



UNIVERSIDADE ESTADUAL DO CEARÁ

MÁRIO ALVES DE MORAES NETO

ESTRATÉGIA HÍBRIDA PARA RECOMENDAÇÃO
PERSONALIZADA UTILIZANDO O GUIA DE PROGRAMAÇÃO
ELETRÔNICO

FORTALEZA - CEARÁ
2011

MÁRIO ALVES DE MORAES NETO

ESTRATÉGIA HÍBRIDA PARA RECOMENDAÇÃO PERSONALIZADA
UTILIZANDO O GUIA DE PROGRAMAÇÃO ELETRÔNICO

Dissertação submetida à comissão examinadora do Mestrado Acadêmico em Ciência da Computação da Universidade Estadual do Ceará – UECE, como requisito parcial para obtenção de título de Mestre em Ciência da Computação.

Orientadora: Prof^a. Dr^a. Mariela Inés Cortés

Co-Orientador: Prof. Dr. Cidcley Teixeira de Souza

FORTALEZA - CEARÁ
2011

M827e Moraes Neto, Mário A.

Estratégia Híbrida para recomendação personalizada
utilizando o Guia de Programação Eletrônico/ Mário Alves de Moraes
Neto. __ Fortaleza, 2011.

106 p. ; il.

Orientadora: Profª. Dra. Mariela Inés Cortés

Dissertação (Mestrado Acadêmico em Ciência da
Computação) – Universidade Estadual do Ceará. Centro de Ciência e
Tecnologia.

1. Sistemas de Recomendação. 2. Estratégia Híbrida. 3.
Regras de Associação. 4. Televisão digital. 5. EPG. I. Universidade
Estadual do Ceará. Centro de Ciência e Tecnologia.

CDD: 001.6

MÁRIO ALVES DE MORAES NETO

ESTRATÉGIA HÍBRIDA PARA RECOMENDAÇÃO PERSONALIZADA
UTILIZANDO O GUIA DE PROGRAMAÇÃO ELETRÔNICO

Dissertação submetida à comissão examinadora do Mestrado Acadêmico em Ciência da Computação da Universidade Estadual do Ceará – UECE, como requisito parcial para obtenção de título de Mestre em Ciência da Computação. Aprovada pela comissão examinadora abaixo assinada.

Aprovada em 31/03/2011

COMISSÃO EXAMINADORA

Profa. Dra. Mariela Inés Cortés – UECE

Prof. Dr. Cidclei Teixeira de Souza – IFCE

Prof. Dr. Jerffeson Teixeira de Souza – UECE

Profa. Dra. Maria Elizabeth Sucupira Furtado – UNIFOR

RESUMO

A vasta gama de informações digitais tem feito com que encontrar aquelas que mais se encaixam no perfil de quem as procura se torne algo cada vez mais difícil. Neste contexto, os Sistemas de Recomendação constituem-se numa solução para o problema acima, uma vez que ajudam o usuário na procura de informações relevantes, atuando baseados em personalização da informação, uma vez que auxiliam determinados indivíduos no processo social de indicar ou receber indicação de itens pertencentes a um domínio específico. De maneira similar, a digitalização dos sinais de TV permite que uma vasta gama de canais possa ser transmitida, fazendo com que os usuários se deparem com uma oferta de programação televisiva bastante volumosa. Ao passo que isso representa um certo benefício, daí extraímos um problema: usuários telespectadores diante de uma oferta tão expressiva que os deixa sem saber quais programas melhor se adéquam ao seu perfil.

Assim sendo, fazendo uso dos benefícios dos Sistemas de Recomendação, este trabalho propõe a utilização das técnicas de Filtragem Baseada em Conteúdo (FBC) e as Regras de Associação para a recomendação de programas no contexto da televisão digital interativa. O sistema aqui apresentado, que permite a criação de diversos perfis e utiliza o formato de um Guia de Programação Eletrônico (EPG), também permite que o usuário avalie a recomendação realizada, tornando possível o aprimoramento constante e evolução do sistema de recomendação em questão.

Palavras-chave: sistemas de recomendação, estratégia híbrida, regras de associação, televisão digital, EPG.

ABSTRACT

The wide range of digital information has become more difficult to find things that best fit the profile of those who demand something. In this context, Recommender Systems represent a solution to the above problem, since they help the user to find relevant information, operating based on information personalization, since it helps certain individuals in the social process of appointing and receiving an indication of items belonging to a specific domain. Similarly, the digitization of TV signals allows a wide range of channels can be transmitted, resulting the users face a range of television programming so extensive. While this represents a benefit, hence draw a problem: users with an expressive range of programs without knowing what those best fit on your respective profile.

Therefore, making use of the benefits of Recommender Systems, this paper proposes the use of techniques for Content-Based Filtering (CBF) and the Association Rules to recommend programs in the context of interactive digital television. The system presented here, which allows create multiple profiles and use the format of an Electronic Program Guide (EPG), also allows the user to evaluate the recommendation process, making possible the continuous improvement of this recommendation system.

Keywords: recommender systems, hybrid strategy, association rules, digital TV, EPG.

LISTA DE TABELAS

TABELA 1: QUADRO CONTENDO AS PRINCIPAIS CARACTERÍSTICAS DOS TRABALHOS RELACIONADOS.	20
TABELA 2: QUADRO CONTENDO INFORMAÇÕES SOBRE A PLATAFORMA DE DESENVOLVIMENTO.	58
TABELA 3: PERCENTUAL DE ACERTO DO SISTEMA DE ACORDO COM A ESTRATÉGIA UTILIZADA.	73

LISTA DE FIGURAS

FIGURA 1: ARQUITETURA DE UM SISTEMA DE RECOMENDAÇÃO.	24
FIGURA 2: MÉTODO PARA EXTRAÇÃO FINAL DAS REGRAS DE ASSOCIAÇÃO – ALGORITMO APRIORI.....	31
FIGURA 3: TOPOLOGIA GERAL DE UM SISTEMA DE TELEVISÃO DIGITAL.	33
FIGURA 4: CICLO DE VIDA DE UM XLET.	37
FIGURA 5: ARQUITETURA DO ISDB-Tb.	39
FIGURA 6: ARQUITETURA DO GINGA.	40
FIGURA 7: ESTRUTURA DO PERFIL EXPLÍCITO DE USUÁRIOS EM XML.	42
FIGURA 8: ARQUITETURA GERAL DA APLICAÇÃO.	47
FIGURA 9: DIAGRAMA DE CLASSES DA APLICAÇÃO.	50
FIGURA 10: DIAGRAMA DE SEQUÊNCIA – COLETA EXPLÍCITA DE DADOS.	53
FIGURA 11: DIAGRAMA DE SEQUÊNCIA – COLETA IMPLÍCITA DE DADOS.....	54
FIGURA 12: DIAGRAMA DE SEQUÊNCIA - GERAÇÃO DO MODELO DO TELESPECTADOR.....	54
FIGURA 13: DIAGRAMA DE SEQUÊNCIA - GERAÇÃO DA RECOMENDAÇÃO DE PROGRAMAÇÃO...55	
FIGURA 14: DIAGRAMA DE SEQUÊNCIA – EXIBIÇÃO DA PROGRAMAÇÃO RECOMENDADA.....	56
FIGURA 15: DIAGRAMA DE SEQUÊNCIA – AVALIAÇÃO DA RECOMENDAÇÃO.	57
FIGURA 16: ESTRUTURA DO ARQUIVO EPG.XML.	59
FIGURA 17: ESTRUTURA DO ARQUIVO DE AVALIAÇÃO DA RECOMENDAÇÃO.....	60
FIGURA 18: DOCUMENTO CONTEUDO PROGRAMAÇÃO ASSISTIDA PELO USUÁRIO.....	61
FIGURA 19: TRECHO DE CÓDIGO – RESTRIÇÕES DE PROGRAMAÇÃO EM XML.	62
FIGURA 20: TRECHO DE CÓDIGO – RESTRIÇÕES DE PROGRAMAÇÃO EM JAVA.....	62
FIGURA 21: GERAÇÃO E OBTENÇÃO DAS REGRAS.	64
FIGURA 22: TELA DE EXIBIÇÃO DE DETALHES DE UM PROGRAMA.....	67
FIGURA 23: TELA DE USUÁRIOS CADASTRADOS.....	68
FIGURA 24: TELA PARA CADASTRO DE PERFIS (INFORMAÇÕES GERAIS).....	69
FIGURA 25: TELA PARA CADASTRO DE PERFIS (INFORMAÇÕES SOBRE FILMES).	69
FIGURA 26: EPG EXIBINDO OS PROGRAMAS RECOMENDADOS.	70
FIGURA 27: FUNCIONALIDADE DE AVALIAÇÃO DO SISTEMA DE RECOMENDAÇÃO.	71
FIGURA 28: AVALIAÇÃO DA RECOMENDAÇÃO - RESTRIÇÕES.	72
FIGURA 29: EVOLUÇÃO DO PERCENTUAL DE ACERTO DO SISTEMA DE ACORDO COM A ESTRATÉGIA UTILIZADA.....	74
FIGURA 30: IDENTIFICAÇÃO DAS PORCENTAGENS DE EVOLUÇÃO DO PROCESSO AVALIATIVO CONFORME APRIMORAMENTO DA ESTRATÉGIA DE RECOMENDAÇÃO.....	75

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
API	Application Programming Interface
ARIB	Association of Radio Industries and Bussinesses
ATSC	Advanced Television Systems Commitee
DSM-CC	Digital Storage Media – Command and Control
DTV	Digital Television
DVB	Digital Video Broadcasting
EIT	Event Information Table
EPG	Electronic Programming Guide
FBC	Filtragem Baseada em Conteúdo
FC	Filtragem Colaborativa
ISDB-Tb	Integrated Services Digital Broadcast – Terrestrial, Brazilian version
ISDTV	Internation Standard Digital Television
JVM	Java Virtual Machine
LCT	Logical Channel Table
MHP	Multimedia Home Platform
MPEG	Moving Pictures Experts Group
NCL	Nested Context Language
NIT	Network Information Table
PAT	Program Association Table
PMT	Program Map Table
SBTV D	Sistema Brasileiro de Televisão Digital
SI	Service Information
STB	Set-Top-Box
TDT	Time and Data Table
TF-IDF	Term Frequency – Inverse Document Frequency
TVDI	Televisão Digital Interativa
UML	Unified Modeling Language
URD	Unidade Receptora Decodificadora
XML	Extensible Markup Language

SUMÁRIO

1. INTRODUÇÃO.....	13
1.1 CONTEXTUALIZAÇÃO.....	13
1.2 JUSTIFICATIVAS	15
1.3 OBJETIVOS.....	16
1.3.1 OBJETIVO GERAL	16
1.3.2 OBJETIVOS ESPECÍFICOS.....	16
1.4 CONTRIBUIÇÕES.....	17
1.5 TRABALHOS RELACIONADOS.....	17
1.6 ESTRUTURA DA DISSERTAÇÃO	20
2. SISTEMAS DE RECOMENDAÇÃO	22
2.1. INTRODUÇÃO.....	22
2.2. ARQUITETURA.....	23
2.3. METODOLOGIA	24
2.3.1 COLETA DE INFORMAÇÕES E MODELAGEM DE USUÁRIO	25
2.3.2 GRAU DE PERSONALIZAÇÃO	25
2.3.3 APRESENTAÇÃO DA RECOMENDAÇÃO	26
2.3.4 DEFINIÇÃO DO PROCESSO DE RECOMENDAÇÃO.....	26
2.3.5 AVALIAÇÃO DE SISTEMAS DE RECOMENDAÇÃO.....	27
2.4. TÉCNICAS DE RECOMENDAÇÃO	27
2.4.1 FILTRAGEM BASEADA EM CONTEÚDO	28
2.4.2 DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS E MINERAÇÃO DE DADOS.....	29
2.4.3 HÍBRIDA	31
2.5. CONSIDERAÇÕES FINAIS	31
3. TELEVISÃO DIGITAL INTERATIVA E O ISDB-TB.....	32
3.1 TOPOLOGIA DE UM SISTEMA DE TELEVISÃO DIGITAL INTERATIVA	32
3.2 DESENVOLVIMENTO DE APLICAÇÕES	34
3.2.1 XLET.....	35
3.3 COMPONENTES / APLICATIVOS PARA TV DIGITAL.....	35
3.3.1 GERADOR DE CARROSSEL	35
3.3.2 TABELAS SI	36
3.3.3 EPGS	37
3.4 O SISTEMA BRASILEIRO DE TELEVISÃO DIGITAL.....	38
3.4.1 ARQUITETURA DO SISTEMA BRASILEIRO DE TELEVISÃO DIGITAL	39

3.4.2	O MIDDLEWARE DO ISDB-TB	39
3.5	CONSIDERAÇÕES FINAIS.....	40
4.	ESTRATÉGIA HÍBRIDA DE RECOMENDAÇÃO.....	41
4.1	ESTRUTURA DA ESTRATÉGIA DE RECOMENDAÇÃO	41
4.1.1	COLETA DE INFORMAÇÕES E MODELAGEM DE USUÁRIO	41
4.1.2	GRAU DE PERSONALIZAÇÃO	42
4.1.3	APRESENTAÇÃO DA RECOMENDAÇÃO	43
4.1.4	DEFINIÇÃO DO PROCESSO DE RECOMENDAÇÃO.....	43
4.1.5	AValiação DE SISTEMAS DE RECOMENDAÇÃO.....	44
4.2	REQUISITOS FUNCIONAIS DO SISTEMA DE RECOMENDAÇÃO	45
4.2.1	DETERMINAR MODELO DO TELESPECTADOR (CRIAR PERFIL).....	45
4.2.2	CRIAR PERFIL ‘GERAL’	46
4.2.3	RECOMENDAR PROGRAMAÇÃO	46
4.2.4	AVALIAR RECOMENDAÇÃO	47
4.3	PROPOSTA ARQUITETURAL DO SISTEMA DE RECOMENDAÇÃO	47
4.4	DIAGRAMAS UML	49
4.4.1	DIAGRAMA DE CLASSES.....	50
4.4.2	DIAGRAMAS DE SEQUÊNCIA.....	52
4.5	CONSIDERAÇÕES FINAIS.....	57
5.	IMPLEMENTAÇÃO E RESULTADOS.....	58
5.1	AMBIENTE DE DESENVOLVIMENTO.....	58
5.2	A ESTRUTURA DE BANCO DE DADOS UTILIZADA	59
5.3	A IMPLEMENTAÇÃO DA AVALIAÇÃO DO SISTEMA DE RECOMENDAÇÃO.....	61
5.4	A GERAÇÃO DAS REGRAS DE ASSOCIAÇÃO.....	63
5.5	A IMPLEMENTAÇÃO DO EPG	65
5.5.1	CLASSE EPGGRADE.....	66
5.5.2	CLASSE BARRA.....	66
5.5.3	CLASSE INFORMAÇÃO	66
5.5.4	CLASSE CADASTRARUSUÁRIO	67
5.5.5	CLASSE TELESPECTADORPERFILFORM	67
5.5.6	CLASSE PROGRAMAÇÃO	68
5.6	TELAS DA APLICAÇÃO.....	68
5.7	PROVA DE CONCEITO	72
5.8	CONSIDERAÇÕES FINAIS.....	75
6.	CONSIDERAÇÕES FINAIS	76

6.1	TRABALHOS FUTUROS.....	77
	REFERÊNCIAS BIBLIOGRÁFICAS	79

1. INTRODUÇÃO

No primeiro capítulo desta dissertação, iremos apresentar, após breve contextualização sobre os assuntos abordados no trabalho, quais os objetivos com o desenvolvimento deste trabalho, bem como as justificativas que nos motivaram para a realização do mesmo.

1.1 CONTEXTUALIZAÇÃO

A vasta gama de informações disponíveis tem feito com que encontrar aquelas que mais se encaixam no perfil de quem as procura se torne algo cada vez mais difícil. Esta enorme quantidade de conteúdo digital, sobretudo quando falamos no escopo da web, tem causado o fenômeno conhecido como *sobrecarga de informação* (*information overload*), que dificulta bastante os processos de busca *online* (HUANG et al., 2002) por parte dos usuários.

Neste contexto, os Sistemas de Recomendação constituem-se numa solução para o problema acima, uma vez que ajudam o usuário na procura de informações relevantes (SCHAFFER; KONSTAN; RIEDL, 2001). Este auxílio também consiste em identificar conteúdos de acordo com alguns critérios, oferecendo, em seguida, informações aos seus respectivos usuários solicitantes (CAZELLA; ALVARES, 2005). Burke (2002) define sistemas de recomendação como qualquer sistema que produza recomendações individualizadas como saída, ou que tenha o efeito de guiar o usuário de forma personalizada a objetos interessantes e úteis, diante de uma grande variedade de opções.

Os sistemas de recomendação atuam baseados em personalização da informação, uma vez que auxiliam usuários no processo social de indicar ou receber indicação de itens pertencentes a um domínio específico. A personalização está relacionada com o modo pelo qual a informação e serviços podem ser ajustados às necessidades específicas de um usuário ou comunidade (CALLAN et al., 2002).

Vale salientar que a mídia televisiva, interativa e digital, uma nova tecnologia da informação e comunicação (NTIC), passa a fazer parte da vida das pessoas, deixando de ser vista como apenas mais uma tecnologia. Ao contrário, passa a ser vista como complemento, como companhia, como continuação do espaço de vida das pessoas. Por meio do que é transmitido pela televisão, as pessoas adquirem informações e transformam seus

comportamentos, tornando-se consumidores ativos, permanentes e críticos de tudo o que é oferecido pelo universo televisivo (VALENTE, 2002).

De forma sucinta, podemos dizer que a televisão digital é um sistema de transmissão, recepção e processamento de sinais em formato digital, que permitem uma alta definição de som e imagens, além de grande interatividade. A televisão que a grande maioria dos brasileiros conhece é a analógica. Uma limitação básica encontrada na televisão analógica ou qualquer outra tecnologia que utilize o espaço como meio de transmissão, é a necessidade de ocupar uma faixa de frequência para que se possa realizar a transmissão dos sinais adequadamente (FERNANDES; LEMOS; SILVEIRA, 2004)(BECKER; MONTEZ, 2004).

Hoje a televisão digital é uma realidade, apresentando inúmeras vantagens quando comparada aos sistemas analógicos. Enquanto num canal analógico existem dificuldades de manipulação e armazenamento do sinal, isso não ocorre num sistema digital. Essa necessidade parte do princípio de que, apesar de aparentar bastante espaço, os 6MHz destinados a uma TV analógica é bastante limitado. Com a digitalização é possível comprimir o sinal de tal modo que se possa acrescentar outros sinais complementares, seja para garantir maior segurança na efetivação da transmissão ou na inserção de novos serviços. Inúmeros são os serviços que podem ser suportados por esta plataforma interativa, tais como *t-banking*, *t-commerce*, *t-gov*, *t-learning* (MORAES NETO, 2006)(MORAES NETO, 2007), dentre outros.

Atualmente, o Brasil se encontra em estágio de implantação do seu modelo nacional de televisão digital interativa, atualmente conhecido por *Integrated Services Digital Broadcasting Terrestrial* (ISDB-Tb), modelo este oriundo de um projeto iniciado em 2003, denominado Sistema Brasileiro de Televisão Digital, SBTVD.

O ISDB-Tb aproveitou o fato da nossa rede nacional de televisão ser a única rede de telecomunicações gratuita, com cobertura nacional, chegando em todas as camadas da sociedade, tendo uma cobertura de mais de 95% (noventa e cinco por cento) dos lares brasileiros. Por se tratar de um instrumento que permite o acesso da população de baixa renda às redes de informações e serviços digitais, o governo brasileiro percebeu a grande chance de alavancar a inclusão digital no Brasil, garantindo a digitalização dos sinais de televisão, permitindo acesso de todas as camadas a um recurso de Tecnologia de Informação e Comunicação de ótima usabilidade, fácil acesso e abrangência sem igual (ANATEL, 2004).

Dessa forma, levando em conta a digitalização dos sinais de TV e os problemas de sobrecarga de informação inerentes a isto, utilizando os princípios, técnicas e características presentes na literatura dos sistemas de recomendação, o objetivo desta dissertação de

mestrado é **propor uma estratégia de recomendação personalizada de programas televisivos a ser utilizada no contexto da televisão digital interativa.**

No que concerne à definição de uma estratégia de recomendação, inúmeras técnicas podem ser utilizadas, de forma isolada ou combinada, para que as recomendações sejam geradas (REATEGUI; CAZELLA, 2009). No nosso caso, utilizamos a Filtragem Baseada em Conteúdo (FBC) (BALABANOVIC; SHOHAM, 1997), que utiliza as preferências do usuário para, após fazer análise dos conteúdos em questão, recomendar itens que o sistema julga ser de seu interesse. Também podemos citar, no escopo das técnicas de mineração de dados, a utilização das Regras de Associação (AGRAWAL; IMIELINSKI; SWAMI, 1993), que representam uma técnica de mineração aplicada para descobrir de forma automática padrões escondidos, tendo seu relacionamento representado como uma regra de associação. Os algoritmos utilizados para a implementação das duas técnicas citadas são, respectivamente, o TF-IDF (*Term Frequency – Inverse Document Frequency*) e o Apriori.

Finalizando, temos a possibilidade do preenchimento de informações específicas do usuário, através da funcionalidade de criação de perfis e também disponibilizamos a funcionalidade de Avaliação de Recomendação, onde a mesma permite que o usuário avalie aquilo que o sistema recomendou (*feedback* do usuário) e permita que o mesmo passe por um contínuo processo de melhoria.

1.2 JUSTIFICATIVAS

Pesquisas e desenvolvimento de projetos e/ou trabalhos que visem à construção de sistemas de recomendação podem ser considerados relativamente importantes, dado a natureza colaborativa inerente a um sistema computacional que visa auxiliar um usuário na escolha de algum item pertencente a um domínio específico. Além disso, a correlação deste tema com outros temas da Computação, tais como Aprendizagem de Máquina e Mineração de Dados, em termos acadêmicos, revelam a importância do assunto em pauta.

Quando falamos em desenvolvimento de aplicações a serem suportadas por uma plataforma televisiva, a contribuição de um trabalho merece destaque específico. Embora não seja escopo do trabalho desenvolver uma estratégia voltada ao modelo brasileiro de televisão digital, quando analisamos dados estatísticos relacionados a esta mídia, sobretudo em nível nacional, atualmente mais de 50 milhões de lares possuem aparelhos de televisão no Brasil. O sistema brasileiro de televisão aberta, que é um dos maiores do mundo, possui seu acesso totalmente gratuito, sendo utilizado por mais de 95% (noventa e cinco por cento) dos lares brasileiros.

Todas estas informações possuem a finalidade de demonstrar a grande responsabilidade e abrangência que este tema exerce na sociedade, especificamente a brasileira, e justificando a contribuição sempre existente em trabalhos que abordem o uso desta mídia. Neste sentido, o governo federal, ao decretar o SBTVD, tinha as seguintes principais premissas (MINICOM, 2003):

- Promover a inclusão social, a diversidade cultural do País e a língua pátria por meio do acesso à tecnologia digital, visando à democratização da informação;
- Propiciar a criação de rede universal de educação a distância;
- Estimular a pesquisa e o desenvolvimento e propiciar a expansão de tecnologias brasileiras e da indústria nacional relacionadas à tecnologia de informação e comunicação.

Falando em televisão digital e na vasta gama de oferta de programação televisiva, facilmente nos reportamos a realidade da tv por assinatura no Brasil. Atualmente são mais de 15 operadoras atuando no país transmitindo mais de 500 diferentes canais. A mesma possui uma abrangência de mais de 10 milhões de lares e de aproximadamente 33 milhões de usuários (TV POR ASSINATURA, 2011). Consideramos válido comentar a realidade desta modalidade de utilização de programação televisiva pois, mesmo que o processo de digitalização do sinal aberto de tv demore a passar a conter uma grande quantidade de canais, esta volumosa oferta já é uma realidade na televisão por assinatura.

Desta forma, facilitar o uso dos recursos da televisão digital interativa, aberta ou paga (por assinatura), bem como desenvolver um trabalho de pesquisa que aborde tecnologias, ferramentas e conceitos neste domínio representam uma contribuição para a comunidade acadêmica.

1.3 OBJETIVOS

1.3.1 OBJETIVO GERAL

Desta forma, o objetivo geral desta dissertação é propor uma estratégia de recomendação personalizada de programas televisivos a ser utilizada no contexto da televisão digital interativa.

1.3.2 OBJETIVOS ESPECÍFICOS

O objetivo geral descrito acima pode ser fragmentado nos seguintes objetivos específicos:

- Desenvolver uma abordagem de recomendação híbrida baseada na combinação de técnicas e algoritmos de recomendação existentes adequada ao domínio da Televisão Digital Interativa.
- Desenvolver um sistema de recomendação de programas para os usuários telespectadores que implemente a abordagem proposta utilizando como base as informações de programação disponibilizadas pelas tabelas SI e do guia de programação eletrônico (MORAES NETO; OLIVEIRA; SOUZA, 2005).

1.4 CONTRIBUIÇÕES

As principais contribuições esperadas como resultados desta dissertação são as seguintes:

- Possibilitar a difusão e discussão sobre a tecnologia, os fundamentos e os problemas dos Sistemas de Recomendação;
- Resultar no desenvolvimento de soluções inovadoras, através da especificação e construção de algoritmos, para os Sistemas de Recomendação;
- Possibilitar amplo e aprofundado estudo acerca de uma mídia de informação e comunicação tão abrangente e difundida, como a televisão;
- Oferecer aos modelos de televisão digital, um sistema que realize recomendações de programas televisivos, fazendo com que seus usuários telespectadores passem a ser auxiliados quando da necessidade de orientação sobre quais programas devem assistir;
- Incentivar a atração de novos pesquisadores interessados nas áreas de atuação deste projeto;
- Ajudar na cooperação com grupos de pesquisa de outras instituições que trabalhem com temas relacionados, por meio da divulgação deste trabalho através de publicações científicas.

1.5 TRABALHOS RELACIONADOS

Ao realizarmos um estudo bibliográfico em trabalhos voltados para a área de Sistemas de Recomendação e Televisão Digital Interativa, percebemos que a literatura fornece uma série de trabalhos relacionados que nos auxiliam a imaginar outras estratégias que contemplem melhorias no que já foi proposto ou então sugerir inovações que ofereçam alguma forma de contribuição para a sociedade.

Iniciando a citação de alguns trabalhos similares, percebemos que o TV Scout (BAUDISCH; BRUECKNER, 2002) é um sistema de recomendação que fornece aos usuários a programação personalizada dos canais de televisão. Entretanto, esta recomendação foi

desenvolvida para a plataforma web. O sistema utiliza uma estratégia baseada na técnica de filtragem de informação utilizando a abordagem colaborativa e também realiza a atualização dinâmica do perfil. A técnica utilizada é bastante simples, não tendo havido a preocupação na geração de estratégias de recomendação mais elaboradas.

Buczak, Zimmerman e Kurupati (2002) propõem um sistema de recomendação que, embora faça uso de interface gráfica do tipo EPG, são mais especificamente voltados para um cenário mais avançado, exigindo a existência de recursos do tipo PVR (*Personal Video Recorder*) e terminais de acesso com elevada capacidade de processamento. A restrição imposta pela utilização de recursos tecnológicos mais avançados pode prejudicar o aspecto de inclusão social, reduzindo a chance de utilização por parte de todos os setores da sociedade. Complementando, aqui não se percebe a preocupação com a criação de perfis e nem com a avaliação do sistema.

