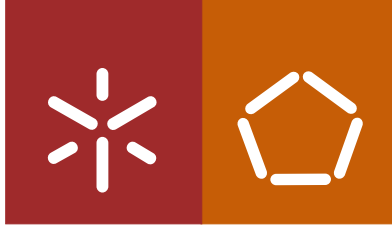


**Universidade do Minho**  
Escola de Engenharia

Luana Geórgia Lopes Telha

***Smart Targeting* de Conteúdos para  
Fidelizações e Ofertas**



**Universidade do Minho**  
Escola de Engenharia

Luana Geórgia Lopes Telha

***Smart Targeting* de Conteúdos para  
Fidelizações e Ofertas**

Dissertação de Mestrado  
Mestrado em Engenharia de Sistemas

Trabalho efetuado sob a orientação do  
**Professor Doutor Paulo Sérgio Lima Pereira Afonso**  
e do  
**Professor Doutor João Miguel Fernandes**

outubro de 2014

## DECLARAÇÃO

Nome: Luana Geórgia Lopes Telha

Endereço eletrónico: luana.telha@gmail.com      Telefone: 919812332

Cartão do Cidadão: 13421094 8 ZZ7

Título da dissertação: *Smart Targeting* de Conteúdos para Fidelizações e Ofertas

Orientadores:

Professor Doutor Paulo Sérgio Lima Pereira Afonso

Professor Doutor João Miguel Fernandes

Ano de conclusão: 2014

Mestrado em Engenharia de Sistemas

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, \_\_\_\_/\_\_\_\_/\_\_\_\_

Assinatura:

## **AGRADECIMENTOS**

Em primeiro lugar gostaria de agradecer à empresa, de uma maneira geral, por ter proporcionado a oportunidade de colaborar neste projeto, onde aprendi o essencial sobre trabalhar numa empresa com uma equipa séria e responsável. Gostaria de agradecer em particular ao meu orientador na empresa, Engenheiro Gonçalo Almeida, que sempre se mostrou disposto a ajudar.

Aos meus orientadores académicos, Professor Paulo Afonso e Professor João Fernandes, pela disponibilidade, compreensão, conhecimento e profissionalismo.

À minha família pela paciência e pelo apoio dado durante todo o meu percurso académico.

Aos meus amigos que me acompanharam durante a elaboração desta dissertação, em especial ao Ricardo Baptista.



## RESUMO

Os utilizadores de serviços informáticos recebem diariamente uma quantidade significativa e crescente de conteúdos não solicitados e muitas vezes não desejados. Este SPAM acontece através de diversos canais de comunicação, tais como: correio eletrónico, redes sociais, *web sites* ou aplicações móveis (via publicidade). Existem vários fatores que podem ajudar a melhorar o direcionamento destes conteúdos para utilizadores potencialmente interessados, tais como: dados demográficos e áreas de interesse obtidos através do perfil do utilizador, dados geográficos obtidos através das coordenadas GPS do *smartphone*, padrões de utilização da aplicação (que conteúdos o utilizador costuma visualizar ou apagar, etc.), entre outros.

É neste contexto que surge este estudo sobre Sistemas de Recomendação de Conteúdos os quais são conjuntos de técnicas que ajudam a filtrar conteúdos tendo em consideração as preferências do utilizador. As técnicas mais utilizadas nestes sistemas são técnicas de Filtragem Híbrida, as quais combinam a Filtragem Colaborativa com a Filtragem Baseada em Conteúdos. O propósito deste projeto focou-se no desenvolvimento de um mecanismo de recomendações de ofertas para uma aplicação específica: a CardMobili.

Primeiramente foi selecionado o algoritmo de Filtragem Híbrida a ser implementado no mecanismo. Na segunda parte do projeto procedeu-se à concepção de um mecanismo adequado à aplicação, para o efeito de recomendações. A aplicação foi testada e foi efetuada a análise crítica das suas vantagens e dos seus pontos fracos. O mecanismo mostrou-se eficiente e satisfatório considerando os resultados obtidos. Porém, o mecanismo poderá ser aperfeiçoado se se proceder à adaptação da interface da aplicação. Foram também identificadas algumas possíveis melhorias tanto na eficácia de apresentação dos resultados, como no enriquecimento dos dados disponíveis para serem processados pelos algoritmos de filtragem.

Palavras-Chave: Sistemas de Recomendação, Filtragem Colaborativa, Filtragem Baseada em Conteúdos, Filtragem Híbrida, Cupões de Oferta, Fidelização



## **ABSTRACT**

Daily, IT services users receive a significant and growing amount of content unsolicited and often unwanted. This SPAM happens through many communication channels, such as email, social networks, web sites and mobile applications (via advertising). There are several factors that can help to improve the targeting of these contents to potentially interested users, such as demographics and areas of interest obtained from the user profile, geographic data obtained from the GPS coordinates of smartphone, usage patterns of the application (which content the user usually view or delete, etc.), among others.

It is in this context that arises this study on Recommender Systems which are techniques that help to filter content taking user preferences into account. In such systems, the most used filtering techniques are hybrid techniques that combine collaborative filtering and content based filtering. The purpose of this project focused on developing an offers recommendations mechanism for specific app: CardMobili.

Firstly, a hybrid filtering algorithm was selected to be implemented in the mechanism. The second part of this project was devoted to the design of an appropriate mechanism to the application, for recommendations effect. The application was tested and a critical analysis of their strengths and their weaknesses has been made. The mechanism proved to be efficient and satisfactory accordingly the obtained results. However the mechanism can be improved with a better preparation of the application interface. It was also identified some possible improvements in both efficiency of reporting and the enrichment of the available data to be processed by the filtering algorithms.

KEYWORDS: RECOMMENDER SYSTEMS, COLLABORATIVE FILTERING, CONTENT-BASED FILTERING, HYBRID FILTERING, CUPPONS, LOYALTY





## ÍNDICE

Agradecimentos .....	iii
Resumo .....	v
Abstract .....	vii
Lista de Figuras .....	xi
Lista de Tabelas .....	xiii
Lista de Abreviaturas, Siglas e Acrónimos .....	xv
1. Introdução .....	1
1.1 Enquadramento .....	1
1.2 Empresa .....	2
1.3 Objetivos .....	3
1.4 Estrutura da Dissertação .....	3
2. Revisão da Literatura .....	5
2.1 Customer Reliatioshiop Management (CRM) .....	5
2.2 Carteira Digital .....	6
2.3 Sistemas de Recomendação .....	10
2.3.1. Tarefas de um Sistema de Recomendação .....	11
2.3.2. Colheita de Dados .....	14
2.3.3. Privacidade em Sistemas de Recomendação .....	16
2.3.4. Tipos de Sistemas de Recomendação .....	16
2.3.4.1. Filtragem Colaborativa .....	17
2.3.4.2. Filtragem Baseada em Conteúdos .....	21
2.3.4.3. Filtragem Híbrida .....	26
2.3.5. Limitações dos Sistemas de Recomendação .....	27
2.3.6. Recomendações Móveis .....	28
3. Metodologia e Tecnologia Utilizadas .....	30
3.1 SCRUM .....	30
3.2 Ferramentas Utilizadas .....	32
3.3 Algoritmos Escolhidos .....	33
4. Caso de Estudo: CardMobili .....	35

4.1	A Empresa .....	35
4.2	Produtos .....	35
4.3	Soluções .....	38
4.4	Tecnologia.....	41
5.	Desenvolvimento do Mecanismo de Recomendação.....	43
5.1	Especificação Funcional.....	43
5.1.1.	Requisitos Funcionais .....	43
5.1.2.	Casos de Uso .....	44
5.1.3.	Diagramas de Sequência.....	46
5.2	Especificação Técnica.....	47
5.2.1.	Arquitetura do Sistema.....	47
4.2.1.	Modelo de Dados .....	49
5.3	Implementação .....	50
5.3.1.	Novas Entidades .....	51
5.3.2.	Triggers .....	51
5.3.3.	Jobs .....	52
5.3.4.	Entidades e Classes Alteradas.....	53
5.4	Fluxo da Implementação.....	55
6.	Análise e Discussão dos Resultados .....	56
6.1	Testes Intermédios .....	56
6.2	Teste final .....	57
7.	Conclusões .....	66
	Bibliografia .....	69
	Anexo I – Modelo de Negócio da CardMobili .....	75
	Anexo II – Cálculo da Similaridade .....	79
	Anexo III – Trigger.....	81
	Anexo IV – Job de Inserção .....	83
	Anexo V – Job de Atualização.....	85
	Anexo VI – Cálculo de Predição .....	87

## LISTA DE FIGURAS

Figura 1 - Lista de artistas relacionados na aplicação Spotify .....	12
Figura 2 - Recomendações baseadas numa pesquisa no site Amazon.com .....	13
Figura 3 - Vídeos Recomendados pelo Youtube .....	14
Figura 4 - Colheita de dados aplicando uma abordagem explícita no site da Cardmobili.....	15
Figura 5 - Representação do processo Scrum .....	31
Figura 6 - Ilustração do algoritmo de filtragem híbrida.....	34
Figura 8 - Caso de Uso para Recomendações de Ofertas .....	44
Figura 9 - Caso de Uso para Ofertas Relacionadas.....	45
Figura 10 - Fluxo de chamadas efetuadas para apresentar o Top de Ofertas .....	46
Figura 11 - Fluxo de chamadas efetuadas quando se visualiza uma oferta .....	47
Figura 12 - Arquitetura do mecanismo desenvolvido .....	48
Figura 13 - Modelo de Base de Dados do mecanismo desenvolvido .....	49
Figura 14 - Fluxo da Implementação do mecanismo de recomendação .....	55
Figura 15 - Top de Ofertas para um novo utilizador.....	57



## LISTA DE TABELAS

Tabela 1 - Tempos de resposta das diferentes tarefas do sistema Pandora Internet Music.....	26
Tabela 2 - Descrição do caso de uso "Recomendações de Ofertas".....	44
Tabela 3 - Descrição do caso de uso "Ofertas Relacionadas" .....	45
Tabela 4 - Descrição dos campos da tabela Interest da base de dados.....	49
Tabela 5 - Descrição dos campos pertencentes à tabela UsersSimilarity da base de dados.....	50
Tabela 6 - Perfis fictícios e similaridades .....	58
Tabela 7 - Resultados do Utilizador 1.....	59
Tabela 8 - Resultados do Utilizador 2.....	60
Tabela 9 - Resultados do Utilizador 3.....	61
Tabela 10 - Resultados do Utilizador 4.....	62
Tabela 11 - Resultados do Utilizador 5.....	63
Tabela 12 - Resultados do Utilizador 6.....	64



## **LISTA DE ABREVIATURAS, SIGLAS E ACRÓNIMOS**

API – Application Programming Interface

BT – Behavioral Targeting

CBF – Content-Based Filtering

CF – Collaborative Filtering

CRM – Customer Relationship Management

HQL – Hibernate Query Language

IAB - Interactive Advertising Bureau

IDE - Integrated Development Environment

MMA - Mobile Marketing Association

ORM - Object-Relational Mapping

POS – Point of Sale

QR – Quick Response

SIG – Sistemas de Informação de Gestão

SMS – Short Message Service

SPAM – Sending and Posting Advertisement in Mass

SQL - Structured Query Language

SR – Sistema de Recomendação

SVN - Apache Subversion





# 1. INTRODUÇÃO

Neste capítulo é apresentado o enquadramento da presente dissertação, no âmbito da distribuição de conteúdos da aplicação CardMobili, identificando o problema de investigação. Seguidamente, é apresentada a CardMobili, empresa na qual foi realizado o projeto de investigação, explicando-se sucintamente quais são os seus produtos e soluções. Por fim, são definidos os principais objetivos deste projeto e apresenta-se uma descrição geral dos conteúdos deste documento.

## 1.1 Enquadramento

A crescente utilização dos meios informáticos e da internet para realizar as tarefas básicas do nosso quotidiano banalizou as compras *online*. Cada vez mais os utilizadores destes serviços fazem compras *online*, o que leva ao aumento da publicidade nestes meios. No entanto, por vezes, este tipo de publicidade é realizada sem ter em conta o público ao qual é endereçada a divulgação do produto ou serviço. O envio para os utilizadores de publicidade ou informações com conteúdos não solicitados ou mesmo não desejados, denomina-se *Sending and Posting Advertisement in Mass* (SPAM)<sup>1</sup>. Diária e sucessivamente é enviado SPAM para os clientes através dos mais variados canais de comunicação, tais como: correio eletrónico, *web sites*, redes sociais, *Short Message Service* (SMS) e publicidade em aplicações móveis. Este fenómeno, além de não servir os reais utilizadores e potenciais consumidores destes serviços, não serve os propósitos dos anunciantes e das suas estratégias de comunicação, podendo até ser contraproducente ou mesmo prejudicar a imagem dos produtos ou serviços que se pretendem promover, afetando, em última instância, negativamente o negócio.

Assim surge a necessidade de ferramentas inteligentes que de uma forma eficiente permitam ao utilizador encontrar o conteúdo mais interessante para ele em contextos *online*. Desta necessidade nasceram os Sistemas de Recomendação, que são conjuntos de técnicas que ajudam a filtrar a informação mediante o perfil do utilizador.

A probabilidade de sucesso na distribuição de conteúdos aumenta em função da relevância, da oportunidade e do interesse que esses conteúdos vão despertar no utilizador que recebe a informação. Como tal, o anunciante deve distribuir a informação para os clientes tendo em conta alguns fatores

---

<sup>1</sup> Consultar "<https://www.emailmanager.com/pt/blog/1/38/email-marketing-e-a-problematICA-da-compra-de-mailings-ou-listas-de-emails.html>"

relevantes, como por exemplo, dados demográficos obtidos através do perfil do utilizador (idade, género, etc.), dados geográficos obtidos através de coordenadas GPS do telemóvel do utilizador, preferências e áreas de interesse do utilizador, hábitos de consumo, como por exemplo que tipo de produtos costuma comprar ou que eventos costuma frequentar ou padrões de utilização da aplicação (que conteúdos o utilizador costuma clicar ou apagar, etc.). Também se pode levar em consideração a época do ano em que é realizada a divulgação, se é Verão ou Inverno, se é o Dia da Mãe ou Natal ou aspetos relacionados com a meteorologia, (se chove, faz sol, calor ou frio). Por exemplo, não fará sentido promover um produto do Dia dos Namorados durante o Verão.

Estes métodos de distribuição de conteúdos apresentam vantagens tanto para os comerciantes, que ao anunciarem a pessoas efetivamente interessadas no produto aumentam as suas vendas, como para os consumidores, que receberão informação que lhes será realmente útil. Mas quando estes métodos são aplicados sem o conhecimento ou aprovação dos utilizadores, tal iniciativa pode ser considerada ilegal e uma violação da segurança, no que diz respeito à privacidade, proteção de dados e defesa do consumidor (Riedl, Frankowski, & Lam, 2006). Esta é uma controvérsia que a indústria de segmentação comportamental está a tentar controlar através da instrução, apoio e restrição de produtos para manter todas as informações não pessoais ou para obter a permissão dos utilizadores finais (EDRi-gram, 2009).

Por outro lado, a potenciação do negócio pode resultar de uma elevada convergência entre os interesses dos utilizadores, clientes e empresas prestadoras do serviço de distribuição de conteúdos. O modelo de negócio deverá ser desenhado ou melhorado tendo em conta todos estes elementos. Neste contexto, o utilizador e a sua interação com os conteúdos e com as ferramentas de distribuição desses mesmos conteúdos, poderão representar também uma fonte de informação relevante, contribuindo recursivamente para o desenho das próprias estratégias de negócio no sentido de um *dynamic smart targeting*.

## **1.2 Empresa**

O projeto de investigação foi desenvolvido na empresa CardMobili – Desenvolvimento de *Software*. Esta empresa de capital português foi fundada em 2009 e possui sede em Leça do Balio. A CardMobili atua no setor de desenvolvimento de aplicações móveis, desenvolvendo e comercializando uma premiada plataforma tecnológica para a carteira digital. A CardMobili é líder no fornecimento de

tecnologia para a carteira digital e no telemóvel. Esta tecnologia transforma a forma como as empresas interagem com consumidores e como os consumidores compram bens e serviços. Há alguns anos que a CardMobili tem vindo a inovar no setor da carteira digital, tornando-se num dos *players* mais experientes neste mercado (CardMobili, 2014).

A empresa apresenta um serviço *online* de adesão simplificada a múltiplos programas de fidelização de empresas, disponível globalmente, o qual proporciona ao utilizador individual a comodidade de ter os seus cartões no telemóvel e às empresas aderentes a possibilidade de virtualizar as suas ofertas de fidelização (cartão, cupões, promoções, etc).

Através dos seus produtos de fidelidade, campanhas, comunicação, pagamentos e compras a CardMobili apresenta soluções não só para o utilizador individual da aplicação, como também para as operadoras móveis, para os retalhistas e para as instituições financeiras.

### **1.3 Objetivos**

O principal objetivo deste projeto de investigação centrou-se no desenvolvimento de um mecanismo de recomendação de conteúdos nomeadamente, ofertas, promoções ou programas de fidelização aos utilizadores da aplicação da CardMobili, com base num número significativo de parâmetros, podendo assim aumentar a relevância das ofertas que são enviadas para o *smartphone* de cada utilizador. Isto permitirá a cada utilizador aceder, no menu inicial, às ofertas que lhe poderão despertar mais interesse, e que variam de utilizador para utilizador, ou seja, cada cliente terá uma vista da página inicial diferente dos restantes. Permitirá ainda aos utilizadores, terem acesso a ofertas adicionais relacionadas com a oferta que eles estejam a visualizar no momento, as quais podem ser potencialmente relevantes para eles. Este processo respeita sempre as políticas de privacidade dos clientes, que abrange a forma como as suas informações são recolhidas, utilizadas, divulgadas e armazenadas.

### **1.4 Estrutura da Dissertação**

Esta dissertação é composta por sete capítulos. Após o enquadramento inicial do projeto, é apresentada, no Capítulo 2, a revisão bibliográfica sobre *Customer Relationship Management* (CRM), Carteira Digital e Sistemas de Recomendação. No Capítulo 3 é explicada a metodologia e as

tecnologias utilizadas, na concepção do mecanismo de recomendação. O caso de estudo e o processo de desenvolvimento do mecanismo são descritos no Capítulo 4 e no Capítulo 5, respetivamente. Segue-se a apresentação e análise dos resultados obtidos, após os testes efetuados ao mecanismo. Por fim, no Capítulo 7, são mencionadas as principais conclusões do projeto, e são identificadas algumas oportunidades de trabalho futuro.

## 2. REVISÃO DA LITERATURA

Neste capítulo apresenta-se uma revisão de conceitos e ferramentas subjacentes ao projeto de investigação, fazendo-se particularmente um enquadramento dos Sistemas de Recomendação enquanto ferramenta CRM, adaptada a dispositivos móveis.

### 2.1 Customer Relationship Management (CRM)

A forma como as empresas se relacionam com os seus clientes é reconhecida como um fator determinante para aumentar o seu valor, assim como a sua participação no mercado. Deste modo, é essencial conhecer as preferências dos clientes, os seus hábitos de compra, identificá-los e segmentá-los através de múltiplos critérios de avaliação, podendo assim a empresa prestar um melhor serviço, não só atingindo o desígnio de qualidade e consistência de serviço prestado, bem como criando oportunidades que geram mais valor para a empresa (Vieira, 2008).

CRM é um sistema de gestão integrado com foco no consumidor, cujo principal objetivo é auxiliar as empresas a atrair e fidelizar clientes de forma a atingir a sua total satisfação, recorrendo a uma melhor compreensão das suas necessidades e expectativas (Vieira, 2008).

Em suma, os processos de gestão que assentam em CRMs estão em vantagem em termos estratégicos, não apenas de *marketing*, mas também a nível económico e financeiro. De facto, se as empresas puderem conhecer intimamente os seus clientes poderão enquadrá-los em diferentes perfis de consumidor, conseguindo criar respostas personalizadas que respondem melhor às necessidades do mercado, aumentando os lucros da empresa (Francisco, Kovalski, & Callegari, 2006).

De acordo com Dyché (2001), a tecnologia CRM pode ser dividida em 3 componentes:

- CRM Operacional: refere-se ao componente que ajuda a melhorar a eficiência das operações do dia-a-dia do consumidor (Peppers & Rogers, 2011). Este componente é o produtor inicial de dados, e inclui funções corporativas de automação de vendas, automação de *marketing* empresarial, gestão de encomendas e serviço de suporte ao cliente (Crosby & Johnson, 2001) (Greenberg, CRM at the Speed of Light, Third Edition: Essential Customer Strategies for the 21st Century, 2004).
- CRM Colaborativo: pode ser visto como um centro de comunicação que providencia a conexão entre as empresas e os seus clientes, fornecedores e parceiros de negócio. O foco é a

obtenção de valor por parte do cliente, fundamentada em conhecimento e interação com o mesmo.

- CRM Analítico: é a análise dos dados coletados e armazenados, de forma a criar interações mais rentáveis e significativas com os clientes. Para alcançar este propósito, os dados são processados, interpretados e reportados (Greenberg, CRM at the Speed of Light, Third Edition: Essential Customer Strategies for the 21st Century, 2004). Tem como principal objetivo identificar as necessidades e entender comportamentos dos utilizadores através do acompanhamento dos seus hábitos.

De acordo com (Reynolds, 2002), cada componente CRM é dependente dos outros. Por exemplo, a componente analítica conduz à tomada de decisão na componente operacional, como preparação de vendas, ações de *marketing* e processos de serviço ao cliente. Por outro lado, sem dados coletados via processos CRM Operacional, o analítico não seria possível. Além disso, os dados processados pelas ferramentas do componente analítico não podem ser usados efetivamente, e a tomada de decisão estratégica não ocorreria sem o CRM colaborativo. Em suma, todos os componentes contribuem conjuntamente para criar valor.

## 2.2 Carteira Digital

Estamos na Era da Tecnologia, cuja evolução traz novos métodos de comunicação e, portanto, novas formas de fazer negócios. Nesta Era em que o telemóvel, mais especificamente o *smartphone*, é um dispositivo que faz parte do dia-a-dia de praticamente 50% dos portugueses<sup>2</sup> surge uma tendência de *mobile commerce*, que consiste, basicamente, em qualquer transação comercial realizada através de um dispositivo sem fios. Esta tendência abrange também os pagamentos *online*, ou *mobile payments*. Assim, o crescimento das transações financeiras via telemóvel induziram o surgimento das *mobile wallets*, ou *m-wallets*, que, segundo Cole, *et al.* (2009), Agarwal, *et al.* (2010), Jefferson (2010), Forum (2011) e Alvarez & Feichtner (2013), podem oferecer serviços como:

- Armazenar cartões de pagamentos (crédito, débito, etc.);
- Armazenar cartões de identificação (cartão do cidadão, carta de condução, passaportes, etc.);

---

<sup>2</sup> “Metade da população portuguesa já utiliza *smartphones*”. Obtido em 15 de Setembro de 2014, de Jornal de Notícias: [http://www.jn.pt/PaginalInicial/Tecnologia/Interior.aspx?content\\_id=4114579](http://www.jn.pt/PaginalInicial/Tecnologia/Interior.aspx?content_id=4114579)

- Armazenar cartões de fidelização de marcas, cartões de sócio, seguros de saúde, bilhetes de avião, entre outros;
- Armazenar cupões de descontos, ofertas ou até cartões de pontos;
- Aceder a todas as informações contidas nos seus cartões guardados;
- Realizar transações financeiras e guardar informação (e.g. registo, talões de compra, saldo de conta);
- Receber publicidade;
- Obter serviços de segmentação baseados em localização ou contextuais;
- Permitir pagamentos móveis remotos, de proximidade ou com cartões de pagamentos (crédito, pré-pagos, débito).

