

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Adaptação da interface de *webmail* às actividades de cada utilizador

Miguel Ângelo Oliveira Lima

VERSÃO DEFINITIVA

Relatório de Projecto
Mestrado Integrado em Engenharia Informática e Computação

Orientador: João José da Cunha e Silva Pinto Ferreira

Julho de 2009

Adaptação da interface de *webmail* às actividades de cada utilizador

Miguel Ângelo Oliveira Lima

Relatório de Projecto
Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: Pedro Alexandre Guimarães Lobo Ferreira do Souto

Arguente: Paulo Cortez

Vogal: João José da Cunha e Silva Pinto Ferreira

14 de Julho de 2009

Resumo

Sistemas adaptativos constituem uma área que tem vindo a ser cada vez mais investigada e explorada, contando já com diversas aplicações práticas em diferentes áreas, como são os casos do comércio electrónico, da educação, da saúde, de sistemas de filtragem de informação, sistemas de recomendação, entre outras. Estes sistemas pretendem detectar as preferências e padrões de comportamento dos seus utilizadores, de forma a adaptar-se automaticamente às preferências e padrões de comportamento destes. A ideia é melhorar a organização, conteúdo e apresentação do sistema de forma a atingir a satisfação dos seus utilizadores.

Sabendo das potencialidades destes sistemas, a Portugalmail como empresa orientada ao fornecimento de serviços de presença e comunicação na Internet, pretendeu integrar essas potencialidades na nova plataforma de *webmail* que se encontra em fase de desenvolvimento. A ideia é oferecer aos seus clientes um serviço de alta qualidade, centrado no conceito de uma secretária virtual inteligente, apostando em características como inovação, dinamismo e personalização. Nesse sentido, pretende-se a criação de uma interface inteligente que consiga proporcionar uma boa experiência no processo de interacção a todos os tipos de utilizadores.

Perante este desafio, decidiu-se aplicar uma abordagem modular, onde se divide o problema de adaptação nas suas etapas fundamentais. No entanto, esta abordagem que consiste na arquitectura proposta para o sistema adaptativo, necessita de ser integrada na arquitectura existente na plataforma de *webmail*. Tendo isto em conta, algumas funcionalidades serão idealizadas e implementadas de forma a comprovar o sucesso dessa integração. Para a implementação de cada uma, também serão propostos métodos e fórmulas que permitem calcular a sua utilidade e interesse para os diversos utilizadores.

Abstract

Adaptive systems is an area that is becoming more investigated and explored, which has several applications in different areas such as electronic commerce, education, healthcare, information filtering systems, recommendation systems, among others. These systems try to detect the preferences and behavioural patterns of their users, by automatically adapting to the users' preferences and behavioural patterns. The purpose of these systems is the improvement of the organization, content and system presentation in order to achieve user satisfaction.

Having into consideration these systems' features, Portugalmail (company that supplies Internet services) tried to integrate these features on the new webmail platform, which is in development phase. The purpose is to offer a high quality service to the users, focused on the virtual intelligent secretary concept, betting in innovation, dynamism and personalization. Based on these facts, we aim to create an intelligent interface that provides a good interactive experience to all kinds of users.

For a challenge like this one, we have decided to apply a modular approach, where the adaptation problems is divided into its main stages. However, this approach (which consists of the proposed architecture for the adaptive system), needs to be integrated on the existent webmail architecture. Having this into consideration, some features will be idealized and implemented in order to prove this integration success. For each feature implementation, methods and formulas will be proposed in order to allow the calculation of each feature utility and relevance to different kinds of users.

Agradecimentos

Um especial obrigado...

...ao Eng. Nuno Lopes (orientador de projecto na Portugalmail) por me colocar à vontade para exercer o meu trabalho, por todo o apoio, paciência e disponibilidade prestadas para conduzir o projecto da melhor maneira.

...ao pessoal em geral na Portugalmail pelo bom ambiente criado, pela companhia, pelo espirito positivo, por estarem sempre prontos para ouvir e ajudar.

...aos meus amigos que constituem uma verdadeira família para mim. Por me fazerem sempre sorrir, por me ouvirem, por me darem forças para enfrentar as adversidades, pelos momentos de diversão que valem todas as horas de trabalho, por me valorizarem pelo que sou e sobretudo, por me fazerem sentir especial. Um especial agradecimento para “Adams Smith” (amigo de todos os dias, horas, minutos...), Pedro e “Oigres7” (amigos de sempre), “Ogun” (pelo que é), “Daddy” (pelo que significa), “Guga Atentado” (só ele para me fazer rir em certos momentos)...

...à minha família pelo enorme esforço despendido na hárdua tarefa que foi educar-me - enquanto profissional, e sobretudo enquanto pessoa. Agradeço pelo apoio, pela confiança depositada nas minhas capacidades, pela liberdade que sempre usufruí para tomar as minhas decisões e seguir o meu caminho, por me encorajarem a acreditar no meu valor e principalmente, por me sentir amado.

Miguel Ângelo Oliveira Lima

Índice

1	Introdução.....	1
1.1	Projecto e Objectivos.....	1
1.2	Motivação.....	2
1.3	Contexto/Enquadramento.....	2
1.3.1	Portugalmail – Comunicações S. A.....	2
1.3.2	Serviço de Webmail da Portugalmail.....	3
1.4	Estrutura do Documento.....	4
2	Sistemas Adaptativos.....	6
2.1	Descrição/Contextualização do Problema.....	6
2.1.1	Sistemas Adaptáveis vs. Sistemas Adaptativos.....	7
2.1.2	Sistemas adaptativos	8
2.1.2.1	Formas de Sistemas Adaptativos.....	8
2.1.2.2	Aspectos a Considerar em Sistemas Adaptativos.....	9
2.2	Modelo do Utilizador.....	10
2.2.1	Classificação do Modelo do Utilizador.....	10
3	Estado da Arte.....	13
3.1	Sistemas Adaptativos em Serviços de Webmail.....	13
3.1.1	Predição de Necessidade de Resposta.....	14
3.1.2	Predição de Falta de Anexo.....	14
3.1.3	Organização/Classificação Automática de Emails.....	15
3.2	Sistemas de Recomendação.....	16
3.2.1	Algoritmos de Recomendação.....	17
3.2.2	Amazon.com.....	18
3.2.3	Recoo – Sistema de Recomendação para Youtube RSS Feeds.....	20
3.2.4	Recomendação de Tags em Folksonomias.....	22
3.3	Educação – Ambiente AdaptWeb.....	23
3.4	SeAN – Server for Adaptive News.....	24
4	Concepção da solução.....	25
4.1	Categorias e Tipos de Utilizadores.....	25
4.1.1	Organização.....	26
4.1.2	Uso do sistema.....	26

4.1.3	Frequência de acesso.....	26
4.1.4	Tipos de Emails.....	26
4.1.5	Volume de emails.....	26
4.1.6	Volume de contactos que comunica (VCC).....	27
4.1.7	Volume de contactos da lista de contactos (VCL).....	27
4.1.8	Rácio envio/recepção.....	27
4.1.9	Modo de navegação.....	27
4.1.10	Experiência Web.....	28
4.1.11	Outros.....	28
4.2	Brainstorming.....	28
4.3	Seleção das Ideias.....	32
4.3.1	Interface Especial para Redes Sociais.....	32
4.3.2	Detecção do Horário de Trabalho.....	32
4.3.3	Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete.....	33
4.4	Requisitos Funcionais.....	34
4.4.1	Actores.....	34
4.4.2	Descrição dos Requisitos.....	34
4.5	Requisitos Não Funcionais.....	35
5	Proposta de Solução.....	37
5.1	Proposta de Arquitectura.....	37
5.1.1	Plataforma de Webmail da Portugalmail.....	37
5.1.2	Módulo de Adaptação Automática.....	39
5.1.3	Etapas fundamentais.....	42
5.1.4	Sub-módulos.....	44
5.2	Abordagem Proposta.....	44
5.2.1	Funcionalidades Recomendadas.....	45
5.2.2	Funcionalidade Não Recomendada.....	48
6	Tecnologias utilizadas.....	51
6.1	Tecnologias de Backend.....	51
6.2	Tecnologias de Frontend.....	52
6.3	Tecnologias de Apoio.....	53
7	Sistema Adaptativo em Serviço de Webmail.....	54
7.1	Monitoração da Interação.....	54
7.2	Interpretação de Informação.....	55
7.3	Modelação.....	55
7.4	Tomada de Decisão	59
7.5	Aplicação de Adaptações.....	64
7.6	Avaliação Automática de Adaptações.....	68
8	Resultados.....	70
9	Conclusões.....	72
9.1	Sobre o Trabalho Desenvolvido.....	72
9.2	Sobre o Trabalho Futuro.....	73

Bibliografia.....	74
Anexo A: Planeamento do Projecto.....	80
Anexo B: Funcionamento da sugestão de grupos de contactos.....	82
Anexo C: Exemplos de resposta do sistema.....	86

Lista de Figuras

Figura 1.1: Portugalmail - Comunicações S. A.....	3
Figura 3.1: Sistema de recomendações na Amazon.com	18
Figura 3.2: Outra forma de recomendar itens pela Amazon.com.....	20
Figura 3.3: Recomendações de vídeos no Youtube	21
Figura 3.4: Sistema de recomendação no Last.fm	22
Figura 3.5: Interface da página de uma disciplina no ambiente AdaptWeb.....	24
Figura 4.1: Todos os tipos de utilizadores identificados e abordados.....	29
Figura 4.2: Exemplo das ideias geradas para um utilizador com uma frequência de acesso ao serviço elevada.....	30
Figura 4.3: Visão geral do resultado do brainstorming.....	31
Figura 4.4: Mapa mental bastante reduzido das ideias geradas.....	31
Figura 4.5: Exemplo da funcionalidade de autocomplete existente no preenchimento do campo de destinatário do Hotmail (os contactos de email foram criados exclusivamente para efeitos de exemplificação).....	33
Figura 5.1: Arquitectura da Plataforma da Portugalmail.....	39
Figura 5.2: Arquitectura proposta para o desenvolvimento do sistema adaptativo.....	41
Figura 7.1: Fluxograma correspondente à tomada de decisão da funcionalidade "Interface especial para redes sociais"	60
Figura 7.2: Fluxograma correspondente à tomada de decisão da funcionalidade "Detecção do horário de trabalho"	62
Figura 7.3: Fluxograma correspondente à tomada de decisão da funcionalidade "Sugestão de grupos de contactos na funcionalidade de autocomplete"	64
Figura B.1: Inserir "user" e escolher terceira opção no IP 192.168.100.32.....	82
Figura B.2: Enviar email para user6@portugalmail.net no IP 192.168.100.32.....	82
Figura B.3: Consultar lista de sugestões para "user" no IP 192.168.100.32.....	83
Figura B.4: Inserir "user" no IP 192.168.100.33.....	83
Figura B.5: Enviar email a user6@portugalmail.net no IP 192.168.100.33.....	83
Figura B.6: Consultar lista de sugestões para "user" no IP 192.168.100.32.....	84
Figura B.7: Escolher sexta opção no IP 192.168.100.32.....	84
Figura B.8: Escolher primeira opção para "user4@portugalmail.net, user" no IP 192.168.100.32	84
Figura B.9: Enviar email tendo escolhido a primeira opção no IP 192.168.100.32.....	85
Figura B.10: Consultar lista de sugestões para "user" no IP 192.168.100.32.....	85

Lista de Tabelas

Tabela 1: Eventos que fazem activar o processo relativo a cada funcionalidade.....	55
Tabela 2: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Interface Especial para Redes Sociais” e seus respectivos campos.....	56
Tabela 3: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Detecção do Horário de Trabalho” e seus respectivos campos.....	56
Tabela 4: Tabela comum às duas primeiras funcionalidades.....	57
Tabela 5: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete” e seus respectivos campos.....	58
Tabela 6: Planeamento do projecto.....	80

Abreviaturas e Símbolos

ADSL	Asymmetric Digital Subscriber Line
AJAX	<i>Asynchronous Javascript And XML</i>
AMCo	<i>Automatic Mail Category Organizer</i>
API	<i>Application Programming Interface</i>
CC	<i>Carbon Copy</i>
CSS	<i>Cascade Style Sheet</i>
DIMP	Dynamic Internet Messaging Program
HTML	<i>HyperText Markup Language</i>
HTTP	<i>HyperText Transfer Protocol</i>
IMAP	<i>Internet Message Access Protocol</i>
IMP	<i>Internet Messaging Program</i>
IP	<i>Internet Protocol</i>
LAMP	Linux, Apache, MySQL and PHP/Perl/Python
LDAP	<i>Lightweight Directory Access Protocol</i>
PHP	<i>Hypertext Preprocessor</i>
POP3	<i>Post Office Protocol version 3</i>
RF	Requisito Funcional
RNF	Requisito Não Funcional
RSS	<i>Really Simple Syndication</i>
SQL	<i>Structured Query Language</i>
TF-IDF	<i>Term Frequency – Inverse Document Frequency</i>
VoIP	<i>Voice Over Internet Protocol</i>
W3C	<i>World Wide Web Consortium</i>
Web	<i>World Wide Web</i>
XHTML	<i>Extensible HyperText Markup Language</i>
XML	<i>Extensible Markup Language</i>

1 Introdução

O presente documento pretende demonstrar o resultado de todas as fases de concepção, análise e desenvolvimento empreendidas no projecto “Adaptação da interface de *webmail* às actividades de cada utilizador”. Neste primeiro capítulo descreve-se a motivação para a escolha do tema de projecto, os objectivos principais em que este se fundamenta, o contexto em que este se insere e os resultados esperados após a sua implementação. Para além do referido, é ainda possível encontrar uma breve apresentação da instituição na qual foi concretizado o projecto - Portugalmail - Comunicações S. A..

1.1 Projecto e Objectivos

O projecto surge baseado na ideia de que uma interface por melhor desenhada que seja, não consegue ir ao encontro das preferências e interesses de todos os utilizadores. Assim, pretende-se proporcionar uma melhor experiência de utilização do serviço de *webmail* da Portugalmail a todos os utilizadores individuais. Alguns tipos de adaptação já se encontram disponíveis em alguns serviços deste género, embora a iniciativa e decisão do que adaptar esteja a cargo do próprio utilizador, ficando o sistema limitado em apenas disponibilizar tais funcionalidades. Muitos utilizadores podem até desconhecer que estas funcionalidades existem ou onde as podem encontrar. No sentido de oferecer um papel mais activo ao sistema, são idealizadas funcionalidades que permitem detectar certos padrões de utilização abrangendo um número maior de adaptações do que as típicas alterações do visual. Basicamente, o pretendido é simplificar ou melhorar o modo de interacção do utilizador com o sistema com base nas suas actividades e padrões de comportamento. O desenho das adaptações não constitui um objectivo deste projecto, mas sim a conclusão que uma determinada adaptação pode ser útil a um utilizador, aumentando assim a sua satisfação.

1.2 Motivação

Como utilizadores de um sistema, é mais gratificante interagir com um sistema que proporcione uma boa experiência no processo de interacção do que um sistema que não satisfaça devidamente os nossos gostos e personalidade, que não esteja adaptado a nós da maneira que desejaríamos (diferentes pessoas possuem diferentes comportamentos). Assim, este processo de análise e adaptação de um sistema interactivo às actividades de cada utilizador é bastante motivador pelo simples facto de se poder tornar um sistema mais interessante e popular, abrangendo todo o tipo de utilizadores. Formas muito simples podem ser adoptadas para alcançar a satisfação dos utilizadores, como por exemplo: alterar certos aspectos da interface, apresentar a mesma informação de uma forma diferente, dar ênfase a algumas aplicações ou partes do sistema, ou até disponibilizar diferentes tipos de ordenação.

Outros factores contribuíram de forma decisiva para a escolha deste projecto: o facto da Portugalmail ser a empresa responsável pelo projecto e a orientação deste estar direccionada para a *Web*.

1.3 Contexto/Enquadramento

O projecto encontra-se inserido principalmente na área de sistemas adaptativos, uma vez que o sistema procura adaptar-se a todos os utilizadores individuais conforme a informação que vai recolhendo destes, tomando a iniciativa na aplicação das adaptações. No entanto, este projecto pode ser enquadrado em outras áreas como é o caso da área Interacção Pessoa-Computador já que toda a envolvente do sistema se encontra à volta da interacção do utilizador com o sistema, e na área de serviços *Web* devido ao sistema ser integrado noutro sistema disponibilizado através da Internet, como é o caso do novo serviço de *webmail* da Portugalmail – Comunicações S. A..

1.3.1 Portugalmail – Comunicações S. A.

A Portugalmail - Comunicações S. A. foi fundada em 1998 consistindo num operador de comunicações electrónicas que se dedica ao fornecimento de serviços de presença e comunicação na Internet. A empresa foi constituída para dar suporte ao lançamento do primeiro serviço de *webmail* gratuito em português inaugurado a 20 de Março de 1999. Devido ao sucesso obtido, a empresa apostou no desenvolvimento de uma plataforma de comunicação mais abrangente e o alargamento dos seus serviços ao mercado empresarial [Portugalmail.pt, 2009]. Soluções de email, soluções de blog (www.blog.com), software de gestão PHC, centrais telefónicas VoIP, acesso à internet ADSL, registo de domínios, alojamento Web e um sítio sobre futebol (www.futebol.com) são alguns dos serviços prestados pela Portugalmail.



Figura 1.1: Portugalmail - Comunicações S. A.

Actualmente, a Portugalmail é uma referência na internet portuguesa, gerindo directa ou indirectamente cerca de 20% das caixas de correio usadas pelos portugueses, orgulhando-se ainda por merecer a confiança das mais proeminentes empresas portuguesas, como são os casos da Media Capital, Grupo Sonae, Grupo M. Coutinho ou Onitelecom [Portugalmail.net, 2009; Portugalmail.pt, 2009; AvilaCosta, 2009].

1.3.2 Serviço de *Webmail* da Portugalmail

O desenvolvimento e gestão de plataformas de correio electrónico constituem o principal negócio da empresa, as quais estão segmentadas em três segmentos alvo: serviço de email gratuito, serviço de email profissional dedicado à área empresarial e um serviço dirigido a portais e empresas de telecomunicações [Portugalmail.net, 2009].

Serviço de Email Gratuito

O serviço de email gratuito é dirigido ao utilizador particular, servindo mais de 370 mil caixas de correio que apresentam 2GB de capacidade e podem ser acedidas a partir do portal portugalmail.pt via *webmail* ou POP3. A rentabilidade deste serviço concentra-se na publicidade *online* apresentada aos seus utilizadores [Portugalmail.net, 2009].

Serviço de Email Profissional

Esta oferta foi criada com o objectivo de melhorar os níveis de produtividade de quem usa o email como instrumento de trabalho e é comercializada em pacotes de caixas de correio com um custo mensal. No entanto, a solução não obriga a qualquer custo de investimento em software ou hardware, permitindo uma gestão personalizada e eficaz dos servidores de email (atribuição de privilégios por colocador da empresa ou definição de políticas de conteúdos são exemplos dessa gestão personalizada) [Portugalmail.net, 2009].

Portal Email ASP

O serviço de email elevou-se a elemento diferenciador, fidelizador e mostruário do nível de qualidade dos serviços prestados pelos portais, ISP e empresas de telecomunicações. Sendo incapazes de desenvolver internamente soluções dentro dos patamares elevados existentes nas ofertas gratuitas globais, aderiram ao *outsourcing* dos serviços de email de grande escala. A Portugalmail como empresa especializada no desenvolvimento de plataformas de

email apercebeu-se desta tendência e desenvolveu uma oferta para responder a esta necessidade do mercado [Portugalmail.net, 2009].

Mais recentemente, foi desenvolvida uma interface em AJAX para o serviço de *webmail*, permitindo à Portugalmail alcançar os patamares de qualidade dos maiores produtores de serviços deste género e provocando a migração das contas dos utilizadores das versões anteriores para a nova interface.

Actualmente, a Portugalmail encontra-se em fase de desenvolvimento de uma plataforma de *webmail* totalmente nova, ambicionando colocar-se como líder em características como inovação e aplicações disponibilizadas. Esta nova plataforma consiste no resultado de um estágio realizado na empresa e apresenta como conceito principal o desenvolvimento de uma secretária virtual que ajuda os utilizadores na organização da informação presente na suite de email. Nesse sentido, pretende-se dotar os diversos módulos existentes na suite (email, contactos pesquisa, calendário, ficheiros e RSS) de alguma inteligência. Esta nova plataforma ainda pretende promover uma forte interligação entre as suas componentes e também entre estas com outros sistemas [AvilaCosta, 2009]. É sobre esta nova plataforma de *webmail* que o sistema adaptativo desenvolvido neste projecto vai ser integrado, de forma a contribuir para o aspecto inteligente que se pretende atribuir à secretária.

1.4 Estrutura do Documento

Para além da introdução, este documento encontra-se organizado em mais sete capítulos essenciais, os quais encontram-se seguidamente descritos.

No segundo capítulo é possível encontrar uma descrição mais detalhada do problema que temos em mãos, especificando as suas fronteiras (distinção existente entre tipos de sistemas onde pode ocorrer ajustes à interface) e classificando o modelo do utilizador, considerando certos critérios.

O terceiro capítulo é dedicado ao estado da arte que corresponde ao estudo e definição do que de mais avançado existe para resolver de uma forma conveniente o problema, apresentando alguns trabalhos relacionados.

No capítulo quarto pretende-se dar a conhecer as fases envolvidas no processo de concepção das ideias para adaptações, que podem ser aplicadas ao serviço de *webmail* inteligente da Portugalmail. Também pode ser consultada informação referente à identificação e descrição dos tipos de utilizadores de um serviço de *webmail*, à selecção e descrição das funcionalidades a implementar, bem como a uma análise de requisitos.

O quinto capítulo tem como objectivo dar a conhecer a proposta de solução, onde se descreve a proposta de configuração da arquitectura do sistema adaptativo e os métodos adoptados nos processos de tomada de decisão e aplicação das adaptações.

No capítulo sexto apresenta-se as tecnologias necessárias para o desenvolvimento do sistema, referindo a sua contribuição no projecto.

O sétimo capítulo corresponde à implementação do sistema adaptativo, descrevendo todas as fases envolvidas no desenvolvimento de cada uma das funcionalidades adaptativas.

No oitavo capítulo apresenta-se os resultados obtidos no final da fase de implementação.

O último capítulo refere-se às conclusões, onde se realiza uma análise geral ao projecto, e à descrição do trabalho que ainda pode ser desenvolvido futuramente.

2 Sistemas Adaptativos

Neste capítulo pretende-se apresentar o problema de uma forma detalhada, especificando as suas fronteiras, pontos fortes e fracos e suas características principais. Também é possível encontrar a definição de modelo de utilizador, bem como a classificação deste, neste caso em específico, segundo cinco critérios.

2.1 Descrição/Contextualização do Problema

O problema de um bom design/interface é baseado em vários factores derivados dos seguintes factos [PerkowitzEtzioni, 1997]:

- Diferentes utilizadores possuem diferentes objectivos e conhecimentos e podem estar interessados em diferentes porções de informação apresentadas numa página, bem como utilizar diferentes *links* para a navegação entre páginas.
- O mesmo visitante pode procurar ou requerir informações diferentes em momentos diferentes.
- Muitos *sites* evoluem do seu estilo original, acumulando *links* e páginas em locais indesejados.
- Um site pode ser desenhado para um determinado tipo de uso, mas na prática ser usado de diferentes formas.
- A informação pretendida nem sempre é fácil de aceder ou encontrar.

Imagine-se um sistema que seria capaz de determinar o que é realmente importante para um utilizador e de fazer com que essa informação se tornasse fácil de encontrar; de colocar as páginas ou aplicações mais importantes cada vez mais visíveis e acessíveis; de colocar aos poucos informação irrelevante ou obsoleta fora do caminho do utilizador. Assim, a criação de sistemas adaptativos tornou-se cada vez mais necessária, importante e factor de diferenciação.

Estes sistemas pretendem melhorar de uma forma automática a sua organização, conteúdo e apresentação, adaptando-se ao utilizador de acordo com o respectivo modelo (representação do conhecimento do utilizador com base no seus padrões de uso). Por outras palavras, os utilizadores acedem a interfaces cujos estilo, conteúdo e apresentação serão dinamicamente seleccionados no sentido de serem apresentadas conforme os objectivos, necessidades, preferências, desejos e características específicas de cada utilizador.

2.1.1 Sistemas Adaptáveis vs. Sistemas Adaptativos

Antes de prosseguir com a apresentação de sistemas adaptativos, convém enquadrar correctamente o leitor no tipo de sistemas que se está a falar para que não exista qualquer tipo de mal entendidos. Existe distinção entre dois tipos de sistemas que são capazes de fornecer algum tipo de personalização. Estes podem ser adaptáveis (costumizáveis/ajustáveis - *Adaptable*) ou adaptativos (*Adaptive*) [Gasparini, 2003]. Em ambos os casos o utilizador possui um papel fundamental para que o sistema possa oferecer a personalização, embora a informação que é fornecida por parte do utilizador ao sistema seja fornecida de forma explícita no primeiro caso ao contrário do segundo, onde o sistema infere essa informação. Seguidamente, apresentam-se as características principais do dois tipos de sistemas acima supracitados.

Sistemas adaptáveis

Nos sistemas adaptáveis [Gasparini, 2003]:

- O utilizador pode fornecer algum tipo de perfil (através de um diálogo ou questionário).
- Com base no perfil, o sistema fornece uma versão da aplicação correspondente.
- Ajustes podem incluir determinadas preferências da apresentação (cores, estilo, etc).

Sistemas adaptativos

Nos sistemas adaptativos [Gasparini, 2003]:

- O sistema adapta-se automaticamente.
- O sistema monitora o comportamento do utilizador e adapta a apresentação adequadamente.
- A evolução das preferências do utilizador e do seu conhecimento pode ser deduzida dos acessos às páginas.
- Adaptação essencialmente baseada nas acções de navegação do utilizador, e possivelmente também no comportamento de outros utilizadores.
- Adaptação pode ser feita modificando-se vários aspectos visíveis do sistema e formas de acesso preestabelecidas.

- Tentam antecipar as expectativas dos utilizadores a partir de modelos que reflectem algumas características dos utilizadores.
- O sistema possui um papel activo no que diz respeito ao processo de decisão da adaptação.
- Conhecidos também como **sistemas dinâmicos**.