Em (DIFINO; NEGRO; CHIAROTTO, 2004), é apresentado um sistema personalizado de guia de programação eletrônico. Este sistema é baseado numa arquitetura multi-agente, com filtragem de informação. O sistema se preocupa com a questão da personalização e da atualização do perfil, entretanto, o trabalho demonstrou um estudo de caso da proposta baseado em um ambiente formado por computadores pessoais, sem ser apresentados resultados relevantes que comprovem o desempenho da proposta. Os autores também informaram da incipiência da proposta.

A realização da recomendação de programas de televisão através da criação de um modelo de usuário baseado apenas na coleta explícita de dados é encontrado no sistema PTV (COTTER; SMYTH, 2000). A estratégia proposta também é baseada na abordagem de filtragem de informação, entretanto, não prevê a interpretação automática do comportamento do usuário (coleta implícita).

A questão da relevância de opinião é perfeitamente abordada em (ZIMMERMAN; KURUPATI; PARAMESWARAN, 2002). Este trabalho apresenta o conceito de Celebidades. Na realidade, a preocupação maior é em verificar o quanto a opinião de pessoas consideradas “celebidades” influencia na decisão de usuários em preferir ou não determinados programas televisivos. O foco deste trabalho consiste na avaliação da relevância da opinião das “celebidades” no processo de recomendação, abstraindo-se dos detalhes específicos da televisão digital e das técnicas presentes dos sistemas de recomendação.

No que se refere ao uso de metadados no processo de recomendação de conteúdos televisivos, Smyth, Wilson e Sullivan (2004) propõem um sistema que realiza a mineração de dados em metadados de programação, objetivando a ocorrência da recomendação. Nenhum

detalhe acerca da análise e interpretação dos metadados é apresentada, bem como não foram abordados detalhes da realidade da televisão digital.

Por último, ainda em nível mundial, temos, em (UCHYIGIT; CLARK, 2002), mais uma vez um guia de programação eletrônico personalizado, baseado em uma infra-estrutura de agentes e em um processo de recomendação colaborativa. Tais agentes também se responsabilizam pela atualização constante do perfil dos usuários, visando uma recomendação eficiente, que é baseada em gêneros. O trabalho fornece detalhes bastante específicos acerca da criação do modelo do telespectador, entretanto, não aborda detalhes sobre os passos executados para a recomendação, logo após a criação do modelo do usuário.

No contexto nacional, Silva (2005) apresentou o SRPTV, um sistema de recomendação de programas de TV cujo algoritmo de filtragem de informação utiliza o Modelo Vetorial. O autor também realizou um estudo comparativo entre os algoritmos Rocchio e KNN (obtendo melhores resultados com o segundo algoritmo), para atualização do perfil do usuário. No trabalho é realizada a coleta de dados tanto da forma implícita quanto explícita, no entanto, a recomendação é feita através de uma página web, e não através da televisão digital. A estratégia não utiliza mineração de dados, ou seja, não realiza a extração de padrões não triviais, assim como também não analisa as informações de perfis de usuários, também não permitindo o aprimoramento constante do sistema através de processos avaliativos da recomendação realizada.

Mais recentemente, Ávila e Zorzo (2009) propõem o RecommenderTV. A implementação do sistema de recomendação procura ser aderente ao middleware brasileiro, e prevê a inclusão de funcionalidades adicionais, de forma específica, no Ginga-NCL. O objetivo do trabalho é comparar a eficiência de três algoritmos pertencentes ao escopo da mineração de dados, no caso, o Naive Bayes, C.45 e Apriori. Esta comparação tem o intuito de escolher um dos três algoritmos já existentes, de forma isolada, para utilização no sistema, sem que seja elaborada uma estratégia híbrida e combinada para este fim. Neste trabalho, o algoritmo Apriori obteve melhor desempenho (72%), fator que determinou a escolha do mesmo. No trabalho em questão, não identificamos a questão da preocupação com a criação de perfis dos usuários (o que nos deixa sem entender o que subsidia a realização da recomendação), avaliação da recomendação e aprimoramento da mesma.

Na Tabela 1 é apresentado um quadro consolidado que resume as principais características dos trabalhos aqui citados. Para alguns trabalhos isto não foi possível, devido à ausência de maiores informações a respeito. Esta mesma tabela subsidiará o comparativo

realizado no Capítulo 5, entre o desempenho (percentual de acerto) da nossa proposta com as demais aqui citadas:

Tabela 1: Quadro contendo as principais características dos trabalhos relacionados.

Nome do autor ou do trabalho	Para TV DIGITAL	Criação de Perfil	Estratégia combinada	Estratégia utilizada	Módulo de Avaliação	Evolução do sistema
<i>TV Scout</i>	Não	Sim	Não	Filtragem colaborativa	Não	Não
<i>Buczak, Zimmerman e Kurupati</i>	Sim	Não	Não	Não informado	Não	Não
<i>Difino, Negro e Chiarotto</i>	Não	Sim	Não	Filtragem de informação (TF-IDF)	Não	Sim
<i>PTV</i>	Sim	Sim	Não	Filtragem de informação (TF-IDF)	Não	Não
<i>Zimmerman, Kurupati e Parameswaran</i>	Não	Não	Sim	Filtragem de informação (TF-IDF) e Filtragem colaborativa	Não	Sim
<i>Uchyigit e Clark</i>	Não	Sim	Não	Filtragem colaborativa	Não	Sim
<i>RecommenderTV</i>	Sim	Não	Não	Apriori	Não	Não

Na sua maioria, os trabalhos citados são baseados em técnicas existentes, utilizadas de forma isolada, como por exemplo, técnicas de Filtragem Baseada em Conteúdo ou Regras de Associação, alguns dos quais são voltados para a recomendação de programas televisivos.

Na busca por uma abordagem de recomendação mais eficiente, o presente trabalho propõe uma estratégia híbrida combinando diferentes técnicas e procedimentos, no contexto da recomendação de programas televisivos através de um EPG. Além disso, dois outros fatores foram associados: as informações presentes no perfil do usuário e as informações fornecidas pelo mesmo durante o processo de avaliação da recomendação. Na esfera nacional, encontramos apenas dois trabalhos similares, no que tange ao objetivo geral, mas bastante diferentes quando analisamos as técnicas utilizadas e a arquitetura da mesma.

1.6 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação é composta de seis capítulos e um apêndice. A divisão dos capítulos procurou ser feita de forma similar à maneira pela qual um sistema de computação é especificado, ou seja, através da apresentação de requisitos, modelagem, arquitetura e projeto. A estrutura é apresentada a seguir:

- **Capítulo 1 – Introdução:** Descrição do problema, contextualização e justificativas, bem como alguns trabalhos relacionados;
- **Capítulo 2 – Sistemas de Recomendação:** São apresentados os principais conceitos envolvidos nesta abordagem, apresentando sua arquitetura, seus módulos e as principais técnicas utilizadas nos processos de recomendação;
- **Capítulo 3 – Televisão Digital Interativa e o ISDB-Tb:** São apresentados os principais conceitos e tecnologias envolvidas, detalhando a arquitetura destes sistemas e os principais modelos comerciais existentes. Também será apresentado o sistema de televisão digital nacional, conhecido por *Integrated Services Digital Broadcasting Terrestrial* (ISDB-Tb);
- **Capítulo 4 – A Estratégia Híbrida de Recomendação:** Além da demonstração da metodologia utilizada para a construção da estratégia proposta, também é feita descrição dos módulos da arquitetura, detalhando a combinação de estratégias e abordando as tecnologias utilizadas. Adicionalmente, neste capítulo, são apresentados os requisitos do sistema desenvolvido e demais artefatos de modelagem;
- **Capítulo 5 – Implementação e Resultados:** É realizada uma prova de conceito arquitetural visando ilustrar o desempenho da estratégia de recomendação fruto desta dissertação e visando demonstrar como ela atende aos requisitos propostos. Antes disso, todos os detalhes de implementação da proposta são devidamente demonstrados;
- **Capítulo 6 – Considerações finais:** É feita uma análise, em caráter de conclusão, sobre os principais resultados alcançados neste trabalho, apresentando sugestões de trabalhos futuros e algumas discussões acerca dos resultados obtidos até então;
- **Apêndice A – Requisitos do Sistema:** São abordados os requisitos do proposto sistema de recomendação utilizado para validação da estratégia híbrida de recomendação. Será feita a especificação funcional e não funcional destes requisitos, realizando o detalhamento dos casos de uso e mostrando os diagramas oportunos para esta etapa do processo de desenvolvimento da Engenharia de Software;

2. SISTEMAS DE RECOMENDAÇÃO

Neste capítulo abordaremos os conceitos voltados a sistemas de recomendação, apresentando os principais conceitos, bem como apresentando sua arquitetura, seus módulos e as principais técnicas utilizadas nos processos de recomendação. Também iremos apresentar alguns trabalhos nesta área e exemplos de algumas aplicações que objetivam a recomendação de informações/conteúdos para diferentes tipos de usuários.

2.1. INTRODUÇÃO

Com a crescente quantidade de informações digitais, encontrar aquelas que mais se adequam ao procurado torna-se cada vez mais difícil. Há mais informações acessíveis do que um ser humano poderia absorver em toda a sua vida. A principal dificuldade é que grande parte desta apresenta-se pouco estruturada e muitas vezes não atende às necessidades do usuário, ocasionando grande investida de esforços e perda de tempo.

A enorme quantidade de informações digitais, em particular no contexto da web, tem causado o fenômeno conhecido como “sobrecarga de informação” (*information overload*) que dificulta bastante os processos de busca *online* (HUANG et al., 2002) por parte dos usuários. Normalmente, usuários com diferentes níveis de conhecimento, experiência e interesses são igualmente providos com a mesma informação, em resposta a uma mesma consulta. Muitas vezes, estes indivíduos possuem pouca ou quase nenhuma experiência para realizar escolhas entre as várias alternativas que lhes são apresentadas.

Os Sistemas de Recomendação constituem-se em uma solução para o problema acima, ajudando o usuário na procura de informações relevantes (SCHAFFER; KONSTAN; RIEDL, 2001). A principal funcionalidade dos Sistemas de Recomendação é identificar itens com base em alguns critérios, para então oferecer as informações mais adequadas aos seus usuários (CAZELLA; ALVARES, 2005).

Burke (2002) define sistemas de recomendação como qualquer sistema que produza recomendações individualizadas como saída, ou que tenha o efeito de guiar o usuário de

forma personalizada a objetos interessantes e úteis, diante de uma grande variedade de opções.

Conforme Reategui e Cazella (2005), os websites de comércio eletrônico são atualmente o maior foco de utilização dos sistemas de recomendação, empregando diferentes técnicas para encontrar os produtos mais adequados para seus clientes e aumentar deste modo sua lucratividade. Sistemas de recomendação fornecem aos consumidores informações de forma a facilitar o processo de tomada de decisão (SCHAFFER; KONSTAN; RIEDL, 2001), provendo ao usuário uma visão personalizada da informação.

A personalização da informação está relacionada com o modo pelo qual a informação e os serviços podem ser ajustados às necessidades específicas de um usuário ou comunidade (CALLAN et al., 2002). Esta funcionalidade pode ser alcançada através da adaptação da apresentação, conteúdo e/ou serviços levando em conta diversas propriedades.

Os proponentes do primeiro sistema de recomendação, que foi denominado Tapestry (RESNICK; VARIAN, 1997), iniciaram classificando esta aplicação como um aplicativo de “filtragem colaborativa”. Entretanto, tal designação não foi bem aceita por demais autores, que preferiram utilizar a expressão sistemas de recomendação, por ser um termo genérico.

Analizando as áreas de atuação onde estes sistemas estão presentes, percebemos a grande contribuição que os mesmos oferecem. Além de serem utilizados com frequência em *sites* de empresas de comércio eletrônico, na área de *marketing*, tais sistemas identificam os tipos de produtos e serviços através do histórico do cliente (SHARDANAND; MAES, 1995).

Na área de educação, tais sistemas têm sido empregados na identificação de sequências de navegação apropriadas para cada estudante, de forma análoga à recuperação de experiências em ambientes interativos de aprendizagem (SCHANK; CLEARY, 1995). Ainda dentro da área da educação, devemos destacar a grande contribuição dos mesmos quando se fala em Sistemas Tutores Inteligentes (MURRAY, 1999)(MITCHELL; GROGONO, 1993). No que diz respeito aos aspectos da Interface Humano Computacional (IHC), existem algumas considerações a serem levadas em conta quando da criação de sistemas para este fim. Os trabalhos de (ALBUQUERQUE et al, 2008)(BECKER, 2006)(KIM et al, 2007) abordam estes aspectos.

2.2. ARQUITETURA

Um sistema de recomendação deve funcionar como um servidor capaz de recuperar, para determinado usuário, suas preferências explícitas, implícitas, conteúdo e ofertas relacionados a itens nos quais já mostrou interesse anteriormente. São dois os principais tipos

de arquitetura possíveis em sistemas de recomendação (REATEGUI; CAZELLA, 2005). No primeiro deles, baseado em técnicas de filtragem de informação, não há componente de mineração de dados, já que o sistema opera através da análise direta dos dados dos usuários, sem a construção de um modelo de recomendação generalizado.

Já no segundo tipo de arquitetura são utilizados os conceitos de descoberta de conhecimento em bases de dados e mineração de dados, onde, através da leitura e interpretação de informações armazenadas, é gerado um modelo de usuário. De acordo com este segundo tipo de arquitetura, a Figura 1 apresenta a arquitetura genérica dos sistemas de recomendação:

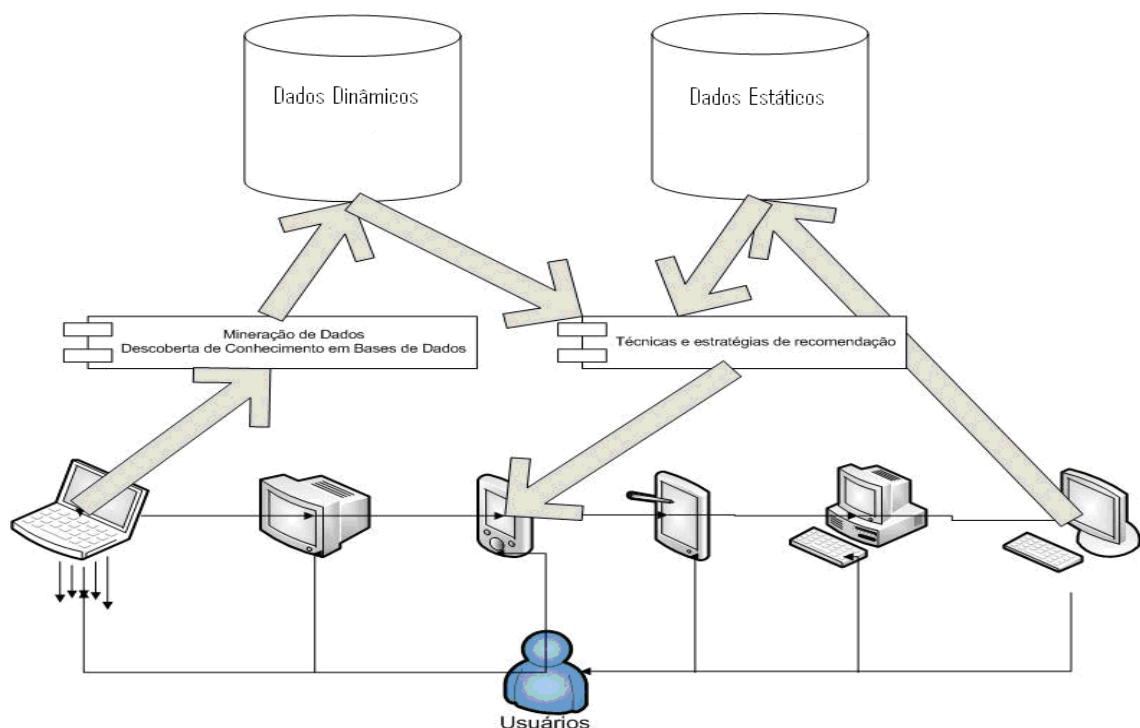


Figura 1: Arquitetura de um sistema de recomendação.

Resumidamente, os Dados Dinâmicos representam as informações coletadas de acordo com a interação do usuário com a aplicação, ao passo que os Dados Estáticos são fornecidos pelo próprio. As técnicas e estratégias de recomendação, bem como os algoritmos de Mineração, tem a função de analisar as fontes citadas anteriormente e, a partir disso, gerar as recomendações para os usuários, independente do tipo de dispositivo que o mesmo utilize.

2.3. METODOLOGIA

A estrutura metodológica contendo os principais passos envolvidos no desenvolvimento de um sistema de recomendação é apresentada em (REATEGUI; CAZELLA, 2005).

2.3.1 COLETA DE INFORMAÇÕES E MODELAGEM DE USUÁRIO

Para que seja possível recomendar itens a um usuário é necessário capturar e armazenar seus dados pessoais e comportamentais relativos, a partir dos quais é definido o modelo do usuário. Esse modelo é essencial para a identificação do usuário no momento em que ele acessa o sistema onde foram implantadas as rotinas de recomendação. Existem duas maneiras de coletar dados sobre o usuário, a saber:

- **Explícita:** Nesta, o usuário precisa informar ao sistema de forma clara e precisa o que lhe é importante. Esta informação pode ser coletada, por exemplo, através do preenchimento de formulários específicos.
- **Implícita:** Através desta técnica, é possível conhecer as preferências dos usuários sem que eles tenham que fornecer informações explicitamente. A partir das ações do usuário, inferem-se informações sobre suas necessidades e preferências. Neste tipo de coleta, a idéia é que o conhecimento extraído sobre o comportamento do usuário seja utilizado não para a recomendação para o mesmo, e sim para os demais usuários cadastrados no sistema.

A criação de um modelo do usuário envolve o estudo das diferentes formas através das quais as informações dos usuários podem ser adquiridas, representadas e usadas por sistemas computacionais. O objetivo desta atividade é permitir que o sistema conheça o seu usuário, podendo assim apresentar para estes, apenas o que é de seu interesse (HAN; KAMBER, 2001)(REATEGUI; CAZELLA, 2005). Este conhecimento é importante devido ao fato de tornar possível uma melhoria na interação do usuário com o sistema. De forma geral, para que a recomendação possa acontecer de forma satisfatória, os Sistemas de Recomendação são integrados com um banco de perfis dos seus usuários contendo as informações referentes a dados cadastrais, a área e o grau de interesse ou de conhecimento do usuário em determinados assuntos (OLIVEIRA et al., 2005).

2.3.2 GRAU DE PERSONALIZAÇÃO

Nesta fase é estabelecido em que medida o sistema de recomendação irá adequar a recomendação das informações em relação às características dos seus usuários. Neste contexto temos os seguintes graus de personalização:

- **Não personalizada:** Quando uma aplicação oferece a mesma recomendação para todos os usuários;

- **Efêmera:** Sistemas que utilizam as informações fornecidas pelo usuário para personalizar a recomendação;
- **Persistente:** Trata cada usuário de acordo com as suas preferências, oferecendo diferentes itens para diferentes usuários, mesmo que eles tenham escolhido os mesmos itens. Isso acontece, pois as demais informações dos perfis destes são levadas em consideração;

2.3.3 APRESENTAÇÃO DA RECOMENDAÇÃO

Esta fase se refere à definição da forma de apresentação da recomendação e ao momento em que a recomendação é apresentada ao usuário, podendo ser durante o uso do sistema (*on line*), ou não (*off line*). Abaixo seguem três dos possíveis momentos:

- **Push:** Itens são recomendados sem que o usuário precise interagir com o sistema, como por exemplo, através do envio de e-mail;
- **Pull:** A recomendação só é apresentada para o usuário quando solicitado explicitamente;
- **Passiva:** As recomendações aparecem no contexto da aplicação, não necessitando de solicitação por parte dos usuários para que sejam exibidas.

2.3.4 DEFINIÇÃO DO PROCESSO DE RECOMENDAÇÃO

Várias tecnologias têm surgido visando à identificação de padrões de comportamento, e utilização destes padrões na personalização do relacionamento com os usuários. Estas, além de estarem diretamente relacionadas, fundamentam o funcionamento dos sistemas de recomendação.

A escolha da estratégia de recomendação influencia também na forma e no tipo de saída que se espera do sistema de recomendação (predição, recomendação Top-N, listas de recomendação, avaliação de usuários, dentre outras). A resposta desejada para o sistema está diretamente relacionada ao problema que se pretende resolver, por isso deve ser definida antes de se partir para as próximas etapas;

Esta etapa é concluída com a definição do algoritmo que implementa o processo de recomendação (Técnica de Recomendação), estabelecendo a forma em que os dados de entrada e os dados armazenados serão combinados pelos algoritmos para gerar recomendações.

2.3.5 AVALIAÇÃO DE SISTEMAS DE RECOMENDAÇÃO

Assim como qualquer produto de software, a qualidade do sistema de recomendação desenvolvido precisa ser avaliada no intuito de verificar se seu funcionamento é adequado. No caso de sistemas de recomendação, a avaliação envolve, principalmente, determinar o grau de precisão das recomendações geradas. Entretanto, existe uma grande dificuldade quando se pretende avaliar recomendações de um sistema em seu estágio inicial de funcionamento, quando ainda não existem dados suficientes para a realização dos testes (RICCI; DEL MISSIER, 2004)(GUERRERO, 2004).

Quando o sistema de recomendação é promovido à produção e passa a conter uma significativa massa de dados, diversas métricas são propostas para avaliar seu desempenho (VOZALIS; MARGARITIS, 2003). De maneira geral, os dados acumulados são divididos em: conjunto de treinamento, que representa os dados prévios do sistema, e conjunto de teste, que simula novas entradas de dados, permitindo comparar a predição feita pelo sistema com a avaliação efetivamente realizada pelo usuário (BURKE, 2002).

Neste sentido, visando obter informações sobre eficácia de um sistema de recomendação, temos a métrica **Precisão**, talvez a mais comumente usada pela sua simplicidade. Esta métrica representa a relação entre o número de itens relevantes sobre o total de itens retornados. Pode ser vista também como a relação entre o número de itens que o usuário considera relevante e o número de itens recomendados. Assume-se que os itens considerados relevantes pelo usuário são aqueles com avaliações iguais ou superiores a quatro (HERLOCKER, 2004). A fórmula abaixo apresenta o cálculo da precisão:

$$P = \frac{N_{rel}}{N_r}$$

Utilizando esta métrica, passa também a ser possível verificar qual o impacto da utilização da relevância da opinião do usuário no que se refere à precisão dos itens recomendados.

2.4. TÉCNICAS DE RECOMENDAÇÃO

Temos uma série de técnicas presentes na literatura de sistemas de recomendação. A leitura de alguns trabalhos acadêmicos (BELVIN; CROFT, 1992)(HERLOCKER, 2000) (BURKE, 2002)(RICCI et al., 2002) poderá dar uma melhor noção a respeito destas. No intuito de não tornar cansativa a leitura desta dissertação, somente as técnicas utilizadas em nossa proposta serão comentadas.

2.4.1 FILTRAGEM BASEADA EM CONTEÚDO

A demanda por tecnologias de filtragem de informação não é algo novo. Segundo Belvin e Croft (1992), a filtragem de informação é o nome utilizado para descrever uma variedade de processos que envolvem a entrega de informação para as pessoas que realmente necessitam delas. Trata-se de uma abordagem que geralmente mantém um perfil dos interesses do usuário, o que veio totalmente de encontro às necessidades e características de nosso trabalho.

Em um sistema de filtragem baseado no conteúdo, o usuário fornece, de forma implícita ou explícita (ou seja, através da descrição de interesses do usuário fornecida por ele próprio ou através de ações, como seleção e aquisição de itens), suas preferências e restrições e o sistema casa estas descrições com os itens contidos em um catálogo de produtos. Estes sistemas podem utilizar o histórico de consultas passadas para construir um perfil do usuário (HERLOCKER, 2000)(ANSARI et al., 2000).

Os sistemas baseados no conteúdo utilizam apenas as preferências do usuário. Eles tentam recomendar itens que são similares aos que o usuário gostou no passado. O foco desses sistemas é *aprender* as preferências do usuário e filtrar, dentre os novos itens, aqueles que mais se adequarem a estas preferências. Nestes sistemas, cada item de um conjunto é definido por características associadas a ele. Por exemplo, um filme pode ter como características: título, diretor, elenco, gênero, etc. Já em um texto, as próprias palavras que o compõe podem ser consideradas como suas características.

É exatamente através destas características que se torna possível a comparação e semelhança entre estes itens. Assim, conhecendo o histórico de preferência do usuário por itens do conjunto e a semelhança entre estes itens, o sistema de recomendação baseado em conteúdo é capaz de recomendar para este usuário outros itens que possam ser de seu interesse.

Neste sentido, temos o TF-IDF, um algoritmo utilizado para a realização da Filtragem Baseada em Conteúdo. O algoritmo *Term Frequency-Inverse Document Frequency* (TF-IDF) (ADOMAVICIUS; TUZHILIN, 2005) é o algoritmo mais utilizado em conjunção com a técnica de filtragem baseada em conteúdo. O algoritmo objetiva o cálculo do número de vezes que uma palavra i aparece em um documento j . Com base no resultado obtido a partir do cálculo, um peso pode ser atribuído às palavras utilizadas no contexto de sistemas de recuperação de informação. Desta forma, palavras muito comuns que podem aparecer em

diversos documentos e consequentemente são pouco úteis para distinguir seu conteúdo podem ter seu peso reduzido.

O peso de uma palavra em um dado documento é calculado como sendo o produto das duas medidas (TF-IDF). Assim, o documento pode ser representado como um vetor de palavras e seus respectivos pesos. Como nos sistemas baseados em conteúdo, o modelo de usuário é obtido com base nas características dos itens previamente vistos e avaliados por ele. Embora a filtragem baseada no conteúdo venha sendo usada com sucesso em algumas aplicações, esta técnica apresenta uma série de limitações (ADOMAVICIUS; TUZHILIN, 2005), a saber:

- Os atributos dos itens geralmente precisam ser manualmente cadastrados, o que representa um sério problema quando estamos tratando com mídias como som e vídeo. Existe uma grande dificuldade na análise destes para extração automática de atributos;
- Impossibilidade de encontrar itens que podem vir a interessar ao usuário, mas não são similares, em termos de conteúdo, com outros itens que ele avaliou, sendo encontrados apenas os itens parecidos com os já conhecidos pelo usuário;
- Incapacidade na avaliação de um conteúdo quanto a dimensões subjetivas, como qualidade. Por exemplo, sistemas de recomendação de notícias de um jornal não conseguem diferenciar um texto mal escrito de um bem escrito com conteúdos muito semelhantes.

2.4.2 *DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS E MINERAÇÃO DE DADOS*

Uma forma de realizar a extração de conhecimento dos bancos de dados dos usuários constitui-se na aplicação de técnicas e ferramentas de mineração de dados. Este processo é conhecido como uma forma de realizar extração não trivial (conhecimento implícito) de informações potencialmente úteis, as quais não são previamente conhecidas e encontram-se implícitas em grandes coleções de dados (ZAIANE, 2000).

