

# Abordagem Combinada para Recomendação Personalizada Utilizando o Guia de Programação Eletrônico

**Mário Alves de Moraes Neto, Diego Sá Cardoso, Cidcley Teixeira de Souza,  
Mariela Inés Cortés**

Mestrado Acadêmico em Ciência da Computação

Universidade Estadual do Ceará

Av. Paranjana, 1700 –60740-000 – Fortaleza – Brasil

[mario@itti.cefetce.br](mailto:mario@itti.cefetce.br), [{diegoccx, cidcley}@gmail.com](mailto:{diegoccx, cidcley}@gmail.com),  
[mariela@larces.uece.br](mailto:mariela@larces.uece.br)

*Abstract. The digitization of TV signals allows a wide range of channels can be transmitted, allowing users find a range of television programming so bulky. Using of the benefits of Recommender Systems, this paper proposes the use of techniques Content-Based Filtering (CBF) and Association Rules to recommend programs in the context of interactive digital television. The system presented here, which allows you to create multiple profiles and use the format of an Electronic Programming Guide (EPG), also allows the User to evaluate the recommendations, enabling the constant improvement of the recommendation system in question.*

*Resumo. 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. 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.*

## 1. Introdução

A televisão digital é uma realidade que apresenta inúmeras vantagens quando comparada aos sistemas analógicos (BECKER; MONTEZ, 2004). A interatividade nesta mídia acontece através da transmissão de áudio e vídeo, e, adicionalmente, o envio de fluxos de dados. Nesse último fluxo, também são transmitidos, na forma de tabelas, informações detalhadas dos fluxos de áudio, vídeo e dados, recebendo a denominação de tabelas SI (ABNT NBR 15603-1, 2008). Estas foram criadas para suprir as necessidades criadas pelo novo sistema digital, onde se torna necessária a informação sobre os programas/canais transmitidos pela mesma. Estes dados podem ser considerados metadados de programação.

No que se refere à programação de televisão, os metadados permitem a construção de aplicações chamadas Guia Eletrônico de Programação (EPGs), que são aplicações baseadas em metadados, armazenando informações pertinentes a certa mídia, provendo a listagem de toda a programação e conteúdo de serviço interativo no televisor digital habilitado para isso (FERNANDES; LEMOS; SILVEIRA, 2004).

Os sistemas digitais de transmissão de sinal de TV contam com o benefício de permitir a transmissão de uma enorme gama de canais televisivos, aumentando, consideravelmente, a oferta de programas para os usuários telespectadores. Entretanto, ao passo que isto representa uma vantagem, os telespectadores podem ficar confusos diante uma oferta de programação televisiva grande o suficiente para deixar-lhe sem ter noção sobre a escolha do programa mais adequado. Neste contexto, os Sistemas de Recomendação constituem-se numa tentativa de solução para problema da sobrecarga de informação, ajudando o usuário na procura de informações relevantes.

Inúmeras técnicas podem ser utilizadas, de forma isolada ou combinada, para que as recomendações sejam geradas de forma satisfatória (REATEGUI; CAZELLA, 2009). Dentre elas a técnica de 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. Outra técnica baseada em mineração de dados considera o uso das Regras de Associação (AGRAWAL; IMIELINSKI; SWAMI, 1993; AGRAWAL et al., 1996), cujo objetivo é descobrir padrões escondidos de forma automática, 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*).

O objetivo deste trabalho é apresentar um sistema de recomendação de programas televisivos, com base nas duas técnicas citadas anteriormente, cuja funcionalidade principal é recomendar (sugerir) programas em sintonia com as preferências do usuário telespectador. As preferências podem ser explicitamente informadas pelo usuário, através da criação de um perfil, ou “descobertas pela aplicação”, de maneira implícita. A recomendação gerada é apresentada através do Guia de Programação Eletrônico (*EPG – Electronic Programming Guide*). Além disso, o sistema prevê um módulo de avaliação da recomendação que, além de informar o percentual de acerto, verifica o retorno do usuário e, através disso, permite que o sistema aprimore a sua recomendação.

## 2. Trabalhos relacionados

Esta seção aborda a apresentação e análise de alguns trabalhos de pesquisa no contexto do desenvolvimento de sistemas de recomendação. 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 a técnica de filtragem de informação utilizando a abordagem colaborativa e também realiza a atualização dinâmica do perfil. O trabalho não apresenta preocupação com os aspectos da interatividade na televisão digital e a técnica utilizada é bastante simples, não tendo havido a preocupação na geração de estratégias de recomendação mais elaboradas.

