REVISTA DE
RADIODIFUSÃO
volume 06 - número 06- 2012

ANAIS DO CONGRESSO DA SOCIEDADE BRASILEIRA DE ENGENHARIA DE TELEVISÃO

2012

ISSN: 2236-9619

Expediente

volume 06 • número 06 • 2012

Presidência
Lilliana Nakonechnyj

Vice Presidente
Olímpio José Franco

Diretoria Editorial
Valderez de Almeida Donzelli
Ana Eliza Faria e Silva

Comitê Editorial
Almir Almas
Francisco S. Husni Ribeiro
José Antonio de Souza Garcia
José Olairson
Marcio Pinto Pereira

Diretoria de Ensino
Carlos Nazareth
Eduardo de Oliveira Bicudo

Comitê de Ensino
Carla Pagliari
Frederico Rehme
Rodrigo Arnault
Tom Jones Moreira

Diretoria de Tecnologia
Raymundo Costa Pinto Barros
Brunei Alexandre Yoshida Sano

Comite de Tecnologia
Cicero Legname Marques
Domingos Stavridis
Jacques Varaschim
Paulo Henrique de Castro Corona
Rodrigo Cascão

Revisores

Alexandre Pohl, Ana Eliza Faria e Silva, Arismar Cerqueira Sodre Junior, Carla Pagliari, Carlos Marins, Carlos Montez, Carlos Ynoguti, Débora, Muchaluat-Saade, Eduardo Bicudo, Eduardo Antonio Barros da Silva, Euzebio Tresse, Geraldo Gomes, Gunnar Bedicks, Jose Rehme, Keiko Fonseca, Luciano Mendes, Luiz Biscainho, Marcio Pereira, Marlene Pontes, Ricardo Salles, Ronald Barbosa, Sandro Fasolo, Tatiana Tavares, Tom Jones Assis, Valdecir Becker, Yuzo Iano

Versão eletrônica, Capa e Diagramação
Solange Lorenzo

SET - Sociedade Brasileira de Engenharia de Televisão
Rio de Janeiro/RJ
Rua Jardim Botânico,700 - Sala 306 | Cep. 22461-000
Tel.: + 55 (21) 2512-8747 - Fax + 55 (21) 2294-2791
São Paulo/SP
Av. Auro Soares de Moura Andrade, 252- Cj. 11
Cep. 01156-001
Tels: +55 (11) 3666 9604
www.set.com.br | set@set.com.br

O conteúdo dos artigos é de inteira responsabilidade de seus autores

SUMÁRIO

ANAIS DO CONGRESSO DA SOCIEDADE BRASILEIRA DE ENGENHARIA DE TELEVISÃO - 2012

ISSN: 2236 - 9619

Conforto Visual Estereoscópico e Determinação de Valores de Paralaxe na Câmera 3D Sony HDR-TD10 06
Helio Godoy de Souza

Uma Arquitetura Baseada em Agentes, Ontologias e Metaheurísticas Aplicada ao Servidor Multimídia do SBTVD 13
Arthur Gómez e Toni Wickert

Um modelo computacional baseado em metaheurísticas aplicado a Transmissão Internet Protocol Television (IPTV) 20
Arthur Gómez e Carlos Alfredo Weissheimer Junior

Panorama da produção de conteúdo interativo no mercado brasileiro de televisão aberta em 2011 27
Matheus Colen

Estudo da Rede de Transmissão 4K dos Jogos da Copa do Mundo de 2014 34
Luana Bravo e Eunézio Antônio Souza

Desenvolvimento de Aplicação Interativa em Ginga para o programa "Som e Prosa" da Televisão Universitária Unesp. 40
Lucas Azevedo

Controle do loudness aplicado ao SBTVD: Metodologia de Aplicação e Avaliação 47
Ricardo Costa e Keiko Fonseca

Uso de Reconhecimento Automático de Voz em Português do Brasil na geração de Closed Caption 54
Luiz Brito, Edilberto Strauss e Flávio Mello

Surround Sound 5.1 for one-seg TV and digital radio: the MPEG Surround standard for the Brazilian broadcasting services 61
Luís Meloni e Rafael Diniz

Arquitetura e Aplicações no contexto do *Mining Knowledge* TV 66
Carlos Patricio Junior, Natasha Correia Queiroz Lino e Ramon Santos

Building and Monitoring Next-Generation Video Transport Infrastructure 73
Chin Chye Koh

Impacto da Cobertura de TV Digital Brasileira na Prestação do Serviço de TV Por Assinatura 80
Mateus Cruz e Keiko Fonseca

TV Digital Interativa: Um Recurso Didático no Processo de Ensino e Aprendizagem da Educação Profissional e Tecnológica 86
Pricila Rodrigues de Souza, Halison Ribeiro Maia, Marlos AndrÃ© Silva Rodrigues, Samir Pereira Batalha, Daniel Albino Sandoval e Orlewilson Maia

MAPS-TV: Uma solução de apoio para a geração dinâmica de conteúdo para a TV Digital Interativa Brasileira 93
Izequiel Norões, Elizabeth Furtado e Vasco Furtado

Elaborada por: Silvania W. Martins

Revista Eletrônica de Radio Difusão / SET. -- v. 6, n. 6 (2012). – São Paulo : Editora Sociedade Brasileira de Engenharia de Televisão - SET, 2012-.

Anual.
Periodicidade do v. 6: anual
Disponível na Internet: : http://www.set.com.br/revistaderadiodifusao/
ISSN 2236-9619

1. Radiofusão - Periódicos. 2. TVDigital - Periódicos. 3. Broadcast – Periódicos. I. Sociedade Brasileira de Engenharia de Televisão. II. SET.

CDD (21. ed.) 384.54

Apresentação

A SET apresenta a Revista de Radiodifusão, uma coletânea dos trabalhos acadêmicos sobre temas relacionados à Televisão Brasileira. Estão sendo contemplados os trabalhos apresentados nos Congressos SET, realizados no Centro de Exposições Imigrantes, em São Paulo.

No intuito de promover a disseminação da ciência brasileira e com o objetivo de ampliar a visibilidade desses trabalhos, a SET está modernizando seu veículo de divulgação, substituindo a modalidade impressa, pelo formato eletrônico. Mantém-se preservada a indexação junto ao ISSN. Além do ISSN, as diretorias de Ensino e Editorial obtiveram recentemente o indexador Qualis, que resultará em maior valorização dos trabalhos acadêmicos apresentados em seus congressos.

Nesta oportunidade, a SET aproveita para registrar seus agradecimentos aos pesquisadores, mestrandos e doutorandos que tem contribuiído crescentemente para as iniciativasa de nossa sociedade, bem como reafirmar seu propósito de fortalecer essa importantíssima parceria.

Liliana Nakonechnyj
Presidente da SET

Conforto Visual Estereoscópico e Determinação de Valores de Paralaxe na Câmera 3D Sony HDR-TD10

Helio Godoy de Souza
(hgodoy@uol.com.br)

Conforto Visual Estereoscópico e Determinação de Valores de Paralaxe na Câmera 3D Sony HDR-TD10.

Prof. Dr. Hélio Augusto Godoy de Souza [1]

[1] Universidade Federal de Mato Grosso do Sul / Photon3D cine & video, Campo Grande, MS, Brasil, hgodoy@uol.com.br

Resumo:

Este artigo descreve a metodologia utilizada para revelar a zona de conforto estereoscópico da câmera 3D Sony HDR-TD10, e indica um procedimento de avaliação da paralaxe durante as filmagens.

Resumo Ampliado:

O lançamento de câmeras de vídeo 3D com duas objetivas vem se destacando no mercado de equipamentos estereoscópicos para televisão. Grandes companhias como Panasonic, Sony e JVC desenvolveram produtos específicos que atendem profissionais e aficcionados.

Todavia, o uso do monitor LCD autoestereoscópico na câmera 3D Sony HDR-TD10, dificulta a determinação da paralaxe na imagem final a ser projetada.

A filmagem de objetos a distâncias controladas e a análise dessas imagens, no software Sony Vegas Pro, permitiu visualizar imagens anaglíficas para determinar precisamente a paralaxe negativa e positiva.

A construção de um gráfico cartesiano, com valores de porcentagem de paralaxe versus a distância dos objetos, permitiu visualizar a zona de conforto estereoscópico da TD10. Com a plotagem de uma curva de tendência, criou-se uma tabela de referência para os valores de paralaxe.

O uso dessa tabela em filmagens com a TD10 permite manter a paralaxe negativa e a paralaxe positiva dentro da zona de conforto estereoscópico.

Palavras chave: estereoscopia, TV3D, câmera 3D, conforto, paralaxe.

1. INTRODUÇÃO

A disseminação da tecnologia para visualização estereoscópica vem sendo seguida por lançamentos de câmeras com duas objetivas para filmagem estereoscópica.

Recentemente tanto a Panasonic como a Sony lançaram modelos profissionais de câmeras estereoscópicas, seguindo tendência já observada desde 2010, quando foram lançadas as primeiras câmeras estereoscópicas semiprofissionais.

As câmeras semiprofissionais carecem de um sistema de monitoramento de paralaxe. Essas câmeras possuem um monitor de LCD autoestereoscópico, que permite a visualização 3D sem a utilização de óculos. Embora essa solução seja interessante do ponto de vista comercial, não resolve o problema fundamental de controle da paralaxe.

A imagem estereoscópica contém uma representação distinta para a visão de cada um dos olhos; a diferença entre uma imagem e outra é denominada de *paralaxe*, que é a distância entre pontos correspondentes ao olho direito e esquerdo na imagem estereoscópica projetada na tela.

A determinação da paralaxe pode ser feita através da visualização de uma imagem estereoscópica anaglífica em tons de cinza.

Na figura 01 é apresentada uma imagem anaglífica na qual se observam as diferentes paralaxes[1] existentes na construção da imagem estereoscópica.

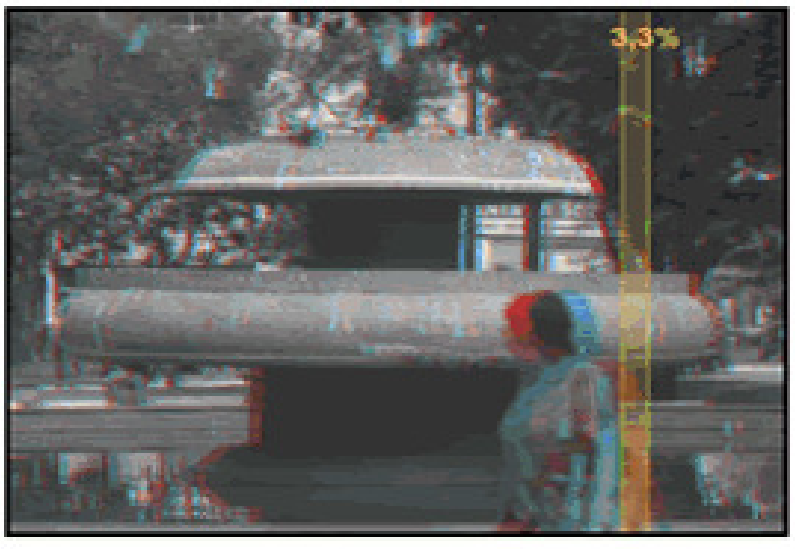

Figura 1 – Exemplo de imagem anaglífica sendo utilizada para monitorar a porcentagem de paralaxe negativa em relação à largura da tela.

Steve Schklair[2] (Mendiburu, 2011) afirma e defende a idéia de que os óculos deveriam ser utilizados apenas para a confirmação de que algum excesso de paralaxe não vai ferir os olhos dos espectadores, mas que a decisão deve ser tomada através da visualização direta (não mediada pelos óculos) das representações imagéticas bidimensionais das paralaxes negativas e positivas observáveis em imagens semelhantes a essa apresentada acima.

---

[1] A paralaxe positiva ocorre quando as imagens são percebidas em um plano localizado atrás da tela; já a paralaxe negativa ocorre quando as imagens são percebidas à frente da tela. A paralaxe zero ocorre quando as imagens são percebidas no plano da tela.

[2] Produtor especializado em S3D, responsável pelos estúdios da 3ality Technica (ex 3ality Digital), uma das empresas que mais investem em P&D para desenvolvimento da tecnologia S3D para transmissões em tempo real e participante de boa parte dos filmes S3D de grande sucesso.

Assim é possível, através de um monitoramento das imagens anaglíficas, a determinação de valores para a paralaxe, definindo com bastante precisão a percepção de profundidade da imagem estereoscópica[3].

A percepção do espaço tridimensional possibilitada pela estereoscopia é determinada pela paralaxe entre as duas imagens. Existem conhecimentos precisos sobre os valores de paralaxe que vão induzir a formação de uma imagem tridimensional no cérebro, e que indicam parâmetros aceitáveis para o conforto visual. Estudos oftalmológicos (Ukai & Howarth, 2008 [1]; Banks, Shibata, Kim & Hoffman, 2011 [2]; Howarth, 2011 [3]) demonstram que é possível determinar a paralaxe mais adequada a partir de valores de porcentagens de paralaxe em relação à largura da tela.

De acordo com Banks, Shibata, Kim & Hoffman (2011), existem inúmeras causas potenciais de desconforto visual ao se visualizar imagens estereoscópicas. Estas incluem desconforto devido aos óculos necessários para separar as duas imagens nos olhos, fantasmas ou crosstalk entre as duas imagens, desalinhamento das imagens, conflitos de vergência e acomodação, dentre outras.

Seguindo-se as premissas de Ukai & Howarth, (2008) considera-se a questão da relação entre vergência e acomodação como sendo a atividade ocular que mais sofre influências dos valores indevidos de paralaxe negativa e positiva, que podem promover desconforto ou fadiga visual.

Acomodação é a atividade de modificação da distância focal do cristalino pela ação da musculatura que o controla, e serve para manter focalizados objetos distantes ou próximos do rosto do indivíduo.

Vergência refere-se ao ângulo formado pelos eixos ópticos dos dois olhos, de modo que, eixos paralelos promovem a fusão de imagens no infinito, a convergência promove a fusão de imagens mais próximas dos olhos e a divergência teoricamente não ocorre no mundo natural. A convergência tem um limite conhecido como o "limite fusional".

Woods, Docherty & Koch (1993) [4] apontam a relação entre a dificuldade de focalização de imagens estereoscópicas e a paralaxe exagerada. Acredita-se que as ações musculares de acomodação e as de vergência estejam relacionadas através de aprendizagem obtida durante o desenvolvimento psicomotor. A focalização de um objeto a poucos centímetros do nariz, necessariamente corresponde a um objeto perto do nariz de modo que há grande convergência dos eixos ópticos dos olhos.

No caso da estereoscopia, uma imagem com elementos com paralaxe negativa, isto é, localizados virtualmente à frente da tela, apesar dos eixos oculares convergirem para esses elementos, a tendência do cristalino é buscar o foco onde a imagem encontra-se focalizada: sobre a tela e não no lugar de vergência. Trata-se de uma informação incoerente com a aprendizagem visual experimentada pelo sujeito ao longo de seu desenvolvimento psicomotor. Isso pode gerar desconforto visual.

Tanto Banks, Shibata, Kim & Hoffman (2011) como Howarth (2011) definem a "Zona de Conforto" da visão binocular. Na figura 2 (hipotética), a linha central representa as exigências de acomodação e vergência de objetos em distâncias diferentes para a visão natural. A linha horizontal (denominada "display estereoscópico") representa as exigências de acomodação e de convergência de uma determinada apresentação estereoscópica. A linha tracejada indica a Zona da Clara Visão Binocular Simples (Zone of Clear Single Binocular Vision - ZCSBV) e a linha pontilhada (hipotética) indica Zona de Conforto Estereoscópico no ZCSBV.

A área situada à esquerda do gráfico representa a imagem estereoscópica situada atrás da tela, enquanto que o lado direito do gráfico representa a área localizada à frente da tela. Os valores nas coordenadas são fracionários por isso aumentam com a diminuição da distância da tela ou dos objetos. O display está colocado a 1/3 de metro.

Na prática, o interior da Zona de Conforto é uma gradação de confortos que vai do máximo conforto próximo à linha contínua central ao mínimo conforto à medida que se aproxima da linha pontilhada no limite da faixa colorida.

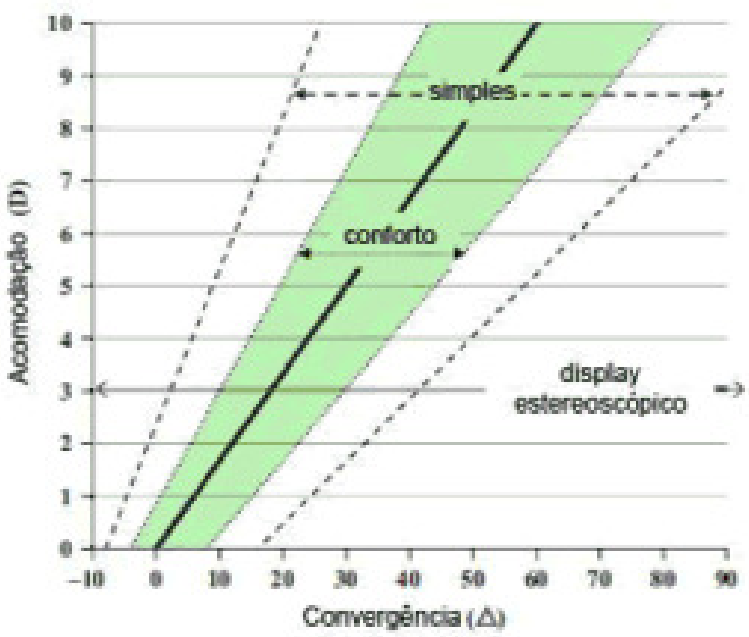

Figura 2 – Gráfico demonstrativo do conceito de Zona de Conforto Estereoscópico (Howarth (2011))

Existe correlação entre a Zona de Conforto Estereoscópico e a Porcentagem de Paralaxe nas imagens estereoscópicas (Banks, Shibata, Kim & Hoffman, 2011). Embora não haja consenso em torno dos valores de porcentagem adequados, há uma espécie de regra da porcentagem em relação à

[3] Existem variadas imagens 2D que buscam representar as diferentes paralaxes: mapas de profundidade, diferenças entre esquerda e direita, mistura 50%-50%, dentre outras. O autor deste artigo optou pela utilização das imagens anaglíficas pois permitem a visualização objetiva das paralaxes negativas e positivas.

largura da tela, que é seguida por muitos estereógrafos. Essa regra apresenta os seguintes valores: entre 2% e 3% para a paralaxe negativa (imagens à frente da tela), e entre 1% e 2% para a paralaxe positiva (imagens atrás da tela). Todavia, através dos testes descritos em Banks, Shibata, Kim & Hoffman (2011), acredita-se ser possível aumentar-se a paralaxe negativa até valores entre 3% e 4% da largura da tela, mantendo-se a paralaxe positiva entre 1% e 2% no máximo. Dessa forma seria ainda possível manter-se a imagem estereoscópica dentro da Zona de Conforto Estereoscópico.

A câmera HDR-TD10 é uma câmera estereoscópica com duas objetivas que grava o sinal de vídeo 3D em memória interna de 64GB no formato AVCHD (MPEG-4 AVC/H.264) - MVC ("multi view coding"), com resolução de 1920 X 1080 pixels e "frame rate" de 60i. O monitor é um LCD autoestereoscópico de 3,5 polegadas, que permite, visualizar imagens 3D sem a utilização de óculos. A câmera possui um botão giratório (manual dial) para o controle de paralaxe. Esse controle é feito através do deslocamento lateral (shift) da área sensível dos sensores (dois sensores de 1/4 de polegada), de modo a variar os eixos ópticos desde uma posição de convergência até uma posição paralela (Figura 3). Através do monitor LCD não é possível determinar-se precisamente os valores da paralaxe positiva e negativa, nem mesmo é possível perceber-se quando os eixos ópticos estão paralelos ou convergentes.

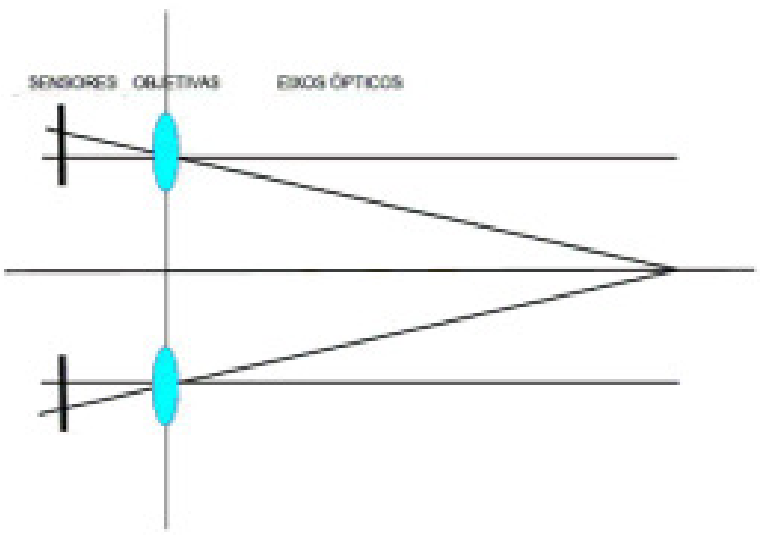

Figura 3 – Esquema demonstrativo do deslocamento de área sensível do sensor da HDR-TD10

A única informação que o manual fornece a respeito do conforto visual estereoscópico é uma referência ao uso do zoom e sua interdependência com a aparência tridimensional da profundidade (*"The distance at wich you can obtain movies with a deep, 3-dimensional appearance varies depending on the zooming level"* ; página US-29 do manual impresso em inglês/espanhol). Assim, sucintamente, o manual define as condições propícias para a atividade de filmagens 3D. A tabela existente no manual (Tabela 1) é transcrita a seguir:

Tabela 1 – Tabela indicativa da relação entre posição do Zoom e a faixa de distâncias na qual é possível obter-se profundidade; transcrita do manual inglês/espanhol da câmera Sony HDR-TD10. W significa "wide"(grande angular), T significa "tele" (teleobjetiva)

Distância do Objeto	Posição do Zoom
Entre 80cm até 6m	W-|||----------T
Entre 2,5m até 10m	W----------|||----------T
Entre 7,5m até 20m	W----------|||-T

Desde os anos 1950 já é conhecido o problema da relação entre a distância focal das objetivas e o resultado em termos de paralaxe em uma imagem estereoscópica (Mckay, 1953 [5]). Várias fórmulas foram desenvolvidas desde então e atualmente são utilizadas em calculadoras que prometem determinar a "base estereoscópica" que é a distância ideal entre os eixos ópticos de câmeras montadas lado-a-lado para se obter uma paralaxe adequada ao conforto visual.

Essas fórmulas possuem como parâmetros inter-relacionados, a distância focal das objetivas e a distância entre os objetos e as câmeras, resultando em um valor para a base estereoscópica que proporcionará uma imagem estereoscópica com um determinado valor de paralaxe. A Fórmula de Bercowitz (1) é um bom exemplo para o caso dos eixos ópticos paralelos (CAHEN, 2011 [6]):

$$B = P/(L-N) * [(LN/F) - ((L+N)/2)] \quad (1)$$

Ainda como exemplo tem-se a fórmula usada na calculadora (Base Calc)[4] de Pierre Meindre (2) que é uma derivação da fórmula de Berkowitz:

$$B = P/F * [LN/ (L - N)] \quad (2)$$

Onde: **B** é Base Estereoscópica (distância entre os eixos ópticos); **P** é o valor da Paralaxe desejada, em milímetros para um filme de 35 mm; **L** é a Maior distância entre objeto e câmeras; **N** é Menor distância entre objeto e câmeras; e **F** é a Distância Focal das Objetivas.

Aplicando-se essa fórmula através da calculadora acima citada é possível deduzir-se o comportamento de lentes com distâncias focais diferentes em diferentes bases estereoscópicas para obter-se uma paralaxe dentro de um valor aceitável de conforto visual estereoscópico. A maior paralaxe negativa aceitável refere-se ao Objeto Mais Próximo da objetiva. Quanto mais próximo estiver o objeto, maior será a paralaxe negativa na imagem desse objeto. Os objetos situados no infinito, para eixos ópticos paralelos será idealmente igual a 0% da largura da tela. Assim, o valor de maior proximidade deverá sempre representar um valor de paralaxe situado dentro do conforto visual estereoscópico.

[4] http://www.stereoscopie.fr/tech/BaseCalc.php

O gráfico derivado da aplicação das fórmulas citadas, permite deduzir que, para uma mesma distância entre eixos ópticos (base estereoscópica), objetivas de distância focal maior irão produzir paralaxes maiores, desde que os objetos estejam situados a uma mesma distância. Para obter-se uma paralaxe adequada, para uma teleobjetiva de distância focal de 300 mm[5], é necessário que o objeto mais próximo fique mais distante do que se fosse utilizada uma objetiva de 135 mm ou de 35 mm (Figura 4).

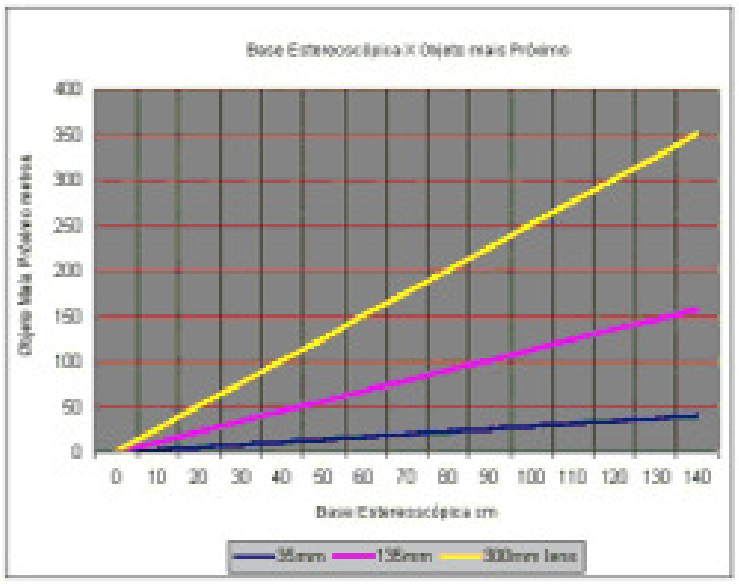

Figura 4 – Gráfico decorrente da aplicação da fórmula de Bercowitz para objetivas com diferentes distâncias focais. (http://nzphoto.tripod.com/stereo/3dtake/fbercowitz.htm)

Assim é possível deduzir-se que para uma mesma distância do objeto, as teleobjetivas (maior que 50mm) produzirão uma paralaxe maior do que as objetivas grande angulares (menor que 50mm); desde que o par de lentes estereoscópicas, estejam colocadas à mesma base estereoscópica. Isto é coerente com a tabela apresentada pelo manual da HDR-TD10. Deve ser lembrado que na câmera analisada, a base estereoscópica é fixada a uma distância de 31 mm e a distância focal da objetiva Zoom varia de 2.9 mm a 29 mm (para um sensor de 1/4 de polegada equivaleria em uma câmera fotográfica de 35mm a uma objetiva Zoom de 16 mm a 160 mm).

Aplicando-se a fórmula de Bercowitz[6] à câmera HDR-TD10, para uma distância focal de 2,9 mm (W – Wide), obtem-se como resultado para a menor distância referente à tabela do manual (80cm), uma paralaxe negativa de aproximadamente 1,75%, que é bastante conservadora vide as considerações de Banks, Shibata, Kim & Hoffman (2011). Já para a distância focal de 29 mm (T - Tele) ao se aplicar a fórmula obtemos um valor de paralaxe negativa de aproximadamente 1,74%.

Como a câmera em questão possui um controle de paralaxe baseado na convergência dos eixos ópticos das objetivas, isso implicaria em ajustes para a aplicação da fórmula de Bercowitz, usada na forma descrita, somente para eixos paralelos.

Assim optou-se por um teste de visualização de imagens da TD10, incluindo a convergência dos eixos ópticos, de modo a se verificar na prática como as paralaxes se comportavam efetivamente.

O teste apresentou-se como a solução correta para o problema e produziu três tabelas para cada uma das três distâncias focais da objetiva Zoom, relacionando a distância dos objetos com a paralaxe da imagem estereoscópica, bem como com a posição do botão manual de controle do ângulo dos eixos ópticos – "stereo knob" ou "manual dial".

2. OBJETIVO

O objetivo do teste foi determinar a zona de conforto estereoscópico para a câmera HDR-TD10, para a obtenção de um sistema de controle de paralaxe que pudesse ser usado em filmagens a campo.

3. MÉTODOS

A partir do manual da câmera (vide tabela 1), para o Zoom na posição Wide (2,9mm), tomou-se a distância mínima e a máxima como pontos de partida para a delimitação de uma faixa de filmagem ao longo dos eixos ópticos, Dividiu-se essa faixa em oito partes e solicitou-se que uma pessoa se deslocasse ao longo da faixa. Repetiu-se o procedimento para a posição Intermediária da objetiva Zoom (aqui denominada Middle) e para a posição Tele. Para cada uma das três posições, realizaram-se também três testes, um para cada uma das posições do botão manual de controle de paralaxe, denominados de: posição (-) minus, (0) zero e (+) plus. O sinal (-) equivale a um "envesgamento" ("toed in") pronunciado dos eixos ópticos, o sinal (0) equivale a um "envesgamento" leve dos eixos ópticos, enquanto que o sinal (+) equivale a posição paralela dos eixos ópticos. Isso somente foi deduzido após a análise das imagens.

A série com as imagens assim obtidas, foi analisada no software Sony Vegas Pro (versão 10e build 737) com o sistema estereoscópico acionado para a visualização de imagens anaglíficas (Figura 4).

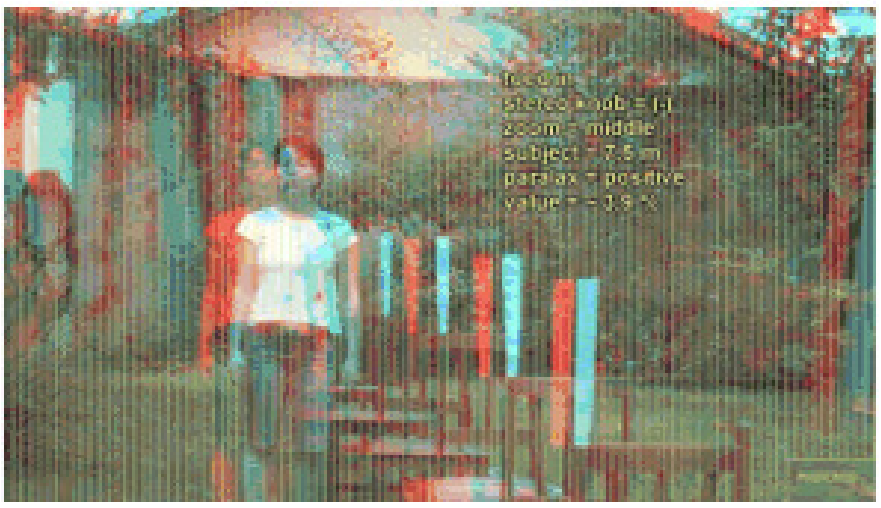

Figura 4 –Imagem anaglífica obtida no teste de paralaxe, parâmetros encontrados: *toed in* (eixos envesgados), *stereo knob = (-), zoom = middle, subject = 7,5m, paralax = positive* (situação da paralaxe na posição da pessoa) e *value = -3,9%*

[5] Os valores de distância focal utilizados neste trecho do artigo, referem-se a objetivas de câmera fotográfica para filmes de 35 mm.

[6] A fórmula foi aplicada a partir de planilha do software Excell fornecida pelo site: http://nzphoto.tripod.com/stereo/3dtake/fbercowitz.htm

Como guia para a medição da paralaxe foi utilizada uma imagem misturada à imagem original, com linhas dispostas a cada 1% do tamanho horizontal das imagens (1920 pixels) ao longo da largura da imagem.

Após a análise visual, os valores de distância e de paralaxe foram lançados em uma tabela no software Microsoft Excell de modo a permitirem posteriormente uma visualização gráfica do comportamento da paralaxe. Abaixo as tabelas obtidas:

Tabela 2, 3 e 4 – Tabelas com os resultados parciais dos testes com as imagens anaglíficas, para as posições 2 ZOOM WIDE (grande angular), 3 ZOOM MIDDLE (intermediário) e 4 ZOOM TELE (teleobjetiva) respectivamente.

2	ZOOM - WIDE			
		toed in >	toed in <	Paralelo
FOTO	DIST meters	KNOB minus	KNOB zero	KNOB plus
1	0,40	0,50%	-2,00%	-5,00%
2	0,80	2,00%	-0,80%	-3,50%
3	1,70	3,00%	-0,20%	-2,10%
4	2,60	3,80%	1,00%	-1,60%
5	3,50	4,00%	1,10%	-1,30%
6	4,40	4,05%	1,30%	-1,20%
7	5,30	4,10%	1,50%	-1,05%
8	6,20	4,20%	1,80%	-1,00%
9	16-18	4,50%	2,00%	-0,60%

3	ZOOM – MIDDLE			
		toed in >	toed in <	paralel
FOTO	DIST meters	KNOB minus	KNOB zero	KNOB plus
1	1,20	-0,20%	-3,00%	-6,00%
2	2,50	2,00%	-0,90%	-3,10%
3	3,75	2,60%	0,00%	-2,60%
4	5,00	3,40%	0,90%	-2,00%
5	6,25	3,80%	1,00%	-1,50%
6	7,50	3,90%	1,10%	-1,30%
7	8,75	4,00%	1,20%	-1,20%
8	10,00	4,10%	1,40%	-1,10%
9	16-18	4,20%	1,60%	-1,00%

4	ZOOM - TELE			
		toed in >	toed in <	paralel
FOTO	DIST meters	KNOB minus	KNOB zero	KNOB plus
1	3,50	-4,00%	-6,00%	-9,00%
2	7,50	0,40%	-2,00%	-4,80%
3	9,60	1,00%	-1,00%	-4,00%
4	11,70	1,70%	-0,50%	-3,40%
5	13,80	2,10%	-0,30%	-3,00%
6	15,90	2,20%	0,00%	-2,50%
7	16-18	2,40%	0,00%	-2,10%

Os dados foram lançados em um gráfico cartesiano relacionando a porcentagem de paralaxe com a distância do objeto em relação à câmera. A partir do próprio software Excell foi automaticamente obtida uma linha de tendência. O melhor resultado de coerência entre a linha de tendência e os dados refere-se a uma equação polinomial de grau 4:

$$y = - ax^4 + bx^3 - cx^2 + dx - e \qquad (3)$$

Esse procedimento foi repetido para as três posições da objetiva ZOOM. Abaixo, como exemplo, apresenta-se o gráfico para a objetiva ZOOM na posição WIDE (Figura 5), com as três curvas referentes às três posições do botão manual de controle de paralaxe, minus(-), zero(0) e plus(+):

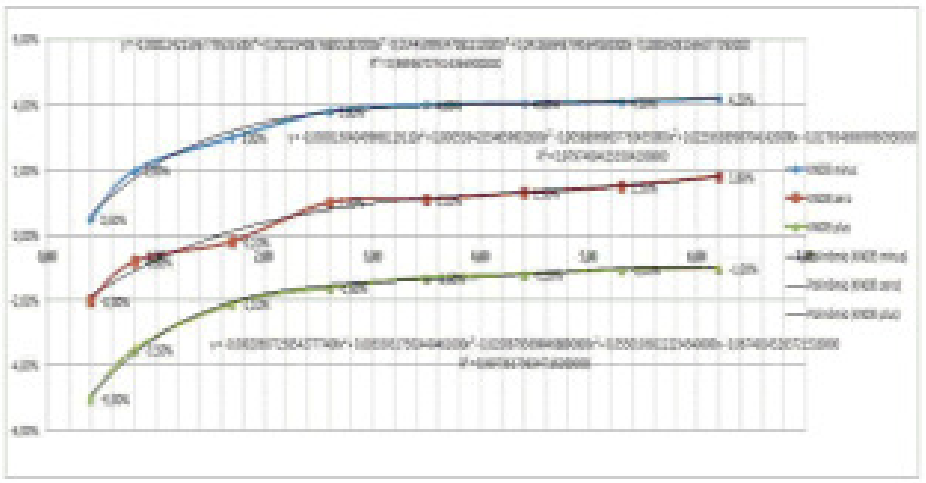

Figura 5 – Exemplo de Gráfico obtido com a plotagem dos dados da tabela 2 (1 ZOOM WIDE). Azul = minus; Vermelho = zero; Verde = plus. A linha escura no interior das coloridas é o resultado da plotagem de uma equação de tendência do tipo polinomial no formato $y = - ax^4 + bx^3 - cx^2 + dx - e$; cada equação com suas constantes é lançada automaticamente sobre o gráfico.