A mineração de dados é um recurso importante para a descoberta de relações entre itens, entre usuários e entre itens e usuários. Através da mineração dos arquivos de log (YANG; PARTHASARANTHY, 2002), por exemplo, podem-se obter conhecimentos aprofundados a respeito dos usuários. Segundo Witten e Frank (2000), a mineração de dados pode ser vista como o processamento de dados, buscando identificar os padrões requisitados.

Neste contexto, temos que as **Regras de Associação** (AGRAWAL; IMIELINSKI; SWAMI, 1993) representam uma técnica de mineração aplicada para descobrir, de forma automática, padrões escondidos, tendo seu relacionamento representado como uma regra de

associação. Tal fato pode ser interpretado como a presença de certos itens em uma transação, baseado na presença de outros itens.

Uma regra de associação é um relacionamento da forma $X \Rightarrow Y$, onde X e Y são conjuntos de itens e a interseção deles, $X \cap Y$, é o conjunto vazio. Cada regra de associação é associada a um fator suporte superior, denominado FSup, e a um fator de confiança, FConf. FSup é definido como a razão do número de tuplas satisfazendo ambos X e Y sobre o número total de tuplas, isto é, $F\text{Sup} = |X \cup Y| / N$, onde N é o número total de tuplas. FConf é definido como a razão do número de tuplas que satisfazem ambos X e Y sobre o número de tuplas que satisfazem X , isto é, $F\text{Conf} = |X \cap Y| / |X|$.

A tarefa de descobrir regras de associação consiste em extrair do banco de dados todas as regras com FSup e FConf maiores ou iguais a um FSup e FConf especificados pelo usuário. A definição desses fatores serve para introduzir uma medida capaz de distinguir associações interessantes, dado que as regras geradas em forma binária elevam em muito o espaço de busca de qualquer algoritmo minerador.

A descoberta de regras de associação é geralmente executada em dois passos. Primeiro, um algoritmo determina todos os conjuntos de itens que têm FSup maior ou igual ao FSup especificado pelo usuário. Estes conjuntos são chamados conjuntos de itens frequentes. Segundo, para cada conjunto de itens frequentes, todas as possíveis regras candidatas são geradas e testadas com respeito ao FConf. Uma regra candidata é gerada por algum subconjunto de itens de um conjunto como antecedente da regra, e os itens remanescentes do conjunto de itens como o conseqüente da regra. Apenas regras candidatas com FConf maior ou igual ao FConf especificado pelo usuário são dadas como saída do algoritmo.

O algoritmo Apriori (AGRAWAL; IMIELINSKI; SWAMI, 1993), é responsável por descobrir o conjunto de itens através de múltiplos passos na base de dados. Cada passo inicia com um conjunto semente de itens, e esse conjunto semente gerará novos conjuntos potenciais, chamados conjunto de itens candidatos. Em linhas gerais, enquanto o conjunto de itens candidatos não ficar vazio, o algoritmo armazena esses conjuntos e para cada tupla do banco de dados testa se um conjunto candidato está ou não contido na tupla (PLASTINO, 2001).

Caso um conjunto candidato esteja contido na tupla, então se incrementa um contador. Se ao final do teste para cada tupla da base de dados uma regra candidata tiver um suporte mínimo especificado, então ela é inserida num novo conjunto semente (itens candidatos). A Figura 2 ilustra, de forma genérica, como se dá o processo de geração da regras de associação

exatamente a partir do momento em que todo o conjunto de dados já foi selecionado e os fatores de Suporte e Confiança foram estabelecidos. Essa etapa nada mais é do que a implementação do método conhecido na literatura (AGRAWAL; IMIELINSKI; SWAMI, 1993) por `genRules()` :

```

procedure genrules( $l_k, a_m$ )

 $A = \{(m-1)\text{-itens } a_{m-1} \mid a_{m-1} \subset a_m\}$  // gera o conjunto A contendo os subconjuntos de  $a_m$  com  $a_{m-1}$ 
elementos
forall  $a_{m-1} \in A$  do
     $\text{conf} = \text{support}(l_k) / \text{support}(a_{m-1})$  // calcula a confiança para a regra de  $a_{m-1}$ 
    if ( $\text{conf} \geq \text{minconf}$ ) then
        output  $a_{m-1} \rightarrow (l_k - a_{m-1})$ , with confidence= $\text{conf}$  and support=  $\text{support}(l_k)$ 
        if ( $m-1 > 1$ ) then // se  $a_{m-1}$  tiver mais de um elemento, chama novamente a função
            genrules( $l_k, a_{m-1}$ ); //se  $a_{m-1}$  tiver mais de um elemento, chama novamente a função
            genrules, passando o conjunto  $l_k$  e o conjunto  $a_{m-1}$  para gerar
            para gerar novas regras com os subconjuntos de  $a_{m-1}$ .
        end
    end
end

```

Figura 2: Método para extração final das Regras de Associação - Algoritmo Apriori.

2.4.3 HÍBRIDA

Uma abordagem híbrida envolve a combinação de dois ou mais tipos de recomendação. A abordagem híbrida procura combinar os pontos fortes de cada técnica a ser utilizada, visando criar um sistema que possa melhor atender às necessidades do usuário (HERLOCKER, 2000)(ANSARI et al., 2000).

2.5. CONSIDERAÇÕES FINAIS

Neste capítulo apresentamos uma visão geral sobre os principais conceitos que envolvem os Sistemas de Recomendação. Inicialmente falamos da origem, principais características e trabalhos relacionados sobre estes sistemas nos mais diversos domínios de aplicação. Em seguida, descrevemos os conceitos relacionados à arquitetura e principais técnicas utilizadas no seu desenvolvimento, bem como mecanismos de avaliação.

3. TELEVISÃO DIGITAL INTERATIVA E O ISDB-TB

Neste capítulo são apresentados os principais conceitos e tecnologias no domínio da Televisão Digital Interativa, apresentando sua arquitetura geral e os principais modelos comerciais existentes. Também será apresentado o sistema de televisão digital nacional, conhecido por *Integrated Services Digital Broadcasting Terrestrial* (ISDB-Tb), percorrendo todo o seu histórico, desde o projeto que determinou o modelo de referência a ser adotado, outrora denominado Sistema Brasileiro de Televisão Digital (SBTVD), e comentando os principais avanços, contribuições e seu atual status de implantação.

A televisão digital é um sistema de transmissão, recepção e processamento de sinais em formato digital, que permite uma alta definição de som e imagens além de grande interatividade (FERNANDES; LEMOS; SILVEIRA, 2004)(BECKER; MONTEZ, 2004).

Hoje, a televisão digital é uma realidade, apresentando inúmeras vantagens quando comparada aos sistemas analógicos. Enquanto em um canal analógico existem dificuldades de manipulação e armazenamento do sinal, isso não ocorre no sistema digital. Essa necessidade parte do princípio de que, apesar de aparentar bastante espaço, os 6MHz destinados a uma TV analógica, são bastante apertados. Com a digitalização é possível comprimir o sinal de tal modo que se possa acrescentar outros sinais complementares, seja para garantir maior segurança na efetivação da transmissão ou na inserção de novos serviços.

3.1 TOPOLOGIA DE UM SISTEMA DE TELEVISÃO DIGITAL INTERATIVA

A topologia de um sistema de TVDI (Figura 3) envolve, de forma geral, os seguintes componentes:

- **Difusor:** Responsável por prover o conteúdo a ser transmitido e suportar as interações com os usuários, sendo composto por:
 - o **Provedor de Serviço de Difusão:** Representando as emissoras e demais provedores de conteúdo, responsáveis pelo envio de serviços, aplicações e conteúdos através do canal de difusão;

- o **Provedor de Serviço de Interação:** Processa as informações interativas, enviadas e recebidas pelos usuários finais (telespectadores), através de um canal de retorno ou interativo;
- **Meio de Difusão:** Composto por um canal de difusão e um canal de retorno (ou canal de interatividade) que habilita a comunicação entre difusor e receptor:
 - o **Canal de Difusão:** Representa o meio pelo qual os Provedores de Serviço enviam informações para os usuários, podendo ser realizado através de Satélite, Cabo ou Radiodifusão (terrestre);
 - o **Canal de Interatividade:** Denomina-se Canal de interação ou de interatividade ou de retorno, o sistema constituído da conexão entre as redes de televisão com as redes de telecomunicações, resultando em dois canais de comunicação: o canal de descida, que estabelece a comunicação no sentido da emissora para o usuário e o canal de retorno, que estabelece a comunicação no sentido usuário para as emissoras;
- **Receptor:** Também conhecido por *Set-Top-Box (STB)*, Unidade Receptora Decodificadora (URD), ou simplesmente conversor. Este dispositivo é responsável por realizar a decodificação do sinal digital recebido, verifica a autenticidade dos direitos de acesso e oferece a interatividade do telespectador com a emissora de televisão. O software responsável pela interatividade do sistema com o usuário é o *middleware*.

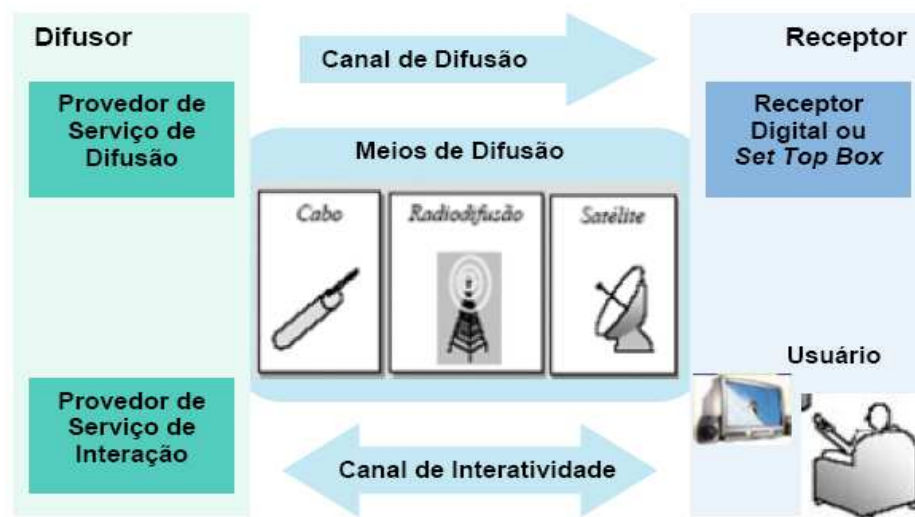


Figura 3: Topologia geral de um sistema de televisão digital (FERNANDES; LEMOS; SILVEIRA, 2004).

O modelo da topologia de um sistema de TV digital é dividido em camadas, onde cada camada oferece serviços para a camada superior, e usa os serviços oferecidos pela camada inferior. Tipicamente o modelo inclui, além da camada dos aplicativos, a camada do *middleware*, de compressão (áudio e vídeo), transporte e transmissão ou modulação.

Cada um dos padrões mundiais existentes para sistemas de televisão digital interativa adota, de forma geral, uma solução diferente para cada camada específica. Entretanto, existe um padrão comum a todos os sistemas na camada de transporte: o MPEG, atualmente utilizado na versão MPEG-2. Maiores detalhes sobre os padrões existentes e suas características tecnológicas podem ser encontrados em (FERNANDES; LEMOS; SILVEIRA, 2004)(OLIVEIRA, 2005).

3.2 **DESENVOLVIMENTO DE APLICAÇÕES**

O processo de desenvolvimento de aplicações para televisão digital é muito semelhante ao processo seguido no caso das aplicações web. A grande diferença está na distribuição e nas características dessa aplicação que, depois de prontas, são transmitidas juntamente com o sinal de televisão (FREED, 2000)(PENG; CESAR; VUORIMAA, 2001).

As aplicações interativas para TV digital, dependendo do tipo de linguagem através do qual as mesmas são desenvolvidas, podem ser classificadas em declarativas ou procedurais. As linguagens declarativas facilitam o desenvolvimento de aplicações, uma vez que não exigem um domínio das ferramentas de programação por parte dos profissionais da área. Desta forma, os custos do desenvolvimento de conteúdo podem ser reduzidos. O conteúdo declarativo é recebido via fluxo de transporte de radiodifusão e armazenado localmente no terminal de acesso, ou acessado remotamente, mediante uma solicitação, via canal de retorno do terminal de acesso.

Baseado em um ambiente de aplicações declarativo, o *middleware* declarativo dá suporte a aplicações desenvolvidas em linguagens declarativas. A linguagem HTML é dominante na Internet devido à sua simplicidade e pelo suporte a diversos recursos de fácil utilização, entretanto, como não foi projetada para ser utilizada em TV digital interativa, apresenta limitações. Outras linguagens declarativas como o SMIL, XMT-O e NCL, foram projetadas para este fim (SOARES; RODRIGUES; MORENO, 2007).

Em contrapartida, o desenvolvimento de aplicações procedurais (desenvolvidas utilizando linguagens não declarativas), requer um *middleware* procedural. Nesse tipo de *middleware*, o programador possui maior poder sobre o programa, mas requer conhecimento dos recursos de implementação (SOUZA FILHO; LEITE; BATISTA, 2007).

Entre as linguagens não declarativas mais utilizadas para esta finalidade temos JAVA TV, que consiste em uma extensão da plataforma Java com bibliotecas de componentes específicos para desenvolvimento de aplicações para televisão digital (JAVATV,

2008)(PENG; CESAR; VUORIMAA, 2001)(FREED, 2000)(CORNEL; HORSTMANN, 1999).

O Java TV consiste em uma máquina virtual Java (JVM - *Java Virtual Machine*) hospedada e executada no STB. Provê, em sua API, recursos de uso geral, sempre em um alto nível de abstração, tornando possível o desenvolvimento de aplicações sem conhecimento específico das camadas de baixo nível, como o protocolo do SI (*Service Information*), os protocolos de transmissão ou mesmo a rede transmissora (TABELAS SI, 2005). Também provê meios de seleção de serviços, acesso a um banco de dados com as informações do SI, controle do *media player* interno do receptor e acesso a dados transmitidos no sinal de televisão.

3.2.1 XLET

Um Xlet é o nome dado a uma aplicação executada em um ambiente de TV Digital que pode vir previamente armazenado no STB, ou pode ser enviada via canal de difusão (JAVATV, 2008)(OLIVEIRA, 2005).

Testes de Xlets podem ser emulados através do XleTView (XLETVIEW, 2010). O ambiente simula uma tela de TV e também disponibiliza um emulador de controle remoto para simular as interações do usuário. Apesar de existirem alguns pré-requisitos para o ambiente executar com sucesso, a instalação do ambiente é considerada simples (PENG; CESAR; VUORIMAA, 2001)(FREED, 2000). Porém, apesar de suas facilidades e das diversas aplicações suportadas pelo XleTView, o ambiente ainda não oferece as ferramentas suficientes para a realização de testes de aplicações com acesso a dados em fluxo de transporte.

3.3 COMPONENTES / APLICATIVOS PARA TV DIGITAL

A seguir são apresentados sucintamente alguns dos aplicativos desenvolvidos para o escopo da televisão digital, bem como alguns componentes cuja conceituação é importante para o entendimento deste trabalho.

3.3.1 GERADOR DE CARROSSEL

No caso da interatividade via televisão digital, além da transmissão de áudio e vídeo, também temos o envio de fluxos de dados, que é viabilizado através do gerador de carrossel. O carrossel é um dos mecanismos utilizados para o encapsulamento e a difusão de dados dentro de um fluxo de transporte, junto com os outros fluxos elementares de áudio e vídeo. Cabe ao gerador de carrossel gerar continuamente estes fluxos de dados.

A implementação do carrossel de dados em fluxos de transporte MPEG-2 está diretamente relacionada à especificação *Digital Storage Media-Command and Control* (DSM-CC), que define uma série de protocolos e fornece funções e operações de controle para gerenciar os fluxos MPEG (FERNANDES; LEMOS; SILVEIRA, 2004)(BECKER; MONTEZ, 2004).

Os elementos do carrossel DSM-CC estão contidos em mensagens, conhecidas como DSM-CC *messages*, que, por sua vez, estão divididas em duas categorias. A primeira é a categoria DSM-CC *Download Data Messages*. Corresponde a uma mensagem, conhecida como *DownloadDataBlock*, que contém os dados pertencentes aos módulos do carrossel.

A segunda categoria é a DSM-CC *Download Control Messages*. É uma mensagem, conhecida por *DownloadInfoIndication* (DII), que informa ao receptor como os blocos estão agrupados nos módulos.

3.3.2 TABELAS SI

Juntamente com os sinais de áudio e vídeo, também são transmitidos dados sobre serviço contendo informações detalhadas dos fluxos de áudio, vídeo e dados. Estas informações são transmitidas no formato de tabela, dentro do STB, denominadas *Service Information* ou, simplesmente, SI. As tabelas SI (TABELAS SI, 2005) foram criadas a partir da necessidade de diferenciação de informações no contexto do novo sistema digital. Elas carregam informações que ajudam e guiam o receptor na localização e interpretação dos dados recebidos.

As tabelas de informação de serviço (SI) foram primeiramente pensadas para o sistema de codificação MPEG (*Motion Picture Experts Groups*), para diferenciar as informações de áudio e vídeo dentro de uma *stream* de dados. Essas tabelas orientam o receptor durante toda a recuperação dos dados recebidos, até que a informação original seja obtida.

As tabelas trabalham em conjunto, umas apontando para as outras, a fim de identificar e denominar cada serviço prestado pela rede. Fica a cargo de cada modelo especificar tabelas que indiquem seus próprios serviços. Para isso cada sistema desenvolveu tabelas específicas inerentes ao seu modelo. A Figura 4 ilustra as principais tabelas com uma breve resenha das informações contidas. No caso do modelo brasileiro, em (ABNT NBR 15603-1, 2008)(ABNT NBR 15603-2, 2008)(ABNT NBR 15603-3, 2008), é apresentada uma especificação completa das tabelas SI do ISDB-Tb.

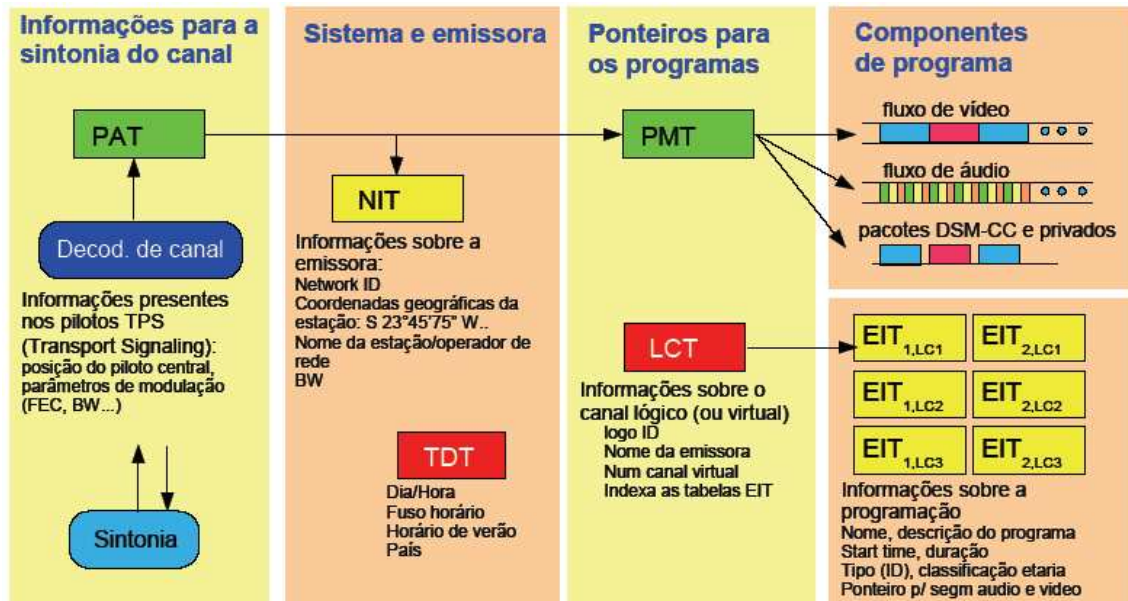


Figura 4: Ciclo de vida de um Xlet (TABELAS SI, 2005).

Segue explicação sobre as tabelas presentes:

- **PAT (Program Association Table):** É a primeira tabela a ser identificada pelo demultiplexador, dentro do STB. Sua principal função é indicar o número de programa e o PID (Program IDentification) das tabelas NIT e PMT;
- **NIT (Network Information Table):** Descreve as informações sobre a rede de transmissão, ou seja, sobre a emissora, tais como rede de transmissão, operadora, dentre outras;
- **TDT (Time and Data Table):** Contem informações sobre data, hora, fuso horário, dentre outros, da emissora em questão;
- **LCT (Logical Channel Table):** Descreve as informações sobre o canal lógico em que os programas são transmitidos;
- **PMT (Program Map Table):** Descreve a composição do serviço, detalhado a lista de todas as *streams* e informando ao receptor que tipo de *stream* está sendo transmitido, assim como o PID de cada elemento de informação que compõe o programa;
- **EIT (Event Information Table):** Descreve um grande número de informações sobre um evento, ou seja, sobre um programa específico. Estas informações são do tipo: nome, descrição, duração, hora início, hora fim, dentre outros.

3.3.3 EPGs

A partir das tabelas SI presentes na base de dados do STB é possível receber dados sobre as informações oriundas do envio *broadcast* via DSM-CC. Uma vez que estes dados transmitem informações sobre outros, podemos falar da existência de *metadados de*

programação. Sem eles, qualquer padrão de televisão digital não seria capaz de mudar o paradigma das TVs tradicionais.

Estes metadados permitem a construção de aplicações chamadas Guia Eletrônico de Programação (EPGs - do inglês, *electronic program guide*). EPG's são aplicações baseadas em um arquivo (metadados) que armazena informações pertinentes a uma determinada mídia ou grupo dessas, provendo a listagem de toda a programação e conteúdo de serviço interativo no televisor digital habilitado para isso. Através destes guias, torna-se possível aos usuários do sistema realizar pesquisas, por exemplo, por ator/atriz, título, gênero, faixa etária, nacionalidade, dentre outros.

3.4 O SISTEMA BRASILEIRO DE TELEVISÃO DIGITAL

O sistema brasileiro de televisão digital, atualmente denominado *Integrated Services Digital Broadcasting Terrestrial* (ISDB-Tb) aproveitou o fato da nossa rede de televisão ser a única rede de telecomunicações gratuita, com cobertura nacional, chegando a todas as camadas da sociedade.

Em se tratando de um instrumento que permite o acesso da população de baixa renda às redes de informações e serviços digitais, o governo brasileiro percebeu a grande chance de alavancar a inclusão digital no Brasil, uma vez garantindo a digitalização dos sinais de televisão, permitindo acesso de todas as camadas a um recurso de Tecnologia de Informação e Comunicação de ótima usabilidade, fácil acesso e abrangência sem igual (ANATEL, 2004).

Como um dos principais marcos do processo de implantação da televisão digital no Brasil temos a instituição do Sistema Brasileiro de Televisão Digital, através do Decreto Presidencial 4.901, de 26 de novembro de 2003 (MINICOM, 2003). A partir de então, o governo lançou 22 editais de pesquisa para o desenvolvimento de componentes que juntos formariam o SBTVD. Terminada a fase de estudos, o governo, no que se refere aos assuntos técnicos referentes à aprovação de inovações tecnológicas, especificações, desenvolvimento e implantação do ISDB-Tb, criou, em novembro de 2006, um Fórum constituído de representantes dos vários segmentos representativos do setor (BRASIL, 2006).

A partir do modelo de referência do ISDB-Tb, uma série de instrumentos visando a normalização do modelo brasileiro de televisão digital foi criada. A Associação Brasileira de Normas Técnicas (ABNT) publicou, em novembro de 2007, e homologou as Normas Brasileiras relacionadas ao padrão de transmissão de televisão digital adotado no Brasil, desenvolvidas pelo Fórum do Sistema Brasileiro de Televisão Digital, disponíveis em: (<http://www.forumsbtvd.org.br/materias.asp?id=112>).

3.4.1 ARQUITETURA DO SISTEMA BRASILEIRO DE TELEVISÃO DIGITAL

O Sistema Brasileiro de Televisão Digital (ISDB-Tb) é estruturado de acordo com a arquitetura apresentada no início deste capítulo, com características próprias que o diferenciam dos existentes (CPQD, 2006). Vale destacar que o Brasil, além de ter sido o único a escolher o MPEG-4 como padrão de compressão de áudio e vídeo, desenvolveu, também, o seu próprio *middleware*, o Ginga (GINGA, 2008). No entanto, o padrão de modulação utilizado é o mesmo (COFDM) utilizado pelo japonês. Analogamente, o padrão de transporte é o MPEG-2 utilizado por todos os demais sistemas existentes. A Figura 5 ilustra as tecnologias utilizadas pelo ISDB-Tb:

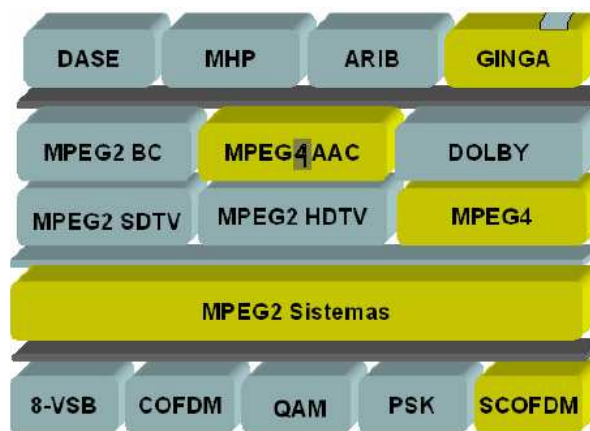


Figura 5: Arquitetura do ISDB-Tb (CPQD, 2006).

3.4.2 O MIDDLEWARE DO ISDB-TB

A função do *middleware* é a de possibilitar que as aplicações possam ser escritas de modo o mais independente possível do *hardware* e do sistema operacional presentes nos receptores digitais, permitindo a um mesmo código de aplicação ser carregado e executado em diferentes equipamentos.

Uma importante inovação tecnológica desenvolvida pelos pesquisadores brasileiros concentra-se nos *middlewares* (CPQD, 2006). A junção dos resultados de pesquisas, que originaram o procedural FlexTV, com os resultados obtidos com o declarativo Maestro, originou o *middleware* brasileiro (GINGA, 2008). Ginga é o nome do *Middleware* Aberto do Sistema Brasileiro de Televisão Digital. É constituído por um conjunto de tecnologias padronizadas e inovações brasileiras, que possibilita o desenvolvimento de aplicações interativas para o Sistema Brasileiro de Televisão Digital Terrestre (ISDB-Tb).

O *middleware* brasileiro provê suporte para ambas as modalidades desta camada de software, seja ela declarativa ou procedural, sendo a primeira através do módulo Ginga-NCL

(SOARES; RODRIGUES; MORENO, 2007) e a segunda através do módulo Ginga-J (SOUZA FILHO; LEITE; BATISTA, 2007)(GINGA-J, 2007).