A preocupação com a sobrecarga de informação, sobretudo no que se refere aos inúmeros programas de televisão numa plataforma digital e interativa, também foi tratada em (BUCZAK; ZIMMERMAN; KURUPATI, 2002). Os autores 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 Vídeo 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.

Uma abordagem baseada em mineração de dados em metadados de programação no processo de recomendação de conteúdos televisivos é apresentada em Smyth, Wilson e Sullivan (2002). Infelizmente, os detalhes acerca da estratégia de recomendação não foram abordados, assim como a eficácia da mesma não foi comprovada através de implementação/estudo de caso. Não foram abordados detalhes da preocupação com a realidade da televisão digital. Talvez seja uma proposta apenas adaptável para a web.

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. Este último obteve melhor desempenho (72%), fator que determinou a escolha do mesmo na abordagem aqui proposta. 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.

### **3. Requisitos do sistema de recomendação**

Nesta seção abordaremos os requisitos funcionais do sistema de recomendação de programas proposto.

#### **3.1 [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 explícita, um modelo de preferências e hábitos do Usuário Telespectador é construído e armazenado no terminal de acesso (STB – *Set-Top-Box*, em inglês). Este modelo possui informações presentes

em duas bases de dados. A primeira contém informações estáticas a partir da criação do perfil do usuário. A segunda base de dados contém informações geradas de forma implícita, de acordo com a utilização do sistema por parte do usuário, à medida que o usuário interage com a programação televisiva. A partir daí, a técnica de mineração de dados chamada Regras de Associação, através do algoritmo Apriori, encontrará demais informações sobre preferências do usuário em questão. Estas informações, juntamente com as informações inseridas pelo usuário nas telas de criação de perfis, comporão o Modelo do Telespectador.

### **3.2 [RF 02] – Criar perfil ‘Geral’**

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).

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, conforme citado posteriormente.

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.

### **3.3 [RF 03] – Recomendar programação**

Este requisito representa a funcionalidade central do sistema e consiste em elaborar a recomendação para o usuário, indicando os programas que mais se adequam de acordo com o Modelo de Telespectador. A partir da base de programação, o sistema verifica quais programas serão recomendados.

A programação recomendada é apresentada através do Guia Eletrônico de Programação, onde todo programa considerado ‘recomendado’ ao telespectador, aparecerá de forma destacada.

### **3.4 [RF 04] – Avaliar recomendação**

A avaliação da recomendação é realizada de forma explícita pelo telespectador, através da solicitação de acesso a tal funcionalidade. A partir da lista de todos os programas que foram recomendados, o telespectador pode aprovar ou não cada uma das recomendações realizadas. A não aprovação de uma recomendação, implica na remoção do programa em questão da recomendação. Adicionalmente, de acordo com informação presente nos metadados de programação, o sistema determina se houve ou não a sintonização no canal de cada programa considerado recomendado. Devido às restrições de tamanho do artigo são apresentados apenas os requisitos funcionais, embora já

tenhamos realizado a especificação não funcional destes requisitos e o detalhamento dos casos de uso, bem como a elaboração dos referidos diagramas.