Depois da diferenciação entre os dois tipos de sistemas, é importante referir que a partir deste ponto apenas serão feitas referências ao segundo tipo de sistemas (sistemas adaptativos) já que o tema de projecto abordado neste documento se concentra neste tipo de adaptação.

2.1.2 Sistemas adaptativos

Muitas questões surgem no contexto deste problema, as quais são respondidas ao longo do documento [PerkowitzEtzioni, 1997; Gasparini, 2003] :

- Como é que é possível construir um *website* que se desenvolve por si próprio em resposta à interacção que este possui com os utilizadores?
- Onde é que sistemas adaptativos podem ser realmente úteis?
- Que tipo de generalizações podemos retirar dos padrões de uso dos utilizadores?
- Que características dos utilizadores podem ser utilizadas como fonte de informação para o processo de decisão na aplicação de uma determinada adaptação?
- Que tipo de alterações é possível efectuar (e como)?
- Como se desenha um sistema para ser adaptável?

Sistemas adaptativos são essencialmente úteis quando há necessidade de disponibilizar informação selectiva e contextual a diversos utilizadores com objectivos e níveis de conhecimento diferentes e em situações em que a rede de nós/links é razoavelmente grande. Sistemas adaptativos constituem a alternativa para a teoria "one-size-fits-all", onde é apresentado o mesmo conjunto de páginas, a mesma interface e o mesmo conteúdo independentemente do utilizador [Gasparini, 2003].

Brusilovsky [Brusilovsky, 1996] identifica seis principais áreas de aplicações onde sistemas adaptativos podem ser utilizados: sistemas educacionais, sistemas de informação *online*, sistemas de ajuda *online*, sistemas de recuperação e busca de informação, sistemas de informação institucional e sistemas para gerir visões personalizadas (o serviço de *webmail* da Portugalmail enquadra-se nesta área).

2.1.2.1 Formas de Sistemas Adaptativos

Os sistemas podem ser adaptativos em duas formas básicas: focando-se em **customização** ou focando-se em **optimização** [PerkowitzEtzioni, 1997].

Customização

- Alterar as páginas em tempo real de forma a ajustar-se aos utilizadores individuais.
- **Path prediction** – Customiza automaticamente tentando adivinhar onde o utilizador quer ir, levando-o até lá mais rapidamente.
 - Prever o próximo passo do utilizador e, principalmente, o objectivo final do utilizador (por exemplo, num *site* se pudermos determinar que página um utilizador está à procura, apresentá-la de imediato).

Optimização

- Alterar o sistema de forma a fazer a navegação mais simples para todos.
- Alterações são efectuadas para todos os utilizadores (generaliza a aprendizagem obtida de todos os utilizadores).
 - Permite novos utilizadores que não conhecem nada saírem beneficiados.
 - **Promotion** – Se uma página popular está à distância de cinco *clicks* da *homepage*, será necessário colocá-la mais acessível (como por exemplo um *link* a partir da *homepage*).
 - **Clustering** – Colocar numa mesma página um conjunto de objectos relacionados (o sistema deve inferir que um grupo de páginas estão relacionadas umas com as outras baseando-se tanto no conteúdo como também na navegação dos visitantes).

2.1.2.2 Aspectos a Considerar em Sistemas Adaptativos

Uma desvantagem deste tipo de sistemas é o facto do sistema não conseguir decifrar qual é a verdadeira intenção do utilizador aquando este assume um determinado comportamento, colocando informação errada no respectivo modelo. Também, devido a não existir uma interface comum entre utilizadores, torna-se complicado para os utilizadores ajudarem-se mutuamente em questões relacionadas com o modo de utilização do sistema.

Outro aspecto negativo está relacionado com o facto de um sistema adaptativo poder ser inconsistente, quando este não define correctamente as suas estratégias de adaptação. Estes sistemas fazem com que algumas vezes a interface do sistema se modifique com demasiada frequência, só para ir de acordo com o modelo do utilizador que se modifica a cada nova descoberta. Isto pode obrigar o utilizador a recomeçar novamente todo o processo de aprendizagem da utilização de um sistema, o que é bastante negativo caso este já se tenha acostumado. No entanto, interfaces adaptativas podem ser úteis para tarefas similares, unificando o modo de interacção e aumentar a consistência entre as aplicações. Também, podemos limitar a capacidade do sistema auto-adaptar-se, assegurando que não piora a interface do *website*, mas também se limita o alcance do melhoramento. Em vez de fazer alterações sem

qualquer tipo de segurança, o sistema pode inteligentemente apresentar sugestões (a um *webmaster* ou ao utilizador), complementadas por uma explicação e justificação [PerkowitzEtzioni, 1997].

2.2 Modelo do Utilizador

O modelo do utilizador representa toda a informação que o sistema conhece sobre o utilizador. A manutenção deste modelo é da responsabilidade do sistema, o qual está atento a todos os acontecimentos e acções realizadas pelo utilizador de forma a detectar padrões de comportamento (como tarefas efectuadas, requisições de ajuda, recepção de um email, etc). O modelo do utilizador apesar de poder ser armazenado no lado do utilizador (*cookies*), este deve ser guardado no lado do servidor devido a: as *cookies* possuem limitações de tamanho e número, e quando estas tornam-se um pouco maiores podem atrasar a transmissão para o servidor. Outro ponto importante é que os utilizadores devem ter conhecimento da existência de um modelo a ser desenvolvido e que o sistema encontra-se personalizado a si de forma a satisfazer melhor as suas necessidades, isto é, o utilizador deve saber que o sistema não aparece exactamente da mesma maneira que aos outros utilizadores. As adaptações levadas a cabo pelo sistema devem ser justificadas para o utilizador compreender melhor o porquê dessa adaptação.

2.2.1 Classificação do Modelo do Utilizador

Os modelos dos utilizadores podem ser classificados segundo cinco critérios: grau de especialização, grau de alteração, extensão temporal, método de uso e número de modelos [Gasparini, 2003]. De seguida, aborda-se cada um desses critérios, indicando o grau de importância na adaptação da interface de *webmail* às actividades de cada utilizador.

Grau de especialização

De acordo com este critério, os modelos podem ser classificados de duas formas diferentes: **genéricos** e **individuais**. Os **modelos genéricos** pressupõem um conjunto homogéneo de utilizadores. Estes modelos devem ser utilizados quando os utilizadores possuam tanta similaridade para que possam ser tratados da mesma forma. Os modelos **individuais** possuem dados correspondentes ao utilizador individual. O problema destes modelos consiste na grande necessidade de memória, ainda mais se a comunidade de utilizadores do sistema for elevada.

A utilização de um modelo misto, através da utilização conjunta das duas abordagens anteriores é também uma solução. O modelo genérico seria responsável pelas características gerais do grupo em questão e o modelo individual pelas características específicas ou até mesmo pelas diferenças de um determinado utilizador no grupo ao qual pertence [Gasparini, 2003].

Na abordagem seguida o modelo do utilizador é numa primeira fase individual (foca-se exclusivamente nos dados individuais de cada utilizador) podendo depressa evoluir para um misto a partir do momento que se identifique os tipos de utilizadores nos quais os utilizadores se

podem inserir. Assim, será possível sugerir a aplicação de uma mesma adaptação a um conjunto de utilizadores com características semelhantes. Também poderá ser possível realizar uma recomendação conjunta conforme o grau de similaridade em relação ao histórico de adaptações, isto é, se um conjunto de utilizadores possuírem o mesmo comportamento em relação às sugestões de adaptações será provavelmente admissível (sob análise) efectuar a mesma sugestão a todos os utilizadores desse conjunto (se uma sugestão de adaptação pode interessar a um utilizador, muito provavelmente também pode interessar aos outros desse conjunto).

Grau de alteração

A possibilidade dos modelos poderem ser modificados, leva a que possam ser considerados **estáticos** ou **dinâmicos**. Os **modelos estáticos** ou explícitos permanecem inalterados durante toda a interacção. O utilizador define numa fase inicial o seu conteúdo e perfil, mantendo-se da mesma forma durante todo o tempo. Os modelos **dinâmicos** ou implícitos são actualizados de acordo com a obtenção de novos conhecimentos. A informação sobre um dado utilizador pode mudar de sessão para sessão, ou mesmo dentro da própria sessão. Entenda-se por sessão, todo o processo de interacção do utilizador com o sistema, desde a autenticação até à saída do sistema [Gasparini, 2003].

Neste projecto o modelo é classificado como dinâmico, uma vez que, todas as adaptações e actualizações da interface são da responsabilidade do sistema, dos padrões de comportamento detectados de cada utilizador ou tipo de utilizador. Embora alguma informação possa ser retirada do perfil de um utilizador, esta não é fornecida de forma explícita com o objectivo de adaptação.

Extensão temporal

Um modelo também pode ser classificado atendendo à sua durabilidade. Se um modelo é criado e armazenado durante toda a sessão e descartado no final, é considerado um modelo a **curto termo**. Se um modelo é preservado de uma sessão para a outra é considerado a **longo termo**.

No desenvolvimento do projecto adopta-se um modelo a longo termo, já que, os conhecimentos apreendidos dos utilizadores são guardados de forma a subsistirem de uma sessão para outra (a apresentação e navegação do sistema é mantida ou actualizada a partir da última sessão). No entanto, dados que já ultrapassaram o seu prazo de validade, isto é, que já não têm qualquer tipo de valor para o processo de decisão na sugestão ou aplicação de uma adaptação, podem ser removidos de forma a economizar espaço de armazenamento.

Método de uso

De acordo com este critério, os modelos do utilizador podem ser **prescritivos** (têm a finalidade de prever ou entender as acções de um utilizador, realizando a simulação do seu comportamento) ou **descritivos** (descrevem o utilizador podendo serem considerados como uma base de informações sobre o mesmo).

No desenvolvimento do projecto adopta-se o modelo descritivo, pois a adaptação é realizada com base na observação do comportamento do utilizador, não tentando simulá-lo ou prevê-lo.

Número de modelos

Um único utilizador pode ser representado por vários modelos, dependendo da situação. Um exemplo óbvio desta situação é: um utilizador quando está no trabalho assume um determinado comportamento e quando chega a casa ou encontra-se em lazer assume outro tipo de comportamento. A diferença de interacção com o sistema está de acordo com o contexto/ambiente em que o utilizador se encontra. Para o sistema o utilizador consiste numa pessoa com diferentes personalidades. O objectivo é saber que personalidade este assume em diversos momentos.

3 Estado da Arte

Neste capítulo pretende-se apresentar o resultado de algumas semanas de pesquisa e investigação acerca do estado da arte no que diz respeito a sistemas adaptativos. Inicialmente, a ideia era organizar o processo de investigação numa só fase: focalizado em sistemas de *webmail* onde existisse algum tipo de adaptação automática. No entanto, devido a esta área ainda não ter sido muito explorada no que diz respeito a sistemas adaptativos, a pesquisa foi alargada a todos os sistemas adaptativos existentes (onde se destaca os sistemas de recomendação existentes em várias aplicações conhecidas como Youtube, Flickr, Amazon ou Del.icio.us).

3.1 Sistemas Adaptativos em Serviços de *Webmail*

Hoje em dia, é possível configurar e personalizar a interface de um serviço de *webmail* conforme o gosto e vontade de cada utilizador individual. Alterar a *skin* da caixa de correio ou a cor de fundo são alguns exemplos deste tipo de personalização. No entanto, todas estas personalizações são disponibilizadas pelo sistema sem qualquer tipo de distinção entre diferentes utilizadores e são levadas a cabo pelo próprio utilizador. O objectivo do projecto “Adaptação da interface de *webmail* às actividades de cada utilizador” é disponibilizar este tipo de personalizações, mas onde o sistema é o principal responsável pela decisão de disponibilizar, recomendar ou aplicar uma determinada alteração ao ambiente interactivo de um determinado utilizador.

Numa análise mais detalhada, é possível encontrar registos de sistemas e técnicas que foram desenvolvidos e estudados com o intuito de fornecer algum tipo de inteligência a um serviço de *webmail*. Essas técnicas passam por:

- Predição de necessidade de resposta a um email (detectar se um determinado email necessita de resposta).
- Predição de falta de anexo (detectar se existe falta de anexo quando se pretende enviar um email).

- Organização/Classificação automática de emails.

3.1.1 Predição de Necessidade de Resposta

A primeira das técnicas referidas (predição de necessidade de resposta a um email) utiliza características relacionais existentes no perfil do utilizador construído ao longo do tempo, como: o número de mensagens enviadas e recebidas de cada contacto, se o contacto que enviou a mensagem está na lista de contactos do utilizador, se o utilizador apenas aparece na lista de CC ou se aparece como destinatário principal, ou ainda se é costume o utilizador responder ao contacto em questão. Nesta técnica são ainda utilizadas características orientadas ao conteúdo da mensagem como: detectar a presença de "question marks" no email ou até pedidos onde não são utilizadas essas "question marks" (dificuldade elevada), e ainda verificar a existência de palavras tipicamente utilizadas em questões. Para esta última funcionalidade são examinadas todas as palavras e identificadas quais as palavras que melhor indicam que se trata de uma questão através de uma classificação TF-IDF (relaciona a número de vezes que uma palavra aparece numa questão e o número total de frases que contenham essa palavra). As 30 palavras com uma classificação TF-IDF mais elevada são nomeadas "question-words". O tamanho do conteúdo da mensagem, a existência de saudações, hora do dia e a presença de anexo são outras características utilizadas para a implementação desta técnica [DredzeBrooks, 2008; DredzeWallach, 2008]. Esta técnica é essencialmente útil para alertar os utilizadores da existência de mensagens na caixa de entrada que necessitam de ser tratadas urgentemente, facilitando todo o processo de identificação e resposta a essas mensagens.

3.1.2 Predição de Falta de Anexo

A transferência de ficheiros através do envio de mensagens com anexo constitui um dos usos mais comuns de um serviço de *webmail*. Assim, a segunda técnica acima supracitada permite que utilizadores não se esqueçam de associar um anexo a uma mensagem caso este seja necessário. O problema de esquecimento de anexo é tão incomodativo que alguns utilizadores desenvolveram o *plugin* Mac OS X Mail que alerta os utilizadores da possível falta de anexo, baseando-se apenas na presença ou não da palavra "attach" no conteúdo. Esta técnica é mais dependente do conteúdo da mensagem do que a técnica anteriormente apresentada, já que verifica a posição da palavra "attach" (primeira ou última frase), analisa a conjugação da palavra "attach" com outras palavras na proximidade e ainda procura a existência de palavras-chave no conteúdo do email como "attach", "attached", "attaching", "attachment". No entanto, é construído e mantido um perfil de utilização para cada utilizador no sentido de guardar a percentagem de mensagens enviada para e por cada utilizador que possuem anexo (capturar padrões de comportamento de forma a saber quem são os utilizadores que enviam muitos anexos e para quem). No sentido de avisar o utilizador surgem algumas soluções de interface: apresentar documentos sugeridos numa sidebar para encorajar o utilizador a seleccionar um anexo ou de uma maneira mais agressiva alertar o utilizador que existe falta de anexo aquando este tenta enviar uma mensagem (similar a alertar falta de assunto). No uso desta técnica há que

ter cuidado na identificação de mensagens enviadas que tenham anexo, porque alguns anexos podem ser do tipo ".html" ou ".eml" que constituem uma alternativa HTML de visualização do conteúdo e que podem ser fruto do reencaminhamento de mensagens (não é claro se o anexo teve origem no reencaminhamento ou no novo remetente) [DredzeBrooks, 2008; DredzeWallach, 2008; DredzeBlitzer, 2006].

3.1.3 Organização/Classificação Automática de Emails

A organização do email é uma actividade onde muitos utilizadores perdem algum do seu tempo e possivelmente dinheiro (em consumos de largura de banda e ligações *dial-up*). Na tentativa de resolução deste problema foram adoptadas técnicas baseadas em algoritmos "data mining" para a classificação das mensagens recebidas. A detecção de emails indesejados ("spam"), a análise de "newsletters" e "mailing lists", e a detecção e separação de mensagens importantes das não importantes tornaram-se hoje em dia problemas com que os utilizadores se deparam com alguma frequência e que gostariam de ver resolvidos ou pelo menos amenizados. Algumas funcionalidades foram desenvolvidas com o objectivo de ajudar os utilizadores a organizar a caixa de email. No entanto, estas necessitam que o utilizador assuma um papel activo para atingir o efeito desejado, na medida em que era requerido ao utilizador que este defini-se filtros (lista de palavras-chave e regras) que eram usadas no reconhecimento de características importantes nas mensagens. Esta abordagem não é a mais correcta já que para grande volume de emails torna-se incomportável encontrar manualmente os pontos fundamentais e também porque o conteúdo dos emails vão-se alterando frequentemente, levando a uma revisão constante dos filtros por parte do utilizador. Existiam várias abordagens para providenciar a organização do email, só que eram baseadas numa classificação supervisionada, isto é, as pastas para onde eram encaminhadas as mensagens tinham que previamente existir, sendo apenas necessário detectar a pasta mais adequada a cada mensagem. O pretendido seria uma abordagem onde as pastas fossem criadas automaticamente, cada uma associada a um determinado tema, a partir de um determinado conjunto de mensagens [MancoMasciari, 2002]. O sistema de organização AMCo (Automatic Mail Category Organizer) foi desenvolvido apoiado nesta abordagem e é apresentado seguidamente.

Numa fase inicial será necessário identificar os termos mais relevantes no conjunto de mensagens. Para este fim são removidas todas as "stopwords" que consistem nos termos que aparecem com maior frequência nas mensagens, os quais não são relevantes já que não distinguem as mensagens. De seguida os termos são reduzidos à sua forma básica (reduzindo o índice de termos), isto é, são retirados as variantes sintácticas das palavras principalmente através da remoção de plurais, gerúndios, tempos, prefixos ou sufixos.

Para definir automaticamente a importância das mensagens e a sua organização em grupos homogéneos ("clusters") é utilizada uma técnica que consiste: na construção de listas com os termos mais relevantes nas mensagens, na definição de um critério de correspondência para as mensagens a partir desses termos e na exploração de métodos de categorização e de "clustering" de forma a rotular cada mensagem. Esta técnica implica a definição de um processo de extracção de conhecimento. O primeiro passo neste processo consiste na extracção de

informação estruturada ou não estruturada de cada mensagem relacionada com os seguintes campos: remetente, destinatário, data/hora, assunto, nome do anexo e conteúdo da mensagem. Informação não estruturada é obtida sobretudo a partir do assunto e do conteúdo (o conteúdo possui muita informação redundante ao contrário do assunto que possui os termos chave de uma mensagem). Com o objectivo de caracterizar um "cluster" de forma apropriada deve ser escolhido um termo que ocorra com muita frequência no conjunto de mensagens de um "cluster" e com pouca frequência nos outros. No sentido de simplificar o problema podia ser escolhidos os três termos mais frequentes de um "cluster" que não aparecem dentro de outros. Agora falando mais propriamente no algoritmo de "clustering" utilizado, foi seguida uma abordagem baseada na integração de outros dois algoritmos: Hierarchical Agglomerative Algorithm e K-Means Clustering Algorithm. Hierarchical Agglomerative Algorithm é usado para construir uma hierarquia de "clusters", onde esta hierarquia é tradicionalmente representada por uma árvore, onde as folhas representam os "clusters" constituídos com os seus termos mais relevantes e a raiz um único "cluster" com o conjunto inteiro de termos [WikiClusterAnalysis, 2009; WikiHierarchical, 2009]. Concretamente neste caso, as folhas representam as mensagens com os seus termos mais importantes. O algoritmo utiliza uma estratégia aglomerativa ("bottom-up") uma vez que junta passo a passo os termos/conjunto de termos com maior associação segundo um critério de distância (termos com uma associação mais forte são agrupados primeiro). Pretende-se com a execução deste algoritmo seleccionar a melhor partição do conjunto de mensagens. K-Means Clustering Algorithm é usado como um técnica de refinamento dos resultados obtidos pelo algoritmo anterior, definindo o número óptimo de "clusters" (valor óptimo para o critério de paragem) [WikiKmeans, 2009].

As três técnicas acima apresentadas têm como principal objectivo amenizar o problema da sobrecarga de emails, na medida em que facilita a gestão da caixa de entrada (identificando e organizando os emails mais importantes e que necessitam de resposta) e em que evita o envio da mesma mensagem várias vezes (para si próprio ou não) por esquecimento do anexo.

3.2 Sistemas de Recomendação

Os sistemas de recomendação são bastante populares pelo seu uso em páginas Web relacionadas com o comércio electrónico, e possuem como propósito a apresentação de uma lista de itens baseada nas preferências e interesses de cada utilizador. Estes sistemas ainda pretendem eliminar o esforço que um utilizador possui quando navega sobre todo o conjunto de dados à procura de algo de interesse, apresentando itens do interesse do utilizador o mais cedo possível. Para além disso, ajudam os utilizadores a encontrarem itens de interesse que de outra maneira não conseguiriam. Na tentativa de conhecer os interesses de cada utilizador são usados determinados conjuntos de dados, tais como: registo dos itens que cada utilizador comprou, explicitamente classificou e visualizou, para além de se usar informação relacionada com dados demográficos, artistas favoritos ou assuntos de interesse.

No ambiente competitivo existente no comércio electrónico é necessário ter em conta certos aspectos que dificultam o processo de recomendação, tais como: os utilizadores de algum

tempo podem possuir excesso de informação acerca de si (milhares de compras ou classificações), novos utilizadores possuem informação limitada acerca de si, a informação disponível de um utilizador está constantemente a ser alterada, e muitas aplicações exigem o retorno dos resultados em tempo real enquanto produzem recomendações de alta qualidade [LindenSmith, 2003].

3.2.1 Algoritmos de Recomendação

No processo de recomendação são normalmente usadas três abordagens/algoritmos: *Traditional Collaborative Filtering*, *Cluster Models* e *Search-Based Methods*.

Traditional Collaborative Filtering

Neste algoritmo cada utilizador é representado através de um vector de dimensão N , onde N é o número de itens distintos relacionados com o utilizador em questão. Os elementos desse vector são positivos para os itens comprados ou classificados positivamente ou negativos se foram classificados negativamente. Este algoritmo pretende dar maior relevância a itens menos conhecidos de forma a compensar os itens mais vendidos. A abordagem seguida para tal efeito consiste na multiplicação dos componentes do vector pela frequência inversa, isto é, pelo inverso do número de utilizadores que compraram ou classificaram um determinado item. As recomendações realizadas por este algoritmo têm ainda por base a similaridade existente entre utilizadores, que pode ser determinada através do cálculo do cosseno entre os ângulos dos vectores correspondentes a cada utilizador. A partir do conjunto de itens existente entre os utilizadores similares podem ser adoptados vários métodos com o objectivo de seleccionar as recomendações. Um método normalmente utilizado consiste na elaboração de um *ranking* de itens de acordo com a quantidade de utilizadores similares que compraram cada item. A aplicação deste algoritmo é computacionalmente cara, principalmente quando aplicado a uma elevada quantidade de dados (depende do número de utilizadores e do número de itens). Embora possam ser aplicadas algumas técnicas de redução da dimensão dos dados (escolhendo aleatoriamente uma amostra de utilizadores ou descartando utilizadores com poucas compras ou até itens muito populares ou muito impopulares), estas afectam negativamente a qualidade da recomendação [LindenSmith, 2003; Krishnakumar, 2007].

Cluster Models

A ideia deste algoritmo é estruturar a base de utilizadores em segmentos e associar cada utilizador ao segmento com características mais similares a si (agrupar os utilizadores mais similares). Estes segmentos são tipicamente criados através do uso de algoritmos de aprendizagem não supervisionada ou *clustering algorithms*. A aplicação destes algoritmos sobre grande quantidade de dados pode ser complicada, fazendo com que se utilize várias formas gananciosas destes algoritmos ou técnicas de redução de dimensão do conjunto de dados. Tipicamente estes algoritmos começam com um conjunto inicial de segmentos que contêm um utilizador aleatoriamente seleccionado. Depois de identificados os segmentos, é calculada a

similaridade de cada utilizador com os vectores que sumarizam cada segmento, associando-se o utilizador ao segmento com a mais forte similaridade com as suas características.

Este algoritmo tem maior performance do que o algoritmo anterior porque cada utilizador é comparado com um número controlado de segmentos e não com todos os utilizadores. No entanto, a qualidade das recomendações é baixa porque considera todos os utilizadores de um segmento similares fazendo as mesmas recomendações por segmento [LindenSmith, 2003].

Search-Based (or Content-Based) Methods

Neste algoritmo, é usado a informação relativa aos itens que um determinado utilizador adquiriu ou classificou, por forma a realizar uma pesquisa por todos os outros itens relacionados (isto é, que contenham o mesmo autor, artista ou director, ou até palavras-chave ou tópicos similares). Este algoritmo resulta bem para utilizadores com poucas aquisições e classificações, já que se torna complicado criar uma *query* baseada em milhares de itens. Neste caso, o algoritmo usa apenas uma parte dos dados, provocando a redução da qualidade da recomendação. Em todos os casos, a qualidade da recomendação é relativamente fraca, já que para além das recomendações serem muito gerais ou muito limitadas, não permite ajudar o utilizador a encontrar novos itens com características diferentes, relevantes e interessantes [LindenSmith, 2003; Krishnakumar, 2007].

3.2.2 Amazon.com

Na área de venda online de produtos, surge um sistema bastante conhecido mundialmente - a Amazon.com.



The screenshot shows the Amazon.com interface with the 'Recommended for You' section. It features two book recommendations:

- Windows NT Server 4.0 Administrator's Bible** by Robert Cowart. Price: \$49.99. Average customer review: 4.5 stars. Recommended because you purchased or rated: [Mastering Windows Nt Server 4.](#)
- Windows NT Server 4 for Dummies** by Ed Tittel. Price: \$24.99. Average customer review: 4.5 stars. Recommended because you purchased or rated: [Mastering Windows Nt Server 4.](#)

Figura 3.1: Sistema de recomendações na Amazon.com

A Amazon.com é um exemplo de sucesso da aplicação deste tipo de sistemas, uma vez que utiliza um algoritmo de recomendação chamado *Item-to-Item Collaborative Filtering* para realizar a recomendação de produtos aos utilizadores com base na monitoração de acções passadas. Assim, a loja online é personalizada e ajustada a cada utilizador e está constantemente a ser modificada com base nas preferências de cada um. Para este efeito a Amazon.com utiliza um conjunto de dados para construir o modelo do utilizador, tais como: registo de compras, anotação dos produtos que o utilizador esteve a visualizar e histórico de navegação. Para além disto, a Amazon.com disponibiliza um sistema de classificação/*ranking* que permite saber as preferências dos utilizadores e assim melhorar cada vez mais o modelo desenvolvido [LindenSmith, 2003].