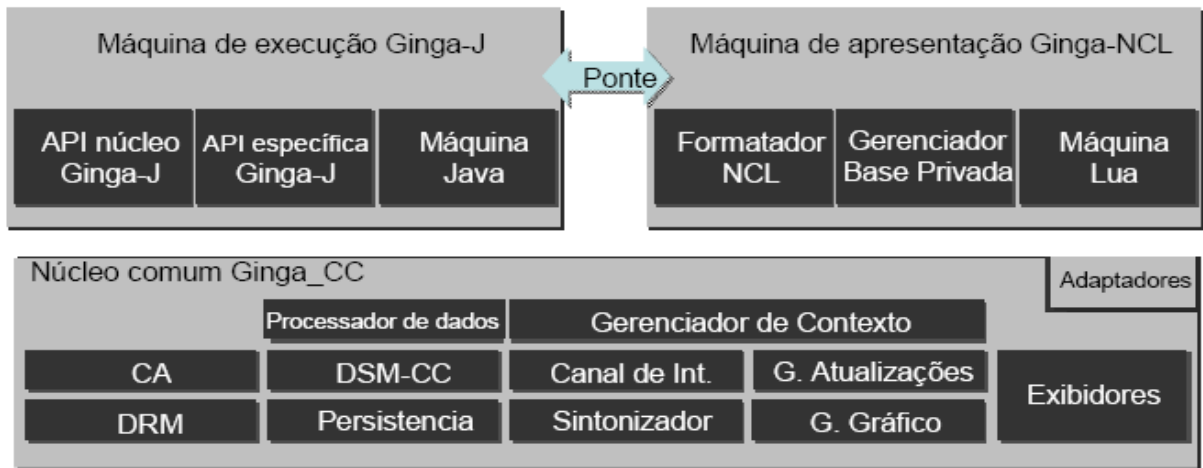


Figura 6: Arquitetura do Ginga (GINGA, 2008).

Conforme Figura 6, a arquitetura do Ginga pode ser dividida em três módulos principais:

- **Ginga-CC:** *Ginga Common-Core*. Oferece o suporte básico para os ambientes declarativo (Ginga-NCL) e procedural (Ginga-J). É o subsistema lógico provedor de todas as funcionalidades;
- **Ginga-J:** Desenvolvido pela UFPB, é o subsistema lógico do *middleware* Ginga responsável pelo processamento de aplicações imperativas escritas utilizando a linguagem Java (SOUZA FILHO; LEITE; BATISTA, 2007)(GINGA-J, 2007);
- **Ginga-NCL:** É o subsistema lógico do *middleware* Ginga responsável pelo processamento de aplicações declarativas, com facilidades para a especificação de aspectos de interatividade, sincronismo espaço-temporal de objetos de mídia, adaptabilidade e suporte a múltiplos dispositivos. NCL possui Lua como sua linguagem de script (SOARES; RODRIGUES; MORENO, 2007).

3.5 CONSIDERAÇÕES FINAIS

Neste capítulo foi apresentada uma visão geral sobre os principais conceitos relacionados com a Televisão Digital Interativa, abordando suas características, vantagens, arquitetura, dentre outros. Com base nesta conceituação, é apresentado o modelo brasileiro ISDB-Tb. As informações aqui presentes são resumidas e necessárias para o correto entendimento das propostas que ainda serão apresentadas no decorrer desta dissertação.

4. *ESTRATÉGIA HÍBRIDA DE RECOMENDAÇÃO*

Neste capítulo são abordados os conceitos e especificações relativas às contribuições principais desta dissertação. Inicialmente é descrita a estratégia de recomendação proposta utilizada como fundamento para o desenvolvimento de um sistema de recomendação para TV digital. A descrição arquitetural do sistema de recomendação é apresentada e documentada através da exibição dos diagramas e demais artefatos de modelagem convenientes, a partir dos requisitos elencados.

4.1 *ESTRUTURA DA ESTRATÉGIA DE RECOMENDAÇÃO*

Conforme mencionado na introdução e publicado em (MORAES NETO et al., 2010), o desenvolvimento do nosso sistema de recomendação levou em consideração quatro aspectos relevantes. Primeiramente, temos a possibilidade de análise das informações explícitas de preferências do usuário, disponibilizadas através da criação de seu perfil. Em segundo, temos o tratamento destas informações através do algoritmo TF-IDF. Em terceiro, temos a coleta implícita, onde, através do algoritmo Apriori, há a análise do uso do aparelho de televisão por parte do usuário. Por último, temos a possibilidade de aprimoramento do sistema, através do módulo de avaliação, permitindo o aprimoramento da avaliação através do *feedback* do usuário. É importante comentar que, além das características citadas neste parágrafo, nossa estratégia também implementou algumas boas práticas presentes em aplicações deste tipo, conforme (KIM et al., 2007).

Nas subseções a seguir, é explicado como iremos atender a cada etapa do processo de construção do sistema de recomendação proposto no presente trabalho.

4.1.1 *COLETA DE INFORMAÇÕES E MODELAGEM DE USUÁRIO*

A estratégia de recomendação prevê que a coleta de informações do usuário aconteça na forma **explícita e/ou implícita, com identificação no cliente**, e não em servidor. Dados presentes na literatura (O’SULLIVAN, 2003) informando que a eficiência de realização de

coleta implícita é igual ou superior à coleta explícita nos serviu de motivação para realizarmos a junção destas duas formas de coleta.

A junção das informações coletadas das duas maneiras permite a construção de um modelo contendo as preferências do usuário, que denominamos de *Modelo do Telespectador*, cujo formato é mostrado na Figura 7:

```
<USER>
<usuario id="1"
  nome="diego"
  profissao="Programador"
  nacionalidade="BRASILEIRO"
  estadoCivil="0"
  filhos="0"
  palavrasChave="Esportes Noticias ciências "
  gêneros=" Filmes documentários " />
</USER>
```

Figura 7: Estrutura do perfil explícito de usuários em XML.

Este arquivo, representando o Modelo do Telespectador armazena todos os perfis explícitos da aplicação, possuindo os atributos 'ID', que é único, o campo 'nome' que é utilizado para mostrar a lista de usuários que estão cadastrados no sistema, os campos 'profissão', 'nacionalidade', 'estadoCivil' e, para finalizar, as 'palavrasChave', que serão a parte principal do registro e de onde o sistema de recomendação irá extrair a maior quantidade de informações e realizar a busca por programas que terão em sua descrição alguma dessas palavras.

A título de informação, neste momento somente os atributos 'ID', 'nome', 'palavras chave' e 'generos' estão sendo usados no processo de recomendação. Espera-se que os demais atributos serão úteis no momento em que seja abordada a recomendação *por perfil*, em um possível trabalho futuro.

4.1.2 GRAU DE PERSONALIZAÇÃO

A maneira pela qual o sistema de recomendação proposto trata os usuários é a **efêmera**, uma vez que as informações sobre preferências do usuário são levadas em conta antes de se realizar a recomendação.

Além disso, como temos a possibilidade de criação de mais de um perfil de usuários, podemos afirmar que nosso sistema possui personalização **persistente e personalizada**, pois fazemos o tratamento para cada tipo de usuário.

4.1.3 APRESENTAÇÃO DA RECOMENDAÇÃO

Analisando os momentos possíveis em que o Usuário Telespectador se depara com a recomendação de programas, podemos informar que a apresentação da mesma é de dois tipos, **Pull** e **Passiva**. A primeira se deve ao fato da recomendação ocorrer somente quando solicitado, através da intenção de visualização do EPG ou da lista de programas recomendados. Diferentemente, no caso da **Passiva**, a recomendação pode ocorrer sem necessitar da solicitação do Usuário Telespectador, uma vez que o EPG possui uma funcionalidade de ‘Lembrete’ que avisa ao usuário que determinado programa terá início em 5 (cinco) minutos.

Em relação à apresentação **Push**, em princípio esta não está sendo utilizada, pois a lista de programas recomendados e o guia de programação apenas residem no terminal de acesso, não enviando nenhum tipo de informação, por exemplo, para o e-mail do telespectador.

4.1.4 DEFINIÇÃO DO PROCESSO DE RECOMENDAÇÃO

Devido às particularidades de um ambiente de televisão digital, aliada ao fato de uso de uma aplicação com as características do EPG, a estratégia híbrida de recomendação utilizada é bastante peculiar em relação às demais presentes na literatura sobre sistemas de recomendação (VOZALIS; MARGARITIS, 2003).

Inicialmente, fica claro que não utilizamos o conceito de **predição**, uma vez que não há a informação de índices ou de valores que classificam determinado programa e sim, o que existe é a **recomendação** propriamente dita. No nosso caso, utilizamos uma lista de programas exibida ao usuário, através do Guia de Programação. Esta estratégia denomina-se **Listas de recomendação**. A diferença é que todos os programas são exibidos, até os que não foram considerados recomendados. Sobre isso reforçamos a idéia de que o processo de recomendação leva em consideração os atributos dos programas, e não os programas em si.

Além disso, nossa proposta também prevê similaridade com o tipo **Avaliação de usuários**, uma vez que, devido à funcionalidade que permite ao Usuário Telespectador realizar a avaliação da recomendação, e como a aplicação também cria um perfil ‘Geral’, as diferentes opiniões de preferências de usuários também são contempladas na nossa proposta.

Considerando a natureza do sistema de recomendação e as características da sua plataforma, a abordagem híbrida proposta para a recomendação envolve a utilização de um conjunto de técnicas, detalhadas a seguir.

A Filtragem baseada em conteúdo (BELVIN; CROFT, 1992) tem a finalidade de recomendar itens de acordo com o registro das preferências do usuário que é realizado pelo armazenamento de todas as preferências de programas. Para que esta filtragem ocorra, fizemos uso do TF-IDF (ADOMAVICIUS; TUZHLIN, 2005), algoritmo de mineração de dados bastante difundido e que implementa a técnica de FBC.

Faz parte das funcionalidades do sistema realizar a análise de conteúdo envolvendo a informação sobre os metadados de programação presentes nas tabelas SI em comparação com as informações presentes no Modelo do Telespectador. Em seguida, é realizada a verificação destas informações e, após isso, gerada uma lista contendo todos os programas recomendados.

Ainda relacionada às técnicas de recomendação, temos que a descoberta de conhecimento em bases de dados consiste na identificação de padrões não triviais, válidos, novos e potencialmente úteis, utilizando o processo de mineração de dados para a descoberta automática de novos relacionamentos. Regras de Associação (ADOMAVICIUS; TUZHLIN, 2005) são utilizadas neste contexto para auxiliar na descoberta de padrões de forma automática. Assim, os STBs dos usuários telespectadores irão processar o algoritmo Apriori, onde, a partir dos atributos de toda a programação assistida, o algoritmo irá extrair as Regras de Associação, ou seja, as descobertas implícitas acerca das preferências dos usuários diante da programação exibida na televisão digital.

Uma grande vantagem da utilização de Regras de Associação em um processo de recomendação é o fato de que as mesmas devem ser utilizadas não para o usuário a partir do qual as mesmas foram geradas, e sim para todos os demais cadastrados no sistema. Relacionado a isso é importante salientar que estas opiniões e o impacto disso no sistema deve-se ao fato de que as Regras de Associação geradas são utilizadas em todo o escopo da aplicação. Ou seja, as regras geradas através da coleta implícita de um usuário são utilizadas no processo de recomendação para todos os demais usuários do sistema. Sobre o período de atualização destas regras, a princípio pode ser estabelecido o intervalo de uma semana, podendo ser facilmente adequado de acordo às necessidades.

4.1.5 AVALIAÇÃO DE SISTEMAS DE RECOMENDAÇÃO

A qualidade do sistema de recomendação desenvolvido precisa ser avaliada no intuito de verificar se seu funcionamento é adequado. Neste caso, a avaliação envolve, principalmente, determinar o grau de precisão das recomendações geradas. No intuito de tornar possível validar a acurácia do sistema desenvolvido, passa a ser necessário escolher métricas que possibilitem encontrar indicadores para esta avaliação.

O processo de avaliação proposto neste trabalho utiliza a técnica de Precisão, a qual estabelece a relação entre o número de itens relevantes sobre o total de itens retornados. Pode ser vista também como a relação entre o número de itens que o usuário considera relevante e o número de itens recomendados.

4.2 REQUISITOS FUNCIONAIS DO SISTEMA DE RECOMENDAÇÃO

Seguem os quatro principais requisitos do sistema de recomendação utilizado para a implementação da proposta apresentada. O objetivo aqui não é realizar a especificação funcional dos requisitos, uma vez que isso já acontece no Apêndice A. Apenas estamos apresentando, no intuito de facilitar o entendimento da proposta.

4.2.1 DETERMINAR MODELO DO TELESPECTADOR (CRIAR PERFIL)

Este requisito constitui-se como premissa para que a recomendação ocorra. Através da coleta de informações, implícita e/ou explícita, um modelo contendo as preferências e hábitos do Telespectador é construído e armazenado no terminal de acesso.

Este modelo possui informações presentes em duas bases de dados. A primeira contem informações estáticas, ao passo que a segunda contem informações dinâmicas. O preenchimento das informações estáticas é feito pelo usuário, através da criação de seu perfil. Para isto, o usuário fornece informações gerais, como Nome, Data de Nascimento, Localização geográfica, Estado civil, dentre outros. Para permitir a construção do estudo de caso que apresentaremos no Capítulo 5, caberá ao usuário também fornecer informações específicas sobre filmes.

O fornecimento das informações solicitadas é de caráter opcional. A presença das informações fará com que o módulo responsável por fazer a recomendação possa utilizá-las como parâmetros de entrada.

Em relação aos dados dinâmicos, as informações a serem armazenadas correspondem às características dos programas mais assistidos pelo telespectador. Estas informações são extraídas pelo sistema através do uso do aparelho de TV por parte do usuário. Toda a programação assistida pelo mesmo é armazenada e todas as informações sobre estes filmes também são devidamente armazenadas para posterior extração das Regras de Associação.

Desta forma, para a criação do Modelo do Telespectador, a cada semana o sistema analisa toda a programação assistida e extrai todas as Regras de Associação possíveis. Além disso, o sistema analisa todas as informações disponibilizadas pelo usuário (quando for o caso) nas telas de criação de seu perfil e, em seguida, analisa toda a programação para saber

quais outros filmes deverão fazer parte da recomendação. Uma vez realizado estes dois processos, temos a criação do Modelo do Telespectador.

Mais uma vez torna-se necessário lembrar que a recomendação levará em conta as informações tanto estáticas como dinâmicas. Entretanto, a primeira a ser analisada é a de dados dinâmicos, tendo as informações de dados estáticos como mandatória em relação a esta primeira. Maiores detalhes serão apresentados na especificação dos casos de uso no Apêndice A. Além disso, uma das informações a serem fornecidas pelos usuários é sobre o desejo ou não de receber recomendações. Além disso, o processo de avaliação da recomendação também é responsável por alterar as informações presentes no Modelo do Telespectador.

4.2.2 CRIAR PERFIL ‘GERAL’

Bastante similar ao requisito anterior, em sua essência, este requisito permite que a aplicação realize recomendações de informação de programação sem necessidade de uma solicitação explícita de criação do perfil por parte do usuário (coleta de dados explícita), gerando o chamado Modelo Geral do Telespectador.

Neste caso, toda vez que o terminal de acesso for ligado, o sistema armazenará, durante todas as horas de todos os dias da semana, as informações sobre a programação assistida naquele momento. Estas informações armazenadas serão posteriormente tratadas de forma similar aos casos em que o usuário cria seu próprio perfil. A diferença aqui é que não existe a necessidade de preenchimento de informações por parte dos usuários para que a recomendação ocorra. Somente informações implícitas serão tratadas pelo processo de recomendação. Este requisito permite que seja realizada uma recomendação única para todos os usuários, uma vez que sabemos que, em uma grande quantidade de lares, o aparelho de televisão é compartilhado por várias pessoas e estas, geralmente, assistem determinados tipos de programas dependendo do horário e/ou do dia em questão.

4.2.3 RECOMENDAR PROGRAMAÇÃO

Este requisito é o principal desta proposta, uma vez que consiste em indicar para o usuário telespectador qual o programa mais recomendado, conforme informações presentes no Modelo do Telespectador, a ser assistido em um dado momento.

O processo de recomendação dar-se-á através do Guia de Programação Eletrônico. Todo programa considerado ‘recomendado’ ao telespectador, aparecerá no referido guia de forma destacada, sinalizando a sua preferência por programas com tais características. A cada semana, o Modelo do Telespectador é atualizado. A visualização da recomendação poderá ser

Em seguida, breve descrição acerca de cada um dos componentes presentes na arquitetura:

- **Base de dados estáticos:** Base de dados em XML (*eXtensible Markup Language*) responsável por armazenar os dados fornecidos pelo Telespectador, quando da criação de seu perfil, através do preenchimento de informações em telas que serão mostradas adiante;
- **Base de dados dinâmicos:** Base de dados em XML com as regras de associação geradas pelo algoritmo Apriori. Outro documento XML, que armazena a programação assistida pelo usuário, funciona como parâmetro de entrada para o Apriori que, após analisar todo este arquivo, descobre relacionamentos implícitos relacionadas ao domínio da aplicação televisiva. Utilizamos os valores¹ de 90% (noventa por cento) para o Fator Suporte (frequência com que ocorrem os padrões em toda a base) e 70% (setenta por cento) para o Fator Confiança (medida da força das regras – quantidade das regras encontradas em relação ao total);
- **Módulo Minerador:** Componente responsável por analisar a interação entre o aparelho de televisão e o Usuário Telespectador, cujo resultado é um documento XML contendo as regras de associação armazenadas na Base de Dados dinâmicos. Estas regras são geradas para cada perfil do sistema, entretanto, o resultado das mesmas permite que a recomendação seja realizada para todos os perfis cadastrados;
- **Módulo Gerador de Perfis:** Este componente analisa as informações presentes nas bases de dados estáticos e dinâmicos, validando as informações que estão na dinâmica de acordo com as da base estática, cuja prevalência é mandatória quando comparada com as informações presentes na base dinâmica. Funcionando como um analisador (*parser*) XML, informações presentes nas bases estática e dinâmica são processadas e o resultado armazenado na base a seguir descrita;
- **Base de dados do Modelo do Telespectador:** Base de dados em XML responsável por armazenar o perfil do Usuário Telespectador;
- **Base de dados de programação (Tabelas SI):** Base de dados, também em XML, que recebe as informações de programação, através das tabelas SI, armazenando no formato XML para posterior alimentação do EPG. É importante ressaltar que, por mais que o objetivo deste trabalho não seja a total aderência ao modelo

¹ Valores de referencia extraídos da literatura (AGRAWAL; IMIELINSKI; SWAMI, 1993).

brasileiro de televisão digital (DTV, 2009), a estrutura das tabelas SI utilizadas já foi aderente ao padrão ABNT NBR 15603 (ABNT NBR 15603-1, 2008). Inclusive, para a própria geração dos dados de programação assistida, desenvolvemos aplicativo JAVA que gera várias relações de programas seguindo a estrutura das tabelas SI;

- **EPG:** Componente responsável pela geração do Guia de Programação através do qual a recomendação será exibida. O EPG possui os programas, considerados favoritos, marcados e disponíveis para visualização por parte dos usuários;
- **Módulo Recomendador:** Componente responsável por analisar todas as informações de programação, presentes na respectiva Base de dados e, através da comparação com as informações presentes na Base de dados do Modelo do Telespectador, realizar a recomendação dos programas considerados favoritos. Uma vez gerada a lista de programas recomendados, este componente realiza a marcação dos programas considerados recomendados;

Os elementos de ligação entre os componentes da arquitetura da aplicação foram numerados no intuito de mostrar os passos envolvidos para a recomendação em um cenário geral, ainda em nível arquitetural. Entretanto, vale ressaltar que não existe a obrigatoriedade de aderência à sequência apresentada para que a solução arquitetural proposta seja viável.

4.4 DIAGRAMAS UML

Seguem os principais diagramas no intuito de documentar a abordagem seguida para o desenvolvimento da aplicação.

4.4.1 DIAGRAMA DE CLASSES

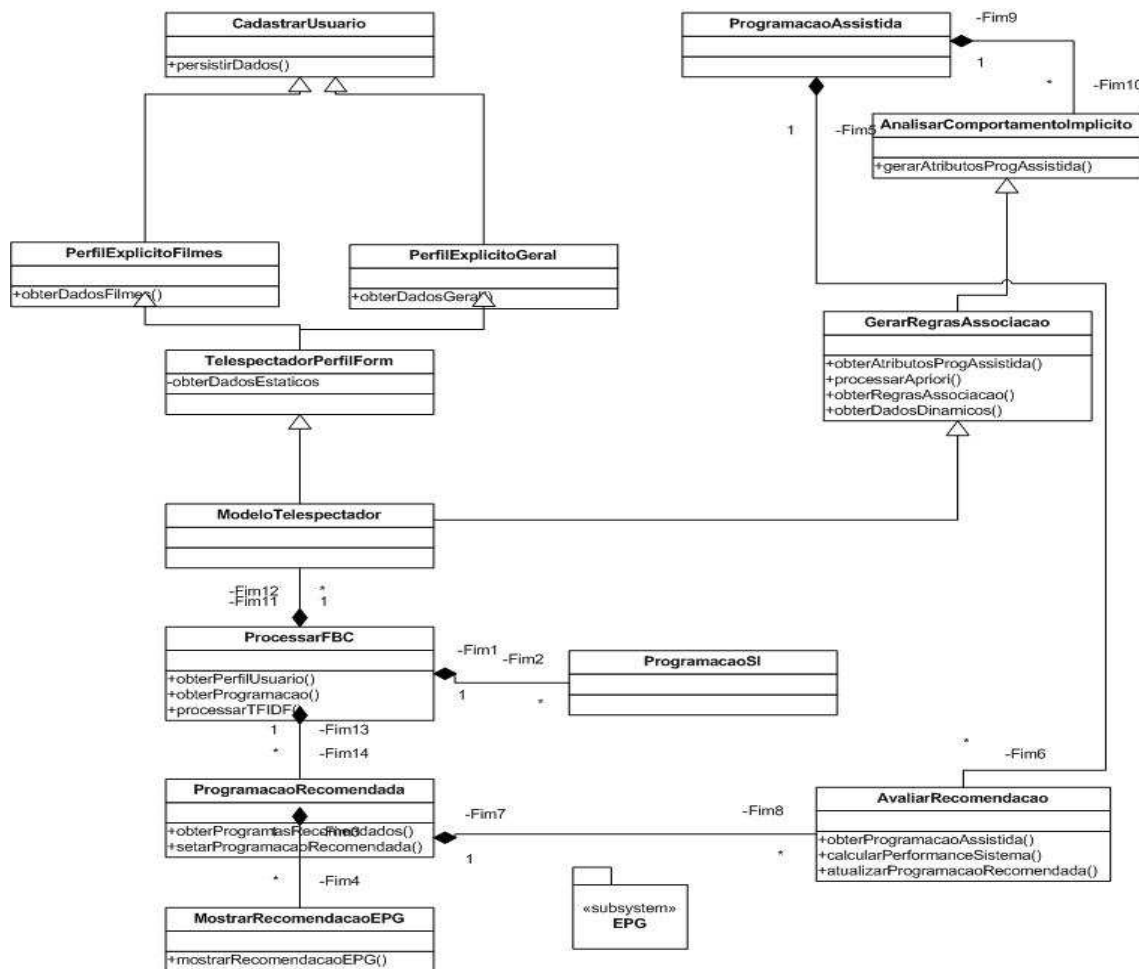


Figura 9: Diagrama de Classes da aplicação.

A seguir, breve descrição das principais finalidades de cada entidade mostrada no diagrama da Figura 9, organizadas de acordo com os componentes arquiteturais:

Base de dados estáticos:

- **CadastrarUsuario:** Esta classe contém as funcionalidades gerais para manter e consultar as informações estáticas do perfil de um usuário, através do uso de parserXML, em documento no formato .XML;
- **PerfilExplicitoGeral:** Esta classe trata, exclusivamente, das informações gerais de perfil do usuário, disponível nas telas de cadastro de perfil (a ser detalhado no próximo capítulo). Através do método `obterDadosGerais()`, é possível que, a qualquer momento, sejam recuperadas as informações gerais do perfil de um determinado usuário;
- **PerfilExplicitoFilmes:** Trata, exclusivamente, das informações específicas de filmes, disponível nas telas de cadastro de perfil (a ser detalhado no próximo capítulo).

Através do método `obterDadosFilmes()`, é possível que, a qualquer momento, sejam recuperadas as informações do perfil de um determinado usuário;

Módulo do Modelo do Telespectador:

- **TelespectadorPerfilForm:** Classe responsável pela consolidação de todas as informações estáticas do perfil dos usuários e disponibilizando método (`obterDadosEstaticos()`) para recuperação destas informações;
- **ModeloTelespectador:** Esta classe consolida as informações estáticas, obtidas através do método `obterDadosEstaticos()`, da classe `TelespectadorPerfilForm` e dinâmicas, obtidas através do método `obterDadosDinamicos()`, da classe `GerarRegrasAssociacao`, gerando assim, o Modelo do Telespectador;

Módulo Gerador de Perfis:

- **GerarRegrasAssociacao:** Esta classe é responsável pela implementação do algoritmo Apriori, obtendo todas as Regras de Associação e, conseqüentemente, os dados dinâmicos do perfil de um determinado usuário;

Módulo Minerador:

- **AnalisarComportamentoImplicito:** Esta classe contém e trata as informações relacionadas ao comportamento do usuário no que diz respeito aos programas assistidos pelo mesmo. Através do método `gerarAtributosProgAssistida()`, é disponibilizada toda a programação assistida pelo usuário e seus respectivos atributos;

Módulo Recomendador:

- **ProcessarFBC:** Utilizando, como parâmetros de entrada, as informações sobre o Modelo do Telespectador e toda a programação extraída através das tabelas SI, esta classe é responsável por processar o algoritmo TF-IDF, iniciando o processo de geração da recomendação de programação;
- **ProgramaçãoRecomendada:** Armazena a saída gerada pela classe `ProcessarFBC`, ou seja, toda a programação recomendada para um determinado telespectador;
- **MostrarRecomendaçãoEPG:** Uma vez conhecida qual a programação recomendada, esta classe realiza a marcação na aplicação EPG, dos programas a serem recomendados, dependendo do perfil de usuário que esteja solicitando visualização e uso do Guia de Programação;
- **AvaliarRecomendação:** Uma vez conhecida a programação assistida e logo após o usuário preencher as informações sobre sua avaliação da recomendação realizada, o

sistema calcula o percentual de desempenho do sistema de recomendação. Além disso, após o usuário, através da funcionalidade de inclusão de restrições, a ser detalhada no próximo capítulo, ter informado a programação da qual ele não quer mais receber recomendação, o método `atualizarProgramacaoRecomendada()` se encarregará de efetuar as devidas atualizações na classe `ProgramacaoRecomendada`;

Base de dados dinâmicos:

- **ProgramacaoAssistida:** Classe que analisa o comportamento do usuário de frente do aparelho de TV, armazenando toda a programação assistida pelos usuário, bem como os atributos desta programação;

Base de dados de programação:

- **ProgramacaoSI:** Através de `parserXML`, extrai as informações sobre a programação, disponíveis nas tabelas SI, tornando-as disponíveis em documento XML e classe JAVA;

4.4.2 *DIAGRAMAS DE SEQUÊNCIA*

Nesta seção procuramos colocar os principais diagramas de sequência que mostram as trocas de informações existentes para a execução do processo de construção dos perfis de usuários, construção do Modelo do Telespectador, geração da recomendação de programação, avaliação da recomendação e atualização da programação recomendada.