#### 4. Arquitetura e implementação do sistema

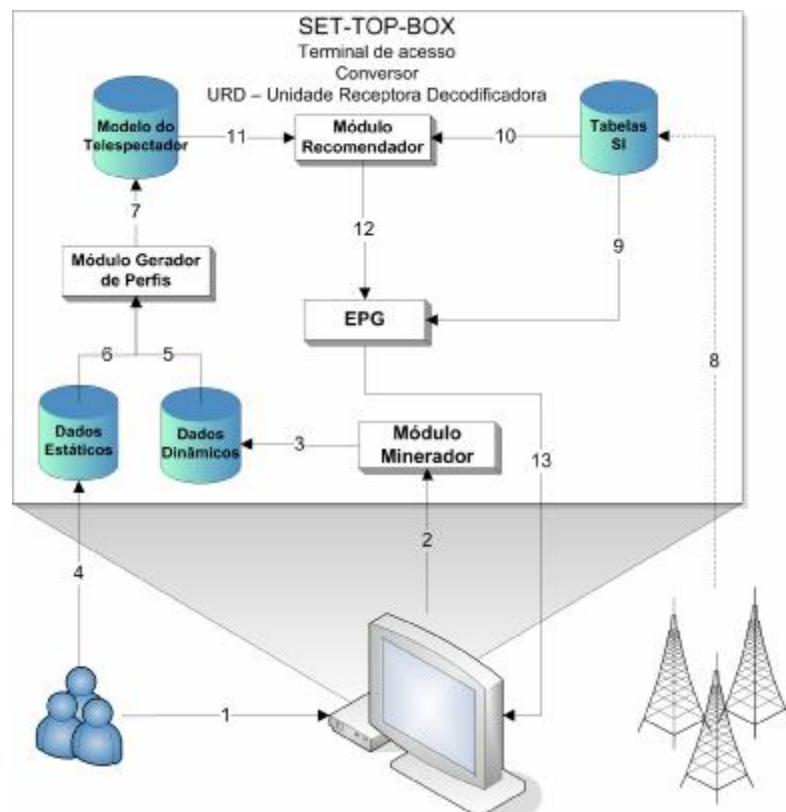
Diferentemente das abordagens seguidas nos trabalhos relacionados apresentados anteriormente, a abordagem proposta no presente trabalho utiliza as técnicas de Filtragem Baseada em Conteúdo e das Regras de Associação de forma combinada, apresentando os resultados da recomendação de programação televisiva através de um Guia de Programação. Adicionalmente, dois outros fatores foram associados a esta combinação de técnicas: as informações presentes no perfil do usuário e as informações fornecidas pelo mesmo durante o processo de avaliação da recomendação. A arquitetura do sistema de recomendação proposto (ilustrada Figura 1) envolve os seguintes componentes:

- **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:** Contem as regras de associação geradas pelo algoritmo Apriori, utilizando como entrada a programação assistida pelo usuário. A partir dessa entrada, o algoritmo determina relacionamentos implícitos relacionados ao domínio da aplicação televisiva. O algoritmo é customizado com os valores<sup>1</sup> de 90% (noventa por cento) para o Fator Suporte (freqüência 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 o uso do aparelho de televisão por parte do Usuário Telespectador, realizando o processamento mencionado no componente anterior, cujo resultado é o documento XML contendo as regras de associação armazenadas na Base de Dados dinâmicos;
- **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, que funciona como um balizador de informações, e as insere na Base de dados do Modelo do Telespectador. Funcionando como um 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, que aqui chamamos de Modelo do Telespectador;

---

<sup>1</sup> Valores de referencia extraidos da literatura (AGRAWAL et al., 1996).

- **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;
- **EPG:** Componente responsável pela geração do Guia de Programação. É nele que a recomendação será exibida;
- **Módulo Recomendador:** Componente responsável por analisar todas as informações de programação, presentes na Base de dados de programação, e, através da comparação com as informações presentes na Base de dados do Modelo do Telespectador, realizar a recomendação dos programas considerados favoritos. Conforme já mencionado, a estratégia de recomendação executada por este módulo consiste em utilizar o algoritmo TF-IDF para pesquisar, na base de programação, todos os atributos que “casam” com as informações presentes no Modelo do Telespectador, inclusive as demais informações contidas nas regras de associação geradas pelo Apriori.