O procedimento seguinte consistiu em usar cada uma das equações polinomiais em uma nova planilha para a obtenção dos valores de paralaxe para uma variação de distância (agora hipotética) de 10 em 10 centímetros do objeto em relação à câmera.

4. RESULTADOS

Para se delimitar a zona de conforto estereoscópico da câmera HDR-TD10 necessita-se escolher quais os valores de porcentagem de paralaxe que serão adequados para uma boa visualização estereoscópica. Escolheu-se como valor máximo para a paralaxe negativa (à frente da tela) a medida de -2,5% e como valor máximo para a paralaxe positiva (atrás da tela) a medida de 1%. São valores abaixo dos preconizados por Banks, Shibata, Kim & Hoffman, (2011), mas esses valores permitem correções na pós-produção em caso de necessidade. Assim os valores obtidos são os seguintes:

Tabela 5, 6 e 8 – Tabelas com os resultados finais do teste de verificação da zona de conforto estereoscópico para a câmera HDR-TD10 5 - ZOOM WIDE (grande angular), 6 - ZOOM MIDDLE (intermediário) e 7 - ZOOM TELE (teleobjetiva) respectivamente.

5 - Objetiva Zoom posição **WIDE** (grande angular)		
Botão Estéreo (stereo knob)	Distância Mínima (m)	Distância Máxima (m)
Minus (-) = muito "vesgo"	**0,00**	**0,50**
Zero (0) = "vesgo"	**0,20**	**3,00**
Plus (+) = paralelo	**1,40**	**∞**

6 - Objetiva Zoom posição **MIDDLE**		
Botão Estéreo (stereo knob)	Distância Mínima (m)	Distância Máxima (m)
Minus (-) = muito "vesgo"	**0,00**	**1,70**
Zero (0) = "vesgo"	**1,50**	**5,90**
Plus (+) = paralelo	**3,70**	**∞**

7 - Objetiva Zoom posição **TELE**		
Botão Estéreo (stereo knob)	Distância Mínima (m)	Distância Máxima (m)
Minus (-) =muito "vesgo"	**4,50**	**9,10**
Zero (0) ="vesgo"	**6,90**	**21,0**
Plus (+) = paralelo	**16,0**	**∞**

Estes resultados permitem a utilização da câmera a campo com maior precisão dos limites de paralaxe a serem obedecidos para cada situação. A implementação deste controle necessitaria de um sistema de medição de distâncias que poderia ser realizado com uma trena do tipo laser para maior agilidade de filmagens.

Não se recomenda exceder o limite máximo, pois poderia provocar divergência dos eixos oculares o que não é aconselhável para a musculatura ocular. Pequenos excessos no limite mínimo podem ser aceitáveis, mas não por longos períodos de tempo.

5. DISCUSSÃO

O limite do método empregado para a obtenção dos valores de paralaxe versus distância do objeto, repousa na equação obtida automaticamente pelo software Microsoft Excell. A equação polinomial (3) apesar de representar o comportamento dos pontos obtidos empiricamente no teste, não se confunde com uma lei dedutível do comportamento da paralaxe em função da distância do objeto em relação à câmera. Sabe-se que no caso dos eixos paralelos, para qualquer posição da objetiva Zoom, a curva assume um comportamento assimptótico, e somente no infinito seria obtido o valor de paralaxe zero. Todavia uma equação polinomial não obedece esse princípio. Assim, tomou-se a decisão de indicar o valor (∞) infinito nas tabelas de eixo paralelo, mais em função daquilo que se conhece do comportamento estereoscópico do que em função dos resultados da fórmula polinomial.

Por outro lado, a fórmula de Berkowitz, mostra consideráveis distorções entre os valores obtidos pelo método empírico e os valores previstos pela fórmula (1 e 2). Isso significa que a fórmula de Berkowitz não é aplicável para o caso da HDR-TD10.

A metodologia aplicada à TD10 poderá ser utilizada em qualquer outro modelo de câmera estereoscópica para a determinação da zona de conforto estereoscópico, bem como em aferições das calculadoras estereoscópicas encontradas no mercado.

6. CONCLUSÕES

Os resultados obtidos pelo método empírico, representados pelos valores das tabelas 5, 6 e 7 são os dados mais confiáveis através dos quais é possível obter-se um resultado estereoscópico dentro da zona de conforto visual.

A restrição de uso da câmera HDR-TD10 é observada para objetos situados a distâncias muito grandes, mesmo nas situações de eixo paralelo, não pelo excesso de paralaxe, mas sim pela falta dele.

Como se sabe, para a visão humana, a acuidade estereoscópica é a capacidade de discernimento de imagens de um mesmo objeto, localizadas em pontos diferentes das duas retinas, isso é denominado estereopsis. De acordo com Valius (Lipton,1982) [7], para uma distância interocular média de 6,5 cm, e uma retina com acuidade estereoscópica de 10 segundos de arco, o observador deverá ter uma percepção estereoscópica até uma distância de aproximadamente 1.340 metros.

Considerando a distância entre eixos ópticos (3,1cm) e a "acuidade estereoscópica" da HDR-TD10, podemos especular a respeito da ausência de paralaxe em objetos situados a grandes distâncias.

Seguindo uma avaliação subjetiva, feita em testes de filmagem de paisagens com horizonte, observou-se que a TD10 tem capacidade estereoscópica até 20 ou 30 metros, para a posição WIDE com o botão manual de controle de paralaxe na posição plus (+) (eixos ópticos paralelos). A partir dessa profundidade os objetos passariam a fazer parte de um mesmo plano de fundo. Todavia, trata-se de uma hipótese, e esse assunto merece uma análise empírica mais controlada.

As fabricantes Sony e Panasonic implementaram o monitoramento de paralaxe nos modelos de câmeras estereoscópicas profissionais. Infelizmente nenhum dos modelos semiprofissionais apresentam opção de monitoramento anaglífico, que proporcionaria tomada de decisão sobre a zona de conforto estereoscópico.

REFERÊNCIAS

[1] K. UKAI, & , P.A. HOWARTH. Visual fatigue caused by viewing stereoscopic motion images: background, theories, and observations. Displays no. 29, 2008, pp. 106–116

[2] M. S. BANKS, T. SHIBATA, J. KIM e D. M. HOFFMAN. The zone of comfort: Predicting visual discomfort with stereo displays. Journal of Vision, 11(8):11, 2011, pp. 1–29

[3] P.A. Howarth. Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review. Ophthalmic & Physiological Optics 31,2011,pp. 111–122

[4] A. Woods, T. Docherty & R. Koch. Image Distortions in Stereoscopic Video Systems. Stereoscopic Displays and Applications IV, Proceedings of the SPIE Volume 1915, San Jose, CA, Feb. 1993

[5] H.C. Mckay. Three-dimensional Photography, Principles Of Stereoscopy, New York, American Photographic Publishing Company, 1953. 333 pag.

[6] O. Cahen. L'image en relief, du film au numerique. Paris, Presses des Mines, 2011. 233 pag.

[7] L. Lipton. Foundations of the Stereoscopic Cinema: A Study in Depth. New York, Van Nostrand Reinhold Company, 1982. 319 pag.

Uma Arquitetura Baseada em Agentes, Ontologias e Metaheurísticas Aplicada ao Servidor Multimídia do SBTVD

Arthur Gómez (atgomezbr@gmail.com)
Toni Wickert (toniwickert@gmail.com)

UMA ARQUITETURA BASEADA EM AGENTES, ONTOLOGIAS E METAHEURÍSTICAS APLICADA AO SERVIDOR MULTIMÍDIA DO SBTVD

Arthur Tórgo Gómez [1], *Toni Ismael Wickert* [2], *Daniel Formolo* [3]

[1] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, atgomezbr@gmail.com
[2] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, toniwickert@gmail.com
[3] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, formolo@gmail.com

Resumo: A televisão analógica aberta atinge cerca de 95% dos lares brasileiros. Com o advento do Sistema Brasileiro de Televisão Digital (SBTVD) os usuários terão a possibilidade de ter um canal de interatividade. Dessa forma, será possível ter acesso ao servidor de aplicações multimídia, ou seja, haverá a possibilidade de enviar ou receber e-mails, acessar aplicações interativas, dentre outros. Esse trabalho propõe o desenvolvimento e a implementação de uma arquitetura que contempla um módulo com o objetivo de sugerir conteúdos ao usuário de acordo com o seu perfil e outro módulo para otimizar o conteúdo que será transmitido. A implementação foi desenvolvida utilizando ontologias, agentes de *software*, Busca Tabu (BT), Algoritmo Genético (AG) e Algoritmo Híbrido (AH). A validação dos resultados é feita através de uma métrica.

Palavras chave: agentes de *software*, ontologias, metaheurísticas.

1. INTRODUÇÃO

Atualmente, no Brasil, a televisão aberta é o mais importante meio de acesso à informação. Dos 58,4 milhões de lares, 55,4 milhões tem pelo menos um aparelho receptor. Ou seja, a televisão aberta chega a atingir em torno de 95% dos lares brasileiros [1].

Com o advento da televisão digital temos também uma evolução em termos tecnológicos, por exemplo, o sinal que era recebido de forma analógica, agora é recebido de forma digital o que melhora a qualidade da imagem evitando o aparecimento de chuviscos na tela e imagens com fantasmas, comumente encontrados em transmissões analógicas [2].

Além das melhorias em relação à qualidade de imagem e transmissão, o Sistema Brasileiro de Televisão Digital (SBTVD) contempla o canal de interatividade no qual é possível ter aplicações interativas. A emissora poderá ter um servidor com aplicações multimídia e o usuário poderá acessá-las através do canal de interatividade. Nesse contexto, o canal de interatividade pode ser usado para transmissão de dados como e-mails, questionários interativos ou até mesmo conteúdos multimídia como aulas interativas à distância, apresentações onde o usuário pode interagir com o conteúdo, dentre outros [3].

O objetivo desse trabalho é a implementação de um sistema dividido em dois módulos. No módulo de sugestão de conteúdo são utilizados agentes de *software* e ontologias para classificar o conteúdo e cruzar com o perfil do usuário para lhe sugerir conteúdo adequado. Já no módulo de transmissão de conteúdo são usados outros agentes de *software* que junto com aplicações da Busca Tabu (BT), do Algoritmo Genético (AG) e de um Algoritmo Híbrido (AH) definirão a melhor política de transmissão, através da otimização dos parâmetros de transmissão, como por exemplo, a taxa de transmissão mais adequada, a qualidade na qual o áudio e o vídeo serão transmitidos, dentre outros.

Esse artigo está dividido como segue. Na seção 2 são apresentados os trabalhos relacionados. Na seção 3 é apresentada a arquitetura da solução proposta. Na seção 4 são mostrados os resultados obtidos. Já na seção 5 são apresentadas as conclusões, contribuições e trabalhos futuros.

2. TRABALHOS RELACIONADOS

Nessa seção serão apresentados trabalhos que utilizaram tecnologias que estão envolvidas nesse trabalho. Serão apresentados trabalhos que usaram ontologias para sugestão de conteúdo para TV Digital juntamente com agentes de *software* e outros trabalhos que usam metaheurísticas para otimização de parâmetros de transmissão. Não foi encontrado na literatura um trabalho único que aborde todas as tecnologias aplicadas neste trabalho.

2.1. Ontologias

Dentro do contexto de televisão digital foram publicados vários trabalhos usando ontologias para a representação dos metadados dos programas, filmes, esportes e diversos conteúdos disponíveis aos usuários.

Na abordagem de [5] os autores propõem um sistema para recomendação automática de conteúdos para programas transmitidos por canais de TV Digital através do uso das tecnologias da Web semântica, ontologias na linguagem OWL e agentes de *software*.

A Figura 1 mostra a estrutura da ontologia que foi montada, pode-se observar uma estrutura parecida com a ontologia montada nesse trabalho, contudo, com uma abrangência maior, pois os autores englobaram vários tipos de conteúdo como cinema,

filmes, séries, entretenimento, shows musicais, dentre outros. Devido ao foco principal do artigo ser a otimização da transmissão do conteúdo optou-se pela diminuição do escopo da área de conhecimento, apenas para cinema.

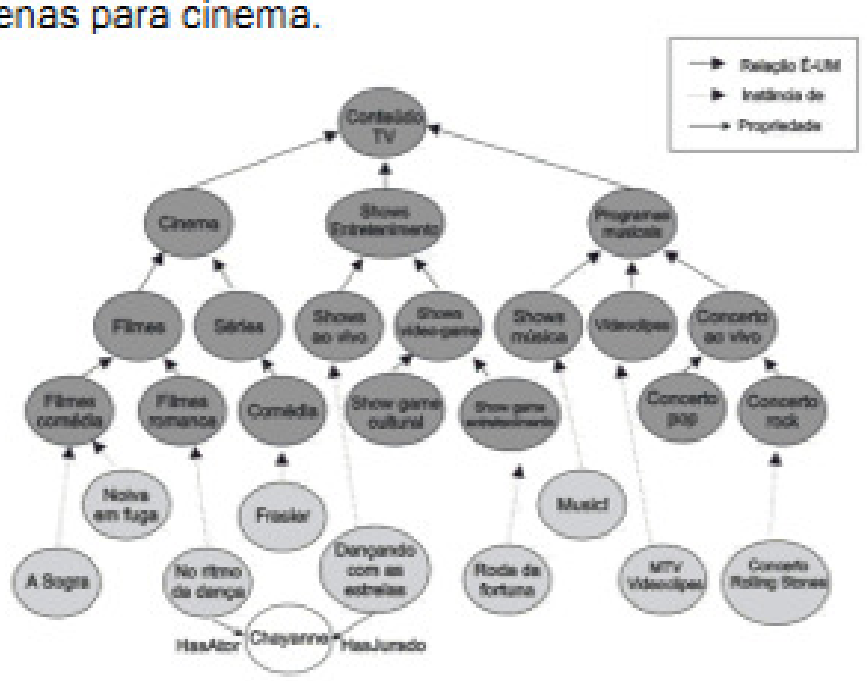

Fig. 1. Estrutura da Ontologia do sistema AVATAR

Na arquitetura de agentes desenvolvida por [5] o Agente Recomendador tem como objetivo recomendar conteúdo de acordo com os conteúdos assistidos no passado pelo usuário. Para isso, são usadas duas estratégias para calcular a similaridade entre os conteúdos assistidos com aqueles que serão sugeridos através da similaridade semântica hierárquica e através da similaridade semântica inferencial.

A similaridade semântica hierárquica leva em consideração a hierarquia É-UM. No caso do ancestral comum ser um nível anterior a similaridade é grande e quanto mais distante menor é a similaridade. No exemplo da Figura 1, pode-se observar que existe similaridade entre os filmes Noiva em Fuga e No Ritmo da Dança, pois possuem o ancestral comum Filme [5].

A similaridade semântica inferencial é usada para descobrir relações implícitas entre os conteúdos. Nesse caso são analisadas características semânticas comuns entre conteúdos da ontologia. No exemplo da Figura 1, é possível observar similaridade entre o filme No Ritmo da Dança com o programa Dançando com as estrelas, pois em ambos há a participação de *Chayanne*, no filme como ator e no programa como jurado [5].

2.2. Metaheurísticas

As metaheurísticas usadas nesse trabalho são a Busca Tabu proposta por Fred Glover em 1986 [5], o Algoritmo Genético proposto por John Holland em 1975 [6] e um Algoritmo Híbrido que faz a combinação entre AG e BT.

O uso de metaheurísticas para otimização de parâmetros de transmissão de dados para IPTV foi abordado por [7]. Nesse trabalho o autor apresenta o desenvolvimento de um modelo computacional baseado na aplicação de metaheurísticas sobre uma plataforma IPTV, com o objetivo de encontrar a melhor configuração dos parâmetros de transmissão considerando os tipos de usuários e dispositivos de recepção.

No trabalho de [8] a autora apresenta um sistema de busca de políticas de configuração dos parâmetros de codificação de vídeo para TV digital usando o padrão H.264. Essa busca é feita através do uso das metaheurísticas Busca Tabu e Algoritmos Genéticos. Nesse trabalho foi desenvolvido um Algoritmo Híbrido, a partir do uso dessas duas metaheurísticas, onde foi utilizado o poder de intensificação da Busca Tabu em conjunto com o poder de diversificação dos Algoritmos Genéticos. Este Algoritmo Híbrido permite a realização de experimentos abordando os problemas de otimização combinatória inerente a arquitetura dos sistemas de codificação e decodificação de sinais de vídeo.

A Tabela 1 mostra uma comparação entre alguns trabalhos encontrados na literatura, pode-se observar que nas abordagens dos demais autores é feita ou a sugestão de conteúdo ou a otimização da transmissão. Já na abordagem desse trabalho é feita tanto a sugestão quanto a otimização da transmissão do conteúdo até o usuário final.

Tabela 1. Comparação de trabalhos relacionados

Autor	Sugestão	Otimização
Rey-Lopez, Diaz-Redondo, Fernández-Vilas and Pazos-Arias (2006)	X	
Fernández, et al. (2006)	X	
Araújo and Ricarte (2010)	X	
Tsinaraki, Polydoros and Christodoulakis (2004)	X	
Kandavanam, Botvich and Balasubramaniam (2009)		X
Weissheimer (2011)		X
Link (2011)		X
Nossa Abordagem	X	X

3. ARQUITETURA DA SOLUÇÃO PROPOSTA

Com o SBTVD é possível termos um canal de interatividade, desde que se tenha conexão de internet para realizar a comunicação com a emissora. Partindo dessa premissa, e que existe uma autenticação do usuário ao conectar no servidor de aplicações multimídia do lado da emissora, é possível estabelecer um canal de comunicação entre o usuário que está assistindo um determinado programa com a emissora que está transmitindo o conteúdo [13]. A Figura 2 ilustra essa dinâmica.

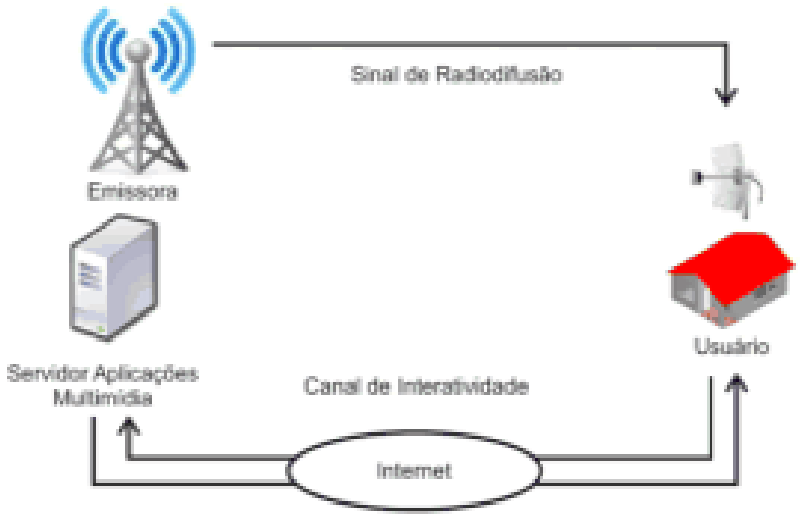

Fig. 2. Comunicação da Emissora com o Usuário

A partir disso, foram criadas duas ontologias, uma ontologia tem como objetivo fazer uma classificação do conteúdo existente no servidor de aplicações multimídia, por exemplo, classificação dos filmes por gênero, principais autores do filme, ano de lançamento, etc. Outra ontologia foi desenvolvida para descrever o perfil do usuário, por exemplo, é possível identificar se o usuário gosta de assistir filmes do gênero ação, comédia, suspense, dentre outros.

Dessa forma, com as duas ontologias é possível cruzar o perfil do usuário com o conteúdo disponível no servidor de aplicações multimídia. Para isso foi criado um agente de *software* que lê o perfil do usuário, pesquisa o conteúdo disponível no servidor de aplicações multimídia e envia essa sugestão para o usuário. Essas sugestões são enviadas em um intervalo de tempo que pode ser determinado pela emissora.

Após o usuário aceitar uma sugestão de conteúdo, é estabelecido um canal de comunicação entre servidor de aplicação multimídia e o usuário. Por outro lado, o agente otimizador, que é executado de tempos em tempos no servidor, determina a forma mais otimizada para transmitir esse conteúdo do servidor até o usuário final. Esse agente roda as metaheurísticas que determinam os parâmetros e a melhor política para a transmissão de acordo com o perfil definido pelas ontologias.

A Figura 3 mostra um exemplo prático do funcionamento da arquitetura proposta. No módulo de sugestão de conteúdo existem os agentes consultores, eles são responsáveis por analisar o perfil do usuário, nesse caso temos dois usuários, cada um com um perfil descrito através de uma ontologia. Após conhecer o perfil cada agente consultor faz uma pesquisa na ontologia do servidor para verificar quais os conteúdos são apropriados para cada usuário. Suponhamos que um usuário goste de assistir filmes de ação, assim o agente irá procurar por conteúdos relacionados a filmes de ação.

Após o aceite de um dos conteúdos sugeridos o controle passa para o módulo de transmissão de conteúdo. Nesse módulo temos o agente coordenador e os agentes otimizadores. O agente coordenador irá acionar os outros agentes otimizadores que por usa vez irão executar as metaheurísticas para analisar o perfil do usuário e de acordo com os recursos que ele tem determinar os melhores parâmetros e a melhor política de transmissão do conteúdo. Ao final cada agente otimizador irá transmitir o resultado encontrado das metaheurística BT, AG e AH para o agente coordenador. O agente coordenador irá analisar qual deles obteve o melhor resultado sendo que esses parâmetros serão os escolhidos para a transmissão do conteúdo entre o servidor de aplicações multimídia e o usuário. Está se levando em consideração que a transmissão de conteúdo se dará através do IPTV, ao final da transmissão o canal de transmissão é encerrado e o perfil do usuário é atualizado com o filme que ele assistiu.

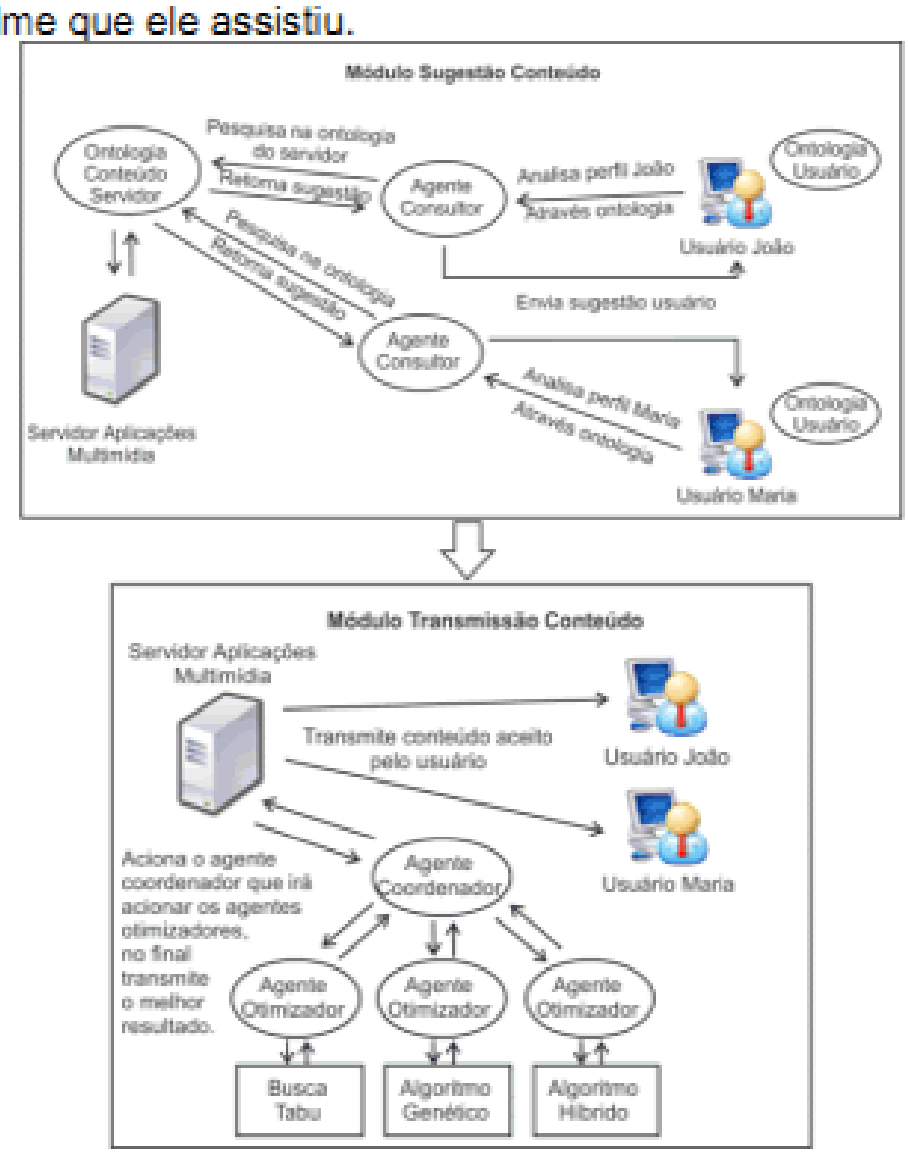

Fig. 3. Comunicação da Emissora com o Usuário

3.1. Ontologia do Servidor de Aplicações Multimídia e do Usuário

A ontologia do servidor de aplicações multimídia tem como objetivo descrever o conteúdo multimídia que está armazenado no servidor. Devido a grande diversidade de conteúdos que um servidor pode armazenar esse trabalho irá restringir o escopo dessa ontologia apenas para filmes.

A ontologia do usuário tem como objetivo descrever o perfil dos usuários que se conectam no servidor. Devido a grande quantidade de dados que podem ser descritos sobre um usuário o escopo foi limitado a algumas informações consideradas relevantes.

A Figura 4 mostra a ontologia do servidor de aplicações multimídia em conjunto com a ontologia do usuário. Os círculos claros indicam as classes da ontologia, já os círculos escuros indicam as instâncias das classes.

Fig. 4. Estrutura da Ontologia

3.2. Formulação Matemática

Nessa seção é apresentada a Função Objetivo (FO) e as restrições relacionadas à FO. Para cada variável de decisão da FO será apresentado o intervalo de valores viáveis. O resultado da formulação é ter o melhor aproveitamento e distribuição possível da Largura de Banda (LB) da emissora para poder atender um número variado de clientes conectados ao serviço. Esse problema pode ser visto como um problema de *Job Shop Scheduler* [13].

$$\text{MAX FO} = \alpha\left(\sum_{n=1}^{nLD} vLD_{[n]}\right) + \beta\left(\sum_{n=1}^{nSD} vSD_{[n]}\right) + \gamma\left(\sum_{n=1}^{nHD} vHD_{[n]}\right) + \delta\left(\sum_{n=1}^{nP1} vP1_{[n]}\right) + \omega\left(\sum_{n=1}^{nP2} vP2_{[n]}\right) + \theta\left(\sum_{n=1}^{nST} aST_{[n]}\right) + \rho\left(\sum_{n=1}^{n51} a51_{[n]}\right) \quad (1)$$

Sujeito a:

$0{,}01 \leq vLD \leq 1{,}00$ (2)

$1{,}00 \leq vP1 \leq 2{,}00$ (5)

$2{,}00 \leq vSD \leq 5{,}00$ (3)

$5{,}00 \leq vP2 \leq 10{,}00$ (6)

$10{,}00 \leq vHD \leq 18{,}00$ (4)

$0{,}096 \leq aST \leq 0{,}256$ (7)

$0{,}384 \leq a51 \leq 1{,}00$ (8)

$LB = 500$ (9)

$\alpha + \delta + \beta + \omega + \gamma + \theta + \rho > 0$ (10)

$nLD, nSD, nHD, nP1, nP2, nST, n51 \in Z^{*}$ (11)

$$\left(\sum_{n=1}^{nLD} vLD_{[n]}\right) + \left(\sum_{n=1}^{nSD} vSD_{[n]}\right) + \left(\sum_{n=1}^{nHD} vHD_{[n]}\right) + \left(\sum_{n=1}^{nP1} vP1_{[n]}\right) + \left(\sum_{n=1}^{nP2} vP2_{[n]}\right) + \left(\sum_{n=1}^{nST} aST_{[n]}\right) + \left(\sum_{n=1}^{n51} a51_{[n]}\right) \leq LB \quad (12)$$

Onde:

- LB = largura de banda total da emissora;
- vLD = qualidade de vídeo LD;
- vP1 = qualidade de vídeo P1;
- vSD = qualidade de vídeo SD;
- vP2 = qualidade de vídeo P2;
- vHD = qualidade de vídeo HD;
- aST = qualidade de áudio estéreo;
- a51 = qualidade de áudio multicanal 5.1;
- nLD = número de clientes conectados como LD;
- nP1 = número de clientes conectados como P1;
- nSD = número de clientes conectados como SD;
- nP2 = número de clientes conectados como P2;
- nHD = número de clientes conectados como HD;
- nST = número de clientes conectados como áudio estéreo;
- n51 = número de clientes conectados como áudio 5.1;
- α = nível de importância da transmissão LD;
- δ = nível de importância da transmissão P1;
- β = nível de importância da transmissão SD;
- ω = nível de importância da transmissão P2;
- γ = nível de importância da transmissão HD;
- θ = nível de importância da transmissão Áudio Estéreo;
- ρ = nível de importância da transmissão Áudio Multicanal 5.1;
- Todas variáveis de decisão aceitam números reais.

4. RESULTADOS

Os experimentos para a validação do módulo de sugestão de conteúdo foram feitos através de simulações e análise dos requisitos para verificar se o sistema está sugerindo conteúdo corretamente de acordo com o perfil do usuário. Foram criados três perfis de usuários com gostos diferentes sobre filmes, também foram cadastrados vários filmes e gêneros de modo que o perfil do usuário possa ser cruzado com o banco de dados dos filmes.

O módulo de transmissão de conteúdo foi validado através da média harmônica calculada sobre os resultados gerados pelas metaheurísticas Algoritmo Genético, Busca Tabu e Algoritmo Híbrido.

4.1. Módulo Transmissão Conteúdo

Para realizar os testes uma lista de filmes foi cadastrada e foram criados alguns usuários com diferentes perfis. A Tabela 2 mostra o perfil do usuário com os gêneros preferidos dele e os filmes que ele já assistiu.

Tabela 2. Lista de Usuários

Usuário	Gêneros Preferidos	Filmes Assistidos
João	Ação Comédia	Velozes e Furiosos Recém Casados
Maria	Terror Drama Aventura	Anjos da Morte Assim é a Vida A Lenda do Tesouro Perdido O Mentiroso
Paulo	Ação, Aventura Comédia Suspense Terror	•

A Tabela 3 mostra os filmes que o sistema sugeriu de acordo com o perfil do usuário. Pode-se observar que o agente consultor sugeriu o conteúdo de acordo com o gênero preferido do usuário.

Tabela 3. Lista de Sugestões

Usuário	Filmes Sugeridos
João	Senhor e Senhora Smith

João	Matrix
João	Ace Aventura
João	Duro de Matar
João	Todo Poderoso
João	Se Beber Não Case
Maria	Alma Perdida
Maria	O Código de Da Vinci
Maria	Carandiru
Maria	Assassinato no Presidio
Maria	Gritos Mortais
Maria	A Bolha
Maria	Assédio
Paulo	Senhor e Senhora Smith
Paulo	Matrix
Paulo	Alma Perdida
Paulo	Ace Aventura

4.2. Módulo Otimização Conteúdo

Esse módulo roda de tempos em tempos no servidor de aplicações multimídia. Ele contém os agentes de *software* que irão coordenar as metaheurísticas para otimizar o conteúdo que é transmitido de acordo com o número de usuários conectados no servidor e de acordo com os perfis dos usuários conectados.

A Tabela 4 mostra o primeiro experimento realizado com o Algoritmo Híbrido com uma instância pequena do problema, nesse caso são 3 clientes conectados de cada tipo (LD, P1, SD, P2 e HD), totalizando 15 clientes. Desses clientes, 5 recebem áudio estéreo e 10 recebem áudio 5.1. A Largura de Banda (LB) do servidor foi limitada a 70 MBit/s. O valor da Função Objetivo (FO) é de 69,504 Mbit/s e a média harmônica é de 0,723 Mbit/s.

Tabela 4. Tabela com as taxas de transmissão

Algoritmo Híbrido						Harmônica	Nr. Clientes
LD	0,732	0,772	0,581			0,685	3
P1	1,216	1,252	1,604			1,337	3
SD	2,754	2,601	2,019			2,414	3
P2	5,106	7,242	5,023			5,628	3
HD	10,100	10,664	10,666			10,470	3
ST	0,180	0,200	0,206	0,199	0,191	0,195	5
A51	0,480	0,554	0,727	0,676	0,576	-	5
A51	0,765	0,703	0,595	0,598	0,522	0,607	5
				FO =>	69,504	0,723	15

A Figura 5 mostra através de um gráfico o valor da métrica, ou seja, a média harmônica dos resultados obtidos pelo AG, BT e AH da mesma instância do problema citado na Tabela 4. Pode-se observar que a BT e o AG obtiveram resultados semelhantes. O AH obteve o melhor desempenho com uma média harmônica da taxa de transmissão de 0,723 Mbit/s.

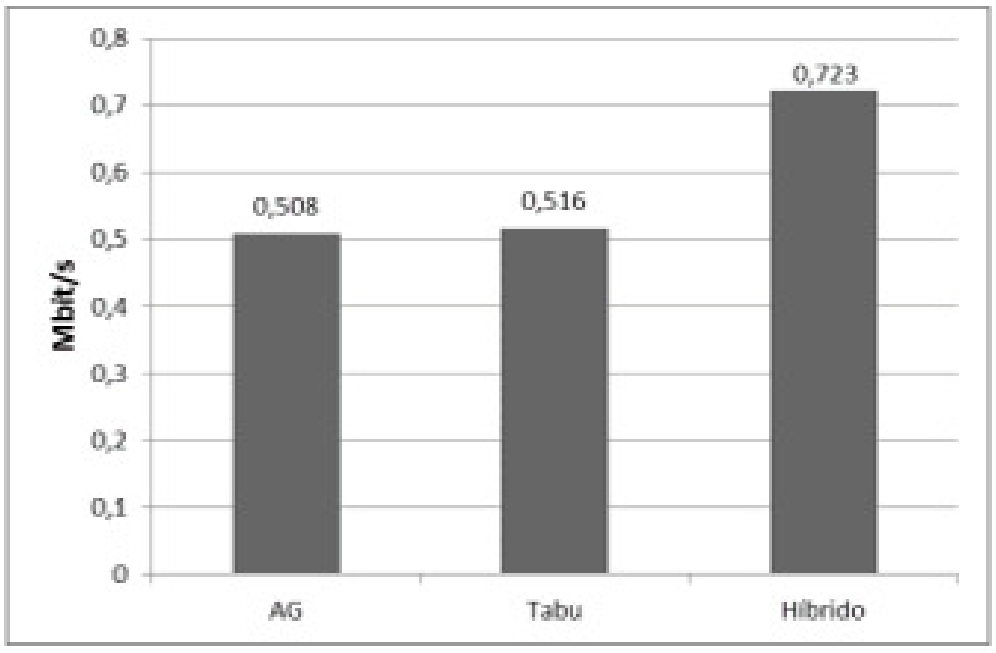

Fig. 5. Métrica aplicada nos resultados

O gráfico da Figura 6 mostra o desvio padrão obtido por cada metaheurística. Novamente, o AH obteve o melhor resultado sendo que desvio padrão do AH foi de 0,023 Mbit/s, da BT de 0,091 Mbit/s e do AG de 0,107 Mbit/s.