No diagrama da Figura 10, temos a representação da coleta das informações explícitas do usuário. Estas informações são coletadas através de duas telas de cadastro (que serão mostradas no próximo capítulo), onde o usuário preenche suas informações gerais e suas preferências relacionadas a filmes (categoria de programação escolhida para o nosso estudo de caso). Os seguintes passos são executados:

- Usuário, ao solicitar a inclusão de um novo perfil, preenche suas informações gerais e sobre filmes;
- Através do método `obterDadosGeral()`, da classe `PerfilExplicitoGeral`, os dados gerais são persistidos em formato .xml;
- Através do método `obterDadosFilmes()`, da classe `PerfilExplicitoFilmes`, os dados sobre filmes são persistidos em formato .xml;

- A classe TelespectadorPerfilForm, através do método obterDadosEstaticos(), consolida as informações estáticas do perfil do usuário, persistindo-as em formato .xml.

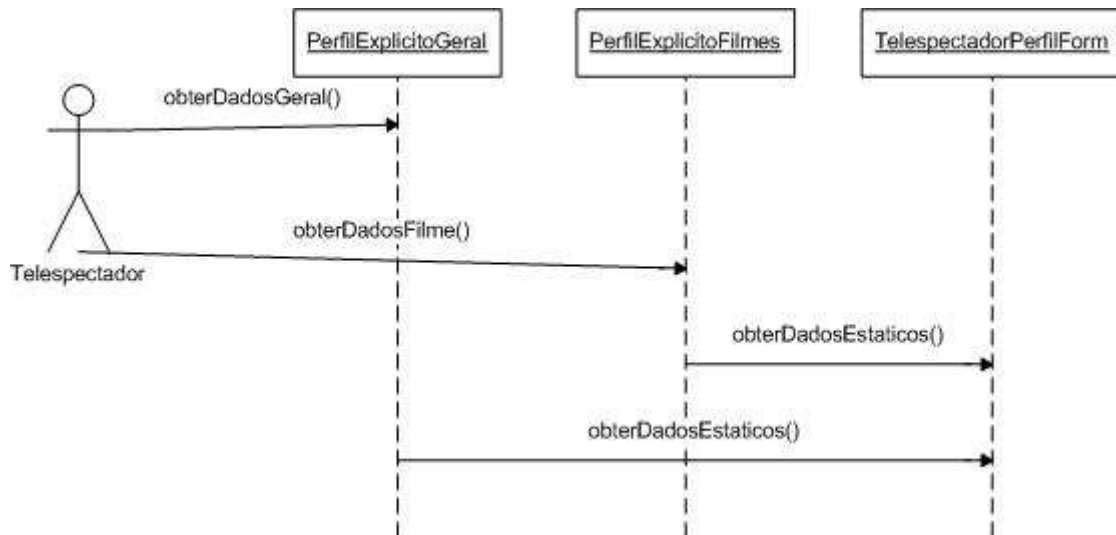


Figura 10: Diagrama de Sequência – Coleta explícita de dados.

Na Figura 11, temos a representação da coleta implícita de dados, onde, a partir da interação do usuário com a programação, seu comportamento é analisado para, posteriormente, o conhecimento implícito ser extraído, da seguinte forma:

- Usuário interage com a programação televisiva e esta interação é armazenada na classe ProgramacaoAssistida;
- A classe AnalisarComportamentoImplicito recebe de ProgramacaoAssistida, através do método gerarAtributosProgAssistida(), as informações dos atributos da programação assistida pelo usuário em questão;
- A classe GerarRegrasAssociacao utiliza estas informações como parâmetros de entrada para o método processarApriori(), que será responsável pela extração das Regras de Associação, conforme será demonstrado com mais detalhes no próximo capítulo;
- Ainda nesta mesma classe, o método obterRegrasAssociacao() disponibiliza as informações implícitas do perfil do usuário.

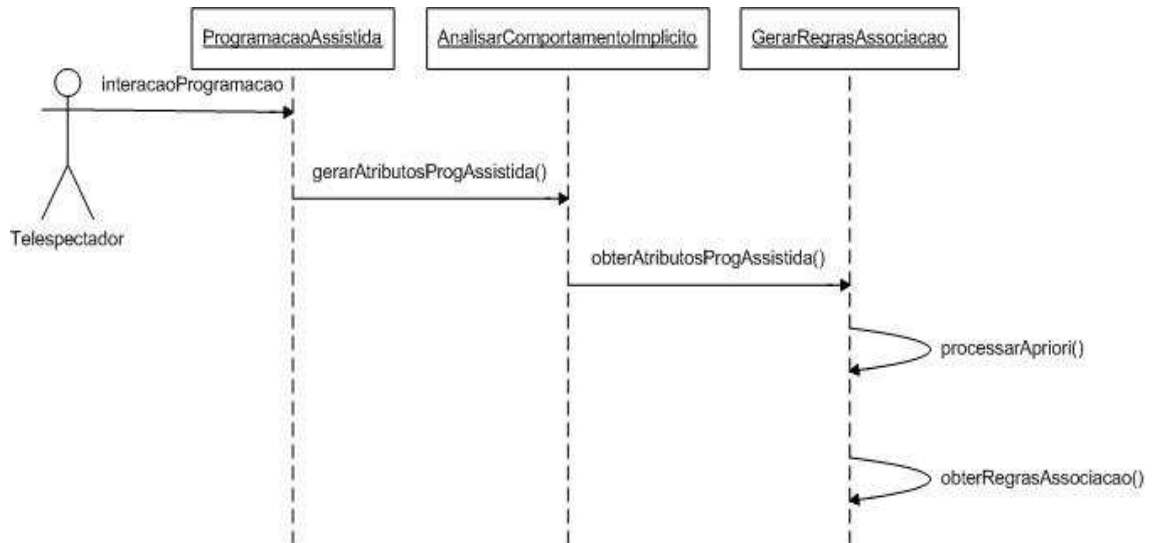


Figura 11: Diagrama de Sequência – Coleta implícita de dados.

Em seguida temos a representação da consolidação do Modelo do Telespectador, onde as classes TelespectadorPerfilForm e GerarRegrasAssociacao fornecem, através dos métodos `obterDadosEstaticos()` e `obterDadosDinamicos()`, estas referidas informações para a classe ModeloTelespectador, para posterior persistência em formato .xml.

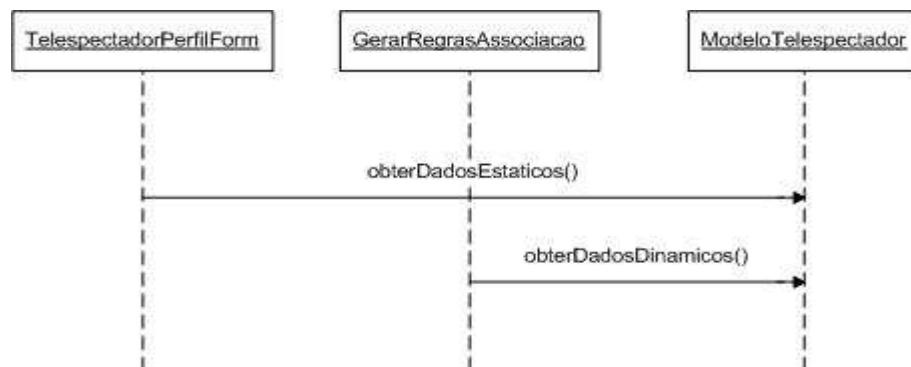


Figura 12: Diagrama de Sequência - Geração do Modelo do Telespectador.

O diagrama de sequência da Figura 13 certamente representa a principal etapa ocorrida no escopo desta aplicação para que a recomendação ocorra. Os seguintes passos são executados:

- A classe ProcessarFBC coleta todas as informações do perfil do usuário, inclusive as Regras de Associação relacionadas, através do método `obterPerfilUsuario()`, da classe ModeloTelespectador;

- ProcessarFBC, conhecendo as preferências do usuário, utiliza o método `obterProgramação()` no intuito de saber a programação que está disponível, informação esta disponibilizada através das tabelas SI, para, em seguida;
- Através do método `processarTFIDF()`, verificar a programação disponibilizada e saber quais se encaixa no perfil do usuário, o que, finalmente, resultará na;
- Geração da lista de programação recomendada, através do método `obterProgramasRecomendados()`, da classe `ProgramacaoRecomendada`.

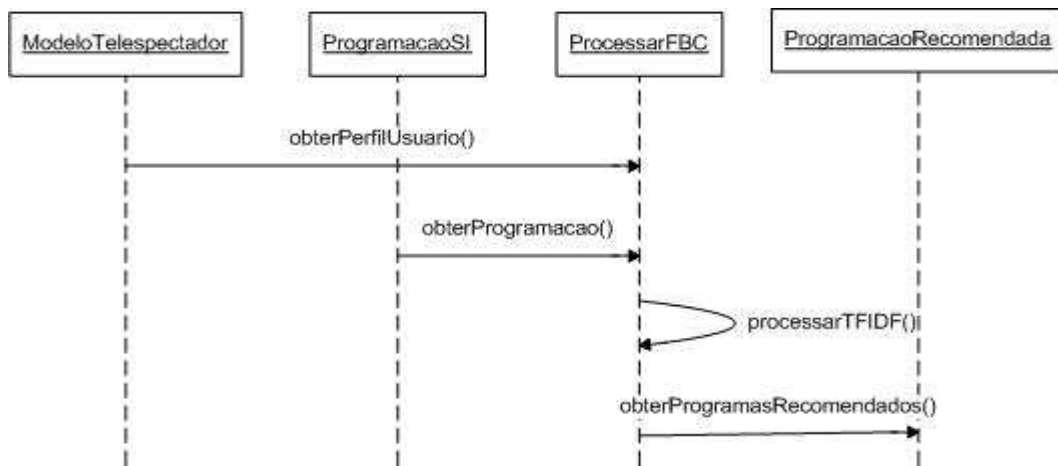


Figura 13: Diagrama de Sequência - Geração da recomendação de programação.

Uma vez que a recomendação da programação é gerada, a qualquer momento que o usuário solicitar a visualização do EPG, este será exibido, conforme será demonstrado no próximo capítulo, destacando a programação recomendada. Isso ocorre após os seguintes passos:

- A relação dos programas recomendados é disponibilizada através do método `obterProgramasRecomendados()`, da classe `ProgramacaoRecomendada`;
- Em seguida, os programas a serem exibidos no EPG possuem seus flags que tratam da recomendação, setados com o valor '1', indicando que devem ser mostrados como recomendados. Isso ocorre através do método `setarProgramacaoRecomendada()`, da classe `MostrarRecomendacaoEPG` que, por sua vez;
- Através do método `mostrarRecomendacaoEPG()`, realiza o destaque da recomendação nesta última aplicação.

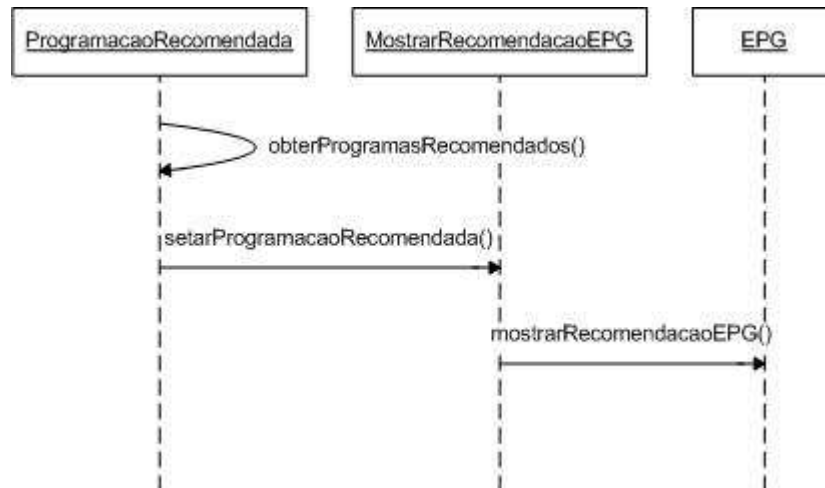


Figura 14: Diagrama de Sequência – Exibição da programação recomendada.

O usuário dispõe da possibilidade de avaliar a recomendação realizada pelo sistema e, com isso, permitir rejeitar algumas recomendações feitas pelo sistema (restrições de recomendação), dando a este a possibilidade de melhoria contínua do seu processo de validação. Estas funcionalidades ocorrem da seguinte forma, conforme ilustrado na Figura 15:

- Usuário, após solicitar a funcionalidade de Avaliação da Recomendação, através das telas disponibilizadas para este fim, o mesmo informa se aprovou ou não a recomendação realizada pelo sistema;
- O sistema, por sua vez, de posse desta informação e da informação da programação assistida pelo usuário, obtida através do método `obterProgramacaoAssistida()`, classe `ProgramacaoAssistida`, calcula o desempenho do sistema através do método `calcularPerformanceSistema()`, da classe `AvaliarRecomendacao`;
- Por último, o método `atualizarProgramacaoRecomendada()` atualiza, de acordo com as restrições inseridas pelo usuário (*feedback* do usuário), os filmes indicados presentes na classe `ProgramacaoRecomendada`.

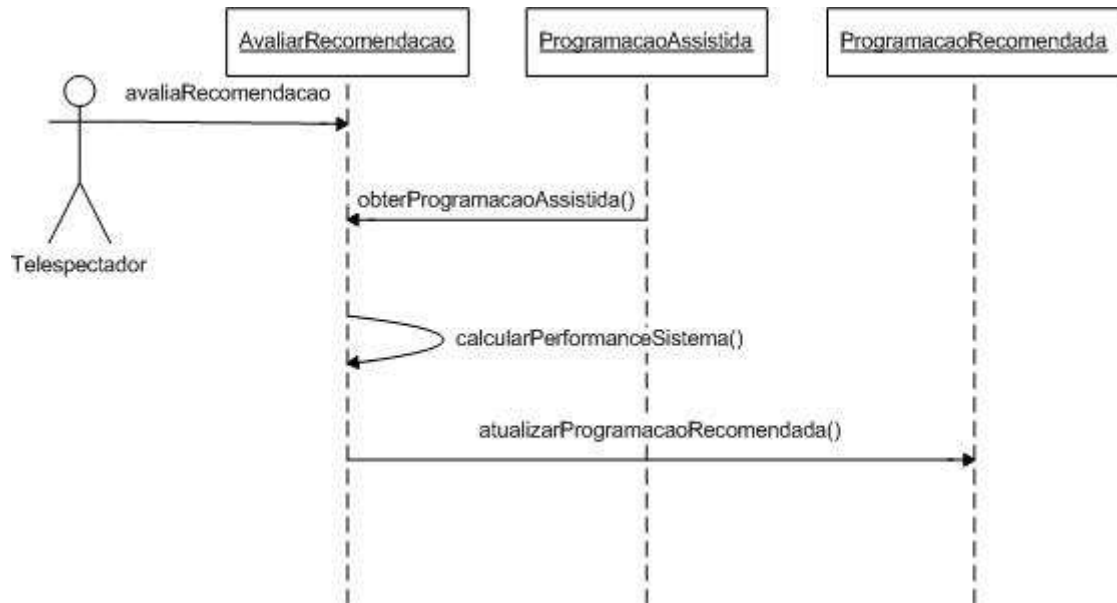


Figura 15: Diagrama de Sequência – Avaliação da recomendação.

4.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os requisitos do sistema desenvolvido, assim como a documentação necessária para um maior aprofundamento técnico sobre os aspectos de nossa proposta. Em seguida mostraremos os detalhes de implementação, telas da aplicação e estudo de caso, no intuito de detalhar totalmente, em nível técnico, como se deu a construção e validação de nossa proposta.

5. IMPLEMENTAÇÃO E RESULTADOS

Uma vez exposta a arquitetura do sistema, agora vamos demonstrar a implementação realizada, exibindo códigos e telas, bem como a avaliação da eficiência da estratégia híbrida de recomendação proposta nesta dissertação, utilizando a forma de prova de conceito.

5.1 AMBIENTE DE DESENVOLVIMENTO

Um aspecto muito importante a ser levado em consideração para o desenvolvimento de um sistema computacional é a escolha pelas ferramentas e demais elementos de infraestrutura que determinam a plataforma na qual o sistema será desenvolvido e implantado. As tecnologias utilizadas para o desenvolvimento dos módulos arquiteturais em questão são apresentada a seguir:

Tabela 2: Quadro contendo informações sobre a plataforma de desenvolvimento.

Nome da ferramenta	Tipo	Finalidade na proposta
Java Sun SDK (JAVA, 2010)	Linguagem e compilador Java	Linguagem Java; Compilador Java.
Eclipse (ECLIPSE, 2010)	IDE de desenvolvimento Java	Criação dos Xlets; Criação dos demais códigos Java da aplicação.
Notepad XML	Editor XML	Criação, edição e manipulação dos documentos XML.
NanoXML (NANOXML, 2010)	Parser SAX XML	Criação de documentos XML a partir de código Java; Manipulação de documentos XML a partir de código Java.
XleTView (XLETVIEW, 2010)	Emulador de Xlets	Testes dos Xlets desenvolvidos, simulando o ambiente de televisão digital interativa;
Gerador de Programação (de autoria)	Aplicativo Java	Geração dos dados de programação para realização das simulações.

5.2 A ESTRUTURA DE BANCO DE DADOS UTILIZADA

O uso de XML nas bases de dados do sistema deu-se devido à sua larga utilização como recurso para armazenamento de dados (GRAVES, 2003).

Seguindo o padrão XML e sua estrutura, os dados utilizados e manipulados em nossa aplicação são persistidos em formato XML, tendo sua respectiva correlação em uma classe Java. A manipulação das informações em formato .xml e sua correlação com as classes Java foi feita através da API nanoXML. As principais classes que possuem sua correlação em documentos XML são: ProgramacaoAssistida.java, PerfilExplícitoGeral.java, PerfilExplicitoFilmes.java, TelespectadorPerfilForm.java, ModeloTelespectador.java, ProgramacaoSI.java e ProgramacaoRecomendada.java.

Uma estrutura de base de dados em XML bastante utilizada no contexto de toda a nossa aplicação refere-se à nossa base de dados de programação, ou seja, a base contendo toda a programação a ser exibida pelo EPG e oriunda das tabelas SI. Neste sentido, a Figura 16 apresenta a estrutura do arquivo EPG.xml, exibindo a relação dos programas e seus atributos:

```
<epg>
  <canal nome="Rede TV" id="02">
    <subcanal id="021" nome="Notícias" favorito="0">
      <programa id="02101" ag="0">
        <nome>Desfrutando a vida diária</nome>
        <data>1808200918082009</data>
        <diasemana>Dom</diasemana>
        <hora>0600</hora>
        <duracao>30</duracao>
        <descricao>Sem descrição</descricao>
        <classificacao>Livre</classificacao>
        <categoria>Notícias</categoria>
        <tipo>Local</tipo>
        <recomendar>0</recomendar>
      </programa>
      <programa id="02102" ag="0">
        <nome>TV Esporte Notícias</nome>
        <data>1808200918082009</data>
        <diasemana>Dom</diasemana>
        <hora>0630</hora>
        <duracao>30</duracao>
        <descricao>Apresentado, ao vivo, pelo jornalista José Carlos
        </descricao>
        <classificacao>12 anos</classificacao>
        <categoria>Notícias</categoria>
        <tipo>Nacional</tipo>
        <recomendar>0</recomendar>
      </programa>
    </subcanal>
  </canal>
</epg>
```

Figura 16: Estrutura do arquivo EPG.xml.

Todos os canais são compostos por subcanais, segregando a programação nas categorias mais comuns de um programa. Quanto aos subcanais, estes possuem todos os programas que serão exibidos para o usuário. Cada programa tem o atributo 'ID', que é único para o programa, mesmo se comparado a programas de subcanais diferentes. O atributo 'AG'

é utilizado para sinalizar que o programa está agendado (identificador '0') ou não agendado (identificador '1'). Além desses dois atributos que estão na raiz de cada registro programas, cada programa também contém o atributo 'nome' relativo ao programa e que será mostrado no guia de programação eletrônico (EPG). Outros atributos utilizados podem ser mencionados abaixo:

- 'diaSemana', usado para mostrar em qual(is) dia(s) da semana o programa é transmitido;
- 'hora', que especifica o horário em que é transmitido o programa;
- 'duração', que mostra quanto tempo é a duração do requerido programa;
- 'descrição', que funciona como um resumo do programa em questão;
- 'classificação', que especifica a faixa etária recomendada para se assistir ao programa (bastante utilizado para o caso de implementação dos mecanismos de censura de programação);
- 'tipo', que pode ser nacional ou internacional;
- 'recomendar', que especifica se o programa está sendo recomendado (identificador '1') ou não está sendo recomendado (identificador '0') pelo sistema de recomendação.

Outro arquivo usado no sistema é o avaliaRecomendação.xml, que irá armazenar, para um determinado perfil, todos os programas que foram reprovados pelo sistema. Sua estrutura é a seguinte:

```
<recomenda>
<perfil id="1"
programasRecomendados=" 112 45466 32267 45635 "
canaisRecomendados="GLOBO"
subCanaisRecomendados="NOTICIAS ESPORTES"
diasRecomendados=""
palavrasChaves="filmes entretenimento"
percentual TFIDF="53"
percentual APRIORI="59"
percentual HIBRIDO="71"
percentual RESTRICAO="76" />
</recomenda>
```

Figura 17: Estrutura do arquivo de avaliação da recomendação.

Na estrutura desse arquivo de avaliação da recomendação são abordados dois aspectos principais: o primeiro é a avaliação básica dos programas que foram considerados recomendados pelo sistema, para os quais o usuário terá a possibilidade de aprovar ou não a recomendação do programa. Caso o programa seja considerado reprovado, ele irá compor o

campo ‘programasNaoRecomendados’, que contem todos os ‘id’s’ dos programas reprovados na avaliação do usuário e que não mais serão recomendados para aquele perfil.

A outra abordagem que é de grande valia para a avaliação é a inclusão de críticas à recomendação, através da inserção pelo usuário de algumas restrições para subprogramas, canais, dias da semana e até mesmo algumas palavras que não fazem parte da sua preferência. No primeiro campo ‘canaisNaoRecomendados’, o usuário irá adicionar os canais que não são de sua preferência. O mesmo ocorre com o atributo ‘subCanaisNaoRecomendados’, que restringe o subcanal da recomendação. O atributo ‘diasNaoRecomendados’ também é usado para restringir dias da semana no qual o usuário não deseja receber recomendação.

O último campo é o ‘palavrasNaoChaves’, representando um ponto de maior importância pois indica as palavras que, se estiverem presentes nos atributos de um determinado programa, fará com que o sistema não realize esta recomendação.

5.3 A IMPLEMENTAÇÃO DA AVALIAÇÃO DO SISTEMA DE RECOMENDAÇÃO

Nesta sub-seção iremos exibir e explicar alguns dos principais códigos utilizados para o desenvolvimento de nosso sistema de recomendação. Durante o uso do terminal de acesso, toda a programação assistida pelo usuário é armazenada em documento XML. Segue exemplo:

```
<ProgramacaoAssistida>
- <!--
  dias da semana numerados de 1 a 7
-->
<assistiu id="1121" canal="10" subcanal="2" diaSemana="1" horaInicio="08:30" duracao="20"
  genero="filmes" />
<assistiu id="1121" canal="10" subcanal="2" diaSemana="1" horaInicio="08:50" duracao="5"
  genero="noticias" />
<assistiu id="1121" canal="10" subcanal="2" diaSemana="1" horaInicio="08:55" duracao="5"
  genero="nacional" />
<assistiu id="1111" canal="12" subcanal="3" diaSemana="1" horaInicio="09:00" duracao="45"
  genero="esportes" />
<assistiu id="1144" canal="02" subcanal="1" diaSemana="1" horaInicio="09:45" duracao="31"
  genero="infantil" />
<assistiu id="1149" canal="08" subcanal="1" diaSemana="1" horaInicio="10:16" duracao="50"
  genero="filmes" />
<assistiu id="1234" canal="12" subcanal="2" diaSemana="1" horaInicio="10:56" duracao="30"
  genero="esportes" />
</ProgramacaoAssistida>
```

Figura 18: Documento contendo programação assistida pelo usuário.

Para a criação do subsistema de avaliação da recomendação, inicialmente foi preciso a criação de uma tela listando todos os programas recomendados, tendo a opção de permitir ao usuário que o mesmo aprove ou não cada um dos programas. Para isso, foi criada a classe AvaliarRecomendação. Esta obtém todos os programas que foram setados como recomendados (programas que possuem o atributo <recomenda> igual a 1). Neste sentido, foi

adicionado à Classe ProgramaçãoRecomendada, o método que percorre todo o arquivo EPG.xml e obtém todos os programas que são recomendados.

Representando a segunda fase da avaliação da recomendação, é disponibilizado um formulário que possibilita ao usuário adicionar restrições à programação, excluindo do processo de recomendação alguns canais, subcanais, dias da semana, dentre outros, podendo adicionar palavras que funcionarão como não chaves, ou seja, que não farão parte da recomendação.

Esses dados de restrições de programação são obtidos durante o processo de avaliação da recomendação e são adicionados como forma de atributos para a classe TelespectadorPerfilForm, cuja representação em formato XML é:

```
<recomenda>
  <perfil id="1"
    programasNaoRecomendados=" 102 134 234 5645 45466 32232 45635 "
    canaisNaoRecomendados="SBT REDE TV BAND"
    subCanaisNaoRecomendados="NOTICIAS ESPORTES"
    diasNaoRecomendados="SEGUNDA TERCA QUINTA"
    filmesIndesejados="COMEDIA ACAO "
    palavrasNaoChaves="" />
</recomenda>
```

Figura 19: Trecho de código – Restrições de programação em XML.

As restrições através de códigos Java são apresentadas conforme Figura 20:

```
private Collection<String> canais = new ArrayList<String>();
private Collection<String> subCanais = new ArrayList<String>();
private Collection<String> diasSemana = new ArrayList<String>();
private String palavrasNaoChaves ;
```

Figura 20: Trecho de código – Restrições de programação em java.

Sempre que o usuário solicita avaliar a recomendação, é criado um registro com o mesmo 'id' do usuário que a realizou, juntamente com os 'id's' dos programas que foram reprovados, os canais, subcanais e os dias da semana que foram restritos da recomendação

Ainda relacionado ao código acima, para que seja realizada, em nível de código fonte, a restrição da programação, foi criado método `setarProgramacaoRecomendada( )`, que antes de setar a recomendação, lê o arquivo `avaliaPerfil.xml` e verifica se o usuário já realizou a avaliação. Em caso afirmativo, terá que levar em consideração as restrições de programação, antes de realizar a recomendação.

Desta forma, sempre que o usuário solicitar novamente realizar a avaliação da recomendação, os dados da avaliação anterior são previamente setados no formulário de restrições e, assim, o usuário poderá visualizar sua avaliação anterior e adicionar ou remover

novas restrições. Com isso, o usuário passa a ser capaz de realizar quantas modificações desejar e assim customizar a recomendação a seu critério.

5.4 A GERAÇÃO DAS REGRAS DE ASSOCIAÇÃO

A efetividade do processo de recomendação depende diretamente das regras de associação geradas. Convém ressaltar que a implementação do algoritmo Apriori que estamos utilizando no contexto deste trabalho é baseada na mesma implementação disponível na ferramenta WEKA (*Waikato Environment for Knowledge Analysis*). Esta ferramenta possui a implementação de uma série de algoritmos de mineração de dados. A razão de utilizarmos esta implementação deveu-se, principalmente, à qualidade da implementação do pacote WEKA (WITTEN; FRANK, 2005).