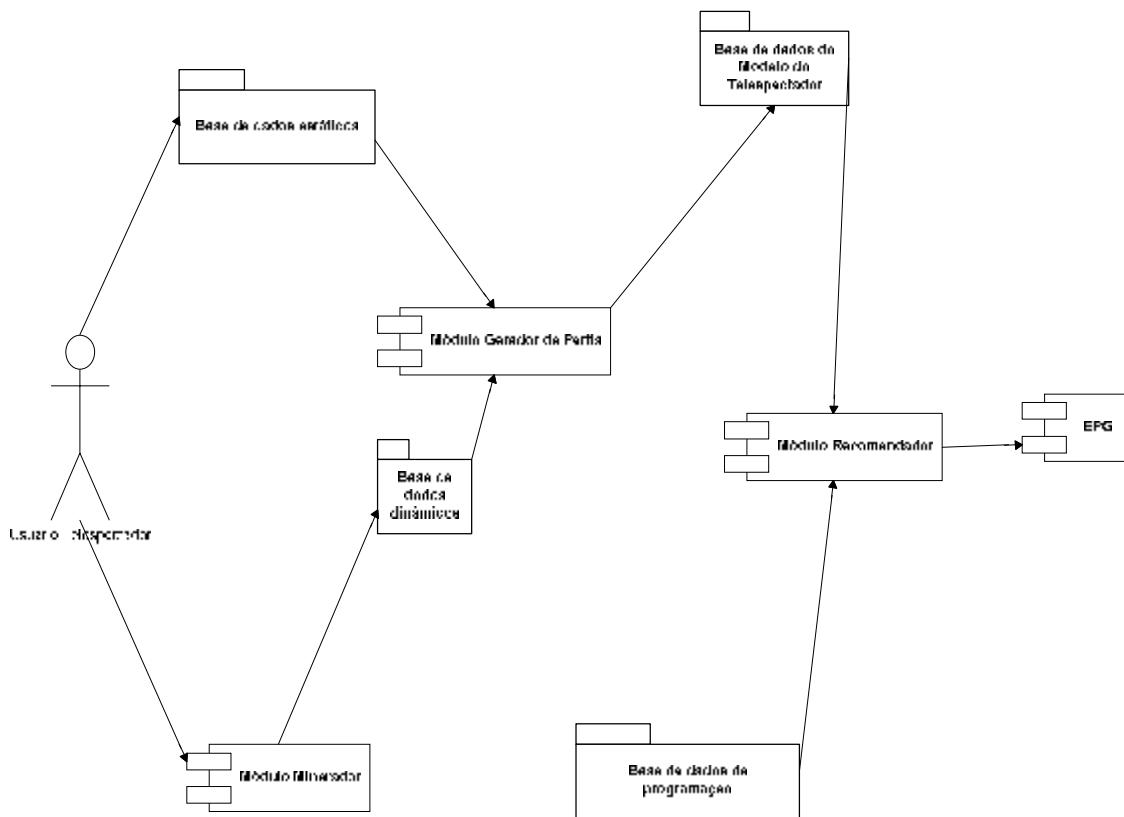
**Figura 1. Arquitetura geral da aplicação.**

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). Além disso, alguns dos autores deste artigo participaram do projeto SBTVD, no ano de 2005, ocasião na qual alguns estudos visando a temática 'Metadados' foram realizados e estes defendiam e utilizavam, em suas propostas, o uso do XML em aplicações voltadas para a TV Digital.

A numeração dada aos elementos de ligação entre os componentes tem o intuito de, ainda em nível arquitetural, mostrar os passos envolvidos para a recomendação em um cenário geral, de forma a tornar mais clara a solução proposta.

1. Usuário Telespectador interage com a programação televisiva;
2. O Módulo Minerador analisa esta interação no intuito de extrair as informações acerca dos gêneros preferidos;
3. Estas informações analisadas são inseridas pelo Módulo Minerador na Base de dados dinâmicos;
4. Usuário Telespectador solicita o preenchimento de suas informações e estas são armazenadas na Base de dados estáticos;
5. Informações dinâmicas são analisadas pelo Módulo Gerador de Perfis;
6. Informações estáticas são analisadas pelo Módulo Gerador de Perfis;
7. Módulo Gerador de Perfis analisa as bases de dados citadas logo acima e cria um Modelo do Telespectador, armazenando informações sobre suas preferências em relação aos gêneros de programação;
8. As informações de programação são transmitidas, através das tabelas SI, pelas estações transmissoras e são armazenadas na Base de dados de programação;
9. Informações de programação são interpretadas para a geração do EPG;
10. O Módulo Recomendador analisa as informações de programação;
11. O Módulo Recomendador analisa as informações presentes na Base de dados do Modelo do Telespectador;
12. O Módulo Recomendador, utilizando a técnica de Regras de Associação e FBC, verifica as preferências do usuário, presentes no Modelo do Telespectador para, a partir da programação disponibilizada da Base de dados de programação, verificar quais são os programas que deverão ser considerados recomendados. Uma vez gerada a lista de programas recomendados, este componente realiza a marcação dos programas considerados recomendados;
13. O EPG possui seus programas, considerados favoritos, marcados e disponíveis para visualização por parte dos usuários.