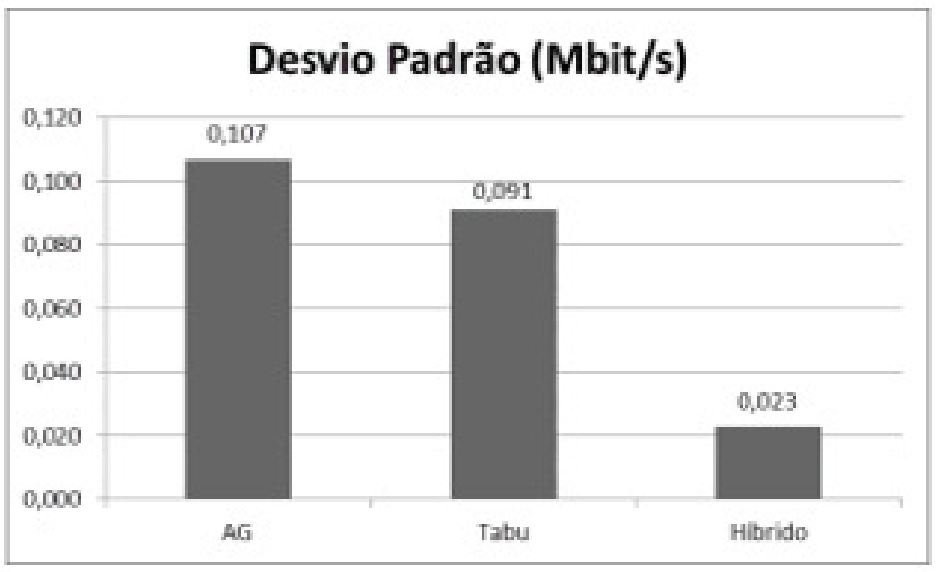

Fig. 6. Desvio padrão das metaheurísticas

Além das instâncias citadas, as metaheuristicas foram testadas com uma instância contendo 15 mil clientes conectados. Sendo 3 mil de cada tipo (LD, P1, SD, P2 e HD). Desses 15 mil, 7,5 mil receberam qualidade de áudio estéreo e 7,5 mil receberam qualidade de áudio 5.1. A Figura 7 mostra a média harmônica dos resultados obtidos por cada metaheuristica, novamente o AH obteve o melhor resultado com taxa harmônica de transmissão de 0,426 Mbit/s, sendo que o AG e BT obtiveram resultados muito próximos com taxas de transmissão harmônicas de 0,407 Mbit/s e 0,406 Mbit/s, respectivamente.

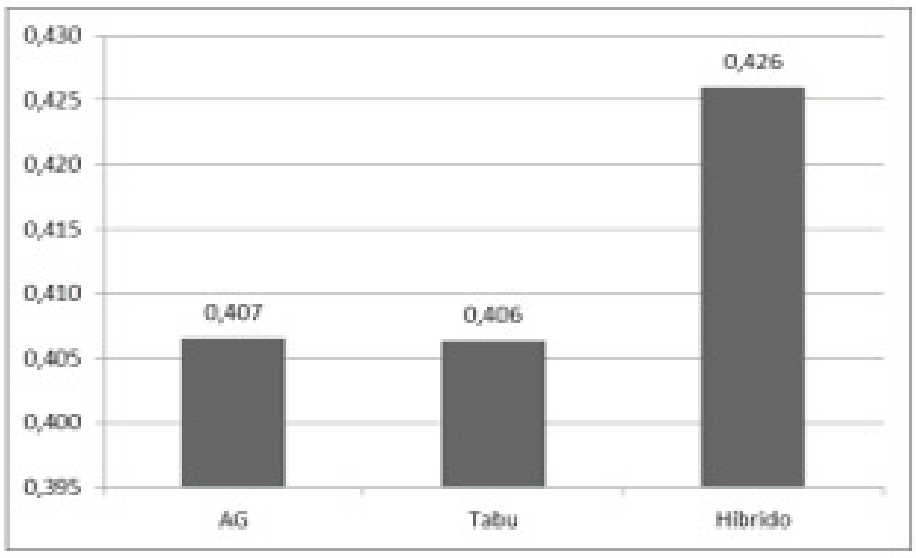

Fig. 7. Métrica dos resultados das metaheurísticas

4. CONCLUSÕES

Esse trabalho apresentou uma arquitetura dividida em dois módulos. O módulo de sugestão teve como objetivo sugerir conteúdo aos usuários. Já o módulo de transmissão de conteúdo teve como função otimizar o conteúdo que será transmitido para o usuário.

Através da implementação e dos experimentos que foram realizados, observou-se que foi possível atender aos requisitos especificados. Os experimentos mostraram que a sugestão de conteúdo ocorreu conforme o perfil do usuário. No módulo de otimização de conteúdo as metaheurísticas obtiveram bons resultados. O Algoritmo Híbrido (AH) obteve os melhores resultados tanto no valor da média harmônica, que foi utilizada como métrica, como no valor do desvio padrão.

Como contribuições desse trabalho pode-se destacar a definição de uma metodologia para sugerir conteúdo adequado aos usuários usando ontologias e agentes de *software* aplicada a uma arquitetura do canal de retorno. O desenvolvimento de um AG, BT e AH aplicado ao modelo de transmissão de conteúdo do servidor de aplicações multimídia, além da avaliação quantitativa feita através da métrica para avaliar a qualidade dos resultados gerados.

Como trabalhos futuros pretende-se desenvolver um algoritmo mais complexo para sugestão de conteúdo conforme proposto na literatura por [5] e [15]. Além disso, pretende-se realizar novos experimentos com as metaheurísticas utilizando novos operadores de mutação, cruzamento e intensificação com o objetivo de melhorar os resultados encontrados até o momento.

REFERÊNCIAS

[1] Bip, 2001. Boletim de Informação Para Publicitários. Available at: <http://comercial2.redeglobo.com.br/biponline/BIP/BIP%20577.pdf>. Acessed in May, 2011.

[2] Carvalho, F. B. S., 2006. Aplicação de Transmissão de Dados via Rede Elétrica Para o Canal de Retorno em Televisão Digital, Campina Grande, PB.

[3] Manhães, M. A. R., Shieh, P J., Lamas, A. C.; Macedo, P. E. O., 2005. Canal de Interatividade em TV Digital., Cad CPqD Tecnologia, páginas 29-36.

[4] Fernández, Y., Arias, J., Nores, M.,; Solla, A., Cabrer, M., 2006 Avatar: An Improved Solution for Personalized TV based on Semantic Inference. IEEE Trans. on Consumer Electronics: Vol. 52(1).

[5] Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Computers and Operations Research, Oxford, UK, v. 13, n. 5, p. 533-549.

[6] Holland, J. H., 1975. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. 1. ed. Michigan: University of Michigan Press.

[7] Weissheimer, C. H., 2011. Aplicação de Metaheurísticas para o modelo Internet Protocol Television. Programa Interdisciplinar de Pós-Graduação em Computação Aplicada – PIPCA. Mestrado Acadêmico. Proposta de Dissertação São Leopoldo.

[8] Link, I. C. D. C., 2011. Um Sistema de Codificação de Vídeo para TV Digital - SBTVD. Programa Interdisciplinar de Pós-Graduação em Computação Aplicada – PIPCA. Mestrado Acadêmico. Proposta de Dissertação São Leopoldo.

[9] Rey-Lopez, M., Díaz-Redondo, R., Fernández-Vilas, A.,Pazos-Arias, J., 2006. Entercation Experiences: Engaging Viewers in Education through TV Programs. Proc. of 4th European Conference on Interactive Television (EuroITV 2006). Athens, Greece.

[10] Araújo, R. C., Ricarte, I. L. M., 2010. Ontologias Para Integração de Metadados da Indústria de Broadcast Multimídia no Contexto da TV Digital. Revista de Radiodifusão, n. 4.

[11] Tsinaraki, C.; Polydoros, P.; Christodoulakis, S., 2004. Integration of OWL ontologies in MPEG-7 and TV-Anytime compliant Semantic Indexing. Proc. of 16th International Conference on Advanced Information Systems Engineering.

[12] Kandavanam, G., Botvich, D., Balasubramaniam, S., 2009. An Optimization based Approach to Maximizing QoS Assurance for IPTV Triple Play Services on the Internet Backbone. IEEE 34th Conference on Local Computer Networks (LCN 2009). Zürich, Switzerland. p. 20-23.

[13] CPqD. 2006. Arquitetura de Referência. Sistema Brasileiro de Televisão Digital Terrestre. Disponível em: <http://sbtvd.cpqd.com.br>. Accessado em Maior de 2011.

[14] Papaioannou, G.,Wilson, J. M., 2008. Fuzzy extensions to Integer Programming models of cell-formation problems in machine scheduling. Annals of Operations Research, 166, 163-182.

[15] De Paula, L., Villaça, R., Magalhães, M., 2011. Organization of multimedia data for conceptual search based on ontologies. Journal of the Brazilian Computer Society. Springer London. Volume 17. Pág. 241-254. Issn: 0104-6500.

Um modelo computacional baseado em metaheurísticas aplicado a Transmissão Internet Protocol Television (IPTV)

Arthur Gómez (atgomezbr@gmail.com)
Carlos Alfredo Weissheimer Junior (carlosawe@gmail.com)

UM MODELO COMPUTACIONAL BASEADO EM METAHEURÍSTICAS APLICADO A TRANSMISSÃO INTERNET PROTOCOL TELEVISION (IPTV).

Arthur Tórgo Gómez [1], *Carlos Alfredo Weissheimer Junior* [2] *Daniel Formolo* [3]

[1] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, atgomezbr@gmail.com
[2] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, carlosawe@gmail.com
[3] Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, formolo@gmail.com

Resumo: Atualmente, o *Internet Protocol Television* (IPTV) é alvo de várias pesquisas, pois pode gerar muitos benefícios para a sociedade, tais como o suporte a interatividade e o aumento da interoperabilidade com redes domésticas. Este trabalho apresenta o desenvolvimento e a implementação de um modelo computacional, baseado na aplicação de metaheurísticas, sobre uma plataforma IPTV com o objetivo de desenvolver um algoritmo híbrido utilizando as metaheurísticas Busca Tabu (BT) e Algoritmo Genético (AG). Esse algoritmo permite que seja feita a análise e o estudo dos seguintes parâmetros: taxa de transmissão, qualidade de áudio, número de clientes e largura de banda visando encontrar a melhor configuração dos mesmos para a transmissão IPTV dadas às características do cliente. Após a validação do algoritmo são realizados experimentos que auxiliam a compreensão da dinâmica do sistema e viabilizam encontrar uma boa solução passível de simulação no simulador de rede Network Simulator 3 (NS3).

Palavras chave: Transmissão, IPTV, Metaheurísticas, Redes de computadores.

1. INTRODUÇÃO

A tecnologia Internet *Protocol Television* (IPTV) é um fator impactante na sociedade. Ela tem sido explorada por diferentes meios de transmissão para a entrega de conteúdo multimídia sobre serviço baseado em *Internet Protocol* (IP) [10-12]. Na Figura 1 é apresentada a uma estrutura básica desta tecnologia.

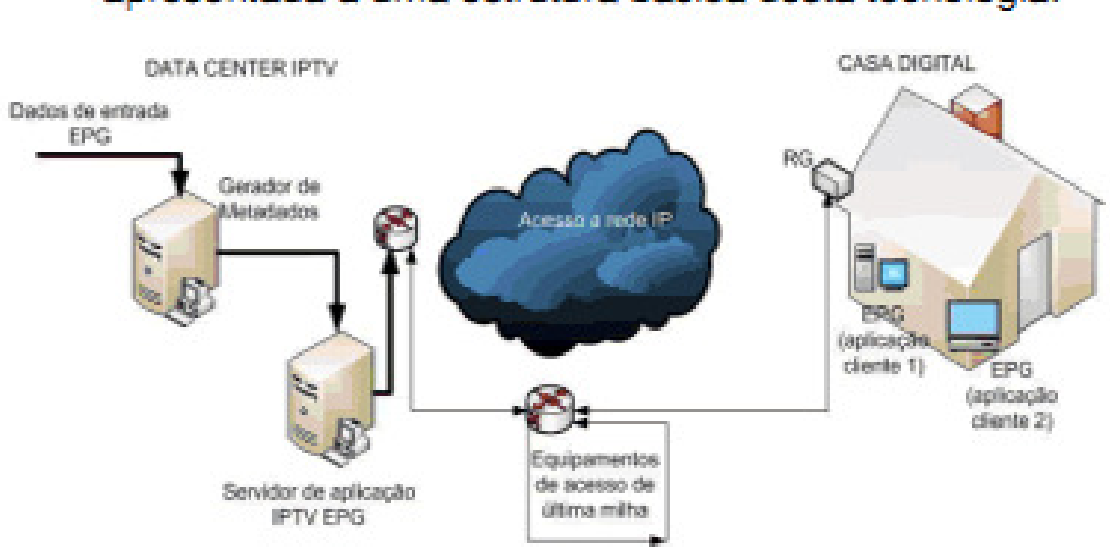

Fig1. Estrutura simplificada IPTV [17].

Na Figura 1, observa-se o Data Center IPTV por onde entram os dados que são codificados e encaminhados para o servidor de aplicação e transmitidos por meio da rede IP, onde encontram-se os equipamentos de rede como switches e roteadores, até o usuário final. Através de equipamentos como o Set-Top-Box o usuário final decodifica esses dados. A casa digital representa um composto de equipamentos que podem suportar essa transmissão, como TVs e computadores [17].

Um dos principais motivos para a aposta em IPTV é que ela pode prover um serviço de TV diferenciado. Essa cesta de serviços de entretenimento, a interação do usuário e o controle do assinante sobre sua TV, podem trazer um diferencial competitivo significativo para plataformas que utilizam a tecnologia IPTV. Porém, como se trata de uma tecnologia que prima pela qualidade, métodos de otimização que possam agregar valor a arquitetura são importantes.

Este trabalho tem como objetivo desenvolver um algoritmo híbrido utilizando as metaheurísticas Busca Tabu (BT) e Algoritmo Genético (AG), que viabilize a análise e estudo dos parâmetros de taxa de transmissão, qualidade de vídeo, qualidade de áudio, número de clientes e largura de banda visando encontrar a melhor configuração dos mesmos para a transmissão IPTV dadas às características de cada cliente.

Esse artigo está estruturado como segue. Na seção 2 são apresentados os trabalhos relacionados. Na seção 3 é apresentado o modelo utilizado, essa seção tem como subseções: a arquitetura do modelo, o algoritmo MCIPTV e a formulação do modelo matemático. Na seção 4 são descritos os resultados obtidos. E por fim, seção 5 são apresentados as conclusões, contribuições e trabalhos futuros.

2. TRABALHOS RELACIONADOS

Essa seção tem como objetivo descrever algumas pesquisas já realizadas nas áreas de atuação deste trabalho. Nas suas subseções serão apresentadas tecnologias e métodos utilizados tanto na otimização de parâmetros como na transmissão IPTV. Os trabalhos relacionados para neste artigo fazem referência aos assuntos abordados em dois momentos distintos. O primeiro é apresentado na seção 2.1 referindo-se aos trabalhos relacionados à transmissão IPTV. O segundo momento é descrito na seção 2.2 onde são apresentados trabalhos que utilizaram metaheurísticas para alguma solução dentro da plataforma do sistema brasileiro de Televisão digital (SBTV).

2.1. *Internet Protocol Television* (IPTV)

Trabalhos como o de Caja [1], e de Neto [13], conseguiram representar, de alguma forma, uma melhora na transmissão IPTV, seja ela feita na otimização do algoritmo utilizado pelo protocolo de transmissão ou, através da substituição do mesmo por outro que obtenha um desempenho melhor.

Neto [13] utilizou um simulador de rede para a realização dos testes de transmissão e tempo de resposta. Sendo possível, através de um ambiente de simulação desenvolvido com base nas características (topologia e ocupação) de uma rede de distribuição real, o desempenho desses mecanismos foi avaliado em diferentes cenários.

Kukhmay, et al., [9], descreveram que a partir de testes realizados em alguns protocolos obteve-se a informação que vários métodos de transmissão podem ser utilizados a fim de otimizar a distribuição IPTV, reduzindo a carga imposta na rede. E ainda, em todos os testes houve perdas de pacotes, porém em nenhum dos casos essa perda foi expressiva.

Uma pesquisa comparativa realizada por Shihab e Cai [20], em tecnologias de distribuição IPTV destinadas às redes domésticas de banda larga, levantou os prós e os contras em termos de taxa de transmissão. Sendo assim possível mencionar que uma das variáveis da formulação do modelo matemático deste trabalho, a taxa de transmissão impacta não só no custo do serviço oferecido, mas também na qualidade do mesmo.

Através da parceria entre as empresas Cisco e Microsoft foi possível melhorar a solução de IPTV. Os clientes não precisam mais se conectar baseados em um formato ou encontrar os servidores corretos para assegurar a melhor qualidade de entrega. O Cisco IPTV *Viewer* traz uma lista unificada de programas Windows Media, MPEG e IP *multicast*. Ao mesmo tempo, os visualizadores IPTV são direcionados ao servidor mais apropriado para otimizar a eficiência da rede. O servidor de controle IPTV 3411 da Cisco gerencia, simplifica e escala instalações de vídeo em rede de serviços múltiplos, controlando o uso da largura de banda, a distribuição de conteúdo para servidores remotos, o planejamento de distribuição de informações baseada em IP e o monitoramento da atividade de servidores. Ele ainda seleciona o melhor servidor de vídeo em proximidade ao visualizador para maior controle da rede em instalações maiores [2].

Estudos sobre as métricas de distância entre nós vêm sendo realizados para tentar melhorar a qualidade dos serviços oferecidos pelo IPTV. Esse fato retrata uma necessidade eminente na melhoria da transmissão IPTV para que ele se torne comercial. Trabalhos como Padmanabhan e Subramanian [18], e Padmanabhan; et al. [19] abordam técnicas para tentar estimar a localização de um nó baseado em seu endereço IP. A ideia é utilizar a localização entre nós da rede para que seja possível realizar a escolha de um serviço mais adequado para o usuário e assim consumindo menos largura de banda. Isto é especialmente interessante para questões de otimização, pois possibilita a construção de árvores de distribuição *multicast* mais eficientes para redes.

2.2. Metaheurísticas

Nesse trabalho são usadas as metaheurísticas Busca Tabu proposta por Fred Glover [4]. O Algoritmo Genético proposto por John Holland [7]. E a partir delas é desenvolvido um Algoritmo Híbrido para a combinação entre o AG e a BT.

Em trabalhos anteriores os autores realizaram experimentos no modelo computacional primeiramente com o AG [22] e posteriormente com o híbrido [21]. Em ambos uma solução foi encontrada a partir de uma formulação matemática e do modelo desenvolvido. Porém não chegaram a ser utilizados simuladores de rede para a análise da solução perante a transmissão IPTV.

Em outro trabalho [11], a autora apresenta um sistema de busca de políticas de configuração dos parâmetros de codificação de vídeo para TV digital. Essa busca é feita, através do uso das metaheurísticas Busca Tabu e Algoritmos Genéticos, as mesmas utilizadas nesse trabalho para encontrar a melhor configuração para transmissão IPTV. A autora desenvolve um Algoritmo Híbrido, a partir do uso dessas duas metaheurísticas. Este Algoritmo Híbrido permite a realização de experimentos abordando os problemas de otimização combinatória conforme a Arquitetura do Modelo Sistêmico do projeto Desenvolvimento de uma Plataforma de Convergência Digital: TV Digital / IPTV / Dispositivos Móveis, financiado pelo FUNTTEL, Chamada Pública FINEP/FUNTTEL, características 02/2009, (DIGICONV) [5].

3. MODELO

Este capítulo apresenta o modelo computacional proposto neste trabalho, através de sua arquitetura e formulação matemática. A arquitetura apresenta atua nos módulos geração de meta-dados e servidor de aplicação ilustrado na Fig.1. A seção 3 descreve o modelo sugerido e apresenta sua formulação, explicando como ambos pretendem atuar sobre o problema abordado.

3.1. Arquitetura do modelo

Um modelo de transmissão de IPTVD é utilizado não só por grandes empresas como a Cisco, que utilizam suas próprias variações do mesmo com intuito de melhorar o enquadramento dentro do seu objetivo, mas também por projetos como o DIGICONV. Na Figura 2 é ilustrado onde o modelo atua.

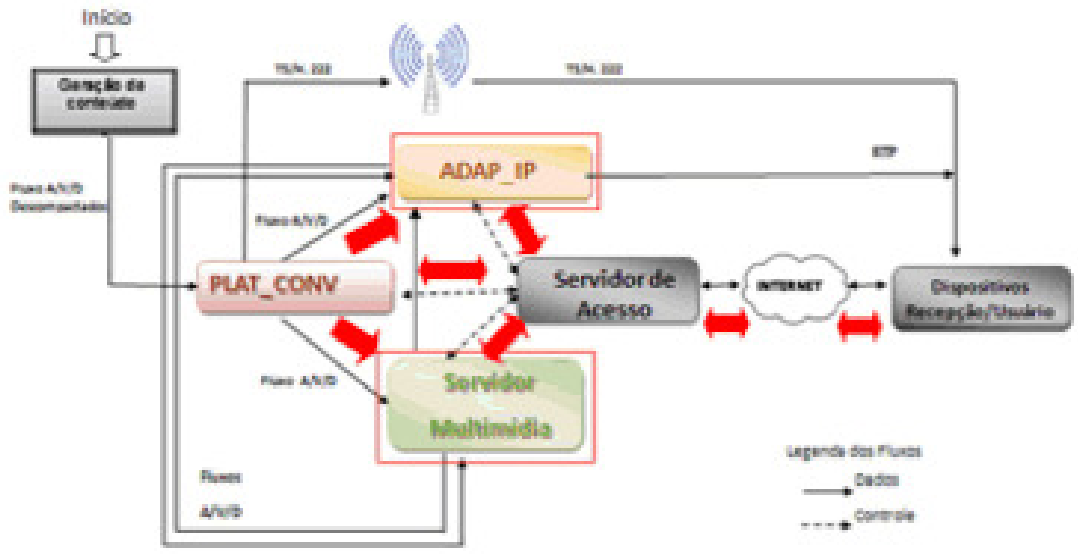

Fig. 2 - Arquitetura da Plataforma de Convergência [5]

Na Figura 2 observa-se que os arquivos de vídeo, dados e áudios entram na arquitetura, pelo processo de geração de conteúdo. A plataforma de convergência (PLAT_CONV) envia os arquivos de áudio vídeo e dados para o servidor de acesso, que classifica o tipo de transmissão quanto à transmissão em tempo real e a transmissão solicitada pelo usuário. Após essa definição, os arquivos, caso seja uma transmissão em tempo real, são codificados pela PLAT_CONV, responsável pela codificação de sinais audiovisuais, a serem empregados em distintos mecanismos de transmissão [5][8].

Em seguida as informações são enviadas pela PLAT_CONV ao Adaptador de Interface de rede (ADAP_IP) e transmitidas para o usuário via protocolo RTP. Caso a transmissão não seja em tempo real, a PLAT_CONV envia as informações para o servidor multimídia onde estão armazenados os arquivos de (áudio vídeo e dados). Estas informações podem ser solicitadas pelo usuário, a partir da sua autenticação no servidor de acesso feita, através da Internet e pelo dispositivo do usuário/cliente. Após a definição do arquivo escolhido no servidor multimídia o mesmo é enviado para o ADAP_IP e transmitido até o usuário/cliente através do RTP [5].

3.2. Algoritmo MCIPTV

O algoritmo MCIPTV considera em sua solução inicial os seguintes parâmetros: taxa de transmissão, qualidade de áudio, número de clientes e largura de banda visando encontrar a melhor configuração dos mesmos para a transmissão IPTV dadas às características do cliente. Uma solução inicial aleatória viável é gerada para a aplicação da BT no MCIPTV. Para as perturbações da Busca Tabu foram gerados movimentos para a geração da vizinhança onde serão explorados todos os intervalos possíveis das variáveis de decisão consideradas no modelo. Este é um problema de seleção de partes que é NP-Completo [6]. A estrutura do MCIPTV é ilustrada pela Figura 3.

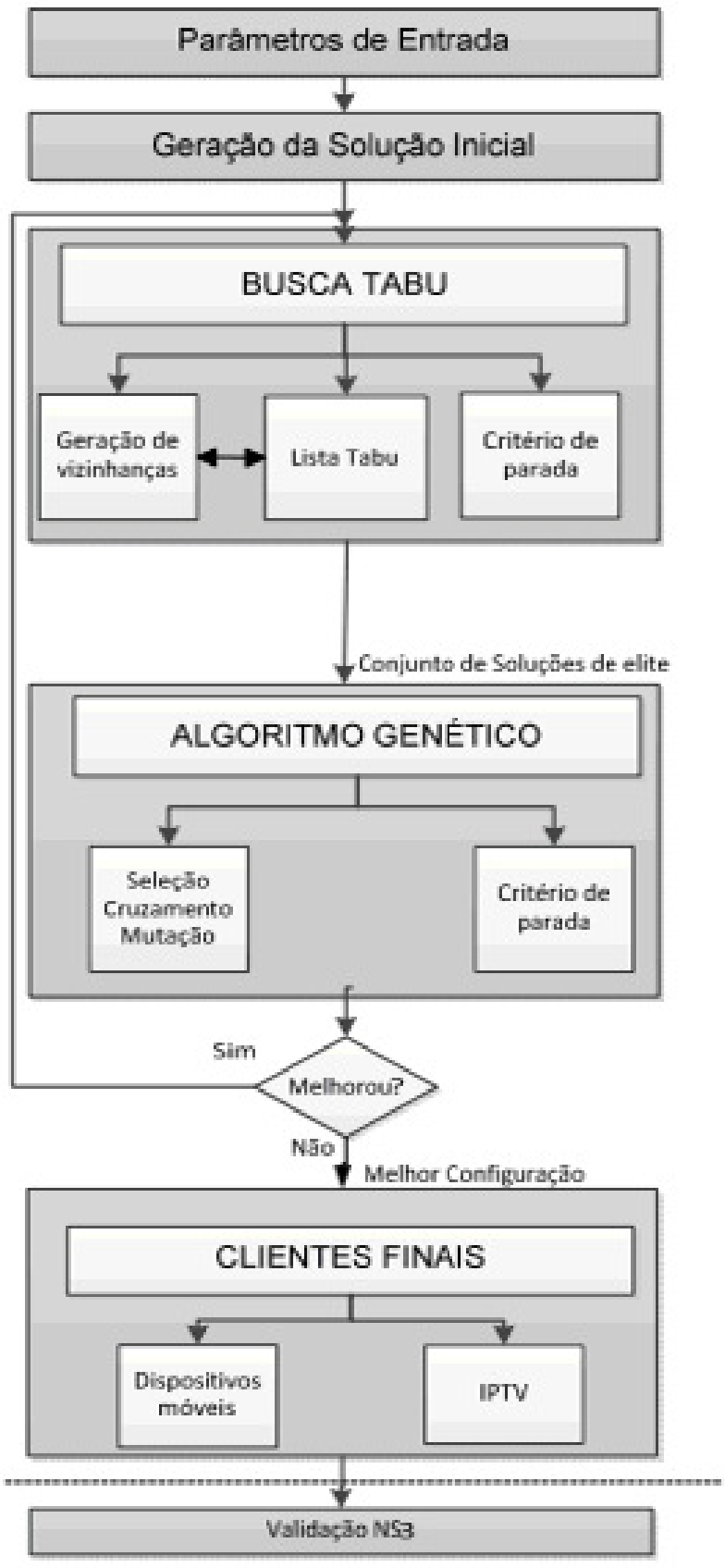

Fig. 3 - algoritmo MCIPTV

Para que o MCIPTV funcione de maneira correta, algumas restrições devem ser respeitadas visando encontrar a melhor configuração dos parâmetros de entrada a serem trabalhadas pelo modelo proposto a partir de uma solução inicial aleatória, sempre respeitando aos limites impostos pelo mundo real. Isso é apresentado na seção 3.2.

3.3. Formulação do modelo matemático

Nesta seção é apresentada a formulação do modelo matemático. Inicialmente representado pela definição das variáveis de decisão que irão compor a Função Objetivo (F.O.) e suas restrições. A F.O. segue ao critério de otimalidade de maximização e tem suas variáveis de decisão baseadas em características da transmissão IPTV.

Deste modo, os requisitos identificados para a construção de uma dada rede, num determinado local, com serviços definidos são otimizados com a escolha dos perfis e parâmetros (variáveis de decisão) mais adequados à situação particular, essas variáveis

podem ser, por exemplo, a taxa de transmissão e largura de banda [3].

O modelo proposto pode ser formalmente descrito com base nas equações desenvolvidas a seguir. Elas levam em conta as tabelas descritas nessa seção que são baseadas em recomendações [8], para que a melhor configuração seja obtida. Segue a formulação matemática proposta:

$$MAX\ FO = Tt + Qv + Q_A + Nc + Lb \quad (1)$$

Onde:

$$Tt = \lambda_{sd}Tt_{SD} + \lambda_{hd}Tt_{HD} + \lambda_{ld}Tt_{LD} + \lambda_{p1}Tt_{P1} + \lambda_{p2}Tt_{P2} \quad (2)$$

$$Qv = \beta_{sd}Q_{SD} + \beta_{hd}Q_{HD} + \beta_{ld}Q_{LD} + \beta_{p1}Q_{vP1} + \beta_{p2}Q_{vP2} \quad (3)$$

$$Lb = \mu_{sd}Lb_{ADSL2+} + \mu_{hd}Lb_{ADSL2} + \mu_{ld}Lb_{VDSL2} + \mu_{p1}Lb_{WINMAX} + \mu_{p2}Lb_{CDMA} \quad (4)$$

$$Nc = \alpha_{sd}Nc_{SD} + \alpha_{hd}Nc_{HD} + \alpha_{ld}Nc_{LD} + \alpha_{p1}Nc_{P1} + \alpha_{p2}Nc_{P2} \quad (5)$$

$$Q_A = \omega_{5.1}Q_{5.1} + \omega_{st}Q_{st} \quad (6)$$

Sujeito a:

$$6 \leq Tt_{HD} \leq 100 \quad (7)$$

$$2 \leq Tt_{SD} \leq 100 \quad (8)$$

$$0.5 \leq Tt_{LD} \leq 100 \quad (9)$$

$$2 \leq Q_{HD} \leq 100 \quad (10)$$

$$0,1 \leq Q_{SD} \leq 100 \quad (11)$$

$$0,01 \leq Q_{LD} \leq 100 \quad (12)$$

$$0,128 \leq Lb_{ADSL2+} \leq 25 \quad (13)$$

$$0,064 \leq Lb_{ADSL2} \leq 12 \quad (14)$$

$$0,128 \leq Lb_{VDSL} \leq 52 \quad (15)$$

$$0,128 \leq Lb_{WINMAX} \leq 15 \quad (16)$$

$$0,01 \leq Lb_{CDMA} \leq 2 \quad (17)$$

$$0,256 \leq Q_{AST} \leq 0,096 \quad (18)$$

$$1 \leq Q_{A5.1} \leq 0,384 \quad (19)$$

$$C \leq TIPTV \quad (20)$$

$$\lambda_{sd}, \lambda_{hd}, \lambda_{ld}, \lambda_{p1}, \lambda_{p2}, \beta_{sd}, \beta_{hd}, \beta_{ld}, \beta_{p1}, \beta_{p2}, \mu_{sd}, \mu_{hd}, \mu_{ld}, \mu_{p1}, \mu_{p2}, \alpha_{sd}, \alpha_{hd}, \alpha_{ld}, \alpha_{p1}, \alpha_{p2}, \omega_{5.1}, \omega_{st} \in Z^* \quad (21)$$

$$Tt, Qv, Q_A, Nc, Lb \geq 0 \quad (22)$$

Para tanto, algumas variáveis devem ser explicadas, a seguir o que cada uma representa.

Tt = Taxa de Transmissão, que pode impactar na qualidade da transmissão.
LB = largura de banda. Define as tecnologias utilizadas pelo cliente podendo até mesmo impactar no não fornecimento da solução enviada para o cliente.
Q_V = qualidade de vídeo. Definida dentro de um range de aceitação, tendo assim uma ligação direta com os formatos de vídeos utilizados.
Q_A = qualidade de áudio;
N_C = Número de Clientes. Impacta na taxa de transmissão e qualidade do serviço prestado que se pretende ter na transmissão de determinado formato de vídeo.
C = Capacidade Total da operadora.

4. RESULTADOS

Essa seção tem como objetivo apresentar experimentos que validem a solução obtida a partir do algoritmo híbrido baseado em metaheurísticas para transmissão IPTV. No primeiro experimento foi utilizado criado um servidor multimídia para a realização dos testes. Para a realização do experimento 2 por ser um evento de simulação foi utilizado o simulador de rede, Network Simulator 3 (NS3). Ambos os experimentos foram rodando em uma máquina com sistema operacional Windows XP 32 bits.

4.1 Experimento 1

O primeiro experimento leva em consideração o comportamento de algumas variáveis do modelo em relação à transmissão IPTV considerando-se os formatos: High Definition (HD) imagem de 1080 linhas na razão de aspecto 16:9 (widescreen), superior ao padrão tradicional de Standard Definition (SD) de 525 linhas e ao Low Definition (LD).

Fig. 4 – Transmissão nos formatos de vídeo (HD,SD e LD).

Na Figura 4 é possível observar a diferença entre a transmissão dos três formatos de vídeo em relação à utilização da rede. Observa-se que o formato LD, considerado o de mais baixa qualidade, apresentou uma vantagem sobre os demais, consumindo menos os recursos da rede.

Através de um servidor multimídia foi constatado que o formato de vídeo HD consome uma largura de banda muito maior que os outros devendo assim ser tratado pelo modelo computacional com mais cuidado.

A transmissão no formato Low-Definition (LD), ocupou apenas 1% da rede, já Standard Definition (SD) ocupou 5% e High Definition (HD) por sua vez ocupou 20% da rede e assim se constatou que as variáveis do modelo onde o formato de vídeo é em HD seriam as mais impactantes na solução proposta.

4.1 Experimento 2

O *Network Simulator* (NS) é um simulador de eventos discretos orientados para a pesquisa em rede. Ele fornece um apoio substancial para a simulação de TCP, roteamento e protocolos *multicast* e *unicast* sobre uma rede (local ou por satélite) e por isso foi utilizado neste trabalho [14][15]. Na Figura 5 é mostrada uma parte do NS3 em execução, tamanho do pacote e IP de destino dentro de uma rede apenas.

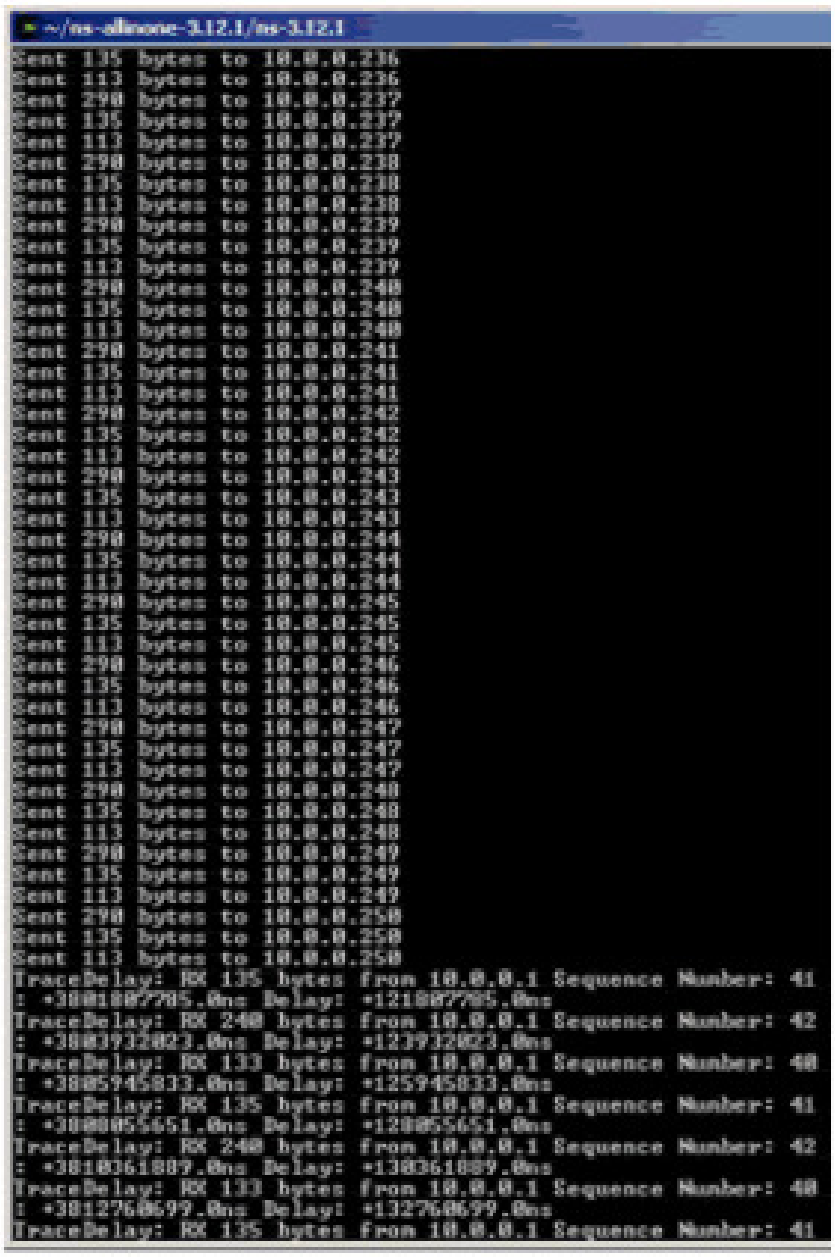

Fig. 5 – Parte do log do NS3

A partir da implementação no simulador de rede NS-3 foram criadas topologias realísticas, utilizam o protocolo RTP, para a simulação e avaliação dos resultados. A topologia de rede utilizada no NS apresenta um servidor multimídia dentro do core IP onde os arquivos de vídeo, áudio e dados são armazenados. Esses arquivos são encaminhados para o roteador da rede a que ele pertence, conforme sugestão e equipamento do cliente. Os arquivos podem ser enviados para inúmeros clientes pertencentes a diferentes redes de computadores. Porém para que houve uma qualidade satisfatória o experimento seguiu o número de conexões e formatos de vídeo da Tabela 1.