Os pagamentos móveis só poderão ser efetuados por *smartphones* devidamente equipados com um componente que, quando em contacto com um leitor, emite uma frequência de baixo nível, permitindo executar pagamentos por cartão de crédito. Este componente também pode permitir o resgate de cupões e vales de desconto (Alvarez & Feichtner, 2013).

### 2.2.1. Vantagens

Logicamente que estes serviços têm uma série de vantagens e desvantagens. Segundo Alvarez & Feichtner (2013), Forum (2011), Jefferson (2010) e Alcatel-Lucent (2010) algumas das vantagens que uma carteira digital oferece são:

- Maior comodidade, pois todos os cartões, dinheiro e todos os recursos físicos normalmente guardados na carteira tradicional, vão estar arquivados no seu dispositivo móvel, do qual os utilizadores geralmente não se separam;
- Portabilidade;
- Versatilidade;
- Fácil gestão de conteúdos;
- Menores custos e possíveis reduções nos custos de transação e custos de processamento de moeda;
- Facilidade de acesso e maior rapidez;
- Facilidade no controlo de gastos;
- Capacidade de *opt-in* para publicidade personalizada.



Em suma, as carteiras digitais apresentam dois pontos fortes, oferecem a conveniência de manter todos os cartões físicos eletronicamente, facilitando a sua gestão e utilização, e traduzem-se numa experiência rica para o utilizador, o qual através da interface pode pedir ajuda sobre como utilizar os serviços (Jefferson, 2010).

### 2.2.2. Desvantagens

Mas segundo Pareek (nd), Wise (2013) e Stephens (2011), a carteira digital pode apresentar alguns inconvenientes ou limitações como por exemplo:

- Problemas legais de segurança e/ou fraude;
- *Malware* que pode invadir a privacidade e dados pessoais do utilizador;
- Potencial roubo do dispositivo móvel;
- Problemas de memória ou de limitação da bateria do telemóvel;
- Falta de cobertura de rede;
- Falta de uma plataforma única e comum para todos os dispositivos móveis.

O facto de muitos retalhistas e lojistas não estarem equipados com estações *Point of Sale* (POS) pode representar uma desvantagem, uma vez que o utilizador não poderá usufruir completamente das vantagens da sua carteira digital (Liyakasa, 2013).

### 2.2.3. Intervenientes

Segundo Cole, *et al.* (2009), dos atores que intervêm no ecossistema de uma carteira digital, devem destacar-se:

- O Utilizador: toda a carteira é desenvolvida à volta deste interveniente, desempenhando ele o papel principal, o qual passa por gerir o conteúdo da carteira, tanto na própria aplicação móvel como através do *browser*, e usá-la em transações *online* ou físicas (nas próprias lojas);
- Os Fornecedores de Conteúdo: São aqueles que fornecem os objetos que desejam incluir na carteira, para serem consumidos pelos Utilizadores. Mas estes têm de superar alguns obstáculos. Primeiro, tem de haver uma maneira de identificar a carteira de um determinado utilizador. De seguida, é preciso adquirir as credenciais necessárias para a interação com essa carteira. Por último, tem de haver uma forma de introspeção sobre a carteira, para que as

regras de acesso à mesma possam ser determinadas antes de dedicar recursos para trabalhar com o seu conteúdo;

- Os Consumidores de Conteúdo: são os intervenientes que leem, atualizam ou extraem um objeto da carteira (por exemplo, leitores de códigos de barras ou QR (*Quick Response codes*), uma estação de pagamento, *Point of Sale*, ou POS);
- *Host*. é a entidade que deve implementar a carteira e torná-la disponível para ser utilizada pelos outros intervenientes;
- Os Fornecedores de Acesso: o fornecedor de acesso principal é o *host*, mas existem outros intervenientes que podem desempenhar um papel de fornecedor secundário de acesso, dentro do contexto da sua função empresarial.

É importante notar que na implementação de uma carteira digital, todos estes constituintes devem ser levados em consideração. Todos eles, com exceção do Utilizador, oferecem soluções tendo em conta quem é o utilizador da carteira, se este deseja a carteira digital e para que fins.

## 2.3 Sistemas de Recomendação

Com a expansão tecnológica e o surgimento da Internet, os sistemas de comunicação assistiram à criação de um novo canal que possibilitou o acesso fácil à comunicação e à informação existente, o que, com o passar do tempo, acabou por levar a uma sobrecarga de informação.

Quando um utilizador está interessado em obter uma determinada informação e fica perdido no meio de intermináveis opções, as quais, na sua maior parte, são completamente irrelevantes para ele, surge a necessidade de filtrar a mesma.

Duas técnicas foram desenvolvidas para auxiliar na obtenção de informações relevantes: Recuperação de Informação e Filtragem de Informação (Rodrigues, 2010). A primeira está associada à maioria dos sites de pesquisa como o Google.com, Bing.com, Yahoo.com ou ainda Sapo.pt, cujo mecanismo de busca na Internet é realizado através das palavras-chave digitadas pelo utilizador e são uma ótima solução para aqueles utilizadores que têm pleno conhecimento daquilo que procuram, caso contrário o sistema irá devolver um vasto universo de resultados referentes à pesquisa efetuada e a dúvida sobre que informação é mais relevante irá persistir (Cazella, 2006). A segunda técnica baseia-se no perfil de um utilizador para realizar as suas pesquisas e geralmente mantém uma base de dados com informação sobre as suas preferências e interesses (Burin, 2007). Esta última é habitualmente aplicada quando se deseja serviços personalizados.

As empresas podem praticar diferentes estratégias de *marketing*, dentre elas (Bezerra, 2004):

- *Marketing* segmentado: talvez ainda o mais aplicado, onde a empresa identifica grupos de indivíduos com características semelhantes e define uma política para cada conjunto identificado;
- *Marketing one-to-one*: onde a empresa considera que cada cliente tem as suas características próprias e direciona as suas ações de maneira diferente para cada indivíduo.

A tecnologia de suporte ao *marketing one-to-one* é conhecida como Sistema de Recomendação (Bezerra, 2004). Os sistemas de recomendação foram originalmente definidos como aqueles em que “as pessoas fornecem recomendações como entrada, as quais o sistema posteriormente agrega e direciona para os receptor adequado” (Resnick & Varian, 1997). Mais recentemente este termo ganhou uma conotação mais ampla, descrevendo os sistemas de recomendação como aqueles que “têm o efeito de guiar o utilizador de forma personalizada para objetos interessantes ou úteis num grande espetro de opções possíveis” (Burke, 2002). Segundo Cazella (2008) os “Sistemas de Recomendação podem ser definidos como sistemas que procuram auxiliar indivíduos a identificarem conteúdos de

interesse num conjunto de opções que poderiam caracterizar uma sobrecarga. São sistemas que procuram facilitar a penosa atividade de busca por conteúdo interessante”.

Um sistema de recomendação é utilizado como uma ferramenta que tenta determinar conteúdos potencialmente interessantes para um determinado utilizador, com base no seu histórico de atividade, e que podem ficar perdidos na larga gama de ofertas existente. Lichtnow, *et al.* (2006) afirmam que o objetivo dos sistemas de recomendação não é apenas recomendar conteúdos através das consultas efetuadas pelos mesmos, mas também ser capaz de prever o interesse de um utilizador por um determinado item que este ainda não conhece.

### 2.3.1. Tarefas de um Sistema de Recomendação

Os Sistemas de Recomendação podem ser vistos, de modo geral, como sistemas de informação especiais que fornecem uma visão personalizada do repertório digital de dados, associados a um determinado sistema, e que pode ser apresentada de diversas formas mediante o objetivo do sistema (Bezerra, 2004). Entre as principais tarefas estão (Bezerra, 2004): apresentar a opinião de utilizadores, apresentar listas de itens mais comprados ou visitados, apresentar listas de produtos semelhantes no contexto visualizado, apresentar lista de produtos relacionados com o produto atual e com o histórico da comunidade, apresentar listas de produtos relacionados com os interesses do utilizador de acordo com o seu perfil.

- Apresentar a opinião de utilizadores

A avaliação de um utilizador, acerca de um produto ou vendedor, é importante para outros utilizadores selecionarem as informações mais relevantes num dado contexto. Dependendo do sistema de recomendação, as avaliações são efetuadas de diferentes formas:

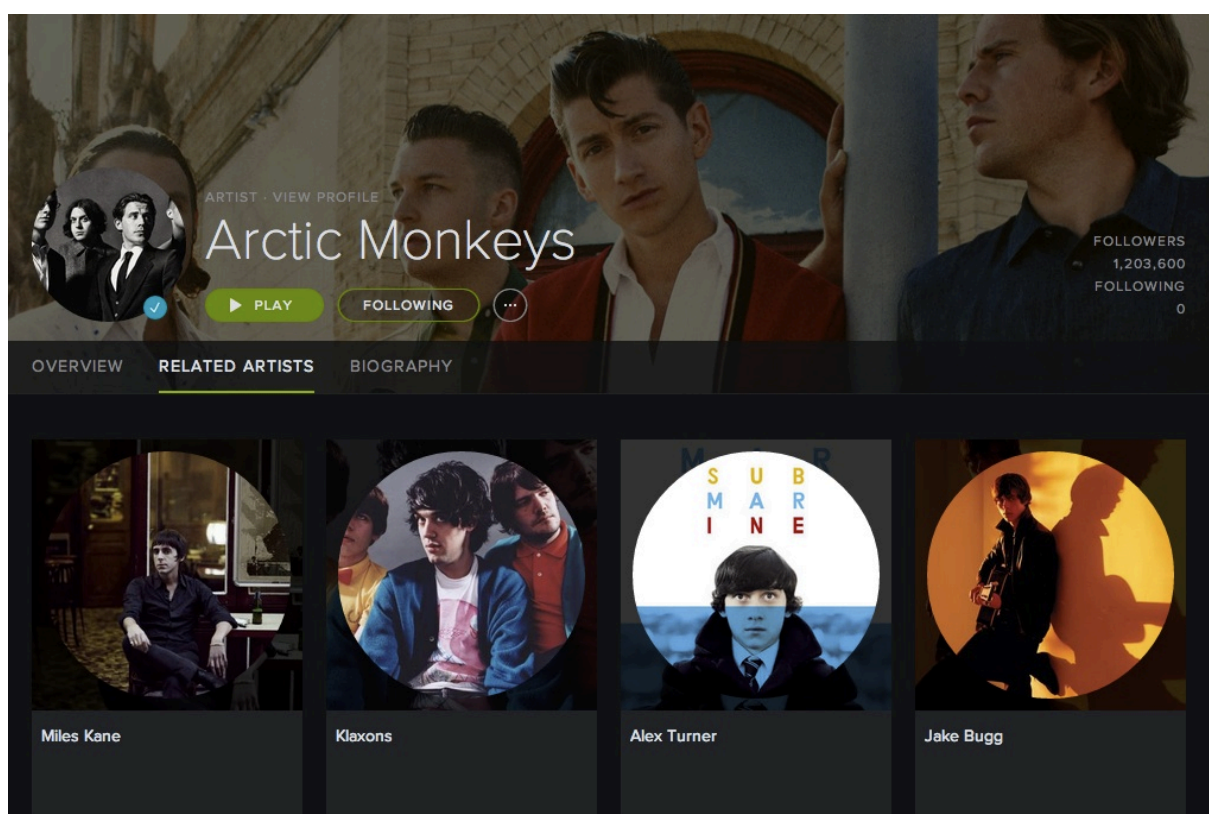
- Através de valores a representar uma nota, definidos pelo sistema (por exemplo, notas de 1 a 5);
- Através de comentários sobre as características, vantagens e desvantagens do item, descrevendo a sua experiência;
- E através de escalas, como por exemplo, de Muito Mau a Muito Bom, ou número de estrelas, etc..

- Apresentar listas de itens mais comprados ou visitados

Esta funcionalidade está presente na maioria dos sistemas de recomendação, pois não necessita de inteligência computacional, nem qualquer tipo de conhecimento sobre o utilizador.

- Apresentar listas de produtos semelhantes no contexto visualizado

Diversos sistemas oferecem esta funcionalidade, como por exemplo os sistemas de recomendação de música. Estes podem exibir listas de músicas do mesmo estilo da música que o utilizador está a ouvir no momento. A Figura 1 mostra a recomendação de artistas considerados semelhantes, ou do mesmo estilo, do artista pesquisado na aplicação Spotify.



**Figura 1 - Lista de artistas relacionados na aplicação Spotify**

- Apresentar lista de produtos relacionados com o produto atual e com o histórico da comunidade

Esta funcionalidade requer a pesquisa não só dos itens semelhantes ao item atual, como também dos itens que foram pesquisados, visualizados ou comprados pelos restantes utilizadores,

quando estes realizavam uma pesquisa igual à que o utilizador realizou no momento. A Amazon (Figura 2), por exemplo, faz uso desta funcionalidade no seu sistema de recomendação.

**Frequently Bought Together**

Price for both: **\$108.84**

[Add both to Cart](#) [Add both to Wish List](#)

Show availability and shipping details

Trade in your item  
Get a **\$21.50** Gift Card. [Trade in](#) [Learn More](#)

Have one to sell? [Sell on Amazon](#)

**Customers Who Bought This Item Also Bought** Page 1 of 15

Product	Price	Prime
Harry Potter: The Complete 8-Film... Daniel Radcliffe ★★★★★ 2,512 DVD \$57.99		Prime
The Hogwarts Library (Harry Potter) J. K. Rowling ★★★★★ 284 Hardcover \$18.35		Prime
J.R.R. Tolkien 4-Book Boxed Set: The Hobbit... J.R.R. Tolkien ★★★★★ 6,213 Mass Market Paperback \$21.72		Prime
Percy Jackson and the Olympians 5 Book... Rick Riordan ★★★★★ 759 Paperback \$25.82		Prime
The The Hunger Games Trilogy Box Set... Suzanne Collins ★★★★★ 8,146 Paperback \$23.38		Prime
Harry Potter: Complete 8- Film Collection... Daniel Radcliffe ★★★★★ 2,512 Blu-ray \$63.96		Prime
The Tales of Beedle the Bard, Standard Edition... J. K. Rowling ★★★★★ 939 Hardcover \$7.40		Prime

Figura 2 - Recomendações baseadas numa pesquisa no site Amazon.com

- Apresentar listas de produtos relacionados com os interesses do utilizador de acordo com o seu perfil

Com esta funcionalidade a recomendação é efetivamente personalizada, pois só são apresentados itens compatíveis com o perfil do utilizador, tendo em conta uma série de variáveis como por exemplo, os itens adquiridos ou visualizados por este, a localização, entre outros. O Youtube, por exemplo, utiliza este tipo de recomendações para sugerir vídeos aos seus espetadores, com base no seu perfil de visualizações, vídeos favoritos e vídeos marcados com um “gosto”. Na Figura 3 pode ver-se um exemplo da secção de recomendações do Youtube.

### Recomendados

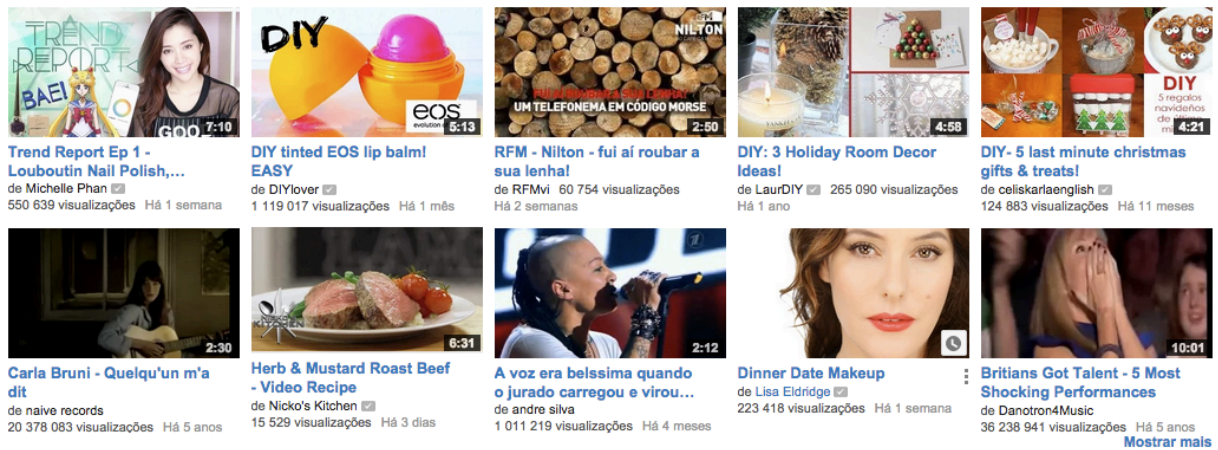


Figura 3 - Vídeos Recomendados pelo Youtube


### 2.3.2. Colheita de Dados

Antes de qualquer processo de colheita de dados é necessário identificar o utilizador, de forma a guardar devidamente as informações obtidas sobre ele. Esta identificação pode ser efetuada de duas formas (Cazella, Nunes, & Reategui, 2005):

- Identificação no servidor: onde são solicitados ao cliente alguns dados pessoais, tais como um *login* e uma *password*;
- Identificação no cliente: através de cookies automáticos, isto é, a máquina é identificada através do endereço IP, o que não é totalmente eficiente, pois não prevê que mais do que uma pessoa possa utilizar a máquina.

Depois de identificar o utilizador procede-se efetivamente à colheita de informação, que também esta pode ser efetuada de duas formas distintas (Cazella, Nunes, & Reategui, 2005) e (Cazella, 2006):

- Explicitamente: onde o utilizador geralmente preenche um formulário indicando deliberadamente as suas preferências. A Figura 4 mostra o formulário do *site* da CardMobili.



Home / Profile

## Profile

**Contact information** ▾

Email  
luana.telha@gmail.com

country  
Portugal ▾

Fiscal number  
[input field]

Name  
Luana Geórgia Lopes Telha

Address  
[input field]

Zip code  
[input field]

City  
[input field]

Mobile phone number  
[input field]

Phone number  
[input field]

Birth date  
18/10/1988

**Basic information** ▾

Gender  
 Female  
 Male

ID document number  
13421094 8 ZZ7

Passport number  
[input field]

Marital status  
 Divorced  
 Married  
 Single  
 Widowed

Number of children  
[input field]

**Education and work** ▾

Highest level of education achieved  
Choose an option ▾

Current employment status  
Choose an option ▾

**Interest areas** ▾

Culture and leisure

- Art and Antiques
- Cinema/Theater
- Gourmet stores
- Music/Concerts
- Restaurants
- Travelling
- Technology

Figura 4 - Colheita de dados aplicando uma abordagem explícita no site da Cardmobili

Deste formulário destacam-se as áreas de interesse, que representam as categorias mais relevantes para o utilizador, e nas quais se devem inserir as ofertas mais significativas para recomendações.

- Implicitamente: através do comportamento do utilizador, como por exemplo, páginas consultadas, produtos visualizados, etc., é possível aprender e perceber os seus interesses, embora haja dificuldade em medir a relevância dos dados.

É possível fazer recomendações comparando as preferências de um utilizador com as de um grupo de outros clientes (recomendações baseadas em utilizadores), ou procurando produtos com características semelhantes aos que este já demonstrou interesse no passado (recomendações



baseadas em produtos). Quanto mais detalhados forem os perfis dos utilizadores e dos produtos, ou seja, quanto mais informação existir, mais eficaz e eficiente será a recomendação.

### 2.3.3. Privacidade em Sistemas de Recomendação

Muitas empresas recorrem à coleta de dados de forma implícita, através da consulta da navegação ou das compras efetuadas pelo utilizador, e portanto sem o conhecimento do mesmo. Esta recolha de dados não apresenta nenhum transtorno, tanto aos utilizadores como à própria empresa, quando associada a ela existem políticas de privacidade adequadas. No entanto, há conhecimento de empresas que vendem os dados dos seus clientes a outras empresas, aumentando a quantidade de SPAM e descredibilizando a empresa em questão (Cardoso, Reimao, & Marques, 2011).

Antes de recolherem os dados, as empresas necessitam criar primeiro uma relação de confiança com os clientes. Enquanto não existe uma lei clara em vigor que assegure a privacidade dos utilizadores da Internet, nos Estados Unidos, algumas entidades executam essa função, nomeadamente a *Federal Trade Commission* (FTC) e o Departamento do Comércio, que em 1999 apresentou um conjunto de princípios que as empresas devem seguir, para proteger a privacidade dos seus clientes (Rohrmann, 2000):

- Notícia - a empresa tem de notificar o utilizador que os seus dados serão recolhidos;
- Escolha - o utilizador terá opção de escolher se permite que os seus dados sejam ou não recolhidos;
- Transferência autorizada - a transferência de dados a terceiros só será efetuada com a permissão do utilizador;
- Segurança - tomar precauções para prevenir a perda de dados de terceiros;
- Integridade - evitar que os dados sejam adulterados;
- Acesso - os utilizadores devem ter acesso a todos os seus dados;
- Mecanismos - estabelecimento de mecanismos que permitam a efetiva fiscalização e cumprimento dos restantes seis princípios.

### 2.3.4. Tipos de Sistemas de Recomendação

Existem diversos tipos de técnicas para filtrar informação, que podem ser categorizadas por (Melville & Sindhvani, Recommender Systems, sd):

#### 1. Filtragem Colaborativa

2. Filtragem Baseada em Conteúdos
3. Filtragem Híbrida

Nas próximas secções explica-se com mais detalhe cada uma delas.

#### 2.3.4.1. Filtragem Colaborativa

O conceito “filtragem colaborativa” surgiu no contexto da comercialização do primeiro sistema de recomendação, o sistema Tapestry (Melville & Sindhvani, Recommender Systems, sd). Goldberg (1992) definiu este sistema como a colaboração de pessoas, que se auxiliavam umas às outras a filtrar informação, através das suas opiniões acerca de determinado documento.

Filtragem Colaborativa é a técnica mais popular e mais vezes implementada nos sistemas de recomendação, e baseia-se no facto de que as opiniões dos outros têm uma influência considerável na tomada de decisão das pessoas. O principal objetivo de um algoritmo de filtragem colaborativa é prever o valor da avaliação que um utilizador daria a um produto, que este ainda não conhece, tendo como base as apreciações de outros utilizadores. Posteriormente, o utilizador poderá informar, direta ou indiretamente, se gostou ou não da recomendação (Oliveira & Coello, 2013).

Ao nível da técnica de filtragem colaborativa existem dois tipos de abordagem, filtragem colaborativa baseada em memória (que pode ser baseada em utilizadores ou em itens) e filtragem colaborativa baseada em modelo (Ekstrand, Riedl, & Konstan, 2011) (Khoshgoftaar & Su, 2009) (Boullé, Meyer, & Candillier, 2007) (Melville & Sindhvani, Recommender Systems, sd):

##### ➤ Filtragem Colaborativa baseada em utilizadores

Nesta abordagem é escolhido o subconjunto de utilizadores com maior valor de similaridade com o utilizador ativo, e é usada uma combinação ponderada das suas avaliações para calcular previsões para este utilizador. Este cálculo passa por três passos:

1. Associar a cada utilizador ( $u$ ) um valor de similaridade ( $w$ ) com o utilizador ativo ( $a$ ),  $w_{a,u}$ . Existem várias fórmulas para calcular essa similaridade, sendo geralmente usada a Correlação de Pearson:

$$w_{a,u} = \frac{\sum_{i \in I} (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i \in I} (r_{a,i} - \bar{r}_a)^2 \sum_{i \in I} (r_{u,i} - \bar{r}_u)^2}} \quad (1)$$

Onde  $I$  representa o conjunto de itens classificados por ambos os utilizadores,  $r_{u,i}$  é a classificação que o utilizador  $u$  atribuiu ao item  $i$ , e  $\bar{r}_u$  representa a média de todas as classificações atribuídas pelo utilizador  $u$ . Alternativamente à Correlação de Pearson, podem tratar-se as classificações de dois utilizadores, como um vetor num espaço  $m$ -dimensional e calcular a similaridade baseada no cosseno do ângulo entre eles:

$$w_{a,u} = \cos(\vec{r}_a, \vec{r}_u) = \frac{\vec{r}_a \cdot \vec{r}_u}{\|\vec{r}_a\|_2 \times \|\vec{r}_u\|_2} \quad (2)$$

2. Seleccionar os  $k$  utilizadores que apresentam similaridade mais alta com o utilizador ativo, também conhecida como *vizinhança*.
3. Calcular a predição com base nas classificações da vizinhança do utilizador ativo. Este cálculo é feito, na maioria das vezes, através da média ponderada dos desvios da média das classificações de cada vizinho:

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u \in K} (r_{u,i} - \bar{r}_u) \cdot w_{a,u}}{\sum_{u \in K} w_{a,u}} \quad (3)$$

Onde  $p_{a,i}$  simboliza a predição para o utilizador ativo  $a$  relativamente ao item  $i$ , e  $w_{a,u}$  equivale ao valor da similaridade entre o utilizador ativo  $a$  e cada utilizador  $u$ , pertencente ao conjunto de vizinhos mais próximos  $K$  (Melville & Sindhvani, Recommender Systems, sd).