O algoritmo referido foi construído de propósito para satisfazer as necessidades da Amazon.com, já que os algoritmos existentes não podiam efectuar uma recomendação de qualidade em tempo real para a imensidade de utilizadores e itens que a Amazon.com possui. Este algoritmo faz corresponder itens similares aos que um determinado utilizador comprou ou classificou, combinando depois esses itens numa lista de recomendação. De forma a encontrar a melhor correspondência para um determinado item, o algoritmo constrói uma tabela com os itens similares a partir da identificação dos itens que os utilizadores tendem a comprar juntamente. Esta tabela é construída *offline* sendo um dos pontos chave para o desempenho demonstrado na aplicação deste algoritmo, mesmo sobre grandes quantidades de dados. O grau de similaridade entre itens pode ser obtido a partir do cálculo do cosseno acima descrito na apresentação do algoritmo *Traditional Collaborative Filtering* com a diferença que os vectores correspondem a itens em vez de utilizadores e a dimensão de cada item corresponde aos utilizadores a si associados. O grau de similaridade calculado é necessário para a aplicação do seguinte método [LindenSmith, 2003].

Para cada item existente (Item1), verificar quem foram os utilizadores que compraram esse item.

Para cada utilizador que comprou Item1, verificar quais são os outros itens comprados por este utilizador.

Para cada item (Item2) comprado por este utilizador registar que este comprou Item1 e Item2.

Para cada Item2 calcular o grau de similaridade com Item1.

Assim, a partir da tabela de itens similares efectua-se a recomendação dos itens mais populares ou correlacionados com os que o utilizador comprou ou classificou. Este cálculo é bastante rápido dependendo apenas do número de itens associados ao utilizador (componente online). O algoritmo possui elevado desempenho mesmo que um utilizador possua apenas dois ou três itens associados a si.

Customers who viewed this item also viewed

Product	Rating	Price
JL421 Badonkadonk Land Cruiser/Tank	★★★★☆ (151)	\$19,999.95
Tuscan Whole Milk, 1 Gallon, 128 fl. oz	★★★★☆ (946)	\$3.99
Elk Carcass	★★★★☆ (4)	\$1,225.00
Fresh Rabbit - Farm Raised All Organic Deboned Filet	★★★★☆ (1)	\$37.00
Relaxman Relaxation Capsule	★★★★☆ (14)	\$39,995.00

► Explore similar items: [Gourmet Food](#) (10) [Sports & Outdoors](#) (4) [Books](#) (4) [Apparel](#) (2) [Toys](#) (2)

Figura 3.2: Outra forma de recomendar itens pela Amazon.com

Como já acima supracitado, este algoritmo foi criado com o objectivo de fornecer grande qualidade na recomendação à Amazon.com que possui mais de 29 milhões de utilizadores e alguns milhões de itens. Os algoritmos existentes tinham sido avaliados sobre pequenas quantidades de dados, como é o caso do MovieLens (<http://movielens.umn.edu>), que em 2003 possuía cerca de 35 mil utilizadores e 3 mil itens [LindenSmith, 2003].

3.2.3 Recoo – Sistema de Recomendação para Youtube RSS Feeds

Muitos *websites* disponibilizam RSS (Really Simple Syndication) *feeds* para publicar as actualizações mais recentes do seu conteúdo. Qualquer tipo de informação que possa ser discriminada em itens discretos pode ser sindicada via RSS (alterações recentes a uma página, histórico de revisões de um livro, os últimos videos carregados para um *site* são exemplos de informação que pode ser disponibilizada via RSS). No entanto, a quantidade de RSS *feeds* e a quantidade de informação transmitida por esse canal tem vindo a crescer demasiadamente depressa, dificultando os utilizadores no processo de filtragem de informação de interesse. Uma solução para resolver este problema consiste na utilização de sistemas de recomendação, uma vez que estes oferecem uma forma inteligente e personalizada de procurar e apresentar nova informação [Krishnakumar, 2007].

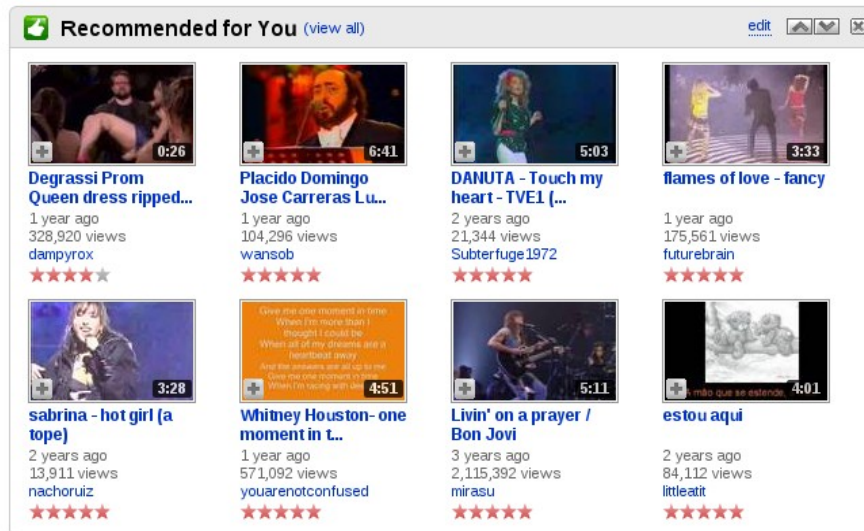


Figura 3.3: Recomendações de vídeos no Youtube

Youtube é um site Web bastante conhecido mundialmente que disponibiliza conteúdo multimédia ao seus utilizadores, já que permite fazer *upload*, visualizar e partilhar videoclipes. O Youtube possui vários RSS *feeds* que os utilizadores podem subscrever e que são dedicados a várias categorias de vídeo, tais como: carregados recentemente, mais vistos, favoritos, melhor classificados, mais discutidos, etc., bem como *feeds* personalizados para utilizadores e *tags*.

Recoo é um sistema de recomendação de *links* para videos a partir da informação provinda das RSS feeds do Youtube e no perfil de cada utilizador. Um utilizador que use o sistema pela primeira vez pode fornecer determinadas *tags* ou palavras-chave que descrevam os tipos de vídeos que esteja interessado em ver. No entanto, com base na utilização que um utilizador faz do sistema é criado um perfil de utilizador. Neste processo, é usada informação explícita, como pedir ao utilizador que classifique um vídeo ou que ordene uma colecção de vídeos por preferência, e implícita, como observar e registar os vídeos que um utilizador visualiza online, analisar o número de vezes que um utilizador visualiza um determinado vídeo/utilizador e procurar por semelhanças entre utilizadores. A partir da informação conhecida sobre um utilizador é efectuada uma comparação com informação similar de outros utilizadores e calculada a lista de recomendações de vídeos. O sistema usa uma abordagem baseada na predição das classificações que supostamente um utilizador dava aos outros vídeos que ainda não visualizou e conforme essa classificação, é efectuada uma ordenação de vídeos que o utilizador possa vir a gostar. Com base no que foi acima supracitado pode-se concluir que o sistema usa uma espécie de abordagem híbrida dos algoritmos *Collaborative Filtering* e *Content-Based Methods* para a realização das recomendações [Krishnakumar, 2007].

3.2.4 Recomendação de *Tags* em Folksonomias

Uma folksonomia é um sistema de classificação derivado da prática de criar e gerir colaborativamente *tags* para anotar e categorizar conteúdo. Esta prática é também conhecida por *collaborative tagging*, *social classification*, *social indexing* ou *social tagging*. Folksonomia é uma analogia à taxonomia, que inclui o prefixo *folks* (em português significa pessoas), de forma a referir-se ao seu carácter social.

A ideia consiste em através de *tags*, os utilizadores poderem aceder ou partilhar informação relacionada, bem como visualizar as *tags* dos outros utilizadores e identificar o grau de popularidade de cada *tag* no sistema.

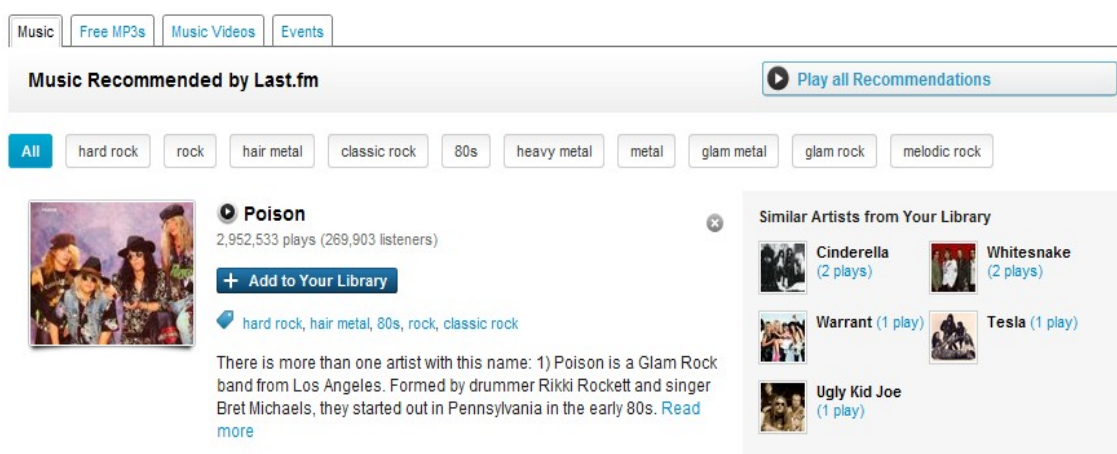


Figura 3.4: Sistema de recomendação no Last.fm

Vários exemplos de sucesso da *Web 2.0* usam folksonomia, como são os casos do Del.icio.us (partilha de *bookmarks*), Flickr (partilha de fotos), Last.fm (partilha de hábitos no que diz respeito a ouvir música) e o Youtube que já foi mencionado acima. Outros sistemas como o CiteULike e Connotea (partilha de referências bibliográficas) ou BibSonomy (partilha de *bookmarks* e publicações literárias) também recorrem a folksonomias. Muitos destes sistemas incluem mecanismos de recomendação de *tags*, facilitando o processo de pesquisa de *tags* e recursos relevantes e consolidando o vocabulário usado nas *tags* pelos utilizadores. O Del.icio.us para além de disponibilizar a recomendação de *tags* aos seus utilizadores, também inclui recomendação dos recursos a que as *tags* estão associadas (no caso do Del.icio.us esses recursos são as *bookmarks*) [JaschkeMarinho, 2007].

No sentido de avaliar a eficiência de alguns algoritmos de recomendação de *tags*, [JaschkeMarinho, 2007] usaram *datasets* reais respectivos aos três seguintes sistemas: Del.icio.us, Last.fm e BibSonomy. Os algoritmos de recomendação que foram o maior alvo da avaliação foram [JaschkeMarinho, 2007]:

- Uma adaptação do *Traditional Collaborative Filtering* a folksonomias.

- Devido à relação ternária existente entre os utilizadores, *tags* e recursos nas folksonomias, o *Traditional Collaborative Filtering* não pode ser aplicado directamente, a não ser que a relação ternária seja reduzida a uma dimensão espacial mais baixa. Consideram-se duas projecções de 2 dimensões representadas por matrizes, as quais permitem preservar a informação dos utilizadores e recomendar com base na ocorrência ou não ocorrência de recursos ou *tags*. Assim:
 - A primeira matriz faz corresponder os pares utilizador-recurso (utilizador que possui aquele recurso). Os pares assumem valor 1 na matriz se existir alguma *tag* associada aos dois elementos do par, e 0 no caso contrário.
 - A segunda matriz faz corresponder os pares utilizador-*tag* (utilizador que possui aquela *tag*). Os pares assumem valor 1 na matriz se existir algum recurso associado aos dois elementos do par, e 0 no caso contrário.
- FolkRank – abordagem baseada em grafos.
 - Adaptação do conhecido algoritmo PageRank, que reflecte a ideia de que uma página *Web* é importante se existir muitas páginas interligadas a si e se essas páginas forem também importantes.
 - Então, a ideia consiste em dar importância a um recurso se este for referenciado com *tags* importantes por utilizadores também importantes. A mesma ideia é aplicada a cada um dos dois outros elementos da relação ternária (utilizadores e *tags*). Assim, ficamos com um grafo onde os vértices possuem diferentes pesos e se reforçam mutuamente distribuindo esses pesos.

No final das experimentações, [JaschkeMarinho, 2007] concluíram que o FolkRank superava os outros métodos testados (pelo menos nos *datasets* utilizados, o quais possuíam diferentes características).

3.3 Educação – Ambiente AdaptWeb

O estudo sobre sistemas adaptativos conduziu a diferentes áreas de aplicação, entre as quais, a educação. Nesta área, mais concretamente no ambiente AdaptWeb, a ideia era o desenvolvimento de um sistema de aprendizagem para alunos, onde se pretendia que o sistema se adapta-se ao nível de experiência dos alunos na navegação pelas páginas e ao nível de conhecimento do aluno. Assim, conforme o estilo de navegação, objectivos e conhecimentos de um utilizador, o sistema fornece apoio aos seus utilizadores, na medida em que utiliza um conjunto de métodos e técnicas de suporte à navegação e apresentação adaptativa, tais como: limitar o espaço de navegação, sugerir os links mais relevantes, ocultar links ou remover links, ou até provendo comentários/observações adaptados aos links visíveis. Por exemplo, se o sistema detecta que um aluno possui pouco conhecimento numa determinada área, apenas vai

permitir acesso a informação básica ou até retirar do campo de visão os links cujo conteúdo seja demasiado complexo para utilizador em questão (torna-se complicado compreender certos conceitos sem primeiro adquirir conhecimentos mais básicos) [Gasparini, 2003].

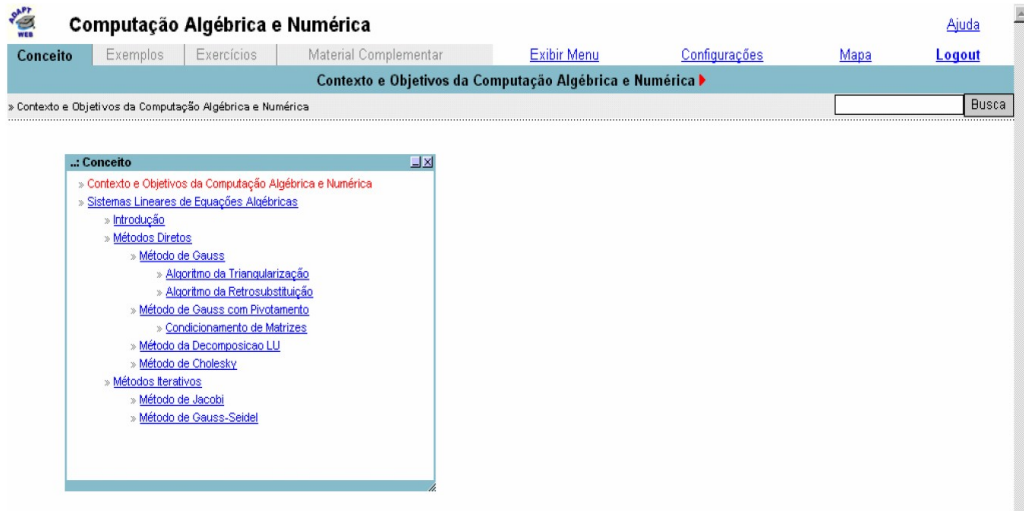


Figura 3.5: Interface da página de uma disciplina no ambiente AdaptWeb

3.4 SeAN – Server for Adaptive News

SeAn é um sistema adaptativo para o acesso personalizado a notícias na *Web*. O objectivo do sistema passa por seleccionar as secções e as notícias que são mais relevantes para o utilizador, por costumizar a nível detalhado cada notícia às características do utilizador e por seleccionar a publicidade mais apropriada para cada página [ArdissonoConsole, 2001].

SeAN foi implementado em Java recorrendo a um conjunto de agentes cooperativos (sistema multi-agente), os quais estabelecem comunicação entre si usando Java RMI (*Remote Method Invocation*). Quanto à arquitectura, esta é de três camadas, a qual foi herdada de um sistema adaptativo para o comércio electrónico (SETA). A primeira camada corresponde a quando um utilizador acede ao servidor usando um *browser*; a segunda camada corresponde quando o *browser* interage com o conjunto de agentes, os quais são responsáveis pelas actividades relacionadas com personalização; e a terceira camada corresponde às bases de dados. As bases de dados foram desenvolvidas usando o MS Access e são acedidas usando JDBC, as quais servem para armazenar o modelo do utilizador, informação relacionada com a publicidade e com as notícias (os itens são classificados numa hierarquia de tópicos e são caracterizados por um conjunto de atributos) [ArdissonoConsole, 2001].

4 Concepção da solução

Este capítulo é dedicado à documentação de todo o processo de concepção das ideias necessárias para a implementação das funcionalidades de adaptação automática no serviço de *webmail* inteligente da Portugalmail. Uma análise de requisitos também é efectuada com base nas funcionalidades seleccionadas.

Numa fase inicial, a concepção das ideias estava direccionada para a individualidade, isto é, o pensamento estava focado em detectar padrões de comportamento para o utilizador individual e efectuar as adaptações necessárias para a satisfação do mesmo. Assim, surgiram muitas ideias com carácter geral, as quais abrangiam utilizadores de vários tipos e personalidades. Numa fase posterior, verificou-se que também fazia sentido pensar em tipos de utilizadores, e até mais que isso, em categorias onde os vários tipos se inseriam. Usando esta abordagem foi elaborado um *mind map* com as ideias que surgiram como resultado de toda a fase de análise e pesquisa. Com o objectivo de estruturar e gerar mais ideias para a adaptação a cada tipo de utilizador foram realizadas várias sessões de *brainstorming* para complementar o trabalho realizado.

4.1 Categorias e Tipos de Utilizadores

Uma mesma pessoa pode assumir vários tipos de utilizador consoante a ocasião (como se trata-se de diferentes personalidades). Um exemplo: um utilizador quando está no trabalho pode querer ler apenas os emails mais importantes ou que estejam relacionados com o seu trabalho, deixando os outros para quando chegar a casa ou tiver em lazer.

Os tipos de utilizadores identificados encontram-se estruturados em várias categorias e são apresentados seguidamente.

4.1.1 Organização

Pastas - Utilizadores que têm por hábito estruturar os emails por pastas/grupos.

Tarefas - Utilizadores que têm por hábito organizar as suas tarefas por importância ou assunto e que utilizam frequentemente funcionalidades relacionadas com o planeamento de tarefas como Task List, Calendar ou lembretes.

Sem organização - Utilizadores que não possuem nenhum tipo de organização.

4.1.2 Uso do sistema

Email - Utilizadores que usam o sistema mais pela aplicação de email em si.

Aplicações - Utilizadores que usam o sistema não tanto pela aplicação de email e mais pelas outras aplicações. Exemplos: Calendário, Contactos, RSS, Ficheiros, Aplicações.

4.1.3 Frequência de acesso

Online - Utilizadores que recorrem ao serviço de *webmail* constantemente.

Daily - Utilizadores que usam o serviço de *webmail* diariamente.

Often - Utilizadores que usam o serviço de *webmail* duas a três vezes por semana.

Sporadic - Utilizadores que utilizam o serviço de *webmail* esporadicamente, isto é, em intervalos de tempo longos (vários dias ou até semanas).

4.1.4 Tipos de Emails

Brincadeira/Lazer - Utilizadores que usam o serviço de *webmail* para fins lúdicos e de lazer. Como forma de identificar estes utilizadores pode-se proceder à verificação dos emails lidos: se estes possuem muitos ficheiros do tipo .ppt ou .wmv em anexo, se existem muitos emails lidos frutos de vários reencaminhamentos ou associados a redes sociais.

Professional - Utilizadores que utilizam o serviço de *webmail* para fins profissionais. Neste tipo, os emails recebidos podem possuir maior urgência de resposta.

Pessoal/Familiar - Utilizadores que pretendem usar o email como forma de comunicação com amigos e familiares, substituindo o serviço de correio tradicional.

Newsletters - Utilizadores que recebem muitos emails correspondentes a *newsletters*.

Mailing Lists - Utilizadores que têm por hábito trocar material com grupos de contactos.

Pornógrafos - Existência de grande quantidade de emails, cujos assuntos e conteúdos estão relacionados com pornografia.

4.1.5 Volume de emails

Elevado - Utilizadores que possuem uma grande quantidade de emails nas suas caixas de entrada ou saída.

Reduzido - Reduzida quantidade de emails nas caixas de entrada ou saída de um utilizador deste tipo.

Notas:

De notar que existe o tipo de utilizador Médio dentro desta categoria que corresponde a um utilizador com um comportamento “normal”, isto é, que não esteja inserido nem no tipo Elevado ou Reduzido (encontra-se entre os dois extremos). No entanto, a sua referência não é considerada relevante. O mesmo acontece nas próximas categorias.

4.1.6 Volume de contactos que comunica (VCC)

Elevado - Utilizadores que enviam ou recebem emails para/de uma grande quantidade de contactos.

Reduzido - Utilizadores que apesar de poderem possuir uma grande quantidade de contactos só recebem ou enviam emails de/para uma pequena parte.

4.1.7 Volume de contactos da lista de contactos (VCL)

Elevado - Utilizadores com uma elevada quantidade de contactos.

Reduzido - Utilizadores com uma reduzida quantidade de contactos.

4.1.8 Rácio envio/recepção

Leitores - Utilizadores que utilizam o serviço de email sobretudo para fins de leitura.

Escritores - A funcionalidade de envio de emails é bastante requisitada por este tipo de utilizadores.

4.1.9 Modo de navegação

Browsers - Utilizadores com algum tempo livre, que geralmente procuram pelos conteúdos de um *site* usando os *links* disponíveis. Estes utilizadores não procuram por nada em específico, mas estão abertos a algo que possa captar o seu interesse.

Searchers - Utilizadores com um objectivo específico e que estão à procura de algo em particular. Normalmente dirigem-se directamente para um formulário de pesquisa, esperando que o mecanismo de pesquisa os ajude a encontrar exactamente o que que eles procuram com o mínimo desperdício de tempo e esforço.

Notas:

Um sistema de alto nível deve fornecer os meios necessários para que possa cativar utilizadores de ambos os tipos.

4.1.10 Experiência Web

Elevada - Utilizadores com grande experiência na utilização de serviços de *webmail*. Estes utilizadores cometem tipicamente poucos erros na navegação e utilização do sistema.

Reduzida - Utilizadores com reduzida experiência na utilização de serviços de *webmail*. Estes utilizadores tendem a cometer erros excessivos na navegação e utilização do sistema (exemplos: recurso exagerado da funcionalidade de *undo*, esquecimento de anexo que dá origem ao envio de mesma mensagem mais que uma vez, uso frequente da funcionalidade que permite cancelar uma operação).

4.1.11 Outros

Viajantes - Utilizadores que por norma acedem ao serviço a partir de locais diferentes (longa distância).

Temas/Interesses - Utilizadores que têm forte ligação com determinados temas/interesses. Estes utilizadores tendem a usar aplicações e a receber/enviar emails relacionados com algum(ns) tema(s) específico(s).

4.2 Brainstorming

A identificação e definição das categorias e tipo de utilizadores de serviços de *webmail* surgiu como resultado da análise realizada juntamente com alguns elementos da equipa de desenvolvimento da Portugalmail. Várias ideias foram pensadas para a detecção de informação a usar e que ajustes a efectuar para cada tipo de utilizador. No entanto, estas ideias mantiveram-se ocultas na realização das sessões de *brainstorming* para não influenciar os intervenientes e dar asas à imaginação. A base do *brainstorming* foram as categorias e tipos de utilizadores já identificados, associando-se uma imagem a cada tipo de utilizador. As ideias originadas ao longo das sessões foram adicionadas devidamente às pré-existentes, caso ainda não tivessem surgido.

O *brainstorming* foi realizado numa sala de reuniões nas instalações da Portugalmail e foi estruturado em duas fases. A primeira fase foi relativamente curta e teve como objectivo discutir o método que devia ser seguido para identificar os vários tipos de utilizador existentes chegando à conclusão que o método até aí seguido tinha sido o mais correcto. Na figura 4.1 pode-se observar o quadro referente às categorias e tipos de utilizadores.