Como o NS é um simulador de rede baseado na transmissão de pacotes de bytes variados, nessa figura pode observar a perda de alguns. Porém o NS3 retransmite esses arquivos de bytes, até certo limite, para que a transmissão seja 100% realizada com qualidade. A perda destes pacotes pode impactar diretamente na qualidade do arquivo recebido pelo cliente [15].

Assim, partindo da solução encontrada pelo modelo construído, execuções foram realizadas no simulador com intuito de validá-lo e ainda obter informações que podem ser de extrema importância para a comunidade científica. A partir da realização dos primeiros testes observou-se que o tempo de transmissão na rede era diferente para os três formatos de vídeos, a Tabela 1 mostra a relação.

Tabela 1. Tempo de transmissão

Formato de Vídeo	Conexões	Tempo
HD	100	31,77s
SD	250	39,02s
LD	500	43,37s

Porém a variação de tempo destas transmissões teve um impacto menor que a perda de pacotes para essa configuração obtida pelo algoritmo hibrido.

4. CONCLUSÃO

Esse trabalho apresentou um modelo computacional que utiliza um algoritmo hibrido baseado nas metaheuristicas BT e AG com objetivo de se encontrar uma melhor configuração para a transmissão IPTV. Parâmetros, normas, arquiteturas e manuais de referência do assunto foram estudados para que fosse possível realizar uma simulação que se aproximasse bem da realidade.

Observou-se, através da implementação e dos experimentos que foram realizados, que, não só foi possível encontrar uma solução considerada boa para atender aos requisitos especificados na transmissão IPTV, como também ela se mostrou bem eficaz.

Como contribuições desse trabalho pode-se destacar a definição de um modelo computacional que utiliza metaheurísticas para encontrar uma solução supostamente boa. Além disso, a sua contribuição pode ser complementada, através dos estudos realizados com os diferentes formados de vídeo em relação aos fatores perda de pacotes, ocupação da rede e tempo de transmissão.

A continuação deste trabalho pode seguir as seguintes direções: aplicar o modelo desenvolvido em uma arquitetura onde a distribuição para as redes de telecomunicações remotas seja feita, através de uma tecnologia de transmissão via satélite. Ou analisar de maneira mais específica o valor que cada componente dentro das variáveis que compõem o modelo.

REFERÊNCIAS

[1] Caja, J. Optimization of IPTV multicast traffic transport over next generation metro networks. Telecommunications Network Strategy and Planning Symposium, 2006. NETWORKS 2006. 12th International, pages 1-6.

[2] Cisco. Cisco solutions for IPTV. Disponivel em: < www.cisco.com/go/iptv/ >. Acesso em: 17 agosto, 2011.

[3] CPqD. Arquitetura de referência. PD.30.12.34A.0001A/RT13/AA. CPqD, Campinas, p. 46, 2006.

[4] Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Computers and Operations Research, Oxford, UK, v. 13, n. 5, p. 533-549.

[5] Gómez, A. T.; Gluz, J. C.; Barbosa J. L. V.; Chiwiacowsky, L.; Ávila R. B.; Crespo, S. C. S. P.; Moreno, E. I. e Martins, M. G. Especificação técnica e sistêmica - plataforma de convergência digital iptv/tv digital. São Leopoldo, 2010. ISSN 2332/09.

[6] HWAN, S.S., e SHOGUN, A.W. Modelling and solving an FMS part selection problem. International Journal of Production Research, 27, 1349-1366, 1989.

[7] Holland, J. H., 1975. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. 1. ed. Michigan: University of Michigan Press.

[8] ITU-T. H.264 - SERIES H: Audiovisual and multimedia systems. 2007.

[9] Kukhmay, Y., Glasman, K., Peregudov, A., e Logunov, A. Video over IP networks: Subjective assessment of packet loss. Consumer Electronics, 2006. ISCE '06. IEEE Tenth International Symposium on, pages 1–6, 2006.

[10] Lee, C. S. IPTV over Next Generation Networks in ITU-T. Broadband Convergence Networks, p. 1–18, 2007

[11] Link, I. C. D. C., 2011. Um Sistema de Codificação de Vídeo para TV Digital - SBTVD. Programa Interdisciplinar de Pós-Graduação em Computação Aplicada – PIPCA. Mestrado Acadêmico. Proposta de Dissertação São Leopoldo.

[12] Marie, J.; Klym, N.; Mirlacher, T. The future of IPTV. Springer Science, 2010.

[13] Neto, Arlindo F. Uma Avaliação dos Mecanismos de Transmissão de Tráfego para IPTV. Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia. Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2009.

[14] NS-3 Project. Background. Disponível em: <http://www.nsnam.org/docs/release/3.10/tutorial/html/tracing.html>. Acesso em: 16 out. 2011.

[15] NS-3 Project. NS-3 Documentation. Disponível em:<http://www.nsnam.org/docs/release/3.12/doxygen/index.html>. Acesso em: 10 out. 2011.

[16] NS-3 Project. NS-3 Model Library. Disponível em: <http://www.nsnam.org/docs/release/3.11/models/html/index.html>. Acesso em: 26 out. 2011.

[17] O'driscoll, G. Next generation IPTV Services and Technologies. New Jersey: Wiley-Interscience, 2008.

[18] Padmanabhan V.N e Subramanian L. An investigation of geographic mapping techniques for Internet hosts. In SIGCOMM '01: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, pp. 173–185, New York, NY, USA, 2001.

[19] Padmanabhan V.N., Wang H.J., Chou P.A., E Sripanidkulchal K. Distributing streaming media content using cooperative networking. In NOSSDAV '02: Proceedings of the 12th international workshop on Network and operating systems support for digital audio and video, pp. 177–186, New York, NY, USA, 2002.

[20] Shihab, E. and Cai, L. IPTV Distribution Technologies in Broadband Home Networks. Electrical and Computer Engineering, CCECE Canadian Conference on , vol., no., pp.765-768, 22-26, 2007.

[21] Weissheimer Júnior, C. A. ; Gómez, A. T . Um modelo computacional híbrido para análise da transmissão sobre o Internet Protocol television (IPTV). In: X Simpósio de Informática da Região Centro/RS, 2011, Santa Maria. X Simpósio de Informática da Região Centro/RS 2011, 2011. v. 1. p. 38-43.

[22] Weissheimer Júnior, C. A. ; Gómez, A. T; Wickert, T I . Aplicação de um algoritmo genético para o modelo Internet Protocolo Television (IPTV). In: 9 ERRC Escola Regional de Redes de Computadores, 2011, São Leopoldo. 9 ERRC Escola Regional de Redes de Computadores. PORTÃO : Sociedade Brasileira de Computação (SBC), 2011. v. 1. p. 95-99.

Panorama da produção de conteúdo interativo no mercado brasileiro de televisão aberta em 2011

Matheus Colen (matheus.cln@gmail.com)

PANORAMA DA PRODUÇÃO DE CONTEÚDO INTERATIVO NO MERCADO BRASILEIRO DE TELEVISÃO ABERTA EM 2011.

Matheus Colen

Fundação Armando Alvares Penteado, São Paulo, Brasil, matheus.cln@gmail.com

Resumo: A interatividade é um dos novos recursos da TV Digital que podem mudar a maneira como as pessoas consomem TV. Esse recurso pode influenciar o desenvolvimento e a comercialização de conteúdos na TV Digital. Este trabalho traçou um panorama da produção de conteúdo interativo no mercado de televisão aberta, tendo como principal fonte de informação os profissionais das emissoras. Por meio de entrevistas, foi possível constatar que a interatividade dá os primeiros passos no mercado e não apresenta um modelo de negócios definido. Casos de sucesso também não existem. O futuro da interatividade ainda é incerto, contudo algumas emissoras investem para oferecer o novo recurso, mesmo sem terem uma previsão de retorno.

Palavras chave: TV digital, interatividade, televisão, panorama.

1. INTRODUÇÃO

Até Dezembro de 2007 a televisão terrestre brasileira era um meio estritamente audiovisual. Apesar de algumas etapas da produção de conteúdo terem sido digitalizadas bem antes dessa data, com a implantação de novas câmeras e ilhas de edição não lineares, as transmissões de TV ainda eram feitas pelo meio analógico. Com a implantação do sistema de transmissão digital, a difusão de conteúdo passa a ser feita através de bits. Essa característica trouxe ao broadcast a possibilidade de emitir qualquer informação digital, seja ela uma foto, um vídeo ou até mesmo um software (CRUZ, 2008). Por conta disso, uma série de serviços podem ser agregados ao conteúdo audiovisual. Recursos como a alta definição, multiprogramação, mobilidade e vídeo sob demanda eram, até então, intangíveis para essa mídia. É importante lembrar que várias emissoras desenvolveram programas interativos muito antes da TV digital existir. Seja por carta, telefone ou SMS o relacionamento do público com a emissora faz parte da história da televisão. Contudo, a digitalização do sistema de transmissão aumentou consideravelmente o leque de possibilidades para o aprimoramento desse diálogo. Agora existem outras maneiras para o telespectador controlar o fluxo da informação, sem precisar recorrer ao tão temido *zapping*.

O sistema de televisão digital escolhido pelo Brasil é de origem japonesa e se chama *Integrated Services Digital Broadcasting Terrestrial* (ISDB–T). Esse sistema possui como principal característica o suporte à multiprogramação e mobilidade. Sobre a multiprogramação, o decreto da TV digital brasileira autorizou somente a TV Cultura a usar o recurso em caráter experimental. Apesar do padrão ser japonês, o *middleware* Ginga foi desenvolvido no Brasil para operar nesse sistema. É o Ginga que torna possível a interatividade na televisão digital aberta. Para que o usuário possa utilizar esse recurso de maneira plena é imprescindível um canal de retorno. De olho na convergência digital, alguns fabricantes de eletrônicos passaram a implantar placas de rede para acesso a internet em seus novos modelos de televisores, além de desenvolverem *middlewares* próprios. Isso deu origem à um mercado que cresce rapidamente (TELETIME, 2011), apesar da baixa penetração de banda larga no país. As TVs conectadas podem ser uma opção de canal de retorno para a interatividade, além de levar conteúdo audiovisual, gratuito ou não, para a tela da TV quase que instantaneamente, a partir de uma solicitação do usuário. É o *vídeo on demand* (VOD) chegando até a televisão através da internet de banda larga sem necessariamente depender dos conteúdos das emissoras. *NetMovies*, *NetFlix* e *SundayTV* são alguns exemplos de serviços de VOD que já estão operando no Brasil.

Apesar do lançamento da TV Digital ter sido feito oficialmente há cerca de quatro anos, na prática, poucos recursos dessa nova mídia se difundiram para o grande público. O caminho em direção à plena utilização da TV Digital ainda está sendo desbravado, tanto pelos produtores de conteúdo, como pelas emissoras e até mesmo pelos consumidores. Contudo, essa jornada já tem data para acabar. Até 2016, pelo menos, viveremos uma fase de transição que exige de diversos atores dessa cadeia de valor uma intensa atualização de *hardware*. Nesse cenário, implantar o recurso de alta definição é relativamente simples para as emissoras, já que ele não exige do público a compra de novos televisores. Quem não possui uma TV HD, verá o programa em definição normal sem problema algum. Para fornecer o serviço, basta a emissora ter dinheiro em caixa para atualizar o parque técnico, investir mais nos recursos para as produções e treinar os colaboradores. Além disso, a alta definição não exige mudança nos modelos de negócios praticados na TV aberta. Agora, quando falamos em implantar interatividade para a TV Digital usando o Ginga, as coisas se complicam um pouco mais. Esse recurso requer uma equipe de desenvolvimento de software trabalhando em conjunto com as produções dos programas e com a programação do canal. A interatividade também traz novas possibilidades para os modelos de negócios, mas para que esses modelos sejam eficientes eles precisam

necessariamente que o público possua televisores compatíveis. Um bom canal de retorno é extremamente desejável. Portanto caímos numa espécie de beco sem saída, na medida em que não há dinheiro da publicidade para financiar a produção de interatividade, uma vez que não existe demanda expressiva do público para esse serviço. A relativa ausência de conteúdos interativos dificulta a formação da demanda. Além disso, se a população não tiver o televisor com o Ginga instalado e não aprender a usar a interatividade, essa demanda dificilmente existirá.

TV Digital é muito mais do que alta definição, mas como o mercado de interatividade poderá se desenvolver nesse cenário?

2. OBJETIVO

Esta pesquisa procurou realizar um panorama da produção de conteúdo interativo na televisão aberta, com a intenção de montar um retrato da evolução deste recurso no mercado. Com um mapa em mãos, pode-se ter uma ideia de qual é o melhor caminho a ser trilhado durante essa fase de pouquíssimas certezas.

3. MÉTODOS

Para reunir em um único documento os conceitos, estratégias e experiências do mercado de interatividade na televisão digital aberta, um representante de cada emissora foi abordado. Todas as entrevistas possuíram a mesma estrutura de perguntas, sofrendo pequenas modificações para atender às especificidades de cada empresa. As questões abordaram cinco temas principais. O primeiro deles diz respeito ao conceito de interatividade. Os três temas seguintes falam sobre como a emissora vê o público consumidor de televisão digital interativa, as possibilidades da tecnologia e, finalmente, o funcionamento do mercado de produção de conteúdo interativo. O último tema abordado ampliou o horizonte e procurou saber a visão do futuro.

A elaboração do questionário obedeceu esses cinco temas e também procurou testar algumas hipóteses. A primeira hipótese a ser testada é a de que o mercado ainda não possui um modelo de negócio definido para trabalhar com produção de conteúdo interativo. A segunda hipótese tenta justificar o fato de não existir conteúdo interativo em larga escala na grade das emissoras. Isso é feito levantando-se quatro pontos. Os altos custos que envolvem a produção de conteúdo interativo, a ausência de um modelo de negócio eficiente, a baixa penetração de aparelhos receptores compatíveis no mercado e os baixos níveis de instrução da população. A terceira e última hipótese diz respeito à tecnologia e procura saber se os canais de retorno que irão prevalecer para a interatividade são os televisores conectados à internet e o SMS.

Para comportar todas essas opiniões no presente artigo, algumas informações foram descritas no corpo do texto e outras agrupadas em tabelas de acordo com o assunto. Com isso teremos uma visão macro do mercado. A única emissora paulistana que optou por não responder a pesquisa foi a Gazeta. A justificativa foi o fato de que a empresa ainda está migrando o parque técnico para equipamentos digitais e não debateu internamente o assunto da interatividade. Além dos representantes das emissoras, a pesquisa procurou um ponto de vista externo ao mercado de broadcast, entrevistando o diretor executivo de uma empresa de desenvolvimento de software para interatividade. Abaixo, a lista dos profissionais entrevistados na ordem de realização.

- Roberto Franco – Diretor de Rede e Assuntos Regulatórios do SBT e atual presidente do Fórum do SBTVD.
- Fábio Angeli – Coordenador do Departamento de Desenvolvimento de Negócios na Record.
- Eduardo Brandini – Gerente de Multimídia na TV Cultura.
- Douglas Barbosa – Diretor de Programação da Gazeta.
- Zico Goes – Diretor de Programação da MTV.
- Kalled Adib – Superintendente de Operações da RedeTV!
- Frederico Nogueira – Vice-Presidente da BAND e ex-presidente do Fórum SBTVD.
- Carlos Fini – Gerente de Suporte e Tecnologia da Globo.
- Salustiano Fagundes – Diretor Executivo da HXD iTV

4. RESULTADOS

Para que a comunicação se dê de maneira eficiente, é fundamental conhecer as capacidades e limitações do receptor da mensagem. Quando se fala em conteúdo interativo para televisão, mídia presente na esmagadora maioria dos lares brasileiros, não se pode deixar de lado o fato de que boa parcela dos telespectadores não possui nível de alfabetização suficiente para interpretar e relacionar diversos significados numa mensagem. Segundo o estudo de 2004 que integra o Projeto do Sistema Brasileiro de Televisão Digital realizado pelo Centro de Pesquisa e

Desenvolvimento em Telecomunicações (CPqD), 8% da população é totalmente analfabeta. Dos alfabetizados, 67% não é capaz de localizar informações em textos longos e estabelecer relações de significado. Com um cenário desses, como exigir do grande público a capacidade de navegar por menus de interatividade na televisão digital? Isso é um entrave para a produção de conteúdo interativo? Segundo os produtores de conteúdo, a resposta para essa pergunta é não. A esmagadora maioria dos entrevistados diz que os baixos níveis de instrução da população não são empecilho significativo para a implantação da interatividade. Contudo, a BAND e a Record dizem levar essa característica da audiência em consideração e pensam em produzir aplicativos que não exijam muito da capacidade dos usuários. Isso é feito desenvolvendo menus bastante intuitivos e que utilizam as cores dos botões do controle remoto como base para interação. Apesar de não considerar a questão do alfabetismo como um problema, o SBT procura obedecer uma inteligência de programação que forneça informações de interesse geral em apenas alguns cliques. Conforme o usuário tiver a necessidade de procurar mais informações, o que demonstra uma maior capacidade para entender o conteúdo interativo, ele deverá utilizar mais cliques e não sofrerá para achar o conteúdo que deseja.

Fernando Crocomo definiu três níveis para a interatividade por meio do sistema digital. A interação de nível 1 ocorre quando a emissora envia um pacote de dados para o receptor, que os armazena e permite ao usuário a navegação dentre os diversos conteúdos interativos sem solicitar novos dados específicos para a emissora. O nível 2 traz o recurso de envio de dados para a emissora, o que possibilita também a participação em votações e enquetes. Na interatividade de nível 3 o envio de dados é feito em tempo real, permitindo ao usuário definir o final de um programa, por exemplo.

Para os indivíduos que não possuem dificuldades em interpretar o conteúdo, quais são os níveis de interatividade que serão oferecidos pelas emissoras? A Tabela 1 traz a relação baseada nos níveis de interatividade definidos por CROCOMO (2007).

Tabela 1. Níveis de interatividade por emissora

TV Cultura	Não usa o Ginga
SBT	1, 2 e 3
Globo	1, 2 e 3
Record	1 atualmente, 2 e 3 no futuro
BAND	1, 2 e 3 de acordo com o aplicativo
MTV	Não usa o Ginga
RedeTV!	1, 2 e 3
HXD iTV	1 e 2

TV Cultura e MTV não acreditam no Ginga como ferramenta eficiente para oferecer a interatividade. O foco da TV Cultura está restrito ao desenvolvimento de aplicativos para aparelhos móveis. Já a MTV trabalha o conceito de interatividade de modo mais amplo. A emissora já teve publicações impressas, está presente na internet com sites e conteúdos nas mídias sociais, além de criar constantemente conteúdos interativos para televisão utilizando meios anteriores ao Ginga.

Uma questão que permeia as discussões sobre interatividade na televisão aberta está relacionada ao hábito de consumo. Muitas vezes, assistir a televisão pode ser considerado uma atividade coletiva. Em contrapartida, a interatividade está mais relacionada ao uso individual da mídia. Para evitar conflitos, o conceito de duas telas se apoia nos aparelhos móveis, como *tablets* e *smartphones*, e oferece os menus da interatividade numa tela alternativa sem interferir na imagem do televisor e no consumo coletivo. Sob esse aspecto, somente Globo, Record e RedeTV! pensam em oferecer o recurso num futuro próximo. O SBT vê essa possibilidade num futuro distante, pois acredita que a tecnologia evoluiu rapidamente e está muito a frente dos hábitos de consumo do público. Apesar de não utilizar a interatividade por meio do Ginga, alguns aplicativos para aparelhos móveis desenvolvidos pela TV Cultura terão a possibilidade de interagir em tempo real com o conteúdo transmitido pela emissora.

Por ser um recurso novo, o *middleware* Ginga é alvo de uma dúvida que persegue a maioria dos produtos que entram num determinado mercado: Vai pegar? Ele está ameaçado de alguma maneira pela nova geração de TVs conectadas e seus *middlewares* proprietários? O *middleware* é eficiente? A Tabela 2 compila algumas opiniões dos entrevistados a respeito da capacidade do *middleware* padrão do SBTVD em se tornar uma ferramenta amplamente utilizada por desenvolvedores e consumidores de conteúdo interativo.

Tabela 2. Sobre o Ginga

TV Cultura	Baixa penetração do Ginga. Público não conhece o recurso.
SBT	Ginga é uma plataforma robusta
Globo	Ginga é forte com interatividade associada ao conteúdo transmitido no momento.

Record	Ginga é forte. Emissora oferece conteúdo baseado na penetração dos perfis de Ginga no mercado.
BAND	Ginga é forte. Proposta das TVs conectadas não afeta o Ginga. São modelos diferentes
MTV	Não usa o Ginga. *Middleware* não trouxe novidade quando comparado a internet
RedeTV!	Ginga é forte e deveria ser usado como padrão para controlar conteúdo.
HXD iTV	Ginga é forte e não é alavancado pela falta de interesse do governo e das emissoras

Como é possível constatar, alguns entrevistados afirmam que acreditam no Ginga como uma ferramenta robusta para permitir a interatividade e até mesmo o controle sobre o conteúdo audiovisual nas TVs conectadas. O contraponto é feito pela TV Cultura e MTV, que se baseiam na baixa penetração do recurso e na relativa falta de novidade que essa ferramenta trouxe para justificar ausência do Ginga em seus planos de relacionamento com a audiência. A HXD iTV foi a única que criticou o mercado de *broadcast* e o governo, apontando para a falta de interesse em difundir a interatividade nesse momento. Ao fazer isso, a empresa chama a atenção para o desempenho do *middleware* brasileiro em mercados internacionais e afirma que o futuro dessa ferramenta depende de ações realizadas no exterior.

Para que a interatividade funcione plenamente é necessário um canal de retorno. Os entrevistados também falaram sobre como se dará esse retorno de informação da casa do telespectador até a emissora. Algum meio físico vai prevalecer? O desenvolvimento dos aplicativos é pautado de alguma maneira pensando no meio de retorno? A Tabela 3 traz algumas respostas.

Tabela 3. Retorno da informação

TV Cultura	Pouca estrutura de retorno. Não usa o Ginga.
SBT	Suporte para todos os meios de retorno
Globo	Suporte para todos os meios de retorno
Record	Aposta na internet como retorno, apesar de usar muito o SMS.
BAND	Suporte para todos os meios de retorno.
MTV	Não usa o Ginga
RedeTV!	Suporte para todos os meios de retorno
HXD iTV	Pela internet por diversos meios físicos

TV Cultura e MTV continuam sustentando opinião de que o Ginga não oferece potencial para mudar o relacionamento com a audiência, dessa vez pela falta de infraestrutura para o canal de retorno. Se isso é um ponto fraco hoje, a BAND acredita que no futuro a evolução das tecnologias usadas para retorno irá diminuir o custo do serviço, ampliando a base de consumidores aptos a trocar informações com a emissora. A grande maioria dos entrevistados irá desenvolver os aplicativos interativos com suporte para todos os meios de retorno disponíveis e não acreditam que um meio específico vá se sobressair sobre os demais. Record e HXD iTV apostam na internet como retorno.

Nesse aspecto as TVs conectadas são uma boa pedida para quem deseja interagir. Contudo, os *middlewares* de alguns fabricantes permitem o acesso irrestrito a internet por meio de *widgets*. Isso dá aos consumidores a possibilidade de adquirir serviços de vídeo sob demanda não controlados pelas emissoras. Essa característica pode danificar o modelo de negócios baseado em publicidade praticado pela TV aberta? De uma maneira geral, a resposta é não. SBT, Globo, Record, BAND, MTV e RedeTV! acreditam que as TVs conectadas são uma nova possibilidade de distribuição do conteúdo e não uma ameaça. Entretanto existem ressalvas. O SBT declarou não fazer parcerias com fabricantes que utilizam *widgets* em seus *middlewares*. A RedeTV! acredita que todas as TVs conectadas deveriam ter o Ginga como *middleware* padrão para dar aos radiodifusores o controle sobre o conteúdo audiovisual consumido no aparelho. Nesse aspecto a HXD iTV encara a realidade de que uma fauna de *middlewares* habita esse mercado e é preciso estar preparado para trabalhar com todos eles. Recentemente a empresa anunciou uma mudança de estratégia para poder atender a demanda de aplicativos em todos os modelos de televisores conectados. Antes o foco de atuação da HXD iTV era restrito ao Ginga.

As emissoras justificam os baixos investimentos em interatividade pelo Ginga com o argumento de que esse recurso ainda não se paga. O desenvolvimento dos aplicativos não é barato e ainda não existem modelos de negócios de eficácia comprovada. Apesar de possuir um portal de interatividade, o SBT alega falta de recursos para desenvolver interatividade para os programas de sua grade. Ainda assim a emissora acredita que o Ginga traz a possibilidade de diversificação no modelo de negócios a partir do momento que permite a prática do marketing de relacionamento para os produtos anunciados. Além disso, será possível utilizar um modelo de remuneração baseado em cliques nos anúncios dos aplicativos, assim como ocorre com a publicidade na internet. A BAND vê essa característica como um forte atrativo para os anunciantes na medida em que permite a mensuração de resultados. Já a Record observa um horizonte de receitas na interatividade com o patrocínio de funcionalidades específicas dos aplicativos. Globo e RedeTV! não acreditam que o recurso será fonte de receitas e encaram a

interatividade apenas como um valor agregado ao conteúdo audiovisual.

O desenvolvimento dos aplicativos interativos está mais relacionado com a área de TI do que com a de comunicação. Por conta disso, a grande maioria das emissoras optou por terceirizar o desenvolvimento das aplicações, controlando apenas a criação e o conteúdo. Somente a Record e a RedeTV! implantaram departamentos internos para desenvolvimento. Nesse aspecto, a Record alega que o novo departamento será capaz de criar aplicativos para os programas da casa, além de dar suporte para anunciantes que desejam usar recursos interativos. A emissora chama atenção para a falta de interesse por parte das produções dos programas sobre os recursos que o Ginga oferece para aprimorar o relacionamento com o público. Os profissionais de criação das produções precisam estar capacitados a pensar conteúdos para a interatividade relacionados diretamente ao programa, o que resultaria em aplicativos inovadores.

A digitalização do sistema de transmissão de televisão trouxe muitas novidades. Ainda vivemos a fase de transição, que tem uma data prevista para acabar. De acordo com decreto da TV Digital, o prazo de transição termina em 2016. Será que esse prazo poderá ser cumprido? SBT, Record e MTV disseram que acreditam na revisão desse prazo. Record e SBT acham que talvez não seja necessário estender o prazo nas principais capitais e chamam a atenção para as dificuldades que existem em atualizar 100% das retransmissoras. Já a BAND e a TV Cultura acreditam que o prazo será cumprido. Globo e RedeTV! não arriscaram uma previsão pois acreditam que há muitas variáveis que podem interferir nesse processo. Até 2016 teremos dois eventos esportivos de grande porte no país, que tradicionalmente são picos de vendas de televisores. Acredita-se que a Copa do Mundo de 2014 e as Olimpíadas de 2016 irão ajudar a atualizar a base de televisores compatíveis com a TV digital e o Ginga. Até lá, como será que estará o consumo de TV digital como um todo? Nessa questão não houve consenso. Cada recurso da televisão digital possui seus próprios defensores no mercado brasileiro de *broadcast*. A Tabela 4 traz as apostas das emissoras para o futuro da TV digital no Brasil.

Tabela 4. Consumo de TV digital em 2016

TV Cultura	HD + VOD
SBT	Interatividade
Globo	HD
Record	Multiprogramação
BAND	Interatividade
MTV	VOD + portabilidade
RedeTV!	HD

5. DISCUSSÃO

Estamos numa fase de transição. A digitalização da TV veio acompanhada de uma série de novas possibilidades, o que também traz muitas mudanças. Para viabilizar produtos interativos no mercado de *broadcast* será preciso muito investimento e criatividade.

Essa nova possibilidade da TV digital exige uma equipe adicional, ligada à área de desenvolvimento de software, trabalhando em conjunto com a produção dos programas. O tema está em fase tão embrionária que cada emissora trata do assunto em um departamento diferente. Já vimos que são vários os fatores que dificultam a produção e o consumo da interatividade. Como uma emissora conseguirá vender publicidade em aplicativos interativos, se não existem números representativos nesse mercado? Se não existe dinheiro para produção, não existe conteúdo e isso impede a formação do hábito de consumo. Sem a demanda do público, não há dinheiro da publicidade para produção. Esse círculo vicioso só pode ser quebrado com investimento direto das emissoras e muita criatividade para bolar estratégias eficazes para a interatividade.

Apesar disso tudo, a interatividade não deve ser vista como um investimento sem retorno, mas sim como uma aplicação com retorno de longo prazo. Esse recurso pode ser uma ferramenta de inclusão digital para centenas de milhares de famílias por todo o Brasil. Muitas pessoas ainda não possuem o hábito de acessar a internet ou usar o computador. A digitalização da mídia de maior penetração no país pode levar desenvolvimento aos lares brasileiros. E esse é um ótimo valor a ser agregado em uma marca. Nos mercados em que a internet já faz parte da rotina das pessoas, como seduzir o telespectador a interagir com a televisão, deixando o computador de lado?

Sob esse aspecto, alguns entrevistados ponderaram que a internet agrega muito mais valor, com uma experiência muito mais rica, para o usuário que deseja interagir com o conteúdo. Isso é verdade quando pensamos em aplicativos que trazem mais informações sobre o programa, colocando menus típicos de internet na tela da televisão. Nesse caso o controle remoto perde de longe para a interação através do computador, *tablet* ou celular. É preciso criar conteúdo interativo relacionado ao negócio principal da televisão: o audiovisual. Programas interativos

para televisão devem oferecer uma experiência que só pode ser vivida plenamente na televisão. Porque não pensar em novas estratégias para *product placement* em programas de dramaturgia interativa, por exemplo? Produtos que podem mudar o rumo de uma história para melhor. Atualmente, não é um absurdo orçamentário pensar na produção de várias trajetórias para o mesmo drama. Talvez isso seja atrativo para os anunciantes, mesmo nesse momento de transição, na medida em que alia um método de divulgação já conhecido pelo mercado, à uma nova tecnologia. Pioneirismo é uma qualidade que diversas marcas procuram aliar a seus produtos.

De qualquer modo, durante a pesquisa foi possível observar um ponto em comum em diversas emissoras – e, como vimos, isso também foi uma questão levantada pela Record. Existem pouquíssimos profissionais de produção envolvidos diretamente na criação de conteúdo interativo para os programas. Grande parte dos aplicativos são criados e desenvolvidos pelas áreas técnicas ou de novas mídias. Para que a interatividade na TV faça sentido, é preciso que ela esteja no DNA dos programas. Os profissionais de produção que atuam na criação devem conhecer um pouco mais sobre a tecnologia e as inúmeras possibilidades que ela oferece. Pessoas capazes de criar conteúdos interativos genuinamente televisivos podem fazer a diferença ao lado de bons desenvolvedores de *software*.

6. CONCLUSÕES

Após realizar todas as entrevistas, foi possível constatar que a interatividade está ainda num estágio muito inicial de evolução. No presente momento, criar conteúdo interativo para televisão aberta é um verdadeiro desafio. Dentro de um planejamento orçamentário a interatividade representa custo e não receita. Apesar de algumas emissoras já fazerem as primeiras tentativas nesse setor, ainda não há cases de sucesso e nem modelos de negócios definidos.

A hipótese de que a baixa penetração dos aparelhos receptores dificulta a implantação do recurso foi comprovada. Para vários radiodifusores, os aspectos econômicos que envolvem a produção de interatividade são problemas muito maiores do que os baixos níveis de alfabetização da população. Desenvolver esse tipo de conteúdo não é algo barato. Foram poucas as emissoras que se arriscaram a montar um departamento interno para trabalhar o assunto. A maioria prefere centralizar a inteligência das aplicações e terceirizar o desenvolvimento. Há também aqueles que não acreditam no Ginga, como é o caso da Cultura e da MTV, e por isso preferem direcionar os investimentos para os aplicativos de aparelhos portáteis e distribuição de conteúdo *on demand*, de olho no mercado móvel e nas TVs conectadas. Sobre o conceito de duas telas, são poucos os que estão investindo para evitar possíveis brigas na família pela posse do controle remoto.

O futuro da TV digital ainda é incerto. Não há consenso se o prazo para o *switch-off* será postergado. Sobre o consumo dos recursos da TV digital em 2016 também não houve consenso. Há uma certa esperança de que os eventos esportivos de 2014 e 2016 contribuam para a renovação da base de televisores compatíveis. Isso ajudaria a criar demanda para interatividade e, quem sabe, mudar a maneira de se consumir informação e entretenimento pela TV.

AGRADECIMENTOS

Aos profissionais entrevistados que gentilmente cederam parte de seu tempo e conhecimento para a pesquisa.

Ao professor Valdecir Becker pelos contatos iniciais e pela a valiosa orientação para a elaboração do questionário.

A todos os professores e colegas do curso de pós-graduação em Produção Executiva e Gestão de TV pela intensa e valorosa troca de experiências.

REFERÊNCIAS

CROCOMO, F. (2007). *TV Digital e Produção Interativa*. Florianópolis, SC, Brasil: Editora da UFSC.

CRUZ, R. (2008). *TV Digital no Brasil*. São Paulo, SP, Brasil: Editora Senac São Paulo.

TELETIME. (2011 em 20/06). *TELETIME News - Samsung vende 2 milhões de TVs conectadas em três meses*. Acesso a TELETIME em 21/07/11: http://www.teletime.com.br/20/06/2011/samsung-vende-2-milhoes-de-tvs-conectadas-em-tres-meses/tt/228699/news.asp

Estudo da Rede de Transmissão 4K dos Jogos da Copa do Mundo de 2014

Luana Bravo (luana_bravo@uol.com.br)
Eunézio Antônio Souza (thoroh@gmail.com)

ESTUDO DA REDE DE TRANSMISSÃO 4K DOS JOGOS DA COPA DO MUNDO DE 2014

L.C. Bravo, C.M.B. Lopes[1], N. Faria Jr[2], L.S.R.M.J. Reis[2] e E.A.T. de Souza

Laboratório de Fotônica, Universidade Presbiteriana Mackenzie, Rua da consolação 896, São Paulo/SP, 01302-907, Brasil, luanabravo@sbt.com.br,

(1) Escola de Engenharia e Tecnologia, Universidade Anhembi Morumbi, Rua Casa do Ator, 275, São Paulo/SP, 04546-001, Brasil,

(2) i9 Rede Globo, Rua Jardim Botânico 674, sala 419 Rio de Janeiro/RJ, Brasil.

Resumo: *Apresentamos um estudo preliminar sobre a rede de transmissão composta pelas doze cidades sedes dos jogos da copa do mundo de 2014 com resolução 4K. Analisamos uma solução utilizando redes fotônicas dedicadas para a conexão das doze cidades sedes de forma que todas possam receber e transmitir os jogos em 4K. A topologia da rede foi dimensionada de maneira a suportar o tráfego de três sinais por cidade, considerando a possibilidade de três jogos simultâneos.*

Palavras chave: Transmissão de vídeo em 4K, rede óptica, rede de transmissão.