➤ Filtragem Colaborativa baseada em itens

Esta técnica calcula a similaridade entre itens, associando a cada um deles um conjunto de vizinhos mais próximos. Posteriormente calcula a previsão do interesse do utilizador por esse item, com base nas avaliações do mesmo sobre os itens vizinhos. Também neste caso a similaridade entre os itens  $i$  e  $j$  pode ser calculada pela Correlação de Pearson:

$$w_{i,j} = \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in U} (r_{u,i} - \bar{r}_i)^2 \sum_{u \in U} (r_{u,j} - \bar{r}_j)^2}} \quad (4)$$

Onde  $U$  representa o conjunto de utilizadores que avaliaram ambos os itens  $i$  e  $j$ ,  $r_{u,i}$  simboliza a avaliação do utilizador  $u$  sobre o item  $i$ , e  $\bar{r}_i$  é a média de todas as avaliações (de todos os utilizadores) efetuadas sobre o item  $i$ . Uma vez reconhecida a vizinhança do artigo, pode calcular-se o valor de avaliação previsto que o utilizador ativo lhe daria, através de uma média ponderada:

$$p_{a,i} = \frac{\sum_{j \in K} r_{a,j} w_{i,j}}{\sum_{j \in K} |w_{i,j}|} \quad (5)$$

Onde  $K$  é o conjunto de itens vizinhos que foram avaliados pelo utilizador ativo  $a$ .

A Amazon é um ótimo exemplo de uma entidade que tem um sistema de recomendação de Filtragem Colaborativa baseada em itens. O seu algoritmo passa por construir uma tabela de similaridade, encontrando itens que os consumidores tendem a comprar em conjunto. Poderia ser construída uma matriz produto-a-produto calculando a medida de similaridade para cada par de itens. No entanto, muitos pares de produtos não têm clientes comuns e, portanto, a abordagem seria ineficiente em termos de tempo de processamento e uso de memória. O seguinte algoritmo fornece uma melhor abordagem (Linden, Smith, & York, 2003):

```

For each item in product catalog,  $I_1$ 
  For each customer  $C$  who purchased  $I_1$ 
    For each item  $I_2$  purchased by
      customer  $C$ 
      Record that a customer purchased  $I_1$ 
        and  $I_2$ 
  For each item  $I_2$ 
    Compute the similarity between  $I_1$  and  $I_2$ 

```

➤ Filtragem Colaborativa baseada em modelo

Esta técnica utiliza a informação da interação parcial para modelar um conjunto de parâmetros para os utilizadores e itens, que juntos podem reconstruir previsões probabilísticas sobre as interações em falta (Bergner, Dröschler, Kortemeyer, Rayyan, Seaton, & Pritchard, 2012).

Os algoritmos nesta categoria têm uma abordagem probabilística e percebem o processo de filtragem colaborativa como o cálculo do valor esperado da predição de um utilizador, dado os valores

das suas avaliações sobre outros itens. O processo de construção do modelo pode ser realizado através de diferentes algoritmos de *machine learning*, tais como redes Bayesianas, *clustering* e abordagens baseadas em regras de associação (Sarwar, Karypis, Konstan, & Riedl, 2001).

As redes Bayesianas pertencem à família dos modelos gráficos probabilísticos. Estas estruturas gráficas servem para representar conhecimento sobre um domínio incerto (Ben-Gal, 2007).

Os modelos de *clusters* tratam a filtragem colaborativa como um problema de classificação, identificando grupos de utilizadores com características semelhantes (Sarwar, Karypis, Konstan, & Riedl, 2001). Uma vez criados os grupos, as previsões para um indivíduo podem ser calculadas com base na média das opiniões de outros utilizadores desse grupo (Xue, et al., 2005).

A abordagem com regras de associação aplica algoritmos de descoberta de associações entre dois conjuntos de itens, em que a presença de alguns numa determinada transação implica que produtos de outro conjunto estejam também presentes na mesma transação, como quando por exemplo alguns produtos são comprados em conjunto (Sarwar, Karypis, Konstan, & Riedl, 2000).

Em conclusão, é possível calcular a semelhança entre dois utilizadores ou artigos de diferentes maneiras, mas dada a grande dimensão da matriz utilizador-item, o processo de identificação de vizinhanças e cálculo de avaliações, diretamente a partir dos dados disponíveis, pode tornar-se computacionalmente dispendioso, com  $O(NM)$  no pior caso (Linden, Smith, & York, 2003), que significa que o tempo computacional gasto apresenta um crescimento de ordem quadrática. De forma a reduzir o tempo necessário para fazer recomendações, uma grande diversidade de algoritmos de filtragem colaborativa baseada em modelo foi sendo desenvolvida. Como estes criam um modelo com os dados disponíveis, em vez de os usarem diretamente, o tempo necessário para fazer uma previsão tende a ser muito menor. Porém, esta técnica necessita de aprender o modelo, antes de fazer as previsões, o que pode ser um processo complexo e demorado. Deve realçar-se que algoritmos baseados em modelos não são tão precisos como os algoritmos baseados em memória. (Oliveira & Coello, 2013)

Os sistemas que utilizam esta técnica têm a vantagem de poder filtrar qualquer tipo de artigo/item, de texto, arte, música, filmes, etc. (Riedl, Konstan, & Herlocker, Explaining Collaborative Filtering Recommendations , 2000). Além disso têm a capacidade de levarem em consideração as qualidades e defeitos de um produto para fazer as sugestões, particularmente nas avaliações explícitas por parte dos utilizadores (Rafsanjani, Salim, Aghdam, & Fard, 2013).

Mesmo sendo uma das técnicas mais usadas atualmente, a filtragem colaborativa apresenta ainda alguns problemas (Melville & Sindhvani, Recommender Systems, sd) (Khoshgoftaar & Su, 2009):

- Esparsidade: dado o grande tamanho das bases de dados atualmente, este método só será eficiente se os utilizadores avaliarem muitos artigos. Este problema ocorre quando a quantidade de avaliações é pequena relativamente ao número de utilizadores e itens existentes no sistema, fazendo com que a matriz de avaliações se apresente esparsa;
- *Cold Start*: este pode ocorrer quando um novo utilizador, ou um novo item, entra no sistema, não havendo dados disponíveis suficientes para calcular uma vizinhança confiável, de maneira a efetuar recomendações;
- Escalabilidade: quando o número de utilizadores e de artigos existentes no sistema cresce consideravelmente, os algoritmos tradicionais de filtragem colaborativa sofrem graves problemas de escalabilidade, exigindo recursos computacionais que excedem o aceitável.
- Sinonímia: refere-se à tendência de itens muito semelhantes apresentarem nomes diferentes, e portanto os algoritmos de filtragem colaborativa não encontrarem qualquer correspondência entre eles, para calcular a similaridade;
- Ovelha negra: refere-se aos utilizadores que não apresentam preferências consistentes com um grupo de pessoas e para os quais, portanto, se torna difícil calcular recomendações;
- *Shilling Attack*: ocorre nos casos em que qualquer pessoa pode fornecer recomendações. Por exemplo, pessoas podem fazer milhares de recomendações positivas aos seus próprios artigos e recomendar negativamente os seus concorrentes. É desejável que sistemas de recomendação com filtragem colaborativa introduzam medidas que evitem estes ataques.

Resumindo, a filtragem colaborativa é uma abordagem mais focada no consumidor e apresenta aspetos positivos, na medida em que evita recomendações repetitivas, apesar de requerer uma grande quantidade de informação sobre o utilizador para funcionar de forma eficiente (Cazella, Nunes, & Reategui, 2005).

#### 2.3.4.2. Filtragem Baseada em Conteúdos

Ao contrário de um sistema de filtragem colaborativa, que recomenda os itens com base na correlação entre pessoas com preferências semelhantes, a filtragem baseada em conteúdos seleciona os itens com base na correlação entre o conteúdo dos itens e as preferências do utilizador (Meteren & Someren, nd).

A abordagem baseada em conteúdo tem raízes na área de recuperação de informação (Cazella, Nunes, & Reategui, 2005) partindo do princípio que os utilizadores tendem a interessar-se por itens semelhantes aos que eles demonstraram interesse no passado, sendo então definida a similaridade entre os itens (Lichtnow, et al., 2006).

Geralmente em sistemas de recomendação baseados em conteúdo é construído, em primeiro lugar, um perfil do item para recomendar, que é o registo das características mais importantes do mesmo. Em casos mais simples, essas características são facilmente detetadas. Por exemplo, considerando que se pretende recomendar um filme, deve ter-se em consideração características como (Ullman, 2011):

- O género do filme (se é comédia, drama, terror, etc.);
- O realizador (algumas pessoas têm preferência por filmes de determinados realizadores);
- O ano em que foi produzido (ou de estreia) do filme (alguns espectadores preferem os filmes mais antigos aos mais recentes ou vice-versa);
- O elenco (para quem prefere ver filmes interpretados pelos seus atores favoritos).

Livros e música têm características semelhantes ao exemplo dado em cima, como o género, o ano e autor, compositor e artista. Mas existem outras classes de itens em que as características não são tão evidentes, tais como coleções de documentos e imagens (Ullman, 2011).

O problema das imagens é que estas são representadas por *arrays* de *pixels*, não fornecendo nenhuma informação sobre os atributos da imagem como um todo, apenas de pequenas características, como a quantidade de uma determinada cor, mas na maior parte dos casos os utilizadores não se interessam por determinadas imagens em função da cor. Nestes casos é possível associar a cada imagem um conjunto de *tags*, que podem ser palavras ou frases simples, para a descrever. O problema reside na vontade do utilizador de colaborar, ou não, com o sistema (Ullman, 2011).

Deste modo, na maioria dos sistemas de recomendação com filtragem baseada em conteúdos, o perfil dos itens são características textuais extraídas de páginas *web*, *emails*, artigos ou descrições de produtos. Porém, as características textuais apresentam alguns problemas de aprendizagem, devido à ambiguidade da linguagem e à incapacidade do sistema de captar a semântica dos interesses do utilizador, pois estes são conduzidos por uma operação de correspondência de palavras ou frases. Se

uma palavra ou frase é encontrada tanto no perfil do utilizador como no documento, a correspondência é feita e o documento é considerado relevante (Semeraro, Gemmis, & Lops, 2011).

Existem várias maneiras de representar os termos de modo a serem usados como base para a componente de aprendizagem. O método usado com maior frequência é o modelo de espaço vetorial, onde um documento  $D$  é representado como um vetor  $m$ -dimensional, em que cada dimensão corresponde a uma duração distinta, e  $m$  é o número total de termos usados na coleção de documentos. O peso associado a cada termo,  $w_i$ , é geralmente calculado pelo método  $tf \cdot idf$  e pode ser representado como (Meteren & Someren, nd):

$$w_i = tf_i \cdot \log \frac{n}{df_i} \quad (6)$$

Onde  $f_i$  é o número de ocorrências do termo  $t$  no documento  $D$ ,  $n$  representa o número total de documentos na coleção,  $df_i$  o número de documentos nos quais o termo  $t$  aparece pelo menos uma vez, e  $w_i$  é o peso do termo, indicando a sua importância, e que é igual a zero caso o documento  $D$  não contenha o termo. Nesta abordagem é atribuído um peso a cada termo, que é baseado nas vezes em que este surge num documento específico e com que frequência isso ocorre em toda a coleção de documentos (Meteren & Someren, nd). Deste modo a similaridade entre dois documentos pode ser calculada através do cosseno do ângulo formado pelos vetores que os representam:

$$sim(d_i, d_j) = \frac{\sum_k w_{ki} w_{kj}}{\sqrt{\sum_k w_{ki}^2} \sqrt{\sum_k w_{kj}^2}} \quad (7)$$

Onde  $k$  indica o número total de palavras no sistema (Semeraro, Gemmis, & Lops, 2011).

Outro método de cálculo da similaridade é o Coeficiente de Jaccard:

$$sim(i, j) = \alpha + \frac{|K_i \cap K_j|}{|K_i \cup K_j|} \quad (8)$$

Onde  $K_i$  representa o conjunto de palavras-chave correspondente ao item  $i$ . O parâmetro adicional  $\alpha$  assegura que todas as palavras-chave são tomadas em consideração, mesmo que o conjunto de palavras-chave seja completamente distinto do produto de cuja previsão se deseja calcular (Wartena, et al., 2011).



Uma estratégia de recomendação baseada em palavras-chave pode ser a dos k-vizinhos mais próximos, que é análoga à estratégia usada nos sistemas colaborativos clássicos, mas a similaridade entre dois itens é baseada no conteúdo e não nas avaliações (Wartena, et al., 2011). Também neste caso é calculada a predição do *rating* que o utilizador dará ao item, sendo esta a média ponderada de todos os produtos avaliados pelo mesmo, onde os artigos mais semelhantes obtêm maior peso:

$$R(u, i) = \frac{\sum_{j \in I_u} sim(i, j) R(u, j)}{\sum_{j \in I_u} sim(i, j)} \quad (9)$$

Onde  $I_u$  representa o conjunto de itens avaliados pelo utilizador  $u$  (Wartena, et al., 2011).

Os sistemas de recomendação baseados em conteúdos têm a vantagem de apenas necessitar das avaliações do utilizador ativo para construir o seu perfil, sendo assim independentes de avaliações de terceiros, o que resolve o problema de esparsidade da filtragem colaborativa (Júnior, 2004). Têm também a capacidade de recomendar itens que ainda não foram avaliados por qualquer utilizador, não sofrendo do problema de “novo item” como a filtragem colaborativa (Júnior, 2004). Além destas vantagens, estes sistemas são transparentes, isto é, têm a opção de mostrar uma explicação sobre o motivo de determinado produto se apresentar na lista de recomendações, listando explicitamente as descrições ou características de conteúdos (Semeraro, Gemmis, & Lops, 2011).

Porém, apesar de ser bastante usada em diversos sistemas de recomendação, a técnica de filtragem baseada em conteúdos apresenta algumas limitações (Júnior, 2004) (Semeraro, Gemmis, & Lops, 2011) (Bezerra, 2004):

- Superespecialização – refere-se ao facto de que as técnicas apenas analisam o texto e recomendam itens com conteúdos semelhantes aos que o utilizador já classificou, tendo um grau limitado de tendência para a novidade (*serendipity*). Estes sistemas não efetuam uma análise gramatical ao texto, não sendo capazes de reconhecer sinónimos. Por exemplo, se um texto fala em “carro” e outro texto em “automóvel” esta técnica pode não considerá-los semelhantes;
- Qualidade do Item – só será avaliada a qualidade de um produto caso essa informação venha descrita no conteúdo do mesmo, caso contrário é possível recomendar um item com defeito se esse apresentar o mesmo conteúdo que um sem defeito algum;

- Conteúdos não estruturados – estes sistemas não são capazes de analisar conteúdos multimédia, como por exemplo música, por não se tratarem de dados estruturados como um texto;
- Escalabilidade – com a adição de novos artigos, a utilização de memória será cada vez maior e a análise de conteúdos será mais demorada;
- Novo utilizador – quando entra um novo utilizador no sistema, este não tem interações com itens suficientes para ser construído um perfil, podendo o sistema não ser capaz de fazer recomendações confiáveis.

O Pandora Internet Radio é um sistema de recomendação baseado em conteúdos que faz recomendações de músicas. Ao contrário da Amazon, que cria uma vizinhança baseada nos utilizadores e nos seus interesses, o Pandora cria uma vizinhança de canções que são similares entre elas. Para descobrir essa semelhança tem implementado um sistema ao qual chamam The Music Genome Project, assumindo que as músicas são compostas por uma série de características, semelhante à forma como o genoma humano descreve uma pessoa, as quais descrevem uma canção. O The Music Genome Project aplica valores para cada canção em cada um dos cerca de 400 atributos musicais, detalhados em harmonia e melodia, ritmo e instrumentação, e forma e performance vocal. Uma vez que as músicas estão claramente classificadas na base de dados, o Pandora compara a descrição do gosto musical de uma estação seleccionada por um utilizador, com a classificação das músicas na base de dados, que devolve um conjunto de canções que impulsionam a lista de reprodução. O algoritmo que realiza esta tarefa deve ser executado de forma suficientemente rápida, para que o utilizador não tenha de esperar para ouvir a próxima seleção. Ele ocorre em tempo polinomial. Na Tabela 1 pode verificar-se os tempos de resposta de algumas tarefas (Howe, 2007):

**Tabela 1 - Tempos de resposta despendidos nas diferentes tarefas do sistema Pandora Internet Music**

$O(n)$	Adição de uma base de dados, onde $n$ é o número de atributos
$O(1)$	Adição de uma música à base de dados
$O(mn)$	Escolha da próxima música, onde $n$ é o número de atributos e $m$ é o número de músicas presentes na base de dados

Esta tabela mostra que o tempo computacional despendido na inserção de uma única música é constante. A inserção de uma base de dados é de tempo linear e leva tão mais tempo quanto mais atributos apresentar. A escolha da próxima música também apresenta tempo linear e depende do número de atributos e da quantidade de músicas presente na base de dados.

#### 2.3.4.3. Filtragem Híbrida

Ao verificar as características das duas técnicas apresentadas anteriormente, pode constatar-se que elas se complementam, na medida em que as vantagens de uma corrigem as falhas da outra. Sistemas de recomendação híbridos trabalham com as características de ambas as filtragens, colaborativa e baseada em conteúdos, de forma a reforçar as vantagens de cada uma ao mesmo tempo que mitigam as suas desvantagens (Cazella, Nunes, & Reategui, 2005). Existem diversas formas de combinar os algoritmos de filtragem colaborativa e baseada em conteúdos, onde as principais diferenças se revelam na apresentação dos resultados. Segundo Rafsanjani et al. (2013) são as seguintes:

- Mista: as recomendações são feitas pelas duas técnicas que no fim são combinadas numa mesma lista final;
- Ponderada: as recomendações são calculadas através de várias técnicas, com pesos diferentes para cada uma delas;
- Combinação de características: as características que estão relacionadas com diversos recursos de dados de recomendação são reunidas num único algoritmo de recomendação;
- Cascata: um dos algoritmos refina as sugestões e recomendações que são, posteriormente, apresentadas por um segundo algoritmo;
- Aumento das características: os resultados de uma abordagem são utilizados como parâmetro de entrada para a segunda abordagem;

- *Meta Level*: a técnica que é aprendida por um sistema de recomendação é utilizada como entrada para a outra técnica;
- *Switching*: este método alterna a técnica de recomendação consoante a situação atual.

Balabanovic e Shoam (1997) definem o Fab como sendo “um sistema de recomendação destinado a ajudar os utilizadores a filtrar a enorme quantidade de informações disponíveis na *World Wide Web* (WWW). Em funcionamento desde Dezembro de 1994, este sistema combina os métodos baseados em conteúdos e colaborativos de recomendação, de uma maneira que explora as vantagens das duas abordagens.

Este sistema apresenta duas fases no seu processamento: recolha de itens para formar a base de dados e a seleção dos mesmos para utilizadores específicos. A colheita dos itens reúne páginas relevantes em pequenos *clusters*, divididos por temas, ou tópicos. O agente de seleção descarta as páginas que o utilizador já viu, selecionando apenas páginas novas para cada cliente. O *feedback* dos utilizadores é armazenado no perfil do seu próprio agente de seleção privado, garantindo que nunca será abafado pelos *feedbacks* dos outros clientes. Quando são apresentadas as recomendações aos utilizadores, estes são obrigados a atribuir classificações apropriadas a partir de uma escala de 7 pontos. Páginas altamente cotadas são automaticamente passadas para os vizinhos mais próximos (Balabanovic & Shoham, 1997).

Para além das vantagens genéricas inerentes a um sistema híbrido, o Fab traz dois benefícios adicionais: trata de dois problemas de escala, comuns a todos os serviços *web*, que são o crescente número de utilizadores e o crescente número de documentos; identificando automaticamente comunidades emergentes de utilizadores com preferências idênticas (Balabanovic & Shoham, 1997).

### 2.3.5. Limitações dos Sistemas de Recomendação

Existem problemas comuns a todos os sistemas de recomendação, além das desvantagens de cada uma das técnicas (Rodrigues, 2010):

- ❖ *Problema de ausência das dimensões temporal e espacial*: há uma variável que raramente é levada em consideração, que se refere à localização do utilizador, tanto geográfica como temporalmente. Por exemplo, uma recomendação realizada no Inverno não fará sentido se for efetuada no Verão, ou para um utilizador que se encontre numa região tropical;

- ❖ *Dificuldade de explicação e apresentação das recomendações*: refere-se à dificuldade de descobrir uma forma de convencer o utilizador de que o produto que está a ser recomendado é de facto o produto que ele necessita, ou deseja. Para contornar este problema, as empresas utilizam diversas formas de comunicação com o cliente, como por exemplo, explicando o porquê da recomendação, quais as vantagens do produto, a opinião de outros clientes acerca do produto, entre outras;
  
- ❖ *Ausência de informações de negócio e integração com outros sistemas*: as informações filtradas por sistemas de recomendação são utilizadas apenas na sua aplicação, não fazendo uma integração com Sistemas de Informação de Gestão (SIG), ou de apoio à decisão. Além da falta de integração, existe uma ausência de informações de negócios, que representam as regras que regem a empresa.

#### 2.3.6. Recomendações Móveis

Os dispositivos móveis, tais como *smartphones* ou *tablets*, estão a evoluir no sentido de serem uma importante fonte de informação. Anteriormente, a maioria dos sistemas de recomendação eram desenvolvidos para computadores. No entanto, estes sistemas não podem ser diretamente aplicados como uma ajuda para utilizadores, uma vez que os sistemas de recomendação precisam de superar muitos desafios, geralmente presentes nos domínios dos dispositivos móveis (Davidsson, 2010).

##### ➤ Usabilidade e Interação

Recomendar num dispositivo móvel introduz novos obstáculos, devido às propriedades do dispositivo em si e ao contexto do seu uso. Comparado com os computadores tradicionais, um dispositivo móvel tem um ecrã mais pequeno e recursos de entrada limitados. O tamanho do ecrã dos dispositivos futuros é improvável que melhore, pois existe uma necessidade de mobilidade. Os utilizadores terão um ambiente diferente, comparado com o computador fixo, pois durante a utilização do *smartphone*, a sua interação e a atenção é condicionada pelo ambiente em que estes se encontram (Davidsson, 2010);

➤ Visualização

A maneira tradicional de apresentar recomendações, bem como os resultados de uma pesquisa regular, é através de uma lista de itens classificados por ordem decrescente. Devido ao tamanho limitado do ecrã de um dispositivo móvel, uma nova forma de exibir mais informação sobre o artigo recomendado torna-se crucial (Davidsson, 2010);

➤ Oportunidades

Os dispositivos móveis não apresentam só limitações. Eles também oferecem grandes oportunidades, como por exemplo, a localização física de um dispositivo e dados sensoriais poderia ser uma fonte valiosa de informação. Este tipo de dados revela o contexto em que o dispositivo é utilizado num determinado momento. Esta informação contextual pode auxiliar o processo de recomendação, fornecendo mais dados para as recomendações (Davidsson, 2010).