A classe `AprioriRecomendacao.java` é responsável pela análise do comportamento implícito do usuário e, a partir daí, a identificação dos padrões não triviais relacionados ao comportamento deste usuário. Nesta classe, através dos atributos ‘ocorrência’ e ‘confiança’, temos a definição dos fatores suporte e confiança, respectivamente, informações inerentes ao contexto das regras de associação e utilizadas pelo algoritmo Apriori para a geração destas. Após realizar a instância de uma classe contendo a programação assistida, o sistema inicializa uma nova geração de regras, logo após reinicializar o valor das variáveis.

Neste primeiro momento as regras de associação geradas em ‘`AprioriRecomendacao.XML`’, estão buscando padrões não triviais somente nas relações entre gêneros, atores e diretores. Nossa atual implementação não informa que todos os usuários que assistem filmes lançados neste ano também assistem filmes de terror. A razão para isso se dá unicamente para fins de facilitar nosso estudo de caso, não representando nenhuma dificuldade a mais em termos de implementação.

Por fim e como a etapa mais importante, temos, a partir do método `aprioriGenRecomendacao()`, a geração das regras propriamente ditas. Todos os programas assistidos por um período superior a cinco minutos irão entrar neste processamento. Todas as relações entre gêneros, atores e diretores da programação assistida serão estabelecidas e, somente aquelas que satisfizerem os fatores de suporte e confiança (método `getRulesRecomendacao()`) irão fazer parte das novas regras descobertas para o usuário telespectador em questão, conforme mostrado na Figura 21:

```

public void aprioriGenRecomendacao() {
    FuncoesUtils funcoes = new FuncoesUtils();
    ArrayList<ProgramaAssistido> programasAssistidos = null;
    programasAssistidos = funcoes.obterProgramasAssistidos();

    ProgramaAssistido programaAnterior = null;
    int totalRelacao;
    for (ProgramaAssistido programaAtual : programasAssistidos) {

        if (programaAnterior != null
            && programaAtual.getTempoAssistido() > 5) {
            for (TipoEnumRelacoesGeneros relacao : relacoesGenero) {
                if (funcoes.intervaloPeriodoHora(
                    programaAnterior.getHora(),
                    programaAtual.getHora(), programaAnterior
                        .getTempoAssistido()) <= periodoValido
                    && (relacao.getGeneroA().equals(
                        programaAnterior.getGenero()) && relacao
                                                                    .getGeneroB().equals(
                                                                    programaAtual.getGenero())) {

                    totalRelacao = relacao.getNumeroOcorrencias();
                    relacao.setNumeroOcorrencias(totalRelacao + 1);

                }
            }

            adicionarAtor(funcoes.colecionaStrings(programaAtual.getAtores()),
                programaAtual.getIdPrograma());
            adicionarDiretor(programaAtual.getDiretores(), programaAtual.getDiretores());
        }
        if (programaAtual.getTempoAssistido() > 5)
            programaAnterior = programaAtual;
    }
}

public Collection<TipoEnumRelacoesGeneros> genRulesRecomendacao() {
    return resultadoRelacoes;
}

public Collection<TipoEnumRelacoesGeneros> obterRelacoesRecomendacao() {
    aprioriGenRecomendacao();
    return genRulesRecomendacao();
}

public Collection<String> obterAtoresApriori(){
    Collection<String> atoresApriori = new ArrayList<String>();
    int tamProgramas = new FuncoesUtils().obterProgramasAssistidos().size();
    Set<Integer> idAtorFinal = atores.keySet();

    for(Integer id1:idAtorFinal){
    }
    int tamAtor=0;
    for(Integer id: idAtorFinal){
        tamAtor = atores.get(id).size();
        if(((double)tamAtor/(double)tamProgramas)>percentagemAtorDiretor){
            atoresApriori.add(atoresNome.get(id));
        }
    }

    Set<Integer> idDiretorFinal = diretores.keySet();
    int tamDiretor=0;
    for(Integer idDir: idDiretorFinal){
        tamDiretor = diretores.get(idDir).size();
    }
}

```



```

        if(((double)tamDiretor/(double)tamProgramas)>percentagemAtorDiretor){
            atoresApriori.add(diretoresNome.get(idDir));
        }
    }
    return atoresApriori;
}

```

Figura 21: Geração e obtenção das regras.

5.5 A IMPLEMENTAÇÃO DO EPG

A avaliação do sistema de recomendação é feita utilizando um módulo de recomendação que, por sua vez, está incorporado ao EPG. Este guia de programação eletrônico é um sistema desenvolvido pelo Instituto Federal de Ciência e Tecnologia (IFCE) que consiste em um aplicativo para a TV Digital com a finalidade de exibir toda a programação através de um guia, com os canais, subcanais e programas. A partir disso, e como contribuição do nosso trabalho, desenvolvemos o processo de criação de perfis, geração das recomendações, exibição destas recomendações no EPG, avaliação das recomendações e inclusão das restrições de recomendação.

Através dos dados de programação, o EPG monta uma grade com todos os canais, subcanais e programas de acordo com o dia e o horário em que eles são exibidos. As principais tabelas (TABELAS SI, 2005) das quais as seguintes informações foram utilizadas no trabalho são a PAT, PMT e EIT.

Em relação ao desenvolvimento da aplicação, conforme já mencionado, esta foi desenvolvida em Java TV, que é uma API para implementação de aplicação na plataforma da TV Digital. Esta interface de programação fornece os seguintes componentes que foram usados na implementação de todo o sistema:

- *javax.tv.graphics*: habilita simples composições e fornece um container para construção de interfaces com componentes AWT;
- *javax.tv.locator*: fornece acesso aos dados e recursos endereçados ao receptor de TV digital;
- *javax.tv.media*: define uma controle e eventos para gerenciamento de conteúdos visuais e de áudio;
- *javax.tv.net*: fornece acesso aos datagramas IP transmitidos pelo stream de broadcast;
- *javax.tv.service*: fornece acesso ao banco de dados do service information (SI) e APIs para seus sub-pacotes;
- *javax.tv.util*: suporta a criação e gerenciamento de eventos de tempo;

- *javax.tv.xlet*: fornece os métodos para gerenciamento do ciclo de vida da Xlet controlados pelo gerenciador de programas.

Todo o EPG é basicamente composto por uma classe que serve de *container* para todas as outras classes. A classe principal da aplicação é a *EPGXletview*, que acopla todos os componentes, onde cada um destes herda da classe *Container*. Nas classes componentes temos como estrutura básica os seus atributos, um construtor e o método *paint*, que é responsável por renderizar os componentes da classe, cujos principais componentes são:

5.5.1 CLASSE *EPGGrade*

Esta classe tem a função de renderizar a grade de programação, mostrando todos os canais, subcanais e programas, juntamente com a linha de tempo que mostra o horário em que determinado programa será exibido. É nessa classe que temos o controle da grade apresentada da navegação pelos diversos canais, subcanais e programas, onde, ao deslocar para os dois lados, os programas são desenhados em seu total, ou apenas uma parte deles (isto é ocasionado pela linha de tempo que analisa o período pelo qual o programa é exibido e a partir disso renderiza os programas que estão contidos naquele trecho de horário). **Esta classe foi desenvolvida para o EPG e extendida no escopo desta dissertação.**

5.5.2 CLASSE *Barra*

Essa classe é responsável por mostrar os botões que ficam na parte inferior da tela da grade de programação. Os botões principais que essa classe mostra são os botões ‘Exibir recomendação’ que é visível na tela de usuários do sistema, o ‘Salvar dados’, que é visível na tela de cadastro do perfil do usuário, onde ao ser pressionada o sistema salva os dados obtidos do usuário no arquivo *usuários.xml*, e o botão ‘Voltar’, que retorna à página anterior. Também temos o botão ‘Avaliar Recomendação’ que é apresentado quando o usuário está logado e deseja avaliar o sistema de recomendação. **Esta classe foi desenvolvida para o EPG e extendida no escopo desta dissertação.**

5.5.3 CLASSE *Informação*

Essa classe é responsável por apresentar os detalhes do programa selecionado. **Esta classe foi desenvolvida para o EPG e extendida no escopo desta dissertação.** Ela tem como entrada o id do programa selecionado na grade de programação e através desse id ele localiza o canal, subcanal e o programa e apresenta os detalhes, como mostrado na Figura 22:



Figura 22: Tela de exibição de detalhes de um programa.

5.5.4 CLASSE CADASTRARUSUÁRIO

É nessa classe que foram implementados todos os formulários de cadastro do usuário no sistema que serão usados para gerar a recomendação de programação. Ela utiliza os métodos de ler e gravar dados em um XML, funcionando como uma classe que é responsável pela persistência dos dados e ao mesmo tempo que é responsável pela parte gráfica da visualização dos formulários no sistema.

Esta classe também é responsável por toda a parte de formulários e também possui a rotina que tem como funcionalidade ler o teclado e a partir daí, inserir os caracteres na caixa de cada registro. Como todos os campos de texto, marcações e de listas de opções não são previamente implementados no pacote JAVA TV, todos esses recursos foram criados utilizando métodos básicos de criação de retângulos e pequenos quadrados (para marcação) e círculos para que assim fosse capaz de se visualizar esses componentes que fazem parte das telas de formulário.

5.5.5 CLASSE TELESPECTADORPERFILFORM

Esta classe encapsula todos os dados referentes aos formulários e que são mapeados para os arquivos XML. Os atributos lidos na classe filtram os programas que coincidem com os dados do usuário e que se encontram setados como recomendados. Para todo novo usuário é inserido uma nova classe para encapsular os dados.

5.5.6 CLASSE PROGRAMAÇÃO

Esta classe é responsável por obter todos os programas dos arquivos XML e, ao mesmo tempo, irá analisar os dados do usuário logado e setar o atributo <RECOMENDA> como programa recomendado. É nessa classe que são obtidos os atributos da classe TelespectadorPerfilForm, todos setados com os dados do usuário.

5.6 TELAS DA APLICAÇÃO

A seguir são apresentadas as principais telas. A Figura 23 representa a primeira tela do sistema e a relação de perfis já cadastrados, tornando possível a inclusão de novos usuários:



Figura 23: Tela de usuários cadastrados.

A partir do momento em que o usuário pressionar o botão de cor azul, a funcionalidade de criação de perfis é iniciada, exibindo a primeira tela contendo informações gerais sobre o perfil do usuário telespectador em questão. Esta primeira tela (Figura 24) contém informações gerais sobre o perfil do usuário. Aqui destacamos a presença de duas informações altamente relevantes para o sistema. É através das informações presentes no campo 'PALAVRAS-CHAVE' que o algoritmo TF-IDF irá analisar os metadados 'nome' e 'descrição', do arquivo EPG.xml (que representa a base de dados de programação), para buscar os programas, neste caso, os filmes, que se enquadram com o perfil do usuário. Além disso, dependendo da informação presente no último item desta tela, o sistema irá ou não realizar o processo de recomendação para o usuário.

Figura 24: Tela para cadastro de perfis (informações gerais).

Após inseridos os dados e pressionado novamente o botão de cor azul, a próxima tela do cadastro do perfil é exibida (Figura 25), contendo informações sobre preferências dos usuários nos casos de filmes. As informações dos campos ‘PALAVRAS-CHAVE’, ‘ATORES’ e ‘DIRETORES’ também serão parâmetros de entrada do algoritmo TF-IDF. Em relação a estes dois últimos, vale ressaltar que ambos permitem a escolha de mais de uma opção.

Figura 25: Tela para cadastro de perfis (informações sobre filmes).

Após a confirmação das informações, passamos a ter mais um perfil de usuário cadastrado na base de dados do sistema. Dessa forma, uma vez selecionado um usuário e pressionado o botão de cor vermelha, neste momento o EPG será exibido, mostrando, de forma destacada, os programas (filmes) considerados recomendados para este usuário, conforme mostrado na Figura 26:

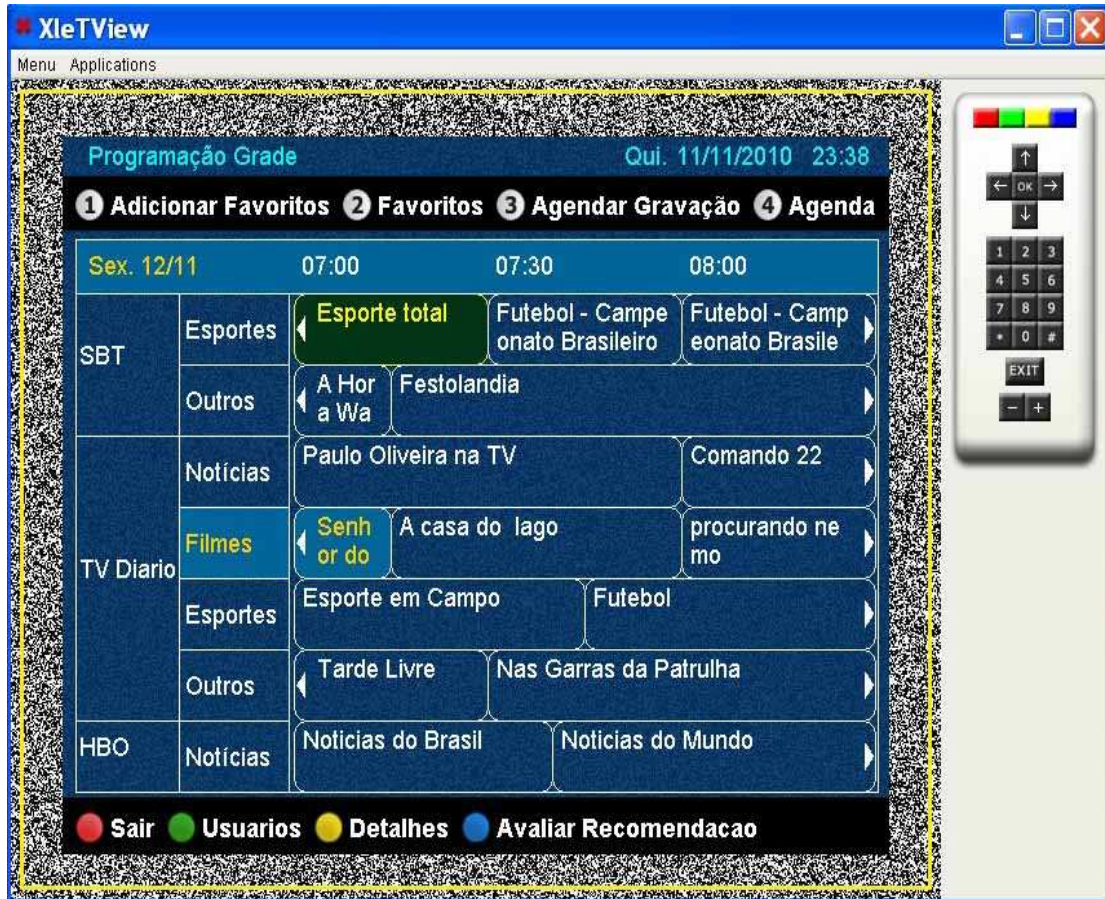


Figura 26: EPG exibindo os programas recomendados.

Sobre o processo de recomendação ocorrido neste exemplo, a saída do algoritmo TF-IDF também contou com o resultado gerado pelo algoritmo Apriori. Em relação a este último, utilizamos um arquivo denominado 'progassistida.xml' (contendo todos os programas assistidos pelo Usuário Telespectador) para simular a programação assistida pelo usuário. O casamento das informações geradas pelos dois algoritmos tornou possível a recomendação dos programas que serão mostrados na tela da Figura 27. Esta é exibida quando o usuário, ao pressionar o botão de cor azul, solicita a funcionalidade de Avaliação da Recomendação:



Figura 27: Funcionalidade de Avaliação do Sistema de Recomendação.

Percebemos que todos os programas considerados recomendados para o perfil são exibidos nesta tela. A segunda coluna permite que o usuário informe, através do botão de cor vermelha, se este aprovou ou não a recomendação. A terceira coluna verifica no metadado ‘sintonização’, do arquivo ‘progassistida.xml’, se o usuário realizou ou não a sintonia do canal onde o programa foi exibido e na hora em questão.

Aqui é importante ressaltar a importância das informações preenchidas pelo usuário para o sistema. Uma vez que ocorra a recorrência da reprovação de um programa recomendado, ou a recorrência da não sintonização do mesmo, este deixará de ser recomendado ao usuário, por mais que seus atributos atendam aos parâmetros considerados “recomendáveis” para o perfil em questão. Sobre esta recorrência, neste primeiro momento definimos que significa a ocorrência de duas reprovações ou de duas não sintonizações. Entretanto, este atributo é parametrizável, de acordo com a necessidade da aplicação.

Outra informação presente nesta tela é a de Percentual de Acerto do sistema, baseando-se na técnica de Precisão (HERLOCKER, 2004). Para cada programa considerado recomendado, este poderá receber 3 valores. Caso o programa tenha sido sintonizado e aprovado, valor 100. Ocorrência de aprovação e ausência de sintonização ou vice-versa, valor 50. Caso o programa não tenha sido sintonizado e nem aprovado, valor nulo. Após a

pontuação de todos os programas, o sistema calcula a média aritmética e informa o percentual de acerto dependendo da abordagem utilizada, em relação às técnicas envolvidas, para fazer a recomendação.

Por último, de forma a garantir que o sistema irá consumir as informações do *feedback* do usuário e permitir que este evolua constantemente, uma vez concluído o processo de avaliação, uma tela de restrições é exibida ao usuário, conforme mostrado na Figura 28:



Figura 28: Avaliação da recomendação - Restrições.

Caso o sistema realize recomendações que não agradaram o usuário com base na estratégia de recomendação utilizada, as restrições informadas também serão analisadas pelo sistema de forma a permitir que este desconsidere, durante o processo de recomendação, determinados programas. Esta funcionalidade pode ser de grande utilidade, sobretudo nos primeiros dias em que um perfil foi criado.

5.7 PROVA DE CONCEITO

A realização dos experimentos de nosso trabalho foi feita considerando o domínio filmes de televisão. Visando analisar a semelhança da abordagem proposta com a de outros trabalhos relacionados, consideramos prudente mencionar que esta mesma abordagem foi

realizada em (SILVA, 2005). No citado trabalho foi utilizada uma coleção composta de 217 documentos, com o objetivo de efetuar uma análise comparativa entre dois algoritmos de recomendação, o *Rocchio* e o *kNN*. Cada documento da coleção correspondia a um filme específico e possuía um texto com dados estruturados, tais como descrição, nome, data, horas e canal.

Voltando ao contexto do nosso trabalho, o mesmo também partiu do princípio de popular os dados de entrada de forma não automatizada, mas fazendo a pesquisa em diversos sites e estruturando os documentos XML com os principais atributos descritivos dos filmes. Ainda neste sentido, para a realização da recomendação, o sistema utiliza as informações cadastradas no perfil do usuário, em específico, no caso da utilização do algoritmo TF-IDF, para a verificação das informações contidas no campo palavras-chave.

Nossa proposta *não* teve um objetivo de avaliar o algoritmo, mas sim a efetividade das estratégias de recomendação utilizadas. Com a interação usuário telespectador x programação de TV, algumas regras de associação são geradas, devido ao uso do algoritmo Apriori, incrementando o processo de recomendação. Além disso, a inclusão das restrições também aprimora o processo, aumentando o percentual de acerto.

Posteriormente foram realizadas algumas simulações, em condições similares às anteriores descritas, com um total de 8 perfis diferentes e mais de 150 registros de programas (filmes), visando analisar o quanto cada algoritmo/técnica utilizada em nossa estratégia aprimora a recomendação realizada. O cálculo do percentual de acerto é realizado com as informações cadastradas no perfil, sem a geração de nenhuma regra de associação. Após isso, é verificado o percentual para usuários com o perfil e com algumas regras geradas e, por fim, analisamos o considerável aumento do percentual de avaliação do sistema quando este utiliza, além dos algoritmos TF-IDF e o Apriori, a inclusão das restrições de recomendação. A tabela a seguir mostra os valores obtidos neste estudo comparativo:

Tabela 3: Percentual de acerto do sistema de acordo com a estratégia utilizada.

Usuários/Estratégia	TF-IDF	APRIORI	TF-IDF + APRIORI	TF-IDF + APRIORI + RESTRIÇÕES
Diego	69.23%	75%	84.61%	88%
Marcelo	42.64%	50.73%	57%	86%
Thiago	47.43%	60.25%	65.38%	84.9%

User	70.31%	67.18%	76.56%	87.5%
João	79.31%	77.58%	82.75%	85.71%
Maria	74.72%	73.62%	83.51%	88.37%
Pedro	57%	60%	69%	83.5%
Geral	61.11%	72.2%	77%	87.5%

Os valores da tabela foram gerados através dos documentos XML contendo o comportamento dinâmico. A análise destes dados nos permite perceber que, conforme gráfico exibido na Figura 29, o acréscimo das Regras de Associação aprimora o processo de recomendação utilizando somente a Filtragem Baseada em Conteúdo e vice-versa.

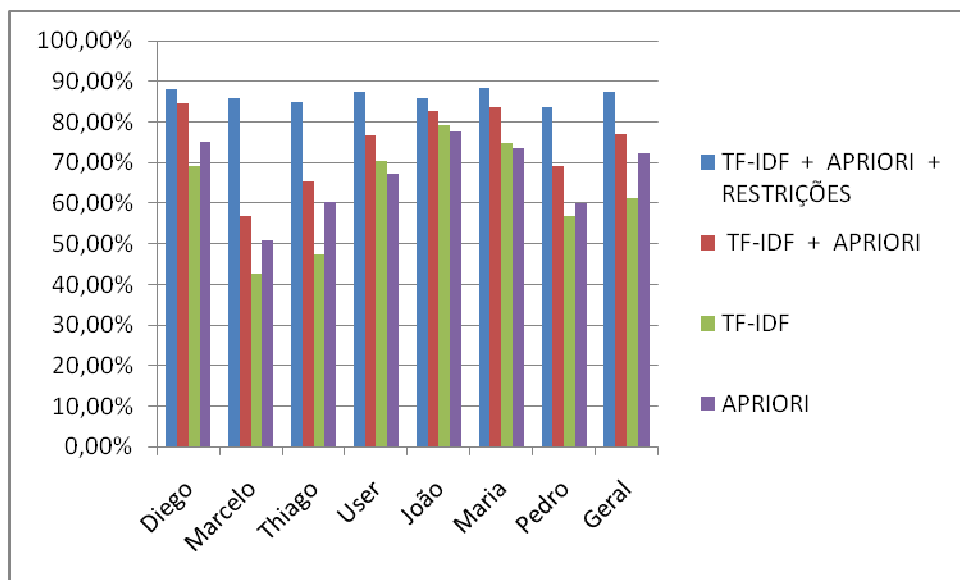


Figura 29: Evolução do percentual de acerto do sistema de acordo com a estratégia utilizada.

Ainda relacionado ao gráfico da figura anterior, percebemos que a inclusão das restrições de recomendação, após o processo avaliativo, também aprimorou o percentual de acerto do sistema, quando este apenas utiliza a combinação dos algoritmos TF-IDF e Apriori.

No intuito de quantificarmos a melhoria do percentual avaliativo do sistema à medida que a estratégia de recomendação é aprimorada, o gráfico da Figura 30 nos mostra, para cada usuário, o percentual de aprimoramento existente em três casos: primeiramente, a inclusão do algoritmo Apriori quando antes só se tinha o TF-IDF. Em seguida, o inverso, ou seja, a inclusão do algoritmo TF-IDF quando só se tinha o Apriori. Por último, a inclusão das

Restrições de Recomendação quando antes só se tinha a combinação da FBC (TF-IDF) com as Regras de Associação (Apriori):

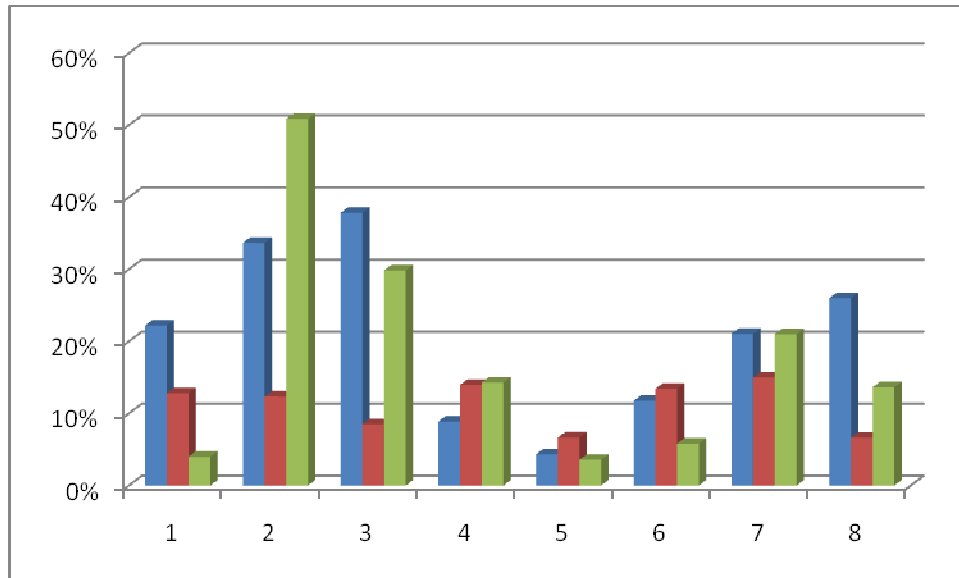


Figura 30: Identificação das porcentagens de evolução do processo avaliativo conforme aprimoramento da estratégia de recomendação.

5.8 CONSIDERAÇÕES FINAIS

Neste capítulo foi possível entender melhor tecnicamente a maneira pela qual a proposta desta dissertação foi implementada. Após apresentação da Prova de Conceito tornou-se mais claro o entendimento de que a combinação de estratégias aumenta o percentual de acerto do sistema, deixando seu processo de recomendação bem mais robusto.

6. CONSIDERAÇÕES FINAIS

Nesta dissertação apresentamos a proposta de uma estratégia híbrida de recomendação a ser utilizada no contexto da televisão digital interativa, através do uso de uma aplicação do tipo Guia de Programação Eletrônico (EPG). A estratégia conta com a combinação de dois algoritmos bastante utilizados no contexto de sistemas de recomendação, o Apriori e o TF-IDF. Além disso, previmos a criação de perfis de usuários e da inclusão de restrições de recomendação, cujas informações presentes são consumidas pelos algoritmos quando do processamento dos dados para posterior geração das recomendações.

No intuito de viabilizar a validação da arquitetura/estratégia proposta, tornou-se necessário o desenvolvimento de um sistema de recomendação que implementasse a referida proposta. Assim, este sistema é capaz de realizar recomendações de maneira individualizada, após perceber, de forma implícita, as preferências de cada usuário. O sistema também possibilita a criação de perfis de forma explícita para cada usuário do sistema, assim seja a vontade de quem o usa. Além do mais, o sistema é capaz de criar um perfil geral, a partir da análise das informações obtidas sobre a programação assistida pelos usuários em todas as horas de todos os dias da semana. A recomendação obtida é chamada de geral uma vez que é construída independentemente de qual usuário telespectador está defronte do aparelho de TV.