1. INTRODUÇÃO

Quando pensamos em imagens de vídeo, um dos aspectos mais importantes é a qualidade por ela apresentada. Ao transferir esse fator para o mundo televisivo, pode-se notar uma grande evolução. No passado, quando havia apenas a televisão analógica colorida, a resolução dessa imagem era de 640 colunas e 525 linhas horizontais, no padrão PAL-M que foi adotado no Brasil [1]. E em 1995 surgiu o DVD, mídia utilizada para o vídeo digital, apresentando uma resolução de vídeo 720x480 pixels, o que de certa forma surpreendia pela qualidade apresentada. Com o desenvolvimento tecnológico, as telas dos dispositivos aumentaram, e conseqüentemente o numero de pixels também. A fim de conseguir cada vez mais a sensação de imersão dos telespectadores nas cenas com a riqueza de detalhes, e atender as novas telas grandes disponíveis ao consumidor, foi necessário aumentar a resolução do vídeo e então desenvolver padrões de alta resolução: o HDTV-720 (*High Definition Television*) com resolução de 1280x720 pixels e o *Full* HDTV ou HDTV-1080 com resolução de 1920x1080 pixels, sendo este último adotado no padrão de Televisão Digital Aberta no Brasil [2].

Seguindo a mesma linha de raciocínio de imersão e riqueza de detalhes para tornar a experiência do público ainda mais real, além de acompanhar o desenvolvimento dos tamanhos displays dos cinemas, surgiram os padrões 2K (2048x1080 pixels) e 4K (4096x2160 pixels), que foram previstos e adotados nas especificações do cinema digital, o *Digital Cinema System Specification* [3].

Os sinais de resolução 4K, presentes em várias telas de cinema estão surgindo também no mundo televisivo. Recentemente, produtos com a tecnologia 4K foram destaques na maior feira mundial para tecnologia de consumo em Las Vegas, EUA, a *Consumer Eletronics Show* (CES) [4].

O objeto de estudo desse trabalho é a análise da rede de transmissão de sinais 4K aplicados à contribuição para eventos esportivos, vislumbrando tal necessidade para copa de 2014.

A Copa do Mundo de Futebol é um evento de importância mundial que acontecerá no Brasil no ano de 2014. Sua contribuição no desenvolvimento tecnológico é eminente dada à evidência e a audiência proporcionada por esse campeonato. Concatenando a grandiosidade da Copa do mundo e a tecnologia 4K, será analisada a topologia de uma rede de transmissão que envolva as doze cidades sedes. Essa rede tem o objetivo de transmitir e receber os sinais de áudio e vídeo de todos os jogos ao vivo utilizando redes fotônicas e tecnologia de super alta definição, 4K.

Nas seções 3 e 4 desse trabalho, veremos que a estrutura é muito parecida entre as transmissões de sinais HD e 4K, porém o grande diferencial é a banda utilizada na rede de contribuição, rede esta que fará a contribuição/distribuição dos sinais captados durante a copa do mundo. Além do aumento de banda, existem alguns pontos que devem ser considerados: a perda de pacotes, o controle de QoS (*Quality of Service*) e *delay/ jitter*. A perda de pacotes e o QoS estão diretamente relacionados. Deverá existir o controle de garantia da banda de transmissão, pois uma vez que é realizada a transmissão de pacotes que contenham vídeo, qualquer variação na banda de transmissão pode implicar na perda de pacotes. Essa perda poderá implicar na existência de macro blocos no ponto de recepção ou mesmo a perda de sincronismo. O *delay/jitter* pode ser causado também pela perda de pacotes, ou pelo próprio processamento intrínseco dos equipamentos de rede.

2. TRANSMISSÃO DE SINAIS 4K

Super High Definition (SHD), ou simplesmente 4K, é um padrão de resolução de vídeo que tem quatro vezes mais a resolução de um vídeo HDTV (1080x1920 pixels), apresenta 2160 linhas, cada uma com 4096 pixels [5]. O formato 4K, é capaz de fornecer imagens com qualidade muito próximas das imagens convencionais de filme 35-mm. Ao contrário da película, as imagens 4K, por serem digitais, podem ser copiadas, transmitidas e manipuladas por várias vezes sem perderem a qualidade [6].

O conceito 4K foi proposto pela primeira vez pela empresa NTT *Corporation* [5], que desde então tem feito vários testes de transmissão com esse tipo de sinal, desde a utilização sem compressão [6] quanto sua utilização e aprimoramento com compressão [7].

No Brasil e no mundo, os vídeos 4K são encontrados principalmente em salas de cinema, porém com uma oferta ainda pequena. É possível vislumbrar o uso dessa tecnologia também para os serviços de televisão aberta, conferências e até mesmo educação à distância.

Testes de transmissão com a utilização de sinais de vídeo de super alta definição já vêm sendo realizado no Brasil e no mundo. Em janeiro de 2006 a NTT em parceria com a Keio DMC (*The Keio University Research Institute for Digital Media and Content* em Mita, Tokyo) fez testes de transmissões de sinais vídeo 4K não comprimidos utilizando uma rede IP (*Internet Protocol*) [7]. Para esse teste, foram utilizadas duas fontes de vídeo: uma câmera digital 4K para sinais ao vivo localizada em Keio DMC, e um servidor de vídeo sem compressão para material de cinema 4K pré gravados na NTT Musashino R&D Center. Cada vídeo, não comprimido, foi encapsulado em pacotes IP e transmitido por uma rede com largura de banda de 10 Gbit/s para uma sala de simpósio no *Sendai International Center* através de um *stream* de 6 Gbit/s podendo ser visualizado a partir de um projetor 4K, conforme demonstra a figura 1.

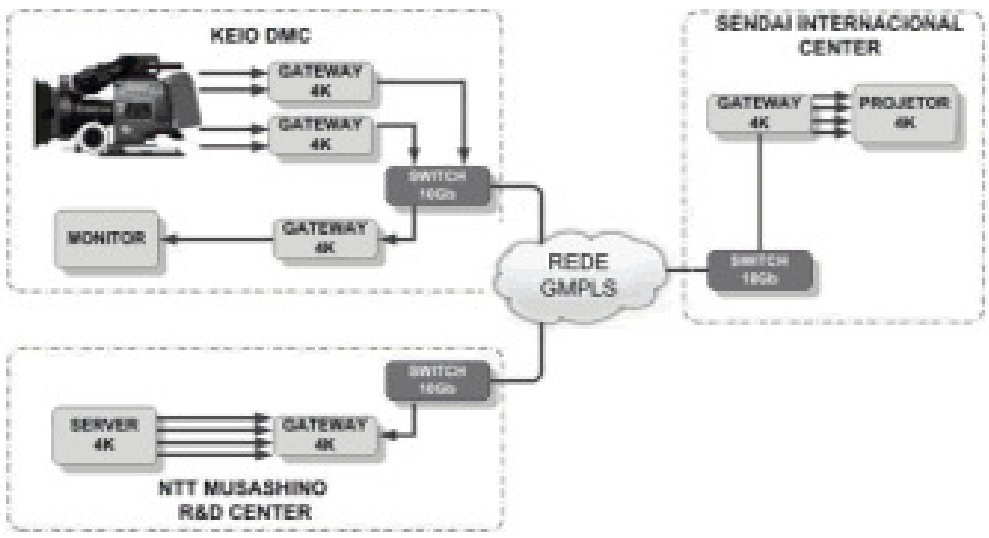

Fig. 1. Configuração do Sistema de teste de transmissão de vídeo 4K não comprimido. Figura adaptada [7].

Nesse teste, o encapsulamento IP de sinais 4K não comprimidos foi feito através do 4K Gateway "i-Visto" desenvolvido pela NTT. O i-Visto é um sistema com entradas de vídeo HD-SDI (*Serial Digital Interface*) que transforma esses sinais de vídeo em pacotes IP em tempo real e transmite o fluxo de vídeo IP sem compressão. O gateway 4K tem quatro placas de processamento de vídeo HD-SDI e duas interfaces de 10-Gbit Ethernet (10-GbE). Inicialmente o sistema divide um frame de vídeo 4K em quatro imagens HD-SDI sincronizadas, converte cada um desses quatro sinais em pacotes IP, com uma taxa de 1,5 Gbit/s em cada pacote.

Todos os i-Visto transmitem um fluxo com dois sinais IP de 3 Gbit/s total através de uma interface 10-GbE usando UDP/IP (UDP: *User Datagram Protocol*). As duas saídas de Ethernet são agregadas por um switch 10-GbE resultando em um fluxo com taxa total de 6 Gbit/s para a rede. O sistema é detalhando na figura 2 a seguir.

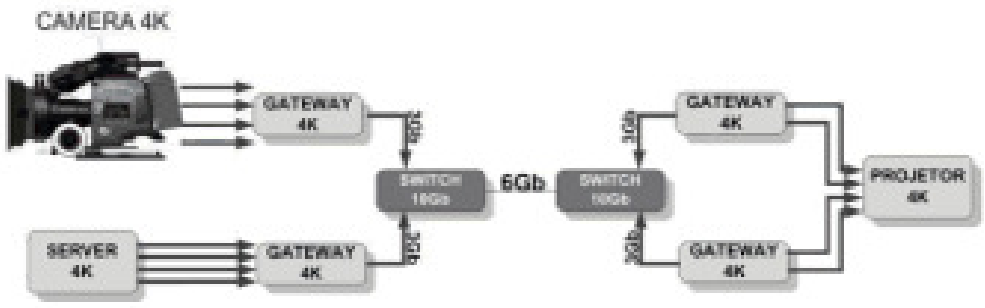

Fig. 2. Sistema de transmissão utilizando gateway 4K. Figura adaptada [7]

Uma rede de 10 Gbit/s pode transmitir apenas um fluxo de 6Gbit/s por vez (taxa conseguida de um vídeo não comprimido 4K) [7]. Para o estudo da NTT, foi necessário comutar entre os vídeos enviados que estavam em sites diferentes, utilizando um comutador óptico (OXCs) em duas localidades de Tokyo e uma em Osaka, apresentando uma distância total de rede de 1500 km. Porém, um sinal de vídeo 4K quando nativo, pode ter uma taxa de dados variando entre 6,3 Gbit/s (formato 4:2:2, 30 bits, 24 quadros/s) e 9,5 Gbit/s (formato 4:4:4, 36 bits, 30 quadros/s), o que torna inviável uma transmissão desse sinal em uma rede Gigabit IP comum.

Para otimizar a taxa a ser transmitida, é necessário que haja a compressão desses dados. O codec JPEG2000, uma das codificações possíveis de serem utilizadas, reduz essa taxa para 200-400 Mbit/s [8].

Na rede proposta pela NTT em 2007 [5], o sinal de vídeo é comprimido, entra na rede IP e é descomprimido por diversos pontos de recepção, conforme mostra a figura 3.

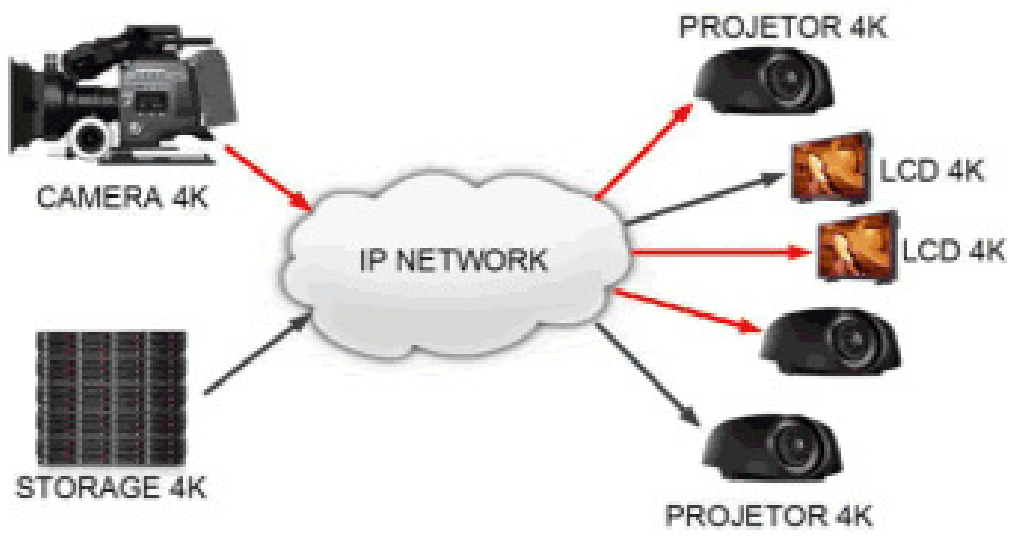

Fig. 3. Esboço de uma rede multiponto para sinais 4K. Figura adaptada [5]

Para tornar essa rede possível, o fluxo de 4K foi replicado em divisores FlexCast e endereçado para vários pontos de recepção que decodificam o JPEG2000 em tempo real. Esse divisor está localizado na rede IP e conectado aos roteadores IP convencionais que trabalham com até 10 GBit/s. Além da codificação, divisão e endereçamento dos pacotes, fez-se necessário o uso de um código de correção de

erro, o LDGM *(Low-Density Generator Matrix)* que recupera os pacotes perdidos. Fundamentalmente, o principio dessa correção de erro, baseia-se na inserção de bits de paridade, que são recuperados caso haja a perda de pacotes durante a transmissão dos dados.

Em 2005, foi realizado um experimento no 19° Festival Internacional de Filme de Tokyo, para promover essas tecnologias [5]. O experimento envolveu cinco locais: Roppongi (site do festival), Mita (DMC, *Keio University*), Akihabara (NTIC: Instituto Nacional de Tecnologia de Informação e Comunicação), Keihanna (NTIC), e Yokosuka (NTT). Eles estavam interligados por uma rede de acesso banda larga fornecida pela NTT Communications, JGN II (Japão Gigabit Network II), e GEMnet 2 (*global enhanced multifunctional network 2*). Nesse experimento, vídeos 4K de uma palestra, um concerto musical e um vídeo de jogo gerados a partir do codificador JPEG2000 em Mita, foram replicados pelo Flexcast e divididos em Akihabara. As imagens foram decodificadas e recebidas em dez pontos distintos sem perda de pacotes ou qualidade.

Em 2009 foi realizado no Brasil um teste de transmissão de sinais 4K entre três países: Brasil, Estados Unidos e Japão [9]. Nesse experimento foi feita a reprodução de um filme 4K com 70 minutos de duração a partir do Brasil para os outros dois países.

Analisando as experimentações de sucesso, é possível identificar vários pontos que precisam de atenção para a transmissão de sinais 4K. Dentre eles destacamos a rede em si utilizada para tal transmissão, as câmeras 4K com streaming, e os switches rápidos o bastante para lidar com este sinal. Neste trabalho focamos nossa atenção apenas nos aspectos da topologia da rede, uma vez que a complexidade dos outros componentes envolvidos merecem igual atenção e podem ser explorados em outros estudos.

3. ANÁLISE DA ARQUITETURA DE TRANSMISSÃO DAS 12 CIDADES SEDES COM RESOLUÇÃO *FULL HD*

Antes de analisarmos a transmissão de sinais de super alta definição, é necessário entender o processo de contribuição utilizado no momento. O processo que envolve a transmissão de um jogo de futebol é composto pela captação, pós-produção e processamento desses sinais. Na fase de captação, são instalados câmeras e microfones ao longo do estádio que enviarão os sinais para uma central técnica que processará o áudio e o vídeo, fazendo o corte de câmera, mixando o áudio, inserindo artes gráficas para o produto final ser transmitido.

As fases de captação e processamento não são o foco desse trabalho, portando a abordagem acontece a partir do sinal de áudio e vídeo finalizado para ser transmitido.

A análise a seguir é baseada no que há de mais moderno nas transmissões televisivas *ao vivo* de eventos esportivos no Brasil, que é o sinal de alta resolução *Full HD e codificação JPEG2000*.

O sinal finalizado, ao entrar na cadeia de transmissão será codificado e convertido a um protocolo IP. Para esse estudo consideramos o padrão de codificação JPEG2000. O Sinal HD, codificado nesse padrão, tem uma taxa de aproximadamente 150 Mbit/s. Para o encapsulamento IP, será acrescido a esse sinal de vídeo, um cabeçalho de aproximadamente 10%, o *transport stream* final terá aproximadamente 165 Mbit/s.

Para ser transmitido através de meios fotônicos, o sinal elétrico é convertido a um sinal óptico. Após a conversão o sinal é enviado através de uma rede de transmissão óptica, que irá endereçar o caminho percorrido pelo *Transport Stream*. No ponto de recepção o TS, fará exatamente o caminho oposto, sendo convertido para sinais elétricos, decodificado recuperando o sinal de áudio e vídeo original. A figura *4* mostra o diagrama em blocos da cadeia de transmissão ponto a ponto.

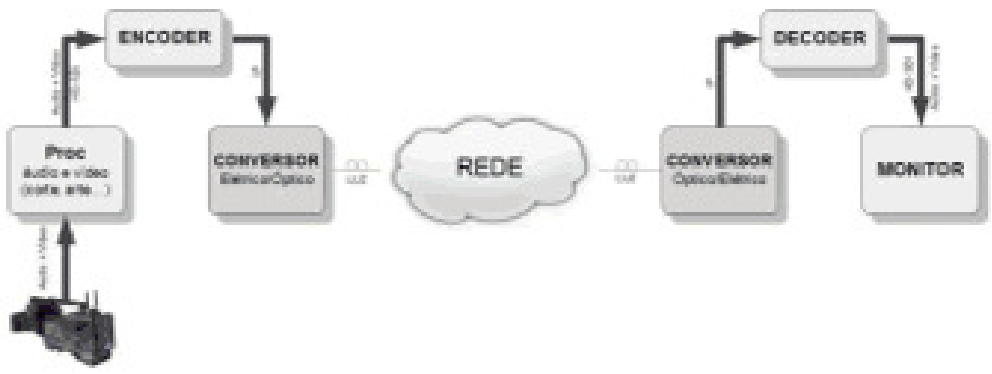

Fig. 4. Ilustração de uma rede de transmissão ponto a ponto

Para os jogos da copa, porém, a rede não será ponto a ponto, e sim haverá uma rede fibra óptica de conexão entre as doze cidades sedes, conforme a figura 5.

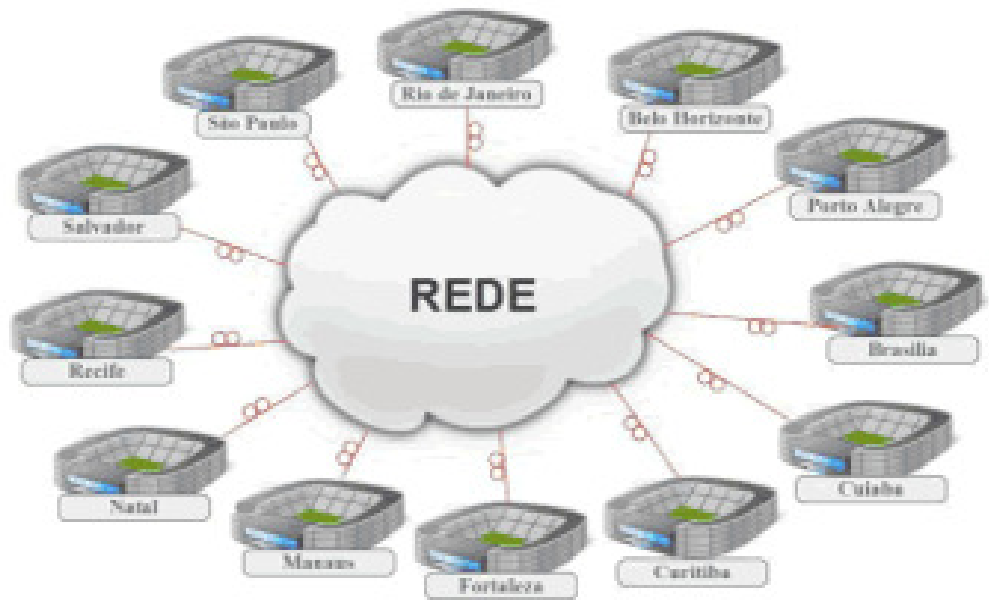

Fig. 5. Conexão de rede das doze cidades sede da Copa do Mundo de 2014.

Para desenhar a topologia da rede, é necessário considerar que na copa acontecem até três jogos simultâneos. Dado este fato, é necessário prever que cada cidade que estiver transmitindo um jogo, estará com um fluxo de três sinais: um sinal de transmissão e outros dois de recepção, não sendo considerada

redundância no primeiro momento. As cidades que não estiverem transmitindo nenhum jogo, entretanto contarão com até três sinais de recepção.

Cada sinal seja ele de recepção ou de transmissão necessita de uma banda de aproximadamente 165 Mbit/s, considerando que haverá o tráfego de três sinais em cada localidade, será necessário uma banda mínima total de aproximadamente 495 Mbit/s na rede das doze cidades. A proposta de arquitetura de rede que compreende as cidades sede foi desenhada considerando as distâncias geográficas, e as principais cidades que contenham mais estruturas para serem as unidades administrativas dessa rede.

A rede baseia-se em quatro anéis, onde um deles é composto apenas pelas cidades que terão o gerenciamento dessa rede, são elas: São Paulo, Rio de Janeiro, Belo Horizonte, Salvador e Brasília.

As distancias consideradas e aproximadas são:

São Paulo – Brasília 925 Km;
Brasília - Salvador: 1075 Km;
Salvador – Belo Horizonte: 950 Km;
Belo Horizonte - Rio de Janeiro: 360 Km;
Rio de Janeiro – São Paulo: 395 Km;
São Paulo – Curitiba: 335 Km;
Curitiba – Porto Alegre: 505 Km;
Porto Alegre – São Paulo: 750 Km;
Salvador – Recife: 670 Km;
Recife – Natal: 250 Km;
Natal – Fortaleza: 435 Km;
Fortaleza – Salvador: 1025 Km;
Brasília – Manaus: 1915 Km;
Manaus – Cuiabá: 1145 Km;
Cuiabá – Brasília: 890 Km.

As distâncias citadas foram retiradas do "Google Earth", considerando uma linha reta entre as cidades.

A figura 6 a seguir, representa essa estrutura de anéis proposto a essa arquitetura.

Fig. 6. Arquitetura de Rede em anéis ópticos proposta para a conexão entre as doze cidades sede da Copa do Mundo de 2014 (a rede está deslocada em relação ao Mapa)

Considerando o trafego dos três sinais simultâneos com um fluxo total de 495 Mbit/s, uma rede de 1 Gbit/s seria suficiente para trabalhar com os sinais HD com compressão JPEG2000, considerando uma margem de banda razoável por questões de segurança e reservas técnicas e sem considerar a redundância de sinais.

Se considerarmos que para cada sinal enviado haverá um sinal de redundância, o Fluxo total então será de 990 Gbit/s. Sendo assim, considerando também a segurança das transmissões é sugerida uma rede de 2 Gbit/s.

4. ANÁLISE DA ARQUITETURA DE TRANSMISSÃO DAS 12 CIDADES SEDES COM RESOLUÇÃO 4K

Por analogia com a análise da transmissão de áudio e vídeo HD, a estrutura necessária para as transmissões 4K é similar, porém com a rede corretamente redimensionada, ou seja, com banda ampliada de forma a suportar o maior *stream* de dados do sinal 4K.

Considerando os estudos citados nesse trabalho, através de um codificador JPEG200, sabemos que é possível reduzir o sinal de vídeo 4K a uma taxa média de 300 Mbit/s [8].

Após as etapas iniciais, quando o sinal já estiver finalizado e pronto para transmissão, o primeiro passo então será a compressão e o encapsulamento IP. Dessa fase teremos então como produto um TS de aproximadamente 330 Mbit/s. Sendo assim, o fluxo na rede dos três sinais deverá ser aproximadamente 1 Gbit/s .

Portanto, considerando a mesma arquitetura de rede apresentada na figura 6, com quatro anéis, onde um deles é composto pelas principais cidades que farão o gerenciamento dessa rede para a transmissão dos jogos da copa do mundo, para as doze cidades com sinais 4K comprimidos em JPEG2000, deverá ter disponível aproximadamente 2 Gbit/s.Essa banda não é suficiente se considerarmos a redundância de sinais, existe nessa estimativa de cálculo apenas as reservas técnicas para segurança das transmissões. No caso de haver redundância em todos os sinais enviados, o fluxo total da rede passará a ser aproximadamente 2 Gbit/s. Considerando também uma sobra de banda, por segurança e valores comerciais de rede, a rede estimada é de então 10 Gbit/s.

5. CONCLUSÃO

A arquitetura da rede de contribuição para os jogos da copa do mundo de 2014 em 4K é a mesma proposta para os sinais HD. A estrutura e equipamento utilizados para ambos os casos são similares, porem com processamentos e capacidades diferentes. Considerando o mesmo padrão de compressão, o JPEG200, e a ausência de

redundância de sinais, enquanto a rede de sinais HD deverá possuir uma banda disponível de 1Gbit/s, para os sinais 4K a rede deverá ser de 2 Gbit/s, sem redundância. Considerando as redundâncias de sinais essas bandas passarão a ser respectivamente 2 Gbit/s e 10 Gbit/s.

Embora, a redução da banda necessária seja o grande diferencial para a estrutura aqui proposta, há outros pontos a serem considerados como os parâmetros da rede, e a garantia de qualidade da mesma para o trafego dos sinais.

Nesse trabalho alguns pontos não foram abordados e que sugerem estudos futuros. Entre eles, podemos citar mecanismos para garantir o menor atraso, com maior qualidade e menor perda de pacotes, identificar a codificação ideal para esse tipo de serviço e resolução, assim como a melhor rede de dados. E uma vez definido a melhor rede, e a melhor codificação, definir também os parâmetros de configuração dessa rede, tais como QoS, e atrasos tolerados.

AGRADECIMENTOS

Gostaríamos de agradecer ao SBT (Sistema Brasileiro de Televisão) pelo apoio na realização desse trabalho

REFERÊNCIAS

[1] ITU-R BT.624-4, "Characteristics of Television Systems". Report. 1990

[2] ABNT NBR 15602-1, "Televisão Digital Terrestre – Codificação de vídeo, áudio e multiplexação. Parte 1: Codificação de vídeo". Pag. 23. Brasil, Abril de 2008.

[3] http://www.dcimovies.com/DCIDigitalCinemaSystemSpecv1_2. Acessado em 18/02/2012 as 15:42:56

[4] http://www.cesweb.org/ Acessado em 07/04/2012 as 17:38:16

[5] H. Takahashi, D. Shirai, T. Murooka, T. Fujii; "Multipoint Streaming Technology for 4K Super-high-definition Motion Pictures", NTT Technical Review, vol. 5, No.5, Tokyo, Japão, Maio de 2007.

[6] K. Shirakawa, M. Nomura, K. Ishimaru, T. Yamaguchi, T. Fujii; "NTT Technology Supporting Digital Cinema", NTT Technical Review, vol. 4, No.7, pp 23-28, Tokyo, Japão, Julho de 2006.

[7] D.Shirai, K. Shimizu, Y. Sameshima, H. Takahashi; "6-Gbit/s Uncompressed 4K Video IP Stream Transmission and OXC Stream Switching Trial Using JNG II", NTT Technical Review, vol. 5, No.1, pp 78-82, Tokyo, Japão, Janeiro de 2007.

[8] D. Shirai, T. Yamaguchi, T. Shimizu, T. Murooka, and T. Fujii, "4K SHD Real-time Video Streaming System with JPEG 2000 Parallel Codec," Proc., APCCAS2006, 2006.

[9] T. Margolis, S. Brown, T. Cornish, H. Bracho, M. Stanton, T. C. M. de B. Carvalho, F. F. Redigolo, F. C. de Castro, K. Kanekof, J. de Almeida, C. I. da Silva, E. A. T. de Souza, "Tri-continental premiere of 4K feature movie via network streaming at FILE" , Future Generation Computer Systems, pp. 924-934, 2011

imagens. Estas imagens foram carregadas na memória do aparelho no início da aplicação, o que aumentou o tempo inicial de carregamento da mesma, porém tornou o uso mais fluído e ágil.

Outro problema detectado foi a ausência do suporte ao botão "voltar" do controle remoto por parte do aparelho. Como as teclas relativas às cores foram usadas ativamente, o botão zero (0) do controle remoto foi alçado a este papel.

O recebimento e envio de dados por parte da aplicação via canal de retorno, utilizado de modo intensivo por esta aplicação, foi implementado através da biblioteca TCP do Lua, disponibilizada pela PUC-Rio.

As chamadas externas são utilizadas para obtenção dos dados das redes sociais *Facebook* e *Twitter*, para o envio da resposta e busca dos resultados da enquete e, principalmente no início da aplicação, para buscar o status do tempo do streaming (*timecode*), necessário para sincronia da aplicação com o *stream* de vídeo, possibilitando a aplicação das legendas e cifras de música, entre outros. Esta sincronia poderia ser implementada facilmente com o uso da cadeia NPT (*Normal Play Time*) [9], contudo a ausência de suporte a este recurso por parte de muitos STB e implementações do Ginga em TVs inviabilizam seu uso pleno. A solução encontrada foi utilizar o canal de retorno, inquirindo um servidor ligado ao *playout* do tempo decorrido do vídeo. Com a resposta, o *script* faz os cálculos necessários e inicia um contador que vigora até o encerramento da aplicação, sincronizando os tempos aos eventos.

A obtenção dos dados das redes sociais e da enquete é feita por chamadas TCP realizadas por *coroutines* (co-rotinas) da linguagem LUA [10]. São buscadas através de conexão de ponte com os servidores de internet da TV Unesp, por razões de segurança, evitando a inserção de senha das devidas redes no código. Assim, quando o usuário acessa a área do Facebook ou do Twitter na aplicação, o aplicativo envia uma chamada ao servidor da TV Unesp, que por sua vez se comunica com os servidores das redes sociais correspondentes, devolvendo os dados em *plain text* (texto puro) para a aplicação, que exibe-os na tela. No caso da enquete, a escolha do usuário é enviada aos servidores da TV Unesp para processamento e armazenamento no banco de dados, e a parcial da enquete é enviada de volta ao aplicativo, que a exibe na tela.

O aplicativo possui por volta de 3 Megabytes, contando os arquivos NCL, Lua, as imagens e o script de auxílio a conexões TCP.

4. CONCLUSÃO

Como conclusão da elaboração da aplicação interativa do programa "Som e Prosa" da TV Unesp, verificou-se como ponto fundamental para o sucesso do projeto o desenvolvimento de processos conjuntos e simultâneos de todas as áreas e profissionais envolvidos. Para isso, foi necessária a sincronização dos cronogramas de desenvolvimento de todos os envolvidos e a realização de reuniões periódicas para que todos os profissionais pudessem se comunicar diretamente. Dessa forma o processo produtivo da aplicação tornou-se dinâmico, dando coesão e qualidade ao produto final.

O ganho mais substancial adquirido pela forma do processo produtivo empregado nesse modelo foi a realização da aplicação em tempo relativamente curto (tempo médio de 5 dias). Esse tempo foi considerado satisfatório para o modelo de produção televisivo, já que é também o tempo médio de produção semanal do programa, podendo ser incorporado ao seu modo de produção, com atualizações semanais, sem perdas de qualidade, desempenho e agilidade.

O aplicativo apresentou uma alternativa de design que preserva o conteúdo audiovisual em exibição, e por ser centralizado no usuário, torna o processo de usabilidade fácil e intuitivo, respeitando os atributos definidos por Nielsen.

Outro resultado expressivo da produção da aplicação foi a verificação de avanços significativos nas três áreas que atuaram na produção (comunicação, design e tecnologia da informação). Toda a equipe envolvida no processo apresentou uma melhora significativa em seu desempenho geral e especifico, além de adquirir um conhecimento sistêmico do processo de produção de aplicações interativas para televisão digital, que os capacita a desenvolver novas aplicações para diferentes gêneros e formatos de programas. Criando um quadro de profissionais qualificado para a emissora.

Em relação a aplicação para o programa "Som e Prosa", obteve-se um produto final de excelente qualidade. Desenvolveu-se um processo de produção ágil e em sintonia com a realidade da emissora. Foram geradas novas qualificações nos profissionais envolvidos com o projeto, o que agregou valores individuais e coletivos à TV Unesp. Além disso, obteve-se integração entre as "novas" áreas da emissora (Tecnologia da Informação) e as áreas "clássicas" da televisão (Criação de Conteúdo e Design) criando-se as competências necessárias para a produção de aplicações interativas em Ginga na emissora.

Através de pesquisas e experimentações a TV Unesp pretende colaborar e estimular estudos na área de televisão digital em emissoras universitárias e integrar-se efetivamente a nova era que vive a transmissão de televisão aberta no Brasil.

AGRADECIMENTOS

Obrigado a todos os colegas de trabalho da TV Unesp envolvidos direta ou indiretamente neste projeto e quem cumprem com muita honra e profissionalismo o papel de prestadores de serviço à comunidade. Agradecemos também a direção da emissora pela oportunidade de realizar pesquisas de tamanha importância.

Desenvolvimento de Aplicação Interativa em Ginga para o programa “Som e Prosa” da Televisão Universitária Unesp.

Lucas Azevedo (lucasazevedo.tv@gmail.com)

DESENVOLVIMENTO DE APLICAÇÃO INTERATIVA EM GINGA PARA O PROGRAMA "SOM E PROSA" DA TELEVISÃO UNIVERSITÁRIA UNESP.

Lucas Silveira de Azevedo[1], *Fabio Cardoso*[2], *Fernando Ramos Geloneze*[3], *Rene Lopez*[4]

[1] Videografista TV Unesp; Mestrando PPGTVD - Unesp, Bauru, Brasil, lucas@tv.unesp.br
[2] Programador TV Unesp; Mestrando PPGTVD - Unesp, Bauru, Brasil, fabiocardoso@tv.unesp.br
[3] Coordenador Multimidia TV Unesp; Mestre PPGTVD - Unesp, Bauru, Brasil, fernando@tv.unesp.br
[4] Produtor TV Unesp; Mestre PPGTVD - Unesp, Bauru, Brasil, rene@tv.unesp.br

Resumo: O presente artigo demonstra o processo de desenvolvimento da aplicação interativa, em Ginga, do programa "Som e Prosa" da Televisão Universitária Unesp. Neste projeto, diferentes áreas trabalharam de modo integrado, o que resultou em uma visão global de todo processo de produção. Além de trazer uma breve discussão dos aspectos positivos e negativos da produção da aplicação em Ginga, este artigo busca elencar as diferentes etapas do desenvolvimento de um produto interativo para a TV. Da elaboração do conteúdo à validação do produto final interativo, optou-se por uma metodologia de trabalho que pudesse corresponder ao dia-a-dia de uma emissora televisiva, tanto em recursos pessoais, quanto em tempo hábil para a finalização do produto.

Palavras chave: televisão digital, televisão universitária, interatividade, ginga, usabilidade.

1. INTRODUÇÃO

A Televisão Universitária Unesp (TV Unesp), sediada na cidade de Bauru em São Paulo, é uma emissora criada em 2006 pela Universidade Estadual Paulista "Julio de Mesquita Filho". Entre seus objetivos estão: a elaboração de programas de televisão culturais e educativos que atendam a comunidade; ser um centro de pesquisas nas áreas de comunicação, tecnologia e educação; colaborar com o desenvolvimento do Sistema Brasileiro de Televisão Digital Terrestre - SBTVD-T (objetivo também impulsionado por se tratar de uma emissora operando desde sua origem nesse padrão); ser uma emissora interativa e conectada a diferentes meios de comunicação.

O SBTVD-T, criado com a intenção de ser um sistema convergente com múltiplas plataformas, tem a interatividade como uma das suas principais aliadas no processo de expansão da televisão digital. No entanto, os conceitos formulados para o uso da interatividade têm sido utilizados de forma bem genérica. "Um termo tão elástico corre o risco de abarcar tamanha gama de fenômenos a ponto de não poder exprimir mais coisa alguma." [1]

Em seu livro "TV Digital e Produção Interativa: a comunidade manda notícias", Fernando Crocomo [2] divide a interatividade em três níveis. No primeiro, nível 1, o usuário "navega" pelas informações armazenadas no terminal de acesso, disponíveis em hipertexto na tela. No nível 2, utiliza-se um canal de retorno, via rede telefônica ou internet, no qual é possível retornar a mensagem, mas não necessariamente em tempo real. Por último, o nível 3 permite enviar e receber conteúdo em tempo real através de conexão à internet.