Em seguida, segue diagrama de componentes exibindo a troca de informações existentes entre os módulos da aplicação:



**Figura 2. Diagrama de componentes da aplicação.**

A análise dos componentes da aplicação nos permite chegar à conclusão da grande importância exercida pelo módulo minerador, uma vez que o mesmo tem a função de extrair os padrões não triviais de comportamento do usuário defronte do aparelho de TV. Abaixo transcreveremos trecho de código da classe AprioriRecomendação, classe esta que utiliza o algoritmo Apriori para, uma vez de posse da programação assistida pelo usuário e com as informações de seu perfil, caso estas estejam disponíveis, gerar a recomendação. O método aprioriGenRecomendação representa o principal método responsável pela geração desta recomendação:

```

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);
                }
            }
        }
        adicionarAtores(funcoes.colecionaStrings(programaAtual.getAtores()),
        programaAtual.getIdPrograma());
    }
}

```

```

        adicionarDiretor(programaAtual.getDiretores(), programaAtual.getDiretores());
    }
    if (programaAtual.getTempoAssistido() > 5)
        programaAnterior = programaAtual;
}
}

```

## 4.1 Ambiente de desenvolvimento

É importante comentar, neste trabalho, que houve, também, a preocupação em escolher um conjunto de ferramentas que melhor se adequasse à nossa proposta e permitisse que a mesma saísse do nível arquitetural e se transformasse em códigos de programação. Assim sendo, abaixo seguem as ferramentas utilizadas no desenvolvimento dos três módulos arquitetrais em questão:

**Tabela 1: Relação das ferramentas utilizadas.**

Nome da ferramenta	Tipo	Finalidade na proposta
Java Sun SDK	Linguagem e compilador Java	Linguagem Java; Compilador Java.
Eclipse	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	Parser SAX XML	Criação de documentos XML a partir de código Java; Manipulação de documentos XML a partir de código Java.
XleTView	Emulador de Xlets	Testes dos Xlets desenvolvidos, simulando o ambiente de televisão digital interativa;
dsmcc-mhp-tools	Gerador de carrossel	Geração dos dados de programação para realização das simulações do processo de recomendação.

## 5. Estudo de caso

O estudo de caso tem por objetivo principal ilustrar o protótipo do sistema de recomendação desenvolvido, de acordo com a abordagem apresentada. A simulação apresentada inicialmente considera exclusivamente à programação de filmes. No entanto, o processo de recomendação é geral, independente de um gênero específico.

A primeira tela do sistema exibe a relação de perfis já cadastrados, tornando possível a inclusão de novos usuários. Ao 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:

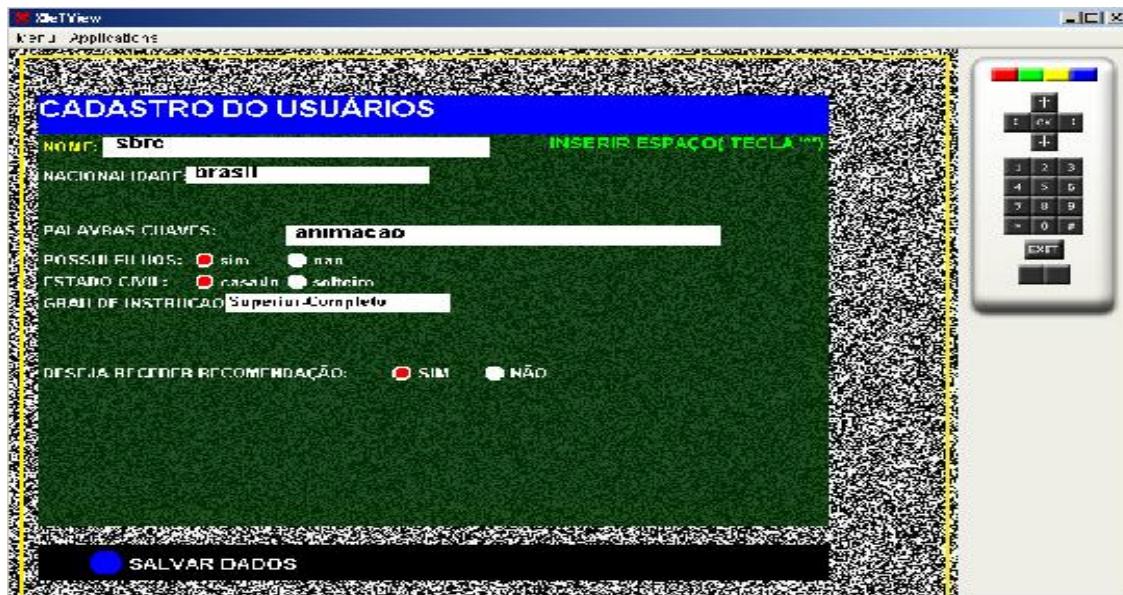


Figura 3. Tela para cadastro de perfis (informações gerais).

