^[tu]. Heurística: Consistência;
- 5.O item A da Seção 5, parte 2 pode ficar em branco, e o sistema não aceita isto. Heurística: Consistência;
- 6.Mensagem de erro “<?php echo (Preencha_todas_as_questoes); ?>” precisa ser melhorada, pois se trata de um código PHP, e ser mais específica, dizendo por exemplo qual questão não foi preenchida. Heurística: Boas mensagens de erro/Previna erros;
- 7.O sistema permite que o exame seja enviado sem data de realização e não exibe nenhuma mensagem de erro^[tu]. Heurística: Boas mensagens de erro/Previna erros;
- 8.O campo “C” está situado em outra pergunta. Heurística: Consistência;

B.6 Tarefa 6: Exame CRQ

9. Para gravar, é necessário visualizar os resultados antes, mas o usuário pode querer gravar sem esta etapa. Heurística: Diálogo natural e simples;
10. Há um botão com somente parte da palavra visualizada. Heurística: Previna erros;
11. Mensagens não explicativas, codificações. Heurística: Ajuda e documentação;
12. O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” deve ser padronizado, pois é confuso e difícil de aprender, causando erros mesmo durante a segunda interação pelo sistema^[tu]. Heurística: Saídas claramente marcadas/Previna Erros/Diálogo simples e natural;

A severidade dos problemas encontrados para essa tarefa está localizada na tabela B.5.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
5.01	2	3	1	2
5.02	2	2	1	1.667
5.03	3	3	2	2.667
5.04	4	4	4	4
5.05	4	4	4	4
5.06	4	3	3	3.333
5.07	4	3	3	3.333
5.08	3	4	2	3
5.09	2	3	1	3.333
5.10	4	3	4	3.667
5.11	3	2	4	3
5.12	3	2	4	3

Tabela B.5: Avaliação de severidade para a tarefa 5

B.6 Tarefa 6: Exame CRQ

1. A informação “Especialidade: Questionários” está irrelevante. Heurística: Diálogo natural e simples;
2. Na seleção das causas de falta de ar, a opção “nenhuma falta de ar” poderia ser padrão^[tu]. Heurística: Atalhos;
3. Quando se digita 7 como valor de falta de ar, o sistema já poderia selecionar automaticamente “nenhuma falta de ar”. Heurística: Atalhos;

4. Mensagens de erro apresentadas em código PHP, o que é praticamente ilegível para o usuário que não consegue entender. Heurística: Boas mensagens de erro/Ajuda e documentação/Fale a língua dos usuários;
5. Não há uma consistência nos dados apresentados na folha de resposta com as entradas no formulário, uma sugestão é que o sistema gere esse formulário impresso^[tu]. Heurística: Consistência;
6. Poderia destacar o que está em branco e precisa ser preenchido, direcionando na página para o usuário. Heurística: Minimize a carga de memória do usuário;
7. O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” deve ser padronizado, pois é confuso e difícil de aprender, causando erros mesmo durante a segunda interação pelo sistema. Heurística^[tu]: Saídas claramente marcadas/Previna Erros/Diálogo simples e natural;

Os valores de severidade para cada problema são encontrados na tabela B.6.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
6.01	2	3	1	2
6.02	3	2	0	1.667
6.03	2	1	1	1.333
6.04	4	3	3	3.333
6.05	3	2	4	3
6.06	3	2	3	2.667
6.07	3	2	4	3

Tabela B.6: Avaliação de severidade para a tarefa 6

B.7 Tarefa 7: Exame Shuttle

1. A informação “Especialidade: Questionários” está irrelevante. Heurística: Diálogo natural e simples;
2. Quando se digita as pressões d/s (antes), antes de selecionar o nível do teste atingido, esta informação é perdida. Tem que informar novamente após. Heurística: Consistência;
3. O preenchimento dos níveis de Shuttle não é intuitivo, e uma explicação poderia ajudar^[tu]. Heurística: Ajuda e documentação;

B.8 Tarefa 8: Exame DP6

4. Por não ter conhecimento nas normas do sistema, foi cadastrada uma ficha de exame incompleta, com a intenção de editá-la posteriormente, mas é apenas permitido para usuários avançados. Heurística: Ajuda e documentação;
5. Apesar de colocar um valor entre o intervalo solicitado, o sistema dizia que deveria ser inserido um valor, ou seja, existiram avaliações que não puderam ser finalizadas. Depois foi descoberto que que eram nas outras colunas que os dados deveriam ser colocados. Não foi possível salvar os dados^[tu]. Heurística: Previna erros;
6. O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” deve ser padronizado, pois é confuso e difícil de aprender, causando erros mesmo durante a segunda interação pelo sistema^[tu]. Heurísticas: Saídas claramente marcadas/Previna Erros/Diálogo simples e natural;

A severidade dos problemas pode ser encontrada na tabela B.7.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
7.01	2	3	1	2
7.02	4	3	2	3
7.03	3	2	1	2
7.04	2	2	3	2.333
7.05	4	2	3	3
7.06	3	2	4	3

Tabela B.7: Avaliação de severidade para a tarefa 7

B.8 Tarefa 8: Exame DP6

1. A informação “Especialidade: Questionários” está irrelevante. Heurística: Diálogo natural e simples;
2. O sistema não deveria aceitar Saturação de Pulso de Oxigênio maior que 90, o que ocorreu. Heurística: Consistência;
3. Deveria alertar sobre os campos faltantes e pedir confirmação de gravação. Heurística: Feedback;
4. A mensagem de erro “Preencha o primeiro nível de freq. cardíaca” não é uma mensagem clara. Heurística: Boas mensagens de erro/Fale a língua dos usuários;

B.9 Tarefa 9: Exame Pressões Respiratórias

- 5.Os formulários digital e impresso são diferentes e não seguem a mesma ordem^[tu]. Heurística: Diálogo natural e simples;
- 6.A tela de consulta de dados cadastrados possui a mesma interface da tela de cadastro, seria recomendado criar uma visualização mais amigável. Heurística: Diálogo natural e simples;
- 7.Existe uma dificuldade ao interpretar os dados do formulário para adicionar os campos^[tu]. Heurística: Previna erros;
- 8.O sistema permitiu que fosse inserido um exame com data anterior a da avaliação. Heurística: Previna erros;
- 9.O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” deve ser padronizado, pois é confuso e difícil de aprender, causando erros mesmo durante a segunda interação pelo sistema^[tu]. Heurística: Saídas claramente marcadas/Previna Erros/Diálogo simples e natural;

Os valores de severidade para cada problema se encontram na tabela B.8.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
8.01	2	3	1	2
8.02	4	3	4	3.667
8.03	3	2	4	3
8.04	3	3	3	3
8.05	3	4	2	3
8.06	3	2	2	2.333
8.07	3	3	3	3
8.08	4	2	3	3
8.09	3	2	4	3

Tabela B.8: Avaliação de severidade para a tarefa 8

B.9 Tarefa 9: Exame Pressões Respiratórias

- 1.A informação “Especialidade: Questionários” está irrelevante. Heurística: Diálogo natural e simples;
- 2.Após cancelar o exame CRQ e então selecionar pressões foi exibida uma mensagem de erro fatal de divisão por zero. Heurística: Boas mensagens de erro;

B.10 Tarefa 10: Evolução do Paciente

3. As tabulações estão passando pelos *label* dos formulários desnecessariamente. Heurística: Atalhos;
4. O comportamento e posicionamento dos botões “gravar”, “finalizar” e “cancelar” deve ser padronizado, pois é confuso e difícil de aprender, causando erros mesmo durante a segunda interação pelo sistema. Heurísticas: Saídas claramente marcadas/Previna Erros/Diálogo simples e natural;

Os valores de severidade para os problemas da tarefa 9 estão na tabela B.9.

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
9.01	2	3	1	2
9.02	4	3	4	3.667
9.03	3	2	2	2.333
9.04	3	2	4	3

Tabela B.9: Avaliação de severidade para a tarefa 9

B.10 Tarefa 10: Evolução do Paciente

1. O link para evolução poderia estar ao lado do ícone de visualizar dados^[tu]. Heurística: Minimizar a carga de memória do usuário;
2. Na apresentação do gráfico do Shuttle, na tabela analítica, o formato da data está *aaaa/mm/dd* e não *dd/mm/aaaa*^[tu]. Heurística: Consistência;
3. A tabela analítica do exame SGRQ está sem data, possivelmente por ter aceito a gravação sem data na tarefa incluir SGRQ. Heurística: Consistência;
4. Após a exibição do gráfico não há botão de retornar. Heurística: Saídas claramente marcadas;
5. Para seleção da variável do DP6 que se deseja visualizar poderia ser dispensável o botão enviar, pois é diferente da lógica inicial presente nos outros gráficos disponíveis^[tu]. Heurísticas: Saídas claramente marcadas/Consistência;
6. No gráfico de pressões a escala está incorreta, e não apresenta os resultados negativos. Heurística: Consistência;
7. Ao selecionar a visualização de todos os gráficos não existe padronização para a apresentação dos resultados do DP6, mostrando somente o primeiro item da lista, distância. Heurística: Consistência.

Os valores de severidade são encontrados na tabela B.10

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
10.01	3	2	1	2
10.02	4	3	3	3.333
10.03	4	3	2	3
10.04	2	2	2	2
10.05	2	3	2	2.333
10.06	4	4	3	3.667
10.07	3	4	2	2.667

Tabela B.10: Avaliação de severidade para a tarefa 10

B.11 Tarefa 11: Relatórios

1. Na construção do filtro, aparecem mais exames do que aqueles que compõem o conjunto de exames da instituição. Heurística: Diálogo natural e simples;
2. Mensagem de erro fatal não tratado do PHP exibida ao enviar o formulário ao selecionar um filtro e não completa os valores do intervalo desejado. Para todos os outros campos ocorre este comportamento indesejável. Heurísticas: Boas mensagens de erro/Previna erros;
3. Ao selecionar 2 filtros, selecionando as variáveis, perde-se os valores do filtro anterior, só mantém o operador. Heurística: Boas mensagens de erro;
4. Não há documentação nem ajuda para este tópico^[tu]. Heurística: Ajuda e documentação;
5. Ao marcar o exibir nos resultados, não está claro que não é obrigatório inserir valores do intervalo^[tu]. Heurística: Diálogo natural e simples.