ORGANIZAÇÃO	USO DO SISTEMA	FREQUÊNCIA DE ACESSO	TIPO DE MAIL	VOLUME DE MAILS	VOLUME DE CONTACTOS QUE ENVIAM (VCE)	VOLUME DE CONTACTOS DA LISTA (VCL)	RÁCIO ENVIOS/RECEBOS	MODO DE NAVEGAÇÃO	OUTROS	EXEMPLOS WEB
PASTAS	# E-MAIL	ONLINE	BRANCA/VERMELHA	REDUZIDO	REDUZIDO	REDUZIDO	LECTORES	BROWSERS	VIAJANTES	REDUZIDA
TAREFAS	CALENDRÁRIO	DAILY	PROFISSIONAL	MÉDIO	MÉDIO	MÉDIO	ESCRITORES	SEARCHERS	INTERESSES	NORMAL
SER ORGANIZADA	CONTACTOS	OFTEN	PERSONAL/FAMILY	ELEVADO	ELEVADO	ELEVADO	NORMAL	NORMAL		ELEVADA
	RSS	SPORADIC	MAILING LIST							
	PICTURES		NEWS LETTERS							
	APLICACÕES		PERSONÁGIOS							

Figura 4.1: Todos os tipos de utilizadores identificados e abordados

Na outra fase era pretendido gerar ideias de adaptação para cada tipo de utilizador, utilizando a técnica de provocação por imagem e uma variação da técnica 6-3-5. A técnica de provocação por imagem consiste na tentativa de suscitar ideias e apelar à imaginação dos intervenientes através da simples observação de uma imagem devidamente pensada. A técnica 6-3-5 baseia-se na seguinte ideia: seis intervenientes devem gerar pelo menos três ideias em cinco minutos. Fala-se em variação desta técnica, uma vez que nem sempre estavam presentes seis intervenientes e nem sempre eram cumpridos os cinco minutos devido às limitações de tempo comparando com a quantidade elevada de tipos de utilizadores

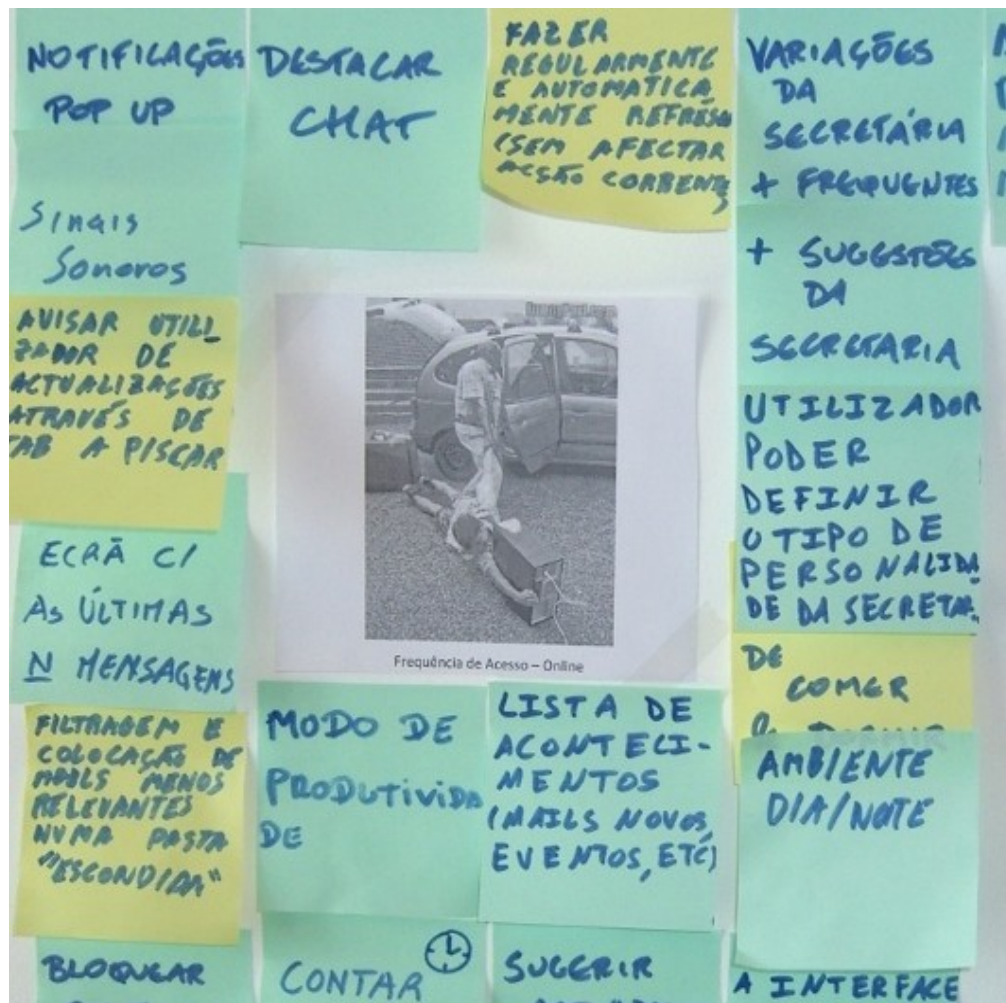


Figura 4.2: Exemplo das ideias geradas para um utilizador com uma frequência de acesso ao serviço elevada

Após toda a discussão gerada nas sessões do *brainstorming* chegou a altura de organizar as ideias. A figura 4.3 é fruto do método usado durante as sessões (as ideias encontram-se escritas em *post-its*) e a figura 4.4 é referente ao uso de um programa de criação de mapas mentais. O programa é o FreeMind que é apresentado mais à frente no capítulo das tecnologias. Devido à enorme quantidade de ideias torna-se impossível apresentar uma figura devidamente legível com todas essas ideias. Assim, optou-se por mostrar os tipos de utilizadores discutidos e, como exemplo, as ideias referentes a apenas um desses tipos (neste caso “Viajantes”).

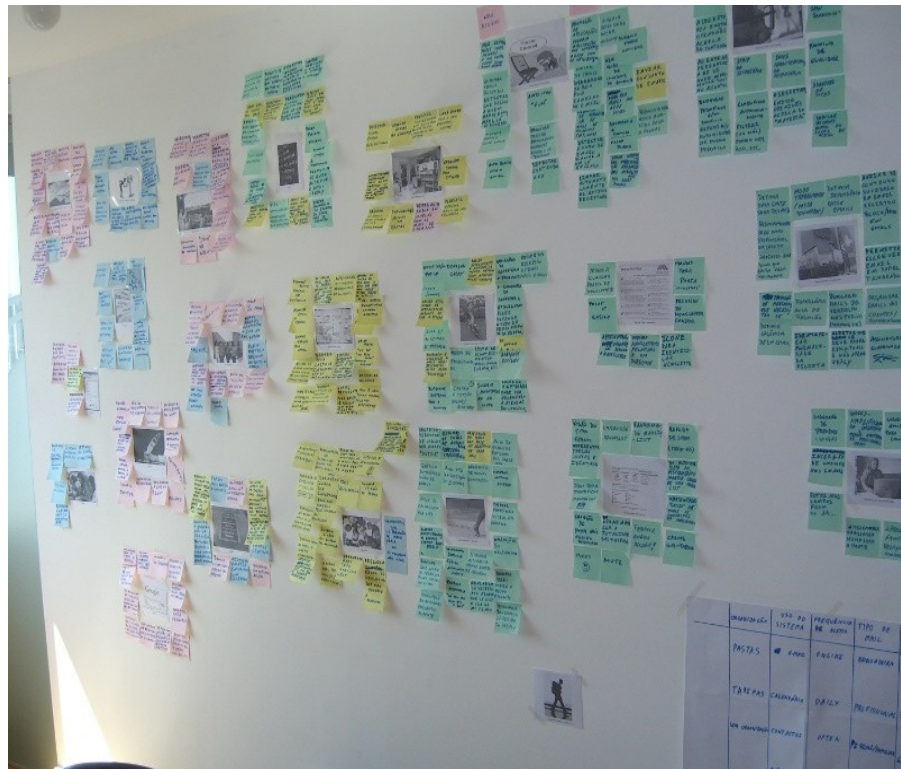


Figura 4.3: Visão geral do resultado do *brainstorming*

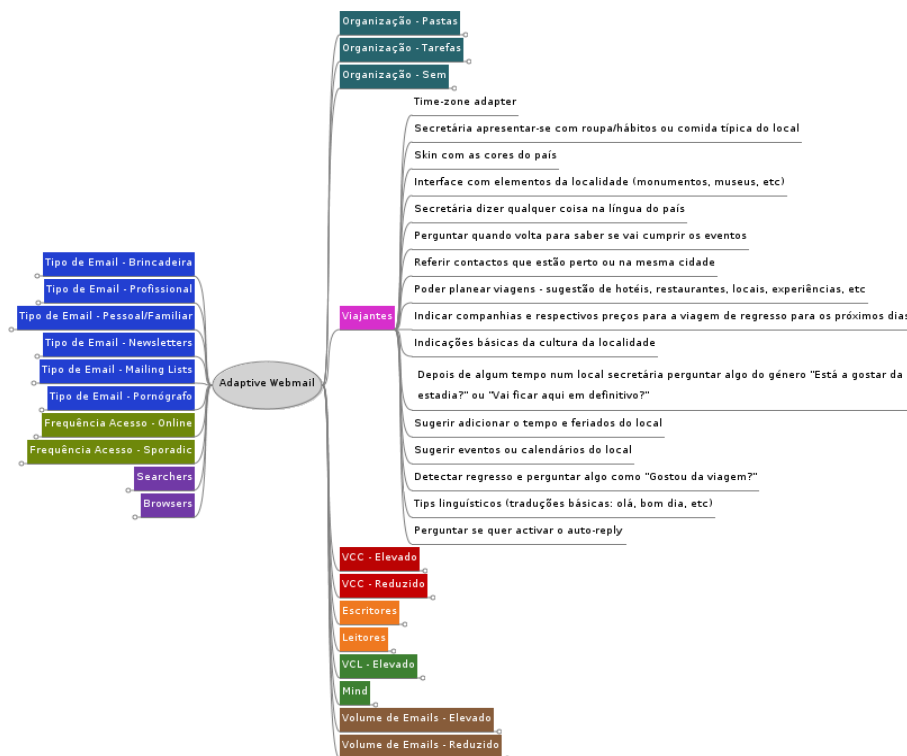


Figura 4.4: Mapa mental bastante reduzido das ideias geradas

4.3 Seleção das Ideias

A fase de selecção das ideias mais importantes seguiu-se à fase de organização. Esta fase foi realizada conjuntamente com os responsáveis pela equipa técnica da Portugalmail de forma a que as ideias para funcionalidades a desenvolver constituíssem uma mais valia para o serviço de *webmail* prestado pela empresa. Assim, chegou-se à conclusão que a implementação das seguintes três funcionalidades constituía uma boa meta para cumprir durante o tempo de projecto: criação de uma interface especial para redes sociais, detecção do horário de trabalho e sugestão de grupos de contactos na funcionalidade de *autocomplete*.

Seguidamente serão apresentadas em maior detalhe as três funcionalidades desenvolvidas.

4.3.1 Interface Especial para Redes Sociais

Muitos utilizadores têm por hábito receber muitos emails de redes sociais como o Hi5, Facebook, MySpace, etc.. O conteúdo desses emails em grande parte dos casos não acrescentam muito mais do que a informação contida no assunto da mensagem. Essa informação é muitas vezes considerada pouco satisfatória já que os utilizadores na maioria das vezes interpretam o email como “olha, tenho algo novo na rede social X”, necessitando de qualquer das maneiras de aceder à página correspondente à rede social em questão (para saberem mais ou procederem a alguma acção). Este tipo de emails são muitas vezes não lidos ou eliminados. Outro aspecto a considerar consiste no facto destes emails poderem encher a caixa de entrada prejudicando tanto a procura de emails realmente importantes, como a navegação pelas páginas da caixa de entrada. Assim, a proposta que agradaria a muitos utilizadores consiste no desenvolvimento de uma interface especial para redes sociais onde a informação contida nos emails seria capturada e apresentada de uma forma mais apelativa nessa interface. Os utilizadores desta maneira não deixavam de ser informados acerca das actualizações nas redes sociais às quais se encontram inscritos, e caso pretendessem saber mais podiam recorrer à interface criada. A interface ainda permitia um acesso bastante mais rápido à informação correspondente a vários emails de várias redes sociais (a informação de vários emails é apresentada de uma só vez na interface).

4.3.2 Detecção do Horário de Trabalho

A leitura de emails de brincadeira ou de qualquer outro assunto fora do contexto de trabalho é um hábito de muitos utilizadores para efeitos de lazer, tratamento de assuntos, comunicação entre amigos ou familiares, etc.. No entanto, devido ao stress do trabalho alguns utilizadores costumam ler um ou outro email com assunto fora do contexto de trabalho como forma de descontração ou necessidade. De modo a simplificar tanto a escrita como a leitura, os emails fora do contexto de trabalho serão referenciados daqui para a frente como emails de

brincadeira. Outros utilizadores são distraídos pela recepção ou presença de emails de brincadeira na sua caixa de entrada. É para este tipo de utilizadores que esta funcionalidade é dirigida, de forma a evitar o excesso de distração.

O objectivo desta funcionalidade reside na detecção do horário de trabalho nos vários IPs em que um utilizador acede ao serviço de *webmail*, através do registo e análise dos dias, horas e locais em que um determinado utilizador leu emails relativos a trabalho. Depois de detectado o horário de trabalho é então estudada a situação do utilizador, e caso se verifique, a adaptação é aplicada. A adaptação consiste no reencaminhamento dos emails de brincadeira para uma pasta “escondida” (ocultar emails de brincadeira) durante o horário de trabalho desse utilizador no respectivo IP. Desta forma o utilizador não se sente tão tentado a perder tempo na leitura de emails de brincadeira (normalmente o impulso é visual).

Importante será referir que esta funcionalidade resulta especialmente bem para utilizadores com horários e locais de trabalho fixos. Para utilizadores que possuem um horário flexível que acessem ao serviço em horas muito dispersas e/ou em locais de trabalho muito alternativos, a aplicação deste funcionalidade revela-se complicada.

4.3.3 Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*

Os serviços de *webmail* actuais possuem a funcionalidade de *autocomplete* (completar automaticamente) no preenchimento do destinatário de um email. No entanto, essa funcionalidade baseia-se apenas na sugestão de contactos de email singulares que fazem *match* com o que o utilizador escreveu, isto é, o que o utilizador escreveu no campo de destinatário deve fazer corresponder/coincidir com o início de um contacto de email ou com o início de um qualquer nome associado a um contacto de email da lista de contactos. Os contactos sugeridos encontram-se ordenados por ordem alfabética.

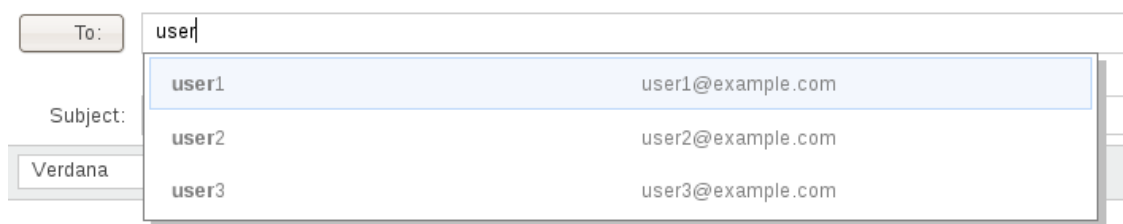


Figura 4.5: Exemplo da funcionalidade de autocomplete existente no preenchimento do campo de destinatário do Hotmail (os contactos de email foram criados exclusivamente para efeitos de exemplificação)

A proposta para melhorar a funcionalidade de *autocomplete* consiste nas seguintes ideias:

- Sugestão de contactos singulares, mas também de grupos de contactos com base no histórico do utilizador relativo ao envio de emails.
- Ordenação baseada primeiramente numa classificação dada aos contactos ou grupos de contactos a quem o utilizador já enviou emails e depois na ordenação alfabética dos contactos singulares. Em ambos os casos, a ordenação e os contactos que são sugeridos dependem do que já se encontra presente no campo referente ao destinatário.
- A classificação, correspondente a cada contacto ou grupo de contactos referida no ponto anterior, depende do IP em que o utilizador acede ao sistema, isto é, pretende-se dar prioridade ao IP em que o utilizador se encontra no momento, mas sem nunca esquecer o comportamento geral do utilizador. Esta prioridade é baseada no facto que os utilizadores assumem comportamentos diferentes dependendo do local em que se encontram. Assim, os contactos a quem um utilizador envia emails pode variar consoante o local, isto é, um utilizador pode só enviar emails para um contacto ou conjunto de contactos caso esteja a aceder ao sistema a partir do IP X ou Y. Contudo, existem aqueles contactos a quem se envia emails independentemente do IP em que se acede. A funcionalidade foi desenvolvida de forma a adaptar-se aos diversos tipos de utilizadores existentes: os que possuem destinatários variantes (consoante o local), os que possuem destinatários fixos (independentemente do local) e os que misturam as duas abordagens.

Mais à frente serão explicadas em pormenor tanto as bases que determinam a classificação e a priorização, como o funcionamento geral desta funcionalidade.

4.4 Requisitos Funcionais

4.4.1 Actores

O principal actor do sistema adaptativo é outro sistema – o novo serviço de *webmail* da Portugalmail. Isto verifica-se, uma vez que é este sistema que usufrui da maior parte das capacidades e funcionalidades oferecidas pelo sistema adaptativo (com recurso a uma API desenvolvida para esse efeito). Os utilizadores finais do serviço de *webmail* são também actores do sistema adaptativo, embora sejam muitas vezes considerados utilizadores indirectos, já que fazem uso das adaptações a partir do serviço de *webmail*.

4.4.2 Descrição dos Requisitos

RF-1: Ao ocorrer um evento que está associado a uma das funcionalidades adaptativas (exemplos de eventos: ler email, enviar email, receber email), o sistema responsável pela plataforma de *webmail* deve invocar os métodos disponibilizados pela API desenvolvida para o

sistema adaptativo. A invocação desses métodos é efectuada com o objectivo de executar toda a lógica de negócio inerente à respectiva funcionalidade.

RF-2: O sistema deve recomendar uma determinada funcionalidade, se se verificar que a aplicação desta pode ser benéfica para o utilizador.

RF-3: O sistema deve informar o utilizador quais foram as razões que levaram a recomendar uma funcionalidade.

RF-4: O sistema deve informar o utilizador que informação está a ser usada para efeitos de adaptação.

RF-5: O utilizador deve ter a possibilidade de aceitar ou rejeitar a recomendação de uma adaptação.

RF-6: O utilizador deve poder cancelar/eliminar uma adaptação que foi recomendada e aceite pelo utilizador.

RF-7: Na funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*”, a cada carácter inserido no campo referente ao(s) destinatário(s) de um email, deve ser apresentada uma lista de sugestões dos contactos ou grupos de contactos que fazem *match* com o que se encontra escrito.

RF-8: Na funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*”, a cada escolha de uma opção da lista mencionada no requisito anterior, o sistema deve preencher o campo de destinatário com os contactos presentes nessa opção e gerar nova lista de sugestões dos contactos ou grupos de contactos que fazem *match* com o que se encontra escrito.

RF-9: A lista mencionada nos requisitos anteriores deve estar ordenada dando prioridade aos contactos ou grupos de contactos a quem o utilizador mais emails enviou no passado recente e aos contactos ou grupos de contactos a quem foram enviados emails no IP de acesso.

RF-10: Os grupos de contactos que aparecem na lista de sugestões devem fazer-se acompanhar pela classificação atribuída nesse IP.

RF-11: O sistema deve ter capacidade de aprendizagem, adaptando o método usado na processo de tomada de decisão.

4.5 Requisitos Não Funcionais

Os requisitos não funcionais descrevem restrições ou atributos de qualidade não directamente relacionados com as funcionalidades oferecidas pelo sistema. Não são, porém, menos importantes, contribuindo decisivamente para o sucesso do produto e para a sua aceitação pelo utilizador. Apresentam-se de seguida os requisitos não funcionais identificados para o sistema adaptativo desenvolvido.

RNF-1

Nome: Eficiência

Descrição: Existem eventos que são responsáveis por accionar todo o processo de análise de informação e conseqüente tomada de decisão. Assim, pretende-se que os tempos de resposta às acções solicitadas pelo utilizador mantenham-se reduzidos.

RNF-2

Nome: Fiabilidade

Descrição: A fiabilidade diz respeito à consistência e robustez de um sistema, de modo a conseguir ganhar a confiança e garantir a satisfação necessária por parte do utilizador. Nesse sentido, pretende-se que a sugestão ou aplicação das adaptações sejam realizadas com um grau de certeza elevado e que o sistema seja desenvolvido com capacidade de aprendizagem, de forma a cometer cada vez menos erros.

RNF-3

Nome: Manutenção

Descrição: Pretende-se que o sistema seja implementado de uma forma que facilite posteriores modificações ou possíveis introduções de novas funcionalidades. Nesse sentido, toda a implementação do projecto segue uma estratégia comum para as suas funcionalidades, onde cada uma encontra-se organizada em etapas.

RNF-4

Nome: Modularidade

Descrição: As funcionalidades devem ser desenvolvidas em módulos ou etapas, de forma a permitir a reutilização das partes que as constituem.

RNF-5

Nome: Segurança

Descrição: Deve ser garantida a confidencialidade e segurança de toda a informação referente aos modelos dos utilizadores. O acesso e modificação dos dados devem ser controlados.

5 Proposta de Solução

Este capítulo tem como objectivo dar a conhecer a proposta de solução. Esta fundamenta-se principalmente na proposta de arquitectura do sistema adaptativo, isto é, descreve-se como se adaptou a abordagem modular proposta para a avaliação de sistemas adaptativos apresentada em [Paramythis, 2001], ao problema que consiste em oferecer adaptação automática aos utilizadores de um serviço de *webmail*. A proposta ainda se concentra nas fórmulas e métodos adoptados nos processos de tomada de decisão e avaliação automática do sucesso ou insucesso das adaptações efectuadas.

5.1 Proposta de Arquitectura

Esta secção pretende dar a conhecer toda a configuração da arquitectura do módulo correspondente ao sistema adaptativo descrevendo as suas etapas fundamentais (as quais encontram-se organizadas em sub-módulos).

5.1.1 Plataforma de *Webmail* da Portugalmail

Antes de prosseguir com a apresentação da arquitectura proposta para desenvolvimento do sistema adaptativo, será importante caracterizar o sistema no qual o sistema adaptativo vai ser integrado. Assim, o leitor possuirá as bases necessárias para uma melhor compreensão da arquitectura proposta.

A plataforma existente no *webmail* da Portugalmail foi desenvolvida sobre uma variação da famosa arquitectura LAMP, usada para o desenvolvimento de servidores e aplicações Web dinâmicas. Este acrónimo advém das seguintes características: sistema operativo Linux, servidor Web Apache, base de dados MySQL e linguagem de programação PHP/Perl/Python. A referida variação desta arquitectura refere-se à alteração do servidor de

base de dados (que passou de MySQL para PostgreSQL), criando assim o acrónimo LAPP. O último P corresponde à linguagem PHP onde se encontra a *framework* Horde. Esta é responsável por toda a comunicação entre as aplicações de mais alto nível e os sistemas de armazenamento de dados. Para além disso, esta *framework* possui como características: ser de licença livre, ser escrita em PHP e facilitar o desenvolvimento *Web*. A sua arquitectura é baseada em aplicações onde se destaca um cliente de *webmail* IMP e a sua versão AJAX (DIMP), permitindo o acesso a contas de email através de IMAP (Internet Message Access Protocol) ou POP3 (Post Office Protocol version 3). Para acesso às contas de email, a Portugalmail utiliza a versão DIMP. [AvilaCosta, 2009].

Existem ainda outras aplicações desenvolvidas pela Horde que a Portugalmail utiliza, como são o caso da aplicação Ingo (gestão dos filtros dos emails) e da Passwd (gestão das palavras-chave das aplicações). O gestor de contactos Turba disponibilizado pela Horde, também já foi usado pela Portugalmail nas suas versões mais antigas. Este foi substituído por outra aplicação (designada Sybil), a qual foi contruída totalmente de raíz, devido à necessidade de responder às novas exigências da nova plataforma. LDAP (Lightweight Directory Access Protocol) consiste noutra aplicação usada pela Portugalmail com o objectivo de controlar o processo de autenticação dos seus utilizadores, bem como guardar as suas preferências.

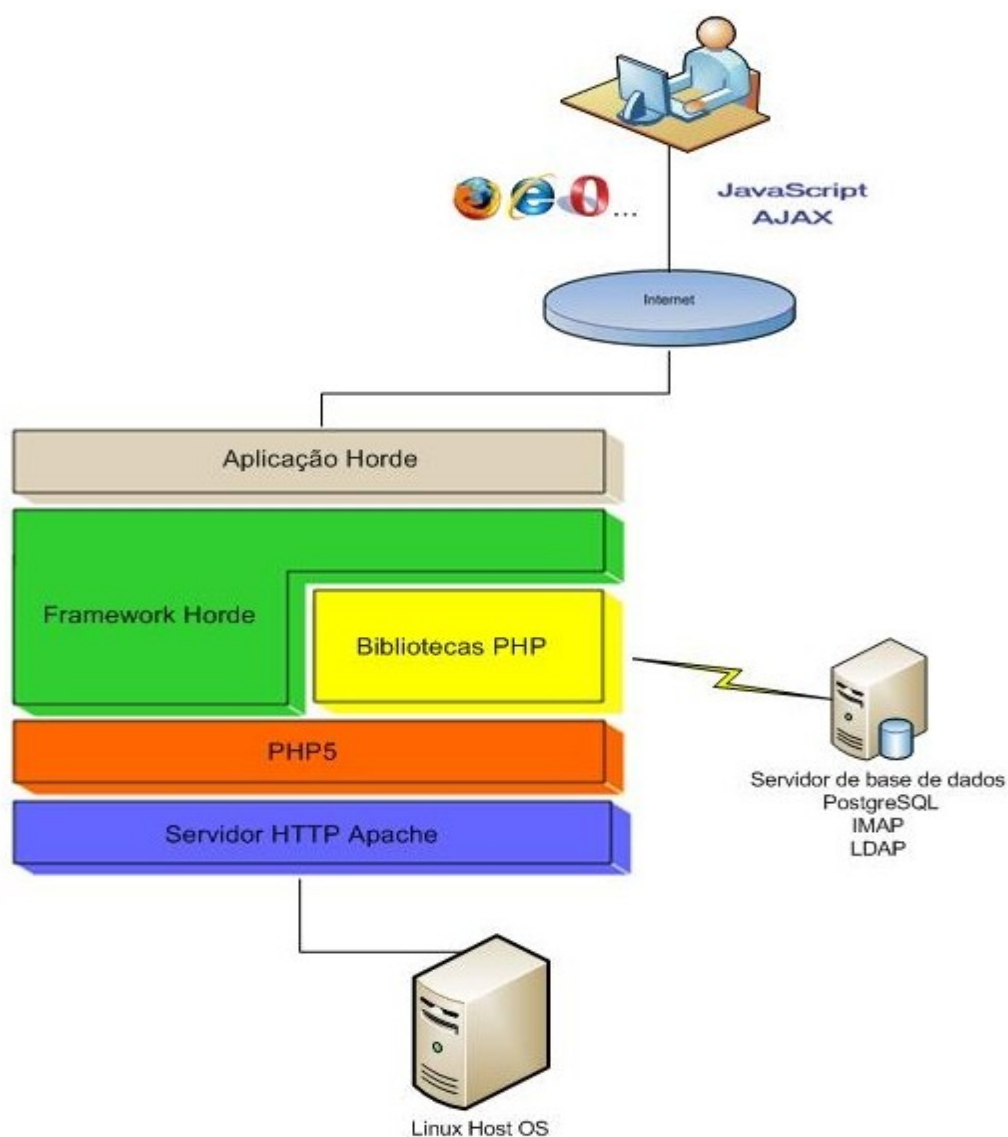


Figura 5.1: Arquitectura da Plataforma da Portugalmail

5.1.2 Módulo de Adaptação Automática

Em relação ao módulo de Adaptação Automática de *Webmail*, a abordagem seguida foi baseada na avaliação das etapas individuais envolvidas nos ciclos de adaptação de sistemas adaptativos. A avaliação destas etapas providencia a derivação de resultados mais pormenorizados, o que por sua vez proporciona um amplo *feedback* para todo o processo de desenho de sistemas adaptativos (o sistema sabe se as opções tomadas estão a surtir o efeito

desejado). O pretendido foi não tratar o problema de adaptação como um processo singular, mas como um problema que devia ser dividido nas suas partes fundamentais. Isto permite a reutilização dessas partes (tanto para outras funcionalidades, como para outros sistemas) e uma forma mais eficaz e simples de alterar certas partes da implementação, isto é, permite introduzir alterações numa parte do código sem com isso afectar as outras partes [Paramythis, 2001].