A aplicação de recomendações móveis recebeu uma forte atenção pela indústria do turismo, uma vez que as atrações de um local podem ser facilmente sugeridas através de uma lista tradicional, sobre um mapa, ou mesmo como um itinerário, partindo da posição em que o utilizador se encontra. Ainda nesta área, podem ser recomendados serviços de viagens, tais como, serviços de transportes locais, postos de turismo, hotéis ou restaurantes. Estes últimos podem ser recomendados a partir das preferências previamente fornecidas pelo utilizador, tais como cozinha tradicional, preços, ambiente, entre outros (Ricci, 2010).

Nos dias de hoje as pessoas utilizam dispositivos móveis em diferentes ocasiões, independentemente do tempo e do espaço, com o objetivo de obter informações relevantes acerca das suas necessidades ou interesses. Como tal, os sistemas de recomendação têm-se tornado cada vez mais populares neste tipo de dispositivos, devido à disponibilidade de diversos serviços de informação móveis (Sotsenko, Jansen, & Milrad, sd).

### 3. METODOLOGIA E TECNOLOGIA UTILIZADAS

Para a realização deste projeto foi necessário selecionar a tecnologia apropriada para a sua elaboração. Como tal, neste capítulo será descrita a tecnologia de suporte ao desenvolvimento do sistema.

#### 3.1 SCRUM

Scrum é o processo ágil mais usado para desenvolvimento de produtos, em particular para o desenvolvimento de software. É um *framework* de gestão de projetos que se pode aplicar a qualquer projeto com prazos agressivos, requisitos complexos e um elevado grau de singularidade<sup>3</sup>.

Este processo enfatiza a ideia do “controlo de processos empíricos”, isto é, utiliza um progresso real de um projeto – não um palpite ou uma previsão – para fazer planos e agendar lançamentos<sup>4</sup>.

Esta metodologia envolve três agentes<sup>5</sup>:

1. Proprietário do Produto, ou *Product Owner*, que estabelece os requisitos e principais características do produto a serem implementadas. Representa a voz do cliente e é geralmente representada pelos *stakeholders* ou alguém da equipa de *marketing*;
2. *ScrumMaster*, representa o gestor de projeto tendo as funções de supervisionar a equipa e o processo de desenvolvimento, garantindo a prática adequada dos princípios da metodologia Scrum;
3. Equipa de desenvolvimento, ou *Development Team*, que trabalham diretamente no desenvolvimento do produto e é tipicamente composta por 5-9 elementos com habilidades multifuncionais de analisar, projetar, desenvolver, testar, documentar, etc..

No Scrum os projetos são divididos em unidades básicas de desenvolvimento, chamadas *sprints*, que duram geralmente de uma semana a 4 semanas<sup>2</sup>; ver Figura 5.

---

<sup>3</sup> “Topics in Scrum” em Mountain Goat Software (<http://www.mountaingoatsoftware.com/agile/scrum/overview>)

<sup>4</sup> “Agile & Scrum Methods” em <https://benanistic.wordpress.com/2012/10/01/agile-scrum-methods>

<sup>5</sup> “Scrum Methodology” em <http://scrummethodology.com/>

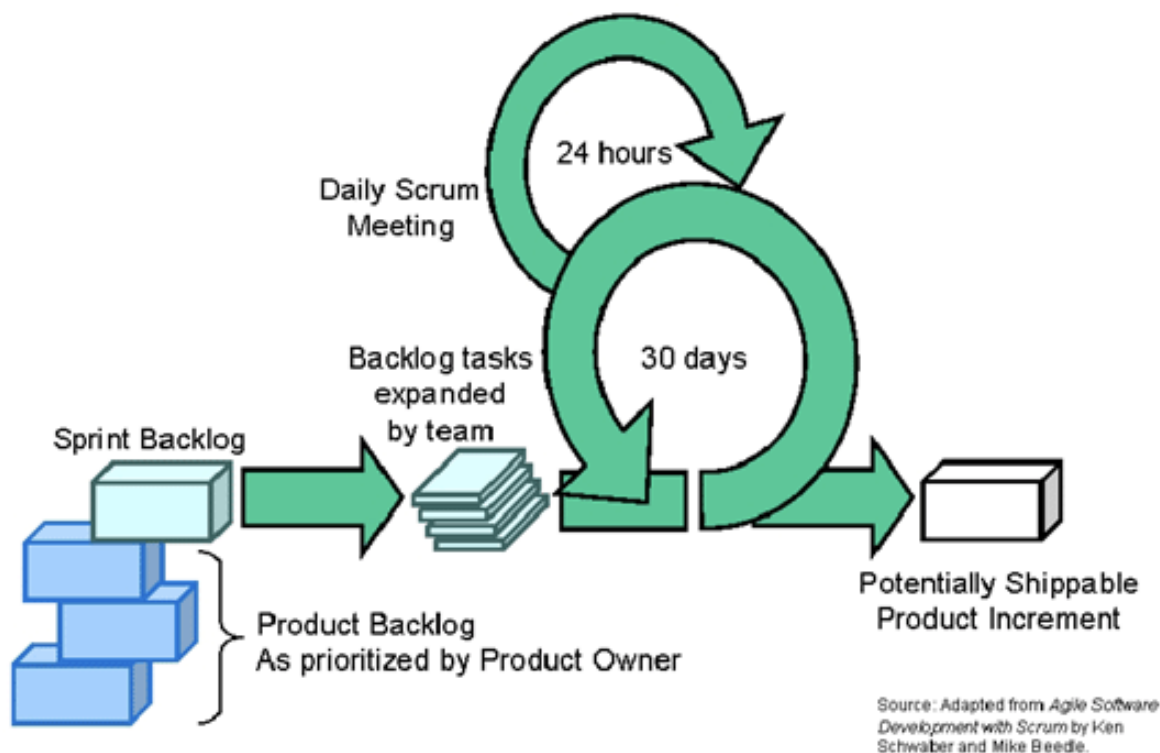


Figura 5 - Representação do processo Scrum (Fonte: <http://www.methodsandtools.com/archive/scrum1.gif>)

A primeira fase do processo é a fase de planeamento, onde o primeiro passo é a criação de um *backlog* do produto, uma lista de propriedades necessárias que terão de ser implementadas durante a fase de desenvolvimento, que é feito pelo dono do produto. De seguida a equipa faz a revisão do *backlog*, pensar que mudanças devem ser feitas para implementar novas propriedades e projetar o processo de implementação. No final são projetadas as reuniões de avaliação, nas quais as equipas trocam informações sobre a evolução do projeto<sup>6</sup>.

A fase de desenvolvimento, também conhecida como fase de *sprint*, é um ciclo iterativo de desenvolvimento, em que cada ciclo é constante em termos de duração. No início de cada *sprint* é realizada uma reunião de planeamento que é dividida em duas partes<sup>5</sup>:

1. Proprietário cria um *sprint backlog* baseado no *backlog* do produto, definindo qual o objetivo do *sprint*;
2. O *sprint backlog* é dividido em unidades (4 – 6 horas/sessões).

<sup>6</sup> “Scrum Phases” em Scrum Methodology (<http://www.scrummethodology.org/scrum-phases.html>)



Antes de iniciar o trabalho é aconselhável fazer-se uma breve reunião na qual cada pessoa da equipa diz o que fez no dia anterior e o que vai fazer no próprio dia, referindo problemas caso eles existam.

Depois do planeamento dá-se início ao processo de desenvolvimento do *sprint*, havendo a possibilidade de mudança de objetivos ou problemas mais complicados de resolver durante o processo. Neste caso, o *Scrum Master* ou o proprietário do produto podem parar o *sprint*, reiniciando-o depois de bem definidos os objetivos ou *backlog*. No final de cada será necessário<sup>7</sup>:

- Atualizar o gráfico de *burndown* (gráfico que representa o progresso do desenvolvimento, isto é, representa a porção de trabalho elaborada diariamente em comparação com o trabalho total planeado);
- Fazer parte da reunião de revisão do *sprint*. Os elementos apresentam os resultados obtidos para o proprietário do produto avaliar se os objetivos foram cumpridos;
- Fazer parte da reunião de retrospectiva do *sprint*. O *Scrum Master* em conjunto com a equipa decidem quais as tarefas que foram realizadas de forma apropriada e discutem maneiras de tornar o processo mais eficaz.

O processo de desenvolvimento do produto termina e este é preparado para ser lançado. Esta fase inclui: integração, testes, documentação do utilizador, formação e preparação do material de *marketing*<sup>8</sup>

Apesar de ser uma metodologia desafiadora a vários níveis, nomeadamente ao nível da mudança na gestão de projetos, exigindo reuniões diárias, definição do proprietário do produto, entre outros aspetos, o Scrum mostra-se particularmente vantajoso no sentido de conferir maior flexibilidade e adaptabilidade a eventuais mudanças nas exigências de trabalho. Isto significa que o projeto pode ser alterado dependendo das necessidades de mudança, de modo a conseguir, no final, estabelecer e entregar a versão mais apropriada do produto<sup>9</sup>.

## 3.2 Ferramentas Utilizadas

O mecanismo foi desenvolvido numa tecnologia de *backend* denominada JBoss Seam, um *framework* para aplicações *web*, através do servidor de aplicação JBoss 7.1. Este servidor de código

---

<sup>7</sup> "Scrum Phases" em Scrum Methodology (<http://www.scrummethodology.org/scrum-phases.html>)

<sup>8</sup> "Scrum Methodology" em <http://www.scrummethodology.org/>

aberto é baseado na plataforma JEE, que é uma plataforma de programação para servidores na linguagem de programação Java.

Para a concepção do mecanismo proposto foi necessária a instalação do ambiente de desenvolvimento da aplicação CardMobili, num computador com o sistema operativo Windows 8. O mecanismo foi implementado no lado do servidor, como tal poderá funcionar em qualquer sistema operativo móvel, utilizando a linguagem de programação Java 7. Esta é uma linguagem de programação orientada a objetos de extrema portabilidade, em que o mesmo código corre em diferentes plataformas, e que suporta processamento paralelo múltiplo.

O *Integrated Development Environment* (IDE) escolhido para trabalhar foi o Eclipse (plataforma Java), um IDE gratuito e que contém as ferramentas necessárias para programar em diversas linguagens, neste caso em Java.

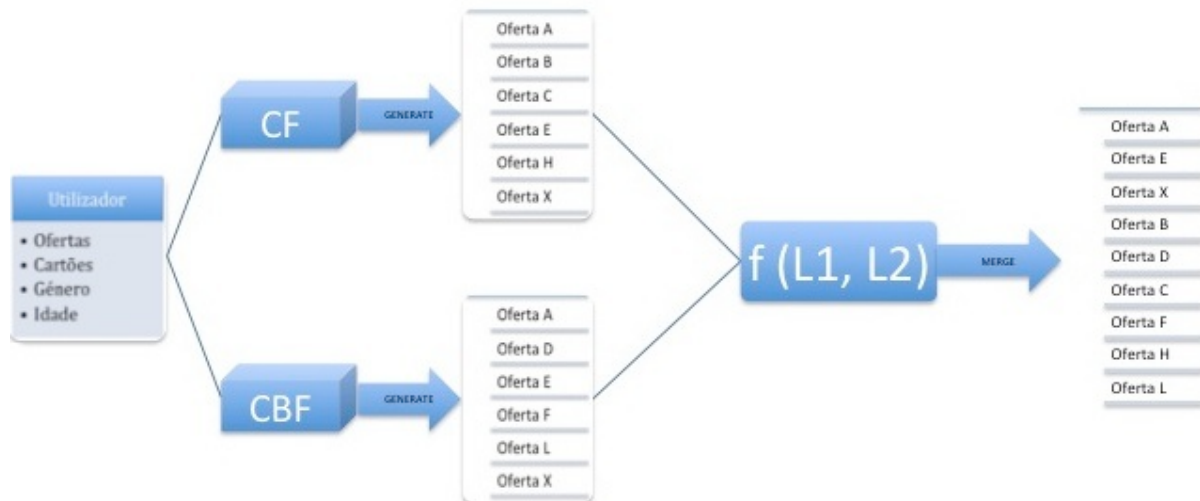
Para lidar com os dados foi instalada a ferramenta MySQL Workbench, que utiliza a linguagem *Structured Query Language* (SQL) como interface, embora no projeto a linguagem utilizada tenha sido *Hibernate Query Language* (HQL), do *Object-Relational Mapping* (ORM), por ser uma linguagem totalmente orientada a objetos, que fornece *Application Programming Interfaces* (APIs) simples para armazenar e recuperar objetos Java diretamente de e para a base de dados.

Foi ainda utilizado um sistema de controlo de versão, Apache Subversion (SVN), pela capacidade de controlar diferentes versões usadas em computadores separados, recuperar versões antigas de documentos e examinar o histórico de alterações ocorridas em cada ficheiro do projeto.

### **3.3 Algoritmos Escolhidos**

A investigação efetuada acerca dos sistemas de recomendação permitiu concluir que um sistema de filtragem híbrida funciona de forma mais eficaz e eficiente do que sistemas com apenas um dos algoritmos de filtragem colaborativa ou baseada em conteúdos. Assim, neste projeto decidiu-se optar por um sistema de filtragem híbrida com combinação mista para o mecanismo a desenvolver.

A Figura 6 mostra o processo do algoritmo de filtragem híbrida adotado, adaptado do algoritmo Fusion apresentado por Torres *et al.* (2004).



**Figura 6 - Ilustração do algoritmo de filtragem híbrida**

Este algoritmo suporta-se em várias etapas:

1. Filtragem Colaborativa e Filtragem Baseada em Conteúdos são realizadas separadamente gerando duas listas distintas, L1 e L2, respetivamente;
  - a. A Filtragem Colaborativa escolhe as ofertas mais relevantes para a vizinhança do utilizador atual. Essa vizinhança é formada pelos utilizadores que apresentam maior similaridade com o utilizador, calculada através do coeficiente de correlação de Pearson, e que já mostraram interesse pela oferta que está a ser avaliada;
  - b. A Filtragem Baseada em Conteúdos gera uma lista de ofertas que apresentam as mesmas palavras-chave que o perfil do utilizador atual. Este perfil é criado com as palavras-chave das ofertas presentes na carteira digital do utilizador e com as palavras-chave e nomes dos cartões que o utilizador possui;
2. As listas geradas são comparadas e combinadas numa só lista;
3. As ofertas que se encontrem presentes em ambas as listas são colocadas no início da lista final. As restantes ofertas são adicionadas à lista final alternadamente;
4. A lista final é apresentada ao utilizador como recomendação.

Após a pesquisa de diversos algoritmos, este aparentou ser o mais apropriado para a aplicação, uma vez que atribui maior relevância a ofertas comuns às duas filtragens distintas, não rejeitando as restantes. Desta forma, nenhuma oferta potencialmente interessante será excluída das recomendações.

## 4. CASO DE ESTUDO: CARDMOBILI

O presente projeto de investigação foi desenvolvido em colaboração com a empresa CardMobili, Desenvolvimento de Software, S.A.. Este capítulo destina-se a apresentar a empresa e os seus produtos e soluções, apoiado na informação existente na página de internet da empresa<sup>9</sup>, os quais traduzem parte do modelo de negócio da CardMobili (Anexo I).

### 4.1 A Empresa

A CardMobili, Desenvolvimento de Software S.A. é uma empresa de desenvolvimento de aplicações móveis de carteira digital, fundada em 2009 por António Murta e Carlos Oliveira, com sede em Leça do Balio, concelho de Matosinhos. Esta empresa opera na Europa, América Latina, Médio Oriente e Norte de África e é membro do *Interactive Advertising Bureau* (IAB) e da *Mobile Marketing Association* (MMA).

Há uma forte ligação entre os consumidores e os novos serviços que podem ser oferecidos pelas Carteiras Digitais. Esta visão impulsiona todos os aspetos do negócio da CardMobili – inovação tecnológica, desenvolvimento de soluções e evolução, para implementações bem sucedidas numa plataforma de carteira digital única baseada na nuvem.

As soluções da CardMobili são projetadas para ajudar a desbloquear novas e importantes fontes de receita a partir de serviços de carteiras digitais. A já significativa experiência da empresa permitiu criar plataformas que proporcionam às operadoras móveis, retalhistas e instituições de serviços financeiros, uma base inovadora para a sua próxima geração de carteiras digitais.

A empresa pretende que a sua base de produtos seja facilmente configurável e personalizável de modo a devolver o resultado mais adequado aos seus clientes.

### 4.2 Produtos

A empresa age como fornecedor-chave de tecnologia para a carteira digital e serviços móveis de valor acrescentado, através de fidelidade, campanhas, comunicação, pagamentos e compras. Com este *portfolio* de produtos, os clientes da CardMobili podem obter vários serviços e vantagens que potenciam os seus lucros.

---

<sup>9</sup> Consultar: <http://corporate.cardmobili.com>

Os produtos da CardMobili são os seguintes:

- M | Loyalty

Este produto fornece um conjunto abrangente de recursos para apoiar a criação e gestão de campanhas móveis e totalmente digitais e programas de ofertas, bem como o alargamento de programas de fidelidade existentes para dispositivos móveis. É um produto flexível projetado para atender às exigências de CRM e de fidelização únicas da empresa, oferecendo uma experiência digital de consumo que fortalece a fidelização à marca. Suporta o ciclo de vida inteiro de um programa de fidelização, desde o registo e aquisição do cliente até à gestão de cartões de clientes, gestão de benefícios e execução do programa.

Este produto pode ser aplicado em:

- Programas de fidelização;
- Programas de coligação;
- Programas de redes de parceiros;
- *Punch Cards*;
- Esquemas de acumulação e resgate de pontos;
- Programas e presentes simples ou por categorias.

- M | Communication

Este produto oferece poder de atendimento ao cliente de uma marca, quer através de uma interface de comunicação autónoma, quer através da integração da sua ampla funcionalidade com os sistemas existentes.

Este produto apoia as operações de gestão de marca e experiência do cliente, construindo fidelidade e apoiando todas as fases do ciclo de vida do cliente, permitindo conectividade em tempo real para dispositivos móveis e fornecendo relatórios sobre as interações com clientes.

O M | Communication pode ser aplicado para:

- Serviços de comunicação e *marketing*;
- *Help Desk* integrado;
- Serviços de atendimento móvel ao cliente;
- *Feedback* para a marca;

- Comunicação *one-to-one*, não invasiva, de valor acrescentado, com o consumidor.

- M|Campaigns

Este produto permite a criação e gestão de campanhas promocionais segmentadas, digitais e móveis – tanto de forma independente, como integrado em sistemas de gestão de campanhas existentes. Oferece um alto nível de convergência digital, com a capacidade de aumentar as transações e compromisso entre comerciantes e marcas. Ao apoiar uma experiência digital centrada no cliente, este produto permite a transferência de ofertas em qualquer momento e lugar.

Este produto pode ser aplicado para:

- Ligação de fidelidade ou ofertas autónomas;
- Ofertas de comerciantes, descontos e cupões;
- Cupões grátis e pré-pagos;
- Campanhas publicitárias;
- Promoções diárias ou campanhas de vendas repentinas;
- *Vouchers* e certificados de presentes.

- M|Shopping

Este produto apoia a criação e gestão de uma experiência de comércio móvel, através da integração com sistemas de tecnologia de negócios *m-commerce* existentes. Combina um conjunto abrangente de recursos e capacidades com características de *merchandising*, possibilitando a comercialização de produtos . M|Shopping pode acrescentar valor tanto à forma como os clientes compram, como à forma de pagamento, com a presença do dispositivo móvel, influenciando as decisões de compra e de prestação de assistência dentro da loja ao cliente e ao comerciante. Pode ser aplicado para:

- Comprar e encomendar;
- Carrinho de compras e *checkout*;
- Digitalizar para comprar;
- Campanhas comercializáveis;
- Vitrina de loja.

- M|Payments

Fornece a capacidade de pagamento e acesso a compras, histórico de pagamentos e notas fiscais a qualquer hora e em qualquer lugar. A empresa apresenta-o como um produto robusto e flexível que integra todas as infraestruturas de pagamento, seja de forma independente ou como uma extensão das redes de pagamento existentes, e que permite ligar pagamentos com programas de *marketing* e estratégias de campanhas, tais como cartões de fidelização e campanhas de ofertas, para fornecer aos clientes de uma empresa uma experiência integrada com mais valor, sempre que estes interagem com a marca. Para a CardMobili este produto representa um catalisador para o comércio móvel

M|Payments pode ser aplicado para:

- Ofertas ligadas a cartões e programas de *cashback*;
- Soluções de pagamento em circuito fechado;
- Cartões de presente e pagamentos digitais;
- Pagamentos *in-app* e compras;
- Transferências de dinheiro e pagamentos.

### 4.3 Soluções

A Carteira Digital representa uma oportunidade de mercado atraente e de rápido crescimento, impulsionando mudanças na experiência de compra do consumidor e na criação de um importante potencial de novas receitas. Velocidade e criatividade são essenciais para o sucesso, exigindo soluções que combinam uma operação comercial eficiente com novos serviços de valor acrescentado, um componente-chave para a realização de uma carteira digital eficiente e implementações de estratégias móveis.

- Operadoras de Redes Móveis (MNO)

É uma solução totalmente *end-to-end*, sustentada por uma plataforma de tecnologia comprovada, a qual encontra-se pronta a ser implementada de imediato. É uma solução que mantém uma posição importante no mercado, combinando três elementos essenciais que são a chave da proposição de valor da carteira digital: uma marca de confiança dos consumidores; uma relação bem estabelecida com clientes e conhecimento do mercado local; vantagem da existência de uma

infraestrutura. Como agregadores de conteúdos e serviços, as operadoras de redes móveis situam-se efetivamente no centro da oportunidade da carteira digital.

O conjunto de produtos da CardMobili considera os seguintes elementos do núcleo da Carteira Digital:

- Fidelidade, recompensas e programas de adesão que acrescentam valor para o consumidor;
  - Conteúdos e recursos baseados em localização: para onde ir, onde comprar, etc.;
  - Instrumentos de comunicação entre comerciantes e consumidores para o compromisso da marca adicional;
  - Ofertas, promoções, cupões, descontos, disponibilizando todas as formas de *marketing* e serviços promocionais para os consumidores;
  - Pagamentos móveis integrados para maximizar a experiência de compras em geral, com foco em todos os serviços que dirigem o valor do consumidor e do comerciante.
- 
- Retalhistas

Com a solução de retalho da CardMobili, os comerciantes podem usufruir de uma carteira digital de marca única, o que irá permitir aos retalhistas oferecer recursos individualmente adaptados a clientes e parceiros de forma a incentivar a atividade de compras *online* ou *offline* e aproximá-los da marca. Também ajuda a criar vantagem competitiva, transformando o curso de compras e a experiência de pagamento.

- *Mobile Loyalty*, apoiando o ciclo de vida completo do programa, desde a aquisição e registo de clientes, até à gestão de cartões de cliente, gestão de benefícios e execução de programas;
- Gestão de Campanhas, oferece aos retalhistas a propriedade do relacionamento com o cliente e um canal de comunicação dinâmico - com entrega de ofertas a qualquer momento e em qualquer lugar;
- Circuito fechado e indicação privada de pagamentos, ligando pagamentos com programas de *marketing* e campanhas estratégicas;
- Conteúdos e recursos baseados em localização: para onde ir, onde comprar, etc.;
- Ofertas, promoções, cupões, descontos: disponibilizando todas as formas de *marketing* e serviços promocionais aos consumidores;
- Oferecendo uma nova experiência de comércio móvel, a ligação com praticamente qualquer comerciante *backend* e possibilitando uma comercialização do produto verdadeiramente eficaz.