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). A partir dessa análise, os programas que se enquadram com o perfil do usuário, neste caso os filmes, serão procurados. Após inseridos os dados e pressionado novamente o botão de cor azul, a próxima tela do cadastro do perfil é exibida, contendo informações sobre preferências de filmes:

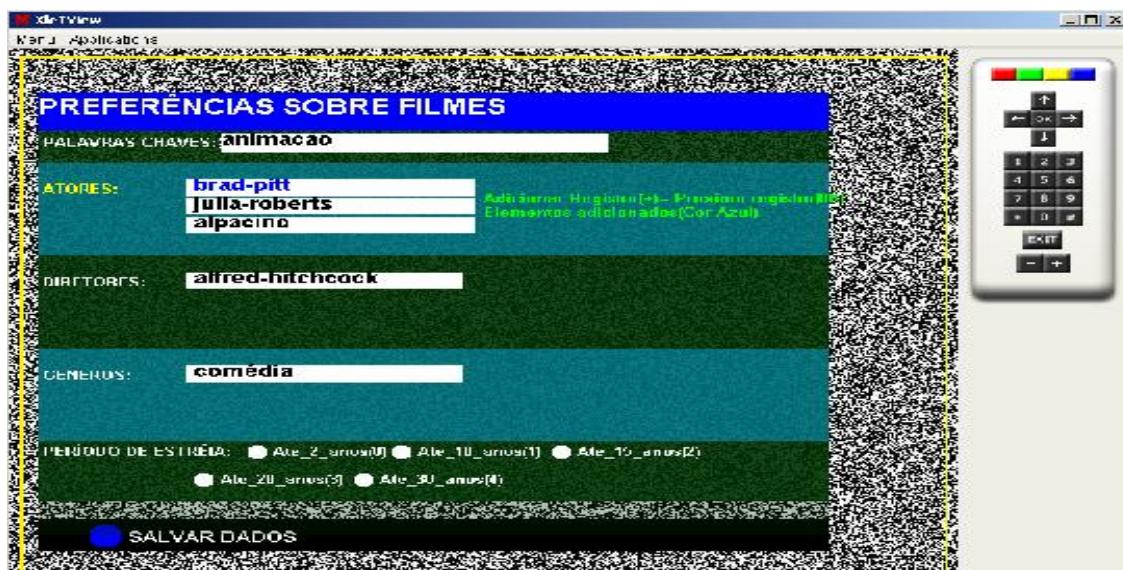


Figura 4. Tela para cadastro de perfis (informações sobre filmes).

Nesta tela, as informações presentes no campo ‘PALAVRAS-CHAVE’, ‘ATORES’ e ‘DIRETORES’ também servem de parâmetros de entrada para o algoritmo

TF-IDF. Em relação a estes dois últimos, vale ressaltar que ambos permitem a escolha de mais de uma opção.

Após a confirmação das informações, o perfil de usuário é construído e cadastrado na base de dados do sistema. Dessa forma, uma vez que o mesmo passa a ser exibido na relação de perfis, uma vez selecionado o usuário ‘sbrc’ (este nome foi escolhido de forma aleatória) e pressionado o botão de cor vermelha, o guia EPG contendo toda a programação televisiva é exibido, mostrando, de forma destacada (preenchimento verde e cor amarela), os programas recomendados para o usuário, conforme mostrado no Figura 4:



Figura 5. EPG exibindo os programas recomendados.

Adicionalmente, a saída do algoritmo TF-IDF também contou com o resultado gerado pelo algoritmo Apriori. Neste contexto, foi utilizado um arquivo denominado ‘progassistida.xml’ (contendo todos os programas assistidos pelo Usuário Telespectador) para simular a programação assistida. O casamento das informações geradas pelos dois algoritmos tornou possível a recomendação dos programas apresentados na tela seguinte (Figura 5). Esta é exibida quando o usuário, ao pressionar o botão de cor azul, solicita a funcionalidade de Avaliação da Recomendação. Através do botão de cor vermelha, o usuário pode aprovar 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.

Outra informação presente nesta tela é a de Percentual de Acerto do sistema, baseando-se na técnica de Precisão. 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 que, no caso do exemplo, obteve valor de 86,11 %. Este valor foi obtido tanto contando com o uso das duas técnicas como após a inclusão das restrições. Apenas o uso combinado das técnicas traria um percentual de acerto de 84,38 %, enquanto que apenas o uso do algoritmo TF-IDF representaria um percentual de 79,22%.