No escopo deste projeto, entende-se por interatividade a ação que possibilite o recebimento e envio de conteúdos complementares à obra audiovisual, no momento de sua exibição. Assim, o protótipo desenvolvido oferece ao usuário informações diversas, relacionadas com o conteúdo, em tempo real e em diferentes níveis - quando conectado ao canal de retorno.

2. OBJETIVOS

A partir de seus objetivos - mais notadamente ao que se refere ao desenvolvimento de pesquisas e produções de produtos interativos - a TV Unesp, por meio dos seus funcionários, deu início ao desenvolvimento de uma aplicação interativa em Ginga, que pudesse ser incorporada ao processo de produção diário da emissora e que atendesse as necessidades básicas de usabilidade para os usuários.

Para tal, foi escolhido um dos destaques da grade de programação - o programa "Som e Prosa". Voltado ao público jovem e com seus temas relacionados ao universo musical, a escolha do programa justificou-se pelo perfil dos usuários, que possuem certa intimidade com as novas mídias, o que facilita o processo de usabilidade.

3. MÉTODOS

3.1. Processo de Produção:

O processo de produção da aplicação interativa se desenvolveu utilizando a metodologia *Agile Development* [3]. Essa metodologia baseia-se no desenvolvimento de pequenos projetos e produtos derivados do produto final (fig. 01). Assim, ao longo do processo, necessidades de alterações e correções de problemas, podem ser detectadas durante o desenvolvimento dos subprodutos e serem executadas e/ou corrigidas sem a necessidade de alteração de todo o produto. Esse método diminui consideravelmente o tempo de desenvolvimento da

aplicação interativa, pois elimina a necessidade de longos processos de prototipagem e correções de erros, adequando-se assim a realidade das emissoras de televisão.

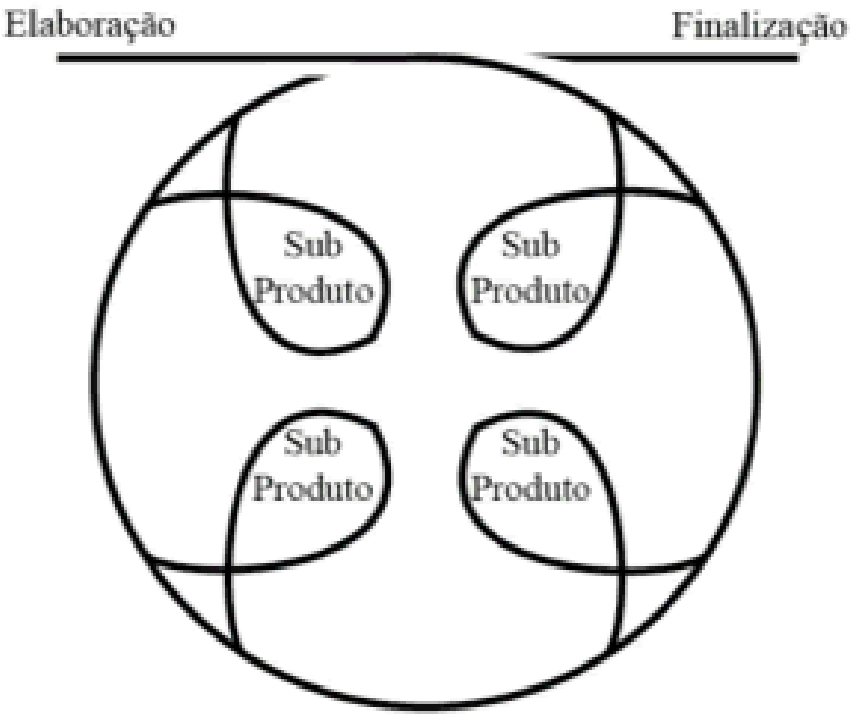

Fig. 01. Processo *Agile Development* para desenvolvimento de produtos

Além disso, o *Agile Development* preconiza que diferentes profissionais das diversas áreas envolvidas no projeto estejam em contato constante durante todo o processo, de forma que, com as trocas de experiência, cada individuo tenha consciência de todo o processo de produção e participe ativamente de todas as etapas do processo. Essa prática traz um ganho considerável na qualidade final da aplicação interativa.

A partir da escolha do modelo de produção, foi montada a equipe que realizou todas as etapas e cobriu todas as áreas do projeto. Durante a escolha dos membros do grupo de produção da aplicação, foi fundamental que todos tivessem, no mínimo, conhecimentos básicos sobre o processo produtivo de uma aplicação interativa, para que o projeto ocorresse de forma rápida e fosse obtido um resultado positivo. Outra característica fundamental da equipe foi seu caráter multidisciplinar. Isso se fez necessário dada a complexidade da produção de uma aplicação interativa e as diferentes áreas do conhecimento que sua concepção, desenvolvimento e conclusão exigem.

Assim, foram reunidos profissionais que, além de já compreenderem o processo básico de desenvolvimento de uma aplicação interativa para televisão digital, tinham conhecimento específico nas áreas fundamentais. São elas: Comunicação, Design e Tecnologia da Informação.

3.2. Conteúdo:

Dividido em execuções musicas e entrevistas, o programa "Som e Prosa" tem como intuito a valorização da produção musical independente, bem como a discussão de importantes temas relacionados ao universo musical. Deste modo, os diversos espaços de apreciação musical foram entendidos como momentos oportunos para o uso das aplicações interativas, uma vez que não exigiriam um maior nível de atenção por parte do usuário.

Assim, ainda na fase de roteirização, os profissionais de comunicação buscaram elaborar o conteúdo da aplicação interativa. Inicialmente, com o auxílio dos demais profissionais das diferentes áreas, foi desenvolvida a ideia da aplicação buscando conceituar todas as suas características básicas através do levantamento das seguintes questões: Quais conteúdos extras são relevantes ao programa? Quais informações dos artistas que participaram do programa merecem destaque? Quais serviços relacionados ao programa poderiam ser oferecidos via aplicação interativa?

A partir destes questionamentos foi decidido que os conteúdos que fariam parte da aplicação seriam: informações gerais do programa, sinopse, enquete, acesso às atualizações do *Facebook* e *Twitter* referentes ao programa e seu conteúdo, informações gerais sobre a banda principal e a banda do quadro "Contraponto", bem como seus integrantes, álbuns, letras, traduções, cifras, agenda e curiosidades, além do conteúdo voltado para orientar o usuário a utilizar a aplicação e informações sobre a emissora.

3.3. Design e Usabilidade

À medida que a interatividade no programa "Som e Prosa" deve corresponder às expectativas do usuário, as interfaces dos aplicativos devem atender aos objetivos de comunicação da televisão, tais como, divertir, entreter, informar, opinar e influenciar.

Baseado na tradição de Design de Interação [4] o processo de design de interface foi composto por quatro atividades básicas: identificação das necessidades dos usuários, desenvolvimento de alternativas de design, construção de versões iterativas e avaliação do design. Três aspectos chaves foram considerados: foco no usuário, definição de metas de usabilidade e iteração.

O foco no usuário é demonstrado no cuidado com o conteúdo e como ele é apresentado ao usuário. A usabilidade é baseada nos cinco atributos definidos por Nielsen em seu livro *Usability Enginnering*, são eles: facilidade de aprendizado, eficiência de uso, facilidade de memorização, baixa taxa de erros e satisfação subjetiva [5]. Os critérios de usabilidade foram identificados, concordados e documentados no início do projeto e são importantes no momento da escolha das alternativas de design (fig. 02).

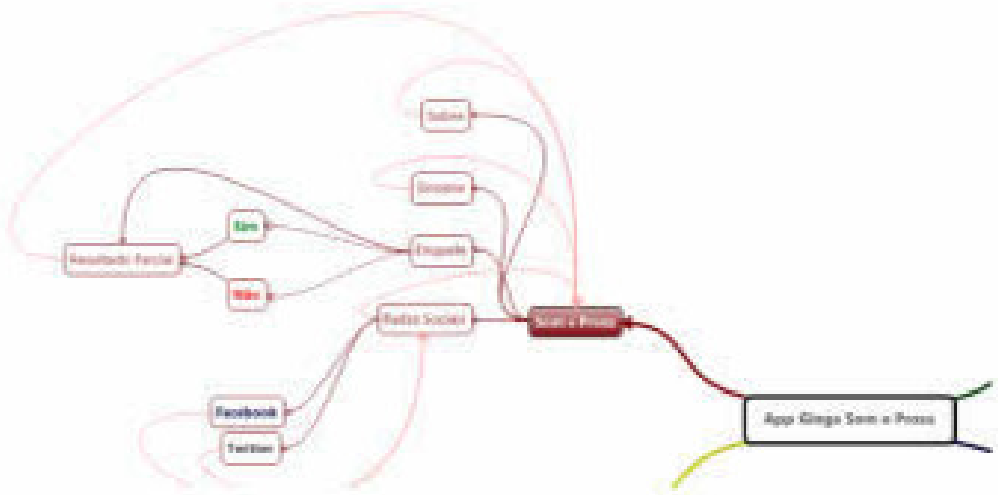

Fig. 02. Parte do mapa de navegação, mostrando um dos 4 menus e seus submenus no aplicativo "Som e Prosa".

Por último, iteração é o processo descrito no *Agile Development*, no qual subprodutos são gerados durante o desenvolvimento o que permite o refino baseado em avaliações feitas nessas versões intermediárias.

A aplicação, quando apresentada ao usuário, mostra quatro menus localizados cada um em um canto da tela. Dessa forma a aplicação não compromete o fluxo de vídeo do programa. Através dos quatro botões coloridos de interatividade pode-se escolher entre uma das opções quando selecionada a cor correspondente. O botão vermelho refere-se ao menu "Som e Prosa", o botão verde ao menu "Banda", o botão amarelo ao menu "Contraponto" e o botão azul dá acesso ao menu "Ajuda" (fig. 03)

Fig. 03. Aplicativo interativo do programa "Som e Prosa"

A navegação é feita pelas setas do controle remoto e a seleção da função pelo botão "ok". O botão que corresponde à função "voltar" não funcionou no modelo de Set-top-box utilizado nos testes, sendo então substituído pelo botão "0". Para sair da aplicação o botão utilizado foi o "*exit*" (fig. 04).

Fig. 04. Botões de navegação do aplicativo

No menu "Som e Prosa" são oferecidas opções como: informações gerais sobre o programa, sinopse do produto em exibição, enquete (fig. 05) onde o usuário é convidado a dar a sua opinião sobre o assunto em questão - tornando-se necessário, apenas nesse caso, conexão à internet - e acesso às atualizações do *Facebook* e *Twitter* referentes ao programa e seu conteúdo (fig. 06).

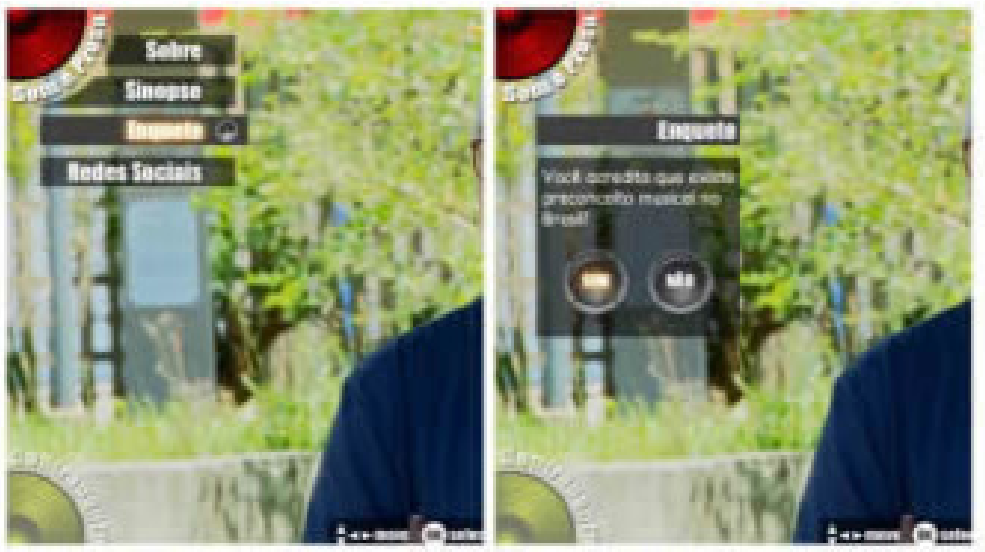

Fig. 05. Enquete no menu "Som e Prosa"

Fig. 06. Atualizações do Twitter

Informações sobre a banda principal do programa, seus integrantes, álbuns, letras, traduções, cifras, agenda e curiosidades são encontradas no menu "Banda". Letras, traduções e cifras podem ser apresentadas automaticamente sempre que uma música é exibida no programa (fig. 07). Basta habilitar a função nos submenus correspondentes.

Fig. 07. Letra e cifra da música em execução.

Houve a necessidade de paginação na apresentação dos integrantes, álbuns e curiosidades, devido a grande quantidade de informações. (fig. 08 e 09)

Fig. 08. Paginação das informações sobre os integrantes.

Fig. 09. Paginação das informações sobre os álbuns.

Conteúdo similar ao do menu "Banda" pode ser encontrado no menu "Contraponto", que corresponde a um dos quadros do programa (fig. 10).

Fig. 10. Menu "Contraponto".

Por último o menu "Ajuda" apresenta informações que ajudam o usuário a utilizar o aplicativo, além de informações sobre a TV Unesp.

Vale ressaltar que o design da interface, que deve seguir a identidade visual proposta para o programa, foi prejudicado por restrições do *middleware* e do *hardware*, o que impossibilitou, em alguns casos, o uso de gradiente de cor e transparência.

3.4. Programação

De acordo com o padrão vigente no Brasil de desenvolvimento de aplicativos para a Televisão Digital Interativa, o aplicativo "Som e Prosa" foi programado utilizando a linguagem NCL (*Nested Context Language*) [6] e com algumas bibliotecas de código e montagem de estruturas em LUA [7].

De início, a aplicação foi testada na máquina virtual Ginga disponibilizada pelo governo, o Set-Top-Box Virtual Ginga-NCL, que consta no Portal do Software Público Brasileiro [8]. O ambiente virtual é satisfatório para o desenvolvimento inicial, porém os vários dispositivos com o *middleware* embarcado respondem de modo diferente ao mesmo programa, devido as implementações distintas dos padrões Ginga, que fogem das normas técnicas vigentes na ABNT. O que inviabiliza ajustes finos da aplicação em ambientes virtuais. Diante a esta peculiaridade, os ajustes e adaptações do produto foram feitos tendo como base um *Set-Top-Box* rodando uma implementação do Ginga da empresa EITv, em sua versão 1.3.

O aplicativo foi rodado no STB usando seu recurso de aplicativos enviados por rede, hospedados em máquina local, com um stream de vídeo em HD rodando no software VLC e enviado por *multicast* para o Set-top Box via seu recurso de canais de IP (*ip channels*).

Assim que o aplicativo é invocado pelo usuário através do botão "ok" do controle remoto, o documento NCL é carregado, que por sua vez carrega em cascata o arquivo LUA responsável pela montagem da aplicação e define o evento de captura das teclas do controle remoto.

O arquivo LUA da aplicação realiza o carregamento da biblioteca responsável pelas chamadas TCP e das imagens que compõem o aplicativo. Logo após, chama um quadro de desenho (elemento *canvas)* de fundo transparente, recurso da linguagem LUA, que é sobreposto ao vídeo. Neste quadro de desenho, as figuras vão sendo encaixadas de acordo com suas posições assim que são necessárias, baseado no apertar dos botões do controle remoto.

A escolha de carregar todas as imagens na memória do aparelho antes da aplicação estar disponível para uso aumentou o tempo de início da aplicação, mas possibilitou um tempo de resposta muito melhor durante o seu uso. O tempo médio de carregamento de uma imagem em PNG com transparência na aplicação rodando no STB da EITv foi de 0.4 segundo. O carregamento inicial de todas as imagens na memória leva em torno de 6 segundos, e toda a chamada desta imagem posteriormente é exibida de forma imediata.

Depois de transportado da máquina virtual para o STB, certas otimizações no código da aplicação foram necessárias. Houve um ajuste no limite das imagens e devido ao desempenho insatisfatório da renderização de tipos (fontes) do hardware, todas as áreas com conteúdo de texto da aplicação foram substituídas por

imagens. Estas imagens foram carregadas na memória do aparelho no início da aplicação, o que aumentou o tempo inicial de carregamento da mesma, porém tornou o uso mais fluído e ágil.

Outro problema detectado foi a ausência do suporte ao botão "voltar" do controle remoto por parte do aparelho. Como as teclas relativas às cores foram usadas ativamente, o botão zero (0) do controle remoto foi alçado a este papel.

O recebimento e envio de dados por parte da aplicação via canal de retorno, utilizado de modo intensivo por esta aplicação, foi implementado através da biblioteca TCP do Lua, disponibilizada pela PUC-Rio.

As chamadas externas são utilizadas para obtenção dos dados das redes sociais *Facebook* e *Twitter*, para o envio da resposta e busca dos resultados da enquete e, principalmente no início da aplicação, para buscar o status do tempo do streaming (*timecode*), necessário para sincronia da aplicação com o *stream* de vídeo, possibilitando a aplicação das legendas e cifras de música, entre outros. Esta sincronia poderia ser implementada facilmente com o uso da cadeia NPT (*Normal Play Time*) [9], contudo a ausência de suporte a este recurso por parte de muitos STB e implementações do Ginga em TVs inviabilizam seu uso pleno. A solução encontrada foi utilizar o canal de retorno, inquirindo um servidor ligado ao *playout* do tempo decorrido do vídeo. Com a resposta, o *script* faz os cálculos necessários e inicia um contador que vigora até o encerramento da aplicação, sincronizando os tempos aos eventos.

A obtenção dos dados das redes sociais e da enquete é feita por chamadas TCP realizadas por *coroutines* (co-rotinas) da linguagem LUA [10]. São buscadas através de conexão de ponte com os servidores de internet da TV Unesp, por razões de segurança, evitando a inserção de senha das devidas redes no código. Assim, quando o usuário acessa a área do Facebook ou do Twitter na aplicação, o aplicativo envia uma chamada ao servidor da TV Unesp, que por sua vez se comunica com os servidores das redes sociais correspondentes, devolvendo os dados em *plain text* (texto puro) para a aplicação, que exibe-os na tela. No caso da enquete, a escolha do usuário é enviada aos servidores da TV Unesp para processamento e armazenamento no banco de dados, e a parcial da enquete é enviada de volta ao aplicativo, que a exibe na tela.

O aplicativo possui por volta de 3 Megabytes, contando os arquivos NCL, Lua, as imagens e o script de auxílio a conexões TCP.

4. CONCLUSÃO

Como conclusão da elaboração da aplicação interativa do programa "Som e Prosa" da TV Unesp, verificou-se como ponto fundamental para o sucesso do projeto o desenvolvimento de processos conjuntos e simultâneos de todas as áreas e profissionais envolvidos. Para isso, foi necessária a sincronização dos cronogramas de desenvolvimento de todos os envolvidos e a realização de reuniões periódicas para que todos os profissionais pudessem se comunicar diretamente. Dessa forma o processo produtivo da aplicação tornou-se dinâmico, dando coesão e qualidade ao produto final.

O ganho mais substancial adquirido pela forma do processo produtivo empregado nesse modelo foi a realização da aplicação em tempo relativamente curto (tempo médio de 5 dias). Esse tempo foi considerado satisfatório para o modelo de produção televisivo, já que é também o tempo médio de produção semanal do programa, podendo ser incorporado ao seu modo de produção, com atualizações semanais, sem perdas de qualidade, desempenho e agilidade.

O aplicativo apresentou uma alternativa de design que preserva o conteúdo audiovisual em exibição, e por ser centralizado no usuário, torna o processo de usabilidade fácil e intuitivo, respeitando os atributos definidos por Nielsen.

Outro resultado expressivo da produção da aplicação foi a verificação de avanços significativos nas três áreas que atuaram na produção (comunicação, design e tecnologia da informação). Toda a equipe envolvida no processo apresentou uma melhora significativa em seu desempenho geral e especifico, além de adquirir um conhecimento sistêmico do processo de produção de aplicações interativas para televisão digital, que os capacita a desenvolver novas aplicações para diferentes gêneros e formatos de programas. Criando um quadro de profissionais qualificado para a emissora.

Em relação a aplicação para o programa "Som e Prosa", obteve-se um produto final de excelente qualidade. Desenvolveu-se um processo de produção ágil e em sintonia com a realidade da emissora. Foram geradas novas qualificações nos profissionais envolvidos com o projeto, o que agregou valores individuais e coletivos à TV Unesp. Além disso, obteve-se integração entre as "novas" áreas da emissora (Tecnologia da Informação) e as áreas "clássicas" da televisão (Criação de Conteúdo e Design) criando-se as competências necessárias para a produção de aplicações interativas em Ginga na emissora.

Através de pesquisas e experimentações a TV Unesp pretende colaborar e estimular estudos na área de televisão digital em emissoras universitárias e integrar-se efetivamente a nova era que vive a transmissão de televisão aberta no Brasil.

AGRADECIMENTOS

Obrigado a todos os colegas de trabalho da TV Unesp envolvidos direta ou indiretamente neste projeto e quem cumprem com muita honra e profissionalismo o papel de prestadores de serviço à comunidade. Agradecemos também a direção da emissora pela oportunidade de realizar pesquisas de tamanha importância.

REFERÊNCIAS

[1] A. Machado; "Pré-cinema & pós-cinema". Campinas: Papirus, 1997.

[2] F. A. Crocomo; "TV Digital Interativa: a comunidade manda notícias". Florianópolis: UFSC, 2007.

[3] L. Reicheld; "Agile Development". Berlin, Web 2.0 Expo Berlin, 2007.

[4] J. Preece, H. Sharp, Y. Rogers; "Interaction design: Beyond human-computer interaction. USA: John Wiley & Sons, Inc, 2002.

[5] J. Nielsen; "Usability Engineering". San Francisco: Academic Press. 1993.

[6] NCL Handbook. Disponível em: < http://handbook.ncl.org.br/doku.php>. Acesso em 08 de agosto de 2012.

[7] R. Ierusalimschy, L. H. de Figueiredo, W. Celes. Lua 5.1 Reference Manual. Lua.org, 2006.

[8] Set-top Box Virtual Ginga-NCL. Portal do Software Público, 2012. Disponível em: <http://www.softwarepublico.gov.br/dotlrn/clubs/ginga/gingancl/xowiki/gingancl_vm>. Acesso em 08 de agosto de 2012.

[9] ABNT NBR 15606, 2007.

[10] R. Ierusalimschy. Programming in Lua. Lua.org, 2006.

[11] V. Becker, C. Montez; "Televisão Digital Interativa: conceitos, desafios e perspectivas para o Brasil". 2 ed. Florianópolis: UFSC, 2005.

[12] C. R. S. Bolaño, V.C. Brittos. "A Televisão Brasileira na Era Digital: exclusão, esfera pública e movimentos estruturantes". Editora Paulus, São Paulo, 2007.

[13] M. Gawalisnk. "Interactive Television Production". Oxford. Focal Press, 2003.

[14] T. V. Ribeiro. "Interatividade na Era Digital: tecnologia extensora do homem". Porto Alegre: 2006.

[15] K. Chorianopoulos. "Animated character likeability revisited: The case of interactive TV". Journal of Usability Studies. UPA Press, 2006.

[16] W. Cybis. "Ergonomia e Usabilidade: conhecimentos, métodos e aplicações". São Paulo, Novatec Editora, 2007.

Controle do loudness aplicado ao SBTVD: Metodologia de Aplicação e Avaliação

Ricardo Costa (rcosta.am@gmail.com)
Keiko Fonseca (keikoveronicaono@gmail.com)

Controle do *loudness* aplicado ao SBTVD: Metodologia de Aplicação e Avaliação

Ricardo de Souza Costa [1], *Keiko Verônica Ono Fonseca* [2]

[1] Universidade Tecnológica Federal do Paraná, Curitiba-PR, Brasil, rcosta.am@gmail.com
[2] Universidade Tecnológica Federal do Paraná, Curitiba-PR, Brasil, keiko@utfpr.edu.br

Resumo: Este artigo aborda o controle do *loudness* no Sistema Brasileiro de TV Digital (SBTVD) em face às recentes normas internacionais vigentes. Uma metodologia para aplicação deste controle é proposta e seus resultados iniciais serão analisados em relação à qualidade do áudio avaliada pelo usuário.

Palavras chave: *Loudness, SBTVD, Controle, Norma Brasileira.*

1. INTRODUÇÃO

As variações de volume de áudio podem acontecer durante o decorrer da programação de uma mesma emissora ou quando há troca de um canal para outro. Este problema é enfrentado pelas emissoras de TV terrestre, mas também podem ocorrer em transmissões de TV a cabo, satélite ou IP [1].

Atualmente as emissoras produzem e exibem diversos gêneros de programas, nos quais podemos citar: jornalismo, auditório, novelas, musical, *reality shows*, entrevistas, filmes, séries, infantil, esportivo, dentre outros. Os gêneros buscam focar em um grande público ou em um nicho de público apenas. Desta maneira, os programas possuem características distintas e visam atender as necessidades de cada telespectador. Enquanto em um filme de suspense as variações abruptas de volume de áudio são essenciais, em um programa de entrevistas, as variações quase não existem e não são perceptíveis pelo telespectador [2].

Para agravar ainda mais a problemática das variações de volume de áudio, a exibição de comerciais e chamadas de programas podem utilizar um nível de volume de áudio acima da média do programa sintonizado com o intuito de chamar a atenção do telespectador [3]. Tais variações de volume de áudio geram reclamações por parte dos telespectadores, pois os mesmos devem ajustar o nível de volume de áudio, toda vez que este sai da sua zona de conforto auditivo [4].

Para diminuir impactos negativos dessa variação, o seu controle é tema de discussão entre pesquisadores e emissoras do mundo inteiro visando à criação de normas regulamentadoras e ferramentas de controle baseadas em tais normas [7], ou seja, relativas ao controle do *Loudness*. O termo *Loudness* se refere à percepção do nível de som que as pessoas ouvem através de um trecho de áudio (música, discurso, efeitos de som, etc) e depende de vários parâmetros, entre eles: frequência, duração, nível, etc. [5][6].

No Brasil, discute-se ainda a regulamentação e implantação do controle de *loudness*. Sua concepção leva em conta as normas internacionais vigentes sobre o assunto [8].

Este trabalho aborda o controle do *loudness* voltado para o Sistema Brasileiro de TV Digital relacionando-o com as normas internacionais vigentes, a ITU-R BS 1770-2[9], A/85[10] e CALM ACT [11]. Uma metodologia para análise do *loudness* e impacto do seu controle através de alguns parâmetros de áudio é proposta e aplicada em ensaios de testes do *loudness* de programas de algumas emissoras brasileiras. Os resultados desses ensaios são analisados visando estabelecer uma proposta de ajuste de parâmetros de áudio codificado no padrão AAC (*Advanced Audio Coding*) [12][13]. Os resultados obtidos serão confrontados em relação ao especificado na recomendação EBU-R 128[14].

Este trabalho está organizado da seguinte maneira: a Seção 2 descreve-se o estado da arte do controle do *loudness*. Na Seção 3 apresentam-se as normas vigentes relacionadas ao *loudness e* formas de medição. A Seção 4 um breve resumo das normas de medição do *loudness* e que serão utilizados neste trabalho. A seção 5 explica os parâmetros do AAC para medição do *loudness*. A seção 6 aborda a metodologia a ser realizada. A seção 7 abrange a discussão deste trabalho. A conclusão e referências seguem ao final deste trabalho respectivamente.

2. LOUDNESS

Nesta seção é abordada a definição e a classificação da avaliação do *loudness* baseado na percepção do sistema auditivo.

Em acústica, *loudness* é uma percepção subjetiva da pressão sonora [15]. É um atributo do Sistema Auditivo e não pode ser medido diretamente [16] [17]. Existem, no entanto duas maneiras de representar o *loudness*. O primeiro é chamado de *"loudness subjetivo"* (LS) e é definido como a sensação auditiva que nos permite ordenar os sons em uma escala desde o nível silencioso ao nível mais alto. O segundo é chamado de *"loudness calculado"* (LC) e é definido com um valor numérico único, isto é, os resultados são obtidos com base em medições nos sinais de áudio, de tal forma que o mesmo pode ser usado para classificar diferentes sons na mesma ordem que a intensidade subjetiva [18].

Na década de 1920, o conceito de volume foi introduzido para fazer a comparação entre diferentes

intensidades de sons [19]. Em 1933, a curva de Flecther-Munson (*Equal-Loudness-Level Contours*) foi estabelecida para medir o nível de volume, medindo pressão sonora em relação ao espectro de frequência sonora bem como uma unidade de medida chamada de *phon* (de percepção de intensidade de som) [20]. As diversas curvas de ponderação desenvolvidas até a década de 1960 basearam-se na curva Flecther-Munson [21].

3. LOUDNESS: METADADOS E LEGISLAÇÃO

O problema das variações do *loudness* entre as emissoras de TV e de suas bruscas mudanças durante a programação televisiva da própria emissora é frequente e muito antigo [22]. Nas últimas décadas, as emissoras têm usado uma variedade de equipamentos e boas práticas na tentativa de medir, investigar e controlar estas variações, e entretanto o problema ainda persiste [23].

Como já discutido na seção 2, a dificuldade de se medir o *loudness* é atribuída à sua natureza subjetiva. Com intuito de melhor medi-lo, diversos métodos e algoritmos de medição de *loudness* têm sido desenvolvidos ao longo dos anos [24].

Na falta de um método aceitável para medição do nível de intensidade de áudio (exceção dos EUA, Austrália e Reino Unido), algumas emissoras utilizam uma medida de liberdade para usar o ajuste de controle de *loudness* conforme suas necessidades e outras simplesmente ignoram esse parâmetro de ajuste do *loudness*, e consequentemente comprometem a Qualidade de Experiência do telespectador [25].

3.1. Metadado de Loudness

A televisão digital nos EUA e no Canadá inclui um mecanismo conhecido como Metadados que pode ser usado para corrigir as diferenças entre os programas de loudness. Metadado é a informação que descreve várias características de um recurso [26]. O mesmo pode ser inserido em um *Elementary Stream* de áudio de um determinado programa. No sistema ATSC, o metadado que indica e controla o *loudness* é o *dialnorm* (pertencente ao padrão de áudio AC3 [27]). Ele informa os dispositivos sobre o nível de intensidade do sinal de entrada, permitindo assim que estes dispositivos (como o decodificador de TV Digital do telespectador) possam realizar correções necessárias do nível de *loudness*. E nem todas as emissoras inserem esse metadado (*dialnorm*) no *Elementary Stream* de áudio ou quando o fazem, os parâmetros definidos no metadado estão de forma errônea [28].

No Brasil, um metadado correspondente ao metadado *dialnorm*, chamado *Dynamic Range Control – DRC* [29] é definido no padrão de áudio AAC. Esse metadado é detalhado nas seções posteriores deste trabalho.

3.2 Leis relacionadas ao loudness

Nos EUA, o CALM Act (*Commercial Advertisement Loudness Mitigation Act*) foi transformado em lei no dia 15 de Dezembro de 2010 em resposta a uma série de reclamações de consumidores sobre o alto nível de volume dos programas de TV, acumuladas desde o início da transmissão da TV comercial há 50 anos [30].

O FCC (*Federal Communication Comission*) é o órgão responsável na implementação do CALM Act e pela fiscalização das empresas de radiodifusão dos EUA, baseando-se no documento de recomendação do padrão ATSC (*Advanced Television Systems Committee*) [31][32], o A/85 (*Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television*).

O A/85 resumidamente aborda as seguintes recomendações [33]: (i) O *Loudness* de todos os segmentos devem ser medidos de acordo com a técnica especificada no ITU BS.1770-2. (ii) O valor *dialnorm (*Metadado de controle de *loudness* do AC3*)* para transmitir com cada segmento deve corresponder ao volume medido dentro de +/- 2LU. (iii) Recomenda-se o *dialnorm* [34] definido a -24 LKFS, mas qualquer valor pode ser usado, desde que ele corresponda ao volume do segmento.

Na Austrália, o FreeTV Austrália (FTVA), órgão que representa as emissoras de televisão, publicou em julho de 2010, o guia operacional OP-59. Este guia estabelece um conjunto de recomendações que descrevem como as emissoras devem gerir o controle de loudness dos programas [35]. O OP-59 recomenda que os programas de televisão utilizem o metadado *dialnorm*, normalizado para -24 LKFS conforme o algoritmo definido na norma ITU-R BS.1770.

No Reino Unido, as emissoras devem seguir regras estabelecidas pelo código do *Broadcast Committee of Advertising Standards* (BCAP) [36]. O código BCAP estipula que: (i) os equipamentos de medição do loudness devem atender a recomendação do ITU; (ii) as normas ITU-R BS1770 e ITU-R BS1771 devem ser atendidas; (iii) não é estabelecido nenhum valor para o nível de *loudness target* para programas e propagandas.

4. NORMAS E RECOMENDAÇÕES PARA MEDIÇÃO E NORMALIZAÇÃO DO LOUDNESS

4.1. ITU-R 1770

A finalidade desse padrão é de estabelecer um algoritmo para a medição do loudness e os níveis *true-peak* dos programas. É um padrão robusto e simples, pois seu algoritmo é de fácil implementação.

A ITU-R 1770 define uma curva de "ponderação K" (versão modificada de um filtro passa-alta de segunda ordem) que constitui uma correlação entre medição subjetiva com a medição objetiva.

Esta curva é aplicada a todos os canais de áudio, exceto o canal de "efeitos de baixa frequência" (LFE), no qual é descartado da medição. A energia total do *valor quadrático médio (mean square)* então é calculada (com diferentes fatores de ganho para os canais frontal e *surround*, como mostra figura 1) e o resultado é exibido como "LKFS" (*Loudness*, Ponderação K, *Full Scale*) [37].

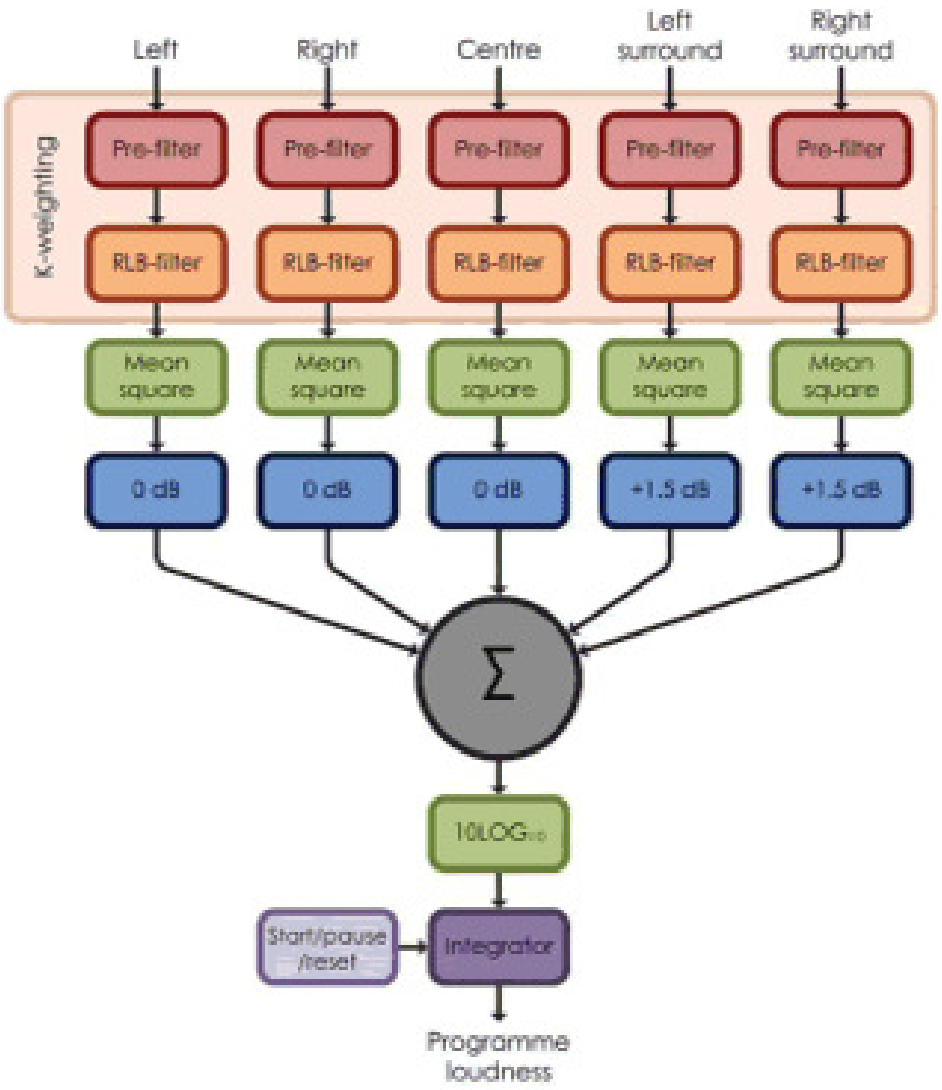

Fig. 1. Cálculo de *loudness* através da soma e processamento de canais de áudio conforme a norma ITU-R 1770 [37].