- Serviços Financeiros

Ao escolher uma solução de Carteira Digital da CardMobili as instituições financeiras podem melhorar as suas atividades bancárias centradas no cliente e avançar para a convergência *online/offline* genuína. Esta solução combina uma série de vantagens exclusivas - incluindo serviços relacionados a comércio de valor acrescentado, pagamentos móveis bem como ofertas ligadas ao cartão e programas de *cashback*.

No centro desta solução está um conjunto de produtos que estabelece os seguintes elementos na carteira digital:

- Lealdade, recompensas e programas de adesão que acrescentam valor para o consumidor;
- Ferramentas para oferecer uma nova geração de experiência de comércio móvel, conectando com praticamente qualquer comerciante de *backend* e possibilitando a comercialização de produtos;
- Conteúdos e recursos baseados em localização: para onde ir, onde comprar, etc.;
- Instrumentos de comunicação entre comerciantes e consumidores para o comprometimento adicional da marca;
- Ofertas, promoções, cupões, descontos: disponibilizar todas as formas de *marketing* e serviços promocionais para os consumidores;
- Pagamentos móveis integrados para maximizar a experiência de compras em geral, com foco em todos os serviços que incrementam o valor do consumidor e comerciante.

A carteira digital representa muito mais do que uma nova plataforma de pagamentos. É uma nova proposta bancária conveniente, que combina atividades bancárias tradicionais com um canal complementar para novos produtos financeiros. É, portanto, a chave para incentivar os clientes a aumentar os gastos do cartão e a mudar decisões de compra de um cartão específico ou métodos de pagamento. Também abre o caminho para oportunidades de prestação de serviços diferenciados com os comerciantes, construindo novas pontes para o comércio futuro.

## 4.4 Tecnologia

A abordagem da CardMobili é permitir um ecossistema completo de carteira digital, com instalações de administração e gestão diferenciados que proporcionam benefícios para todos os participantes. Na Figura 7 apresenta-se a plataforma da CardMobili.

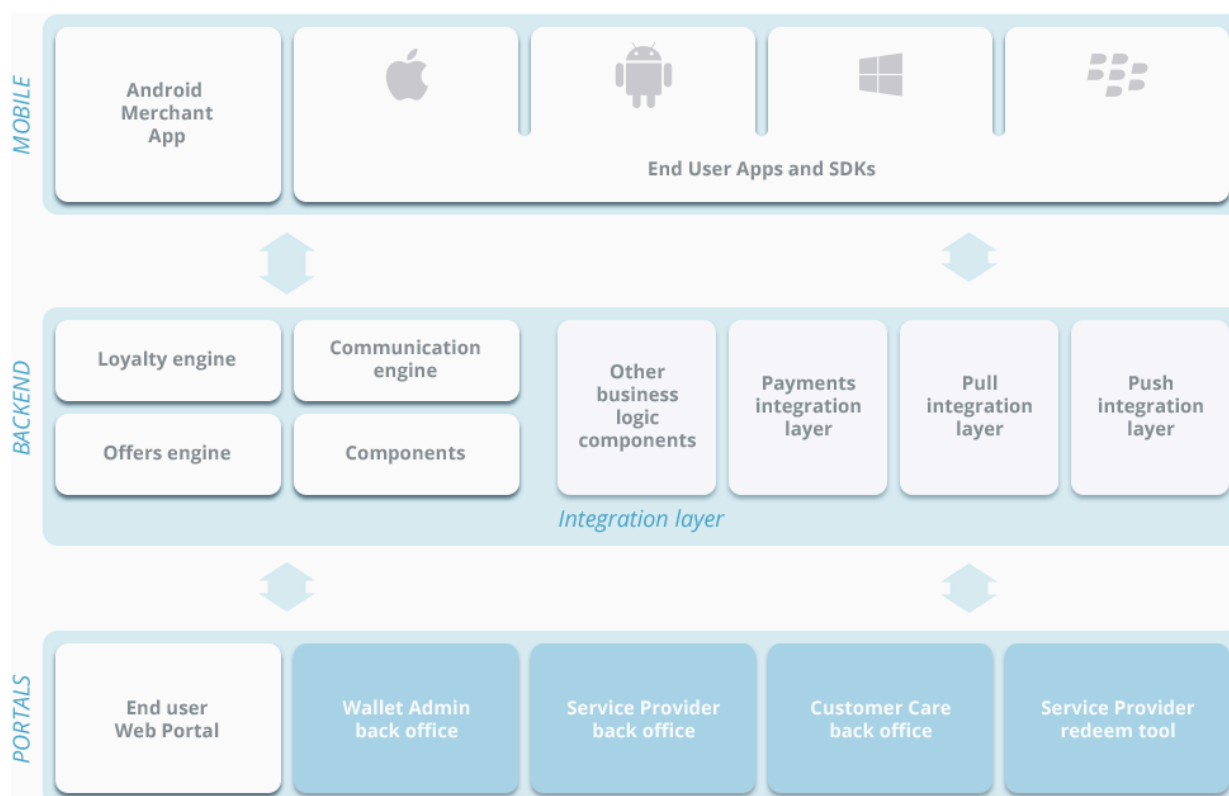


Figura 7 – Tecnologia da Plataforma de Carteira Digital (Fonte: <http://corporate.cardmobili.com/en/technology>)

Como um dos núcleos de tecnologia mais avançada disponível atualmente, a plataforma da CardMobili incorpora uma multiplataforma *web* e aplicações móveis para o utilizador final, um servidor de *backend* robusto baseado na nuvem, aplicações móveis e *web* para prestadores de serviços de terceiros, bem como mecanismos de integração completos, ferramentas de *backoffice* e ligações para qualquer fonte externa de cupões. A plataforma oferece ainda suporte a uma grande variedade de cenários de implementação, desde o *rebranding* completo, ou recursos de incorporação numa aplicação já existente, até uma solução de *backend* completa ou um simples projeto de portal *web*;

permitted to companies manage all aspects of the digital portfolio – through loyalty and rewards, offers and coupons, electronic receipts, product information and payments<sup>10</sup>.

The company's approach offers brands the tools to accept the change of their customers to a mobile system considering it as economical, exponential and perfectly possible, providing a platform and a set of solutions with which it is possible to prolong marketing programs and reach commercially viable digital portfolio implementations.

---

<sup>10</sup> <http://corporate.cardmobili.com/en/technology>

## 5. DESENVOLVIMENTO DO MECANISMO DE RECOMENDAÇÃO

A recomendação de conteúdos é uma técnica de CRM à qual recorre um crescente número de empresas para oferecer aos seus consumidores produtos e serviços adequados às suas preferências.

Na secção “Top Offers” da aplicação da CardMobili, o facto de as ofertas aparecerem em função do número de *downloads*, faz com que surjam frequentemente ofertas repetidas. Deste modo, os utilizadores tendem a considerá-las como SPAM e acabam por não disfrutar da experiência que a aplicação pode oferecer. Considerando este problema, surge a necessidade de criar um sistema de recomendação de conteúdos, ofertas e, posteriormente, programas de fidelização, que sejam potencialmente interessantes para cada utilizador. Um sistema com estas características, num primeiro momento apenas direcionado para as ofertas de top da aplicação, a experiência de utilização da aplicação vai decerto tornar-se mais satisfatória e intuitiva, assumindo que não será necessário qualquer esforço adicional para aceder às ofertas que serão mais apelativas para o consumidor.

O utilizador terá acesso a ofertas ordenadas por ordem de relevância e cada oferta terá uma lista de ofertas relacionadas, considerando o seu conteúdo.

Neste capítulo, explica-se a implementação do algoritmo escolhido e do desenvolvimento do mecanismo de recomendação, mostrando as funcionalidades do mesmo.

### 5.1 Especificação Funcional

A especificação funcional descreve com detalhe as funcionalidades, comportamentos e interações com o utilizador, que o sistema terá, e é usada como referência pelos programadores que irão desenvolver o sistema.

#### 5.1.1. Requisitos Funcionais

Os requisitos funcionais definem as funções de um sistema de software. Eles podem ser cálculos, detalhes técnicos, manipulação de dados e outras funcionalidades específicas que definem um sistema.

O mecanismo desenvolvido apresenta os seguintes requisitos funcionais:

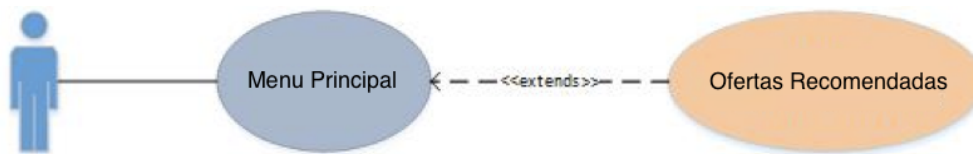
1. Mostrar top de ofertas recomendadas no menu inicial da aplicação móvel;
2. Mostrar ofertas relacionadas na descrição de cada oferta na aplicação móvel;
3. Calcular a similaridade entre os utilizadores através dos interesses de cada um;

4. Calcular a predição de interesse do utilizador pelas ofertas que ainda não conhece;
5. Mostrar as ofertas ordenadas por relevância.

### 5.1.2. Casos de Uso

Os casos de uso descrevem cenários de possíveis interações com o utilizador ou com um outro sistema. Cada caso de uso deve descrever somente uma funcionalidade. Aqui são descritas as duas principais funcionalidades do mecanismo desenvolvido, recomendação de ofertas e ofertas relacionadas.

A Figura 8 representa a interação do utilizador com o sistema quando requisita uma lista de ofertas recomendadas, seguida da Tabela 2 onde essa interação é detalhada.



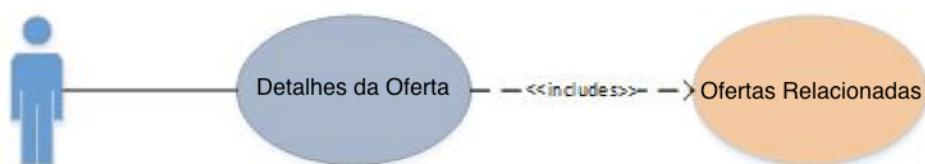
**Figura 8 - Caso de Uso para Recomendações de Ofertas**

**Tabela 2 - Descrição do caso de uso "Recomendações de Ofertas"**

<i>Nome do Use Case</i>	Recomendações de Ofertas
<i>Descrição</i>	O utilizador seleciona a opção Pesquisa de Ofertas e automaticamente a aplicação apresenta o Top de Ofertas.
<i>Atores</i>	Utilizador final
<i>Pré-condições</i>	Haver pelo menos duas ofertas O utilizador já usou a aplicação por algum tempo (tem cartões, adquiriu ofertas, tem um perfil relativamente rico em informações, etc.) Existe conectividade
<i>Pós-condições</i>	Se não houver ofertas disponíveis no sistema ou conectividade no dispositivo surge uma notificação de erro Se houver ofertas estas são segmentadas de acordo com os interesses do utilizador e apresentadas na secção Top de Ofertas Se o utilizador não tiver histórico de utilização da aplicação as ofertas apresentadas no Top de Ofertas são ordenadas por ordem decrescente de

	<i>downloads</i>
<i>Basic Flow</i>	<ol style="list-style-type: none"> <li>1. Utilizador seleciona Pesquisa de Ofertas</li> <li>2. São disponibilizadas 3 opções de pesquisa</li> <li>3. Seleciona Top Offers</li> <li>4. Verifica que as ofertas que existem no sistema são realmente relevantes</li> <li>5. Mostra a lista final de recomendações</li> </ol>

O caso de uso de Ofertas Relacionadas está representado na Figura 9, e detalhado na Tabela 3.



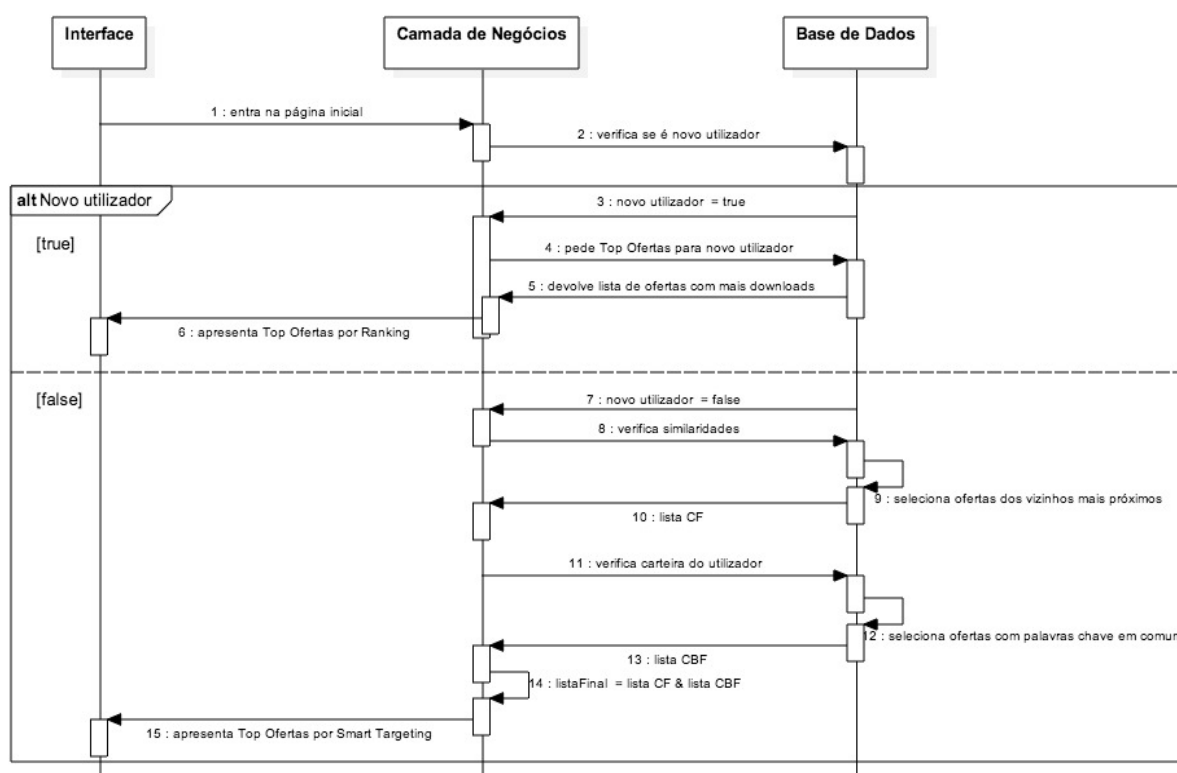
**Figura 9 - Caso de Uso para Ofertas Relacionadas**

**Tabela 3 - Descrição do caso de uso "Ofertas Relacionadas"**

<i>Nome do Use Case</i>	Ofertas Relacionadas
<i>Descrição</i>	Enquanto visualiza uma oferta o utilizador pode ter acesso a ofertas que estejam relacionadas com a atual
<i>Atores</i>	Utilizador final
<i>Pré-condições</i>	Existência de conectividade Existência de ofertas relacionadas
<i>Pós-condições</i>	Se não houver conectividade aparece no ecrã uma mensagem de erro Se não houver ofertas relacionadas aparece um aviso Caso contrário aparece uma lista de ofertas com conteúdo semelhante ao da atual
<i>Basic Flow</i>	<ol style="list-style-type: none"> <li>1. O utilizador visualiza uma oferta</li> <li>2. A lista de ofertas semelhantes aparece no final da página</li> </ol>
<i>Alternative Flow</i>	<ol style="list-style-type: none"> <li>1. O fluxo alternativo consiste em adicionar um botão chamado "Ofertas Relacionadas" para o menu de contexto da oferta</li> </ol>

### 5.1.3. Diagramas de Sequência

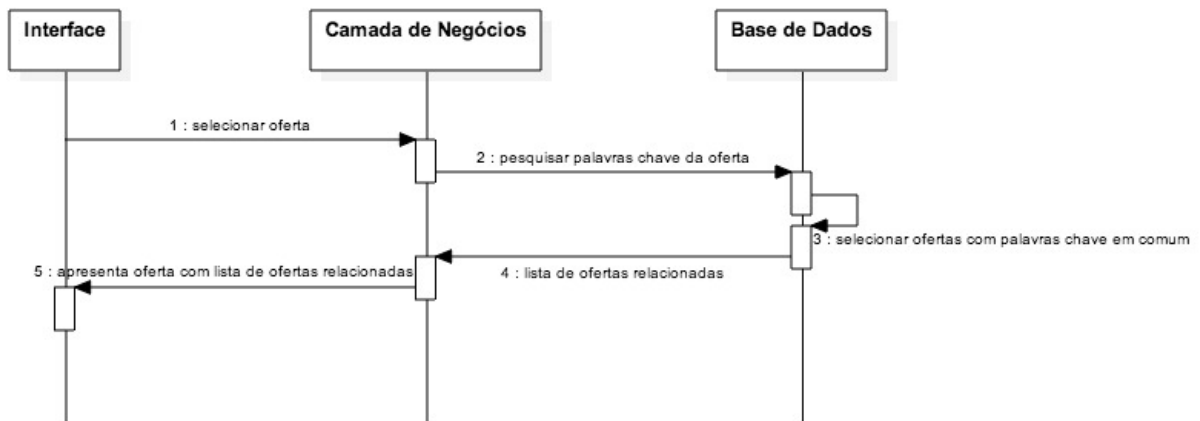
Os diagramas de sequência representam a sequência de processos e mensagens trocadas entre os objetos intervenientes num caso de uso, seguindo uma ordem temporal. A Figura 10 representa o diagrama de sequência relativo ao caso de uso “Recomendações de Ofertas”.



**Figura 10 - Fluxo de chamadas efetuadas para apresentar o Top de Ofertas**

Quando um utilizador entra na página inicial o sistema verifica automaticamente se se trata de um novo utilizador ou não. Caso se trate de um novo utilizador as ofertas apresentadas no Top correspondem às ofertas com maior número de downloads. Caso contrário, são calculadas duas listas, uma lista de ofertas por filtragem colaborativa e outra lista de ofertas baseadas em conteúdos. Estas duas listas são combinadas gerando uma lista final que é apresentada no ecrã do telemóvel.

A Figura 11 corresponde ao diagrama de sequência do caso de uso “Ofertas Relacionadas”.



**Figura 11 - Fluxo de chamadas efetuadas quando se visualiza uma oferta**

Ao visualizar uma oferta o sistema procura automaticamente ofertas que lhe estejam relacionadas, através das palavras-chave que lhe estão associadas, nome do comerciante, categoria, etc. Esta lista de ofertas relacionadas é apresentada no final da página que apresenta a oferta.

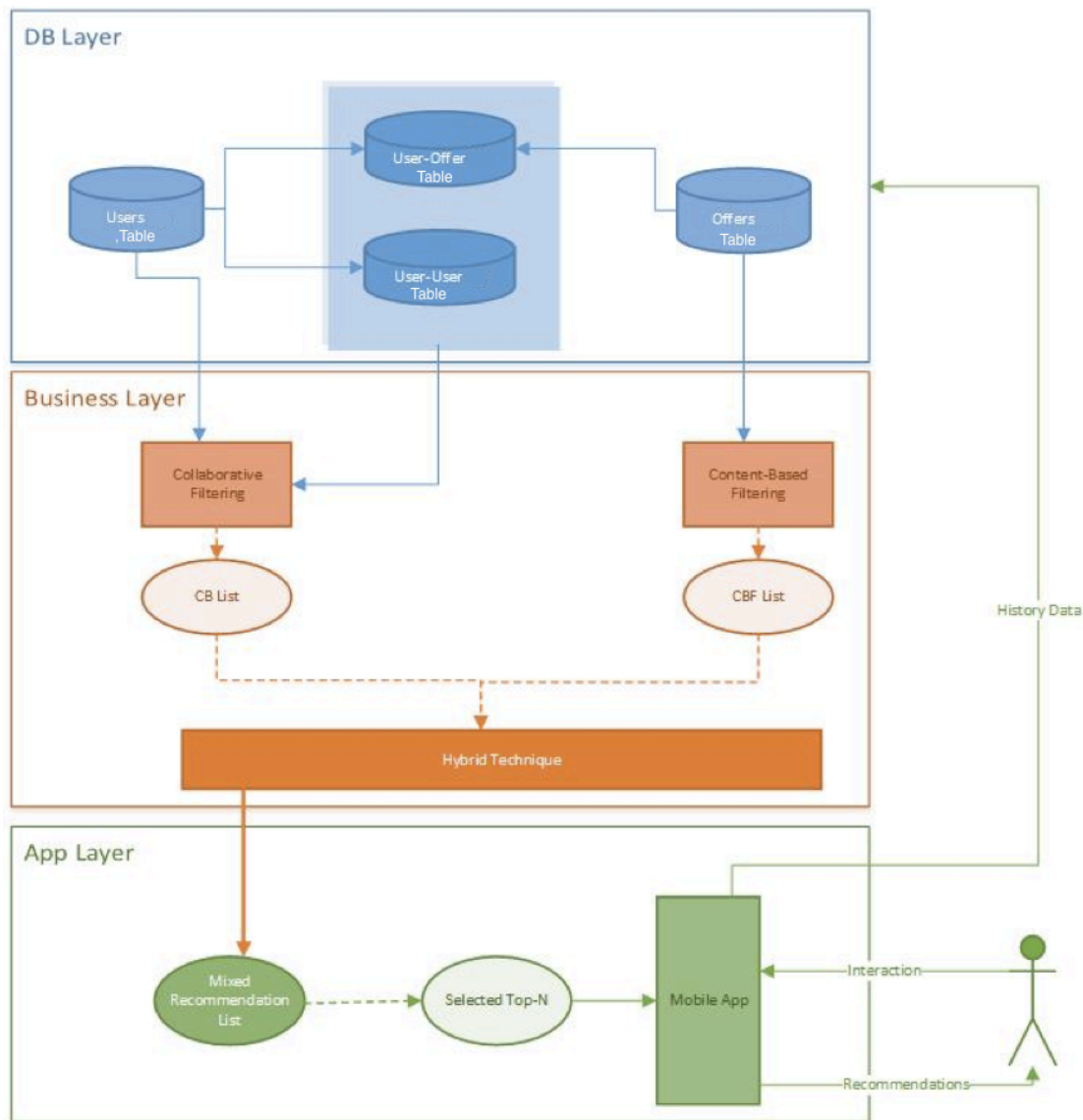
## 5.2 Especificação Técnica

Com o apoio da especificação funcional, a especificação técnica detalha a implementação do sistema, principais métodos, classes, armazenamento de dados, etc.

### 5.2.1. Arquitetura do Sistema

Na Figura 12 está representada a arquitetura geral do mecanismo desenvolvido. Estão apresentadas três camadas: camada de dados, camada de negócios e camada de apresentação.