Como se poderá observar, o módulo adaptativo (Adaptação Automática de *Webmail*) encontra-se a par com a Aplicação Horde, que representa os módulos correspondentes a cada aplicação da plataforma de *webmail*. Isto deve-se ao sistema adaptativo relacionar-se com todas essas aplicações, já que o objectivo do sistema adaptativo é adaptar a interface ao utilizador, não só a nível do email, mas também nos contactos, calendário, entre outros.

As etapas fundamentais envolvidas dividem-se tanto pelo módulo adaptativo como pelos módulos correspondentes às aplicações de *webmail* da Portugalmail. Mais à frente são apresentadas cada uma destas etapas, onde será possível compreender o porquê desta divisão. Entretanto, apresenta-se graficamente a arquitectura desenvolvida, a qual permite ao leitor ter um primeiro contacto com cada umas das etapas fundamentais, demonstrando ainda os componentes importantes de arquitectura de alto nível.

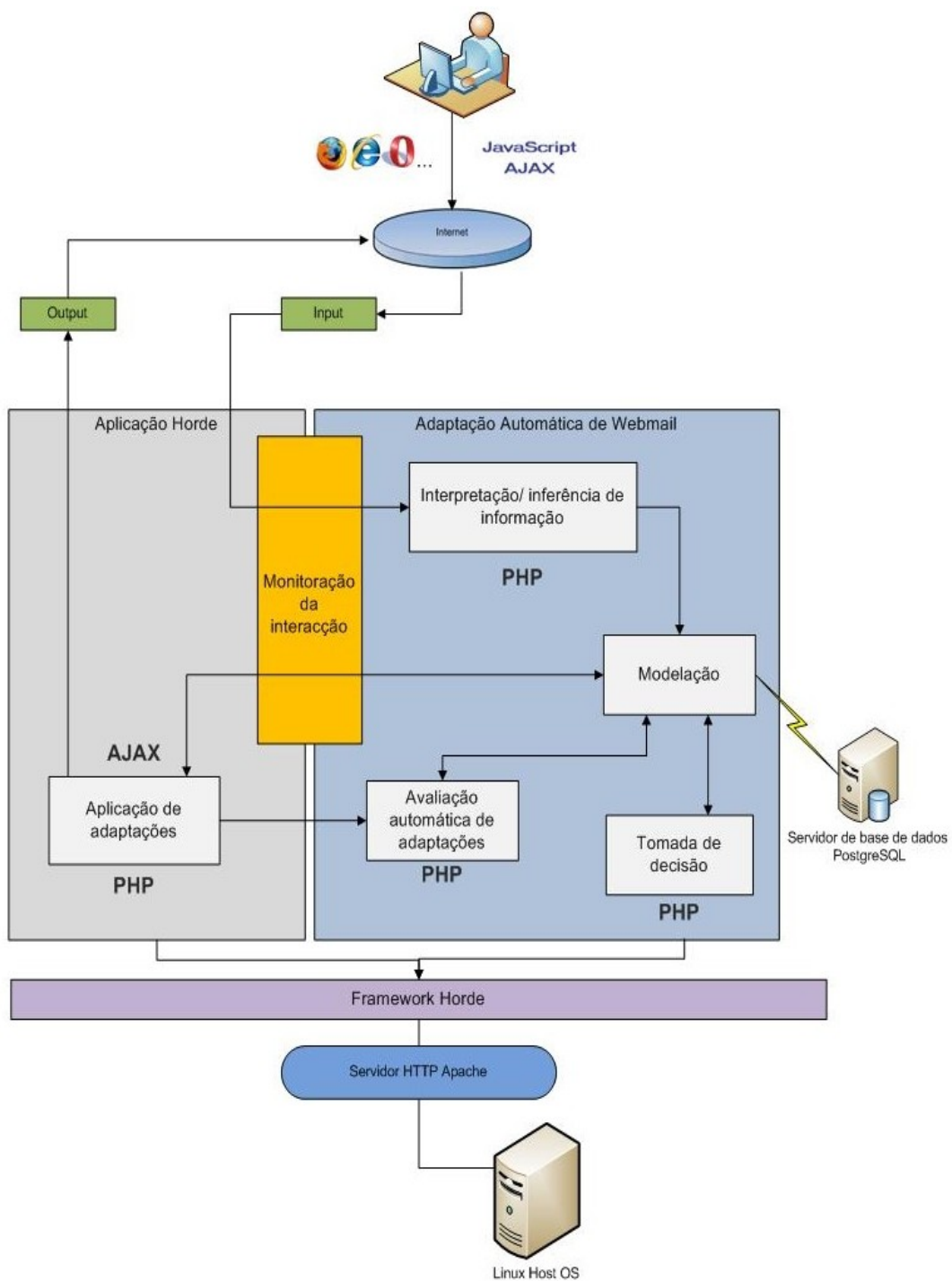


Figura 5.2: Arquitectura proposta para o desenvolvimento do sistema adaptativo

5.1.3 Etapas fundamentais

As etapas fundamentais são [Paramythis, 2001]:

- Monitoração da interacção
- Interpretação/inferência de informação (dados/comportamentos)
- Modelação
- Tomada de decisão
- Aplicação de adaptações
- Avaliação automática de adaptações

Monitoração da Interacção

A etapa de **monitoração da interacção** pretende capturar as trocas existentes entre o utilizador e a interface a diferentes níveis. Será necessário proceder a certos ajustes nos módulos desenvolvidos pela equipa técnica da Portugalmail respectivos às aplicações, de forma a cada nova acção do utilizador ou acontecimento detectados, proceder à invocação dos métodos disponibilizados pela interface programática criada (API). A API é usada com o intuito de realizar a interpretação da informação provida desse evento. A informação fornecida ao sistema pode ser implícita (detectada e inferida pelo sistema) ou explícita (caso o utilizador tenha fornecido directamente ao sistema informação sobre as suas preferências, características, interesses, entre outros, através de formulários e questionários).

Interpretação de Informação

A etapa de **interpretação de informação** é responsável por interpretar, estruturar e organizar a informação provida da etapa anterior. A informação fica assim com um formato adequado para a inserção na base de dados, e disponível para o processo de tomada de decisão retirar conclusões com base na detecção de padrões de comportamento. Esta etapa é implementada utilizando linguagem PHP com recurso a funções SQL, de forma a executar toda a lógica inerente à interpretação e comunicação com a base de dados.

Modelação

Na etapa de **modelação**, o modelo referente a cada utilizador é dinamicamente actualizado baseado no conhecimento adquirido durante o tempo de execução. Assim, o modelo do utilizador representa toda a informação que o sistema conhece sobre o utilizador. Este modelo é armazenado numa base de dados PostgreSQL e a sua manutenção está a cargo do sistema. Na base de dados PostgreSQL é igualmente armazenada informação não só relativa aos modelos respectivos a cada utilizador, mas também relativa às funcionalidades adaptativas desenvolvidas, como por exemplo: número de vezes que os utilizadores aceitaram ou rejeitaram a sugestão de uma determinada funcionalidade, ou nos casos em que não exista qualquer tipo de sugestão, regista-se se a adaptação não foi utilizada conforme o sistema tinha previsto. Estes dados

vão permitir ao sistema retirar conclusões acerca dos factores existentes nas fórmulas adoptadas no processo de tomada de decisão.

Em suma, esta etapa funciona como um intermediário entre outras etapas, uma vez que a transmissão de informação é realizada através da informação contida na base de dados.

Tomada de Decisão

A etapa de **tomada de decisão** tem como objectivo analisar os dados e detectar padrões de comportamentos (exemplos: não leitura de emails de brincadeira num determinado horário, recepção de emails de redes sociais, entre outros) mantidos no modelo do utilizador, para assim decidir se existe necessidade de efectuar algum tipo de adaptação. As decisões de adaptação daqui resultantes pretendem sempre melhorar a interacção do utilizador com o sistema. Utiliza-se mais uma vez PHP para a lógica de negócio e SQL, tanto para aceder à informação existente na base de dados, como para inserir os dados relativamente às conclusões tiradas.

Aplicação de Adaptações

Com base nas decisões concluídas na etapa anterior, esta etapa tem como objectivo sugerir ao utilizador a aplicação das novas configurações de interface ou mesmo aplicá-las, caso se trate de uma adaptação que favorece inequivocamente a interacção e faça sentido ser apresentada como uma funcionalidade primária a todos os utilizadores. Esta etapa situa-se nos módulos correspondentes às aplicações Horde, já que a responsabilidade de apresentar ao utilizador uma recomendação ou adaptação está a cargo destas.

Quando uma adaptação é aplicada à interface de um determinado utilizador é necessário verificar quais são os dados que vão ser registados na base de dados. Algumas adaptações apenas requerem a inserção do contacto de email do utilizador e o seu estado, enquanto outras podem requerir para além do referido outro tipo de dados como é o caso do IP do computador que o utilizador acede ao serviço de *webmail*. Guardar o IP faz sentido nas situações em que a presença de uma adaptação seja dependente do local onde o utilizador acede ao serviço. O armazenamento destes dados revela-se de extrema importância, visto que desta maneira o sistema mantém-se informado de quando deve adaptar e de quando deve efectuar o processo de tomada de decisão.

Para a implementação desta etapa recorre-se à tecnologia PHP combinada com AJAX de forma a tornar as páginas mais interactivas e personalizadas a cada utilizador e a SQL para comunicar com a base de dados.

Avaliação Automática de Adaptações

Esta etapa funciona como uma adaptação de segundo nível (a qual pode ser chamada “Adaptação da Adaptação”), uma vez que pretende avaliar os efeitos dos métodos e fórmulas adoptados (ou seja, avaliar o sucesso ou insucesso das adaptações efectuadas). Se os efeitos obtidos não forem os pretendidos, esta etapa é responsável por verificar onde é que o processo de tomada de decisão errou, realizando as alterações necessárias no sentido de errar cada vez menos. Esta etapa revela-se ainda mais importante, devido a que a abordagem seguida para cada uma das funcionalidades não teve qualquer tipo de suporte ou sustentação, devido à inexistência

de funcionalidades do género nos serviços de *webmail* actuais. Assim, de forma a gerar maior confiança nas opções tomadas e no código produzido desenvolveu-se esta etapa.

As alterações que esta etapa aplica sobre o processo de tomada de decisão, está a nível do factor de decisão existente nas fórmulas utilizadas nas diferentes funcionalidades. Este factor varia conforme as aceitações ou rejeições das sugestões aos diversos utilizadores ou com as escolhas efectuadas por um utilizador acerca dos grupos de contactos sugeridos (no caso da funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*”). Basicamente, o sistema pode tirar uma conclusão geral baseada nas decisões de todos os utilizadores (factor de decisão varia igualmente para todos os utilizadores – factor por funcionalidade), ou baseada apenas num único utilizador (factor de decisão varia de modo diferente entre utilizadores – factor por utilizador).

5.1.4 Sub-módulos

Apesar destas etapas poderem ser avaliadas isoladamente, estas devem ser conjugadas para chegar-se a resultados mais concretos e seguros [Paramythis, 2001]. Assim, à luz da arquitectura apresentada, faz sentido agrupar as etapas em dois sub-módulos, que são apresentados seguidamente.

Interpretação e Decisão

Este sub-módulo reúne as quatro primeiras etapas, as quais são executadas devido à ocorrência de um evento (acontecimento ou acção do utilizador). A monitoração da interacção capta o evento, enviando os dados respectivos para a etapa de interpretação e organização. Após o modelo do utilizador ser actualizado, a etapa de tomada de decisão toma lugar para concluir acerca desses dados. A actualização da base de dados pode ainda ser necessária (caso assim se verifique), fornecendo a informação a ser usada pelo módulo seguinte.

Adaptação

Neste sub-módulo estão reunidas as etapas aplicação de adaptações, monitoração da interacção, modelação e avaliação automática de adaptações. A primeira das etapas referidas acede à base de dados para decidir se deve sugerir ou aplicar uma adaptação. O utilizador pode interagir com o sistema quer seja para aceitar/rejeitar uma sugestão, quer seja para cancelar ou assumir um determinado comportamento sobre uma adaptação. No final, a etapa avaliação automática de adaptações verifica o sucesso da acção levada a cabo pelo sistema. Conforme as opções do utilizador e do sistema, a actualização da base de dados pode ser necessária.

5.2 Abordagem Proposta

A aplicação das funcionalidades “Interface Especial para Redes Sociais” e “Detecção do Horário de Trabalho” será efectuada com o consentimento do utilizador, isto é, será efectuada uma recomendação nesse sentido, ficando a decisão de aceitar ou rejeitar a cargo do

utilizador. A aplicação da funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*” é efectuada sem qualquer tipo de sugestão ou recomendação, uma vez que se trata de uma funcionalidade que não faz muito sentido sugerir, já que pode ser incluída no novo serviço de *webmail* como uma actualização à típica funcionalidade de *autocomplete*. Pelos motivos acima referidos, a análise da eficiência da lógica de negócio inerente a cada funcionalidade é realizada de forma diferente (entre a terceira e as duas primeiras funcionalidades).

5.2.1 Funcionalidades Recomendadas

Interface Especial para Redes Sociais

A ideia será detectar e registar a recepção de emails de redes sociais num espaço de tempo definido (considera-se os últimos 40 dias mais que suficientes para se concluir sobre esta funcionalidade). Quando o número de emails de redes sociais ultrapassar o factor de decisão que está sujeito a ser alterado, a recomendação da funcionalidade é efectuada. O factor de decisão nesta funcionalidade assume o valor 60 como pré-definição, o que corresponde a receber 60 emails de redes sociais num espaço de 40 dias. O processo poderia ter seguido outro rumo, como entrar com os seguintes dados: verificar quantos dos emails de redes sociais recebidos não foram lidos, ou foram eliminados sem ser lidos. Isto poderia ser uma forma de saber se um utilizador dá valor ao conteúdo desses emails ou se eles servem apenas para informar que têm uma actualização da respectiva rede social. No entanto, o facto de não ler esses emails pode não significar exactamente isso. Essa atitude pode apenas informar de certos aspectos da personalidade do utilizador em relação à gestão e organização dos emails.

Depois da decisão estar consumada, a recomendação deve ser realizada ao utilizador da próxima vez que este procede à leitura de um email recebido de uma rede social ou que inicia uma sessão no sistema (*login*). A abordagem adoptada depois da recomendação é apresentada mais à frente juntamente com a funcionalidade “Detecção do Horário de Trabalho”, uma vez que existem muitos aspectos em comum entre as duas funcionalidades.

Detecção do Horário de Trabalho

A ideia será detectar e registar os eventos de *login*, *logout* e leitura de email nos últimos 45 dias (mês e meio permite possuir uma boa fonte de dados acerca do utilizador, servindo ainda para detectar o seu comportamento mais recente). Nesta funcionalidade, o objectivo será verificar se o utilizador lê muitos emails de trabalho nos vários IPs em que acede. De referir que o número de emails de trabalho lidos num determinado IP deve ser igual ou maior que 65, de forma a existir informação suficiente para suportar uma decisão com segurança (dá uma média a rondar a leitura de dois emails de trabalho por dia útil).

A detecção da horário de trabalho do utilizador num determinado IP, vai depender se este lê emails de brincadeira no respectivo IP.

- Se se verificar que não lê, o horário será determinado com base nos tempos em que o utilizador a partir do IP em questão realizou qualquer um dos três eventos (presume-se

que sempre que acede ao serviço de *webmail* naquele IP é exclusivamente para tratar de assuntos de trabalho) .

- No caso contrário, o horário será determinado com base nos tempos em que o utilizador a partir do IP em questão leu emails de trabalho. Os tempos dos eventos de *login* e *logout* já não entram nesta tarefa, porque o utilizador lê emails de brincadeira naquele IP, logo presume-se que todo o tempo que está a aceder ao serviço de *webmail* naquele IP, não é apenas para tratar de assuntos de trabalho. No entanto, a determinação do horário neste caso, também pode basear-se nos tempos em que o utilizador leu emails de brincadeira naquele IP, caso a leitura destes ocorrer normalmente durante o horário de trabalho e durante um intervalo de tempo em específico (intervalo relativo a almoço ou lanche possivelmente). Neste intervalo não deve existir registos de mais de 5% dos emails de trabalho lidos. Assim, o sistema ainda é capaz de detectar para além do horário, uma pausa/intervalo onde os utilizadores aproveitam para relaxar e consultar emails não relacionados com trabalho. Esta situação não deve aparecer com muita frequência, mas é mais uma capacidade extra para o sistema.

O horário detectado não deve possuir tempos de entrada e saída muito dispersos (diferença de onze ou mais horas) ou muito próximos (menor do que duas horas de diferença). Em ambos os casos será difícil retirar qualquer tipo de conclusão. No primeiro, a elevada diferença de horas pode indicar que se trata de um utilizador com horário flexível, dificultando muito a decisão de quando se deve aplicar a adaptação. No segundo, devido à diferença ser mínima, a aplicação da adaptação por tão pouco período de tempo não se justifica.

No caso em que o horário é determinado com base nos tempos em que o utilizador lê emails de trabalho e os tempos de entrada e saída são muito dispersos, deve ser dada uma tolerância de 5% de forma a calcular o horário, se possível, dentro das onze horas referidas (considera-se que o utilizador pode ter feito horas extras uma vez ou outra).

Depois de todos os dados reunidos e as restrições anteriores serem contornadas com sucesso, é aplicada a fórmula 5.1 que permite concluir sobre a aplicação da funcionalidade.

$$\text{Resultado} = \text{Factor} * \text{EmailsFunAll} - \text{EmailsFunIP} \quad (5.1)$$

Legenda:

Resultado – Resultado da fórmula adoptada.

Factor – Simboliza o factor de decisão actual.

EmailsFunAll – Número de emails de brincadeira que o utilizador leu em todos os IPs, incluindo o IP onde se encontra mas não durante o horário de trabalho.

EmailsFunIP – Número de emails de brincadeira que o utilizador leu naquele IP durante o horário de trabalho.

Repare-se que os emails de trabalho não entram na fórmula, visto que o objectivo é saber se ocultar os emails de brincadeira durante o horário de trabalho constitui uma mais valia. Além disso, os emails de trabalho já tiveram um papel importante para a detecção do horário e

detecção dos IPs de trabalho. Assim, a tomada de decisão é baseada na diferença entre os emails de brincadeira lidos naquele IP durante o horário de trabalho e todos os restantes emails de brincadeira lidos (uma parte destes determinada pelo factor de decisão).

Se o *Resultado* for maior ou igual a zero, significa que o número de emails lidos durante o horário de trabalho no correspondente IP é irrelevante em comparação com os emails de brincadeira lidos fora do trabalho. Por outras palavras, os emails de brincadeira que lê durante o trabalho não são assim muitos comparando com o seu comportamento habitual. A leitura de emails de brincadeira no trabalho pode ser devido ao utilizador receber emails desse género no preciso momento que está a consultar o email, por ver emails de brincadeira não lidos na caixa de entrada (devido ao stress sente-se tentado), ou por não apreciar ter a caixa de entrada com mensagens não lidas. Nesse sentido a recomendação da funcionalidade deve ser realizada.

Se o *Resultado* for menor do que zero, conclui-se que a recomendação não deve ser realizada pois a leitura de emails de brincadeira é um hábito já assumido pelo utilizador e este não o faz por mera distração ou descontração (ocultar os emails de brincadeira iria provocar o descontentamento deste utilizador). Estes utilizadores não passam sem a leitura deste tipo de emails no horário de trabalho ou porque não têm possibilidade de o fazer noutra local ou altura, ou porque não conseguem concentrar-se no trabalho.

Como *default* o valor do factor de decisão é 0.25, o que significa que se um utilizador ler um número de emails de brincadeira fora do trabalho pelo menos quatro vezes superior aos emails de brincadeira que lê no trabalho, a recomendação é realizada. No entanto, este valor está sujeito a alteração conforme o sucesso ou insucesso obtido nas recomendações efectuadas aos utilizadores.

Depois da decisão estar consumada, a recomendação deve ser realizada ao utilizador da próxima vez que este efectuar *login*, mas unicamente no IP onde se detectou que esta adaptação devia ser aplicada.

Comum às duas funcionalidades (após recomendar)

Perante a recomendação o utilizador possui duas opções: aceitar ou rejeitar. De notar que a recomendação da aplicação é efectuada em duas tentativas. Caso um utilizador rejeite à primeira, será realizada uma segunda tentativa após um intervalo de tempo definido (daí ser importante guardar a data das sugestões). A repetição da recomendação justifica-se porque o utilizador quando confrontado pela primeira vez com a recomendação da funcionalidade pode estar sem disposição para tomar atenção ao assunto, pode ter carregado sem intenção no botão de rejeitar, pode ainda não ter ouvido falar das capacidades da funcionalidade, entre outras possibilidades. A opção “Decidir mais tarde” também pode ser uma possível alternativa, embora esta ainda não seja considerada. Depois de aplicada a adaptação, o utilizador tem sempre a possibilidade de cancelar/eliminar essa aplicação.

Considera-se que quando o número de recomendações atinge as 100, já se possui dados suficientes para se começar a tirar conclusões acerca do factor de decisão adoptado no algoritmo de tomada de decisão. Por outras palavras, menos de 100 amostras considera-se que não existem dados suficientes para se alterar o factor de decisão com segurança. O pretendido é que o sistema tome a decisão correcta pelo menos 95% das vezes (margem de erro de 5%). Assim,

devem ser contadas as vezes que a adaptação foi aceiteada e rejeitada para verificar então a eficácia do algoritmo implementado. Se o número de rejeições corresponder a 5% ou mais do número total de recomendações, deve-se:

- No caso da funcionalidade “Interface Especial para Redes Sociais”:
 - Deve-se somar 0.5 ao factor de decisão actual, porque verificou-se que o número de emails de redes sociais recebidos pelos utilizadores (representado pelo factor de decisão), ainda não se revelava suficiente para recomendar a funcionalidade com o sucesso pretendido. Noutras palavras, revelou-se que o número que representa o factor de decisão ainda não correspondia a um número suficiente de emails para justificar a aplicação de uma interface que aglomera a informação sobre as redes sociais. O valor 0.5 serve para alterar o factor de decisão gradualmente até encontrar-se o valor ideal.
- No caso da funcionalidade “Detecção do Horário de Trabalho”:
 - O factor de decisão deve ser decrementado porque os emails de brincadeira que os utilizadores lêem durante o horário de trabalho num determinado IP continuam a ser muitos em relação aos emails de brincadeira que estes utilizadores lêem fora desse horário. Assim, pretende-se aumentar a diferença entre os emails de brincadeira lidos no trabalho e fora deste, de modo a efectuar uma recomendação mais segura. O valor 0.005 serve para alterar o factor de decisão decrementalmente até encontrar-se o valor ideal.

O número de aceitação deve ser decrementado de uma unidade quando algum utilizador cancela a respectiva adaptação, pois a aceitação da adaptação fica sem efeito (o que faz sentido porque afinal aquela adaptação não foi apreciada ou não foi útil para o utilizador). O facto de retirar um valor ao número de aceitação, faz com que aumente o peso do número de rejeições.

5.2.2 Funcionalidade Não Recomendada

Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete

A ideia será detectar e registar o envio de emails num período de tempo determinado (últimos 60 dias). Nesta funcionalidade interessam dados referentes a um espaço de tempo maior do que nas duas primeiras funcionalidades, devido à recepção de emails e à leitura de emails ocorrerem com maior frequência do que o envio.

Quando o utilizador se encontra a preencher o destinatário de um email, é sugerida a lista de sugestões de grupos de contactos. Esta deve ser actualizada a cada inserção de um carácter (sem ser vírgula) no campo de preenchimento de destinatário. Os grupos que são apresentados na lista, passam por um processo de correspondência (*match*) com o que se encontra escrito no campo de destinatário (este processo é explicado detalhadamente no capítulo referente à implementação do sistema adaptativo). A ordenação desses grupos é baseada numa classificação que é calculada a partir da fórmula 5.2.

$$\text{Classificação} = \text{ContactosIP} + \text{Factor} * \text{ContactosOutrosIP} \quad (5.2)$$

Legenda:

Classificação – Simboliza a classificação atribuída a um contacto ou grupo de contactos para efeitos de ordenação, dependendo do IP que o utilizador acede.

ContactosIP – Número de emails que um utilizador enviou do IP onde se encontra para aquele contacto ou grupo de contactos.

Factor – Factor que define quanto deve ser a importância dada aos emails enviados a partir de outros IPs em relação aos emails enviados do IP que o utilizador acede.

ContactosOutrosIP – Número de emails que um utilizador enviou de outros IPs para aquele contacto ou grupo de contactos.

Depois de enviar um email, deve-se verificar onde recaiu a escolha do utilizador em relação aos contactos sugeridos no *autocomplete*. Esta verificação é necessária para averiguar se a ordenação dos contactos sugeridos encontra-se correctamente adaptada ao utilizador e tomar medidas caso não se confirme esse cenário. Caso tenha recaído sobre:

- A primeira opção com classificação, nenhuma medida é tomada pois antecipou-se correctamente o comportamento do utilizador.
- Uma opção sem qualquer tipo de classificação (contactos que se encontram ordenados por ordem alfabética), nenhuma medida é tomada pois foi escolhido um contacto que não estava presente no histórico do utilizador. Desta forma, o sistema não pode fazer nada já que se trata de um dado novo, sem antecedentes. No entanto, esta informação é registada podendo ser utilizado da próxima vez.
- Outra opção com classificação, é necessário averiguar onde é que o factor de decisão falhou.