Os valores de severidade para a tarefa 11 se encontram na tabela B.11

Tarefa.Problema	AV ₁	AV ₂	AV ₃	Média
11.01	3	2	3	2.667
11.02	4	4	4	4
11.03	4	3	3	3.333
11.04	3	2	2	2.333
11.05	3	3	1	2.333

Tabela B.11: Avaliação de severidade para a tarefa 11

APÊNDICE C – Guia para teste de usabilidade do SacarWeb

Esse documento é um guia geral para execução das tarefas para o teste de usabilidade do SacarWeb. As tarefas a serem realizadas se encontram no apêndice D.

Basicamente deverão ser inseridas duas avaliações para um determinado paciente, somente do exame de sua especialidade. As fichas devem ser selecionadas com base nas informações desse guia, sendo duas fichas (avaliações) por exame. No caso de um paciente possuir somente uma avaliação para um exame, e dois para outro, selecione um segundo paciente com duas avaliações do primeiro exame.

C.1 Importante

- Os dados dos exames devem ser inseridos estritamente no período de 26/03/2011 a 01/04/2011;
- Não é permitido qualquer tipo de ajuda de outra pessoa durante a execução das tarefas, nem execução em conjunto;
- Insira somente os exames de sua especialidade, o acesso a outros exames pode gerar dados discrepantes e interferir na avaliação;
- Deve-se utilizar o navegador Mozilla Firefox versão 3.6.x.

Os dados podem ser inseridos em qualquer computador, preferencialmente no laboratório do LFIP ou em sua casa. Mas lembrando, sem nenhum tipo de ajuda de terceiro e dentro do período estabelecido.

Os dados serão totalmente anônimos, serão repassados somente para o grupo de pesquisa parceiro do LFIP. Apenas a aplicação SacarWeb será avaliada, nunca o desempenho, correteude ou eficiência das ações executadas pelo usuário será avaliada.

Antes de iniciar a inserção dos dados no SacarWeb, responda ao questionário pré-avaliação.

Após terminar a inserção dos dados, responda ao questionário pós-avaliação.

No caso de um paciente das fichas disponibilizadas já estar cadastrado no sistema, será necessário cadastrá-lo novamente, mas com uma identificação diferente. Utilize um nome com alguma adição de modo que seja possível identificar que ele é um segundo cadastro, por exemplo, João da Silva poderia ser João da Silva 2, assim como um código diferente de outros já presentes no sistema.

As tarefas a serem executadas no sistema se encontram no apêndice D.

APÊNDICE D – Tarefas a serem realizadas durante teste de usabilidade do SacarWeb

Com o navegador Mozilla Firefox, acesse o SacarWeb utilizando o seguinte endereço: www.protem.dc.uel.br/sacarweb-alt/

Quando o sistema solicitar login e senha utilize:

- login: avaliador _ ux;
- senha: heuristica.

Dentro do sistema e com as fichas dos pacientes em mãos deverão ser executadas as seguintes tarefas, na ordem estabelecida:

- 1.Insira um paciente no sistema¹;
- 2.Preencha a avaliação inicial do paciente;
- 3.Crie uma nova avaliação para o paciente;
- 4.Adicione os dados do exame de sua especialidade;
- 5.Verifique a evolução do paciente para o exame de sua especialidade;
- 6.Obtenha um relatório com informações relevantes para pesquisa no exame de sua especialidade.

D.1 Observações importantes

Os passos 3 e 4 deverão ser executados duas vezes, com duas fichas de avaliação diferentes para o mesmo paciente, de modo que seja possível acompanhar sua evolução.

¹No caso de um paciente das fichas disponibilizadas já estar cadastrado no sistema, será necessário cadastrá-lo novamente, mas com uma identificação diferente. Utilize um nome com alguma adição de modo que seja possível identificar que ele é um segundo cadastro, por exemplo, João da Silva poderia ser João da Silva 2, assim como um código diferente de outros já presentes no sistema

APÊNDICE E – Síntese dos problemas encontrados durante os testes de usabilidade

Este documento reúne os problemas encontrados pelo mantenedor da aplicação em todas as avaliações, piloto e avaliações com usuários reais. O documento utiliza duas convenções que ajudam a compreender melhor a gravidade e validade dos problemas, como mostrado abaixo:

- Problemas em itálico são considerados graves e necessitam de correção urgente;
- Problemas marcados com o sinal gráfico^[h] ao final coincidem com resultado da avaliação heurística

E.1 Tarefa 1: cadastrar paciente

1. Há uma dificuldade em encontrar o link para inserção dos pacientes, que fica na parte inferior da página de pacientes como um botão. Pode ser mais fácil posicioná-lo na parte superior da tela como um link, de modo semelhante ao “Procurar pacientes”. 100% das avaliações piloto apresentaram o problema assim como 50% das avaliações com usuários reais. No total, o problema se apresentou em 67% das avaliações;

2. A tela de cadastro de usuários segue a convenção de campos obrigatórios em negrito e com “*” indicando isso, mas a mensagem explícita dessa condição se encontra somente no final da tela, tornando a identificação/confirmação desse comportamento difícil. 50% dos usuários da avaliação piloto tiveram problemas relacionados, assim como 75% dos usuários reais. Os problemas encontrados foram:

- falta de preenchimento de campos obrigatórios que se encontravam em diferentes fichas impressas;

E.1 Tarefa 1: cadastrar paciente

- dúvida no preenchimento de campos não obrigatórios que o usuário considerou importante, o que ocorreu somente com usuários reais.

-

O ideal nesse caso é que a mensagem seja transferida para a parte superior do formulário;

3. O padrão para o campo data de nascimento ajuda os usuários a inserir a informação da maneira correta. Apesar disso, ao ganhar foco o campo mantém o valor de exemplo, exigindo que o usuário remova manualmente, deixando a operação mais lenta. 100% dos usuários enfrentam esse problema, apesar de nenhum reclamar explicitamente;
4. Problema semelhante ao da data de nascimento ocorre com os campos peso e altura, que possuem um valor 0 como padrão, e a unidade de medida aceita logo após o campo. O ideal seria que o valor exemplo fosse removido assim que o campo ganhar foco;
5. *A informação código do paciente é confusa, pois não há uma ajuda sobre o que deve ser inserido no campo. 100% dos usuários reais tiveram problemas, e 50% dos usuários da avaliação piloto. Na avaliação com usuários reais foi constatado que a mudança do nome para "Código do HU" resolveria o problema, uma vez que esse é o identificador do paciente utilizado dentro do projeto. Uma segunda opção seria gerar automaticamente^[h];*
6. O campo CEP parece ser importante para os usuários reais, pois acreditam que seja uma informação relevante para entrar em contato com o paciente. Apesar do campo não ser obrigatório, 75% deles tentou inserir, e desses 100% falhou pois não há uma informação do padrão aceito. O campo deveria aceitar tanto com hífen, quanto sem ou informar o padrão de entrada aceito, como ocorre com o campo de data de nascimento;
7. *O campo de raça não está disponível no formulário impresso, entretanto, não existe um valor nulo no formulário digital. 100% dos usuários tiveram dificuldades ao chegar nesse campo, e 75% dos usuários reais questionaram a sua presença^[h];*
8. *Ao submeter o formulário com dados errados, a tela de cadastro é retornada sem os dados informados pelo usuário anteriormente, o que deixa os usuários muito frustrados. Esse comportamento ocorreu principalmente com os usuários que tentaram inserir o CEP, e acabaram fazendo de maneira errada. Uma solução alternativa é verificar a validade de todos os campos, inclusive os não obrigatórios, assim que o usuário submeter os dados^[h];*
9. A altura na ficha impressa utiliza a unidade centímetros, mas o sistema utiliza metros. Apenas um usuário da avaliação piloto teve problema nesse campo, mas logo percebeu

seu erro e corrigiu. Apesar de poucos problemas, é uma boa prática padronizar ambos formulários em termos de dados necessários e suas unidades de medida;

10.O formulário digital não está totalmente igual ao formulário impresso. As informações solicitadas no formulário digital são encontradas em dois formulários impressos diferentes. Alguns usuários realizam a busca dessas informações, mesmo que não obrigatórias, o que deixa o preenchimento mais lento. Uma padronização total de ambos formulários seria interessante^[h].

E.1.1 Boas práticas observadas

- A conversão transparente para os usuários do delimitador de casas decimais, que aceita tanto vírgula quanto ponto ajuda a evitar erros e deixa o sistema mais fácil de aprender;
- A mensagem de sucesso no cadastro não deixa nenhuma dúvida quanto ao sucesso da operação. O mesmo ocorre com o botões na página de sucesso;
- Os campos obrigatórios tem sua validade verificada antes do formulário ser submetido, o que evita erros;
- A tabulação entre os campos, apesar de utilizada somente por 1 usuário, ajuda na velocidade de preenchimento.

E.2 Tarefa 2: Avaliação Inicial

- 1.Existe uma dificuldade para encontrar a tela de avaliação inicial após a tela de cadastro de paciente. Como o fluxo natural de execução é cadastrar o paciente e então sua avaliação inicial seria interessante que houvesse uma maneira de acessar após o cadastro de paciente. 100% dos usuários tiveram dificuldades para encontrar essa tela. Apesar disso, todos conseguiram realizar a ligação entre a seleção de pacientes e avaliação inicial. Destacar mais a área do menu pode ser interessante para uma visualização mais fácil;
- 2.Algumas informações solicitadas no formulário eletrônico não estão disponíveis no formulário impresso, apesar de estarem em outras fichas impressas. Os usuários reais trocaram rapidamente de ficha e conseguiram preencher, entretanto, 100% dos usuários piloto tiveram que ser orientados a fazer isso. Padronizar totalmente ambos os formulários deixaria mais fácil o preenchimento da avaliação inicial. Uma dessas informações

- é o peso atual, que poderia ser resgatado automaticamente pelo sistema do registro do paciente;
- 3.A questão 6 “fuma ou já fumou” solicita dois dados ao ser respondida de modo afirmativo, “quantidade e período”. Pode ser interessante fornecer dois campos ao invés de somente um, no caso, um para cada dado. Os usuários reais também notaram que seria interessante adicionar a informação anos/maço, pois é algo importante para eles;
 - 4.Os campos obrigatórios não estão explícitos nesse formulário, o que dificulta o seu preenchimento. Explicitar os campos obrigatórios como no formulário de pacientes seria uma boa prática;
 - 5.*Pelo fato de não informar os campos obrigatórios, usuários tiveram dificuldades ao submeter o formulário com dados incompletos. Além dos dados não serem verificados ao submeter, o formulário volta sem os dados informados pelos pacientes, o que é muito frustrante para eles. 75% dos usuários reais informaram todos os dados, mesmo quando não disponíveis no formulário impresso, com receio de perderem os dados ao submeter e por não saber quais eram obrigatórios^[h];*
 - 6.50% dos usuários reais acreditaram que deveriam inserir as informações de peso com a unidade de medida. O campo não informa qual o padrão necessário de inserção, e não verifica o erro no campo antes de submeter. Para solucionar esse caso seria interessante colocar a unidade de medida logo após o campo, como no formulário de inserção de paciente;
 - 7.Um (17% total) usuário da avaliação real imaginou que a avaliação já estava preenchida apesar, de serem os valores padrão para as perguntas que estavam informados. Apesar de agilizar, caso o usuário não fique atento isso pode fazer com que o exame seja preenchido erroneamente;
 - 8.Dois usuários reais acreditam que os valores dos campos de detalhamento quando uma pergunta é respondida de modo afirmativo eram obrigatórios, apesar disso não estar explícito no formulário. Isso deve ser informado para evitar a ocorrência de erros;
 - 9.Existe somente o botão de confirmação no final do formulário, não sendo fornecida nenhuma opção de cancelamento no caso de operação errada. Seria interessante seguir o padrão do sistema e utilizar um botão cancelar^[h];
 - 10.As perguntas do questionários estão distribuídas em duas colunas, e deve ser realizada por linha. É mais comum aos usuários somente uma coluna, com uma pergunta por

linha, o que pode deixar o preenchimento mais intuitivo ^[h];

- 11.O ícone para selecionar um paciente é muito pequeno e inicialmente confuso. É melhor inserir o clique no id e nome do usuário para realizar a sua seleção na página de pacientes.