**Figura 12 - Arquitetura do mecanismo desenvolvido**

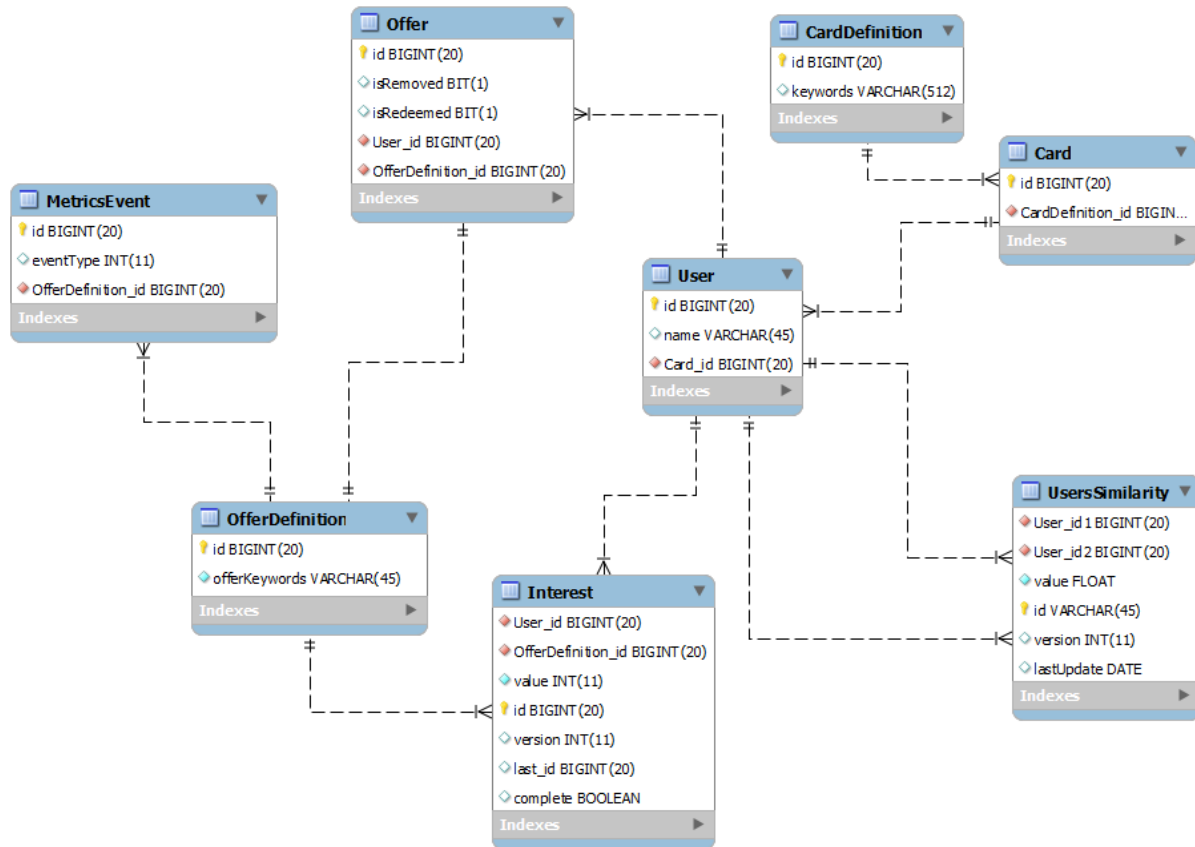
Na primeira camada estão envolvidos os dados acerca dos utilizadores, das ofertas, uma tabela que relacionam os utilizadores entre si, de similaridade, e uma tabela que relaciona os utilizadores com as oferta, de interesse.

A camada lógica de negócios é aquela onde todos os cálculos são realizados para assegurar o correto funcionamento do mecanismo. É vista como uma ligação entre os dados e a interface, uma vez que suporta as funções e as regras que trabalham sobre a informação proveniente da camada de dados e cujos resultados são apresentados na camada de apresentação.

A camada de apresentação é a camada que interage diretamente com o utilizador disponibilizando-lhe todas as funcionalidades do mecanismo. Neste caso, esta camada já estava desenvolvida pelo que só foi utilizada para apresentação dos resultados.

#### 4.2.1. Modelo de Dados

Na Figura 13 pode consultar-se o esquema remodelado da base de dados que permite o funcionamento do mecanismo de recomendação de ofertas.



**Figura 13 - Modelo de Base de Dados do mecanismo desenvolvido**

Neste caso, foi apenas necessário proceder à criação de duas entidades, a tabela *Interest*, apresentada na Tabela 4, e a tabela *UsersSimilarity*, apresentada na Tabela 5.

**Tabela 4 - Descrição dos campos da tabela Interest da base de dados**

Interest	
<b>id</b>	Identificador da entrada na tabela, é gerado automaticamente, incrementado a cada nova entrada
<b>version</b>	Identificador de controlo da versão que está a aceder à entrada em questão.
<b>user_id</b>	Valor de identificação do utilizador.
<b>offerDefinition_id</b>	Valor de identificação da oferta.

<b>value</b>	Valor do interesse demonstrado pelo utilizador identificado pela oferta identificada.
<b>last_id</b>	Valor correspondente ao identificador da tabela que foi verificado em último lugar na última vez que foi calculada a similaridade entre os utilizadores. Pára de atualizar quando a variável <i>complete</i> for igual a 1.
<b>complete</b>	Se este valor for igual a zero significa que o utilizador desta entrada ainda não foi comparado com todos os utilizadores da tabela (auxiliar para calcular valor da tabela de similaridade). Se o valor for 1 significa que está completo o processo de cálculo de similaridade para este utilizador.

**Tabela 5 - Descrição dos campos pertencentes à tabela UsersSimilarity da base de dados**

UsersSimilarity	
<b>id</b>	Identificador da entrada na tabela, é gerado automaticamente, incrementado a cada nova entrada
<b>version</b>	Identificador de controlo da versão que está a aceder à entrada em questão
<b>user1_id</b>	Valor de identificação do utilizador 1
<b>user2_id</b>	Valor de identificação do utilizador 2
<b>value</b>	Valor da similaridade entre o utilizador 1 e o utilizador 2
<b>lastUpdate</b>	Data da última atualização da entrada na tabela

Na primeira coluna, de cada tabela, encontra-se o nome do campo que é descrito posteriormente na segunda coluna.

### 5.3 Implementação

Como o projeto incluía a implementação de um mecanismo de recomendação que será posteriormente integrado numa aplicação já existente, foi necessário criar novas entidades e alterar outras.

### 5.3.1. Novas Entidades

Neste projeto cada entidade corresponde, no fundo, a uma tabela na base de dados. Esta entidade é representada numa classe em java, cujas variáveis de instância representam as colunas da tabela.

#### **Interest**

É a entidade que representa o interesse do utilizador X pela oferta Y. As variáveis que descrevem esta entidade já foram referidas e explicadas na secção anterior. Como foi explicado na secção 2.3. deste documento, a filtragem colaborativa passa por um processo de recolha da opinião dos utilizadores acerca de um produto, traduzindo-se esta opinião num valor de interesse do utilizador por esse item. No entanto, a aplicação da CardMobili ainda não tem essa opção implementada pelo que foi necessário recorrer a uma alternativa para perceber os interesses do utilizador.

Através da aplicação, o utilizador pode ter várias ações sobre as ofertas disponíveis. Pode visualizar, comprar, descarregar ou resgatar. A cada ação destas foi associado um valor, que pode simbolizar a relevância da oferta para o utilizador atual, por exemplo, se o interesse atual de um utilizador por determinada oferta for  $x$ , é incrementado um certo valor estipulado dependendo do evento que ocorrer:

- Visualização da oferta →  $x + 1$
- Download da oferta →  $x + 2$
- Partilha a oferta →  $x + 3$
- Compra da oferta →  $x + 4$
- Resgate da oferta →  $x + 5$

#### **UsersSimilarity**

Esta entidade representa a similaridade entre dois utilizadores, isto é, o grau de semelhança entre dois utilizadores tendo em conta as ofertas pelas quais eles mostraram interesse. A similaridade é calculada através do Coeficiente de Correlação de Pearson (Anexo II) e mostra que dois utilizadores são tão mais semelhantes quanto mais próximo de 1 for o valor do coeficiente, e conseqüentemente tão mais distantes quanto mais próximo de -1 for o coeficiente.

### 5.3.2. Triggers

Em Sistemas de Gestão de Bases de Dados, um *trigger*, ou gatilho, é um procedimento em SQL que inicia uma ação quando ocorre um evento que altere uma tabela na base de dados (*Insert*, *Delete* ou *Update*). Um *trigger* não pode ser chamado ou executado (como uma função), este é ativado automaticamente pelo sistema, como resultado de uma modificação na tabela à qual ele está associado. Os *triggers* são usados para manter a consistência dos dados ou para divulgar as alterações em determinados dados de uma tabela para outras.

### **actionOfferInterests**

Este *trigger* (Anexo III) é acionado cada vez que há uma inserção na tabela *MetricsEvent*. Sempre que acontece um evento este é guardado na tabela *MetricsEvent*, caso este evento corresponda a alguma ação do utilizador sobre uma oferta (visualizar, guardar, resgatar, etc.) este *trigger* identifica o tipo de ação e atualiza a tabela *Interest*. Caso não exista nenhum interesse associado a um utilizador e determinada oferta, uma nova entrada é inserida na tabela *Interest*. Caso já exista, apenas o interesse é atualizado, somando o valor correspondente à ação efetuada.

### 5.3.3. Jobs

Jobs são programas que são executados sem que o utilizador final se aperceba da sua execução, i.e. ocorrem em *background*, e sem intervenção humana. Os Jobs podem ser programados para começar em determinado momento e para serem executados periodicamente. Por outro lado, podem ser utilizados para não sobrecarregar o sistema quando existe uma grande quantidade de informação armazenada e que deve ser processada.

### **InsertUserSimilarity**

Este *job* (Anexo IV) foi criado para inserir na tabela *UsersSimilarity* a similaridade de um novo utilizador com os restantes, isto é, sempre que um utilizador que nunca tenha mostrado interesse por alguma oferta, mostra pela primeira vez interesse, este job calcula a similaridade com todos os outros utilizadores que já mostraram interesse por essa e outras ofertas, guardando o seu valor na tabela *UsersSimilarity*.

Ele recebe a primeira entrada da tabela *Interest* que apresente a coluna *complete* igual a zero e verifica o seu *last\_id*. Depois vai buscar as X entradas que seguem o *last\_id* (neste caso X = 5, mas pode ser escolhido outro valor), e calcula a similaridade entre os utilizadores associados a essas

entradas, apenas se os utilizadores forem diferentes e ainda não tenha sido calculada a sua similaridade, guardando o valor na tabela *UsersSimilarity*.

Para efeitos de testes este *job* foi iniciado 1 minuto depois da execução da aplicação, para correr de 2 em 2 minutos.

### **UpdateUserSimilarity**

Este *job* (Anexo V) calcula a similaridade de pares de utilizadores que já existem na tabela *UsersSimilarity* e atualiza-o caso este seja diferente. Primeiro escolhe as X entradas (neste caso X = 5, podendo este valor ser alterado) que foram atualizadas há mais tempo, escolhidas através da coluna *lastUpdate*, e para cada par de utilizador calcula a similaridade entre eles, tendo em conta as ofertas que se mostraram relevantes para eles (na tabela *Interest*), e caso a similaridade seja diferente este *job* altera o valor da respetiva entrada na tabela.

Para efeitos de testes este *job* foi iniciado 2 minutos depois da execução da aplicação, para correr de 5 em 5 minutos.

#### 5.3.4. Entidades e Classes Alteradas

Nesta secção apenas são indicados os métodos criados em entidades ou classes que já existiam, indicando os parâmetros e explicando a finalidade de cada método.

### **OfferDefinition**

Esta entidade corresponde ao *template* de uma oferta, contendo as características gerais que uma oferta deve possuir, enquanto que a entidade Offer corresponde a uma instância desta classe.

- `getKeywordsListByUser(EntityManager, User)` que devolve como resultado o conjunto de palavras chave pertencentes a todas as ofertas que o utilizador, passado como parâmetro, possui na sua carteira.

### **CardDefinition**

Identicamente à entidade anterior, esta corresponde ao *template* de um cartão, contendo os requisitos gerais que um objeto necessita ter para representar um cartão, sendo a entidade Card uma instância desta classe.

- `getKeywordsListByUser(EntityManager, User)` que devolve como resultado o conjunto de palavras chave, nome do cartão, e comerciante pertencentes a todos os cartões que o utilizador, passado como parâmetro, possui na sua carteira.

### **FindOffersUtils**

- `getOffersDefinitionBySimilarity` aplica alguns filtros às ofertas disponíveis no sistema, como filtros de localização, ordenando as ofertas pela sua proximidade ao local onde o utilizador se encontra, filtros de data, dando mais relevância às ofertas criadas mais recentemente, ou ainda filtragem por género ou idade do utilizador, destacando as ofertas mais relevantes a utilizadores da mesma faixa etária ou do mesmo género que o utilizador atual. Depois de aplicados os filtros é calculada a predição (Anexo VI) de interesse do utilizador pelas ofertas filtradas. Este cálculo envolve os interesses do utilizador atual e da sua vizinhança. As ofertas são ordenadas por ordem decrescente de predição.
- `buildQueryForContent` devolve uma lista de ofertas que contêm as palavras-chave associadas ao utilizador, no seu nome, comerciante ou palavras-chave também.
- `filterByGender` ordena as ofertas colocando as ofertas mais relevantes para utilizadores do género oposto no final da lista.
- `filterByAge` organiza as ofertas colocando no início da lista as ofertas mais relevantes ao grupo de utilizadores pertencentes à mesma faixa etária que o utilizador atual.
- `getSmartTargetingOffers` é o método que junta os dois algoritmos implementados, Filtragem Colaborativa e Filtragem Baseada em Conteúdos. Primeiro verifica quais as ofertas comuns aos resultados de cada algoritmo colocando-as em primeiro lugar, e juntando as restantes ofertas alternadamente.

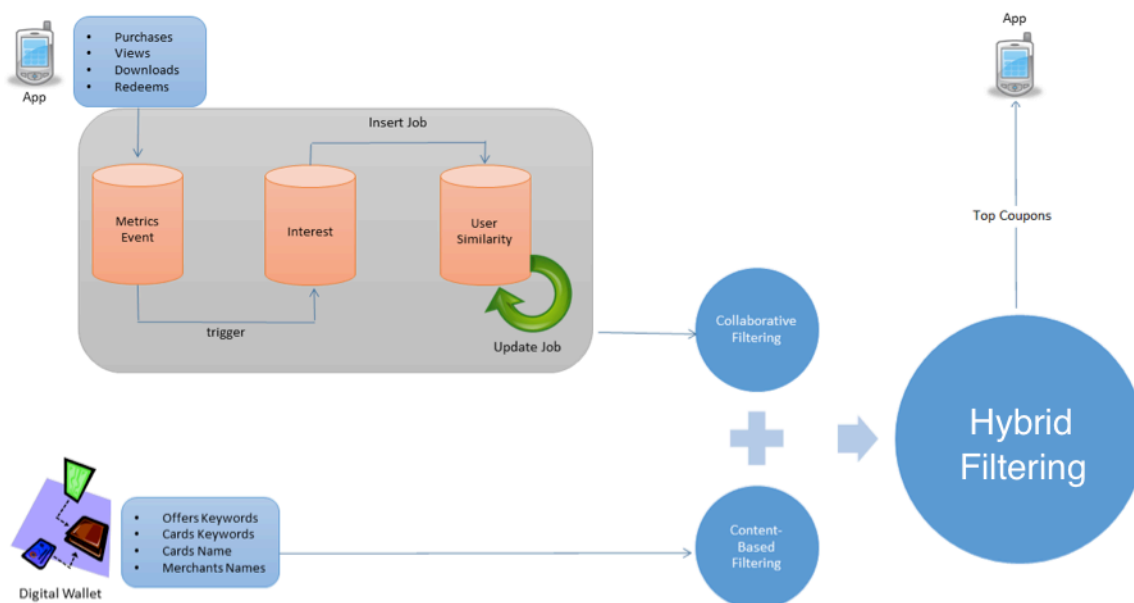
### **CouponsControllerBean**

- `topCouponsBySmartTargeting` tem como resultado uma lista de Top de Ofertas, diferentes para cada utilizador. Caso seja um novo utilizador, como este não tem histórico de utilização da aplicação, logo não tem registo dos seus interesses, é gerada uma lista de ofertas ordenada

por número de downloads, ou seja, em primeiro lugar surge a oferta ativa com maior número total de downloads no sistema e assim consecutivamente.

## 5.4 Fluxo da Implementação

A Figura 14 representa o fluxo da implementação do mecanismo de recomendação desenvolvido, ajudando a perceber melhor como este funciona. Esta foi elaborada no final do projeto, a fim de incluir num documento de especificação da empresa.



**Figura 14 - Fluxo da Implementação do mecanismo de recomendação**

Como se pode comprovar pela figura, os dados utilizados para efetuar as recomendações são colhidos da carteira digital do utilizador e do seu histórico de utilização da aplicação. A informação presente na carteira digital, tal como ofertas adquiridas e cartões de fidelidade são utilizadas diretamente no algoritmo de filtragem baseada em conteúdos, através de palavras chave. A informação do histórico de utilização da aplicação, como por exemplo ofertas visualizadas, adquiridas ou resgatadas, passa por um processo de conhecimento, cálculo de similaridade e cálculo de predição para poder efetuar recomendações viáveis. No fim as duas listas geradas são combinadas numa só, sendo esta apresentada na aplicação do dispositivo do utilizador.



## 6. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo é realizada a análise dos resultados obtidos com o desenvolvimento do mecanismo de recomendação. Esta análise de resultados baseia-se em duas fases de testes do mecanismo. Em primeiro lugar referem-se os testes efetuados durante a construção do mecanismo e em seguida referem-se os testes finais, com o modelo completo já implementado na aplicação.

### 6.1 Testes Intermédios

O mecanismo desenvolvido sofreu diversas alterações durante o seu processo de evolução até devolver recomendações satisfatórias num tempo admissível. Essas alterações destinaram-se essencialmente a melhorar o tempo de processamento do mecanismo e a evitar uma potencial sobrecarga de dados. Por exemplo, foi necessário melhorar a atualização da tabela de similaridades, que no início percorria todas as entradas existentes, e que se tornaria muito dispendioso computacionalmente quantas mais entradas fossem adicionadas à tabela.

Os testes efetuados durante o desenvolvimento do mecanismo podem ser considerados limitados, uma vez que foram efetuados fora do tronco principal da aplicação, ou seja, as ofertas e utilizadores disponíveis para testes eram restritos. Foi necessário criar ofertas e utilizadores fictícios para disponibilizar um número considerável de dados de forma a testar os algoritmos concebidos para efetuar recomendações.

Como a quantidade de dados existente na base de dados da aplicação CardMobili é consideravelmente grande, os algoritmos de filtragem necessitaram passar por algumas modificações de forma a executarem num tempo aceitável, ou seja reduzido.

A filtragem baseada em conteúdos, filtra as ofertas através de palavras-chave, e executa tão mais rápido quanto mais restrições aplicar na *query* SQL que vai procurar as ofertas à base de dados. Desta forma resultam menos ofertas sobre as quais efetuar os cálculos de ordenação de ofertas.

No caso da filtragem colaborativa, que considera a similaridade entre os utilizadores e os seus interesses, foi necessário aperfeiçoar o algoritmo de cálculo da similaridade, dada a vasta quantidade de utilizadores no sistema. Os *jobs* criados servem exatamente para diminuir o volume de informação a ser tratado de cada vez que executam, tornando o processamento mais rápido, distribuído no tempo e eficaz. Desta forma as recomendações são apresentadas ao utilizador quase em tempo real, aumentando a satisfação e fidelidade do mesmo.

## 6.2 Teste final

Não tendo havido oportunidade de escolher um grupo de utilizadores da aplicação CardMobili, o teste final foi efetuado numa demonstração ao Chefe de Desenvolvimento e Investigação da CardMobili, e a um Software Developer, que têm pleno conhecimento da aplicação e o seu funcionamento, já com o mecanismo integrado numa versão anterior da aplicação CardMobili.

O primeiro teste efetuado foi o teste do novo utilizador, que se revela ser um dos maiores problemas dos mecanismos de recomendação. Aos novos utilizadores, que ainda não têm um padrão de utilização definido nem cartões ou ofertas na sua carteira digital, não é possível efetuar recomendações baseadas nos seus interesses. Neste caso, foi mantido o método anterior, que mostra as ofertas com maior número de *downloads*. Assim, sempre que é criado um novo utilizador as ofertas recomendadas são apresentadas como se mostra na imagem patente na Figura 15.

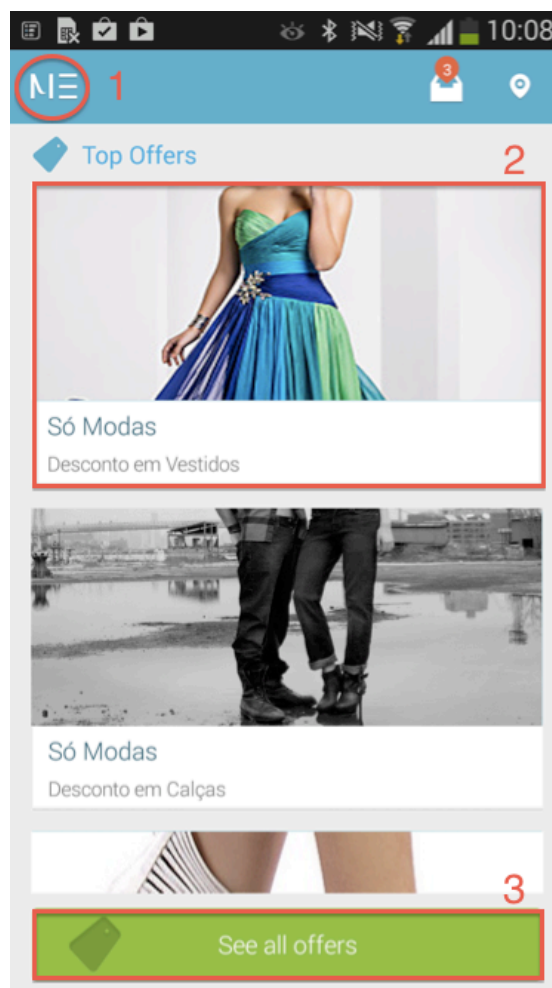


Figura 15 - Top de Ofertas para um novo utilizador

O “Top Offers” da aplicação é apresentado no menu inicial, quando o utilizador efetua o *login*, e apresenta um número pré-definido de ofertas recomendadas, ordenadas por número de *downloads*. Assim, pode concluir-se que, de todas as ofertas existentes no sistema, a oferta “Desconto em Vestidos” do comerciante “Só Modas”, assinalada com o número 2 na Figura 15, é a oferta mais descarregada pelos utilizadores. O campo assinalado com o número 1 corresponde ao botão de visualização do menu principal, que contém o acesso a conteúdo da carteira do utilizador, às definições de conta, mensagens e *logout*. O campo assinalado com o número 3 corresponde ao botão de visualização de toda a lista de ofertas recomendadas.

Foram criados seis utilizadores fictícios através da aplicação móvel, onde o perfil é definido através da utilização da aplicação, pelo que foi necessário adquirir diferentes ofertas para os diferentes utilizadores. A Tabela 6 apresenta a relação de similaridade entre os utilizadores fictícios criados.

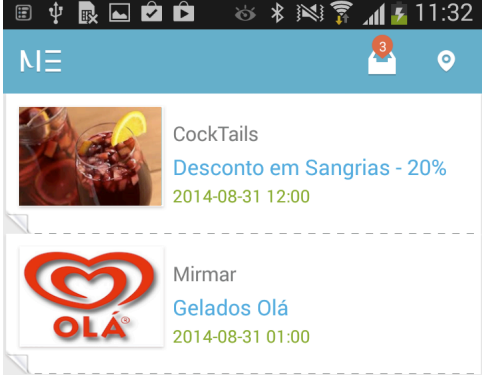

**Tabela 6 - Perfis fictícios e similaridades**

	Utilizador 1 (masc.)	Utilizador 2 (fem.)	Utilizador 3 (fem.)	Utilizador 4 (fem.)	Utilizador 5 (masc.)	Utilizador 6 (masc.)
Utilizador 1	-					
Utilizador 2		-		X		
Utilizador 3		X	-			
Utilizador 4		X		-		X
Utilizador 5					-	X
Utilizador 6				X	X	-

As tabelas seguintes apresentam as ofertas pelas quais o utilizador mostrou interesse e o respetivo Top de Ofertas, mais especificamente as 3 primeiras ofertas mais relevantes, depois da similaridade entre os utilizadores ter sido calculada. De seguida serão apresentados os resultados das recomendações para cada um dos utilizadores criados, em tabelas separadas.