Finalmente, o sistema propicia uma evolução contínua da qualidade da recomendação através do módulo de avaliação, que permite que o usuário possa dar um retorno acerca das informações recomendadas. Desta forma, eventuais falhas na recomendação ocorridas no início do uso do sistema por um novo perfil podem ser corrigidas com o decorrer do seu uso.

Adentrando mais nas contribuições do trabalho, é possível identificar importantes benefícios no mesmo. Analisando segmentos de gestão da informação, passa a ser possível que grandes volumes de dados multimídia distribuídos (programação televisiva), possam ser acessados de forma seletiva, através de metadados de programação. Por outro lado, podemos falar em viabilização do acesso participativo e universal dos cidadãos ao conhecimento e à

informação. A abordagem aqui proposta tem o objetivo de amenizar o problema já citado no início desta dissertação, usuários defronte de uma vasta gama de canais televisivos sem saber quais melhor se adequam ao seu perfil, sem exigir para isso nenhum conhecimento tecnológico específico. Isto se deve devido ao fato da estratégia aqui proposta contar com o uso de um sistema de fácil manipulação e de resposta bastante satisfatória.

Ainda citando o início deste trabalho, não foi nosso objetivo desenvolver uma proposta voltada para o modelo brasileiro. Entretanto, sabemos que, de acordo com as tecnologias utilizadas em nossa arquitetura e o uso aderente às tabelas SI do modelo brasileiro, a estratégia especificada é totalmente passível de ser abrigada pelo GINGA.

Finalmente, o estudo de caso apresentado ilustrou o uso do sistema desenvolvido por 8 usuários e mais de 150 registros através de uma série de simulações. Inicialmente foi avaliada a combinação dos algoritmos TF-IDF e Apriori, permitindo visualizar uma evolução do percentual de acerto do sistema na ordem de 16%. Quando adicionamos os procedimentos de restrição de recomendação à combinação anteriormente citada, o percentual teve um acréscimo em média de 18%.

6.1 TRABALHOS FUTUROS

Uma série de trabalhos futuros pode ser vislumbrada como continuação do que foi produzido nesta dissertação. Entre esses, inicialmente pretende-se, uma vez que o desenvolvimento de aplicações para um ambiente de tv digital interativa já possui, por si só, algumas restrições no que diz respeito ao processo de testes e simulação da aplicação, transpor os módulos deste sistema para um ambiente real de transmissão digital de TV.

Na prática, pode ser interessante que a recomendação possa ser realizada através da recepção do sinal e, com este, as informações da tabela SI, tendo um terminal de acesso como recurso responsável pela execução das funcionalidades da aplicação. No entanto, existem diversas dificuldades técnicas por causa da incipiência de trabalhos neste sentido e o difícil acesso a recursos deste tipo, sobretudo no estado do Ceará.

Visando ampliar a abrangência desta aplicação é interessante analisar e efetuar os ajustes de forma que este sistema seja totalmente compatível com as especificações técnicas do modelo brasileiro de televisão digital. Para isso, se torna necessário a realização de novos testes e implementações, contando para isso, com terminais de acesso dotados do *middleware* brasileiro GINGA.

Outro trabalho que consideramos bastante ousado, e também por isso bem interessante, consiste em adaptar a estratégia híbrida aqui proposta para uma arquitetura de dispositivos móveis sensíveis a contexto. A idéia é que usuários do sistema tenham seus perfis criados através de dispositivos móveis que sejam capazes de aferir preferências implícitas inclusive de acordo com o ambiente no qual estes usuários se encontram. Com isso, podemos passar a ter o processo de recomendação sendo realizado para perfis de usuários em comum, e não para cada usuário do sistema, individualmente. Teríamos uma recomendação POR PERFIL, o que acarretaria na necessidade de definição dos diversos possíveis tipos de perfis de usuários para tv digital, como por exemplo infantil, jovem, esportista, etc.

Dada a quantidade de interfaces utilizadas pela aplicação, torna-se interessante, embora não mandatório, uma maior preocupação com aspectos de Interface Humano Computacional (IHC). Neste sentido, pretendemos consultar grupos de pesquisa que atuem nesta área no intuito de aplicar algumas boas práticas de IHC às telas da aplicação desenvolvida.

Por último, dada a menor importância e maior esforço operacional, realizaremos estudos de caso do processo de avaliação em escopos diferentes dos de filmes, envolvendo vários outros gêneros de programas, tais como, Esportes, Notícias, Educação, dentre vários outros.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABNT NBR 15603-1. ABNT NBR 15603-1:2007, Televisão digital terrestre - Multiplexação e serviços de informação (SI) - Parte 1: Serviços de informação do sistema de radiodifusão. Versão corrigida em 2008. Disponível em: < http://www.abnt.org.br/imagens/Normalizacao_TV_Digital/ABNTNBR15603-1_2007Vc_2008.pdf >. Data de acesso em 18 set. 2008;
- ABNT NBR 15603-2. ABNT NBR 15603-2:2007, Televisão digital terrestre - Multiplexação e serviços de informação (SI) - Parte 2: Sintaxes e definições da informação básica de SI. Versão corrigida em 2008. Disponível em: < http://www.abnt.org.br/imagens/Normalizacao_TV_Digital/ABNTNBR15603-2_2007Vc_2008.pdf >. Data de acesso em 18 set. 2008;
- ABNT NBR 15603-3. ABNT NBR 15603-3:2007, Televisão digital terrestre - Multiplexação e serviços de informação (SI) - Parte 3: Sintaxe e definição da informação estendida do SI. Versão corrigida em 2008. Disponível em: < http://www.abnt.org.br/imagens/Normalizacao_TV_Digital/ABNTNBR15603-3_2007Vc_2008.pdf >. Data de acesso em 18 set. 2008;
- ADOMAVICIUS, G., TUZHILIN, Alexander. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, Piscataway, v. 17, n. 6, p. 734-749, 2005;
- AGRAWAL, R., IMIELINSKI, T., SWAMI, A. Mining Association rules between sets of items in large databases. *International Conference Management of Data (SIGMOD-93)*. 1993;
- ALBUQUERQUE, J.P., FIGUEIRA FILHO, F.M., GEUS, P.L. *Sistemas de Recomendação e Interação na web social. I Workshop de Aspectos de Interação na Web Social*. Porto Alegre-RS, 2008;
- ANATEL. Agência Nacional de Telecomunicações. *A TV digital terrestre no Brasil: panorama atual*. Brasília, 2004;
- ANSARI, A., ESSEGAIER, S., KOHLI, R. Internet Recommendation Systems. *Journal of Marketing Research*, v.37, n.3, p. 363-375, 2000;
- AVILA, P.M., ZORZO, S.D. A personalized TV Guide System Compliant with Ginga. In: *XV Simpósio Brasileiro de Sistemas Multimídia e Web*. Fortaleza, 2009;

- BALABANOVIC, M., SHOHAM, Y. Fab: Content-Based, Collaborative Recommendation. Communications of the ACM, New York, v.40, n.3, p. 66-72, 1997.
- BAUDISCH, P., BRUECKNER, L. TV Scout Lowering the entry barrier to personalized TV program recommendation. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2002;
- BECKER, V. Recomendações de Usabilidade para TV Digital Interativa. In: II Workshop de TV Digital Interativa. Curitiba-PR, 2006;
- BECKER, V., MONTEZ, C. TV digital interativa: conceitos, desafios e perspectivas para o Brasil. 1. ed. Florianópolis: I2TV, 2004;
- BELVIN, N. J., CROFT, W. B. Information Filtering and Information Retrieval: two sides of the same coin?. Communications of the ACM, New York, v.35, n.12, p. 29, 1992;
- BRASIL. Decreto n 5.820, de 29 de Junho de 2006. Implantação do Sistema Brasileiro de Televisão Digital Terrestre - SBTVD-T. DOU de 27/11/2006. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Decreto/D5820.htm>. Data de acesso em: 18 set. 2008;
- BUCZAK, A. L., ZIMMERMAN, J., KURUPATI, K. Personalization: Improving Ease-of-Use, Trust and Accuracy of a TV show Recommender. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2002;
- BURKE, R. Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction, vol. 12, Issue 4, Novembro 2002. pp. 331-370;
- CALLAN, J. et al. Personalisation and Recommender Systems in Digital Libraries. Joint NSF-EU DELOS Working Group Report, 2002;
- CAZELLA, S. C., ALVARES, L. C. Modeling user's opinion relevance to recommending research papers. In: UM 2005 User Modeling: The Proceedings of the Tenth International Conference. Springer's LNAI, 2005;
- CORNEL, G., HORSTMANN, C.S. Core Java. SunSoft press, 1999;
- COTTER, P., SMYTH, B. PTV: Intelligent Personalised TV Guides. In: Proceedings of the 12th Innovative Applications of Artificial Intelligence IAAI Conference. AAAI Press, 2000;
- CPQD. Centro de Pesquisas e Desenvolvimento em Telecomunicações. Arquitetura de referência. São Paulo, 2006. Disponível em: <<http://sbtvd.cpqd.com.br>>. Data de acesso em: 18 set. 2008;
- DIFINO, A., NEGRO, B., CHIAROTTO, A. A Multi-Agent System for a Personalized Electronic Program Guide. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2004;

- DTV. Site Oficial da TV Digital Brasileira. Disponível em: <<http://www.dtv.org.br>>. Data de acesso em: 02 set. 2009
- DVB. Digital Video Broadcasting Project. Disponível em: <<http://www.dvb.org>>. Data de acesso em: 18 set. 2008;
- ECLIPSE. The Eclipse Foundation open source community. Disponível em: <<http://www.eclipse.org/>> Data de acesso em: 12 set. 2010;
- FERNANDES, J., LEMOS, G., SILVEIRA, G. Introdução à televisão digital interativa: arquitetura, protocolos, padrões e práticas. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 24, Salvador, 2004;
- FREED, K. Desenvolvimento de aplicações para TV Digital Interativa, 2000. Disponível em: <<http://www.media-visions.com/itv-newbies.html>>. Data de acesso em: 18 set. 2008;
- GINGA. Middleware Ginga – TV interativa se faz com Ginga. Disponível em: <<http://www.ginga.org.br>>. Data de acesso em: 18 set. 2008;
- GINGA-J. ABNT/CEET-00:001.85. Televisão digital terrestre - Codificação de dados e especificações de transmissão para transmissão digital – Parte 4: Ginga-J - Ambiente para a execução de aplicações procedurais. Setembro de 2007;
- GRAVES, M. Projeto de Banco de Dados com XML. Editora Makron Books. ISBN: 8534614717. Primeira Edição, 2003;
- GUERRERO, J. D. Determinación de tendencias en un portal web utilizando técnicas no supervisadas: Aplicación a sistemas de recomendaciones basados en filtrado colaborativo. Tese. Doutorado em Engenharia Eletrônica - Departament d'Enginyeria Electrònica, University of València, Espanha, 2004. Disponível em: <http://www.uv.es/jdmg/tesis_jdmartin.pdf>. Data de acesso em: 28 ago. 2008;
- HAN, J., KAMBER, M. Data mining: concepts and techniques. Morgan Kaufmann Publishers, San Francisco, USA, 2001;
- HERLOCKER, J. L. Evaluating Collaborative Filtering Recommender Systems. In: ACM Transactions on Information Systems, Vol. 22, No. 1, 2004;
- HERLOCKER, J. L. Understanding and Improving Automated Collaborative Filtering Systems. Tese. Doutorado em Ciência da Computação, University of Minnesota, Minnesota, 2000;
- HUANG, Z., CHUNG, W., ONG, T., CHEN, H. A Graph-based Recommender System for Digital Library. In: JCDL'02. Portland, Oregon, 2002;
- JAVA. Java + you. Disponível em: <http://www.java.com/pt_BR/> Data de acesso em: 12 set. 2010;
- JAVATV. Java TV Application Programming Interface. Disponível em: <<http://java.sun.com/products/javatv>>. Data de acesso em: 30 set. 2008;

- KIM, M.H., KO, S.M., MUN, J.S., JI, Y.G., JUNG, M.R.: A Usability Study on Personalized EPG (pEPG) UI of Digital TV. In HCI (3) 892-901, 2007;
- MINICOM. Ministério das Comunicações. Decreto que institui o Sistema Brasileiro de Televisão Digital (SBTVD), em 26/11/2003. Disponível em: <http://sbtvd.cpqd.com.br/downloads/decreto_4901_2003.pdf> Acesso em: 18 set. 2008;
- MITCHELL, P. D., GROGONO, P.D. Modelling Tecniques for Tutoring Systems. Computers and Education, 20(1),55-61, 1993;
- MORAES NETO, M. A. Novas Tecnologias da Informação e Comunicação: A Educação a Distância no Sistema Brasileiro de Televisão Digital. Monografia para obtenção do título de tecnólogo em Telemática. CEFET-Ce. Fortaleza, 2006;
- MORAES NETO, M. A. Gerenciamento de serviços de TI com o modelo ITIL: Proposta de curso a distância utilizando uma plataforma de televisão digital interativa. Monografia para obtenção do título especialista em Educação a Distância. SENAC-Ce. Fortaleza, 2007;
- MORAES NETO M.A.; CARDOSO D. ; SOUZA, C. T. ; CORTÉS, M. I. .Abordagem Combinada para Recomendação Personalizada Utilizando o Guia de Programação Eletrônico. In: XXXVII Seminário Integrado de Software e Hardware (SEMISH), 2010, Belo Horizonte. Anais do XXXVII Seminário Integrado de Softwaree Hardware, 2010.
- MORAES NETO, M.A., OLIVEIRA, C. T., SOUZA, C. T. . Uma Arquitetura Distribuída para o Gerenciamento de Informações Sobre Programação em Televisão Digital. In: XI Webmedia: Simpósio Brasileiro de Sistemas Multimídia e Web, 2005, Fortaleza. XI Webmedia: Simpósio Brasileiro de Sistemas Multimídia e Web, 2005.
- MURRAY, T. Authoring Intelligent Tutoring Systems: An analysis of the state of the art. International J. of Artificial Intelligence in Education, Vol. 10, pp. 98-129. 1999;
- NANOXML. DEVKIX - Marc De Scheemaeker. Disponível em: <<http://devkix.com/nanoxml.php>> Data de acesso em: 12 set. 2010;
- O’SULLIVAN, D., WILSON, D., SMYTH, B., MCDONALD, K., SMEATON, A. Improving the Quality of the Personalised Electronic Program Guide. In: L. Ardisonno, M. Maybury and Alfred Kobsa (eds.): UMUAI Special Issue on User Modeling and Personalization for Television, In Press. 2003;
- OLIVEIRA, C. T. Um estudo sobre padrões de Middleware para TV digital interativa. Monografia para obtenção do título de tecnóloga em Telemática. CEFET-Ce. Fortaleza, 2005;
- OLIVEIRA, J. P. M., GALANTE, R. M., MUSA, D. L., EDELWEISS, N. Uma proposta para editoração, indexação e busca de documentos científicos em um processo de avaliação aberta. I Workshop em Bibliotecas Digitais, 2005;

- PENG, C.; CÉSAR, P.; VUORIMAA, P. Integration of Applications into Digital Television Environment. In: INTERNATIONAL CONFERENCE ON DISTRIBUTED MULTIMEDIA SYSTEMS, Taipei, 2001. Disponível em: <<http://lib.tkk.fi/Diss/2002/isbn9512261723/article7.pdf>>. Data de acesso em: 18 set. 2008;
- PLASTINO, A. Regras de Associação e Algoritmos de Mineração de Dados. Mini-Curso 2, SBB D, 2001. Instituto Militar de Engenharia. Rio de Janeiro: Rio de Janeiro, 2001;
- PRESSMAN, R. Software Engineering: A Practitioner's Approach. 6ª edição, Mc Graw Hill, 2005;
- REATEGUI, E. B., CAZELLA, S. C. Sistemas de Recomendação. In: XXV Congresso da SBC, 2005, São Leopoldo. São Leopoldo, 2005. Disponível em: <<http://www.sbc.org.br/bibliotecadigital/download.php?paper=415>>. Data de acesso em: 28 ago. 2008;
- RESNICK, P., VARIAN, H. R. Recommender Systems. Communications of the ACM New York v. 40, n. 3, 1997, p. 56-58;
- RICCI, F., ARSLAN, B., MIRZADEH, N., VENTURINI, A. ITR: A case-based travel advisory system. In: 6th European Conference on Case Based Reasoning, 2002. Aberdeen. Proceedings of the seventh European Conference on Case Based Reasoning. Heidelberg: Springer Berlin, 2002, p. 613-627. Disponível em: <<http://dietorecs.itc.it/Papers/eccbr2002final.pdf>>. Data de acesso em: 28 ago. 2008;
- RICCI, F., DEL MISSIER, F. Supporting Travel Decision Making Through Personalized Recommendation. Designing personalized user experiences in eCommerce. Kluwer Academic Publishers, Norwell, MA, 2004. Disponível em: <<http://sra.itc.it/tr/RD04.pdf>>. Data de acesso em: 28 ago. 2008;
- SCHANK, R., CLEARY, M. Engines for Education. Lawrence Erlbaum Associates, 1995;
- SCHAFFER, J.B., KONSTAN, J. A., RIEDL, J. E-Commerce Recommendation applications. Data Mining and Knowledge Discovery. Kluwer Academic Publishers, Hingham, MA, USA, Vol 5. Issue 5, 2001;
- SHARDANAND, U., MAES, P. Social information filtering: Algorithms for automating “word of mouth” In: Conference on Human Factors in Computing Systems, 1995, Denver. Proceedings of the SIGCHI conference on Human factors in computing systems. New York: ACM Press/Addison-Wesley Publishing Co., 1995, p. 210-217. Disponível em: <<http://www.cs.ubc.ca/~conati/532b/papers/chi-95-paper.pdf>>. Data de acesso em: 28 ago. 2008;
- SILVA, F. S. Personalização de Conteúdo na TVDI Através de um Sistema de Recomendação Personalizada de Programas de TV (SRPTV). In: III Fórum de Oportunidades em Televisão Digital Interativa, 2005, Poços de Calda/MG. III Fórum de Oportunidades em Televisão Digital Interativa, 2005;

- SMYTH, B., WILSON, D., SULLIVAN, D. Improving the Quality of the Personalized Electronic Programme Guide. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2004;
- SOARES, L. F. G., RODRIGUES, R. F., MORENO, M. F. Ginga-NCL: the Declarative Environment of the Brazilian Digital TV System. In: Journal of the Brazilian Computer Society. No. 4, Vol. 13. p.37-46. ISSN: 0104-6500. Porto Alegre-RS, 2007;
- SOUZA FILHO, G. L., LEITE, L. E. C., BATISTA, C. E. C. F. Ginga-J: The Procedural Middleware for the Brazilian Digital TV System. In: Journal of the Brazilian Computer Society. No. 4, Vol. 13. p.47-56. ISSN: 0104-6500. Porto Alegre-RS, 2007;
- TABELAS SI. SBTVD - Tabelas de SI. Versão AA. PD.20.12.34A.0002A/RT-27-AA. CPqD, Campinas-SP, 2005;
- TV POR ASSINATURA. A realidade da TV por assinatura no país. Mais de 10 milhões de residências tem tv por assinatura no Brasil. Disponível em: <<http://www.diariodepernambuco.com.br/nota.asp?materia=20110328182203>>. Data de acesso em: 20 fev. 2011;
- UCHYIGIT, G., CLARK, K. An Agent Based Electronic Program Guide. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2002;
- VALENTE, J. A. A espiral da aprendizagem e as tecnologias da informação e comunicação: repensando conceitos. In JOLY, M. C. (Ed.) Tecnologia no ensino: implicações para a aprendizagem. São Paulo: Casa do Psicólogo Editora, p. 15-37, 2002;
- VOZALIS, E., MARGARITIS, K. Analysis of recommender systems algorithms. In: Proceedings of the Sixth Hellenic-European Conference on Computer Mathematics and its Applications - HERCMA 2003, 2003. Disponível em: <<http://macedonia.uom.gr/~mans/papiria/hercma2003.pdf>>. Data de acesso em: 28 ago. 2008;
- VUORIMAA, P. Digital television service architecture. Proceedings of IEEE International Conference on Multimedia and Expo, ICME2000, New York City, 2000;
- WITTEN, I. H.; FRANK, E., Data mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2th edition, 2005;
- XLETVIEW. XleTView SourceForge Project. Disponível em: <<http://sourceforge.net/projects/xletview>>. Data de acesso em: 18 set. 2010;
- YANG, H., PARTHASARANTHY, S. On the use of constrained association for Web log mining. In: Web Mining for Usage Patterns & User Profiles, WebKDD, 4., 2002;
- ZAIANE, O. R. Web Mining: Concepts, Practices and Research. In: Simpósio Brasileiro de Banco de Dados, SBBD, 15., João Pessoa, 2000;
- ZIMMERMAN, J., KURUPATI, K., PARAMESWARAN, L. Celebrity Recommender. In: AH'2002 Workshop on Personalization in Future TV. Málaga, Spain, 2002;

Apêndice A – Requisitos do Sistema

Neste apêndice abordaremos os requisitos do proposto sistema de recomendação de programas para a televisão digital. Faremos a especificação funcional e não funcional destes requisitos, detalhando os casos de uso e mostrando os diagramas oportunos para esta etapa do processo de desenvolvimento dentro da Engenharia de Software.

1. Especificação funcional de requisitos

O Capítulo 4 desta dissertação já apresentou os requisitos funcionais citados nas primeiras páginas deste apêndice. Entretanto, no intuito de construir uma “linha de raciocínio” sobre a especificação de requisitos, tomamos a liberdade de repetir algumas destas informações apenas no início desta subseção, já acrescentando outras informações relevantes, tanto nesta, como no decorrer das próximas seções. Seguem os quatro principais requisitos:

1.1 [RF 01] – Determinar Modelo do Telespectador (criar perfil)

Este requisito constitui-se como premissa para que a recomendação ocorra. Através da coleta de informações, de maneira implícita e/ou explícita, um modelo de preferências e hábitos do Usuário Telespectador é construído e armazenado no terminal de acesso (STB).

Este modelo possui informações presentes em duas bases de dados. A primeira contém informações estáticas, ao passo que a segunda contém informações dinâmicas. O preenchimento das informações estáticas é feito pelo usuário, através da criação de seu perfil. Quando isto ocorre, duas telas são disponibilizadas ao mesmo. Uma contém informações gerais, como Nome, Data de Nascimento, Localização geográfica, Estado civil, dentre outros. Uma segunda tela, mais específica para o estudo de caso, disponibiliza informações específicas sobre filmes.

Dessa forma, o preenchimento destes campos é realizado através de interface específica para esta funcionalidade e exibida sob demanda. Todos os campos são de preenchimento opcional. A presença de informações em alguns destes fará com que o módulo responsável por fazer a recomendação, antes de tal ação, consulte-os e utilize-os, como parâmetros de entrada.

Em relação aos dados dinâmicos, as informações a serem armazenadas correspondem às características (gênero, atores, diretor, sinopse) dos filmes mais assistidos pelo telespectador. Estas informações são extraídas pelo sistema através do uso do aparelho de TV

por parte do usuário. Toda a programação de filmes assistida pelo mesmo é armazenada e todas as informações sobre estes filmes que estão presentes na tela de cadastro de perfis também são devidamente armazenadas para posterior extração das Regras de Associação.

Desta forma, para a criação do Modelo do Telespectador, a cada semana o sistema analisa toda a programação assistida e, através do algoritmo Apriori, extrai todas as Regras de Associação possíveis. Além disso, utilizando o algoritmo TF-IDF, o sistema analisa todas as informações disponibilizadas pelo usuário (quando for o caso) nas telas de criação de seu perfil e, em seguida, analisa toda a programação para saber quais outros filmes deverão fazer parte da recomendação. Uma vez realizado estes dois processos, temos a criação do Modelo do Telespectador.

Mais uma vez torna-se necessário lembrar que a recomendação levará em conta as informações presentes nas duas bases de dados. Entretanto, a primeira a ser analisada é a de dados dinâmicos, tendo a de dados estáticos em caráter mandatório para a primeira. Mais detalhes serão ilustrados na especificação dos casos de uso nas próximas seções. Além disso, uma das informações presentes nas telas de criação de perfil é sobre o desejo ou não do usuário de receber recomendações. Além disso, o processo de avaliação da recomendação também é responsável por alterar as informações presentes no Modelo do Telespectador.

Casos de uso associados

UC 01 e UC 02.

Prioridade do requisito

Essencial.

1.2. [RF 02] – Criar perfil ‘Geral’

Bastante similar ao requisito anterior, em sua essência, este requisito permite que a aplicação realize recomendações de informação de programação sem que para isso haja a necessidade de que algum usuário, em algum momento, tenha solicitado a criação de algum perfil (coleta de dados explícita). Em outras palavras, trata-se da criação de um Modelo Geral do Telespectador.

Neste caso, toda vez que o terminal de acesso for ligado, o sistema armazenará, durante todas as horas de todos os dias da semana, as informações sobre a programação assistida naquele momento. Estas informações armazenadas serão posteriormente tratadas de forma similar aos casos em que o usuário cria seu próprio perfil. A diferença aqui é que as

informações que existem somente quando informadas pelo usuário na coleta explícita não serão tratadas neste caso, uma vez que as mesmas não foram preenchidas. Somente informações implícitas serão tratadas pelo processo de recomendação, através das técnicas de FBC e Regras de Associação, ou seja, através do processamento dos algoritmos Apriori e TF-IDF.

Este requisito não torna obrigatório o preenchimento/criação de um perfil por parte de um usuário e permite também que seja realizada uma recomendação única para todos os usuários, uma vez que sabemos que, em uma grande quantidade de lares, o aparelho de televisão é compartilhado por várias pessoas e estas, geralmente, assistem determinados tipos de programas dependendo do horário e/ou do dia em questão.

Casos de uso associados

UC 01 e UC 02.

Prioridade do requisito

Essencial.

1.3 [RF 03] – Recomendar programação

Este requisito é o principal desta proposta, uma vez que consiste em indicar para o usuário telespectador qual o programa mais recomendado, informações presente no Modelo do Telespectador, a ser assistido em um dado momento.

O processo de recomendação dar-se-á através do Guia Eletrônico de Programação. Todo programa considerado ‘recomendado’ ao telespectador, aparecerá no referido guia de forma destacada, sinalizando a este último a sua preferência por programas com tais características.

A cada semana, nas primeiras horas da segunda-feira, o Modelo do Telespectador é atualizado após processamento do algoritmo Apriori e do TF-IDF. Assim sendo, cada vez que o telespectador solicitar a visualização do Guia de Programação, este último, antes de ser exibido, terá seus filmes que tiverem sido considerados recomendados, destacados de forma a sinalizar a recomendação para o telespectador.

Vale informar que a visualização da recomendação poderá ser feita de duas maneiras:

- Através da solicitação de visualização do Guia de Programação;
- Através da solicitação de visualização da Lista de programas recomendados.

Casos de uso associados

UC 03, UC 05 e UC 06.

Prioridade do requisito

Essencial.

1.4 [RF 04] – Avaliar recomendação

Trata-se de um requisito não mandatório para a nossa proposta, entretanto, com uma grande valia para a melhoria contínua dos sistemas de recomendação e fortemente indicado pela literatura que aborda o tema.