No sentido de averiguar onde é que o factor de decisão falhou, iguala-se a classificação da opção escolhida com a classificação da primeira opção (fórmula 5.3). A fórmula 5.4 é a resolução da fórmula 5.3 em relação ao *Factor*, a qual é utilizada para saber quanto é que deveria ser o factor para a classificação da opção escolhida ser pelo menos igual à primeira opção. Convém referir que o valor pré-definido do factor de decisão é 0.8 (dá maior relevância ao IP em que se acede mas sem esquecer o comportamento geral do utilizador). De notar também, que antes de aplicar a fórmula 5.4, deve-se verificar se o denominador é diferente de zero (isto é, se os números de emails enviados de outros IPs para os grupos da opção escolhida e da primeira opção, são diferentes para a variação do factor de decisão resultar em modificações de classificação). Caso seja zero, o *Factor* mantém-se inalterado (visto que a partir da sua alteração não resulta nada de diferente). Caso seja diferente de zero, deve-se verificar se o *Factor* calculado é maior ou menor que o seu valor antigo. Em ambos os casos, deve-se actualizar o factor no sentido correcto (incrementar ou decrementar o factor em 0.005 unidades).

$$\text{EscIP} + \text{Factor} * \text{EscOutrosIP} = \text{MaxIP} + \text{Factor} * \text{MaxOutrosIP} \quad (5.3)$$

$$Factor = (EscIP - MaxIP) / (MaxOutrosIP - EscOutrosIP) \quad (5.4)$$

Legenda:

Factor – Simboliza quanto deveria ser o factor para a classificação entre a opção escolhida e a primeira opção serem equivalentes.

EscIP – Número de emails que um utilizador enviou do IP onde se encontra para os contactos correspondentes à opção escolhida.

EscOutrosIP - Número de emails que um utilizador enviou de outros IPs para os contactos correspondentes à opção escolhida.

MaxIP - Número de emails que um utilizador enviou do IP onde se encontra para os contactos correspondentes à primeira opção.

MaxOutrosIP - Número de emails que um utilizador enviou de outros IPs para os contactos correspondentes à primeira opção.

6 Tecnologias utilizadas

Neste capítulo são apresentadas e descritas as tecnologias necessárias para o desenvolvimento do projecto com o sucesso pretendido. A descrição de cada tecnologia será acompanhada por uma referência à utilidade que a respectiva tecnologia teve no curso do projecto.

As tecnologias utilizadas encontram-se seguidamente organizadas em tecnologias de *backend*, como é o caso do PostgreSQL, tecnologias de *frontend*, como são os casos do HTML, PHP, CSS e SQL (AJAX será referida porque vai ser necessária para a próxima fase do projecto), e tecnologias de apoio ao desenvolvimento do projecto, como é o caso do gedit ou a Wiki utilizada na Portugalmail.

6.1 Tecnologias de *Backend*

PostgreSQL é um dos sistemas de gestão de base de dados de código aberto mais avançados e utilizados actualmente, possuindo como características principais a extensibilidade, estabilidade, fiabilidade, independência de plataforma e desenho para armazenamento de grande volume de dados. Indexação por texto, linguagem procedural em várias linguagens (PL/pgSQL, PL/Java, entre outras) para procedimentos de armazenagem, views, triggers, controle de concorrência multi-versão, consultas complexas ou integridade transaccional são alguns dos recursos disponibilizados pelo PostgreSQL que contribuíram para a sua escolha enquanto sistema de gestão de base de dados [Postgresql, 2009; WikiPostgresqlpt, 2009; WikiPostgresqlen, 2009]. No projecto, as capacidades de gestão de base de dados disponibilizadas por esta tecnologia serão especialmente úteis no armazenamento da informação respectiva às funcionalidades adaptativas e à informação dos diversos utilizadores (modelo do utilizador).

6.2 Tecnologias de *Frontend*

HTML (HyperText Markup Language) é uma linguagem de marcação utilizada para a criação de páginas *Web*, a qual descreve a formatação e estrutura da informação presente num documento, denotando texto como *links*, cabeçalhos, parágrafos, listas, etc., e enriquecendo esse texto com formulários interactivos, imagens ou outros objectos. Todo o documento HTML é escrito e apresentado sob a forma de etiquetas (*tags*), as quais se encontram entre parênteses angulares (sinais de maior e menor). As etiquetas são os elementos responsáveis pela formatação da linguagem [WikiHTMLen, 2009; WikiHTMLpt, 2009]. Esta é uma das tecnologias responsáveis pelo conteúdo que os utilizadores poderão visualizar nas páginas respectivas às adaptações.

PHP (Hypertext Preprocessor) é uma linguagem de programação livre e amplamente usada na criação de soluções *Web* dinâmicas e interactivas, que é o seu principal propósito. PHP é normalmente usado conjuntamente com o servidor *Web* Apache em vários sistemas operativos e pode ser embebida em código HTML. As suas características principais são: velocidade, robustez, portabilidade (independente da plataforma), orientação a objectos e sintaxe similar às linguagens C/C++ e PERL [PHPGroup, 2009; WikiPHPpt, 2009; WikiPHPen, 2009]. PHP constitui a linguagem principal deste projecto, tanto pelas características mencionadas, como por ser responsável por toda a lógica de negócio inerente ao projecto e pela sua presença em todas as etapas correspondentes à implementação do sistema adaptativo.

SQL (Structured Query Language) é uma linguagem de pesquisa declarativa concebida para a gestão de informação em sistemas de gestão de base de dados relacional, permitindo a recuperação, inserção, actualização e eliminação de dados. SQL foi padronizada primeiro pela ANSI (*American National Standards Institute*) e mais tarde pela ISO. (*International Organization for Standardization*) [WikiSQLpt, 2009; WikiSQLen, 2009]. Esta linguagem revela-se importante na medida em que a interacção com a base de dados é realizada através da execução singular de chamadas SQL.

CSS (Cascading Style Sheets) é uma linguagem de estilo usada para descrever a apresentação (aparência e formatação) de um documento ou página escrita através de uma linguagem de marcação (como por exemplo HTML ou XHTML). O principal benefício desta linguagem é prover a separação entre o conteúdo (escrito numa linguagem de marcação) e a apresentação (definição de elementos como cores, fontes e *layout*) de um documento. CSS pode ainda permitir que uma mesma página seja apresentada em diferentes estilos para diferentes métodos de renderização, o que pode ser bastante útil na adaptação a diferentes utilizadores [WikiCSSpt, 2009; WikiCSSen, 2009]. Esta é uma tecnologia bastante importante para o desenho da interface das adaptações (o modo como será apresentado o conteúdo).

AJAX (Asynchronous Javascript and XML) consiste na conjugação de várias tecnologias conhecidas de desenvolvimento *Web*, como Javascript e XML, e é usado com o intuito de criar aplicações *Web* mais dinâmicas, interactivas e criativas [WikiAJAXpt, 2009; WikiAJAXen, 2009]. Esta tecnologia revelar-se-á de extrema importância numa próxima fase para realização das recomendações das adaptações, bem como na aplicação de algumas destas (casos em que as capacidades dinâmicas sejam requeridas).

6.3 Tecnologias de Apoio

O projecto foi desenvolvido utilizando os sistemas operativos Linux e Windows. Em Linux recorreu-se a aplicações como o OpenOffice.org 2.4 para a escrita da tese, o gedit para a escrita do código relativo à implementação das funcionalidades, o FreeMind para a construção do mapa mental que consiste numa forma visual de representar o conjunto de ideias surgido durante as sessões do *brainstorming*, e o GIMP como editor de imagens. Em Windows recorreu-se ao Microsoft Visio para desenhar a arquitectura do sistema. Em ambos os sistemas operativos utilizou-se o Firefox tanto como ferramenta de apoio à pesquisa e estudo sobre a matéria, como para visualizar os resultados à medida que as funcionalidades iam sendo implementadas.

Wiki é uma página ou conjunto de páginas Web que permite a criação e edição colectiva dessas páginas usando uma linguagem de marcação muito simples e eficaz. A grande maioria das Wikis são baseadas na *Web* e podem ser acedidas através de um *browser*. Uma única página numa Wiki é designada como “única página”, enquanto o conjunto global de páginas que se encontram altamente interligadas é que se designa Wiki. As Wikis são maioritariamente utilizadas para suportar trabalho colaborativo, onde qualquer utilizador que tenha acesso pode contribuir e modificar o seu conteúdo. A Portugalmail possui uma Wiki que mantém toda a informação acerca dos seus projectos. Este projecto como sendo parte integrante da Portugalmail não fugiu à regra e todos os passos seguidos na criação e desenvolvimento do projecto foram registados [WikiWikip, 2009; WikiWikien, 2009; AvilaCosta, 2009].

7 Sistema Adaptativo em Serviço de *Webmail*

Este capítulo é totalmente dedicado à documentação de todas as fases envolvidas na implementação das funcionalidades desenvolvidas. Aconselha-se, caso o leitor não se recorde no que consiste cada funcionalidade, uma pequena revisão ao subcapítulo “4.3 Seleccção de Ideias”.

A abordagem seguida foi baseada na proposta de solução apresentada no capítulo 5, mais concretamente com base nas etapas fundamentais identificadas, e nas fórmulas propostas para apoiar a decisão e avaliação.

Notas:

A partir deste ponto quando se falar em primeira, segunda ou terceira funcionalidade está-se a referir respectivamente às funcionalidades pela seguinte ordem: interface especial para redes sociais, detecção do horário de trabalho e sugestão de grupos de contactos na funcionalidade de *autocomplete*.

7.1 Monitoração da Interacção

Esta etapa é implementada do lado das aplicações existentes no serviço de *webmail*. Para cada funcionalidade adaptativa existe um evento (acção ou acontecimento) que acciona a comunicação dos módulos referentes a essas aplicações com o módulo referente ao sistema adaptativo desenvolvido. A comunicação é baseada na invocação dos métodos existentes na API desenvolvida para o efeito. A informação encontra-se no formato *string* onde os dados necessários encontram-se separados por *tags*. De seguida apresenta-se um quadro onde são

referidas os eventos que fazem accionar todo o processo adaptativo relativo a uma funcionalidade.

Tabela 1: Eventos que fazem activar o processo relativo a cada funcionalidade

Funcionalidade	Evento (acção do utilizador ou acontecimento)
Interface especial para redes sociais	Receber um email
Detecção do horário de trabalho	Efectuar <i>login</i> , <i>logout</i> ou ler um email
Sugestão de grupos de contactos na funcionalidade de <i>autocomplete</i>	Enviar um email

Os eventos presentes no quadro dizem respeito à fase de recolha de dados para a tomada de decisão. No entanto, interessa também monitorar a interacção com o utilizador no momento da recomendação e mesmo depois da adaptação estar realizada. Como esta etapa encontra-se interligada com a etapa “Aplicação de Adaptações”, na apresentação desta será demonstrado a resposta do sistema às acções do utilizador.

7.2 Interpretação de Informação

O processo de interpretação da informação é realizada em traços gerais da mesma forma para as três funcionalidades, visto que todas elas fazem a descodificação da informação recebida da etapa anterior através de um *parsing*, e caso assim se confirme, inserem na base de dados os dados relativos ao evento e iniciam o processo de tomada de decisão (verificar se os novos dados produzem alguma alteração).

As diferenças que existem entre as três funcionalidades são descritas seguidamente.

Na primeira funcionalidade é necessário verificar se o email que o utilizador recebeu pertence a alguma rede social, e caso se confirme é registado na base de dados esse acontecimento com a informação necessária para a tomada de decisão.

Na terceira funcionalidade é necessário verificar se existe contactos de email duplicados aquando um envio de um email e retirá-los.

7.3 Modelação

Nesta secção pretende-se apresentar as tabelas criadas na base de dados, por forma a armazenar todos os dados necessários ao correcto funcionamento das respectivas funcionalidades. A apresentação de cada tabela será acompanhada por uma descrição, onde se indica a razão da criação da tabela, e pela indicação e descrição dos campos correspondentes.

Interface Especial para Redes Sociais

Para a implementação da primeira funcionalidade foram criadas duas tabelas: *socialnetwork* e *socialnetwork_adaptations*. A primeira tabela pretende registar a informação correspondente à recepção de emails de uma determinada rede social. A segunda tem a finalidade de indicar ao sistema a quem já foi recomendada a aplicação desta funcionalidade, quem rejeitou, quem aceitou, quem cancelou a funcionalidade depois de aplicada ou para quem a recomendação deve ser realizada. O campo *date* presente em ambas tabelas pode ser útil caso se pretenda realizar estatísticas sobre os dados em diferentes intervalos de tempo.

Tabela 2: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Interface Especial para Redes Sociais” e seus respectivos campos

<i>Interface Especial para Redes Sociais</i>		
Tabela	Campos	Descrição
<i>socialnetwork</i>	<i>social_network</i>	Nome da rede social que enviou o email.
	<i>user_</i>	Utilizador que recebeu o email.
	<i>date</i>	Data que foi recebido o email. Este campo será importante para verificar qual é o comportamento mais recente do utilizador. Para efeitos de decisão apenas interessam os emails recebido até 40 dias.
<i>socialnetwork_adaptations</i>	<i>user_</i>	Identificação do utilizador.
	<i>status</i>	Indicação sobre o estado actual da funcionalidade em relação ao utilizador: aplicada, rejeitada (uma ou duas vezes), cancelada ou pronta para sugestão.
	<i>date</i>	Data em que o estado acima mencionado foi actualizado a última vez.

Detecção do Horário de Trabalho

Na segunda funcionalidade também foram criadas duas tabelas, que são: *worktime* e *worktime_adaptations*. A primeira tabela pretende registar a informação correspondente a um dos seguintes eventos: *login*, *logout* ou leitura de um email. A tabela *worktime_adaptations* é muito parecida com a segunda tabela referida em cima para a primeira funcionalidade. A única alteração é o facto que em vez de associar-se os campos *status* e *date* a um utilizador, associa-se ao par utilizador-IP (a um utilizador num determinado IP de acesso).

Tabela 3: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Detecção do Horário de Trabalho” e seus respectivos campos

<i>Detecção do Horário de Trabalho</i>		
Tabela	Campos	Descrição

<i>worktime</i>	<i>event</i>	Identificação do evento. O evento pode ser: leitura de um email relacionado com trabalho, leitura de um email de brincadeira, <i>login</i> ou <i>logout</i> .
	<i>user_</i>	Utilizador responsável pelo evento.
	<i>ip</i>	Indicação do IP em que o utilizador se encontrava a aceder ao serviço no momento do evento.
	<i>time_</i>	Data e hora do evento. Para efeitos de decisão apenas interessa o comportamento do utilizador no passado recente (45 dias).
<i>worktime_adaptations</i>	<i>user_</i>	Identificação do utilizador.
	<i>ip</i>	IP de acesso.
	<i>status</i>	Indicação sobre o estado actual da funcionalidade em relação ao utilizador: aplicada, rejeitada (uma ou duas vezes), cancelada ou pronta para sugestão.
	<i>date</i>	Data em que o estado acima mencionado foi actualizado a última vez.

Interface Especial para Redes Sociais & Detecção do Horário de Trabalho

A seguinte tabela (*suggestion_history*) está associada às duas primeiras funcionalidades, e pretende armazenar os dados relativos às recomendações. Estes dados vão permitir averiguar se o processo de tomada de decisão existente na aplicação das respectivas funcionalidades está a surtir efeito positivo ou negativo junto dos utilizadores. O factor de decisão também pode ser consultado a partir desta tabela, o qual é baseado nos dados das recomendações.

Tabela 4: Tabela comum às duas primeiras funcionalidades

<i>Interface Especial para Redes Sociais & Detecção do Horário de Trabalho</i>		
Tabela	Campos	Descrição
<i>suggestion_history</i>	<i>adaptation</i>	Identificação da funcionalidade: <i>socialnetwork</i> ou <i>worktime</i> .
	<i>accept</i>	Número de vezes que os utilizadores aceitaram a recomendação da respectiva funcionalidade.
	<i>deny</i>	Número de vezes que os utilizadores rejeitaram ou cancelaram a recomendação da respectiva funcionalidade.
	<i>date</i>	Data que se iniciou a contagem dos dois campos acima mencionados. Nesta data, o factor de decisão que está representado pelo campo <i>var1</i> assume o valor pré-definido.

	<i>var1</i>	Representa o factor de decisão de uma determinada funcionalidade, que vai sendo alterado conforme o sistema vai aprendendo. Este factor varia igualmente para todos os utilizadores da funcionalidade correspondente.
--	-------------	---

Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete

Na terceira funcionalidade as tabelas criadas foram: *autocomplete*, *autocomplete_groups*, *autocomplete_recipients*, *autocomplete_groups_recipients*, *autocomplete_factors* e *autocomplete_suggestion_temp*. A primeira tabela é semelhante às tabelas *socialnetwork* e *worktime* mencionadas em cima, visto que foi criada para o mesmo efeito - registar informação associada a um evento, neste caso enviar um email. A segunda, terceira e quarta tabela têm a finalidade de armazenar as conclusões retiradas durante o processo de tomada de decisão e estão relacionadas entre si, uma vez que existe uma relação de muitos para muitos entre a segunda e terceira tabela, servindo a quarta como intermediário entre as duas (um grupo de contactos possui vários contactos e um contacto pode está incluído em vários grupos de contactos). A informação presente nestas tabelas, será utilizada para efectuar a ordenação e respectiva sugestão dos grupos de contactos na funcionalidade de *autocomplete*. A quinta tabela (*autocomplete_factors*) tem como objectivo guardar os factores de decisão para cada utilizador. A última tabela foi criada simplesmente para guardar informação temporária de forma a servir de apoio à etapa “Avaliação Automática de Adaptações”.

Tabela 5: Apresentação das tabelas criadas para efeitos de desenvolvimento da funcionalidade “Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete” e seus respectivos campos

<i>Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete</i>		
Tabela	Campos	Descrição
<i>autocomplete</i>	<i>user_</i>	Utilizador responsável pelo evento.
	<i>ip</i>	Indicação do IP em que o utilizador se encontrava no momento que realizou o evento.
	<i>recipients</i>	Endereços de email que são destinatários do email enviado.
	<i>date</i>	Data em que foi realizado o evento.
<i>autocomplete_groups</i>	<i>id</i>	Identificação do grupo de contactos.

	<i>user_</i>	Utilizador responsável pelo envio de emails para esse grupo.
	<i>ip</i>	Indicação do IP.
	<i>count_ip</i>	Número de emails enviados para esse grupo pelo utilizador e IP referidos acima.
	<i>count</i>	Classificação responsável pela ordenação dos grupos. Cada grupo possui uma classificação consoante o <i>user_</i> e <i>ip</i> .
<i>autocomplete_recipients</i>	<i>id</i>	Identificação do endereço/contacto de email.
	<i>recipient</i>	Contacto de email.
<i>autocomplete_groups_recipients</i>	<i>group_id</i>	Identificação do grupo de contactos.
	<i>recipient_id</i>	Identificação do contacto de email.
<i>autocomplete_factors</i>	<i>user_</i>	Utilizador.
	<i>factor</i>	Factor de decisão associado ao utilizador.
<i>autocomplete_suggestion_temp</i>	<i>user_</i>	Utilizador.
	<i>recipients_max</i>	Contactos correspondentes à primeira opção, aquando a escolha de uma opção da lista de sugestões apresentadas pela funcionalidade de <i>autocomplete</i> .
	<i>recipients_max_temp</i>	Contactos correspondentes à primeira opção da lista de sugestões apresentadas pela funcionalidade de <i>autocomplete</i> , mas sem escolha de qualquer opção.

7.4 Tomada de Decisão

Esta etapa consiste na etapa principal num sistema adaptativo, pois é aqui que se encontra o núcleo da lógica de negócio. O sistema assume um papel preponderante, realizando decisões com base na análise dos dados conhecidos de cada utilizador e de cada funcionalidade.

Interface Especial para Redes Sociais

O método seguido para a tomada de decisão na primeira funcionalidade é o mais simples. Inicialmente é realizada uma consulta à base de dados para obter o factor de decisão actual. De seguida, também recorrendo à base de dados faz-se uma contagem do número de emails que o respectivo utilizador recebeu de redes sociais nos últimos 40 dias. Se esse número for igual ou superior ao factor de decisão, actualiza-se o campo *status* da tabela *socialnetwork_adaptations* (de modo a informar à etapa “Aplicação de adaptações” que pode prosseguir com a recomendação da adaptação ao utilizador).

Para uma melhor compreensão de todo o raciocínio, apresenta-se na figura 7.1 um fluxograma.

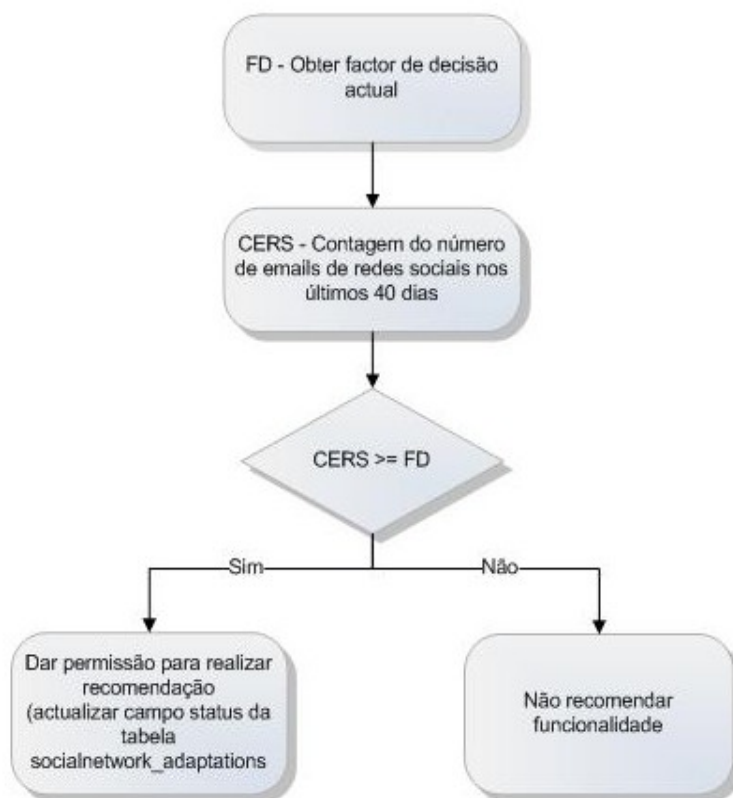


Figura 7.1: Fluxograma correspondente à tomada de decisão da funcionalidade "Interface especial para redes sociais"

Detecção do Horário de Trabalho

Na segunda funcionalidade, o método adoptado já é bastante mais complexo. Primeiro, são consultados e guardados todos os registos de um utilizador em qualquer IP relativos a *login*, *logout* ou leitura de emails nos últimos 45 dias. A partir destes registos verifica-se se o utilizador leu emails relacionados com trabalho. Se não, o sistema não pode fazer qualquer tipo

de inferência e termina aí o processo de decisão. Se sim, é feita outra verificação – se leu emails de trabalho a partir de um ou mais IPs. Se for a partir de mais do que um, verifica-se quais são os IPs onde o utilizador leu um maior número de emails de trabalho (interessam os IPs que possuem o maior e o segundo maior número de emails lidos, sendo o limite inferior de 65 emails de trabalho). A partir da altura que se sabe quais são os IPs que interessam analisar é seguido um processo paralelo para cada um. Esse processo é executado pela seguinte ordem:

- Se a funcionalidade já se encontra recomendada, aplicada ou cancelada àquele utilizador naquele IP, terminar processo imediatamente. Caso contrário o processo continua.
- Consultar base de dados para obter o factor de decisão actual.
- Guardar os dias em que o utilizador leu emails de trabalho naquele IP. Permite saber os dias que um utilizador normalmente costuma aceder ao serviço a partir daquele IP.
- Guardar tempos em que o utilizador a partir do IP em questão procedeu à leitura de emails de trabalho, à leitura de emails de brincadeira, e à realização de qualquer um dos seguintes eventos: *login*, *logout* e leitura de emails.
- Contagem do número de emails de brincadeira lidos, tanto para o IP em questão como para todos os IPs.
- Se a contagem efectuada no ponto anterior para o respectivo IP for igual a zero:
 - A determinação do horário será efectuada com base nos tempos em que o utilizador a partir do IP em questão realizou qualquer um dos três eventos.
- Se a contagem for maior que zero:
 - A determinação do horário será efectuada com base nos tempos em que o utilizador a partir do IP em questão leu emails de trabalho.
 - Verificar se existe intersecção de mais de 5% dos emails de trabalho lidos com os tempos máximo e mínimo em que o utilizador leu emails de brincadeira naquele IP. Se não se verificar, é detectado um intervalo no horário correspondente aos tempos máximo e mínimo acima referidos.
- Verificar se os tempos de entrada e saída do horário detectado não possuem uma diferença de onze ou mais horas ou menos do que duas horas. Se possuem uma diferença de onze ou mais horas (no caso da determinação do horário com base nos tempos dos emails de trabalho), verifica-se com uma tolerância de 5% se é possível encaixar o horário dentro das onze horas. Se se conseguir e existir várias alternativas, é escolhida a que possui menor diferença entre entrada e saída.
- Passando com sucesso na verificação anterior, aplica-se a fórmula 5.1 que permite concluir sobre a aplicação da funcionalidade.