A Tabela 6 corresponde aos resultados apresentados para o Utilizador 1, do sexo masculino, sem similaridade com nenhum outro utilizador.

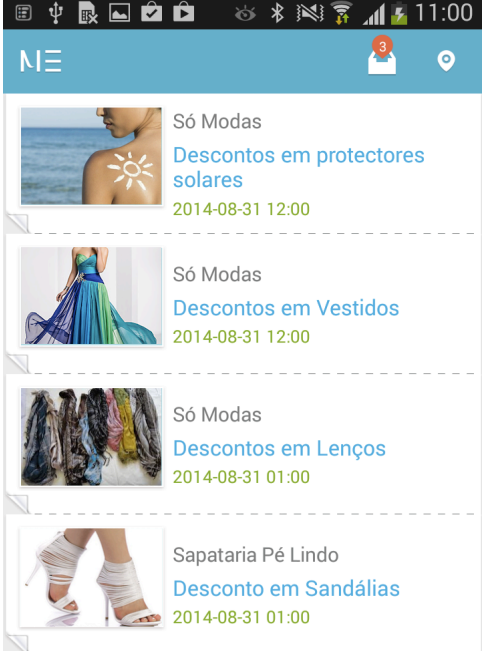
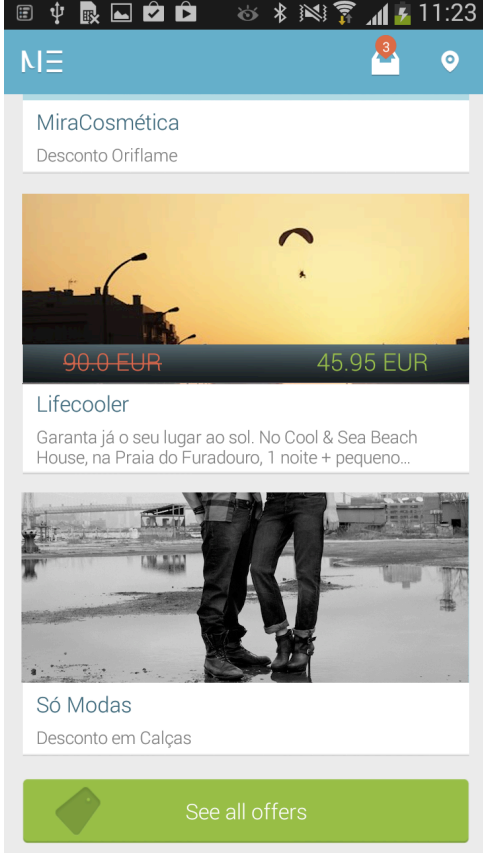
**Tabela 7 - Resultados do Utilizador 1**

Ofertas Adquiridas	Top de Ofertas
 <p>The screenshot shows a mobile app interface with a blue header. Below the header, there are two offer cards. The first card features a photo of cocktails and text: 'CockTails', 'Desconto em Sangrias - 20%', and '2014-08-31 12:00'. The second card features the 'OLA' logo and text: 'Mirmar', 'Gelados Olá', and '2014-08-31 01:00'.</p>	 <p>The screenshot shows a mobile app interface with a blue header. Below the header, there are three offer cards. The first card has a photo of a restaurant interior and text: 'SO Modas', 'Desconto em Calças', '46.7 EUR', and '24.0 EUR'. The second card has a photo of blue sneakers and text: 'Sapataria Pé Lindo', 'Desconto em Sapatilhas'. At the bottom, there is a green button with a diamond icon and the text 'See all offers'.</p>

A primeira e a terceira ofertas, “Desconto em Calças” e “Desconto em Sapatilhas”, foram recomendadas por serem as ofertas mais relevantes para utilizadores do sexo masculino. A segunda oferta, da Lifecooler, apresenta palavras-chave em comum com as ofertas do utilizador, como por exemplo “Sangria”.

A Tabela 7 apresenta os resultados das recomendações efetuadas ao Utilizador 2, do sexo feminino, que apresenta similaridade com o Utilizador 4.

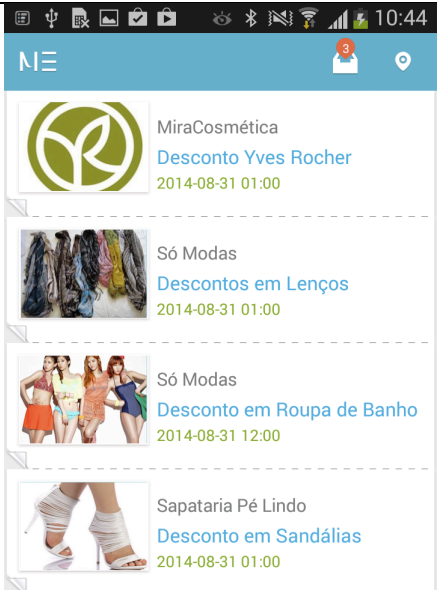
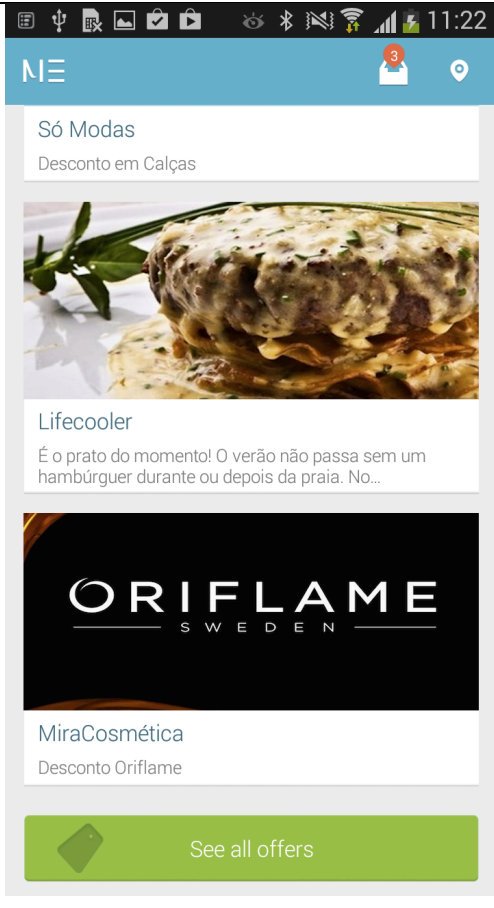
**Tabela 8 - Resultados do Utilizador 2**

Ofertas Adquiridas	Top de Ofertas
 <p>Ofertas Adquiridas</p> <ul style="list-style-type: none"> <li>Só Modas: Descontos em protectores solares (2014-08-31 12:00)</li> <li>Só Modas: Descontos em Vestidos (2014-08-31 12:00)</li> <li>Só Modas: Descontos em Lenços (2014-08-31 01:00)</li> <li>Sapataria Pé Lindo: Desconto em Sandálias (2014-08-31 01:00)</li> </ul>	 <p>Top de Ofertas</p> <ul style="list-style-type: none"> <li>MiraCosmética: Desconto Oriflame</li> <li>Lifecooler: Garanta já o seu lugar ao sol. No Cool &amp; Sea Beach House, na Praia do Furadouro, 1 noite + pequeno...</li> <li>Só Modas: Desconto em Calças</li> </ul> <p>See all offers</p>

A primeira oferta, “Desconto Oriflame” da MiraCosmetica, aparece devido à similaridade com o Utilizador 4, por ser a oferta com maior valor de interesse para este. A segunda oferta, da Lifecooler, é recomendada através das palavras-chave “Sol” e “Praia”, pertencentes à oferta “Desconto em protectores solares” da carteira do utilizador. A terceira oferta, “Desconto em Calças”, é relevante por pertencer ao mesmo comerciante da maioria das ofertas adquiridas por este utilizador, Só Modas.

Os resultados para o Utilizador 3 são apresentados na Tabela 8. Este utilizador é do sexo feminino, e apresenta similaridade com o Utilizador 2.

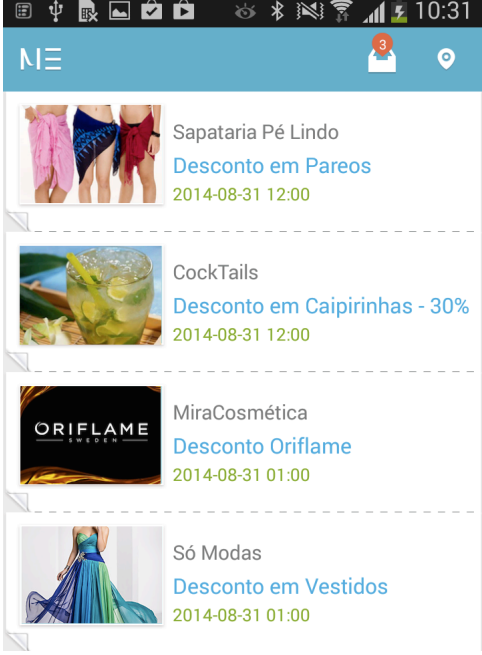
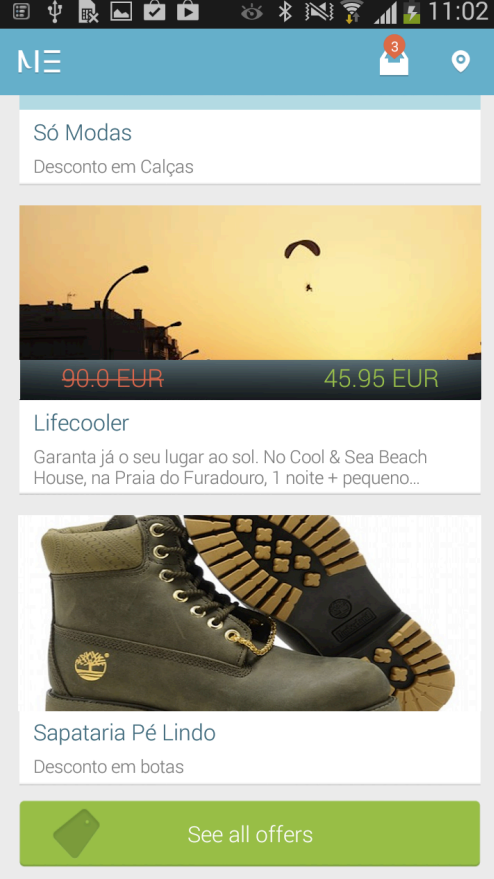
**Tabela 9 - Resultados do Utilizador 3**

Ofertas Adquiridas	Top de Ofertas
 <p>MiraCosmética Desconto Yves Rocher 2014-08-31 01:00</p> <p>Só Modas Descontos em Lenços 2014-08-31 01:00</p> <p>Só Modas Desconto em Roupas de Banho 2014-08-31 12:00</p> <p>Sapataria Pé Lindo Desconto em Sandálias 2014-08-31 01:00</p>	 <p>Só Modas Desconto em Calças</p> <p>Lifecooler É o prato do momento! O verão não passa sem um hambúrguer durante ou depois da praia. No...</p> <p>MiraCosmética Desconto Oriflame</p> <p>See all offers</p>

A primeira e terceira ofertas, “Desconto em Calças” e “Desconto Oriflame” respetivamente, são consideradas relevantes por pertencerem aos mesmos comerciantes, Só Modas e MiraCosmética por esta ordem, de algumas das ofertas adquiridas pelo utilizador. A segunda oferta é recomendada por apresentar as mesmas palavras-chave, como por exemplo “Verão” e “Praia”, que as ofertas do utilizador, como “Desconto em Roupas de Banho”.

Na Tabela 9 apresentam-se os resultados do Utilizador 4, do sexo feminino, que apresenta similaridade com o Utilizador 2 e com o Utilizador 6.

**Tabela 10 - Resultados do Utilizador 4**

Ofertas Adquiridas	Top de Ofertas
 <p><b>Ofertas Adquiridas</b></p> <p>Sapataria Pé Lindo Desconto em Pareos 2014-08-31 12:00</p> <p>CockTails Desconto em Caipirinhas - 30% 2014-08-31 12:00</p> <p>MiraCosmética Desconto Oriflame 2014-08-31 01:00</p> <p>Só Modas Desconto em Vestidos 2014-08-31 01:00</p>	 <p><b>Top de Ofertas</b></p> <p>Só Modas Desconto em Calças</p> <p>90.0-EUR 45.95 EUR</p> <p>Lifecooler Garanta já o seu lugar ao sol. No Cool &amp; Sea Beach House, na Praia do Furadouro, 1 noite + pequeno...</p> <p>Sapataria Pé Lindo Desconto em botas</p> <p>See all offers</p>

A segunda oferta, da Lifecooler, contém as mesmas palavras-chave que algumas ofertas do utilizador, tais como “Sol” e “Praia” (relativamente à oferta “Desconto em Pareos”). A primeira e terceira ofertas, “Desconto em Calças” e “Desconto em botas” respetivamente, podem ser relevantes para este utilizador por pertencerem à carteira do Utilizador 6, em relação ao qual este apresenta similaridade.

Os resultados relativos ao Utilizador 5 podem ser consultados na Tabela 10. Este utilizador apresenta similaridade com o Utilizador 6 e é do sexo masculino.

**Tabela 11 - Resultados do Utilizador 5**

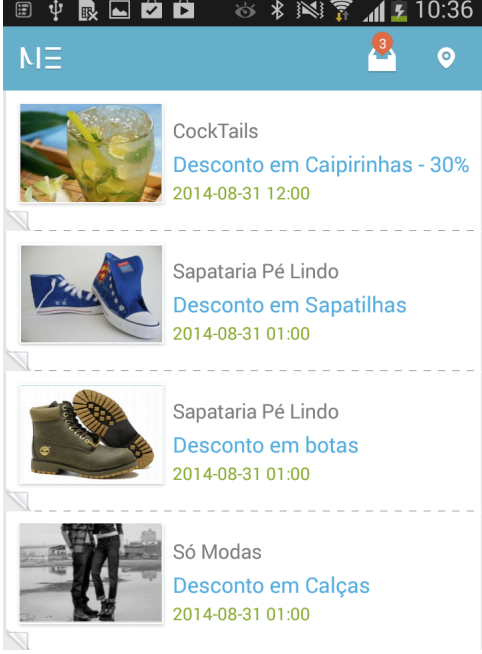
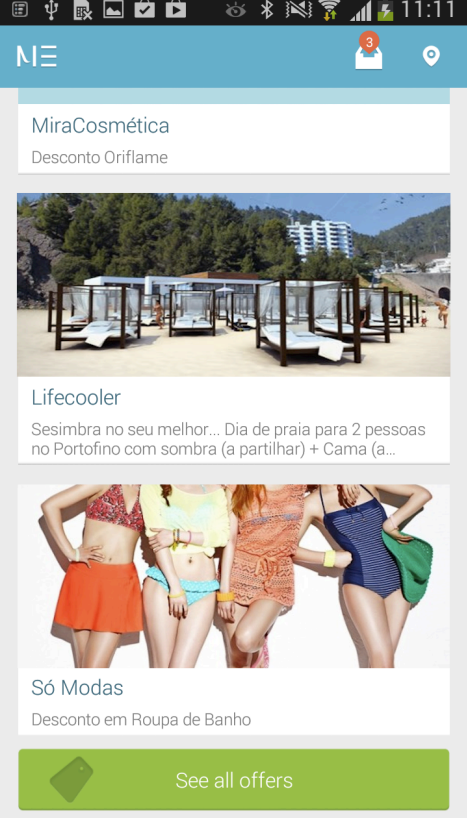
Ofertas Adquiridas	Top de Ofertas
<p>The screenshot shows a mobile app interface with a blue header containing a menu icon and a notification badge with the number '3'. Below the header, there are three offer cards separated by dashed lines. Each card features a small image, the merchant name, the offer title, and the expiration date. The offers are: 1. Sapataria Pé Lindo: Desconto em Sapatilhas, 2014-08-31 01:00. 2. Só Modas: Desconto em T-shirts, 2014-08-31 01:00. 3. Só Modas: Desconto em Calças, 2014-08-31 01:00.</p>	<p>The screenshot shows a mobile app interface with a blue header containing a menu icon and a notification badge with the number '3'. Below the header, there are three offer cards. Each card features a large image, the merchant name, and the offer title. The offers are: 1. Sapataria Pé Lindo: Desconto em botas. 2. Lifecooler: É o prato do momento! O verão não passa sem um hambúrguer durante ou depois da praia. No... 3. Mirmar: Menu do Dia. At the bottom, there is a green button with a diamond icon and the text 'See all offers'.</p>

A primeira oferta, “Desconto em botas”, é recomendada pela similaridade com o Utilizador 6, pois este mostrou interesse por esta oferta. A segunda e terceira ofertas, da Lifecooler e Mirmar respetivamente, apresentam palavras-chave, tais como “Verão”, presentes nas ofertas adquiridas por este utilizador, como por exemplo na oferta “Desconto em T-shirts”.



A Tabela 11 contém os resultados das recomendações para o Utilizador 6, do sexo masculino e com similaridade com os Utilizadores 4 e 5.

**Tabela 12 - Resultados do Utilizador 6**

Ofertas Adquiridas	Top de Ofertas
	

A primeira oferta, “Desconto Oriflame”, foi considerada relevante pela similaridade existente com o Utilizador 4, que contém esta oferta na sua carteira. A segunda oferta, da Lifecooler, apresenta palavras-chave em comum, tais como “Praia” com algumas ofertas do utilizador, por exemplo “Desconto em Caipirinhas”. A terceira oferta, “Desconto em Roupas de Banho”, pertence ao mesmo comerciante, Só Modas, que a oferta “Desconto em Calças” adquirida pelo utilizador.

Perante a empresa mostrou-se que este mecanismo funciona corretamente. Para medir o nível de serviço do mecanismo, teve-se apenas em consideração dois indicadores de medição: qualidade e tempo de resposta. Assim, o mecanismo apresentou resultados considerados relevantes para cada utilizador e um tempo de resposta quase instantâneo, de aproximadamente 2 segundos, pelo que foi aprovado e provavelmente será integrado no lançamento da próxima versão da aplicação.

Apenas foi testado o mecanismo para o Top de Ofertas, pois a interface da aplicação não estava preparada para testar as Ofertas Relacionadas. No entanto, apesar de não estar aplicado na interface, os métodos necessários para implementar essa funcionalidade estão já configurados para serem aplicados.

## 7. CONCLUSÕES

Esta dissertação reflete o trabalho de desenvolvimento de um mecanismo de recomendação de ofertas para a aplicação móvel da CardMobili, mostrando no dispositivo de cada utilizador as ofertas disponíveis potencialmente relevantes para este. Durante a sua elaboração procurou-se fazer uma explicação teórica dos fundamentos que suportam o estudo inerente ao problema, dos quais se destacam os algoritmos de recomendação mais utilizados na atualidade, por diversas empresas, seguindo-se a especificação do processo de desenvolvimento do mecanismo.

*Smart Targeting* não é mais do que uma tecnologia inteligente que direciona os conteúdos considerados interessantes para cada utilizador, cada vez mais requisitada pelas organizações de venda de produtos e prestação de serviços *online*, tais como Amazon.com, Ebay, Spotify ou Netflix. Isto permite ganhar a confiança e fidelidade dos utilizadores que procuram um serviço livre de publicidade ou conteúdos não desejados. Esta tecnologia caracteriza-se essencialmente por englobar certos domínios do conhecimento preditivo, ou seja, é capaz de descobrir padrões através do histórico de utilização dos consumidores para prever os gostos de cada um, e não só.

O mecanismo proposto pretende apresentar a cada utilizador da aplicação CardMobili a vasta gama de ofertas disponíveis que lhes possam ser atraentes. O seu desenvolvimento foi baseado em algoritmos já existentes, adaptados à aplicação própria para o efeito. Foram ainda aplicados mais alguns filtros, tais como filtros demográficos e filtros de género, à lista de ofertas de forma a tornar as recomendações mais direcionadas.

Os algoritmos relacionados com estes mecanismos são tão mais eficientes e eficazes quanto mais informação conseguirem obter sobre os utilizadores. É portanto necessário que um número considerável de utilizadores use a aplicação da CardMobili, de preferência preenchendo o seu perfil, obtendo cartões de fidelidade e adquirindo ofertas.

Neste projeto, o principal problema dos sistemas de recomendação, conhecido como *Cold Start*, referente às recomendações para novos utilizadores foi diminuído. Para os novos utilizadores, sobre os quais não há informação suficiente para criar um perfil e fazer recomendações creíveis, é apresentada a lista de ofertas com maior número de downloads.

Foram realizados diversos testes à aplicação, tanto durante como após o seu desenvolvimento. Deste modo, apesar de serem necessários alguns ajustes, o algoritmo mostrou-se eficiente, sendo o tempo de resposta bastante aceitável, e os resultados revelaram-se suficientemente satisfatórios, com recomendações diferentes e relevantes para cada utilizador e tempos de resposta reduzidos.

A solução desenvolvida oferece confiança nas recomendações geradas e portanto, o mecanismo concebido será implementado na próxima versão da aplicação.

Porém, embora considerados satisfatórios, os resultados obtidos não constituem uma perspectiva de recomendações com a máxima qualidade. Será possível melhorar a eficácia do mecanismo, tanto em termos de enriquecimento da informação processada pelos algoritmos, como em termos de apresentação dos resultados. Apresentam-se de seguida as melhorias que poderão ser aplicadas no futuro.

A interface da aplicação não permite apresentar a lista de ofertas relacionadas com uma oferta que está a ser visualizada. É portanto fundamental que se altere o design de visualização de uma oferta para que seja possível mostrar uma lista de ofertas semelhantes à oferta atual.

O algoritmo da filtragem baseada em conteúdo tem em consideração as palavras-chave associadas à oferta, o seu nome e comerciante, comparando estas palavras com as de outras ofertas, bastando haver uma palavra em comum que a oferta já é considerada como relevante. Este método poderia ser melhorado se fosse verificado o número de ocorrências de um determinado termo na descrição de duas ofertas para determinar se estas são semelhantes e se devem ser recomendadas. Palavras como “o”, “a”, “de”, “ou”, “em”, etc. não devem ser consideradas, sendo portanto necessário definir uma lista de termos de paragem, mais conhecidas como *stop words*.

Ainda no âmbito das palavras-chave, é possível melhorar o mecanismo de forma a fazer coincidir palavras semelhantes, isto é, cuja distância de Hamming<sup>11</sup> seja baixa. Por exemplo, se uma oferta tiver a palavra “Hotel” e outra tiver a palavra “Hotéis” deve haver uma relação entre as duas. Para concretizar este propósito é necessário recorrer a tecnologias de pesquisa textual que a CardMobili não possui atualmente.

O *trigger* que foi criado associado à tabela MetricsEvents está preparado para atualizar a tabela Interests caso um utilizador remova uma oferta da carteira, mostrando que perdeu interesse na oferta, ou faça o resgate de uma oferta ou a visualize apenas. No entanto, quando uma destas ações tem lugar nenhum registo é adicionado à tabela MetricsEvents logo, os interesses dos utilizadores só são incrementados quando um utilizador guarda ou faz o download de uma oferta, perdendo o algoritmo de filtragem colaborativa alguma eficácia. Propõe-se então o aperfeiçoamento da atualização da tabela MetricsEvents, de modo a que toda a atividade do utilizador com a aplicação seja registada e aproveitada para o mecanismo de recomendação.

---

<sup>11</sup> Consultar: [https://www.princeton.edu/~achaney/tmve/wiki100k/docs/Hamming\\_distance.html](https://www.princeton.edu/~achaney/tmve/wiki100k/docs/Hamming_distance.html)

Tal como foi referido anteriormente, o algoritmo tem em consideração diversas variáveis do utilizador e todas elas são igualmente consideradas. Uma possível melhoria seria aplicar ponderações ou pesos, por configuração, a cada uma delas. Assim poderíamos ter uma implementação em que se dá mais destaque à idade e ao género das pessoas para fazer recomendações e uma outra em que se dá mais destaque às palavras-chave, por exemplo.