E.2.1 Boas práticas observadas

- O detalhamento das perguntas com resposta afirmativa é exibido e ganha foco automático ao responder a pergunta com “sim”. Isso agiliza o preenchimento pois não é necessário utilizar o mouse, diminui o tamanho do formulário na tela e foi elogiado por 1 usuário real.
- O formulário real e o impresso são muito semelhantes, o que agiliza muito o preenchimento dos dados.

E.3 Tarefa 3: Avaliação

- 1.*Um grave problema ocorre ao criar uma avaliação, mas é visível somente a partir da segunda avaliação criada. A avaliação recém criada não é selecionada por padrão, mas a de menor data e por isso ao criar uma segunda avaliação, e apresentar uma tela com exames já preenchidos o usuário acredita que a avaliação não foi criada com sucesso. 100% dos usuários tiveram problemas com esse comportamento. Um usuário do teste real tentou criar a avaliação novamente e recebeu um erro de data base já existente, o que o deixou ainda mais confuso^[h];*
- 2.O campo de seleção de data bases de avaliação é pouco visível, o que dificulta a seleção de diferentes avaliações mesmo para usuários que procuram por ele. Trocar a localização do campo para um lugar melhor ou então dar um destaque maior para ele. Somente um usuário (17%) identificou rapidamente que as avaliações são separadas por data base, e então começou a buscar pelo campo onde seria possível selecionar as diferentes datas fornecidas;
- 3.*100% dos usuários piloto tiveram problemas com o termo “data base”, pois imaginavam que estava ligado a algum conceito de banco de dados. 75% dos usuários reais tiveram algum problema com o termo. O fato de não estar explícito que é um campo de data confunde os usuários, o que é muito problemático para um campo que diferencia as avaliações. Um usuário real imaginava que era o código do paciente, e um usuário, em sua segunda passagem pela tela informou o padrão de data errado, uma vez que ele não*

é explicitado nessa tela. Somente o padrão de entrada no campo já é capaz de resolver o problema. Uma explicação adicional do campo seria interessante para o usuário saber para que serve. Trocar a terminologia também resolveria o problema, utilizando por exemplo “data da avaliação”. Um dos usuários reais compreendeu somente na segunda passagem pela tela que era o campo que diferenciava avaliações. Para que um usuário progredisse no sistema o conceito teve que ser explicado^[h];

4. Peso e altura da avaliação poderiam ser exportados da avaliação inicial ou do perfil do usuário, pois não são informações que se alteram com facilidade. Também pelo fato de não existir uma ficha de avaliação, agiliza a tarefa para os usuários e confunde menos;
5. A mensagem de sucesso ao se criar uma nova avaliação parece com uma mensagem de erro devido ao recurso utilizado para exibí-la, e assusta os usuários no primeiro momento. Após ler a mensagem eles percebem que se trata de sucesso. Nesse caso é interessante alterar o comportamento para o padrão do sistema^[h];
6. Não é informado no formulário de criação de avaliação quais são os campos obrigatórios, o que evitaria erros. Isso poderia ser realizado da mesma forma que no cadastro de pacientes;
7. O termo “avaliação” é muito confuso sendo que 100% dos usuários se confundiam principalmente por já terem preenchido a avaliação inicial. Eles acreditavam estarem criando a avaliação novamente. Seria interessante ter uma explicação melhor sobre o termo, explicando as diferenças entre a avaliação inicial e as avaliações, que contém os exames, o que seria ideal na página de criação de avaliação. Informar valores do paciente por padrão nessa tela já deve agilizar o processo, mas não trazer a compreensão. Os usuários reais acreditavam que os exames já deveriam estar disponíveis após a avaliação inicial, e somente perceberam que estavam realizando a operação correta após a apresentação da tela com os exames. Somente um (17%) usuário real não foi capaz de perceber que havia feito a operação corretamente. Alguns usuários, mesmo na segunda oportunidade não compreendem o conceito;
8. O botão “exportar dados” parece ser somente uma função a mais na página da avaliação, e poderia ser eliminado ou inserido em outro local. Um usuário real ao utilizar a tela de avaliação inicial acreditava que o botão “exportar dados” iria disponibilizar os exames, mesmo sem ter certeza disso;
9. Dois usuários acreditavam que o botão finalizar avaliação era obrigatório ao finalizar os exames para que ela fosse gravada, o que não é correto;

10. O botão “nova avaliação” não é muito visível, deixar ele mais destacado ou trocar sua posição. Somente um usuário (17%) conseguiu identificá-lo rapidamente;
11. Os campos informativos, ou somente de leitura, dos dados do usuário devem ter um padrão diferente dos campos onde é possível inserir dados. Esses padrões devem ser seguidos em todo o sistema. Nesse caso apesar de ajudar a identificar que está no paciente correto, pode induzir o usuário a alterar os dados onde não é possível;
12. Um usuário real teve um sério problema ao identificar o que era uma avaliação e onde realizar os exames. Ao criar a avaliação não percebeu que a tela exibida era a dos exames, e então procurou por eles na área de relatórios. Na tela de relatórios tentou preencher os dados do exame, mas ficou bastante confuso pois não compreendia a tela. Apesar da tela apresentar os dados do exame, caso fosse ela não seria de fácil preenchimento pois é muito diferente do formulário impresso. Alguns conceitos como valor inicial e final podem ser interpretados como dados do exame. O usuário não percebeu os botões ao final da tela de relatórios, mas mesmo ao visualizá-los ainda acreditava que estava na tela correta. Por não ter preenchido corretamente a tela de relatórios, pois achava que era a tela de um exame, foi exibido um erro ao final da operação, o que deixou o usuário muito confuso. Achou os relatórios eram o único lugar onde os exames poderiam estar então foi necessária a intervenção do avaliador, explicando como encontrar os exames. O usuário acredita que o conceito dos exames dentro da avaliação está correto, mas acha que está confuso de encontrar. Ao visualizar a tela do exame logo identificou que estava na tela correta.

E.3.1 Boas práticas observadas

- A distinção de colunas de exames realizados e a realizar ajuda os usuários a compreender que realizaram a operação correta e a verificar seus próximos passos dentro da avaliação;
- Devido a clara terminologia do botão “nova avaliação” logo após a avaliação inicial, a maioria dos usuários reais (66%) identificou rapidamente o botão de criação de avaliação e o utilizou. Somente um deles teve dúvida no processo.

E.4 Tarefa 4: Shuttle

1. 100% dos usuários piloto tiveram problemas com o preenchimento do nível de shuttle, acreditando que deveriam fazer como na ficha impressa, selecionando todos os níveis

percorridos ao invés do nível atingido. Mesmo quando o sistema desmarcava o nível selecionado anteriormente ao selecionar um novo, os usuários continuaram acreditando nisso, repetindo o comportamento inclusive no segundo preenchimento. Por isso, para usuários sem conhecimento uma mensagem curta na tela de preenchimento pode ajudar. A cor verde nos níveis preenchidos dá a sensação de corretude do processo. Nenhum usuário real teve problemas nesse ponto, todos preencheram diretamente somente o nível atingido pelo paciente^[h];

2. Existe uma diferença entre a Pressão Arterial no formulário impresso e digital que trouxe um pouco de insegurança aos usuários. O formulário digital solicita em todo nível de shuttle, mas não é realizada a medição durante o teste. Além disso, no formulário digital está dividida entre sistólica e diastólica, enquanto no impresso está somente em um campo^[h];
3. Algumas diferenças entre os formulários impresso e digital deixam o usuário com receio pois ele acredita que pode cometer erros. Entre elas a terminologia do último nível, presente na área “depois” no formulário impresso e na área do nível atingido no formulário digital. Isso não foi um grande problema para os usuários reais, que logo identificaram o que deveria ser feito sem auxílio. O campo “antes” também sofre o mesmo problema e poderia ser normalizado entre os formulários;
4. Os dados de recuperação estão diferentes entre os formulários, o que dificulta o preenchimento. Normalmente eles são esquecidos e uma mensagem de erro ao submeter o formulário evita que eles sejam enviados com erros e os dados perdidos. Esses campos podem ser negligenciados pois não possuem destaque como os campos dos níveis em que o paciente passou;
5. Os usuários reais alertaram para a diferença entre os formulários relacionadas ao nível de recuperação, nas duas avaliações preenchidas. Deve ser deixado somente um nível de recuperação, e não dois;
6. Apesar do campo de data de realização do exame estar presente em uma área com vários campos somente leitura, uma diferenciação entre os campos seria interessante, o que tornaria esse campo mais visível e evitaria erros, apesar da falta de data ser verificada ao enviar o formulário. Outra prática interessante nesse campo é fornecer o padrão de entrada;
7. 100% dos usuários piloto tiveram problemas com os botões “gravar”, “finalizar” e “cancelar”, pois a terminologia é confusa. O correto seria que sempre fosse repetida em

todo sistema a terminologia, posicionamento e comportamento desses botões. Outra recomendação para melhor compreensão é eliminar o botão “finalizar”, pois a operação “gravar” é inicialmente mais segura para os usuários e localizada mais a direita, posição comum de operações de “salvar” em outros sistemas^[h].