Recomendacao de programas		total de Programas:36	
PROGRAMA		APROVAÇÃO	SINTONIZADO
Cine Maximo - Academia de Gênios		aprovado	SIM
Cine Maximo - Deb e Loide		aprovado	SIM
Cine Maximo - Conah o barbaro		aprovado	SIM
Academia de Gênios		aprovado	SIM
Cine Maximo - O Mundo Perdido		aprovado	NÃO
O Mundo Perdido		aprovado	NÃO
Debate Bola		aprovado	SIM
Esporte Record 2 edicao		aprovado	SIM
Debate Bola 2 edicao		aprovado	SIM
Bola na rede		aprovado	NÃO
Atraíço esportiva		aprovado	SIM
Mulheres e Esporte		aprovado	SIM

PORCENTAGEM DE ACERTO: 86.11%  
 SEM APRIORI || SEM RESTRIÇÕES  
 79.22% 84.38%

Figura 6. Funcionalidade de Avaliação do Sistema de 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 Figura 6:



Figura 7. Avaliação da recomendação - Restrições.

Caso o sistema realize recomendações que não agradaram o usuário, além da reprovação dada através da tela presente na Figura 5, as restrições informadas, conforme Figura 6 (Palavras Não chave, canais e sub-canais indesejados, gêneros indesejados, dentre outras), também serão analisadas pelo sistema de forma a permitir que este desconsidere programas com características indesejadas.

Para a realização da recomendação, o sistema utiliza as informações cadastradas no perfil do usuário, mais especificamente, as informações contidas no campo palavras-chave são verificadas pelo algoritmo TF-IDF. Com a interação usuário telespectador *versus* programação de TV, algumas regras de associação são geradas como consequência do uso do algoritmo Apriori, enriquecendo o processo de recomendação.

Além disso, a inclusão das restrições, quando do momento de avaliação da recomendação, contribui no aprimoramento do processo, aumentando o percentual de acerto.

Tomando por base estas informações, foram realizadas algumas simulações visando analisar em que medida cada algoritmo/técnica utilizada na abordagem proposta aprimora a recomendação realizada. Em uma primeira instância, somente as informações cadastradas no perfil são utilizadas no cálculo do percentual de acerto. Posteriormente, juntamente com o perfil são geradas as regras de associação. Finalmente, além dos algoritmos TF-IDF e Apriori, são consideradas as restrições de recomendação, onde é visível um aumento considerável do percentual de avaliação do sistema. A tabela a seguir mostra os valores obtidos, com aproximação e sem uso de casas decimal, neste estudo comparativo:

**Tabela 2. Percentual de acerto do sistema de acordo com a estratégia utilizada.**

Usuários / Estratégia	Marcelo	Tiago	Geral	SBRC
<b>TF-IDF</b>	35%	41%	53%	79%
<b>TF-IDF + Apriori</b>	60%	60%	53%	84%
<b>TF-IDF + Apriori + Restrições</b>	76%	72%	53%	86%

As simulações acima foram realizadas, no caso dos usuários ‘Marcelo’ e ‘Tiago’, em um período correspondente a uma semana. No caso dos usuários ‘Geral’ e ‘SBRC’, utilizamos um período de 2 semanas. Os valores da tabela foram gerados após o final do período e após a realização de 7 avaliações, ou seja, após 7 diferentes tipos de inclusão de restrições. No caso do usuário ‘Geral’, uma vez que este é gerado automaticamente pelo sistema, só temos um valor de percentual de acerto, que, no caso, é referente somente ao uso das Regras de Associação. Estes tempos foram utilizadas por percebermos que até estes os valores dos percentuais mudavam consideravelmente. Aumentando estes tempos, a diferença nos valores dos percentuais passava a ser insignificante.

As simulações foram realizadas utilizando dados de programação obtidos através da estrutura das tabelas SI, principalmente a Tabela de Associação de Programas (PAT), responsável por indicar os valores de IDs nos fluxos de transporte, a Tabela de mapeamento de programas (PMT), responsável por identificar e indicar a localização das transmissões que compõem cada serviço e, por fim, a Tabela de informação de eventos (EIT), responsável por informações em ordem cronológica sobre os eventos existentes em cada serviço.

É importante ressaltar que, por mais que o objetivo deste trabalho não seja a total aderência ao modelo brasileiro de televisão digital (DTV, 2009), a estrutura das tabelas SI utilizadas é aderente ao padrão ABNT NBR 15603 (ABNT NBR 15603-1, 2008). Para a obtenção deste ambiente de simulação incluindo a geração das citadas tabelas, foi utilizado o dsmcc-mhp-tools (2009), que consiste em um gerador de carrossel de dados, disponibilizando fluxos elementares de dados no formato padronizado do MPEG-2 system. Uma vez geradas as tabelas foi utilizada a API nanoXML, que desempenha a função de *parser*, retirando as informações das tabelas SI, geradas em ambiente de simulação pelo carrossel de dados, e inserindo-as no documento XML contendo a Base de dados de programação.