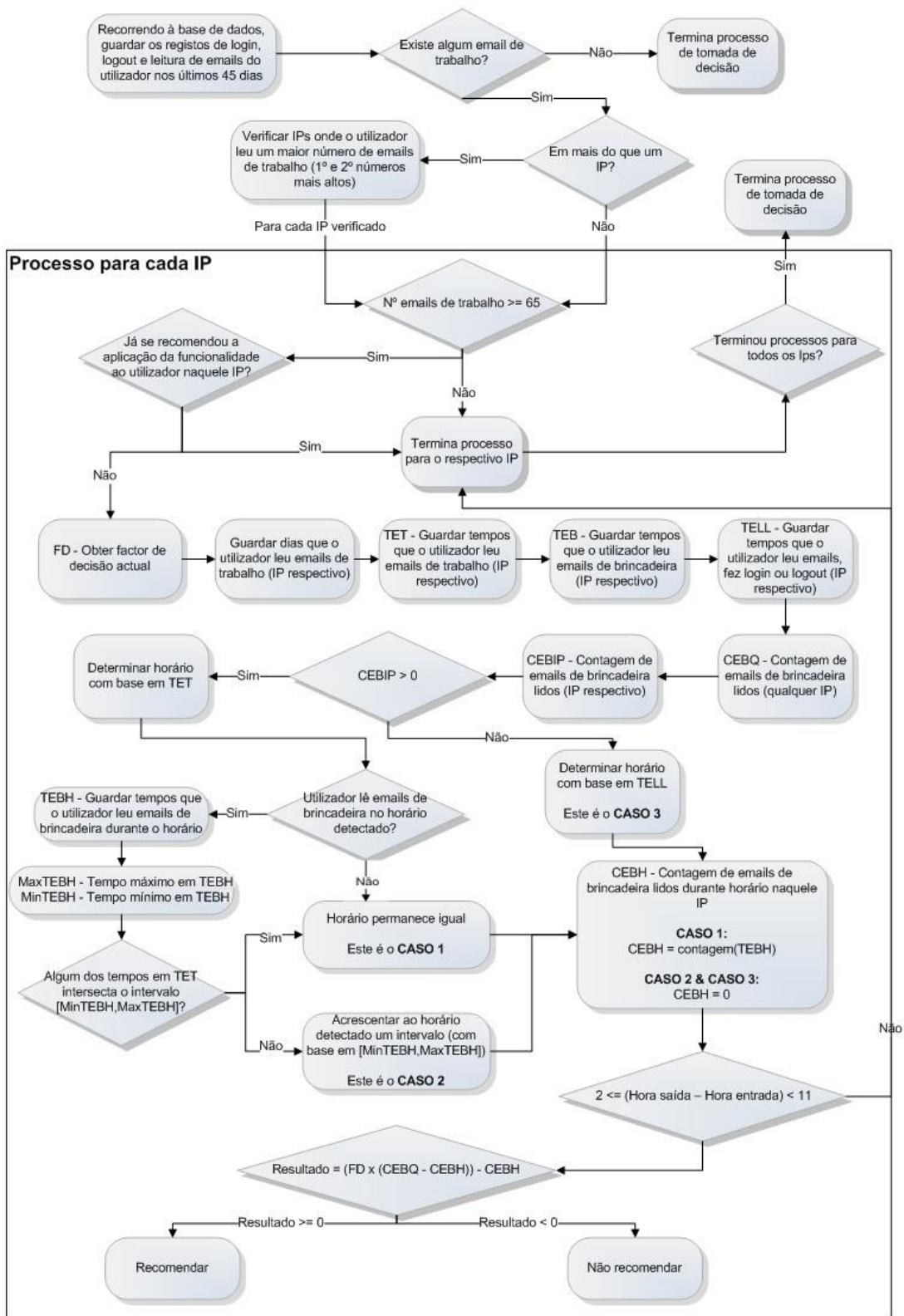


Figura 7.2: Fluxograma correspondente à tomada de decisão da funcionalidade "Detecção do horário de trabalho"

Sugestão de Grupos de Contactos na Funcionalidade de *Autocomplete*

Em relação à terceira funcionalidade, a abordagem adoptada é apresentada seguidamente. Primeiro é necessário realizar uma consulta à base de dados para obter o factor de decisão actual correspondente àquele utilizador. De seguida, também recorrendo à base de dados são consultados e guardados os IPs onde o utilizador enviou emails, e para cada um desses IPs, os grupos de contactos que constituíram destinatários desses emails (acompanhados pelo número de emails enviados). A partir destes dados, o processo de tomada de decisão resume-se aos seguintes passos:

- Para cada IP:
 - Guardar os grupos de contactos a quem foram enviados emails a partir doutros IPs e respectivas contagens.
 - Guardar grupos de contactos comuns entre este IP e os outros IPs.
 - Guardar grupos de contactos a quem foram enviados emails só deste IP.
 - Guardar grupos de contactos a quem foram enviados emails só de outros IPs.
 - Calcular classificação de cada grupo de contactos, com base na formula 5.2.

- Depois da fase de análise e organização dos dados, segue-se a fase correspondente à inserção da nova informação na base de dados (ou actualização, caso seja apenas necessário efectuar *update* ao campo correspondente à classificação de um grupo de contactos).

O próximo fluxograma ajuda a perceber melhor todo este processo.

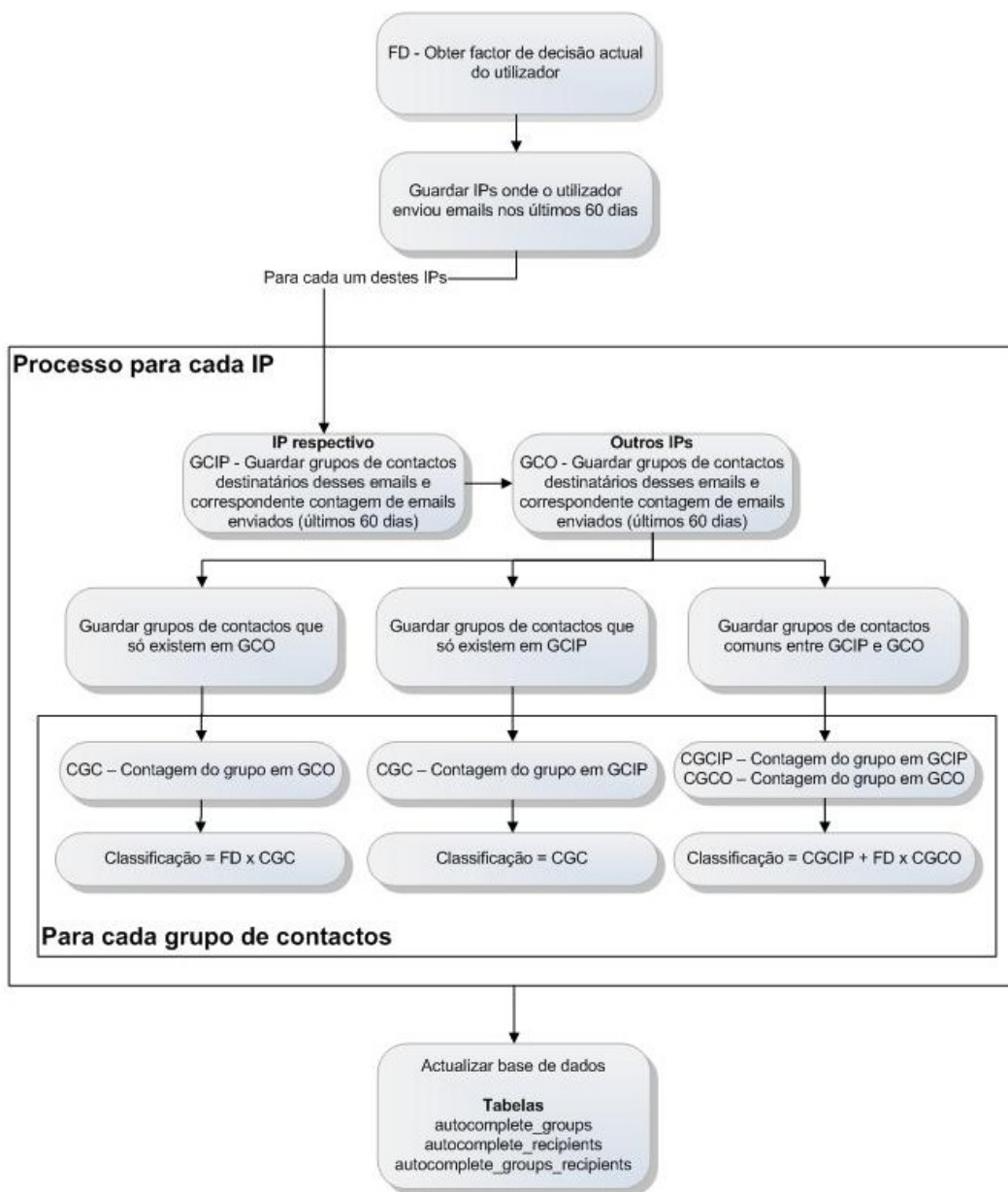


Figura 7.3: Fluxograma correspondente à tomada de decisão da funcionalidade "Sugestão de grupos de contactos na funcionalidade de autocomplete"

7.5 Aplicação de Adaptações

Nesta etapa pretende-se apresentar a fase seguinte à tomada de decisão.

Interface Especial para Redes Sociais

Inicialmente, será necessário verificar na tabela *socialnetwork_adaptations* se contém alguma informação do utilizador. Se não possuir, quer dizer que a funcionalidade nunca foi recomendada ao utilizador, nem está de momento. O processo termina logo por aí. Se possuir, o rumo seguido depende do campo *status*. Então, se o *status* for:

- *no* – Significa que a funcionalidade nunca foi recomendada, mas que chegou o momento de recomendar.
- *tried once* – Significa que já se recomendou a funcionalidade uma vez ao utilizador e este rejeitou. Recomendar segunda vez caso já tenha passado pelo menos cinco dias depois da primeira tentativa.
- *tried twice* – Significa que já se recomendou duas vezes a funcionalidade ao utilizador e este rejeitou as duas. Não recomendar mais nenhuma vez.
- *yes* – Significa que a funcionalidade já se encontra aplicada. Não recomendar.
- *cancelled* – Significa que a funcionalidade já foi recomendada, aplicada e cancelada. De momento não possui a funcionalidade. Não recomendar de novo, pois se o utilizador cancelou é porque não ficou satisfeito com a adaptação efectuada.

Em suma, a recomendação é efectuada ao utilizador apenas no caso do *status* ser igual a *no* ou *tried once*. Seguidamente, apresenta-se a resposta do sistema às várias acções do utilizador. Assim:

No caso de aceitar uma recomendação:

- Colocar campo *status* da tabela *socialnetwork_adaptations* a *yes*.
- Incrementar o número de aceitações (actualizar campo *accept* correspondente a esta funcionalidade da tabela *suggestion_history*).

No caso de rejeitar uma recomendação com o campo *status* a *no*:

- Colocar campo *status* da tabela *socialnetwork_adaptations* a *tried once*.

No caso de rejeitar uma recomendação com o campo *status* a *tried once*:

- Colocar campo *status* da tabela *socialnetwork_adaptations* a *tried twice*.
- Incrementar o número de rejeições (actualizar campo *deny* correspondente a esta funcionalidade da tabela *suggestion_history*).
- Processar etapa “Avaliação Automática de Adaptações”.

No caso de cancelar uma adaptação (campo *status* tem que estar *yes*):

- Colocar campo *status* da tabela *socialnetwork_adaptations* a *cancelled*.
- Decrementar o número de aceitações (actualizar campo *accept* correspondente a esta funcionalidade da tabela *suggestion_history*).
- Processar etapa “Avaliação Automática de Adaptações”.

Detecção do Horário de Trabalho

Em relação à funcionalidade “Detecção do Horário de Trabalho”, a implementação desta etapa é muito idêntica à funcionalidade anterior. Em traços gerais, todas as funcionalidades que são recomendadas implementam esta etapa de uma forma muito semelhante. No entanto, surgem algumas diferenças devido ao facto de que é necessário ter em conta para além do utilizador, o IP de acesso. Como é obvio, as tabelas e os campos não se mantêm todos os mesmos. Assim, descreve-se essas pequenas alterações:

- O evento onde é necessário verificar se chegou a altura de realizar a recomendação é quando um utilizador faz *login*, mas unicamente no IP onde se detectou que esta adaptação devia ser aplicada.
- As tabelas (excepto a *suggestion_history*) não são as mesmas da funcionalidade anterior. No entanto, as tabelas usadas são equivalentes em termos de uso (a tabela *worktime* equivale à tabela *socialnetwork*, e a tabela *worktime_adaptations* equivale à tabela *socialnetwork_adaptations*).
- A pesquisa de dados na tabela *worktime_adaptations* é referente a um utilizador e a um IP (não só ao utilizador como acontecia na funcionalidade anterior).
- As actualizações na tabela *suggestion_history* são a respeito desta funcionalidade (campo *adaptation* igual a *worktime*).

Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete

Nesta funcionalidade, a lista de sugestões de grupos de contactos é actualizada a cada inserção de um carácter (sem ser vírgula) no campo de preenchimento de destinatário ou a cada escolha de uma opção dessa lista (inserir enter ou vírgula selecciona a opção da lista que se encontra destacada, ou clicar com o rato na respectiva opção). O método adoptado baseia-se no seguinte:

- Verificar se foi escolhida uma opção da lista de sugestões ou se foi inserido carácter.
- No caso de escolher opção da lista de sugestões:
 - Guardar o grupo de contactos que encontrava na primeira posição da lista apresentada ao utilizador antes da escolha (utilização da tabela *autocomplete_suggestion_temp*).
 - Guardar o grupo de contactos escolhido.

- Preencher campo de destinatário com os contactos referentes ao grupo de contactos que foi seleccionado.
 - Estes contactos são acrescentados aos outros contactos que já se encontravam definidos (contactos que se encontram antes da última vírgula), caso existissem. Por outras palavras, se o utilizador ainda se encontrava a escrever o primeiro contacto de email, a variável que representa o campo de destinatário é completamente substituída pelo grupo seleccionado; se não, é removido à variável o que se encontra escrito depois do último contacto definido (o que se encontra depois da última vírgula), acrescentando os contactos seleccionados.
- No caso de inserir vírgula:
 - E existir lista de sugestões: é escolhida o grupo que se encontrava destacado dos outros. O resto do processo é idêntico ao caso de escolher opção da lista de sugestões.
 - E não existir lista de sugestões: declarar como um contacto de email (nem que seja um contacto que não exista), o que se encontra escrito depois da última vírgula ou caso não existam vírgulas, o que se encontra escrito no campo.
- No caso de inserir um carácter diferente de vírgula:
 - Guardar todos os grupos de contactos a quem o utilizador já enviou emails, acompanhados pelas classificações correspondentes ao IP em que acede (acedendo às tabelas *autocomplete_groups*, *autocomplete_recipients* e *autocomplete_groups_recipients*). Como forma de identificação destes grupos utiliza-se G1.
 - No caso de já existir alguma vírgula:
 - Guardar contactos que aparecem no campo de destinatário até à última vírgula. Como forma de identificação destes contactos utiliza-se C1 .
 - Verificar quais são os grupos de G1 que possuem todos os contactos de C1.
 - Se existir grupos que atendem à verificação:
 - Os grupos são guardados com a respectiva classificação.
 - Para cada um dos grupos guardados no ponto anterior, retira-se os contactos que existem em C1. Assim, fica-se apenas com os contactos que o utilizador ainda não escolheu, nem escreveu (as classificações continuam sempre a acompanhar os respectivos grupos).
 - Ordenar esses grupos por ordem decrescente de classificação. Como forma de identificação, estes grupos ordenados serão referenciados como G2.

- Verificar quais são os grupos de G2 que possui pelo menos um contacto que faz *match* com o que o utilizador escreveu depois da última vírgula. O *match* também é efectuado para os nomes associados a um contacto, e não apenas para o email. Isto é possível recorrendo ao *address book* do respectivo utilizador.
 - Os grupos de G2 que atenderem à verificação são guardados pela mesma ordem (identificação destes grupos: G3). Adiciona-se ainda a G3 os contactos singulares ordenados alfabeticamente, em que: o respectivo email ou nome atribuído fazem *match* com o que se encontra escrito depois da última vírgula, e já não se encontra num dos grupos de G3 que possua apenas um contacto. Esta é a lista de sugestões a ser apresentada ao utilizador (os grupos de G3 devem fazer-se acompanhar pela respectiva classificação, ao contrário dos contactos singulares).
- Se não existir:
 - A lista de sugestões é composta apenas pelos contactos singulares ordenados alfabeticamente, em que: o respectivo email ou nome atribuído fazem *match* com o que se encontra escrito depois da última vírgula.
- No caso de não existir vírgulas:
 - Ordenar os grupos de G1 por ordem decrescente de classificação. Como forma de identificação, estes grupos ordenados serão referenciados como G4.
 - Verificar quais são os grupos de G4 que possui pelo menos um contacto que faz *match* com o que o utilizador escreveu (com o nome ou email).
 - Os grupos de G4 que atenderem à verificação são guardados pela mesma ordem (identificação destes grupos: G5). Adiciona-se ainda a G5 os contactos singulares ordenados alfabeticamente, em que: o respectivo email ou nome atribuído fazem *match* com o que se encontra escrito, e já não se encontra num dos grupos de G5 que possua apenas um contacto. Esta é a lista de sugestões a ser apresentada ao utilizador (os grupos de G5 devem fazer-se acompanhar pela respectiva classificação, ao contrário dos contactos singulares).

7.6 Avaliação Automática de Adaptações

Interface Especial de Redes Sociais

Esta etapa só será necessária nos casos de o utilizador rejeitar uma recomendação com o campo *status a tried once* ou de cancelar uma adaptação (campo *status a yes*). Esta fundamenta-se em:

- Verificar se o número de recomendações ultrapassa as 99.
 - Se sim, é necessário verificar se o número de rejeições equivale a 5% ou mais do número total de recomendações.
 - Se sim, somar 0.5 ao factor de decisão actual (representado pelo campo *var1*).
 - Se não, não actualizar factor de decisão.
 - Se não, não verificar nada.

Deteção do Horário de Trabalho

A única alteração a acrescentar à abordagem demonstrada na funcionalidade anterior, consiste em decrementar ao factor de decisão 0.005 unidades em vez de somar 0.5.

Sugestão de Grupos de Contactos na Funcionalidade de Autocomplete

Quando um email é enviado, é necessário verificar onde é que recaiu a escolha do utilizador. Esta verificação é necessária para averiguar se a ordenação dos contactos sugeridos encontra-se correctamente adaptada ao utilizador e tomar medidas, caso não se confirme esse cenário. Caso tenha recaído sobre:

- A simples escrita dos contactos no campo de preenchimento do destinatário sem recorrer às sugestões:
 - Neste caso nada é efectuado, pois o utilizador não chegou a interagir com a funcionalidade.
- A primeira opção com classificação, também nenhuma medida é tomada pois antecipou-se correctamente o comportamento do utilizador.
- Uma opção sem qualquer tipo de classificação (caso dos contactos que se encontram ordenados por ordem alfabética), nenhuma medida é tomada pois foi escolhido um contacto que não estava presente no histórico do utilizador. Desta forma, o sistema não poderia fazer nada já que se trata de um dado novo, sem antecedentes. No entanto, esta informação é registada podendo ser utilizado da próxima vez.
- Outra opção com classificação, é necessário averiguar onde é que o factor de decisão falhou, usando a fórmula 5.4 e as metodologias apresentadas no subcapítulo 5.2.2.

8 Resultados

Neste capítulo pretende-se apresentar os resultados obtidos no final da fase de implementação para cada uma das funcionalidades desenvolvidas.

Em relação às duas primeiras funcionalidades, em que praticamente todo o processo é realizado em *background*, torna-se difícil retirar resultados conclusivos acerca da abordagem proposta, já que estes dependem:

- De outro sistema que ainda se encontra em fase de construção (nova plataforma de *webmail* da Portugalmail), não permitindo a realização de testes junto dos utilizadores finais. Assim, não foi possível testar a reacção destes tanto às recomendações efectuadas, como às adaptações em si.

No entanto, foram simulados comportamentos de vários tipos de utilizadores (introduzindo informação diversa na base de dados relativa a vários utilizadores), de forma a cobrir todos os casos possíveis. Assim, verificou-se que a implementação foi realizada com o sucesso pretendido, uma vez que o sistema analisava essa informação correctamente e reagia conforme a abordagem proposta no capítulo 5. Alguns exemplos da resposta do sistema, podem ser visualizados no anexo C.

Quanto à terceira funcionalidade, a situação é a mesma. A dependência acima supracitada continua a existir, se pretendermos averiguar se os métodos e fórmulas propostos para a implementação desta funcionalidade, foram os mais adaptados para corresponder a todas as personalidades e comportamentos existentes.

No entanto, devido à componente relativa à etapa “Avaliação automática de adaptações”, foi criada uma interface simples, de forma a verificar a correcção da sua implementação (conforme a proposta de solução).

No anexo B, apresentam-se alguns *screenshots* de um exemplo do funcionamento desta funcionalidade, os quais são descritos aqui neste capítulo. Importante será referir que a interface criada pretendia apenas avaliar o funcionamento, não tendo sido investido muito tempo na sua elaboração, visto que futuramente será necessário criar outra para ser integrada no novo serviço de *webmail*.

Recorrendo às figuras existentes no anexo B, pode-se notar que na lista de sugestões os grupos são precedidos por um número entre parênteses, o qual simboliza a classificação atribuída a esse grupo de contactos.

A figura B.1 diz respeito a quando o utilizador user1@portugalmail.net encontra-se no IP 192.168.100.32 e insere “user” no campo relativo ao destinatário. Depois, o utilizador escolhe o terceiro grupo de contactos da lista de sugestões, o qual possui uma classificação igual a 31, fruto de já ter enviado 31 emails para aquele grupo a partir daquele IP. De seguida, envia o email a esse grupo de contactos e verifica como ficou a lista de sugestões. Repare-se que a classificação do contacto correspondente passou de 31 para 32 como era previsto a partir da fórmula proposta.

A figura B.5 corresponde à consulta da lista de sugestões utilizando no campo de destinatário a mesma expressão “user”, mas a partir de outro IP. Repare-se agora que as classificações são diferentes para os mesmos grupos. De seguida, o utilizador envia o email para o mesmo grupo que mandou anteriormente no outro IP.

A figura B.7 corresponde a quando o utilizador vai consultar novamente a lista de sugestões para a mesma expressão “user” no primeiro IP em que enviou o email. O tal grupo escolhido que se encontrava com uma classificação 32 naquele IP passou a 32.66. Isso deviu-se ao facto do factor de decisão na altura encontrar-se a 0.66 (os emails enviados a partir doutros IPs assumem o valor desse factor).

A figura B.8 representa a escolha do sexto grupo de contactos por parte do utilizador. Este grupo é inserido no campo do destinatário e é introduzida uma vírgula, gerando uma nova lista de sugestões. Note-se que o grupo de contactos que aparecia em primeiro lugar antes, continua a ser o mesmo, devido a que o grupo escolhido é um sub-grupo desse. Perante este cenário, o utilizador escolheu o primeiro grupo de contactos e enviou o email. Na última figura (B.11) pode-se verificar que a classificação do primeiro grupo ficou inalterada após o envio, uma vez que o sistema adaptou-se correctamente ao utilizador.

9 Conclusões

Neste capítulo pretende-se apresentar um resumo do trabalho realizado, concluir acerca da satisfação dos objectivos e dos resultados obtidos a partir da solução proposta, e ainda identificar as direcções para trabalho futuro.

9.1 Sobre o Trabalho Desenvolvido

O problema proposto como tema de projecto, consistia em dotar a nova plataforma de *webmail* da Portugalmail de inteligência, de forma a esta possuir a capacidade de detectar automaticamente padrões de comportamento nas actividades dos utilizadores e disponibilizar ou recomendar certas funcionalidades que possam favorecer o processo de interacção do utilizador com o sistema. Nesse sentido, desenvolveu-se o sistema adaptativo para tornar-se parte integrante dessa plataforma.

A solução proposta para o referido problema foi baseada na arquitectura apresentada em [Paramythis, 2001] para sistemas adaptativos, uma vez que é definida uma abordagem modular, a qual permite ser reutilizada por várias aplicações, providenciando ainda um amplo *feedback* para todo o processo referente à lógica de negócio. Essa arquitectura foi analisada de forma a verificar se era possível a sua integração com a arquitectura existente na plataforma de *webmail*. Verificou-se que a adaptação dessa abordagem foi conseguida, uma vez que as três funcionalidades adaptativas (seleccionadas após intensa fase de concepção de ideias), foram desenvolvidas nesses moldes com o sucesso desejado. Essas três funcionalidades consistem em: detectar se a criação de uma interface nova e apelativa, para apresentar a informação existente nas mensagens de redes sociais pode ser útil para o utilizador; detectar se a ocultação de emails de brincadeira (emails fora do contexto de trabalho) durante o horário de trabalho do utilizador num determinado IP, pode consistir numa mais valia; e sugerir e ordenar os grupos de contactos a quem o utilizador mais emails enviou, quando este preenche o destinatário de um email. As

duas primeiras funcionalidades são recomendadas ao utilizador, enquanto a outra é disponibilizada a todos os utilizadores sem excepção.

Não foram possíveis recolher qualquer tipo de resultados acerca da eficiência dos métodos e algoritmos adoptados junto dos utilizadores, uma vez que o novo serviço de *webmail* encontra-se ainda em fase de desenvolvimento e, conseqüentemente, não se encontra a ser utilizado. Várias baterias de testes foram criadas para verificar se o sistema estava a decidir correctamente a partir dos dados recolhidos. No entanto, é impossível simular o comportamento e personalidade dos diversos utilizadores de forma a prever se estes vão apreciar a adaptação ou aceitar a recomendação. O facto das interfaces das adaptações não estarem efectuadas (as quais não eram objectivo do projecto), constitui outro factor que não permitia testar os efeitos das tomadas de decisão.

9.2 Sobre o Trabalho Futuro

Actualmente, o sistema adapta-se às características do utilizador individual apesar desses utilizadores se poderem enquadrar em diferentes tipos. Futuramente, a classificação dos utilizadores quanto aos tipos em que se inserem constitui uma tarefa interessante para ser levada a cabo. A adopção de um método baseado no conceito implementado pelo algoritmo *Cluster Models* pode ser uma opção para recomendar as mesmas adaptações a utilizadores similares.

A componente respectiva ao desenvolvimento das interfaces das funcionalidades não faz parte integrante dos objectivos do projecto, já que o conceito principal deste baseia-se na detecção das adaptações que podem aumentar a satisfação dos utilizadores. No entanto, o trabalho aqui desenvolvido só vai ter verdadeiro valor para a nova plataforma de *webmail* e principalmente para os utilizadores, quando estes puderem usufruir realmente das adaptações que foram decididas ser recomendadas ou aplicadas. No sentido de criar as interfaces será necessário discutir com o elemento responsável pelo *design* da equipa técnica da Portugalmail, de forma a estas adaptações encaixarem-se da melhor maneira no visual e estilo existente nas aplicações disponibilizadas no serviço de *webmail*.