E.4.1 Boas práticas observadas

- Houve uma grande diferença comportamental entre os usuários piloto e reais. Com certeza nesse exame o conhecimento específico da área é um diferencial;
- A tabulação do formulário ajudou usuários mais experientes a realizar o preenchimento mais rapidamente;
- A mensagem de sucesso da segurança aos usuários que acreditam ter realizado a tarefa corretamente;
- Um usuário real achou muito interessante a disponibilização do cálculo automático de algumas variáveis, que no caso do formulário impresso devem ser calculadas manualmente;
- Em 100% das avaliações o segundo preenchimento mais rápido que o primeiro. A familiaridade com o exame fez muita diferença no preenchimento das telas.

E.5 Tarefa 5: SGRQ

1. 100% dos usuários acreditavam que o formulário impresso e digital eram iguais, o que é um grande auxílio ao preenchimento dos dados. Mesmo sendo iguais, uma atenção redobrada é necessária nesse aspecto, pois é comum que ao perceber isso o usuário simplesmente transcreva os dados sem ler, o que pode causar preenchimento errado devido a inconsistências entre os formulários;
2. Existe uma inconsistência entre os campos facultativos do formulário impresso e digital, pois não existe diferença entre eles e os obrigatórios. Na primeira parte o sistema exige que todas sejam preenchidas, enquanto uma delas é opcional. Algumas delas possuem um texto alertando, mas nenhuma diferença visual, o que seria interessante para uma percepção mais rápida^[h];
3. O sistema não indicou erro na falta da data de realização do exame, que deveria ser um dado obrigatório. O esquecimento pelo usuário se deve a sua localização, próximo a

dados somente de leitura, e não há nenhuma diferença entre os campos, o que ajudaria a diferenciar e identificar o campo que deve ser preenchido. Para esse campo também é interessante fornecer o padrão de entrada^[h];

4. Os botões “gravar”, “finalizar” e “cancelar” causam confusão nos usuários, pois não possuem comportamento nem posicionamento padronizado pela aplicação, erro que também ocorre neste formulário. Nesse caso, confundiram um usuário (17%) por estarem localizados também em etapas intermediárias do exame. Outro caso agravante nesse exame é a terminologia “visualizar dados para gravação” da última tela de preenchimento, que foi estranha para um usuário, que preferiu o botão “finalizar”. Ao receber a mensagem de alerta de perda de dados, então utilizou o botão em questão^[h];
5. A pergunta 8 está com as opções de resposta muito distantes, o que pode dificultar a visualização delas;
6. *Um usuário real (17%) apresentou receio ao não preencher campos facultativos, pois já perdeu dados em outros formulários por preenchimento errado e falta de avisos do sistema.*

E.5.1 Boas práticas observadas

- A padronização entre os formulários agilizou bastante o preenchimento do maior exame do SacarWeb, ponto bastante positivo.

E.6 Tarefa 6: DP6

1. A diferença entre os formulários impresso e digital dificulta muito a realização dessa tarefa de maneira satisfatória, principalmente para usuários sem conhecimento específico ou experiência, como foi observado na avaliação com usuários piloto. A principal ocorrência é relacionada aos campos “Frequência Cardíaca” e “Saturação de Pulso de Oxigênio” que se encontram invertidos entre os formulários. Além disso possuem uma faixa de valores aceitos semelhante não sendo possível diferenciá-los por valor e não são identificados no formulário impresso, exigindo uma grande atenção do usuário para não cometer erros durante o preenchimento^[h];
2. Uma vez que o sistema não define um padrão utilizado por todo ele para campos desabilitados, é possível que os usuários não tenham percebido que os campos da “Escala de Borg” estão desabilitados, provavelmente acreditando que eles estavam diferenciados

- dos demais. Somente ao tentar inserir os dados e não obter sucesso é que os usuários visualizam os *checkboxes* de ativação. Usuários reais questionaram a necessidade dessa da função, que para eles parecia desnecessária;
3. Os campos da “Escala de Borg” não são úteis nos níveis intermediários do exame, pois a medição não é realizada durante o exame nesses pontos. Esse ponto também foi questionado por usuários reais;
 4. *O campos da “Escala de Borg” aceitam somente entradas de valores inteiros, enquanto é possível inserir o valor 0.5. Esse comportamento foi questionado por usuários reais;*
 5. A ordem da tabulação (*tabindex*) dos dados deve ser alterada pois passa por campos onde não é possível inserir dados, o que acaba sendo sem utilidade para o usuário;
 6. A diferença entre os formulários para as áreas “antes”, “depois” e “recuperação” dificulta o preenchimento, deixando ele lento principalmente para os usuários inexperientes ou sem conhecimento específico, o que foi observado na avaliação piloto. Erros devido a falta de preenchimento de um dos campos, que é obrigatório, são evitados através de mensagens de erro. Apesar disso, seria uma boa prática evitar que ele ocorra. Devido ao problema, formulários impressos com dados não obrigatórios informados acabaram não sendo transcritos na interface, por falta de percepção visual semelhante entre ambos formulários;
 7. Terminologias diferentes são utilizadas para “Escala de Borg” entre os exames, o que dificulta uma rápida associação;
 8. O campo “data do exame” está próximo a campos somente leitura, e não há nenhuma diferença visual entre eles, dificultando sua identificação. Apesar do erro de enviar o exame sem data de realização ser evitado com uma mensagem de erro, é interessante evitar a ocorrência do erro;
 9. Os botões “gravar”, “finalizar” e “cancelar” causam confusão nos usuários, pois não possuem comportamento nem posicionamento padronizado pela aplicação, erro que também ocorre neste formulário^[h];
 10. O fato de dois exames estarem presentes nas fichas impressas dificulta o preenchimento, pois o usuário dificilmente se atentará a colocar os dados do melhor resultado, mas ao invés disso os que visualizar primeiro, comumente o localizado a esquerda;
 11. Na ficha de avaliação existe uma grande dificuldade para identificação dos minutos que o exame possui;

12. Ao se habilitar os valores para “Escala de Borg”, informá-los, desabilitar os campos e então submeter o formulário os dados informados são gravados, mesmo desabilitados.

E.6.1 Boas práticas observadas

- O recurso de tabulação agiliza bastante o preenchimento do formulário para usuários experientes;
- Os gráficos apresentados após o exame foram interessantes para todos os usuários, e muito elogiados por usuários reais.

E.7 Tarefa 7: CRQ

1. 100% dos usuários sofreram com a falta de padronização da disposição das cinco primeiras questões. Não existe uma ordenação lógica das opções, que não estão ordenadas conforme a ordenação do formulário impresso, nem em ordem alfabética. Outro fator complicante é que as terminologias são diferentes. Essa grande diferença irritou todos os usuários e dificultou muito o preenchimento, inclusive diminuindo drasticamente a velocidade de execução da tarefa, pois deve ser realizada uma associação entre as respostas disponíveis nos formulários. O usuários reais tiveram desempenho melhor que os usuários piloto, o que é provável devido ao seu conhecimento específico^[h];
2. Os campos para inserção dos valores de falta de ar não possuem um identificador que deixe explícito que os valores devem ser informados neles. Um usuário real não tinha certeza que os campos eram para esse fim;
3. Apesar de ser a resposta padrão para as questões sobre falta de ar, a opção “nenhuma falta de ar” não é disponibilizada como tal, o que pode agilizar e facilitar o preenchimento^[h];
4. O campo “data do exame” está próximo a campos somente leitura, e não há nenhuma diferença visual entre eles, dificultando sua identificação. Apesar do erro de enviar o exame sem data de realização ser evitado com uma mensagem de erro, é interessante evitar a ocorrência do erro;
5. Os botões “gravar”, “finalizar” e “cancelar” causam confusão nos usuários, pois não possuem comportamento nem posicionamento padronizado pela aplicação, erro que também ocorre neste formulário^[h];

- 6.O alinhamento dos campos a direita não é comum na maioria das aplicações, e chamou a atenção de um usuário;
- 7.O sistema recomenda para exames que não possuem data de realização utilizar a data base da avaliação. Um botão com essa função pode ser interessante, pois deixa a opção mais clara para os usuários ao invés de subentendida;
- 8.*Existem diferenças entre os resultados apresentados pelo sistema e atingidos pelos usuários no formulário impresso, o que foi preocupante para um usuário real.*

E.7.1 Boas práticas observadas

- A tabulação foi bastante útil nesse exame e muito bem utilizada por todos os usuários, realmente agilizando o preenchimento;
- Apesar dos formulários digital e impresso serem bastante diferentes, nesse caso não representou um problema pois os usuários fizeram uma rápida analogia entre eles e não tiveram dificuldades durante a transcrição.

E.8 Tarefa 8: PFP

- 1.Há uma grande diferença entre os formulários, o que foi causa de problemas para usuários sem experiência na área, observado na avaliação piloto. Usuários reais não tiveram problemas com as diferenças. Um exemplo de problema ocorre com as colunas PRE, equivalente a “Obtida” e POS, equivalente a POS BD, outro exemplo é o campo PFE, que se encontra fora de ordem. Um dos meios utilizados pelos usuário para transpor esse problema foram as unidades de medida, fornecidas em ambos formulários para os campos^[7];
- 2.Há grandes diferenças na terminologia utilizada entre os formulários impresso e digital, que se encontram em idiomas diferentes;
- 3.A informação “padrão do exame” parece estar sem utilidade no sistema, e cada usuário a interpreta de uma maneira. Um usuário piloto imaginou que era americano devido ao formulário impresso estar em inglês, ao se deparar com um formulário impresso em português na segunda avaliação ficou confuso. Já o usuário real não sabia qual o padrão utilizado. É importante verificar essa informação pois ela tem grande impacto nos resultados do exame;

- 4.Os campos calculados, que são somente leitura, deveriam ser apresentado de um modo diferenciado e padronizado com o restante do sistema, evitando que o usuário tente interagir com eles e possibilitando identificação ágil do comportamento esperado deles;
- 5.Os labels dos campos deveriam ter titles que ajudem a identificar seu significado;
- 6.Os campos obrigatórios não estão explícitos. O usuário real teve grande receio ao tentar gravar, pois acreditava que perderia todos os dados caso estivessem errados, pois já havia passado por essa experiência antes no sistema;
- 7.Os dados de ventilometria estão sem utilidade no formulário digital, pois nem mesmo existem no formulário impresso. Eles poderiam ser retirados ou diferenciados de modo a mostrar que não são obrigatórios^[h];
- 8.*Houve diferenças entre os resultados calculados do sistema, e obtidos pelos usuários no formulário impresso;*
- 9.Os botões “gravar”, “finalizar” e “cancelar” causam confusão nos usuários, pois não possuem comportamento nem posicionamento padronizado pela aplicação, erro que também ocorre neste formulário. Nesse caso adicionalmente a terminologia “gravar dados e emitir laudo” é utilizada ao invés de “gravar”, o que é mais um nível de falta de padronização^[h];
- 10.*A sigla PFP não é clara para todos os usuários do sistema, inclusive os reais. Um deles não conseguiu associar Espirometria, o nome mais comum do exame, com PFP. Nesse caso foi necessário inclusive intervenção do avaliador para que a avaliação pudesse continuar. Alterar a terminologia nesse caso é de grande necessidade;*
- 11.Apesar de ter grande quantidade de dados não utilizados, e dos campos mais importantes estarem presentes, o usuário real acreditam que o formulário está incompleto em termos de variáveis;
- 12.O campo “data do exame” está próximo a campos somente leitura, e não há nenhuma diferença visual entre eles, dificultando sua identificação. Apesar do erro de enviar o exame sem data de realização ser evitado com uma mensagem de erro, é interessante evitar a ocorrência do erro;
- 13.*Segundo o usuário real é possível que o protocolo implementado do exame esteja incorreto, pois não são duas medições para um exame (pré e pós BD) mas sim uma medição para cada.*

E.8.1 Boas práticas observadas

- Dados inseridos com vírgula e ponto como separador decimal são aceitos e convertidos para o padrão do sistema de maneira transparente para os usuários;
- Os cálculos automáticos foram bastante interessantes para o usuário real, que gostou pois agilizam o preenchimento e facilitam a inserção.