4.2. EBU 128 [14]

Os membros da *European Broadcast Union* (EBU) estabeleceram um grupo de estudo chamado P/LOUD com o objetivo de proporcionar níveis de *loudness* mais consistentes observados pelo ouvinte. O trabalho do grupo inclui também a investigação de novas formas de medição do *loudness* e de medição do sinal de pico para complementar ou até substituir os equipamentos VU e PPM. Em 2010, a EBU publicou a Recomendação R128 "normalização do *Loudness* e nível máximo permitido de sinais de áudio". A EBU R128 recomenda que o volume do áudio possa ser medido utilizando o método da norma ITU-R BS.1770 complementado por um método chamado "*gating*" para considerar as passagens silenciosas e suaves no sinal de áudio. Além disso, os programas devem ser normalizados para um nível de *loudness* de -23 LKFS com uma tolerância de ± 1 LU.

Os documentos de apoio incluem EBU Tech. 3341 [38], 3342 [39], 3343 [40] e 3344 [41] que descrevem em maiores detalhes, os aspectos: (i) métodos de medição, (ii) unidades de medidas, (iii) implementação de algoritmos e (iv) diretrizes de produção, distribuição e transmissão.

5. METADADOS DO PADRÃO DE ÁUDIO HE-AAC PARA CONTROLE DO *LOUDNESS*

Esta seção descreve resumidamente os parâmetros de controle do *loudness* do padrão de áudio HE-AAC, baseado do *elementary stream* multicanal (5.1 canais). A figura 2 mostra o frame FILL-ELEMENT e seus *sub-frames*.

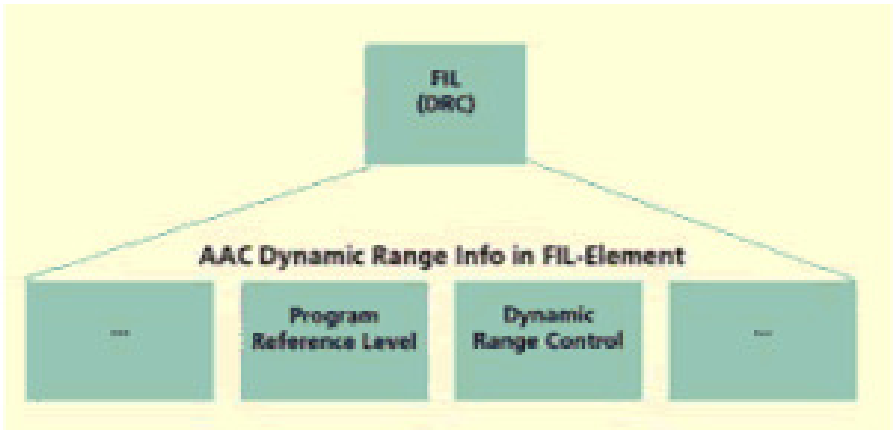

Fig. 2. Ilustração do frame AAC *Dynamic Range Info* expandido [29].

5.1. Program Reference Level

Este parâmetro pode ser utilizado para sinalizar o *loudness* médio de um programa transmitido. O parâmetro é representando em um campo de 7 bits que representa o nível de referência de *loudness* de um programa com *steps* de ajustes de 0,25 dB e com um *range* de 32dB. Durante a decodificação, este parâmetro permite que o nível de saída de áudio seja alterado, a fim de corresponder ao nível de *loudness* desejado [29].

O cálculo de ajuste do *loudness* é administrado pelo decodificador do usuário. Os resultados da mudança de nível (g) com a diferença entre o nível *target* de *loudness* do programa (L_{target}) e o nível de referência de *loudness* (L_{ref}), é aplicado aos dados espectrais anteriores para a decodificação de um *frame* de áudio [29].

$$g = L_{target} - L_{ref} \qquad (1)$$

A equação 2 mostra o cálculo da reconstrução do nível de *loudness* antes da realização do ajuste desejado. O nível "*full scale*" é 32767 e o (prog_ref_level) é um valor relativo de nível de áudio, conforme descrição na norma ISO/IEC 14496-3:2005, Subseção 4.5.2.7.2.

$$level = 32767 \cdot 2^{\frac{-prog_ref_level}{24}} \quad (2)$$

A Figura 3 ilustra o ajuste de um nível de referência de *loudness* de um programa de -31 dBFS a um nível *target* de -23 dBFS. DBFS refere-se a decibel relativo ao nível *full scale*.

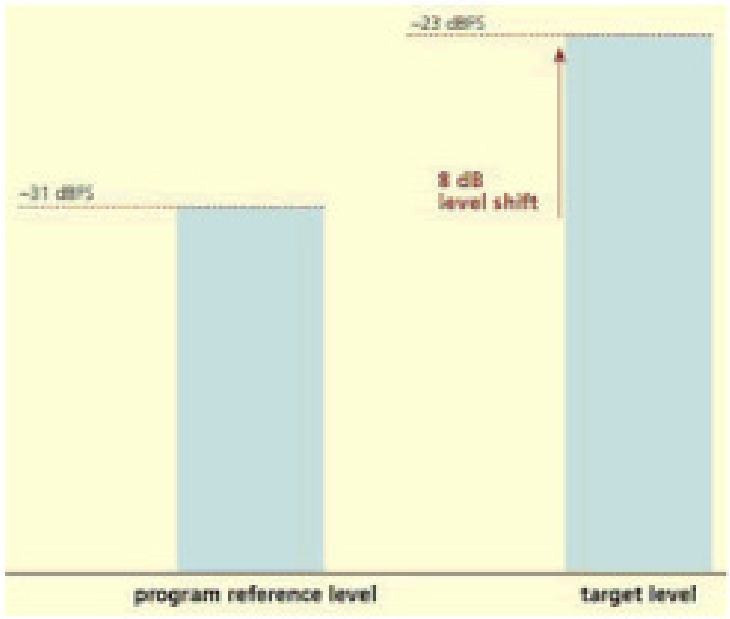

Fig. 3. Ilustração do ajuste nível de referência de loudness de um programa para o nível *target* [29].

Se nenhum nível de referência de *loudness* de um programa estiver presente no *Elementary Stream* de áudio, o nível de saída permanece inalterado.

5.2. Dynamic Range Control (DRC)

É o *frame* que contém todos os parâmetros necessários para o ajuste do nível de *loudness* de um programa além do mais importante já citado anteriormente (prog_ref_level) *Program Reference Level* conforme a norma ISO IEC 14496-3:2005, seção 4.5.2.7.

5.3. Matrix-Mixdown

É um indicador do coeficiente a ser utilizado no *downmix* para realizar a transição de áudio multicanal para estéreo. A figura 4 mostra onde o indicador de *downmix* pode ser encontrado. Detalhes do funcionamento do *Matrix-Mixdown*, está descrito na ISO/IEC 14496-3:2005, Subseção 4.5.1.2.2 e Tabela 4.70

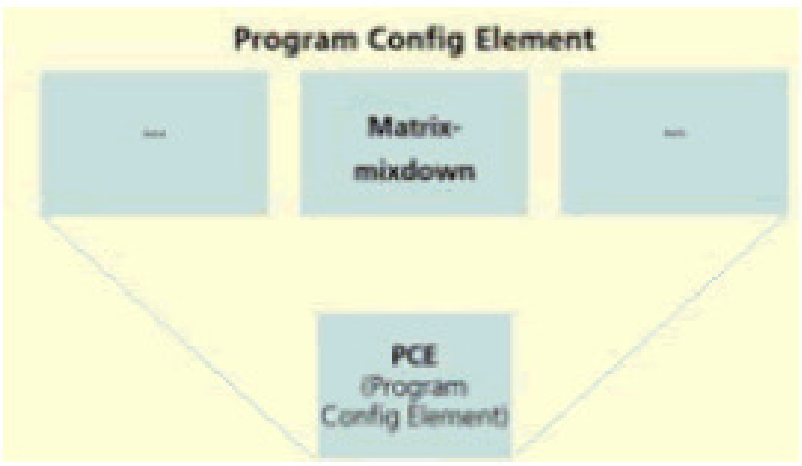

Fig. 4. Ilustração do frame PCE expandido [29].

A norma brasileira ABNT-NBR 16602-2 afirma que o uso *downmix* é obrigatório quando um programa em modo multicanal for transmitido.

Quando o parâmetro *matrix_mixdown_idx_present* estiver ativo em "1" no respectivo PCE do programa (PID), deve-se obrigatoriamente utilizar o sistema de *downmix*.

6. METODOLOGIA PROPOSTA PARA ESTE TRABALHO

Para a realização deste trabalho, serão necessários os seguintes passos, a saber:

a) Obtenção de *Transport Streams Over The Air*;

b) Extração e conversão dos *Elementary Streams* de Áudio;

c) Inserção de parâmetros para ajuste de *loudness* no *Elementary Stream* de áudio;

d) Avaliação objetiva e subjetiva do áudio antes e depois da inserção dos parâmetros;

6.1 Obtenção de Transport Streams Over The Air

A aquisição dos *TS* serão realizados através de equipamentos gravadores de TS ISDB-T (no Laboratório de Comunicação de Dados – LCD da Universidade Tecnológica Federal do Paraná), de duas emissoras apenas, A e B, que estão transmitindo TV Digital na região de Curitiba-PR, focando principalmente os horários de pico, com aquisição de 30 minutos e intervalos de 5 minutos entre as aquisições. A tabela abaixo, mostram os horários de aquisições relacionados aos dias da semana:

Tabela 1. Horário para aquisição de TS Over the Air

Dia da Semana	Horário 1	Horário 2
Segunda-feira	19:00 – 19:30	19:35 – 20:05
Quarta-Feira	19:00 – 19:30	19:35 – 20:05
Sexta-feira	19:00 – 19:30	19:35 – 20:05

6.2 Extração e conversão dos ES de Áudio

A extração dos *ES de áudio* serão realizadas através de ferramentas *freeware* ou de avaliação obtidos na internet. A ferramenta de extração deve extrair o áudio do TS e montar o arquivo HE-AAC LATM/LOAS, conforme a norma brasileira ABNT NBR-15602-2.

Para leitura do ES de áudio na análise pela ferramenta MATLAB, é necessário realizar a conversão

do padrão HE-AAC v2 LATM/LOAS para formato PCM (formato reconhecido pela ferramenta).

6.3 Inserção de parâmetros para ajuste de *loudness* no *Elementary Stream* de áudio

Para inserção dos parâmetros, será realizada através de um software de leitor de arquivos em hexadecimal, baseando-se em um TS que exista áudio HE-AAC 5.1 com parâmetros do Dynamic Range Info presente no TS.

6.4 Avaliação objetiva e subjetiva do áudio antes e depois da inserção dos parâmetros

A avaliação objetiva e subjetiva do *loudness* áudio será analisado conforme a norma EBU R128.

A avaliação subjetiva será realizada com 20 pessoas, utilizando um decodificador de TVD capaz de decodificar áudio HE-AAC 5.1 utilizando o recurso de *downmix* ou um software freeware, código livre ou de avaliação para PC.

A avaliação objetiva será realizada através da ferramenta chamada MATLAB e os seus resultados serão confrontados com os resultados da avaliação subjetiva do *loudness* e analisar se ambos resultados são equivalentes.

7. DISCUSSÃO

Diversas dificuldades estão sendo enfrentadas para a realização deste trabalho, nos quais podemos citar:

a) Aquisição de um TS 5.1 no formato HE-AAC V2 LATM/LOAS Over The Air. Os TS capturados e identificados como 5.1 estavam no padrão Dolby AC3;

b) Aquisição de um decodificador ISDB-Tb com suporte ao áudio HE-AAC V2 5.1 + DRC + Downmix. Os decodificadores ISDB-Tb disponibilizados no laboratório, não decodificavam o áudio HE-AAC v2 LATM/LOAS em 5.1 canais;

c) Aquisição de e/ou implementação de uma ferramenta que transforme áudio estéreo para 5.1 para o *ES* de áudio ser multiplexado no TS;

No entanto, os testes no momento estão sendo realizados através de TS obtidos pelas emissoras, com áudio HE-AAC e LC-AAC LATM/LOAS e ajustados conforme a recomendação EBU 128.

8. CONCLUSÃO

Este trabalho propõe que o ajuste do *loudness* seja realizada pelo decodificador do usuário, com a utilização dos parâmetros (metadados) definidos na norma de áudio do *Advanced Audio Coding*, a ISO IEC 14496-3:2005, referenciado pela norma brasileira ABNT-NBR 16602-2 – *Multiplexação de Áudio*.

A vantagem dessa solução é que o decodificador ajuste o *loudness* automaticamente baseado no nível alvo (usuário) com o nível de referência (emissora).

A desvantagem da solução é que se os decodificadores do padrão ISDB-Tb existentes no mercado brasileiro, não possuem suporte para decodificar áudio 5.1 no padrão AAC ou realizar o *matrix-mixdown* do áudio para estéreo isso dificultaria a implementação proposta neste trabalho.

Atualmente, não é do conhecimento dos autores se as emissoras estejam produzindo conteúdos com 5.1 no padrão AAC, o que dificulta a implementação desta proposta, ficando este limitado a testes de parâmetros de TS obtidos de conteúdo obtidos da Internet.

AGRADECIMENTOS

A Universidade Tecnológica Federal do Paraná – UTFPR, ao Programa de pós-graduação em Engenharia Elétrica e Informática Industrial – CPGEI, a CAPES através do projeto "Formação de Recursos Humanos em TV Digital" (PE-RHTVD) e pela Positivo Informática pela oportunidade de realizar este trabalho.

REFERÊNCIAS

[1] Poers, Peter. "Loudness RollerCoaster and how to survive it.", disponível em: http://junger-audio.com/technology/the-loudness-rollercoaster.

[2] Souza Brito, L. F de; et al. "Loudness: Oportunidade para solucionar o problema de variações de volume de áudio". Revista SET, ed. 118.

[3] FTVA. "Loudness in Advertisements", Australia. disponível em: http://www.freetv.com.au/content_common/pg-loudness-in-advertisements.seo.

[4] Moerman, Jean Paul: "Program Loudness: Nuts & Bolts". 118th AES Convention, 2005.

[5] Lund, Thomas. "Control of Loudness in Digital TV". NAB BEC Proceedings. Risskov, Denmark, 2006.

[6] EBU-EUR. "10 Things You Need to Know about Loudness". EBU Technical Media Technology & Innovation, 2010.

[7] Lund, Thomas. "ITU-R BS.1770 Revisited", TC Electronics A/S, Denmark.

[8] Fórum SBTVD. "SET apresenta soluções para diminuir o Loudness na TV", disponível em: http://www.forumsbtvd.org.br/materias.asp?id=735

[9] ITU-R BS 1770-2. "Algorithms to measure audio programme loudness and true-peak audio level", Geneva 2011.

[10] A/85. "Techniques for Establishing and Maintaining Audio Loudness for Digital Television", USA, 2009.

[11] CALM ACT,. "The Commercial Advertisement Loudness Mitigation Act ", USA, 2010.

[12] ISO/IEC 14496-3. "Information Technology – Coding of Audio-Visual Objects – Part 3", Third Edition, 2005.
[13] ABNT NBR-15602-2. "Televisão digital terrestre – Codificação de vídeo, áudio e multiplexação, Parte 2: Codificação de áudio", 2007 atualização 2008.
[14] EBU R128. "Loudness Recommendation", 2011.
[15] Poulsen, T. "Acoustic Communication, Hearing and Speech". 2 edition, 2005.
[16] Musiek, F. E. et. al. "The auditory System: Anatomy, Phisiology and Clinical Correlates", 2006.
[17] Moore, B. "An Introduction to the Psychology of Hearing." Great Britain: Academic Press, 2003.
[18] Liejon, A. "Sound Perception: Introduction and Exercise Problems". Stockholm, Royal Institute of Technology, 2007.
[19] Nygren,P. "Achieving Equal Loudness between Audio Files – evaluations and improvement of loudness algorithm". Royal Institute of Technology.
[20] Letcher, H. Munson, W. A. "Loudness, Its Definition, Measurement and Calculation". Bell Labs, 1933.
[21] Florentine, M. et al. "Loudness". Springer Handbook of Auditory Research, 2011.
[22] FCC Docket 14904, 27 Fed. Reg. 12681, Dec, 1962.
[23] Robjohns, H. "Metering". Sound on Sound Article FAQ, 2000. Disponível em: http://www.soundonsound.com/sos/jun00/articles/metring.htm.
[24] Corliss, E. L. R., Winzer E. George. Study of Methods for Estimating Loudness. USA, July 1965.
[25] Pyykkō, S. J., Häkkinen Jukka, et. Al. "Experienced Quality Factors – Qualitative Evaluation Approach to Audiovisual Quality", Jan, 2007.
[26] The University of Queenland. An introduction to Metadata. Disponível em: http://www.library.uq.edu.au/iad/ctmeta4.html
[27] A/53. "ATSC Digital Television Standard: Part 5 – AC-3 Audio System Characteristics". 2010.
[28] Norcross S., Lavoie M. "A Status Report on Loudness Control Technologies and Standardization for Broadcasting". Canada, March, 2011.
[29] Franhoufer. "White Paper. HE-AAC Metadata for Digital Broadcasting". Sept. 2011.
[30] FCC 11-84. "Implementation of the Commercial Advertisement Loudness Mitigation (CALM) Act Document". Washington D.C.
[31] NAB. "FCC Seeks Comment on Implementing the CALM Act". Newsletter, 2011. Disponível em: http://www.nab.org/xert/scitech/pdfs/tv061311.pdf.
[32] ATSC. "Advanced Television System Committee - Standards". Disponível em: <http://www.atsc.org/cms/index.php/standards>.
[33] QUALIS Audio. "Loudness Measurements and the CALM Act". Tech Note. Disponível em: < www.qualisaudio.com/documents/TechNote-1.pdf>.
[34] ETVCOOKBOOK. "Understanding Dialnorm". Disponível em: <http://etvcookbook.org/audio/dialnorm.html>.
[35] FTVA. OP-59 - Operational Guide. Australia, July, 2010.
[36] The BCAP Code: The UK Code of Broadcast Advertising Edition 1, Broadcast Committee of Advertising Practice, London, June 29, 2010.
[37] Camerer, F. "On the way to Loudness nirvana – audio leveling with EBU R128". EBU TECHNICAL REVIEW. Sept, 2010.
[38] EBU Tech Doc 3341, "Loudness metering: 'EBU Mode' metering to Supplement Loudness Normalization in Accordance with EBU R128", Geneva, 2010.
[39] EBU Tech Doc 3342, "Loudness Range: A Descriptor to Supplement Loudness Normalization in Accordance with EBU R128", Geneva, 2010.
[40] EBU Tech Doc 3343, "Practical Guidelines for Production and Implementation in Accordance with EBU R128", Geneva, 2011.
[41] EBU Tech Doc 3344, "Practical Guidelines for Distribution Systems in Accordance with EBU R128", Geneva, 2011.

Uso de Reconhecimento Automático de Voz em Português do Brasil na geração de Closed Caption

Luiz Brito (luiz.fausto@tvglobo.com.br)
Edilberto Strauss (strauss@poli.ufrj.br)
Flávio Mello (fmello@poli.ufrj.br)

USO DE RECONHECIMENTO AUTOMÁTICO DE VOZ EM PORTUGUÊS DO BRASIL NA GERAÇÃO DE CLOSED CAPTION

Luiz Fausto de Souza Brito [1], *Edilberto Strauss* [2], *Flávio Luis de Mello* [2]

[1] Rede Globo, Rio de Janeiro - RJ, Brasil, luiz.fausto@tvglobo.com.br
[2] Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ, Brasil, {strauss,fmello}@poli.ufrj.br

Resumo: Este trabalho tem a finalidade de apresentar a necessidade de utilização de legenda oculta (*closed caption*) na programação da televisão, tanto como recurso de acessibilidade, quanto para cumprimento de legislação específica. Para viabilizar tal recurso durante a ocorrência de fala espontânea ao vivo, é apresentada a tecnologia de reconhecimento automático de voz como uma alternativa à estenotipia com menor custo operacional. Em seguida, são apresentados alguns exemplos de utilização dessa abordagem relatados por diversos autores ao redor do mundo. Por fim, são apresentadas as dificuldades específicas de implementação desse tipo de sistema utilizando o idioma português do Brasil e alguns caminhos possíveis para a superação de tais dificuldades.

Palavras chave: *Closed Caption*, Reconhecimento Automático de Voz, Português do Brasil.

1. CLOSED CAPTION E ACESSIBILIDADE

De acordo com os resultados do Censo Demográfico 2010 do IBGE [1], cerca de 5,1% da população brasileira (quase 10 milhões de pessoas) possuem deficiência auditiva. Para que essa população tenha acesso ao conteúdo dos programas de televisão, é necessário que tais programas disponham de legendas, contendo não apenas a transcrição textual do que é falado, como também descrições de músicas e efeitos sonoros que sejam necessários para a melhor compreensão do conteúdo sendo exibido. Tais legendas, porém, não devem estar visíveis para todos os telespectadores, apenas para aqueles que assim desejarem, ativando tal recurso no receptor de TV. Isso é possível, tanto na transmissão de TV analógica [2], quanto na transmissão de TV digital [3], através de um recurso denominado "legenda oculta" ou "*closed caption*" (CC).

Embora a maior motivação para utilizar legendas ocultas na programação da televisão seja para garantir a acessibilidade aos telespectadores com deficiência auditiva, tal recurso também é interessante para o público em geral em algumas situações específicas, a saber: em ambientes nos quais a necessidade de silêncio não permite que o som da TV seja reproduzido ou pelo menos não com volume suficiente para uma compreensão adequada (e.g. hospitais); em ambientes muito barulhentos em que não seria possível compreender o áudio da televisão, ainda que reproduzido com volume elevado (e.g. bares e restaurantes); em locais com sonorização musical ambiente (e.g. academia); para permitir que diferentes televisores reproduzam programas diferenciados em um mesmo ambiente para atender a um público heterogêneo; em ambiente residencial para permitir que se assista TV à noite sem incomodar os demais moradores; para aprendizagem da língua por parte de estrangeiros ou pessoas em processo de alfabetização.

2. LEGISLAÇÃO

No Brasil, a lei nº 10.098 de 2000 [4], criada para estabelecer normas de acessibilidade, já previa que os serviços de radiodifusão de sons e imagens deveriam adotar medidas técnicas com o objetivo de permitir o uso de subtitulação, para garantir o direito de acesso à informação às pessoas portadoras de deficiência auditiva. Contudo, foi somente em 2006, que a Norma Complementar 01, aprovada pela Portaria nº 310 do Ministério das Comunicações [5] estabeleceu prazos, obrigando os radiodifusores a uma utilização progressiva de legendas ocultas nas suas programações. De acordo com o cronograma estabelecido, em 2017 a totalidade da programação diária das emissoras e retransmissoras deverá dispor de legendas ocultas, como pode ser verificado na Tabela 1.

Tabela 1. Prazos e cargas horárias mínimas para disponibilização de *closed caption* por faixa horária

Prazo	Faixa Horária	Carga Horária
27/06/2008	08:00 - 14:00	1 h
	20:00 - 02:00	1 h
27/06/2009	08:00 - 14:00	2 h
	18:00 - 02:00	2 h
27/06/2010	08:00 - 14:00	3 h
	18:00 - 02:00	3 h
27/06/2011	08:00 - 14:00	4 h
	18:00 - 02:00	4 h
27/06/2012	06:00 - 14:00	6 h
	18:00 - 02:00	6 h
27/04/2014	06:00 - 02:00	16 h
27/04/2015	Dia Inteiro	20 h
27/06/2017	Dia Inteiro	Totalidade da Programação

As estações transmissoras ou retransmissoras analógicas que não suportarem a transmissão de *closed caption* devem se adequar de acordo com um prazo estabelecido conforme a população da cidade onde a estação estiver localizada, como pode ser observado na Tabela 2. As estações transmissoras ou retransmissoras digitais devem suportar a transmissão de *closed caption* desde o princípio da operação.

Tabela 2. Prazos de adequação da infraestrutura das estações analógicas para transmissão de *closed caption*, de acordo com a população da cidade onde a estação estiver localizada

População	Prazo
> 1.000.000	27/06/2008
> 500.000	27/06/2010
> 200.000	27/06/2012
> 100.000	27/06/2014
Qualquer	27/06/2016

Além disso, a Norma Complementar 01 de 2006 do Ministério das Comunicações estabelece que a produção e veiculação de legendas ocultas deve atender os critérios e requisitos técnicos da ABNT NBR 15290:2005. Dentre

esses requisitos, consta que o texto das legendas deve ter 100% de acerto acompanhando o tempo exato da cena para programas com legenda pré-gravada e deve ter no mínimo 98% de acerto com no máximo 4 segundos de atraso para programas com legenda ao vivo [6].

Tais exigências afetam todas as emissoras de televisão do Brasil, independente do porte. Mesmo as retransmissoras ou afiliadas das redes de televisão, deverão veicular legendas ocultas para sua programação e comercialização, inclusive local, atendendo os requisitos técnicos especificados. O descumprimento das disposições da norma sujeita as emissoras ou retransmissoras às penalidades prescritas no Código Brasileiro de Telecomunicações.

Maximizar a taxa de acerto das legendas ocultas é importante não apenas para cumprir a legislação, como também para viabilizar a compreensão das mensagens sendo veiculadas. Por vezes, apenas uma palavra errada é suficiente para modificar profundamente o significado de um texto.

O maior desafio está na produção de legendas ocultas ao vivo, sobretudo quando há fala espontânea (sem que haja um texto pré-definido em roteiro). Essa situação ocorre com bastante frequência na programação televisiva: programas de entrevistas, de auditório, religiosos, de variedades, *reality shows* etc. Existem atualmente duas formas de produzir as legendas ocultas que se aplicam à transcrição da fala espontânea ao vivo: estenotipia e reconhecimento automático de voz. Quando a fala ao vivo utiliza a leitura de um *teleprompt*, o texto do *teleprompt* pode ser utilizado como legenda oculta. Para a parte da programação que não é produzida ao vivo, a legenda oculta pode ser inserida por digitadores durante o processo de pós-produção.

3. ESTENOTIPIA

Trata-se do uso de um profissional especializado que digita as palavras na velocidade em que as ouve, utilizando símbolos fonéticos em um equipamento com teclado especial (estenótipo), conectado a um computador, que utilizando um dicionário específico converte as sequências de símbolos em palavras.

Um curso de formação profissional em estenotipia dura de seis meses a um ano, mas normalmente é necessário pelo menos dois anos de prática para obter a velocidade necessária para legendagem ao vivo (180 palavras por minuto). É uma mão-de-obra relativamente cara, escassa e difícil de repor. De acordo com matéria publicada na revista Galileu de maio de 2012, existem apenas 400 profissionais desses no Brasil [7].

Erros no processo de estenotipia podem ocorrer tanto por falha humana na digitação quanto pela ocorrência de palavras fora do dicionário.

4. RECONHECIMENTO AUTOMÁTICO DE VOZ

O reconhecimento automático de voz é o processo de conversão do sinal acústico da voz em uma transcrição textual correspondente. Um sistema que implemente tal processo inclui a captação do sinal de voz, sua digitalização, processamento digital de sinais e modelos estatísticos. Uma representação esquemática de um sistema de reconhecimento automático de voz pode ser visualizada na Figura 1.

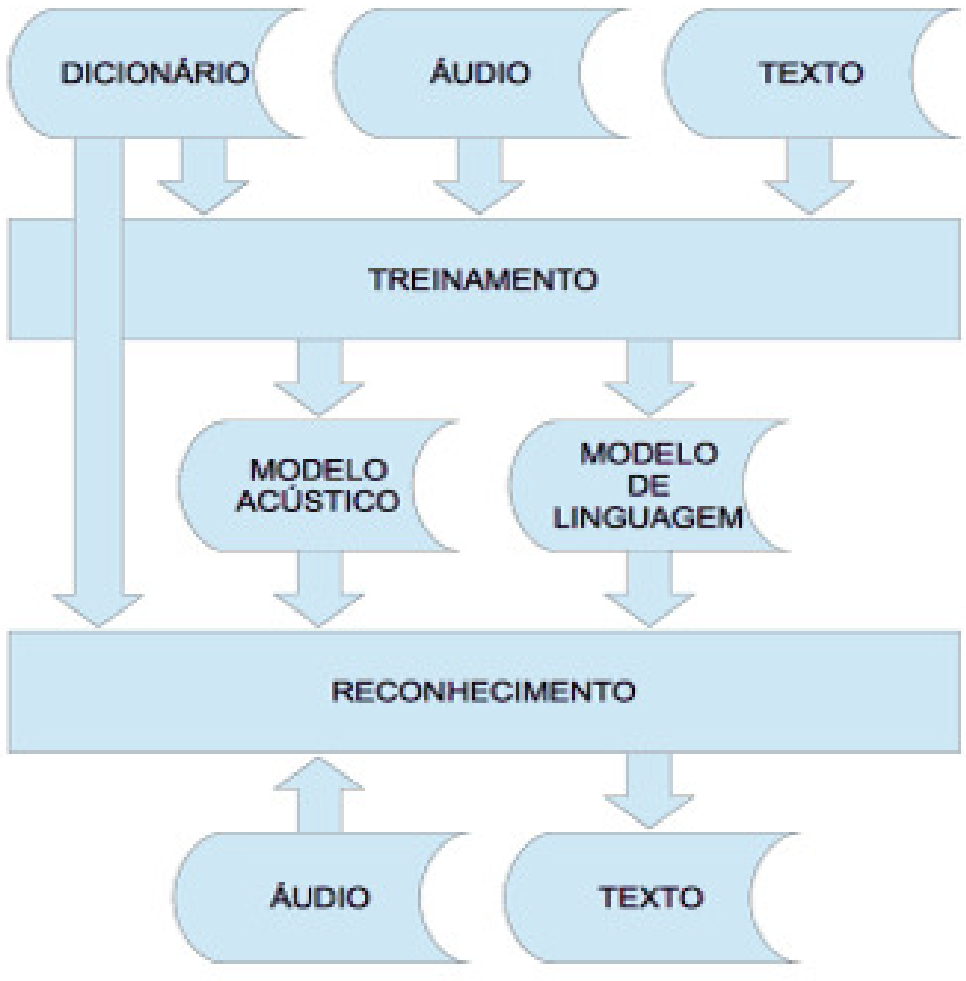

Figura 1. Reconhecimento Automático de Voz

A captação do sinal de voz é realizada através de um microfone e a digitalização consiste em um processo de amostragem no tempo e quantização da amplitude do sinal. O processamento digital de sinais empregado se destina a extrair do sinal de áudio correspondente à voz captada e digitalizada, parâmetros que possam ser úteis no processo de reconhecimento de padrões fonéticos. Os modelos estatísticos são treinados para cada idioma a partir de uma base de dados que deve conter: um dicionário com o vocabulário que o sistema deve utilizar contendo a transcrição fonética de cada palavra, um conjunto de gravações de fala com transcrição textual e um conjunto de textos.

Como palavras fora do dicionário não serão reconhecidas pelo sistema, é comum a necessidade de atualização frequente do dicionário, incorporando novas palavras, sobretudo nomes próprios. Contudo, deve-se destacar que quanto maior o vocabulário considerado pelo sistema, maior a probabilidade de confusão entre palavras. Portanto, para maximizar a acurácia do sistema, também pode ser interessante retirar do vocabulário palavras que não sejam mais utilizadas. Uma estratégia comum é extrair o vocabulário a partir da análise das palavras mais comuns do conjunto de textos. Dessa forma, também podem ser construídos dicionários especializados, segmentando o conjunto de textos por assunto. A transcrição fonética das palavras do dicionário pode ser realizada de forma automática, utilizando um conjunto de regras de conversão grafema-fonema. Eventualmente, uma mesma palavra no dicionário pode conter mais de uma transcrição fonética, o que é particularmente útil se o sistema precisar suportar variações regionais de pronúncia.

As transcrições textuais das gravações de fala são convertidas para uma sequência de fonemas de acordo com o dicionário, para que sejam identificados nos parâmetros extraídos do sinal de áudio, pela etapa de processamento digital de sinais, padrões correspondentes aos fonemas empregados no dicionário. A modelagem estatística desses padrões fonéticos é denominada

"Modelo Acústico". Diferentemente do dicionário, as gravações de fala utilizadas no treinamento não precisam conter todas as palavras do vocabulário do sistema, mas devem conter exemplos suficientes de todos os fonemas, preferencialmente em todos os contextos de vizinhança fonética possíveis (uma vez que cada fonema pode ter suas características alteradas de acordo com o antecessor e o sucessor). Os modelos acústicos, se treinados com a voz de diferentes locutores, são ditos independentes de locutor. Contudo, a acurácia é máxima quando o sistema pode ser treinado com a voz a ser reconhecida, produzindo modelos acústicos dependentes de locutor. Mas como, na prática, é difícil obter gravações suficientes para treinar adequadamente o modelo acústico, uma solução de compromisso comum é o que se denomina adaptação de locutor, em que há um treinamento independente de locutor, que é posteriormente adaptado para a voz de um locutor específico utilizando uma base de treinamento relativamente pequena. Se a base de treinamento independente de locutor for suficientemente grande e diversificada, também é possível segmentá-la, gerando modelos acústicos independentes de locutor, mas agrupados por sexo, idade, sotaque etc. A adaptação de locutor também pode ser feita a partir dos modelos acústicos desses agrupamentos.

O conjunto de textos é utilizado na modelagem da probabilidade de palavras e sequências de palavras, no que se denomina "Modelo de Linguagem". Para maximizar a acurácia, os textos devem possuir similaridade de vocabulário e estilo de linguagem com a fala a ser reconhecida. Para um sistema com uso previsto em situações com assuntos (e, portanto, vocabulários e estilos de linguagem) muito distintos, é comum segmentar a base de dados de textos por assunto, produzindo modelos de linguagem diferenciados, a serem selecionados em cada situação. Assim como no caso do dicionário, há necessidade de atualização frequente dos modelos de linguagem, para incorporação de novos assuntos e atualização do vocabulário.

Erros no processo de reconhecimento de voz podem ser atribuídos a problemas com a qualidade da captação e digitalização do sinal de voz, às limitações dos algoritmos de processamento de sinais e modelos estatísticos empregados, à restrição e/ou baixa qualidade da base de dados empregada no treinamento do sistema, bem como às diferenças entre a base de dados utilizada no treinamento e a fala a ser reconhecida (diferenças acústicas, de vocabulário, de linguagem etc.) [8].

5. USO AO REDOR DO MUNDO

Na aplicação de reconhecimento automático de voz para a geração de *closed caption*, normalmente é utilizado um relocutor, que ouve a fala espontânea através de um fone de ouvido e a repete em um estúdio [9] [10]. Com o uso da relocução, o modelo acústico do sistema pode ser adaptado à voz do relocutor, reduzindo a complexidade do reconhecimento de voz, por não precisar considerar a variabilidade fonética entre indivíduos. Além disso, como a relocução se dá em um ambiente acusticamente controlado, proporciona uma relação sinal/ruído melhor para o sistema de reconhecimento. Outra vantagem de utilizar a relocução é que o relocutor pode reformular a fala, corrigindo as disfluências comuns na linguagem oral, tornando-a mais adequada a uma transcrição para a linguagem escrita. Para facilitar o processo de reconhecimento, o relocutor pode ainda falar os sinais de pontuação que devem ser utilizados e acrescentar pausas entre as palavras. Para realizar a relocução não é necessário um curso de formação profissional específico, sendo uma mão-de-obra mais barata, fácil e rápida de repor do que um estenotipista.