## 6. Considerações finais

O presente trabalho apresenta uma abordagem para recomendação no domínio da TV. Digital com base em uma estratégia híbrida, fazendo uso dos algoritmos Apriori e TF-IDF (Filtragem Baseada em Conteúdo), ambos bastante difundidos e utilizados em diversos domínios de aplicação. De acordo com a tabela comparativa apresentada, percebemos que as técnicas, quando utilizadas de forma isolada, apresentam um percentual de acerto menor do que quando combinadas. A recomendação obtida a partir desta combinação de técnicas, pode ser ainda aprimorada levando em conta a opinião do usuário sobre a recomendação fornecida (maior percentual de acerto).

O sistema elabora a recomendação com base na construção do perfil do usuário. Com isso o sistema pode realizar recomendações de maneira individualizada, após perceber, de forma implícita, as preferências de cada usuário. O sistema possibilita a criação de perfis de forma explícita por parte do usuário. No entanto, e independente da criação explícita de um ou mais perfis, 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.

No contexto dos desafios definidos pela SBC, é possível identificar importantes contribuições do presente trabalho. Por um lado, uma clara contribuição no campo da *gestão da informação* possibilitando que grandes volumes de dados multimídia distribuídos (programação televisiva), possa ser acessada de forma seletiva através de metadados de programação. Por outro lado, o *acesso participativo e universal do cidadão brasileiro ao conhecimento* e à informação, que já é viabilizado pelo middleware brasileiro, pode ser melhor aproveitado uma vez que a grande maioria dos atuais usuários de aparelhos de televisão, no Brasil, são pessoas que não detêm grandes conhecimentos tecnológicos. Esta constatação pode nos levar a crer que grandes dificuldades serão encontradas por estes quando imersos na vasta gama de programação oferecida pela televisão digital. A abordagem aqui proposta tem o objetivo de amenizar este problema sem exigir nenhum conhecimento tecnológico específico. Desta a população usuária poderá interagir de forma facilitada, e com isso melhor conhecer, os benefícios desta nova tecnologia.

## Referências

- 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](http://www.abnt.org.br/imagens/Normalizacao_TV_Digital/ABNTNBR15603-1_2007Vc_2008.pdf)>. Data de acesso em 18 set. 2009
- AGRAWAL, R. et al. Fast discovery of Association Rules. In: Advanced in knowledge discovery and data mining. AAAI/MIT Press, Menlo Park, Calif., USA, 1996, chap.12, pp. 307-328.

- AGRAWAL, R., IMIELINSKI, T., SWAMI, A. Mining Association rules between sets of items in large databases. International Conference Management of Data (SIGMOD-93). 1993.
- AVILA, P. M., ZORZO, S.D. A personalizad TV Guide System Compliant with Ginga. In: XV Simpósio Brasileiro de Sistemas Multimídia e Web (Webmedia). Fortaleza-CE, 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., MONTEZ, C. TV digital interativa: conceitos, desafios e perspectivas para o Brasil. 1. ed. Florianópolis: I2TV, 2004;
- 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.
- DTV. Site Oficial da TV Digital Brasileira. Disponível em: <<http://www.dtv.org.br>>. Data de acesso em: 02 set. 2009.
- 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.
- GRAVES, M. Projeto de Banco de Dados com XML. Editora Makron Books. ISBN: 8534614717. Primeira Edição, 2003.
- REATEGUI, E. B., CAZELLA, S. C. Sistemas de Recomendação. In: XXV Congresso da Sociedade Brasileira de Computação, 2005, São Leopoldo. Anais do Encontro Nacional de Inteligência Artificial. São Leopoldo, 2005, p. 306 - 348. Disponível em: <<http://www.sbc.org.br/bibliotecadigital/download.php?paper=415>>. Data de acesso em: 28 ago. 2009.
- 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 Caldas/MG. III Fórum de Oportunidades em Televisão Digital Interativa, 2005.
- SMYTH, B., WILSON, D., SULLIVAN, D. Improving the Quality of the Personalized Electronic Program Guide. In: AH'2002 Workshop on Workshop on Personalization in Future TV. Málaga, Spain, 2002.