E.9 Tarefa 9: Pressões Respiratórias

1. O padrão utilizado no exame não parece ter utilidade, os usuários mantêm o padrão já selecionado e nem mesmo o usuário real sabe qual é utilizado;
2. O valor de P_i é multiplicado por -1 caso seja positivo, o que foi muito estranho para os usuários piloto, mas para o usuário real foi normal, pois faz parte do protocolo. Poderia ter um texto explicitando isso no local;
3. *O usuário real alega que o protocolo do exame está parcialmente implementado, e pode não atender a necessidade deles, sendo melhor se estivessem disponíveis as cinco medições de pressões para inserção. Apesar disso, soube identificar que a maior deveria ser inserida;*
4. O campo “data do exame” está próximo a campos somente leitura, e não há nenhuma diferença visual entre eles, dificultando sua identificação. Apesar do erro de enviar o exame sem data de realização ser evitado com uma mensagem de erro, é interessante evitar a ocorrência do erro;
5. Os botões “gravar”, “finalizar” e “cancelar” causam confusão nos usuários, pois não possuem comportamento nem posicionamento padronizado pela aplicação, erro que também ocorre neste formulário;

E.9.1 Boas práticas observadas

- O usuário real elogiou a função interessante pois realiza os cálculos automaticamente;
- Novamente o conhecimento sobre o exame foi importante pois o protocolo de execução não surpreendeu o usuário real, somente os pilotos.

E.10 Tarefa 10: Evolução do paciente

- 1.A área de dados tabulares possui uma limitação de tamanho que dificulta a sua visualização, obrigando o usuário a utilizar a barra de rolagem. É interessante remover e deixar que somente a barra de rolagem da janela do browser seja utilizada. A função também não é tão visível, pois pode ficar abaixo da rolagem do browser e o usuário normalmente não verá. Uma solução seria posicionar ao lado do gráfico;
- 2.A exibição de todos os gráficos não exibe os valores tabulares, o que seria interessante também e poderia ser realizado ao lado de cada gráfico;
- 3.A função não é fácil de localizar, pois 50% dos usuários tiveram problema. Mais agravante ainda foi o fato de 50% dos usuários reais encontrarem com dificuldades para encontrar a função^[h];
- 4.50% dos usuários reais acredita que se os gráficos deveriam ser separados para cada variável, pois facilitaria a interpretação e comparação, pois algumas variáveis não precisam ser comparadas. Essa afirmação é visível em alguns gráficos em que a variação dos valores é muito alta;
- 5.Para o DP6 os gráficos poderiam ser todos gerados, sem necessidade de selecionar a variável desejada, o que facilitaria e agilizaria a interação^[h];
- 6.Não há uma versão de impressão dessa página, o que pode ser útil para os usuários;
- 7.A data na tabela de dados está apresentada em formato errado. Corrigir para o padrão brasileiro, utilizado no sistema^[h].

E.10.1 Boas práticas observadas

- A função foi vista como útil por todos os usuários pois permite verificar com facilidade a evolução dos principais dados de um exame entre as diferentes avaliações, assim como as tabelas de dados analíticos;
- A exibição somente dos gráficos disponíveis evita erros como solicitação de comparações não disponíveis no sistema por falta de dados.

E.11 Tarefa 11: Relatórios

- 1. Essa tarefa pode ser classificada como a mais problemática do sistema, uma vez que 75% dos usuários reais não foram capazes de completar a tarefa. No único caso de usuário real que obteve sucesso houve contato com a tela quando estava tentando encontrar a avaliação, e por isso tinha um pouco mais de experiência. Mesmo os usuários que conseguiram tiveram que realizar várias tentativas, sendo 4 para o usuário real;*
- 2. Os usuários reais não compreendem os campos apresentados, principalmente por estarem relacionados a termos computacionais, como “ordenação”, “operador”, “valor inicial”, “valor final”;*
- 3. O campo “utilizar nos resultados” parece ser redundante, pois, uma vez selecionado o campo é provável que o usuário queira verificá-lo nos resultados. A presença desse campo obriga duas seleções para que um campo seja exibido nos resultados^[h];*
- 4. O botão “adicionar filtro” é confuso, não se sabe se irá realizar a pesquisa dos dados, se aplicará o filtro definido ou possibilitará adicionar um novo filtro. Além disso está localizado no local padrão do sistema para ação principal da tela, o que não é o caso dessa tela, que seria “pesquisar”. Alguns usuários insistiram antes de ler o botão pesquisar;*
- 5. Não existe nenhum tipo de ajuda ou documentação para essa função, o que dificulta muito a sua operação. Pelo menos um texto breve na própria página auxiliaria o aprendizado. Apesar dos usuários piloto terem identificado rapidamente a função da tela e como operá-la, eles não representam o público alvo da aplicação, pois tem conhecimento específico e profundo com computação. É possível que a maior parte dos problemas nessa tela sejam por falta de documentação ou ajuda, uma vez que os usuários não sabem o que fazer e nem o que esperar como resultado^[h];*
- 6. Todos os usuários ao verificar os resultados ficaram confusos pois muitos registros eram exibidos, uma vez que esperavam que a pesquisa fosse realizada somente com pacientes de sua instituição, e não de todo o banco de dados;*
- 7. Não é possível remover um filtro anteriormente criado, adicionar a função é de grande interesse;*
- 8. Valor inicial e final para sexo não são seleções válidas, e deixa, o usuário bastante confuso;*
- 9. A função que um filtro de exame seja selecionado para que sejam exibidos resultados, mas em nenhum local isso está explícito;*

- 10.A página de resultados deveria ter um botão voltar para a página de definição de um relatório;
- 11.*Uma vez que é muito fácil cometer erros nessa página, poderia haver uma verificação de erros prévia a submissão do formulário;*
- 12.Alguns usuários confundem o relatório com a evolução do paciente, o que evidencia que os termos são confusos;
- 13.*Usuários reais chegaram a acreditar que todos os campos deveriam ser utilizados para que a função funcionasse, o que pode ser devido a experiência ruim anterior durante a operação de gravar exames com dados errados e retornar com o formulário vazio;*
- 14.Período foi um filtro nunca utilizado, os usuários não tem nenhuma idéia do que se trata;
- 15.Por não existir um padrão no sistema, o usuário não consegue identificar campos desabilitados nessa página, mesmo que estejam diferentes dos campos comuns;
- 16.A função de impressão dos dados resultantes do filtro é interessante para os usuários, mas não disponível.

E.11.1 Boas práticas observadas

- Após a explicação da função da tela, todos os usuários achavam ela útil.

APÊNDICE F – Definição das tarefas para a abordagem automatizada proposta

Este capítulo apresenta os arquivos definidos para cada tarefa analisada com a abordagem automatizada de avaliação de usabilidade proposta. Apesar de ser possível condensá-los em apenas um arquivo, são apresentados separadamente para que os valores selecionados sejam justificados de maneira separada para cada tarefa.

Todos os arquivos possuem as URLs definidas com base na raiz do servidor, utilizando o caminho a partir dela. A definição dos parâmetros segue o padrão apresentado em 3.4.

F.1 Tarefa 1: Cadastrar Paciente

A tarefa cadastrar paciente possui somente duas telas, uma onde os dados são informados e uma segunda com a mensagem de retorno para a tarefa, mostrando sucesso ou falha no cadastro. Uma vez que não há tópico de ajuda nem de documentação para essa tarefa, somente uma URL intermediária foi considerada de modo a verificar para onde o usuário vai no sistema, e desse modo a quantidade de URLs entre o após o início da tarefa é no máximo 2, incluindo a final. O arquivo utilizado se encontra no quadro F.1.

```
/sacarweb-alt/cadastro_paciente.php?local=pac /sacarweb-alt/  
insere_paciente.php?local=pac&sub=cons 2  
insere_dados_paciente
```

Quadro F.1: Arquivo de entrada para tarefa cadastrar paciente

F.2 Tarefa 2: Preencher avaliação inicial

A tarefa preencher avaliação inicial idealmente é realizada em duas etapas, pois também se trata de um cadastro simples. A tarefa não possui nenhum tópico de ajuda ou documentação, e por isso duas URLs foram definidas como a quantidade máxima após a URL inicial, de modo a verificar para onde o usuário se dirige após não finalizar a tarefa. O arquivo utilizado pode ser visualizado no quadro F.2.

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini /sacarweb-  
alt/grava_avaliacao_inicial.php 2 avaliacao_inicial
```

Quadro F.2: Arquivo de entrada para a tarefa avaliação inicial

F.3 Tarefa 3: Exame PFP

A tarefa PFP é uma das mais utilizadas no sistema, pois é de grande importância para os usuários. Ela foi definida com um máximo de 3 URLs após o início da tarefa, uma vez que é possível alterar o padrão do exame desejado, o que provoca acaba carregando a página novamente, e ela possui uma página de ajuda e documentação. O arquivo está foi definido como mostra o quadro F.3.