Finalmente deve ser referido que o mecanismo desenvolvido no âmbito deste projeto não esteve sujeito à carga total de informação existente no sistema da empresa. Apesar de ter sido feito um esforço para utilizar o máximo de dados possível, a fim de testá-lo em termos de eficiência e eficácia, apenas quando este estiver implementado numa versão futura da aplicação é que se poderá obter o *feedback* dos utilizadores da aplicação. Sugere-se assim que este mecanismo seja testado por um grupo de pessoas, escolhido pela empresa, antes de o incluir no próximo lançamento.

## BIBLIOGRAFIA

- Agarwal, A., Almashan, M., & Patali, R. (2010). *Mobile Wallet Using FeliCa*. Ohio: Ohio State University.
- Alcatel-Lucent. (2010). *The mobile phone as a wallet: Appealing to the new generation of consumers*. Alcatel-Lucent.
- Alvarez, P. P., & Feichtner, D. (2013). *Riiing, riiing! Cha-ching!: The Dawn of the Mobile Wallet*. Fribourg: Université de Fribourg.
- Balabanovic, M., & Shoham, Y. (1997). Fab: Content-Based, Collaborative Recommendation. *Communications of the ACM*, 66-72.
- Ben-Gal, I. (2007). Bayesian Networks. In *Encyclopedia of Statistics in Quality and Reliability*. Wiley & Sons.
- Bergner, Y., Dröschler, S., Kortemeyer, G., Rayyan, S., Seaton, D., & Pritchard, D. E. (2012). *Model-Based Collaborative Filtering Analysis of Student Response Data: Machine-Learning Item Response Theory*. Cambridge: Massachusetts Institute of Technology.
- Bezerra, B. L. (2004). *Uma Solução Em Filtragem De Informação Para Sistemas De Recomendação Baseada Em Análise De Dados Simbólicos*. Recife: Universidade Federal de Pernambuco.
- Boullé, M., Meyer, F., & Candillier, L. (2007). *Comparing State-of-the-Art Collaborative Filtering Systems*. Lannion, France: France Telecom R&D.
- Burin, P. (2007). *Sistema de Recomendação Baseado na Arquitetura Multiagente*. Rio Grande do Sul: Universidade Luterana do Brasil.
- Burke, R. (2002). *Hybrid Recommender Systems*. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.

Cardoso, C. M., Reimao, S., & Marques, J. (2011). Privacidade e ética nos Sistemas de Recomendação para e-commerce: uma breve reflexão. *Amazônia e o direito de comunicar*. Belém: Mídia Cidadã.

Cazella, S. C. (2006). *Aplicando a relevância da opinião de usuários em sistema de recomendação para pesquisadores*. Porto Alegre: Universidade Federal do Rio Grande do Sul.

Cazella, S. C. (2 de Abril de 2008). Sistemas de Recomendação Para E-Commerce. *Mini-Course: E-Commerce and Recommender Systems*. Florianópolis, Brasil: ERDB.

Cazella, S. C., Nunes, M. S., & Reategui, E. B. (2005). A Ciência da Opinião: Estado da Arte em Sistemas de Recomendação. *XXV Congresso da Sociedade Brasileira da Computação*. São Leopoldo: .

Cole, A., McFaddin, S., Narayanaswami, C., & Tiwari, A. (2009). *Toward a Mobile Digital Wallet*. New York: IBM Research Division.

Crosby, L., & Johnson, S. (Novembro de 2001). Technology: Friend or Foe to Customer Relationships? *Marketing Management*, p. 10.

Cuffie, M., Hydari, M. Z., Yu, C., & Zitzmann, M. (s.d.). *Recommendation Systems and Targeted Advertising for Online Video*. Obtido em 2014 de Março de 14, de <http://www.andrew.cmu.edu/user/mhydari/ov/paper.html>

Davidsson, C. (2010). *Mobile Application Recommender System*. Uppsala: Uppsala University.

Dyché, J. (2001). *The CRM Handbook: A Business Guide to Customer Relationship Management*. USA: Addison-Wesley.

EDRi-gram. (8 de Abril de 2009). Behavioral Targeting. 7 (7).

Ekstrand, M., Riedl, J. T., & Konstan, J. A. (2011). Collaborative Filtering Recommender Systems. *Now the essence of knowledge*, 4, 81-173.

Forum, M. (2011). *Mobile Wallet - Definition and Vision Part 1*. Mobey Forum.

Francisco, A. C., Kovalski, J. L., & Callegari, N. M. (2006). CRM – Uma ferramenta tecnológica inovadora. *XIII SIMPEP*. Bauru, SP, Brasil.

Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). *Using Collaborative Filtering to weave an information Tapestry*. : Gale Group.

Greenberg, P. (2004). *CRM at the Speed of Light, Third Edition: Essential Customer Strategies for the 21st Century*. New York, NY, USA: The McGraw-Hill Companies.

Greenberg, P. (2004). *CRM at the Speed of Light, Third Edition: Essential Customer Strategies for the 21st Century*. USA: The McGraw-Hill Companies.

Howe, M. (2007). *Pandora's Music Recommender*. Washington: University of Washington.

Júnior, R. D. (2004). *Combining Collaborative and Content-Based Filtering to Recommend Research Papers*. Universidade Federal do Rio Grande do Sul, Instituto de Informática. Porto Alegre: .

Jefferson, T. (25 de Maio de 2010). Making Mobile Wallets a Success. *Technology In-Depth* , pp. 16-21.

Khoshgoftaar, T. M., & Su, X. (2009). A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence* , 19.

Lichtnow, D., Garin, R. S., Palazzo, L. A., Loh, S., Kampff, A. J., Primo, T., et al. (2006). O Uso de Técnicas de Recomendação em um Sistema para Apoio à Aprendizagem Colaborativa. *Revista Brasileira de Informática na Educação* , 14, 49-59.

Linden, G., Smith, B., & York, J. (2003). *Amazon.com Recommendations: Item-to-Item Collaborative Filtering*. IEEE Computer Society.



Liyakasa, K. (Abril de 2013). *Adoption Problems Plague Digital Wallets*. Obtido em 2014, de Destination CRM: <http://www.destinationcrm.com/Articles/Columns-Departments/Insight/Adoption-Problems-Plague-Digital-Wallets-88274.aspx>

Melville, P., & Sindhvani, V. (nd). *Recommender Systems*. New York: IBM T.J. Watson Research Center.

Melville, P., & Sindhvani, V. (sd). *Recommender Systems*. New York: IBM T.J. Watson Research Center.

Meteren, R. v., & Someren, M. v. (nd). *Using Content-Based Filtering for Recommendation*. Amesterdam: .

Oliveira, A. S., & Coello, J. M. (2013). *Desenvolvimento de Algoritmo Híbrido para Sistemas de Recomendação: Filtragem Colaborativa e Etiquetagem Social*. Campinas.

Pareek, D. *The value of the mobile wallet*.

Pareek, D. (nd). *The value of the mobile wallet*. Obtido em 2014, de Geocities: [http://www.geocities.ws/fina\\_tech/ppt/vmw.pdf](http://www.geocities.ws/fina_tech/ppt/vmw.pdf)

Pazzani, M. J., & Billsus, D. (2007). Content-based Recommendation Systems. In P. K. Brusilovski, *The Adaptive Web: Methods and Strategies of Web Personalization* : Springer.

Peppers, D., & Rogers, M. (2011). *Managing Customer Relationships: A Strategic Framework*. New Jersey: Johyn Wiley & Sons, Inc.

Rafsanjani, A. H., Salim, N., Aghdam, A. R., & Fard, K. B. (2013). Recommendation Systems: a review. *International Journal of Computational Engineering Research* | , 47-52.

Resnick, P., & Varian, H. R. (1997). *Recommender systems*. Communications of the ACM.

Reynolds, J. (2002). *A Practical Guide to CRM*. New York: CMP Books.

Ricci, F. (2010). *Mobile Recommender Systems*. Free University of Bolzano.

Riedl, J., Frankowski, D., & Lam, S. K. (2006). *Do You Trust Your Recommendations? An Exploration Of Security and Privacy Issues in Recommender Systems*. Minneapolis: University of Minnesota.

Riedl, J., Konstan, J. A., & Herlocker, J. L. (2000). *Explaining Collaborative Filtering Recommendations*. University of Minnesota, Dept. of Computer Science and Engineering. Minneapolis, USA: GroupLens Research.

Rodrigues, T. (2010). *Protótipo de um Sistema de Recomendação de Presentes Baseado no Perfil do Usuário*. Chapecó: Universidade Comunitária da Região de Chapecó.

Rohrmann, C. A. (2000). *Notas Acerca do Direito à Privacidade na Internet: A Perspectiva Comparativa*. : Instituto Online Para o Direito e Informática.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). *Analysis of Recommendation Algorithms for E-commerce*. Minneapolis: GroupLens Research Group / Army HPC Research Center.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). *Item-based Collaborative Filtering Recommendation Algorithms*. Minneapolis: GroupLens Research Group/Army HPC Research Center .

Semeraro, G., Gemmis, M. d., & Lops, P. (2011). Content-based Recommender Systems: State of the Art and Trends. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor, *Recommender Systems Handbook* (pp. 73-105). : Springer.

Sotsenko, A., Jansen, M., & Milrad, M. (sd). *Using a Rich Context Model for a News Recommender System for Mobile Users*. Linnaeus University, Växjö.

Stephens, D. (2011). The Problem With Mobile Wallet No One is Talking About. *Retail Prophet* .

Torres, R., McNee, S. M., Abel, M., Konstan, J. A., & Riedl, J. (2004). Enhancing Digital Libraries with TechLens. *JCDL '04:Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries* , 228-236.

Ullman, J. D. (2011). Recommendation Systems. In J. Ullman, & A. Rajaraman, *Mining of Massive Datasets* (pp. 307-341). Cambridge: Cambridge University Press.

Vieira, A. (2008). *Otimização de uma plataforma CRM*. Lisboa: Faculdade de Ciências da Universidade de Lisboa.

Wartena, C., Slakhorst, W., Wibbels, M., Gantner, Z., Freudenthaler, C., Newell, C., et al. (2011). *Keyword-Based TV Program Recommendation*. Barcelona, Spain: .

Wise, L. (2013). Mobile wallet technology raises privacy, security concerns. *McClatchy Newspapers* .

Xue, G.-R., Lin, C., Yang, Q., Xi, W., Zeng, H.-J., Yu, Y., et al. (2005). Scalable Collaborative Filtering Using Cluster-based Smoothing. *ACM SIGIR Conference* (pp. 114-121). Salvador, Brazil: .

## ANEXO I – MODELO DE NEGÓCIO DA CARDMOBILI

### *Value Propositions*

A plataforma funciona como carteira digital, onde se podem guardar todos os cartões que se guarda na carteira física, desde identificação, carta de condução, cartão de contribuinte, até cartões de fidelização, cupões, descontos. A carteira fica mais leve, sem cartões nem papéis e não se corre o risco de se perder cartões ou de se esquecer deles nas lojas, pois é muito difícil uma pessoa esquecer-se do telemóvel.

Também se pode guardar um cartão de crédito digital (@bes) para efetuar compras *online*, com o máximo de segurança, protegido com um código PIN, como aconteceria numa situação normal.

É possível criar os próprios cartões de fidelização, caso estes ainda não estejam disponíveis na base de dados da aplicação. Para as empresas esta plataforma oferece mais aderência por parte do consumidor final, pois não sendo obrigado a transportar um cartão físico, torna-se mais apelativa a fidelização a essa marca e o uso do cartão de adesão, fortalecendo portanto a relação do cliente com a marca.

A plataforma é totalmente grátis e está disponível para todos os sistemas operativos, iOS, Android, Windows Mobile e Blackberry. É ainda possível fazer o *download* da aplicação em várias línguas, como português, inglês, espanhol, francês e holandês, com cartões desses países disponíveis.

A plataforma disponibiliza ainda uma quantidade significativa de ofertas, fazendo o *targeting* por localização e mostrando as lojas da marca que se pretende.

O objetivo não é criar uma carteira digital da CardMobili, mas sim carteiras digitais de empresas com a tecnologia da CardMobili.

### *Key Partners*

Os parceiros chave são a Portugal Telecom, que complementa a plataforma de cartões de fidelização e cupões com a área de pagamentos móveis, a AdvanceCare, que disponibiliza a todos os clientes dos seus parceiros o acesso a toda a sua informação apenas utilizando a plataforma da CardMobili, o Banco Espírito Santo, que criou um cartão de crédito totalmente digital para ser utilizado

exclusivamente em transações à distância e que pode ser utilizado e consultado através da plataforma da CardMobili, e a EDP, que escolheu esta aplicação para promover o seu cartão de cliente digital, que oferece aos clientes acesso a uma rede de parceiros usufruindo de descontos em atividades comerciais e lazer.

Os fornecedores, neste caso investidores, que contribuem para o desenvolvimento da empresa são a Panthena, Capital Criativo e o empresário Carlos Oliveira.

Para o negócio funcionar é fundamental ter empresas aderentes, isto é, têm acordos comerciais com a empresa e garantem a aceitação total deste serviço nas suas lojas, usam o serviço para comunicar com os seus clientes, e os seus clientes atualizam os seus dados pessoais através da plataforma CardMobili, podendo ainda consultar os seus pontos, descontos, promoções, etc.

### *Key Activities*

É importante ter conhecimento dos cartões de fidelização existentes para facilitar a sua criação por partes dos consumidores finais. Depois é essencial que as marcas desses cartões tenham conhecimento da plataforma da CardMobili e a aceitem de forma a o utilizador não ser obrigado a andar com os cartões na carteira.

Os cartões são identificados por número de cartão ou por um código de barras, que é depois mostrado de forma a poder usá-lo.

### *Key Resources*

Os principais recursos incluem habilidades de computação web e móvel, pesquisa de tecnologia móvel e inovadora, experiência com tecnologia de retalho, área de marketing, e gestão financeira.

É necessário um vasto equipamento (computadores, servidores, etc.) para o desenvolvimento e manutenção da plataforma.

### *Cost Structure*

Custos de Investigação e Desenvolvimento:

- Equipamento
- Equipa de Programadores
- Equipa de testes
- *Marketing*
- Investigação

Custos de Infraestruturas

### *Customer Relationships*

Suporte ao Cliente

Cocriação

### *Customer Segments*

Mercado em massa, pois é para qualquer utilizador que esteja interessado em ter os seus cartões todos digitalizados.

### *Channels*

Toda a comunicação entre a empresa e o cliente, compra, entrega, apoio ao cliente é feita através de um único canal: web.

### *Revenue Streams*

A CardMobili contacta os comerciantes apresentando a aplicação e as vantagens que ela oferece à empresa, caso esta se ligue a ela, propondo substituir SMS, papel, plástico e toda uma estrutura para estar a cuidar disso por uma ligação digital a uma plataforma operada pela CardMobili. Quanto maiores forem os volumes de dados a colocar na plataforma, menor custo terá o serviço.



## ANEXO II – CÁLCULO DA SIMILARIDADE

```
public static double mean(List<Interest> aux1) {
    double m = 0.0;
    for (Interest i : aux1) {
        double d = i.getValue();
        m += d;
    }
    m /= aux1.size();
    return m;
}

public double calculateSimilarity(EntityManager entityManager, User
user1, User user2) {
    VerifyArgs.isNotNull(entityManager, "entityManager");
    VerifyArgs.isNotNull(user1, "user1");
    VerifyArgs.isNotNull(user2, "user2");

    List<Interest> aux1 = Interest.findByUser(entityManager, user1);
    List<Interest> aux2 = Interest.findByUser(entityManager, user2);

    double meanX = mean(aux1);
    double meanY = mean(aux2);

    double sumXY = 0.0, sumX2 = 0.0, sumY2 = 0.0;
    List<OfferDefinition> offersUser1 = new ArrayList<OfferDefinition>();
    List<OfferDefinition> offersUser2 = new ArrayList<OfferDefinition>();
    for (Interest i1 : aux1)
        offersUser1.add(i1.getOfferDefinition());
    for (Interest i2 : aux2)
        offersUser2.add(i2.getOfferDefinition());
    for (OfferDefinition o : offersUser1) {
        if (offersUser2.contains(o)) {
            sumXY += 0.5 +
            ((Interest.findByOfferDefinitionAndUser(entityManager, o, user1).getValue()
            - meanX) * (Interest.findByOfferDefinitionAndUser(entityManager, o,
            user2).getValue() - meanY));
        }
    }
}
```



```

                sumX2                +=                0.5                +
Math.pow(Interest.findByOfferDefinitionAndUser(entityManager,                0,
user1).getValue() - meanX, 2.0);
                sumY2                +=                0.5                +
Math.pow(Interest.findByOfferDefinitionAndUser(entityManager,                0,
user2).getValue() - meanY, 2.0);
        }
    }
    this.similarValue = (sumXY / (Math.sqrt(sumX2) * Math.sqrt(sumY2)));
    if (Double.isNaN(this.similarValue))
        this.similarValue = 0.0;
    return this.similarValue;
}

```

## ANEXO III – TRIGGER

A cada evento corresponde um valor:

Adicionar uma oferta	=	13
Resgate de uma oferta	=	14
Partilhar uma oferta	=	16
Visualizar uma oferta	=	15
Comprar uma oferta	=	17

```
CREATE TRIGGER actionOfferInterests AFTER INSERT ON MetricsEvent FOR EACH
ROW
BEGIN
    declare counter int;
    declare oldvalue int;
    SELECT count(*) into counter from Interest i where
i.offerDefinition_id = new.offerDefinitionId AND i.user_id = new.userId;
    select i.interestValue into oldvalue from Interest i where
i.offerDefinition_id=new.offerDefinitionId AND i.user_id = new.userId;
    IF (new.eventType = 14) THEN
        IF (counter > 0) THEN
            UPDATE Interest SET interestValue = oldvalue + 5 WHERE
offerDefinition_id = new.offerDefinitionId AND user_id = new.userId;
        ELSE INSERT into Interest (interestValue, offerDefinition_id,
user_id) VALUES (5, new.offerDefinitionId, new.userId);
        end if;
    end if;
    IF (new.eventType = 16) THEN
        IF (counter > 0) THEN
            UPDATE Interest SET interestValue = oldvalue + 2 WHERE
offerDefinition_id = new.offerDefinitionId AND user_id = new.userId;
        ELSE INSERT into Interest (interestValue, offerDefinition_id,
user_id) VALUES (2, new.offerDefinitionId, new.userId);
```

```

        end if;
    end if;
    IF (new.eventType = 17 OR new.eventType = 13) THEN
        IF (counter > 0) THEN
            UPDATE Interest SET interestValue = oldvalue + 4 WHERE
offerDefinition_id = new.offerDefinitionId AND user_id = new.userId;
            ELSE INSERT into Interest (interestValue, offerDefinition_id,
user_id) VALUES (4, new.offerDefinitionId, new.userId);
            end if;
        end if;
        IF (new.eventType = 15) THEN
            IF (counter > 0) THEN
                UPDATE Interest SET interestValue = oldvalue + 1 WHERE
offerDefinition_id = new.offerDefinitionId AND user_id = new.userId;
                ELSE INSERT into Interest (interestValue, offerDefinition_id,
user_id) VALUES (1, new.offerDefinitionId, new.userId);
            end if;
        end if;
    END;
    $$
DELIMITER ;

```

## ANEXO IV – JOB DE INSERÇÃO

```
public String taskMethod() {

    UserSimilarity userSimilarity = new UserSimilarity();
    InsertUserSimilarity insertUserSimilarity = (InsertUserSimilarity)
Component.getInstance(InsertUserSimilarity.class, true);

    double similar = 0.0;
    Date currentDate = Calendar.getInstance().getTime();
    int limit = 5;
    try {
        Interest lastInterest =
Interest.getLastUserForSimilar(entityManager);
        User user1 = lastInterest.getUser();
        Long lastId = lastInterest.getLastId();
        Long interestId = lastInterest.getId();
        if (!Interest.alreadyInserted(entityManager, user1)) {
            List<User> users =
Interest.getUsersToInsert(entityManager, limit, lastId);
            int size = users.size();
            if (size != 0) {
                for (User user2 : users) {
                    if (user1.getId() != user2.getId()) {
                        similar =
userSimilarity.calculateSimilarity(this.entityManager, user1, user2);
                        if
(!userSimilarity.alreadyExistsSimilarity(this.entityManager, user1, user2))
{

                            userSimilarity.saveSimilarityValue(this.entityManager, user1, user2,
similar, currentDate);

                                if (size < limit || (lastId +
size) >= lastInterest.getId()) {

                                    Interest.updateInterestLastId(entityManager, user1, lastId + size,
```

```

interestId, true);
                                } else {

        Interest.updateInterestLastId(entityManager, user1, lastId + size,
interestId, false);
                                }
                                }
                                } else {

        Interest.updateInterestLastId(entityManager, user1, lastId + size,
interestId, false);
                                }
                                }
                                } else {
                                Interest.updateInterestLastId(entityManager, user1,
lastId + size, interestId, true);
                                }
                                } else {
                                Interest.updateInterestLastId(entityManager, user1,
interestId, interestId, true);
                                }
                                }
                                catch (Exception ex) {
                                this.log.warn("Error calculating InsertUserSimilarity: " +
ex.getMessage(), ex);
                                }
                                // }
                                return " finished with success";
                                }
}

```

## ANEXO V – JOB DE ATUALIZAÇÃO

```
public String taskMethod() {

    UserSimilarity userSimilarity = new UserSimilarity();
    UpdateUserSimilarity updateUserSimilarity = (UpdateUserSimilarity)
Component.getInstance(UpdateUserSimilarity.class, true);

    double similar = 0.0;
    Date currentDate = Calendar.getInstance().getTime();
    try {
        List<UserSimilarity> aux =
UserSimilarity.getSimilarityToUpdate(this.entityManager, 10);
        for (UserSimilarity us : aux) {
            User user1 = us.getUser1();
            User user2 = us.getUser2();
            similar = us.calculateSimilarity(this.entityManager,
user1, user2);
            us.updateSimilarityValue(this.entityManager, user1,
user2, similar, currentDate);
        }
    }
    catch (Exception ex) {
        this.log.warn("Error calculating UpdateUserSimilarity: " +
ex.getMessage(), ex);
    }
    // }

    return " finished with success";
}

}
```



## ANEXO VI – CÁLCULO DE PREDIÇÃO

```
public static int calculatePrediction(EntityManager entityManager, User
user, OfferDefinition offerDefinition) {
    double pred = 0.0;
    User auxUser = null;
    double meanX = mean(Interest.findByUser(entityManager, user));
    double meanN = 0.0;
    double a = 0.0, b = 0.0;
    int interestUserOffer;

    ArrayList<User> usersList = new ArrayList<User>();
    List<User> list1 = findNeighborByUser(entityManager, user);
    List<User> list2 = findUsersByInterestOfferDefinition(entityManager,
user, offerDefinition);
    Set<User> allLists = mergeLists(list1, list2);
    Iterator<User> it = allLists.iterator();
    while (it.hasNext()) {
        auxUser = it.next();
        double sim = 0.0;
        meanN = mean(Interest.findByUser(entityManager, auxUser));
        Interest interest = new Interest();
        interestUserOffer =
interest.findByOfferDefinitionAndUser(entityManager, offerDefinition,
auxUser).getValue();
        if (interestUserOffer != 0) {
            sim = UserSimilarity.findByUsers(entityManager, user,
auxUser).getSimilarValue();
            a += 0.2 + sim * (interestUserOffer - meanN);
            b += Math.abs(sim);
        }
    }
    if (b != 0) { pred = (a / b) + meanX; }
    else { pred = 0.0; }
    return (int) pred;
}
```