Após o serviço de *webmail* ser lançado para o mercado e as interfaces das adaptações estarem completamente concluídas, deve-se então avançar para a avaliação do sistema desenvolvido junto dos utilizadores, e implementar as devidas alterações (caso necessário).

Bibliografia

- [ArdissonoConsole, 2001] Liliana Ardissono, Luca Console e Ilaria Torre. An Adaptive System for the Personalized Access to News. *AI Communications*. Vol. 14. Nº 3. Pp. 129-147. IOS Press. 2001.
- [AvilaCosta, 2009] Felipe Ávila da Costa. New Paradigm of Webmail Interfaces. Março 2009.
- [Balabanovic, 1997] Marko Balabanovic. An Adaptive Web Page Recommendation Service. *AGENTS '97: Proceedings of the first international conference on Autonomous agents*. Pp. 378-385. 1997.
- [BorkurZwol, 2008] Borkur Sigurbjornsson e Roelof van Zwol. Flickr Tag Recommendation Based on Collective Knowledge. *WWW '08: Proceeding of the 17th international conference on World Wide Web*. 327-336. ACM. 2008.
- [Brusilovsky, 1996] Peter Brusilovsky. *User Modeling and User Adapted Interaction*. 1996.
- [Brusilovsky, 2002a] Peter Brusilovsky, Charalampos Karagiannidis e Demetrios Sampson. The Benefits of Layered Evaluation of Adaptive Applications and Services. 2002.
- [Brusilovsky, 2002b] Peter Brusilovsky e Mark T. Maybury. From Adaptive Hypermedia to the Adaptive Web. *Communications of the ACM*. Vol. 45. Nº 5. Pp. 30-33. 2002.
- [CaoKirshner, 2001] Li Cao, Michael Kirshner, Matthew Tevenan e Carolyn Wei. *Adaptivity*. 2001. Disponível em <http://eserver.org/courses/s01/tc510/adaptivity/index.html>, acessado a última vez em 4 de Maio de 2009.

- [ChenChen, 2001] Hung-Chen Chen e Arbee L. P. Chen. A Music Recommendation System Based on Music Data Grouping and User Interests. CIKM '01: Proceedings of the Tenth International Conference on Information and Knowledge Management. Pp. 231-238. ACM. 2001
- [CrawfordKay, 2002] Elisabeth Crawford, Judy Kay e Eric McCreath. An Intelligent Interface for Sorting Electronic Mail. 2002.
- [Dacor, 2009] Art Dacor. Two Types of Web Users. 2009. Disponível em <http://www.great-web-design-tips.com/web-site-design/150.html>, acessido a última vez em 4 de Maio de 2009.
- [DredzeBlitzer, 2005] Mark Dredze, John Blitzer e Fernando Pereira. Reply Expectation Prediction for Email Management. CEAS. 2005.
- [DredzeBlitzer, 2006] Mark Dredze, John Blitzer, e Fernando Pereira. "Sorry, I Forgot the Attachment:" Email Attachment Prediction. CEAS. 2006.
- [DredzeBrooks, 2008] Mark Dredze, Tova Brooks, Josh Carroll, Joshua Magarick, John Blitzer e Fernando Pereira. Intelligent Email: Reply and Attachment Prediction. IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces. 321-324. 2008.
- [DredzeWallach, 2008] Mark Dredze, Hanna M. Wallach, Danny Puller, Tova Brooks, Josh Carroll, Joshua Magarick, John Blitzer e Fernando Pereira. Intelligent Email: Aiding Users with AI. AAAI. Pp. 1524-1527. 2008.
- [Eugénio, 2009] Eugénio Oliveira. Inteligência Artificial: Aprendizagem Simbólica Automática - MIEIC - FEUP. Disponível em http://www.fe.up.pt/~eol/IA/IA0809/APONTAMENTOS/6_ASA.pdf, acessido a última vez em 4 de Maio de 2009.
- [FEUProposta, 2009] FEUP. Proposta de Projecto/Dissertação - Adaptação da interface de webmail às actividades de cada utilizador. 2009. Disponível em https://www.fe.up.pt/si/estagios_empresas.ver_dados_proposta?p_id=25135, acessido a última vez em 4 de Maio de 2009.
- [GalvãoDias, 2009] Teresa Galvão Dias. Investigação Operacional: Teoria da Decisão - MIEIC - FEUP. Disponível em https://www.fe.up.pt/si/conteudos_service.conteudos_cont?pct_id=45116&pv_cod=30iLDJJjWrFH, acessido a última vez em 4 de Maio de 2009.
- [Gasparini, 2003] Isabela Gasparini. Interface Adaptativa no ambiente AdaptWeb: navegação e apresentação adaptativa baseada no modelo do usuário. Junho 2003.

- [Gilliland, 2008] Art Gilliland. Intelligent E-Mail Archiving: Classification, Filtering, Retention and Discovery. Novembro 2008.
- [Horde, 2009] Horde. Create. Communicate. Collaborate. 2009. Disponível em <http://www.horde.org/about/>, acessado a última vez em 22 de Junho de 2009.
- [JaschkeMarinho, 2007] Robert Jaschke, Leandro Marinho, Andreas Hotho, Lars Schmidt-Thieme e Gerd Stumme. Tag Recommendations in Folksonomies. Knowledge Discovery in Databases: PKDD 2007. Pp. 506-514. 2007.
- [Krishnakumar, 2007] Anita Krishnakumar. Recoo: A Recommendation System for Youtube RSS Feeds. 2007.
- [Langley, 1999] Pat Langley. User Modeling in Adaptive Interfaces. UM '99: Proceedings of the seventh international conference on User modeling. Pp. 357-370. 1999.
- [LangleyFehling] Pat Langley e Michael Fehling. The Experimental Study of Adaptive User Interfaces. 1998.
- [LindenSmith, 2003] Greg Linden, Brent Smith e Jeremy York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. Internet Computing, IEEE. Vol. 7. Nº 1. Pp. 76-80. 2003.
- [MancoMasciari, 2002] Giuseppe Manco, Elio Masciari, Massimo Ruffolo e Andrea Tagarelli. Towards An Adaptive Mail Classifier. In Italian Association for Artificial Intelligence Workshop Su Apprendimento Automatico: Metodi Ed Applicazioni. 2002.
- [Moore, 2009] Andrew Moore. Statistical Data Mining Tutorials. 2009. Disponível em <http://www.autonlab.org/tutorials/>, acessado a última vez em 4 de Maio de 2009.
- [Paramythis, 2001] A. Paramythis, A. Totter e C. Stephanidis. A modular approach to the evaluation of Adaptive User Interfaces. 2001.
- [PazzaniBillsus, 2007] Michael J. Pazzani e Daniel Billsus. Content-Based Recommendation Systems. The Adaptive Web. Pp. 325-341. 2007.
- [PazzaniBillsus, 2009] Michael Pazzani, Jack Muramatsu e Daniel Billsus. Syskill & Webert: Identifying interesting web sites. Disponível em <http://www.ics.uci.edu/~pazzani/RTF/AAAI.html>, acessado a última vez em 4 de Maio de 2009.
- [PengSilver, 2007] Xiaoyan Peng e Daniel L. Silver. Interface Adaptation based on User Expectation. AINAW '07: Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops. Pp. 264-269. 2007.

- [PerkowitzEtzioni, 1997] Mike Perkowitz e Oren Etzioni. Adaptive Web Sites: an AI Challenge. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence. Pp. 16-23. 1997.
- [PerkowitzEtzioni, 1999] Mike Perkowitz e Oren Etzioni. Towards adaptive Web sites: Conceptual framework and case study. Artificial Intelligence. Vol. 118. Pp. 245-275. 1999.
- [PHPGroup, 2009] The PHP Group. PHP: Hypertext Preprocessor. 2009. Disponível em <http://php.net/>, acessido a última vez em 4 de Maio de 2009.
- [Portugalmail.net, 2009] Portugalmail - Comunicações, S.A.. Portugalmail.net - plataformas de inovação. Disponível em <http://www.portugalmail.net/>, acessido a última vez em 4 de Maio de 2009.
- [Portugalmail.pt, 2009] Portugalmail - Comunicações, S.A.. Portugalmail.pt. 2004. Disponível em <http://www.portugalmail.pt/>, acessido a última vez em 4 de Maio de 2009.
- [Postgresql, 2009] PostgreSQL Global Development Group. PostgreSQL: The world's most advanced open source database. 2009. Disponível em <http://www.postgresql.org/>, acessido a última vez em 4 de Maio de 2009.
- [RuckerPolanco, 2005] James Rucker e Marcos J. Polanco. Siteeer: Personalized Navigation for the Web. Communications of the ACM. Vol. 40. Nº 3. Pp. 73-76. ACM. 1997.
- [W3SchoolsAJAX, 2009] W3Schools. AJAX Tutorial. 2009. Disponível em <http://www.w3schools.com/Ajax/Default.Asp>, acessido a última vez em 4 de Maio de 2009.
- [W3SchoolsPHP, 2009] W3Schools. PHP Tutorial. 2009. Disponível em <http://www.w3schools.com/PHP/DEfaULT.asP>, acessido a última vez em 4 de Maio de 2009.
- [WikiAJAXen, 2009] Wikipedia. Ajax (programming). Disponível em [http://en.wikipedia.org/wiki/Ajax_\(programming\)](http://en.wikipedia.org/wiki/Ajax_(programming)), acessido a última vez em 4 de Maio de 2009.
- [WikiAJAXpt, 2009] Wikipedia. AJAX (programação). Disponível em [http://pt.wikipedia.org/wiki/AJAX_\(programação\)](http://pt.wikipedia.org/wiki/AJAX_(programação)), acessido a última vez em 4 de Maio de 2009.
- [WikiClusterAnalysis, 2009] Wikipedia. Cluster analysis. Disponível em http://en.wikipedia.org/wiki/Data_clustering, acessido a última vez em 9 de Junho de 2009.

- [WikiCSSen, 2009] Wikipedia. Cascading Style Sheets. Disponível em http://en.wikipedia.org/wiki/Cascading_Style_Sheets, acessido a última vez em 15 de Junho de 2009.
- [WikiCSSpt, 2009] Wikipedia. Cascading Style Sheets. Disponível em http://pt.wikipedia.org/wiki/Cascading_Style_Sheets, acessido a última vez em 15 de Junho de 2009.
- [WikiHierarchical, 2009] Wikipedia. Hierarchical clustering. Disponível em http://en.wikipedia.org/wiki/Hierarchical_clustering, acessido a última vez em 9 de Junho de 2009.
- [WikiHordeen, 2009] Wikipedia. Horde (software). Disponível em [http://en.wikipedia.org/wiki/Horde_\(software\)](http://en.wikipedia.org/wiki/Horde_(software)), acessido a última vez em 22 de Junho de 2009.
- [WikiHTMLen, 2009] Wikipedia. HTML. Disponível em <http://en.wikipedia.org/wiki/HTML>, acessido a última vez em 14 de Junho de 2009.
- [WikiHTMLpt, 2009] Wikipedia. HTML. Disponível em <http://pt.wikipedia.org/wiki/HTML>, acessido a última vez em 14 de Junho de 2009.
- [WikiKmeans, 2009] Wikipedia. k-means clustering. Disponível em http://en.wikipedia.org/wiki/K-means_clustering, acessido a última vez em 9 de Junho de 2009.
- [WikiPHPen, 2009] Wikipedia. PHP. Disponível em <http://en.wikipedia.org/wiki/PHP>, acessido a última vez em 4 de Maio de 2009.
- [WikiPHPpt, 2009] Wikipedia. PHP. Disponível em <http://pt.wikipedia.org/wiki/PHP>, acessido a última vez em 4 de Maio de 2009.
- [WikiPostgresqlen, 2009] Wikipedia. PostgreSQL. Disponível em <http://en.wikipedia.org/wiki/PostgreSQL>, acessido a última vez em 4 de Maio de 2009.
- [WikiPostgresqlpt, 2009] Wikipedia. PostgreSQL. Disponível em <http://pt.wikipedia.org/wiki/PostgreSQL>, acessido a última vez em 4 de Maio de 2009.
- [WikiSQLen, 2009] Wikipedia. SQL. Disponível em <http://en.wikipedia.org/wiki/SQL>, acessido a última vez em 22 de Junho de 2009.
- [WikiSQLpt, 2009] Wikipedia. SQL. Disponível em <http://pt.wikipedia.org/wiki/SQL>, acessido a última vez em 22 de Junho de 2009.
- [WikiTdidf, 2009] Wikipedia. Tf-idf. Disponível em <http://en.wikipedia.org/wiki/Tf-idf>, acessido a última vez em 4 de Maio de 2009.

- [WikiWikien, 2009] Wikipedia. Wiki. Disponível em <http://en.wikipedia.org/wiki/Wiki>, acessido a última vez em 14 de Junho de 2009.
- [WikiWikip, 2009] Wikipedia. Wiki. Disponível em <http://pt.wikipedia.org/wiki/Wiki>, acessido a última vez em 14 de Junho de 2009.
- [WoodenRobot, 2009] Wooden-Robot.net. "Taking Email to Task" - PARC HCI research. Disponível em <http://wooden-robot.net/2007/11/03/taking-email-to-task-parc-hci-research/>, acessido a última vez em 4 de Maio de 2009.

Anexo A: Planeamento do Projecto

O planeamento e organização de todas as etapas envolventes no projecto foi realizado numa fase inicial, logo após ao estudo e análise de toda a envolvente do projecto e definição das ideias a implementar. O planeamento do projecto foi discutido com a equipa de desenvolvedores da Portugalmail de forma a estabelecer metas e estimativas mais seguras do tempo que seria necessário dispensar para cada etapa do projecto.

Planear o projecto revela-se de extrema importância devido ao tempo limitado de 20 semanas imposto pela Faculdade de Engenharia e Universidade do Porto para o seu desenvolvimento. Um plano geral das actividades torna-se assim bastante útil para se saber se uma etapa está a demorar mais ou menos tempo que o previsto e, possivelmente, necessário para a sua realização e também para se saber se é necessário efectuar ajustes em alguma etapa.

Tabela 6: Planeamento do projecto

Data de Início	Nome da Fase	Descrição
25 Fevereiro 2009	Estado da arte	Pesquisa e análise do estado da arte sobre sistemas adaptativos e serviços de <i>webmail</i> relativamente a adaptação automática às actividades de cada utilizador.
19 Março 2009	<i>Brainstorming</i>	Sessões de <i>brainstorming</i> (do projecto abordado neste relatório e dos projectos dos colegas que se encontram a estagiar na Portugalmail) com o objectivo de definir correctamente o caminho a seguir, discutir ideias e gerar outras novas. Organização de toda a informação.
2 Abril 2009	Arquitectura e Tecnologias	Estudo e definição da arquitectura a usar. Estudo das tecnologias necessárias para o desenvolvimento das

		funcionalidades.
13 Abril 2009	Seleccção de ideias	Seleccção das ideias mais importantes a serem implementadas durante o projecto e após a conclusão deste.
15 Abril 2009	Implementação	Implementação das funcionalidades escolhidas.
1 Junho 2009	Validação, testes e alterações	Verificação se as funcionalidades encontram-se correctamente implementadas, simulando a utilização do serviço de <i>webmail</i> com a introdução de dados na base de dados. Implementação das modificações necessárias baseadas nos resultados obtidos pelos testes realizados.
8 Junho 2009	Relatório	Escrita e revisão do relatório.
29 Junho 2009	Entrega relatório	Entrega do relatório e fim do projecto.

Anexo B: Funcionamento da sugestão de grupos de contactos

USER: user1@portugalmail.net ||| IP: 192.168.100.32



Figura B.1: Inserir “user” e escolher terceira opção no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.32

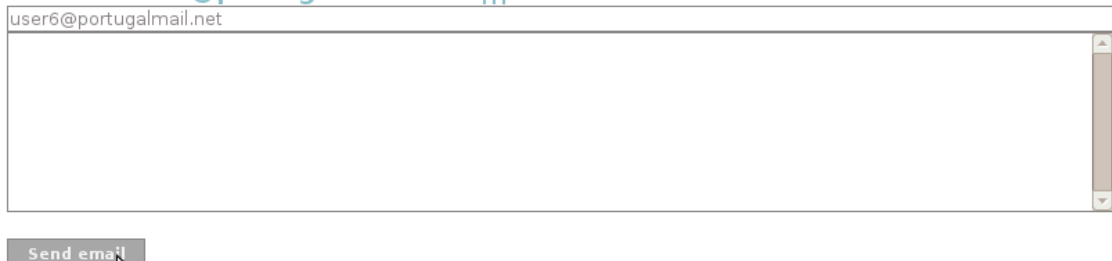


Figura B.2: Enviar email para user6@portugalmail.net no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.32

```
user
(109.37) "2user4" <user4@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@portugalmail.net>
(39.72) "2user3" <user3@portugalmail.net>
(32) "uuser6" <user6@portugalmail.net>
(6) "2user5" <user5@portugalmail.net>, "2user4" <user4@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@
(5.67) "2user5" <user5@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@portugalmail.net>
(1) "2user4" <user4@portugalmail.net>
(1) "7user7" <user7@portugalmail.net>
(1) "uuser8" <user8@portugalmail.net>
"2user2" <user2@portugalmail.net>
"2user5" <user5@portugalmail.net>
```

Send email

Figura B.3: Consultar lista de sugestões para "user" no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.33

```
user
(78.35) "2user4" <user4@portugalmail.net>, "2user3" <user3@portugalmail.net>, "2user2" <user2@portugalmail.net>
(34.11) "2user3" <user3@portugalmail.net>
(21.44) "uuser6" <user6@portugalmail.net>
(4.02) "2user5" <user5@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@portugalmail.net>
(0.67) "7user7" <user7@portugalmail.net>
(0.67) "uuser8" <user8@portugalmail.net>
(0.67) "2user4" <user4@portugalmail.net>
"2user2" <user2@portugalmail.net>
"2user5" <user5@portugalmail.net>
```

Send email

Figura B.4: Inserir "user" no IP 192.168.100.33

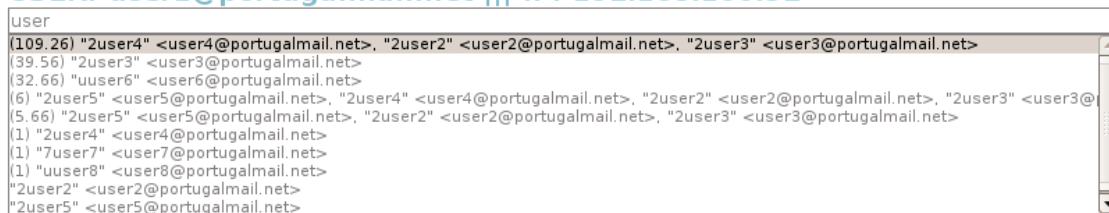
USER: user1@portugalmail.net ||| IP: 192.168.100.33

```
user6@portugalmail.net
```

Send email

Figura B.5: Enviar email a user6@portugalmail.net no IP 192.168.100.33

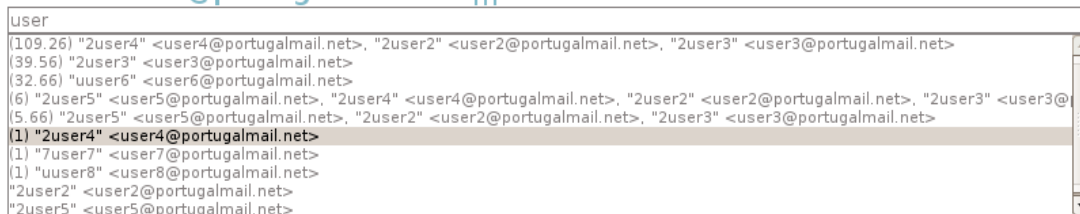
USER: user1@portugalmail.net ||| IP: 192.168.100.32



Send email

Figura B.6: Consultar lista de sugestões para "user" no IP 192.168.100.32

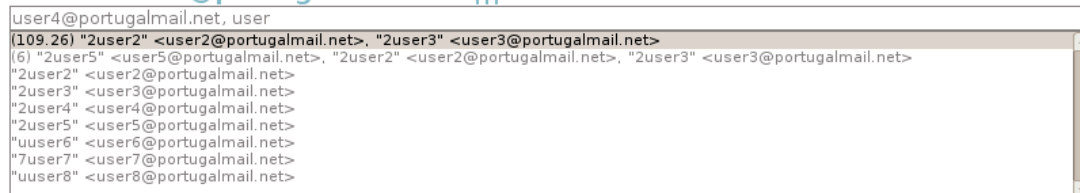
USER: user1@portugalmail.net ||| IP: 192.168.100.32



Send email

Figura B.7: Escolher sexta opção no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.32



Send email

Figura B.8: Escolher primeira opção para "user4@portugalmail.net, user" no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.32

```
user4@portugalmail.net, user2@portugalmail.net, user3@portugalmail.net
```

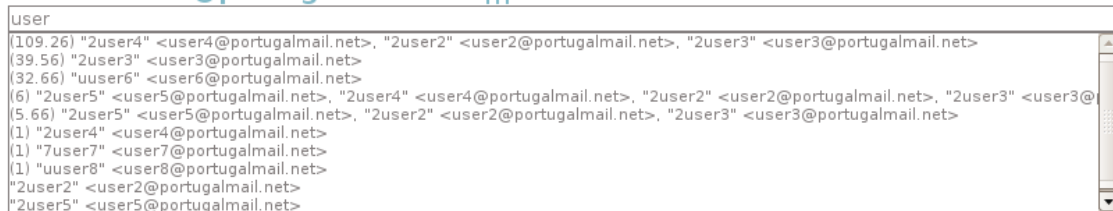


Send email

Figura B.9: Enviar email tendo escolhido a primeira opção no IP 192.168.100.32

USER: user1@portugalmail.net ||| IP: 192.168.100.32

```
user
(109.26) "2user4" <user4@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@portugalmail.net>
(39.56) "2user3" <user3@portugalmail.net>
(32.66) "uuser6" <user6@portugalmail.net>
(6) "2user5" <user5@portugalmail.net>, "2user4" <user4@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@
(5.66) "2user5" <user5@portugalmail.net>, "2user2" <user2@portugalmail.net>, "2user3" <user3@portugalmail.net>
(1) "2user4" <user4@portugalmail.net>
(1) "7user7" <user7@portugalmail.net>
(1) "uuser8" <user8@portugalmail.net>
"2user2" <user2@portugalmail.net>
"2user5" <user5@portugalmail.net>
```



Send email

Figura B.10: Consultar lista de sugestões para "user" no IP 192.168.100.32

Anexo C: Exemplos de resposta do sistema

Interface Especial para Redes Sociais

USER: user1@example.com
FROM: info@hi5.com
DATE: Mon, 29 Jun 2009 10:19:14 -0700

FACTOR: 70.5
NUMBER of emails of social networks last 40 days: ≥ 70.5
DECISION: APPLY adaptation.

Figura C.1: Recomendar "Interface especial para redes sociais"

A figura C.1 representa a decisão do sistema após o utilizador user1@example.com receber um email do hi5 na data de 29 Junho de 2009.

USER: user5@example.com
FROM: info@hi5.com
DATE: Mon, 29 Jun 2009 18:19:14 -0700

FACTOR: 70.5
NUMBER of emails of social networks last 40 days: 4
DECISION: NOT apply adaptation.

Figura C.2: Não recomendar "Interface especial para redes sociais"

A figura C.2 representa a decisão do sistema após o utilizador user5@example.com receber um email do hi5 na data de 29 Junho de 2009.

Detecção do Horário de Trabalho

user5@portugalmail.net

IP	Number of work emails
192.168.100.33	30
IP	Number of work emails
192.168.100.35	34
Number of fun emails (any place): 71	

IP: 192.168.100.35
Insufficient user data to make a decision.

IP: 192.168.100.33
Insufficient user data to make a decision.

Figura C.3: Resposta do sistema a informar que ainda não tem informação suficiente do utilizador

O sistema informa que detectou dois IPs onde o utilizador user5@portugalmail.net recebeu emails de trabalho. No entanto, estes não são suficientes para retirar conclusões com segurança.

user5@portugalmail.net

IP	Number of work emails
192.168.100.32	11
IP	Number of work emails
192.168.100.33	76
IP	Number of work emails
192.168.100.35	75
Number of fun emails (any place): 84	

IP: 192.168.100.33
Number of work emails: 76
Number of fun emails: 29

DAYS OF WEEK: Tue Mon
MIN TIME WORK: 09:39:34
MAX TIME WORK: 17:46:34

Number of fun emails during working hours: 20

CONCLUSION:
READ fun emails on work place AND during working hours.

RESULTADO:
-4

IP: 192.168.100.35
Number of work emails: 75
Number of fun emails: 16

DAYS OF WEEK: Tue Fri Wed Thu
MIN TIME WORK: 11:32:34
MAX TIME WORK: 19:21:34
INTERVAL: 13:13:34 - 13:39:34

Number of fun emails on interval: 16

CONCLUSION:
READ fun emails on work place BUT not during working hours (interval).

RESULTADO:
21

Figura C.4: Recomendar a funcionalidade ao utilizador
user5@portugalmail.net no IP 192.168.100.35

O sistema detectou que utilizador user5@portugalmail.net possui 76 emails de trabalho e 29 emails de brincadeira lidos no IP 192.168.100.33. Desses emails de brincadeira lidos naquele IP, só 20 é que são durante o horário de trabalho detectado (09:39:34 – 17:46:34). No entanto, este número revela-se muito grande para aplicar a adaptação (resultado da fórmula 5.1 deu menor que zero).

Em relação ao IP 192.168.100.35, o utilizador possui 75 emails de trabalho e 16 emails de brincadeira lidos. Estes emails de brincadeira são lidos apenas durante um período pequeno,

onde não é costume ler emails de trabalho. Assim, não contam para os emails lidos durante o horário de trabalho (11:32:34 – 19:21:34) com intervalo a 13:13:34 – 13:39:34. O resultado é maior do que zero e a funcionalidade deve ser recomendada.