```
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava /sacarweb-alt/  
grava_emite_laudo_pfp.php?local=pac&sub=ava 3 PFP
```

Quadro F.3: Arquivo de entrada para a tarefa PFP

F.4 Tarefa 4: Exame SGRQ

A tarefa SGRQ é uma das mais extensas do sistema, e possui quatro passos intermediários obrigatórios e ajuda/documentação disponível em todos eles. Desse modo, inserindo os passos intermediários, a página de status e uma possível consulta a ajuda chegamos ao valor de 6 URLs, como visto no quadro F.4.

```
/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava /sacarweb-alt/  
grava_dados_sgrq.php?local=pac&sub=ava 6 SGRQ
```

Quadro F.4: Arquivo de entrada para a tarefa SGRQ

F.5 Tarefa 5: Exame CRQ

A tarefa CRQ possui uma tela intermediária que mostra os resultados do exame com base nos dados inseridos para o usuário revisar logo após submeter os dados. Caso o usuário confirme a operação, então o exame é gravado. Adicionalmente ajuda/documentação são disponibilizados em todas as etapas. Dessa maneira, a quantidade de URLs foi definida para quatro, como mostra o quadro F.5.

```
/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false /
sacarweb-alt/grava_crq.php?local=pac&sub=ava 4 CRQ
```

Quadro F.5: Arquivo de entrada para a tarefa CRQ

F.6 Tarefa 6: Exame Shuttle

Na tarefa shuttle existe uma consulta intermediária que provoca um recarregamento da página após a seleção do nível de Shuttle, que atualiza a interface com os valores para o nível selecionado. Além disso, é disponibilizada a consulta a ajuda/documentação, o que levou-nos a definir a quantidade de URLs intermediárias para três, como pode ser visto no quadro F.6.

```
/sacarweb-alt/tela_shuttle.php?local=pac&sub=ava /sacarweb-alt/
tela_shuttle.php?local=pac&sub=ava 3 Shuttle
```

Quadro F.6: Arquivo de entrada para a tarefa Shuttle

F.7 Tarefa 7: Exame DP6

O exame DP6 possui todo o formulário disponibilizado na primeira tela, e não necessita de chamadas intermediárias para consulta de dados ou atualização da interface. Por possuir documentação/ajuda é possível que mais uma requisição aconteça durante o exame, sendo assim a quantidade de URLs foi definida para três, como é visto no quadro F.7.

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava /sacarweb-alt/
grava_dp6.php?local=pac&sub=ava 3 DP6
```

Quadro F.7: Arquivo de entrada para a tarefa DP6

F.8 Tarefa 8: Exame Pressões Respiratórias

O exame Pressões Respiratórias é um dos mais compactos do sistema, entretanto nele é possível selecionar o padrão de interpretação desejado além de estar disponibilizada ajuda/documentação, o que fez com que a quantidade de URLs fosse definida para três nessa tarefa. O arquivo utilizado pode ser visualizado no quadro F.8.

```
/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava /sacarweb-alt
/grava_pressoes.php?local=pac&sub=ava 3 Pressoes
```

Quadro F.8: Arquivo de entrada para a tarefa Pressões Respiratórias

F.9 Tarefa 9: Evolução do Paciente

A tarefa evolução do paciente inicialmente mostra uma lista com os gráficos disponíveis com base nos dados do paciente selecionado. O usuário deve selecionar o gráfico desejado para que o sistema mostre ele na tela. Em ambas operações a página é carregada novamente, e por não ter nenhum tipo de ajuda ou documentação a quantidade de URLs foi definida para três, como pode ser visto no quadro F.9.

```
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=
sacarweb-alt/exibe_grafico.php?local=pac&sub=graf 3 Evolucao
```

Quadro F.9: Arquivo de entrada para a tarefa Evolução do Paciente

F.10 Tarefa 10: Relatórios

A tarefa Relatórios pode recarregar a página várias vezes, uma vez que a cada filtro de exame adicionado a página é recarregada. Entretanto, no caso da avaliação o usuário possuía somente uma especialidade, o que pode indicar que a página será recarregada somente uma vez. Ainda prevendo dificuldades na execução da tarefa, a quantidade de URLs máxima foi definida para quatro, como mostra o quadro F.10.

```
/sacarweb-alt/tela_relatorio2.php?local=rel /sacarweb-alt/
processa_relatorio.php?local=rel 4 Relatorio
```

Quadro F.10: Arquivo de entrada para a tarefa Relatórios

APÊNDICE G – Resultados da avaliação utilizando a abordagem proposta

Este capítulo apresenta os resultados do protótipo desenvolvido utilizando a abordagem proposta por esse trabalho para cada tarefa. Os dados consistem dos arquivos de saída do protótipo, detalhando o conteúdo de cada *cluster* e ao final do arquivo, fornecendo os valores de qualidade para os *clusters*.

G.1 Tarefa 1: Cadastrar Paciente

---- Cluster 1 ----

Representa 5.26315789474 por cento

----- Acesso 1 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

289 seg

----- Acesso 2 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

331 seg

----- Acesso 3 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

286 seg

---- Cluster 2 ----

Representa 31.5789473684 por cento

G.1 Tarefa 1: Cadastrar Paciente

----- Acesso 1 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/tela_relatorio2.php?local=rel

51 seg

----- Acesso 2 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=crq

81 seg

----- Acesso 3 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

12 seg

----- Acesso 4 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=036
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

6 seg

----- Acesso 5 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=0&nome=

70 seg

----- Acesso 6 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac

9 seg

----- Acesso 7 -----

/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

10 seg

----- Acesso 8 -----

G.1 Tarefa 1: Cadastrar Paciente

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
15 seg
----- Acesso 9 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
26 seg
----- Acesso 10 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
22 seg
----- Acesso 11 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
12 seg
----- Acesso 12 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
22 seg
----- Acesso 13 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
36 seg
----- Acesso 14 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
61 seg
----- Acesso 15 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
```

G.1 Tarefa 1: Cadastrar Paciente

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
11 seg
----- Acesso 16 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
4 seg
----- Acesso 17 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/perfil_usuario.php?local=perf_usu
6 seg
----- Acesso 18 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
6 seg

---- Cluster 3 ----
Representa 10.5263157895 por cento
----- Acesso 1 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/inicio.php?local=ini
187 seg
----- Acesso 2 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
153 seg
----- Acesso 3 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
169 seg
----- Acesso 4 -----
```

G.1 Tarefa 1: Cadastrar Paciente

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
166 seg
----- Acesso 5 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
158 seg
----- Acesso 6 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
150 seg

---- Cluster 4 ----
Representa 1.75438596491 por cento
----- Acesso 1 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=16
319 seg

---- Cluster 5 ----
Representa 7.01754385965 por cento
----- Acesso 1 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac
/sacarweb-alt/logout.php?local=ini
83 seg
----- Acesso 2 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
/sacarweb-alt/tela_grafico.php?id=527&exame=ConsultaDp6
159 seg
----- Acesso 3 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788
```

G.1 Tarefa 1: Cadastrar Paciente

/sacarweb-alt/consultar_dp6.php?local=pac&sub=ava

3 seg

----- Acesso 4 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf

/sacarweb-alt/tela_grafico.php?local=pac&sub=graf&login=Gilberto253788_&exame=Dp6&v

1 seg

---- Cluster 6 ----

Representa 31.5789473684 por cento

----- Acesso 1 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

0 seg

----- Acesso 2 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

59 seg

----- Acesso 3 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

7 seg

----- Acesso 4 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

0 seg

----- Acesso 5 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

0 seg

----- Acesso 6 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

42 seg

----- Acesso 7 -----

/sacarweb-alt/cadastro_paciente.php?local=pac

0 seg

----- Acesso 8 -----

G.1 Tarefa 1: Cadastrar Paciente

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
83 seg
----- Acesso 9 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
0 seg
----- Acesso 10 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
67 seg
----- Acesso 11 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
0 seg
----- Acesso 12 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
0 seg
----- Acesso 13 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
0 seg
----- Acesso 14 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
0 seg
----- Acesso 15 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
91 seg
----- Acesso 16 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
95 seg
----- Acesso 17 -----
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
88 seg
----- Acesso 18 -----
```

G.1 Tarefa 1: Cadastrar Paciente

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
65 seg
```

---- Cluster 7 ----

Representa 7.01754385965 por cento

----- Acesso 1 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
199 seg
```

----- Acesso 2 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
213 seg
```

----- Acesso 3 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
208 seg
```

----- Acesso 4 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
202 seg
```

---- Cluster 8 ----

Representa 5.26315789474 por cento

----- Acesso 1 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
510 seg
```

----- Acesso 2 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
359 seg
```

----- Acesso 3 -----

```
/sacarweb-alt/cadastro_paciente.php?local=pac
```

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/insere_paciente.php?local=pac&sub=cons

372 seg

--- Quality Measurements ---

Within groups variance: 1.39685392365

Between groups variance: 1.81980463414

Total variance: 5.79451245641

F-value: 1.30278807492

Wilk's Lambda value: 0.241064961748

G.2 Tarefa 2: Preencher avaliação inicial

---- Cluster 1 ----

Representa 40.6504065041 por cento

----- Acesso 1 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16

24 seg

----- Acesso 2 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

39 seg

----- Acesso 3 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=16&nova_avaliacao=tr

9 seg

----- Acesso 4 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16

/sacarweb-alt/cadastro_paciente.php?id=16&alterar=true&local=pac&sub=dados

26 seg

----- Acesso 5 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/info_help/i_avaliacao.html

40 seg

----- Acesso 6 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=50

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=50

20 seg

----- Acesso 7 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=33

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=33

15 seg

----- Acesso 8 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 9 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=1

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

11 seg

----- Acesso 10 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

36 seg

----- Acesso 11 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava

25 seg

----- Acesso 12 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=1

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=1

8 seg

----- Acesso 13 -----

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 14 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
2 seg
----- Acesso 15 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 16 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 17 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 18 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
1 seg
----- Acesso 19 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 20 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 21 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
1 seg
----- Acesso 22 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 23 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
```

G.2 Tarefa 2: Preencher avaliação inicial

```
2 seg
----- Acesso 24 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/perfil_usuario.php?local=perf_usu
2 seg
----- Acesso 25 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
5 seg
----- Acesso 26 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 27 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 28 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 29 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
25 seg
----- Acesso 30 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 31 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 32 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
5 seg
----- Acesso 33 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
```

G.2 Tarefa 2: Preencher avaliação inicial

----- Acesso 34 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 35 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 36 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 37 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 38 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

22 seg

----- Acesso 39 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 40 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 41 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 42 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 43 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 44 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

0 seg

----- Acesso 45 -----

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 46 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 47 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16
1 seg
----- Acesso 48 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16
11 seg
----- Acesso 49 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
0 seg
----- Acesso 50 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
44 seg