A avaliação da recomendação é realizada pelo telespectador, através da solicitação de acesso a tal funcionalidade. Este visualizará uma lista de todos os programas recomendados para o mesmo durante a semana corrente. Uma das informações presentes nesta lista, e a única a qual o mesmo poderá preencher, é a relacionada com a pergunta se o usuário aprovou ou não a recomendação feita pelo sistema. Outra informação presente é sobre a sintonização ou não, por parte do usuário, do referido filme.

A recorrência de valor ‘Não’ na pergunta sobre a aprovação da recomendação deverá fazer com que o sistema não mais recomende o filme em questão ao usuário. Vale ressaltar, entretanto, que apenas o filme deixará de ser recomendado, e não demais filmes que possuam atributos semelhantes ao mesmo.

Outra informação presente nesta funcionalidade é o percentual de acerto do sistema. Esta informação, que ficará visível ao usuário telespectador é calculada da seguinte forma: para cada programa considerado recomendado, este poderá receber 3 valores. Caso o programa tenha sido sintonizado e aprovado, valor 100. Ocorrência de aprovação e ausência de sintonização ou vice-versa, valor 50. Caso o programa não tenha sido sintonizado e nem aprovado, valor nulo. Após a pontuação de todos os programas, é calculada a média aritmética e informado o percentual de acerto.

Casos de uso associados

UC 03 e UC 04.

Prioridade do requisito

Desejável.

2. Especificação não funcional de requisitos

Nesta seção é apresentada a especificação dos requisitos não funcionais, também conhecida como especificação suplementar de requisitos. A Especificação Suplementar consiste em capturar as necessidades do sistema que não são mapeadas nas especificações dos casos de uso, e que, de forma geral, lidam com aspectos voltados a qualidade. Abaixo, seguem os requisitos não funcionais para o sistema de recomendação proposto:

2.1 [RNF 01] – Sistemas Operacionais suportados

Devido às dependências do sistema a ser desenvolvido de acordo com a API Java, a aplicação aqui proposta deverá residir em um terminal de acesso dotado de um sistema operacional que possua uma JVM (Java Virtual Machine).

Como a aplicação possui arquitetura específica para a plataforma de televisão digital interativa, nossas especificações suplementares não prevêem a necessidade de portabilidade para sistemas operacionais Linux, Windows, Unix, dentre outros.

Casos de uso associados

Todos.

Prioridade do requisito

Essencial.

Categoria do requisito

Requisito de sistema.

2.2 [RNF 02] – Interatividade do terminal de acesso

Por tratar-se de uma aplicação onde existe a troca de informações entre o Usuário Telespectador e a própria aplicação, esta última deverá, então, estar residente em um terminal de acesso dotado de middleware interativo, com API de programação baseado na linguagem Java.

Além de ser dotado de middleware, este deverá, obrigatoriamente, ser ou o Ginga (modelo brasileiro), ou o MHP, do modelo europeu, ou o ARIB (japonês), uma vez que a aplicação aqui proposta é suportada pela plataforma Java, plataforma esta suportada pelos middlewares em questão.

Casos de uso associados

Todos.

Prioridade do requisito

Essencial.

Categoria do requisito

Requisito de sistema.

2.3 [RNF 03] – Inicialização automática

Levando em conta as características da aplicação, de acordo com as informações dos requisitos funcionais, esta deverá ter suas funcionalidades executadas assim que o terminal de acesso for ligado. A aplicação será executada como um *Thread* do sistema, possuindo baixíssima prioridade neste.

Vale destacar que as *Threads* serão utilizadas para a execução das seguintes funcionalidades:

- Preenchimento da base de dados de informações dinâmicas;
- Geração da lista de programas recomendados;
- Marcação, no Guia de Programação, dos programas que devem ser recomendados ao telespectador;
- Reavaliação da lista de programas recomendados, de acordo com a avaliação feita pelos usuários telespectadores.

Casos de uso associados

Todos.

Prioridade do requisito

Essencial.

Categoria do requisito

Requisito de sistema.

2.4 [RNF 04] – Dependência das Tabelas SI

Como a aplicação proposta faz recomendação de programas, e esta é feita baseada nos gêneros dos programas, fica evidente a necessidade de que haja a transmissão das tabelas de

SI (Service Information) e posterior processamento destas por parte do terminal de acesso para que, enfim, ocorra a recomendação.

A impossibilidade de manipulação das referidas tabelas inviabiliza, completamente, tanto a geração das informações de programação, como a recomendação propriamente dita destas.

Casos de uso associados

Todos.

Prioridade do requisito

Essencial.

Categoria do requisito

Requisito de sistema.

2.5 [RNF 05] – Dependência do Guia de Programação

Nossa proposta prevê que a única maneira de representar a recomendação ao usuário é através do Guia de Programação. Portanto, um terminal de acesso não dotado de uma aplicação EPG não permitirá que haja recomendação de programação.

Casos de uso associados

Todos.

Prioridade do requisito

Essencial.

Categoria do requisito

Requisito de sistema.

2.6 [RNF 06] – Controle remoto ergonômico

O controle remoto deve possuir um design capaz de trazer conforto ao Usuário Telespectador. Deve possuir formato, dimensões e disposições dos botões que se adequem de forma satisfatória à mão humana.

Ainda, é desejável que este possua teclas numéricas, haja vista estarmos prevendo a entrada de dados do tipo ‘Nome do telespectador’. Entretanto, a não obrigatoriedade de

preenchimento destas informações não inviabilizará a proposta recomendação de programas em um controle remoto tradicional.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de ergonomia.

2.7 [RNF 07] – Tamanho da aplicação

Tanto a aplicação como as informações a serem armazenadas por conta desta deverão ocupar pouco espaço no sistema de armazenamento do terminal de acesso. Juntas, estas deverão ocupar um espaço total inferior a 300 Kilobytes. Vale destacar que essa premissa visa o não comprometimento de demais aplicações que possuem maior prioridade no terminal de acesso.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de desempenho.

2.8 [RNF 08] – Processamento da aplicação

As *Threads* a serem executadas, bem como todo o restante do processamento de informações decorrentes da especificação dos casos de uso da aplicação proposta, em um dado instante, o somatório das que estiverem executando não deverá corresponder a um consumo de processamento superior a 5 (cinco) por cento da capacidade total do processador do terminal de acesso.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de desempenho.

2.9 [RNF 09] – Tempo de resposta

Além de atender às restrições de processamento ditas acima, a aplicação, para os momentos em que seu processamento ocorre devido a uma solicitação do telespectador, deverá apresentar um tempo de resposta para a solicitação inferior a 5 (cinco) segundos.

Casos de uso associados

UC 03, UC 05 e UC 06.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de desempenho.

2.10 [RNF 10] – Integridade das informações

A integridade das informações que circulam pelos módulos arquiteturais desta proposta deverá ser preservada a todo momento, não sendo permitida a quebra desta integridade nem por módulos desta arquitetura e nem por módulos ou componentes da arquitetura de outras aplicações.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de confiabilidade.

2.11 [RNF 11] – Execução segura

Mesmo sabendo que os dados que são processados pelos módulos da arquitetura desta aplicação não representam informações de alto caráter de confidencialidade, ainda assim, os processos executados pelo sistema de recomendação aqui proposto não deverão corromper os dados presentes em nenhuma das aplicações presentes no terminal de acesso.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de segurança.

2.12 [RNF 12] – Controle de permissões

A aplicação proposta não poderá carregar, inicializar, executar, parar ou pausar qualquer outra aplicação. Além disso, por tratar-se de uma aplicação com baixa prioridade no sistema, não é permitido a esta concorrer com recursos compartilhados do sistema.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de segurança.

2.13 [RNF 13] – Tolerância a falhas

A aplicação deverá continuar em execução, de forma confiável, mesmo após a ocorrência de erros internos e de reinicialização abrupta do terminal de acesso.

Casos de uso associados

Todos.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de robustez.

2.14 [RNF 14] – Interface intuitiva

A aplicação deve ser de simples utilização, contando com o auxílio de ícones, figuras e textos explicativos. A disposição das informações e a maneira de acessá-las devem ser provisionadas de maneira que permita acessibilidade às pessoas de diferentes níveis de familiaridade com a tecnologia.

Casos de uso associados

UC 01, UC 03, UC 04, UC 05, UC 06.

Prioridade do requisito

Importante.

Categoria do requisito

Requisito de usabilidade.

3. Modelo de casos de uso

O Modelo de Caso de Uso deve servir como um meio de comunicação e uma espécie de contrato entre o cliente, os usuários e os desenvolvedores do sistema. A modelagem responsável pela construção do citado modelo visa, através da exibição dos casos de uso e atores, mostrar o completo relacionamento existentes entre estes, permitindo uma visão geral sobre a interação entre atores, requisitos (casos de uso), e aplicações externas.

A figura 1 apresenta o modelo de casos de uso para o nosso proposto sistema de recomendação:

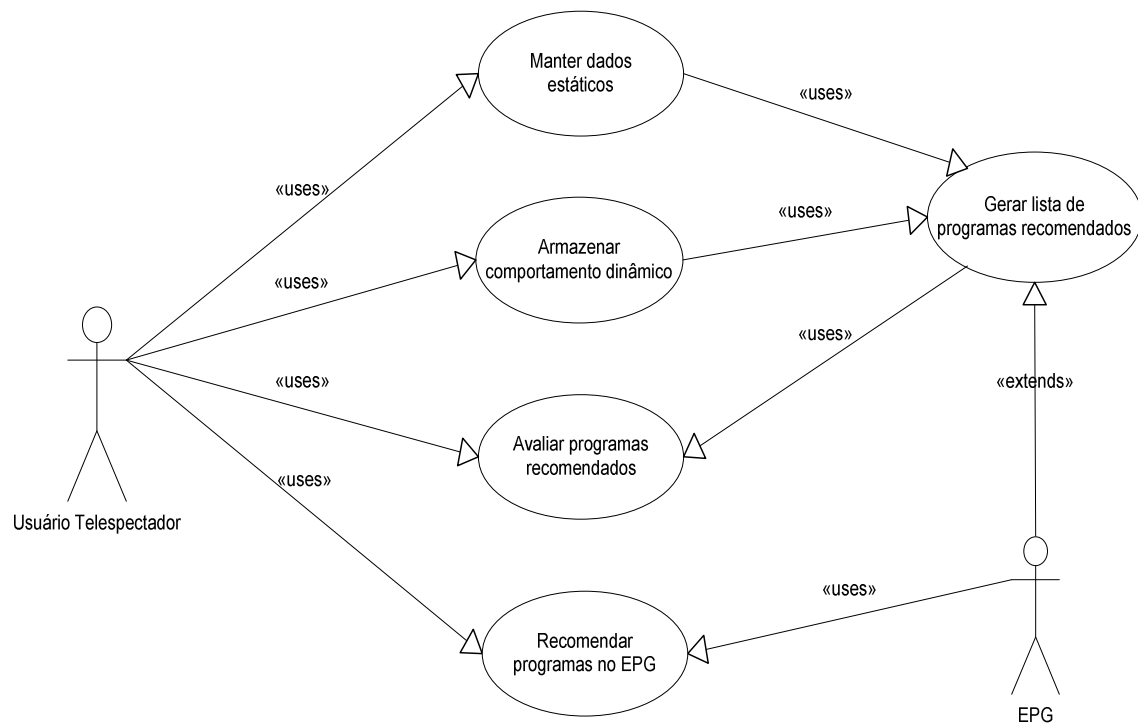


Figura A-1: Modelo de casos de uso da aplicação.

3.1 Casos de uso

A especificação dos casos de uso de um sistema é uma maneira eficiente de expressar os requisitos funcionais deste de uma maneira mais detalhada. Os requisitos funcionais explicitam, de forma geral e com linguagem de alto nível, o que um determinado sistema deve fazer. Os casos de uso, por sua vez, embora ainda utilizem linguagem de alto nível, realizam um detalhamento sobre determinada funcionalidade.

Em termos de prioridade, usamos a mesma utilizada para os requisitos funcionais e não funcionais. Abaixo, segue a especificação dos casos de uso do sistema de recomendação proposto nesta dissertação:

3.1.1 [UC 01] – Manter dados estáticos

Descrição

Este caso de uso permite que as informações estáticas do Usuário Telespectador possam ser incluídas, excluídas, alteradas ou consultadas, por este último, permitindo, assim, o preenchimento da tabela de banco de dados específica para estas informações, que posteriormente serão utilizadas para a criação do Modelo do Telespectador.

Atores

Usuário telespectador.

Pré-condições

Não se aplica.

Pós-condições

As informações estáticas do usuário telespectador são inseridas, excluídas, consultadas ou alteradas.

Fluxo básico

1. O caso de uso inicia quando o Usuário Telespectador necessita fazer a manutenção (inclusão, alteração, exclusão ou consulta) de seus dados estáticos (criação de seu perfil);
2. De acordo com o tipo de operação de manutenção desejado pelo Usuário Telespectador, um dos subfluxos é executado;
 - a. Se o Usuário Telespectador deseja incluir seus dados estáticos, o subfluxo ‘Incluir’ é executado.
 - b. Se o Usuário Telespectador deseja alterar as informações dos seus dados estáticos já cadastrados, o subfluxo ‘Alterar’ é executado.
 - c. Se o Usuário Telespectador deseja excluir seus dados estáticos já cadastrados, o subfluxo ‘Remover’ é executado.
 - d. Se o Usuário Telespectador deseja consultar informações sobre seus dados estáticos, o subfluxo ‘Consultar’ é executado.

Subfluxo Incluir

1. Este subfluxo inicia quando o Usuário Telespectador solicita incluir as informações referentes aos seus dados estáticos;

2. O sistema solicita ao Usuário Telespectador o preenchimento dos atributos presentes nas telas de criação de perfil;
3. O Usuário Telespectador preenche os atributos acima e confirma a inclusão;
4. O sistema realiza a inclusão dos dados informados pelo Usuário Telespectador no passo 3;
5. O sistema exibe uma mensagem informando que a inclusão dos dados estáticos do Usuário Telespectador foi efetivada com sucesso.

Subfluxo Alterar

1. Este subfluxo inicia quando o Usuário Telespectador solicita alterar as informações referentes aos seus dados estáticos;
2. O sistema solicita ao Usuário Telespectador a alteração dos atributos presentes nas telas de criação de perfil;
3. O Usuário Telespectador altera qualquer um dos atributos acima e confirma a alteração;
4. O sistema realiza a alteração dos dados informados pelo Usuário Telespectador no passo 3;
5. O sistema exibe uma mensagem informando que a alteração dos dados estáticos do Usuário Telespectador foi efetivada com sucesso.

Subfluxo Remover

1. Este subfluxo inicia quando o Usuário Telespectador solicita excluir as informações referentes aos seus dados estáticos;
2. O Usuário Telespectador solicita a exclusão dos dados;
3. O sistema solicita a confirmação para exclusão;
4. O Usuário Telespectador confirma a exclusão;
5. O sistema remove os dados e exibe mensagem informando que a remoção dos dados estáticos foi efetuada com sucesso.

Subfluxo Consultar

1. Este subfluxo inicia quando o Usuário Telespectador solicita consultar as informações referentes aos seus dados estáticos;

- O sistema apresenta os respectivos valores para os atributos estáticos presentes em seu perfil;

Validações e regras de negócio

- Esta regra se aplica a todos os subfluxos. **Atributos com valores não permitidos.** Se algum atributo for preenchido com valor não permitido, o sistema não completará a operação e notificará ao Usuário Telespectador a ocorrência, solicitando o correto preenchimento das informações.

Fluxos alternativos

FA 01.

Fluxos de exceção

FE 01, FE 02 e FE 03.

Diagrama de casos de uso

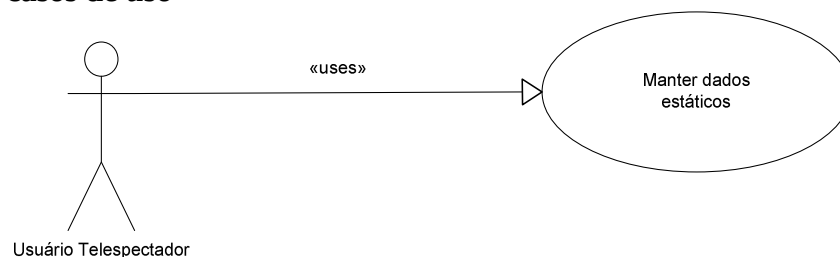


Figura A-2: Diagrama de caso de uso do UC 01.

3.1.2 [UC 02] – Armazenar comportamento dinâmico

Descrição

Este caso de uso tem a finalidade de armazenar, através do comportamento dinâmico do Usuário Telespectador, as informações preferenciais acerca dos atributos dos filmes assistidos pelos usuários.

Atores

Usuário Telespectador.

Pré-condições

Não se aplica.

Pós-condições

As informações dinâmicas do Usuário Telespectador são incluídas.

Fluxo básico

1. O caso de uso inicia a cada vez que o Usuário Telespectador iniciar o seu terminal de acesso;
2. O sistema armazena, em uma tabela temporária, os atributos do filme que está sendo exibido;
3. A cada mudança na sintonização dos canais televisivos, o sistema armazena, na mesma tabela temporária, a quantidade de tempo que o usuário passou assistindo a determinado filme. Somente se este tempo for superior a quinze minutos, os atributos deste filme ficarão armazenados no intuito a compor, posteriormente, o Modelo do Telespectador;
4. Todos estes passos são executados enquanto o terminal de acesso permanecer ligado;
5. Ao desligar o terminal de acesso, o caso de uso é encerrado.

Fluxos alternativos

Não se aplica.

Fluxos de exceção

FE 01, FE 02 e FE 03.

Diagrama de casos de uso

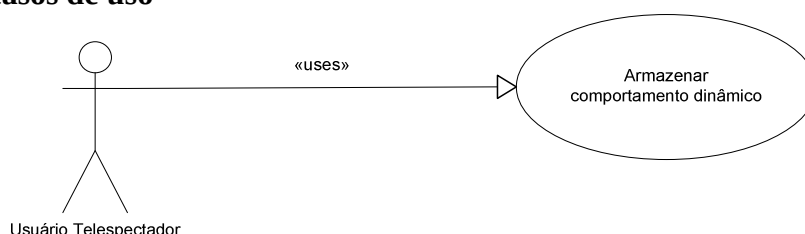


Figura A-3: Diagrama de caso de uso do UC 02.

3.1.3 [UC 03] – Gerar lista de programas recomendados

Descrição

Este caso de uso tem a finalidade de gerar, sempre no início de cada semana, uma lista atualizada, baseada no comportamento do Usuário Telespectador na semana anterior, de todos os programas (filmes) considerados recomendados.

É através deste caso de uso que é possível a realização da recomendação propriamente dita, uma vez que é a partir deste que o EPG realiza a marcação dos programas considerados recomendados. Além disso, a partir deste caso de uso é que se torna possível ao Usuário Telespectador a avaliação do referido sistema de recomendação.

Atores

- Usuário telespectador;
- EPG.

Pré-condições

- O Modelo do Telespectador já deverá estar disponível, através do preenchimento das tabelas contendo as informações dinâmicas e estáticas deste;
- O Guia de Programação já deverá ter sido gerado contendo toda a programação da semana corrente.

Pós-condições

O sistema passa a ter a lista de programas a serem recomendados, para a semana corrente, totalmente preenchida, possibilitando, então, a ocorrência da recomendação.

Fluxo básico

1. O caso de uso inicia nas primeiras horas do domingo de cada semana;
2. O sistema percorre a base de dados contendo o Modelo do Telespectador, onde estas conterão as informações inseridas pelo usuário através da criação de seu perfil e as informações geradas pelo processamento dos algoritmos TF-IDF e Apriori;
3. Com a disponibilização das tabelas SI e posterior montagem do Guia de Programação, o sistema irá percorrer todo o guia, verificando toda a programação de todos os canais. Em cada programa de cada canal (de toda

a semana), o sistema verificará se os atributos do mesmo coincidem com as informações presentes no Modelo do Telespectador. Em caso afirmativo, o programa será indicado como recomendado, recebendo este flag positivo na tabela que armazena a programação televisiva;

4. Desta forma, na aplicação EPG (Guia de Programação), a tabela contendo a programação da semana terá seu campo ‘Indicado como Favorito’, atualizado para todos os programas da semana em questão. Além disso, o campo booleano, da mesma tabela, denominado ‘Recomendação atualizada’, terá seu valor setado como *true*.

Fluxos alternativos

FA 02.

Fluxos de exceção

FE 01, FE 02 e FE 03.

Diagrama de casos de uso

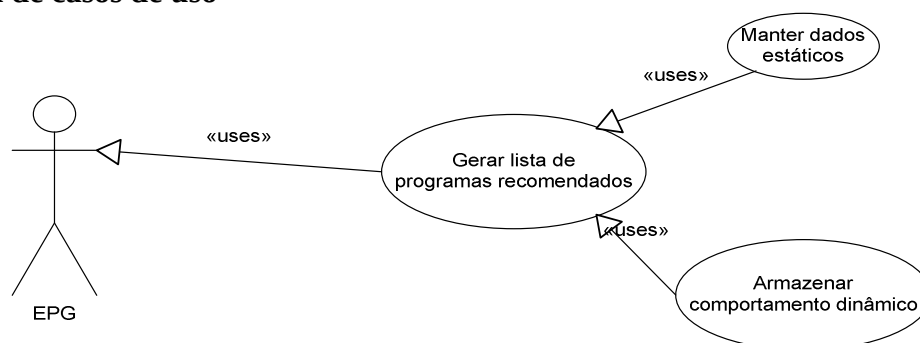


Figura A-4: Diagrama de caso de uso do UC 03.

3.1.4 [UC 04] – Avaliar programas recomendados

Descrição

Este caso de uso tem a finalidade de permitir que o sistema de recomendação aqui proposto possa ser avaliado pelo Usuário Telespectador, tanto de forma direta, como indireta.

A avaliação de forma direta ocorre através da disponibilização, por parte do Usuário Telespectador, de sua opinião sobre o referido sistema, ao passo que, a forma indireta, é realizada através da análise, por parte do próprio sistema, se houve ou não sintonização do recomendado programa.

Atores

Usuário Telespectador.

Pré-condições

- A lista de programas recomendados já deve estar completamente preenchida.

Pós-condições

O sistema de recomendação possuirá seu percentual de avaliação atualizado.

Fluxo básico

1. O caso de uso inicia quando o Usuário Telespectador solicita realizar a avaliação do sistema de recomendação aqui proposto;
2. O sistema exibirá a lista de todos os programas recomendados durante a semana, disponibilizando, a cada programa recomendado, para preenchimento com valores 'Sim' ou 'Não', o campo 'Recomendação aprovada';
3. O telespectador preenche a avaliação e solicita calcular o percentual de acerto;
4. O sistema analisa os valores dos campos 'Recomendação aprovada' e 'Programa sintonizado' para, de acordo com a proposta de cálculo da avaliação presente no capítulo seguinte desta dissertação, disponibilizar um percentual de acerto, correspondente à avaliação do sistema de recomendação proposto;
5. O valor da avaliação é disponibilizado para o telespectador;
6. O caso de uso é encerrado.

Validações e regras de negócio

Esta regra se aplica a todos os subfluxos. **Atributos com valores não permitidos.** Se algum atributo for preenchido com valor não permitido, o sistema não completará a operação e notificará ao Usuário Telespectador a ocorrência, solicitando o correto preenchimento das informações.

Fluxos alternativos

FA 01.

Fluxos de exceção

FE 01, FE 02 e FE 03.

Diagrama de casos de uso



Figura A-5: Diagrama de caso de uso do UC 04.

3.1.5 [UC 05] – Recomendar programas no EPG

Descrição

Este caso de uso tem a finalidade de, a partir do Guia de Programação, exibir para o Usuário Telespectador os programas considerados Recomendados para este.

Atores

- Usuário telespectador;
- EPG.

Pré-condições

- O Modelo do Telespectador já deverá estar disponível, através do preenchimento das tabelas contendo as informações dinâmicas e estáticas deste;
- A lista de programas recomendados já deve estar completamente preenchida;
- O Guia de Programação já deverá ter sido gerado contendo toda a programação da semana corrente.

Pós-condições

O EPG geral será exibido ao telespectador destacando os programas recomendados a este.

Fluxo básico

1. O caso de uso inicia quando o Usuário Telespectador solicita a visualização do Guia de Programação geral;

2. O sistema, antes de disponibilizar o guia, consultará, para cada programa a dispor na tela do telespectador, o valor do campo 'Indicado como Favorito', na tabela contendo a programação da semana;
3. Para o programa com valor **true** no referido campo, este aparecerá de forma destacada no guia geral da programação da semana;
4. O Guia de Programação geral é exibido ao usuário destacando os programas recomendados;
5. O caso de uso é encerrado.

Fluxos alternativos

Não se aplica.

Fluxos de exceção

FE 01, FE 02 e FE 03.

Diagrama de casos de uso

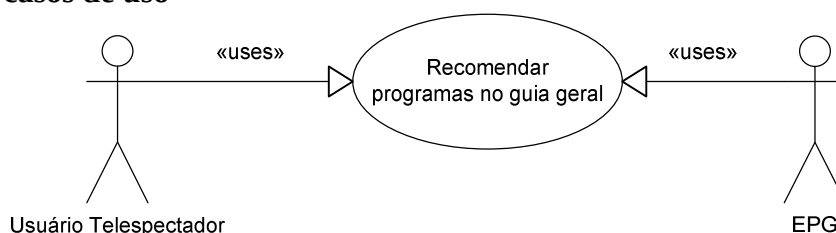


Figura A-6: Diagrama de caso de uso do UC 05.

3.2 Fluxos alternativos

Em seguida, seguem os principais fluxos alternativos citados durante a especificação dos casos de uso:

3.2.1 Fluxos alternativo FA 01 - Cancelamento de operação

1. O Usuário Telespectador solicita cancelar a operação;
2. O sistema solicita confirmação de cancelamento da operação;
3. O Usuário Telespectador confirma o cancelamento;
4. O caso de uso é encerrado.

3.2.2 Fluxo alternativo FA 02 – Terminal de acesso desligado abruptamente

1. Este fluxo alternativo inicia quando, durante a geração da lista de programas recomendados, o terminal de acesso é desligado abruptamente;

2. O sistema consulta a lista de programas recomendados e verifica que, embora exista a informação de programas recomendados nesta, a tabela contendo a programação da semana possui o campo 'Recomendação atualizada' setada como false;
3. A geração da lista de programas recomendados é iniciada novamente;
4. O caso de uso é encerrado.

3.3 Fluxos de exceção

Em seguida, seguem os principais fluxos de exceção citados durante a especificação dos casos de uso:

3.3.1 Fluxo de exceção FE 01 – Falha no acesso à base de dados

1. Este fluxo de exceção inicia quando há uma falha no acesso à base de dados;
2. A aplicação propaga a mensagem de erro retornada pelo sistema;
3. O caso de uso é encerrado.

3.3.2 Fluxo de exceção FE 02 – Falha no acesso a componentes externos

1. Este fluxo de exceção inicia quando há uma falha no acesso a componentes externos;
2. A aplicação propaga a mensagem de erro retornada pelo sistema;
3. O caso de uso é encerrado.

3.3.3 Fluxo de exceção FE 03 – Falha desconhecida

1. Este fluxo de exceção inicia quando há uma falha desconhecida no sistema;
2. A aplicação propaga a mensagem de erro retornada pelo sistema;
3. O caso de uso é encerrado.