---- Cluster 2 ----
Representa 6.50406504065 por cento
----- Acesso 1 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
291 seg
----- Acesso 2 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
230 seg
----- Acesso 3 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
```

G.2 Tarefa 2: Preencher avaliação inicial

328 seg

----- Acesso 4 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

349 seg

----- Acesso 5 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

255 seg

----- Acesso 6 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

239 seg

----- Acesso 7 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

225 seg

----- Acesso 8 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

307 seg

---- Cluster 3 ----

Representa 18.6991869919 por cento

----- Acesso 1 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/perfil_usuario.php?local=perf_usu

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

3 seg

----- Acesso 2 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=808&nova_avaliacao=t

15 seg

----- Acesso 3 -----

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
10 seg
----- Acesso 4 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=808
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=808
10 seg
----- Acesso 5 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=808
/sacarweb-alt/consulta_pacientes.php?local=conspac&origem=avaliacao
2 seg
----- Acesso 6 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=
9 seg
----- Acesso 7 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=808
1 seg
----- Acesso 8 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=808
1 seg
----- Acesso 9 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=36
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
10 seg
----- Acesso 10 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=36
```

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

8 seg

----- Acesso 11 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=5

11 seg

----- Acesso 12 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=431

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

23 seg

----- Acesso 13 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=78

3 seg

----- Acesso 14 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=78

4 seg

----- Acesso 15 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=431

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=431

27 seg

----- Acesso 16 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=431

8 seg

----- Acesso 17 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava

8 seg

----- Acesso 18 -----

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3
/sacarweb-alt/perfil_usuario.php?local=perf_usu
8 seg
----- Acesso 19 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=253788
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
1 seg
----- Acesso 20 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao.php?local=pac&sub=ava
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava
3 seg
----- Acesso 21 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=760349
27 seg
----- Acesso 22 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=760349
26 seg
----- Acesso 23 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=0
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava
8 seg

---- Cluster 4 ----
Representa 5.69105691057 por cento
----- Acesso 1 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
148 seg
----- Acesso 2 -----
```

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=16
54 seg
----- Acesso 3 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
56 seg
----- Acesso 4 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
58 seg
----- Acesso 5 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
77 seg
----- Acesso 6 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
67 seg
----- Acesso 7 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/grava_avaliacao_inicial.php
156 seg

---- Cluster 5 ----
Representa 26.0162601626 por cento
----- Acesso 1 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
19 seg
----- Acesso 2 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=22222
```

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=22222&nova_avaliacao=
13 seg
----- Acesso 3 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3333
/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=
8 seg
----- Acesso 4 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3333
/sacarweb-alt/tela_consulta_pfp.php?local=pac&sub=ava
7 seg
----- Acesso 5 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3333
/sacarweb-alt/tela_consulta_pfp.php?local=pac&sub=ava
25 seg
----- Acesso 6 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
36 seg
----- Acesso 7 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/error_report.php
/sacarweb-alt/tela_relatorio2.php?local=rel
4 seg
----- Acesso 8 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3
/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=
7 seg
----- Acesso 9 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=127670
```

G.2 Tarefa 2: Preencher avaliação inicial

```
/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=127670&nova_avaliaca
170 seg
----- Acesso 10 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=127670
/sacarweb-alt/removido_sucesso.php?code=127670&local=pac&sub=cons
10 seg
----- Acesso 11 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229
/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=696229&nova_avaliaca
5 seg
----- Acesso 12 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=696229
5 seg
----- Acesso 13 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/error_report.php
/sacarweb-alt/tela_relatorio2.php?local=rel
3 seg
----- Acesso 14 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=105
/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=105&nova_avaliacao=t
21 seg
----- Acesso 15 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=105
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=105
41 seg
----- Acesso 16 -----
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=105
```

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=105

8 seg

----- Acesso 17 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=431

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=431&nova_avaliacao=t

8 seg

----- Acesso 18 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=431

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=431

11 seg

----- Acesso 19 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=253788&nova_avaliaca

26 seg

----- Acesso 20 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

4 seg

----- Acesso 21 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788

3 seg

----- Acesso 22 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/alterar_dp6.php?local=pac&sub=ava

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788

17 seg

----- Acesso 23 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=253788

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229

3 seg

----- Acesso 24 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229

/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=696229

10 seg

----- Acesso 25 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

7 seg

----- Acesso 26 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

135 seg

----- Acesso 27 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=696229&nova_avaliaca

19 seg

----- Acesso 28 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=696229

/sacarweb-alt/cadastro_paciente.php?id=696229&alterar=true&local=pac&sub=dados

15 seg

----- Acesso 29 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

6 seg

----- Acesso 30 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=1111

G.2 Tarefa 2: Preencher avaliação inicial

/sacarweb-alt/tela_nova_avaliacao.php?local=pac&sub=ava&codigo=1111&nova_avaliacao=

5 seg

----- Acesso 31 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=126

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

7 seg

----- Acesso 32 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=100

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

15 seg

---- Cluster 6 ----

Representa 2.43902439024 por cento

----- Acesso 1 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

460 seg

----- Acesso 2 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

576 seg

----- Acesso 3 -----

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/grava_avaliacao_inicial.php

460 seg

--- Quality Measurements ---

Within groups variance: 2.08268100465

Between groups variance: 1.18854038295

Total variance: 16.307281747

F-value: 0.570678073258

Wilk's Lambda value: 0.127714786374

G.3 Tarefa 3: Exame PFP

```
---- Cluster 1 ----
Representa 7.40740740741 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava
199 seg
----- Acesso 2 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava
182 seg

---- Cluster 2 ----
Representa 3.7037037037 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava
141 seg

---- Cluster 3 ----
Representa 59.2592592593 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 2 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 3 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 4 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 5 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
```

G.3 Tarefa 3: Exame PFP

```
0 seg
----- Acesso 6 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 7 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 8 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 9 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 10 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 11 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 12 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 13 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 14 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 15 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
----- Acesso 16 -----
/sacarweb-alt/tela_pfp.php?local=pac&sub=ava
0 seg
```

G.3 Tarefa 3: Exame PFP

---- Cluster 4 ----

Representa 7.40740740741 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=16

30 seg

----- Acesso 2 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/error_report.php

/sacarweb-alt/tela_relatorio2.php?local=rel

265 seg

---- Cluster 5 ----

Representa 7.40740740741 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5

/sacarweb-alt/tela_shuttle.php?local=pac&sub=ava

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5

17 seg

----- Acesso 2 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

23 seg

---- Cluster 6 ----

Representa 11.11111111111 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava

75 seg

G.4 Tarefa 4: Exame SGRQ

----- Acesso 2 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava

19 seg

----- Acesso 3 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava

77 seg

---- Cluster 7 ----

Representa 3.7037037037 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pfp.php?local=pac&sub=ava

/sacarweb-alt/grava_emite_laudo_pfp.php?local=pac&sub=ava

616 seg

--- Quality Measurements ---

Within groups variance: 1.28472467762

Between groups variance: 3.72397287492

Total variance: 15.5066544854

F-value: 2.89865442752

Wilk's Lambda value: 0.08284989382

G.4 Tarefa 4: Exame SGRQ

---- Cluster 1 ----

Representa 18.1818181818 por cento

----- Acesso 1 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=33

24 seg

----- Acesso 2 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

0 seg

G.4 Tarefa 4: Exame SGRQ

---- Cluster 2 ----

Representa 27.2727272727 por cento

----- Acesso 1 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5

/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

/sacarweb-alt/grava_avaliacao.php?local=pac&sub=ava

88 seg

----- Acesso 2 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=696229

/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=696229

/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=696229

20 seg

----- Acesso 3 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

201 seg

---- Cluster 3 ----

Representa 45.4545454545 por cento

----- Acesso 1 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

G.4 Tarefa 4: Exame SGRQ

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

223 seg

----- Acesso 2 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

174 seg

----- Acesso 3 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

197 seg

----- Acesso 4 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

174 seg

----- Acesso 5 -----

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava

/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava

/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava

195 seg

---- Cluster 4 ----

Representa 9.09090909091 por cento

----- Acesso 1 -----

G.5 Tarefa 5: Exame CRQ

```

/sacarweb-alt/tela1_sgrq.php?local=pac&sub=ava
/sacarweb-alt/tela2_sgrq.php?local=pac&sub=ava
/sacarweb-alt/tela3_sgrq.php?local=pac&sub=ava
/sacarweb-alt/tela4_sgrq_resultado.php?local=pac&sub=ava
/sacarweb-alt/grava_dados_sgrq.php?local=pac&sub=ava
133 seg

```

```

--- Quality Measurements ---
Within groups variance: 0.937546811119
Between groups variance: 5.0585546855
Total variance: 0.246581015122
F-value: 5.39552225607
Wilk's Lambda value: 3.80218570621

```

G.5 Tarefa 5: Exame CRQ

```

---- Cluster 1 ----
Representa 33.3333333333 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false
/sacarweb-alt/consulta_crq.php?local=pac&sub=ava
/sacarweb-alt/grava_crq.php?local=pac&sub=ava
376 seg
----- Acesso 2 -----
/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false
/sacarweb-alt/consulta_crq.php?local=pac&sub=ava
/sacarweb-alt/grava_crq.php?local=pac&sub=ava
353 seg

---- Cluster 2 ----
Representa 16.6666666667 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false
/sacarweb-alt/consulta_crq.php?local=pac&sub=ava
/sacarweb-alt/grava_crq.php?local=pac&sub=ava

```

G.6 Tarefa 6: Exame Shuttle

236 seg

---- Cluster 3 ----

Representa 33.3333333333 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false

0 seg

----- Acesso 2 -----

/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5

/sacarweb-alt/tela_nova_avaliacao.php?nova_avaliacao=true&local=pac&sub=ava&codigo=

/sacarweb-alt/grava_avaliacao.php?local=pac&sub=ava

70 seg

---- Cluster 4 ----

Representa 16.6666666667 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_crq.php?local=pac&sub=ava&consultar=false

/sacarweb-alt/consulta_crq.php?local=pac&sub=ava

/sacarweb-alt/grava_crq.php?local=pac&sub=ava

141 seg

--- Quality Measurements ---

Within groups variance: 10405.1

Between groups variance: 3791.53333333

Total variance: 10261.69

F-value: 0.364391820678

Wilk's Lambda value: 1.01397528087

G.6 Tarefa 6: Exame Shuttle

Não há resultados a serem apresentados para essa tarefa uma vez que seu fluxo de dados impede que seja avaliada devido a limitação da abordagem proposta em tarefas com as URLs inicial e final iguais.

G.7 Tarefa 7: Exame DP6

```
---- Cluster 1 ----
Representa 26.6666666667 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=808
12 seg
----- Acesso 2 -----
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
/sacarweb-alt/inicio.php?local=ini
/sacarweb-alt/tela_login.php?local=login
5 seg
----- Acesso 3 -----
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/login.php
/sacarweb-alt/inicio.php?local=ini
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
40 seg
----- Acesso 4 -----
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=1
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=1
17 seg

---- Cluster 2 ----
Representa 33.3333333333 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
246 seg
----- Acesso 2 -----
```

G.7 Tarefa 7: Exame DP6

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808  
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=808
```

366 seg

----- Acesso 3 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=808  
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini  
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
```

318 seg

----- Acesso 4 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
```

238 seg

----- Acesso 5 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
```

230 seg

---- Cluster 3 ----

Representa 13.3333333333 por cento

----- Acesso 1 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/inicio.php?local=ini
```

2 seg

----- Acesso 2 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava  
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
```

168 seg

---- Cluster 4 ----

Representa 13.3333333333 por cento

----- Acesso 1 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
```

0 seg

G.8 Tarefa 8: Exame Pressões Respiratórias

----- Acesso 2 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=5
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
13 seg
```

---- Cluster 5 ----

Representa 13.3333333333 por cento

----- Acesso 1 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=16
172 seg
```

----- Acesso 2 -----

```
/sacarweb-alt/tela_dp6.php?local=pac&sub=ava
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/grava_dp6.php?local=pac&sub=ava
212 seg
```

--- Quality Measurements ---

Within groups variance: 0.925790753933

Between groups variance: 1.56076594207

Total variance: 0.192130780197

F-value: 1.68587333092

Wilk's Lambda value: 4.81854470682

G.8 Tarefa 8: Exame Pressões Respiratórias

---- Cluster 1 ----

Representa 14.2857142857 por cento

----- Acesso 1 -----

```
/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava
/sacarweb-alt/grava_pressoes.php?local=pac&sub=ava
```

G.8 Tarefa 8: Exame Pressões Respiratórias

153 seg

---- Cluster 2 ----

Representa 14.2857142857 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

/sacarweb-alt/grava_pressoes.php?local=pac&sub=ava

37 seg

---- Cluster 3 ----

Representa 42.8571428571 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

0 seg

----- Acesso 2 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

0 seg

----- Acesso 3 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

0 seg

---- Cluster 4 ----

Representa 14.2857142857 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

/sacarweb-alt/grava_pressoes.php?local=pac&sub=ava

83 seg

---- Cluster 5 ----

Representa 14.2857142857 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_pressoes.php?local=pac&sub=ava

/sacarweb-alt/grava_pressoes.php?local=pac&sub=ava

39 seg

```

--- Quality Measurements ---
Within groups variance: 1661.05
Between groups variance: 881.3
Total variance: 554.6025
F-value: 0.530568014208
Wilk's Lambda value: 2.99502797048

```

G.9 Tarefa 9: Evolução do Paciente

Não há resultados a serem apresentados para essa tarefa uma vez que seu fluxo de dados impede que seja avaliada devido a limitação da abordagem proposta em tarefas com as URLs inicial e final iguais.

G.10 Tarefa 10: Relatórios

```

---- Cluster 1 ----
Representa 5.08474576271 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=pfp
/sacarweb-alt/tela_relatorio2.php
/sacarweb-alt/processa_relatorio.php?local=rel
119 seg
----- Acesso 2 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=pfp
/sacarweb-alt/processa_relatorio.php?local=rel
85 seg
----- Acesso 3 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/processa_relatorio.php?local=rel
71 seg

---- Cluster 2 ----
Representa 5.08474576271 por cento

```

G.10 Tarefa 10: Relatórios

----- Acesso 1 -----

/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=3
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=3
6 seg

----- Acesso 2 -----

/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
6 seg

----- Acesso 3 -----

/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=dp6
/sacarweb-alt/retorna_campos_select.php?exame=shuttle
/sacarweb-alt/processa_relatorio.php?local=rel
24 seg

---- Cluster 3 ----

Representa 6.77966101695 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=sgrq
/sacarweb-alt/error_report.php
/sacarweb-alt/envia_report.php
/sacarweb-alt/help.php
79 seg

----- Acesso 2 -----

/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=dp6
/sacarweb-alt/processa_relatorio.php?local=rel
61 seg

----- Acesso 3 -----

G.10 Tarefa 10: Relatórios

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/retorna_campos_select.php?exame=crq

/sacarweb-alt/processa_relatorio.php?local=rel

54 seg

----- Acesso 4 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/retorna_campos_select.php?exame=sgrq

/sacarweb-alt/processa_relatorio.php?local=rel

20 seg

---- Cluster 4 ----

Representa 3.38983050847 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

/sacarweb-alt/cadastro_paciente.php?local=pac

/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

65 seg

----- Acesso 2 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=253788

/sacarweb-alt/perfil_usuario.php?local=perf_usu

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

7 seg

---- Cluster 5 ----

Representa 37.2881355932 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/cadastro_paciente.php?local=pac

G.10 Tarefa 10: Relatórios

25 seg

----- Acesso 2 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

0 seg

----- Acesso 3 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/retorna_campos_select.php?exame=pfp

/sacarweb-alt/tela_relatorio2.php

/sacarweb-alt/processa_relatorio.php?local=rel

13 seg

----- Acesso 4 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/logout.php?local=ini

/sacarweb-alt/inicio.php?local=ini

3 seg

----- Acesso 5 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

2 seg

----- Acesso 6 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

0 seg

----- Acesso 7 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

5 seg

----- Acesso 8 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

0 seg

----- Acesso 9 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

1 seg

----- Acesso 10 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

G.10 Tarefa 10: Relatórios

/sacarweb-alt/error_report.php

8 seg

----- Acesso 11 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

6 seg

----- Acesso 12 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/perfil_usuario.php?local=perf_usu

3 seg

----- Acesso 13 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

/sacarweb-alt/help.php

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

17 seg

----- Acesso 14 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/perfil_usuario.php?local=perf_usu

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

6 seg

----- Acesso 15 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

4 seg

----- Acesso 16 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/inicio.php?local=ini

/sacarweb-alt/perfil_usuario.php?local=perf_usu

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

7 seg

----- Acesso 17 -----

G.10 Tarefa 10: Relatórios

```
/sacarweb-alt/tela_relatorio2.php?local=rel
0 seg
----- Acesso 18 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/processa_relatorio.php?local=rel
14 seg
----- Acesso 19 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/processa_relatorio.php?local=rel
7 seg
----- Acesso 20 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
0 seg
----- Acesso 21 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
6 seg
----- Acesso 22 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
0 seg

---- Cluster 6 ----
Representa 1.69491525424 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/retorna_campos_select.php?exame=pfp
/sacarweb-alt/tela_relatorio2.php
197 seg

---- Cluster 7 ----
Representa 5.08474576271 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/cadastro_paciente.php?local=pac
```

G.10 Tarefa 10: Relatórios

```
/sacarweb-alt/insere_paciente.php?local=pac&sub=cons
/sacarweb-alt/dados_paciente.php?id=16&local=pac&sub=cons
397 seg
----- Acesso 2 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/cadastro_paciente.php?local=pac
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
340 seg
----- Acesso 3 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/envia_report.php
/sacarweb-alt/error_report.php
/sacarweb-alt/envia_report.php
354 seg

---- Cluster 8 ----
Representa 15.2542372881 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/help.php
/sacarweb-alt/dados_exame.php?exame=dp6
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
54 seg
----- Acesso 2 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=36
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=36
16 seg
----- Acesso 3 -----
```

G.10 Tarefa 10: Relatórios

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=78
19 seg
----- Acesso 4 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/cadastro_paciente.php?local=pac
21 seg
----- Acesso 5 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/help.php
/sacarweb-alt/dados_exame.php?exame=pfp
20 seg
----- Acesso 6 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/perfil_usuario.php?local=perf_usu
16 seg
----- Acesso 7 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/inicio.php?local=ini
/sacarweb-alt/logout.php?local=ini
/sacarweb-alt/inicio.php?local=ini
2 seg
----- Acesso 8 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
```

G.10 Tarefa 10: Relatórios

/sacarweb-alt/perfil_usuario.php?local=perf_usu

12 seg

----- Acesso 9 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=696229

5 seg

---- Cluster 9 ----

Representa 8.47457627119 por cento

----- Acesso 1 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/retorna_campos_select.php?exame=pfp

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=50

/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=50

62 seg

----- Acesso 2 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

/sacarweb-alt/perfil_usuario.php?local=perf_usu

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=3333

13 seg

----- Acesso 3 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

/sacarweb-alt/error_report.php

/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons

/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=22222

20 seg

----- Acesso 4 -----

/sacarweb-alt/tela_relatorio2.php?local=rel

/sacarweb-alt/error_report.php

G.10 Tarefa 10: Relatórios

```
/sacarweb-alt/perfil_usuario.php?local=perf_usu
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=253788
16 seg
----- Acesso 5 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/error_report.php
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=696229
10 seg

---- Cluster 10 ----
Representa 11.8644067797 por cento
----- Acesso 1 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=3
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=3
9 seg
----- Acesso 2 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=3
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=3
/sacarweb-alt/dados_paciente.php?local=pac&sub=dados&id=3
28 seg
----- Acesso 3 -----
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=3
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf&codigo=3
/sacarweb-alt/exibe_grafico.php?local=pac&sub=graf
29 seg
```

G.10 Tarefa 10: Relatórios

----- Acesso 4 -----

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=105
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=105
```

53 seg

----- Acesso 5 -----

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=253788
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=253788
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava
```

10 seg

----- Acesso 6 -----

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=1111
/sacarweb-alt/tela_avaliacao.php?local=pac&sub=ava&codigo=1111
```

12 seg

----- Acesso 7 -----

```
/sacarweb-alt/tela_relatorio2.php?local=rel
/sacarweb-alt/consulta_pacientes.php?local=pac&sub=cons
/sacarweb-alt/consulta_pacientes.php?ordem=&inicio=15&qtd_paginas=2&nome=
/sacarweb-alt/seleciona_paciente.php?local=pac&sub=cons&codigo=760349
/sacarweb-alt/avaliacao_inicial.php?local=pac&sub=ini
```

148 seg

--- Quality Measurements ---

Within groups variance: 1.09959799571

Between groups variance: 1.11933156882

Total variance: 0.157862707971

F-value: 1.01794617049

Wilk's Lambda value: 6.96553359463