No Japão, de acordo com Miyasaka [11], desde 2000 a emissora NHK utiliza um *software* de reconhecimento automático de voz para geração de legenda oculta ao vivo em telejornais, por ela desenvolvido para atender os requisitos técnicos desejados por essa emissora (acurácia mínima de 95% e latência máxima de 2 segundos), porque o *software* comercial disponível (IBM ViaVoice em japonês) não atendia a tais requisitos. O sistema utilizava modelos acústicos adaptados para cada apresentador de telejornal, sem utilizar relocução. Antes de cada telejornal, o modelo de linguagem era atualizado utilizando os rascunhos das matérias disponíveis. Embora o *software* desenvolvido atendesse o requisito de 95% de acurácia, para que não houvesse erro algum na legenda oculta, o sistema incluía quatro operadores: dois para revisar frases alternadas marcando as palavras erradas e, para cada revisor, um operador para corrigir as palavras erradas marcadas. De acordo com Homma [12], desde 2001 a NHK passou a empregar, sobretudo em emissoras afiliadas à sua rede, um sistema com apenas um ou dois operadores para realizar as correções, em que os as palavras erradas são selecionadas em um monitor sensível ao toque e corrigidas pelo mesmo operador. Com dois operadores, no reconhecimento automático de voz de telejornais com apenas um apresentador, o sistema chegou a obter 99,9% de acurácia e 99,8% com apenas um operador. A emissora passou também a utilizar reconhecimento automático de voz para geração de legenda oculta ao vivo em outros programas, além de telejornais, empregando relocução. Sako e Ariki [13] relatam em 2005 a implementação de um sistema de reconhecimento automático de voz para produzir *closed caption* para a transmissão de jogos de beisebol ao vivo. O sistema não emprega relocução, porém os modelos acústicos são adaptados para o locutor do jogo. Além disso, os modelos acústicos são segmentados de acordo com alguns estados emocionais pré-definidos. Assim, o sistema deve primeiro classificar o estado emocional da fala, para a seguir escolher qual modelo acústico empregar. Também foi incorporado ao modelo de linguagem uma memória de estado, com significado específico para o jogo (ex: o placar não pode regredir, mudanças no placar normalmente estão associadas a uma mudança no estado emocional da voz do locutor etc.). A melhoria de acurácia proporcionada pela análise de estado emocional e memória de estado do jogo foi de cerca de 1%.

Em Portugal, Meinedo *et alli* [14] relatam em 2003 o desenvolvimento de um *software* de reconhecimento automático de voz para telejornais. A acurácia obtida foi de 85%, com modelo acústico independente de locutor. Martins *et alli* [15] relatam em 2010 a implementação de um sistema que atualiza diariamente o dicionário e o modelo de linguagem, de forma automática, utilizando textos de *sites* de notícias. Com essa abordagem,

conseguiram reduzir em 88% os erros ocasionados pela ocorrência de palavras fora do vocabulário do sistema.

No Reino Unido, Evans [16] relata em 2003 o uso na BBC de um sistema de reconhecimento automático de voz desenvolvido internamente utilizando o *software* IBM ViaVoice e empregando relocução, para a geração de *closed caption* em programas ao vivo. Lambourne *et alli* [17] relatam em 2004 a adoção de um sistema semelhante em três outras grandes emissoras britânicas, também utilizando o IBM ViaVoice, empregando relocução e modelos de linguagem diferenciados de acordo com o assunto. A acurácia obtida foi de 95 a 98%.

No Canadá, de acordo com Boulianne *et alli* [18], desde 2004 utiliza-se regularmente na geração de legenda oculta ao vivo um *software* de reconhecimento automático de voz para a língua francesa, desenvolvido em conjunto por uma emissora e um instituto de pesquisa canadenses. O sistema utiliza relocução, com modelos acústicos adaptados a partir de modelos independentes de locutor agrupados por sexo. Utiliza ainda dicionários e modelos de linguagem diferenciados por assunto. Os dicionários, os modelos de linguagem e os modelos acústicos são atualizados semi-automaticamente. O sistema busca diariamente textos de sites de notícias correspondentes aos assuntos modelados no sistema, para atualizar os modelos de linguagem. Quando ocorrem palavras novas, o relocutor precisa validá-las (juntamente com suas pronúncias e associação aos assuntos) na pré-produção dos programas para que sejam inseridas nos dicionários. Palavras inativas por muito tempo também são removidas do vocabulário do sistema. Ainda na pré-produção, o relocutor faz uma breve gravação para atualizar seu modelo acústico. Durante a produção, o relocutor pode comutar o assunto, indicar a mudança da pessoa cuja fala está sendo transcrita e inserir pontuação e outros símbolos utilizando um *joystick*. Após a produção, o relocutor pode corrigir os erros do reconhecimento para que a transcrição corrigida seja utilizada na atualização dos modelos acústico e de linguagem. A acurácia obtida pelo sistema variou entre 89 e 93%. Cardinal *et alli* [19] relatam em 2007 o desenvolvimento de um sistema de correção em tempo real de legendas ocultas para ser usado em conjunto com um sistema de reconhecimento automático de voz. Além de permitir a digitação das palavras corrigidas, o sistema sugere, para cada palavra, uma lista de alternativas baseada nos erros de reconhecimento mais comuns observados.

Na Eslovênia, devido à falta de gravações de fala com transcrição textual em quantidade suficiente para um treinamento adequado dos modelos acústicos, Zgank [20] propôs em 2010 uma metodologia para incorporar gravações sem transcrição na base de dados, treinando o sistema inicialmente apenas com as gravações com transcrição disponíveis, utilizando o sistema para fazer uma transcrição automática das gravações sem transcrição, associando a essas transcrições um escore de confiança, e aproveitando em um novo treinamento do sistema as gravações com transcrição automática que obtiveram escores de confiança suficientemente altos. Pelos resultados experimentais que ele obteve, o acréscimo dessas gravações transcritas automaticamente na base de treinamento do sistema melhorou significativamente a acurácia do mesmo. A acurácia obtida foi de 82%, mas trata-se de um modelo acústico independente de locutor, e portanto, pode-se esperar obter uma acurácia significativamente mais alta empregando-se adaptação de locutor.

6. SITUAÇÃO BRASILEIRA

O único *software* de reconhecimento automático de voz comercial existente para o português do Brasil e empregado na geração de *closed caption* é o IBM ViaVoice. Este produto foi desenvolvido e comercializado pela IBM até 2003, quando foi vendido para a ScanSoft, posteriormente adquirida pela Nuance [21]. Nem a ScanSoft, nem a Nuance deram continuidade ou desenvolveram produtos para substituir o ViaVoice em Português do Brasil. Com isso, o *software* não dispõe mais de suporte corretivo ou evolutivo, podendo eventualmente deixar de funcionar com sistemas operacionais mais novos. Além disso, a quantidade de licenças existentes é limitada.

Note-se que o IBM ViaVoice nem era um produto otimizado para aplicação na geração de legenda oculta: tratava-se de um sistema de reconhecimento automático de voz para utilização como interface homem/máquina. Emissoras de televisão e outras empresas desenvolveram interfaces para a geração de legendas ocultas utilizado o ViaVoice. Como já citado, parte dos erros do reconhecimento de voz se devem às limitações dos algoritmos de processamento de sinais e modelos estatísticos empregados. Pode-se imaginar que hoje seria possível utilizar algoritmos e modelos mais precisos e complexos, uma vez que a tecnologia de reconhecimento de voz evoluiu nesse período e que a capacidade computacional aumentou desde a descontinuidade do ViaVoice. Além disso, não há necessidade de ficar restrito a computadores de uso doméstico, pode-se utilizar até mesmo um conjunto de servidores realizando processamento distribuído.

De fato, alguns dos sistemas de reconhecimento de voz mais bem sucedidos atualmente utilizam o paradigma de computação em nuvem (*cloud computing*), como, por exemplo, a pesquisa por voz do Google [22] e o Apple Siri [23]. Dessa forma, estão livres da limitação da capacidade computacional dos usuários (o que é especialmente útil para aplicações embarcadas em dispositivos móveis).

Apesar da falta de *softwares* comerciais que possam ser empregados no reconhecimento automático de voz do português do Brasil, existem diversas opções de *software* livre que podem ser consideradas: HTK [24] da Universidade de Cambridge, CMU Sphinx [25] da Universidade Carnegie Mellon, Julius [26] da Universidade de Kyoto, Simon [27] da Organização "Simon Listens", RWTH ASR [28] da Universidade Técnica de Aachen, iATROS [29] da Universidade Politécnica de Valência, SHoUT [30] da Universidade de Twente, ISIP ASR [31] da Universidade Estadual do Mississippi etc. Esses *softwares* implementam os algoritmos de processamento de sinais e modelos estatísticos utilizados no reconhecimento automático de voz, mas para serem aplicados ao português do Brasil necessitam de treinamento com bases de dados brasileiras. O treinamento desses sistemas utiliza ferramentas distribuídas juntamente com os mesmos, além de outras distribuídas separadamente, como o SRILM [32] da SRI International.

O treinamento é um aspecto crítico para o bom funcionamento de um sistema de reconhecimento

automático de voz. Sendo o português, ao contrário do inglês, uma língua altamente flexiva (em que as palavras podem sofrer flexão de gênero, número, modo, tempo e pessoa), o vocabulário a ser considerado aumenta substancialmente, o que tende a ter um impacto negativo na acurácia do reconhecimento. A gramática do português, novamente ao contrário do inglês, também é relativamente livre quanto à ordem das palavras, o que torna os modelos de linguagem mais complexos, o que também tende a ter um impacto negativo na acurácia do sistema. Essas duas diferenças entre os idiomas português e inglês apontam para a necessidade de bases de dados para o treinamento de um sistema de reconhecimento automático de voz em português maiores do que seria necessário em inglês, para obter-se a mesma acurácia. Entretanto, o que ainda se observa é a disponibilidade de bases de dados significativamente menores. Uma das iniciativas que buscam disponibilizar as bases de dados necessárias para o português do Brasil é a do site VoxForge [33]. Outra iniciativa importante de ser citada é a do grupo FalaBrasil [34], da Universidade Federal do Pará. Além de disponibilizar suas bases de dados, esse grupo já possui uma experiência considerável no treinamento de diversos *softwares* livres para o português do Brasil e na integração desses *softwares* para algumas aplicações específicas. Também pode-se considerar que o acervo das próprias emissoras de televisão interessadas em utilizar o sistema de reconhecimento automático de voz poderia contribuir significativamente para o desenvolvimento das bases de dados necessárias.

7. CONSIDERAÇÕES FINAIS

Uma opção para o desenvolvimento de alternativas viáveis à falta de ferramentas comerciais de reconhecimento automático de voz para o português do Brasil aplicáveis na geração de legenda oculta é realizar um levantamento das opções de *software* livre disponíveis. Em um levantamento criterioso de tais *softwares* pode-se avaliar as opções dos parâmetros de configuração disponíveis quanto aos impactos sobre a acurácia e o custo computacional dos sistemas, podendo ser adotadas opções tais que necessitem de processamento distribuído para viabilizar a execução em tempo real, se isso resultar em uma melhoria significativa da acurácia do sistema. É necessário ainda avaliar as bases de dados disponíveis publicamente para treinamento dos sistemas de reconhecimento automático de voz em português do Brasil e, eventualmente, aproveitar o acervo das emissoras para desenvolver e/ou incrementar tais bases de dados, além de melhor adequá-las à aplicação em questão. Por fim, algumas otimizações implementadas nesse tipo de sistema em emissoras de televisão ao redor do mundo (e.g. sistema de atualização automática ou semi-automática de dicionário, modelos de linguagem e modelos acústicos, sistema de correção manual em tempo real etc.) podem ser experimentadas no Brasil.

Algumas dificuldades que podem surgir na tentativa de desenvolver sistemas de reconhecimento automático de voz para o português do Brasil para a aplicação na geração de legenda oculta a partir de ferramentas disponíveis publicamente são as seguintes:

- A licença de ferramentas livres pode não permitir a aplicação desejada;
- O desempenho dos *softwares* livres a serem testados pode estar abaixo do desejado;
- Os *softwares* podem não possuir documentação suficiente;
- As bases de dados disponíveis publicamente podem ser insuficientes para um treinamento adequado do sistema.

A grande quantidade de *softwares* livres disponíveis para reconhecimento automático de voz e a reputação das instituições que os desenvolveram, sugere que sejam obtidas algumas opções viáveis. Adicionalmente, uma vez que os códigos-fonte são abertos, é possível desenvolver otimizações se necessário, ou mesmo, é possível estudar esses códigos como subsídio para o desenvolvimento de novos sistemas.

As bases de dados para treinamento dos sistemas serão provavelmente o principal limitador da acurácia que pode ser obtida. Porém, além da opção de utilizar o acervo das emissoras, pode-se considerar a realização de acordos com universidades para o desenvolvimento de bases públicas e a utilização de bases comerciais, pré-existentes ou desenvolvidas sob encomenda. A abordagem de aproveitar áudio sem transcrição na base de dados para treinamento dos modelos acústicos também pode ser útil.

REFERÊNCIAS

[1] INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2010:** Resultados gerais da amostra, Rio de Janeiro, 2012. Disponível em: <http:// www.ibge.gov.br/ home/ presidencia/ noticias/ imprensa/ ppts/ 00000008473104122012315727483985.pdf>. Acesso em: 12 maio 2012.

[2] CONSUMER ELECTRONICS ASSOCIATION. **ANSI/CEA-608-E: Line 21 Data Services**. Arlington, p. 127. 2008.

[3] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 15610-1: Televisão digital terrestre - Acessibilidade - Parte 1: Ferramentas de texto**. Rio de Janeiro, p. 23. 2011.

[4] BRASIL. Lei nº 10.098, de 19 de dezembro de 2000. Estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil, Poder Legislativo**, Brasília, DF, p. 2-3, Seção 1, 20 dez. 2000.

[5] BRASIL. MINISTÉRIO DAS COMUNICAÇÕES. **Portaria nº 310**, 27 jun. 2006. Disponível em: <http:// www.mc.gov.br/ images/ o-ministerio/ legislacao/ portarias/ portaria-310.pdf>. Acesso em: 12 maio 2012.

[6] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 15290: Acessibilidade em comunicação na televisão**. Rio de Janeiro, p. 10. 2005.

[7] SANT'ANA, T. Eu sou: Estenotipista. **Revista Galileu**, maio 2012. Disponível em: <http:// revistagalileu.globo.com/ Revista/ Common/ 0,EMI281040-17773,00-EU+SOU+ESTENOTIPISTA.html>. Acesso em: 12 maio 2012.

[8] HUANG, X.; ACERO, A.; HON, H.-W. **Spoken Language Processing:** A Guide to Theory, Algorithm, and System Development. Upper Saddle River: Prentice-Hall, 2001.

[9] AHMER, I. **Automatic Speech Recognition for Closed Captioning of Television: Data and Issues**. Thesis (Master of Engineering) - University of South Australia. Adelaide, p. 244. 2002.

[10] INTERNATIONAL TELECOMMUNICATION UNION. **Report ITU-R BT.2207-1: Accessibility to broadcasting services for persons with disabilities**. Geneva, p. 17. 2011.

[11] MIYASAKA, E. Development of information systems in Japan Broadcasting Corporation. **Journal of the Center for Information Studies**, n. 5, p. 52-57, 2004.

[12] HOMMA, S. et al. **New Real-Time Closed-Captioning System for Japanese Broadcast News Programs**. In: Proceedings of the 11th International Conference on Computers Helping People with Special Needs. Linz: Springer. 2008. p. 651-654.

[13] SAKO, A.; ARIKI, Y. **Structuring baseball live games based on speech recognition using task dependent knowledge and emotion state recognition**. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. Philadelphia: IEEE. 2005. p. 1049-1052.

[14] MEINEDO, H. et al. **AUDIMUS.media- A Broadcast News Speech Recognition System for the European Portuguese Language**. In: PROPOR 2003: Proceedings of the 6th International Workshop on Computational Processing of the Portuguese Language. Faro: Springer. 2003. p. 9-17.

[15] MARTINS, C.; TEIXEIRA, A.; NETO, J. Dynamic language modeling for European Portuguese. **Computer Speech and Language**, v. 24, n. 4, p. 750-773, out. 2010.

[16] EVANS, M. J. BBC R&D White Paper 065. **Speech Recognition in Assisted and Live Subtitling for Television**, 2003. Disponível em: <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP065.pdf>. Acesso em: 18 maio 2012.

[17] LAMBOURNE, A. et al. Speech-Based Real-Time Subtitling Services. **International Journal of Speech Technology**, v. 7, n. 4, p. 269-279, out. 2004.

[18] BOULIANNE, G. et al. **Computer-Assisted Closed-Captioning of Live TV Broadcasts in French**. In: Interspeech 2006: Proceedings of the International Conference of Spoken Language Processing. Pittsburgh: ISCA. 2006. p. 273-276.

[19] CARDINAL, P. et al. **Real-Time Correction of Closed-Captions**. In: Proceedings of the ACL 2007 Demo and Poster Sessions. Prague: ACL. 2007. p. 113-116.

[20] ZGANK, A. Three-Stage Framework for Unsupervised Acoustic Modeling Using Untranscribed Spoken Content. **ETRI Journal**, v. 32, n. 5, p. 810-818, out. 2010.

[21] IBM. **IBM Desktop ViaVoice**. Disponível em: <http:// www-01.ibm.com/ software/ pervasive/ viavoice.html>. Acesso em: 13 maio 2012.

[22] GOOGLE. **Introducing Voice Search:** Now on your computer. Disponível em: <http:// www.google.com/ insidesearch/ features/ voicesearch/ index.html>. Acesso em: 13 maio 2012.

[23] APPLE. **Learn more about Siri**. Disponível em: <http:// www.apple.com/ iphone/ features/ siri-faq.html>. Acesso em: 13 maio 2012.

[24] CAMBRIDGE UNIVERSITY. **HTK**. Disponível em: <http:// htk.eng.cam.ac.uk>. Acesso em: 14 maio 2012.

[25] CARNEGIE MELLON UNIVERSITY. **CMU Sphinx**. Disponível em: <http:// cmusphinx.sourceforge.net>. Acesso em: 14 maio 2012.

[26] KYOTO UNIVERSITY. **Julius**. Disponível em: <http:// julius.sourceforge.jp/ en_index.php>. Acesso em: 14 maio 2012.

[27] SIMON LISTENS. **Simon**. Disponível em: <http:// simon-listens.org/ index.php?id=122&L=1>. Acesso em: 14 maio 2012.

[28] RWTH AACHEN UNIVERSITY. **RWTH ASR**. Disponível em: <http:// www-i6.informatik.rwth-aachen.de/ rwth-asr>. Acesso em: 14 maio 2012.

[29] UNIVERSIDAD POLITÉCNICA DE VALENCIA. **iATROS**. Disponível em: <https:// prhlt.iti.upv.es/ page/ projects/ multimodal/ idoc/ iatros>. Acesso em: 14 maio 2012.

[30] UNIVERSITY OF TWENTE. **SHoUT**. Disponível em: <http:// shout-toolkit.sourceforge.net>. Acesso em: 14 maio 2012.

[31] MISSISSIPPI STATE UNIVERSITY. **ISIP ASR**. Disponível em: <http:// www.isip.piconepress.com/ projects/ speech/ index.html>. Acesso em: 14 maio 2012.

[32] SRI INTERNATIONAL. **SRILM**. Disponível em: <http:// www.speech.sri.com/ projects/ srilm>. Acesso em: 14 maio 2012.

[33] VOXFORGE. **VoxForge**. Disponível em: <http:// www.voxforge.org/ pt_br>. Acesso em: 14 maio 2012.

[34] UNIVERSIDADE FEDERAL DO PARÁ. **FalaBrasil**. Disponível em: <http:// www.laps.ufpa.br/ falabrasil>. Acesso em: 14 maio 2012.

Surround Sound 5.1 for one-seg TV and digital radio: the MPEG Surround standard for the Brazilian broadcasting services

Luís Meloni (meloni@decom.fee.unicamp.br)
Rafael Diniz (rafael@riseup.net)

Surround Sound 5.1 for one-seg TV and digital radio: the MPEG Surround standard for the Brazilian broadcasting services

Luis Geraldo P. Meloni [1] e *Rafael Diniz* [2]

[1] UNICAMP, Campinas, Brazil, meloni@decom.fee.unicamp.br
[2] Eldorado Research Institute, Campinas, Brazil, rafael@riseup.net

Abstract: The paper presents a brief introduction to the MPEG Surround technology and how it can be used for the one-seg profile of the Brazilian Digital TV and also for digital radio in order to provide spatial surround sound for the broadcasting services.

In the paper, some important characteristics of the MPEG Surround standard are covered, as well as the real world scenario of the MPEG Surround used by broadcasters.

An amendment to the ABNT/NBR standard that would allow the use of the MPEG Surround in Brazil is proposed in the paper.

Key words: Digital TV, one-seg, surround audio, Digital Radio.

1. Introduction

With the launch of Brazilian ISDB-Tb and the upcoming Brazilian Digital Radio System the introduction of a new kind of media was made available for the Brazilian Broadcasters: the 5.1 Surround Sound.

However, the ABNT NBR 15602-2 standard only allows the transmission of more than two channels in the full-seg fixed-reception service. In the case of the one-seg mobile-reception service, no more than two channels can be transmitted. The digital radio standards evaluated in Brazil allows surround transmissions.

The assumption that surround sound has no application in the portable and mobile services is not true as there is an increasing demand for high-end 5.1 surround sound configuration for cars and also for binaural rendering capable [1] mobile players.

The purpose of this paper is to propose and to emphasize the importance for inclusion of the MPEG Surround (MPS) [2] standard inside the body of the ABNT/NBR 15602-2 regarding the one-seg service. The paper also explains how the technology can be used including compatibility issues with legacy devices that do not support MPEG Surround.

2. The MPEG Surround technology

2.1 MPEG Surround ISO/IEC 23003-1 standard

The MPEG Surround standard is described in ISO/IEC 23003-1, approved in early 2007.

The MPEG Surround standardizes the parametric multichannel extension technology and can be used in combination with any mono or stereophonic audio codec (see Figure 1), like MPEG-1/2 Layer II and III, and MPEG-4 AAC LC and HE-AAC. However in this text the focus is on the use of MPEG Surround together with the MPEG-4 HE-AAC codec, which is the codec used by the Brazilian ISDB-Tb. It is important to mention that this technology is backward compatible, i.e., it can be used to upgrade the existing distribution infrastructures while retaining full compatibility with existing receivers.

MPEG Surround is based on a principle called spatial audio coding. Spatial audio coding is a multi-channel compression technique incorporating a backwards compatible downmix, and exploiting the perceptual inter-channel redundancy in multi-channel audio signals to achieve higher compression rates.

In this way an ISDB-Tb one-seg transmission using the MPEG Surround technology on top of a stereo HE-AAC transmission will be decoded as a normal stereo program by a legacy device (the legacy device just ignores the unknown MPEG Surround information), and as a multi-channel program by a MPEG Surround capable receiver.

2.2 Bitrate Range and Performance

Extensive listening tests were conducted by 3GPP on MPEG Surround using movie, music, sports and radio drama material.

The majority of test sites found that the overall quality of the MPEG Surround at 96 kbps is in the "Excellent" range, in a MOS (Mean Opinion Score) test with five categories (excellent, good, fair, poor or bad).

At 64 kbps the scores are in the "Good" range for all test sites. Even for 48 kbps the majority of the conducted tests resulted in perceived quality in the "Good" range [3]. Furthermore, the quality of MPEG

Surround binaural decoding for headphone playback was tested. Specifically the benefit of adding MPEG Surround data to the stereo bitstream for binaural rendering was evaluated. The overall test results show significant improvement of MPEG Surround over stereo for the same overall bitrate of 64 kbps.

Concerning the possible bitrates for the audio program in the one-seg service, in which not much more than 400 kbit/s or 300 kbit/s (depending on the robustness of the transmission) should be shared between audio and video, the MPEG Surround bitrate range clearly allows the use of this technology in a one-seg service. Also for digital radio systems, for example the DRM (Digital Radio Mondiale), one of the standards which is been considered as candidate for adoption in Brazil, it allows bitrates ranging from 40 kbps up to 180 kbps.

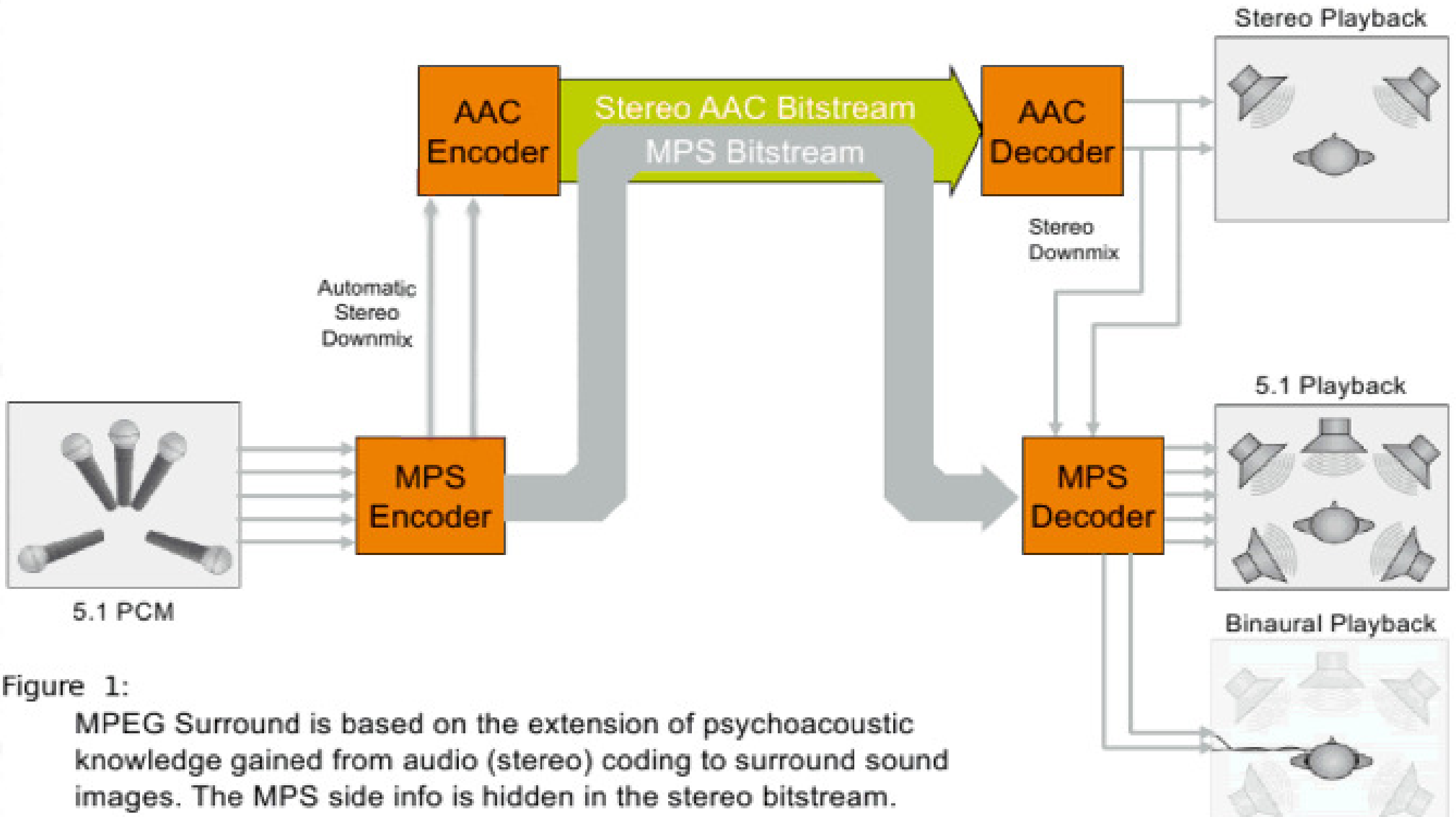

Figure 1:
MPEG Surround is based on the extension of psychoacoustic knowledge gained from audio (stereo) coding to surround sound images. The MPS side info is hidden in the stereo bitstream.

3. MPEG Surround use for Broadcasting

3.1. Radio Broadcasting

Digital radio is one of the key targets for MPEG Surround as the bitrate that a digital radio station can use is very limited, so MPEG Surround for multi-channel radio transmission is a natural choice.

The Digital Radio standard in the US is the Ibiquity HD Radio, and the MPEG Surround can be used in combination with the HDC codec. In Japan the new ISDB-Tsb Digital Radio System also allows the use of the MPEG Surround technology, as it does the DAB, DAB+ and DRM Digital Radio Systems.

The first radio station that officially started broadcasting 5.1 content using the MPEG Surround technology for a trial during NAB was the CBS WZLX of Boston, US in April 17, 2007[4].

After that, in October 2007 the German broadcaster "Rock Antenne" [5] conducted a MPEG Surround trial over the DAB (Digital Audio Broadcast) digital radio standard during the IFA consumer electronics trade show.

In this case, the MPEG Surround technology is used in combination with the MPEG-2 Layer II codec.

In July 2009, the first public Surround broadcast over a DRM+ (Digital Radio Mondiale) transmission was conducted in Paris by the *Syndicat Nacional des Radios Libres* [6]. The MPEG Surround technology in this case is used in combination with the MPEG4 HE-AAC.

In Brazil, trials with both HD Radio and DRM Digital Radio systems are ongoing and one of the two standards may be elected as the basis of the national standard [7].

Both digital radio standards being evaluated by Brazil allows the use of the MPEG Surround technology.

3.2. Mobile TV Broadcasting

In Japan the new ISDB-Tmm Multimedia Broadcast System for Mobile-TV already allows the use of the MPEG Surround.

MPEG Surround may also be used to establish multi-channel services for Mobile-TV broadcasts such as the European DVB-H.

3.3. Integration with already existing multichannel Digital TV workflow

After almost five years of the first digital TV transmission in Brazil, all the major broadcasters have 5.1 multichannel programs on air, ranging from music programs to football matches. The integration of a one-seg surround encoder in the transmission chain will be as easy as disconnecting the old one-seg encoder (or just upgrading its firmware by a MPEG Surround capable one) and configuring the input SDI to carry the desired 5.1 audio channels.

The currently stereo-only one-seg service can be considered as a technology that misuses the content already produced in 5.1 multichannel audio by the broadcasters.

3.4. ABNT/NBR 15602-2 amendment for MPEG Surround conformance

The standardization of MPEG Surround for Brazilian ISDB-Tb depends on approval by the SBTVD Forum [8].

The current proposal makes as minor modifications as possible in the Brazilian standard in order to allow the introduction of the MPEG Surround technology for one-seg service.

There is also the possibility to allow the use of the MPEG Surround technology for the full-seg service, but as the full-seg service is not as bitrate constrained as the one-seg service, the importance of the MPEG Surround technology in the full-seg is reduced, however it would be straightforward to make the inclusion of the MPEG Surround for the full-seg service.

Considering the English version of the standard ABNT/NBR 15602-2 [9], the suggested modifications for MPEG-Surround allowance are described below:

Page 1, Section 2,

After:
"ISO/IEC 14496-3:2005, Information technology -- Coding of audio-visual objects - Part 3: Audio"

Add:
"ISO/IEC 23003-1:2007, Information technology -- MPEG audio technologies -- Part 1: MPEG Surround"

Page 2, Section 4,

After:
"LOAS Low Overhead Audio Stream"

Add:
"MPS MPEG Surround"

Page 7, Section 8.2,

After:
"The profile and level of the MPEG-4 AAC coder shall be adequately signalized according to ABNT NBR 15602-3 e ABNT NBR 15603-2."

Add:
"In addition to the MPEG AAC coder, the MPEG Surround coder is allowed to be use, according to the ISO/IEC 23003-1.

When the MPEG Surround coder is used, the Baseline MPEG Surround Profile shall be used, and only levels 2 and 3 are allowed.

The MPEG Surround use is only allowed on top of the MPEG-4 AAC HE version 1 or 2, Level 2 only."

Page 8, Section 8.3,

After:
"The receivers shall be capable of processing the SBR tool.
The SBR presence signaling shall be explicit using the non-backward compatible explicit signalization mechanism, according to ISO/IEC 14496-3."

Add:
"When the MPEG Surround coder is used, the MPEG Surround payload shall use Implicit Backwards Compatible signaling, according to ISO/IEC 23003-1."

Page 10, Section 9.2.1,

Inside Table 6, below "Restriction", Substitute:
"Monoaural (1/0), stereo (2/0)"
by:
"Monoaural (1/0), stereo (2/0), multichannel (3/2) + LFE *"
"* Multichannel is possible when using the MPEG Surround coder."

Substitute:
"The audio decoder shall be capable to process any of the recommended audio modes."
by:
"The audio decoder shall be capable to process the monoaural and stereo audio modes. Multichannel decoding is recommended."

Inside Table 7, after the last line, add (the lines of the table are comma separated) :

"Multichannel 5.1 | 6 | <SCE1><CPE1><CPE2> | SCE1 = C,
| | | CPE1 = L and R,
(3/2 + LFE) | | <LFE><TERM> | CPE2 = LS and RS"

Page 11, Section 9.2.2,

In Table 8, Substitute the corresponding lines (second and third) by:

Allowed profiles and levels | High-efficiency (HE): level 2 (HE-AAC v2@L2), Baseline MPEG Surround Profile: level 2 or 3.

Maximum number of encoded channels | 2 channels per bitstream (stereo or 2 monaural channels) | 6 channels per bitstream (5.1) when MPS is used.

Substitute:

"The signals may be decoded with any rate and any sample rate supported by the profile and level of Table 8."

by:

"The signals may be decoded with any rate and any sample rate supported by the profile and level of Table 8, with exception to the MPS extension, where is decoding is recommended but not obligatory."

4. Conclusion

The MPEG Surround technology and standards are ready to be used in the Brazilian one-seg service allowing great quality 5.1 surround sound. It is also of great importance its introduction in the future Brazilian digital radio broadcasting.

We hope this paper gives motivations for the SBTVD Forum towards the inclusion of the MPEG Surround technology in the body of the ABNT NBR 15602-2 standard, aiming to improve the ISDB-Tb, offering users the high quality of present days audio technologies.

REFERENCES

[1] Jeroen Breebaart, Jurgen Herre, Lars Villemoes, Craig Jin, Kristofer Kjorling, Jan Plogsties and Jeroen Koppens, Multi-Channel goes mobile: Mpeg Surround Binaural Rendering, presented at 29th AES International Conference, Seoul, Korea, 2006 September 2-4.

[2] ISO/IEC 23003-1 – MPEG-D (MPEG audio technologies), Part 1: MPEG Surround, 2007.

[3] 3GPP Specification TR 26.950, Study on Surround Sound codec extension for Packet Switched Streaming (PSS) and Multimedia Broadcast/Multicast Service (MBMS), 2010.

[4] Trial broadcast by CBS WZLX of Boston during NAB, USA, First 5.1 radio station broadcast using MPEG Surround, *http://www.businesswire.com/portal/site/google/index.jsp?ndmViewId=newsview&newsId=20070417005711&newsLang=en* April 2007.

[5] Harald Fuchs, Olaf Korte and Johannes Hilpert of Fraunhofer IIS, Digital Broadcasting with MPEG Surround, 2009.

[6] DRM Consortium, DRM+ in Band I successfully tested in Paris, *http://www.drm.org/old/index.php?id=352* July 2009.

[7] Ministry of Communications, text that contains the information of the standards being considered for the Brazilian Digital Radio, *http://www.mc.gov.br/noticias-do-site/21137-belo-horizonte-inicia-testes-de-transmissao-de-radio-digital* 03/02/2010.

[8] SBTVD Forum is the Brazilian Digital TV standards maintainer, *http://www.forumsbtvd.org.br*.

[9] ABNT/NBR 15602-2, Digital terrestrial television – Video coding, audio coding and multiplexing – Part 2: Audio coding, *http://www.dtv.org.br/download/en-en/ABNTNBR15602 2D2 2007Ing 2008.pdf* 2008.

[10] The Fraunhofer Institute for Integrated Circuits IIS, *http://www.iis.fraunhofer.de/*.

Arquitetura e Aplicações no contexto do *Mining Knowledge* TV

José Carlos Patricio Junior (jcapjunior@gmail.com)
Natasha Correia Queiroz Lino (natasha@ci.ufpb.br)
Ramon Santos (ramonob13@gmail.com)

Arquitetura e Aplicações no contexto do *Mining Knowledge TV*

José Carlos Almeida Patrício Júnior[1], Natasha Correia Queiroz Lino[2], Ramon Nóbrega dos Santos[3]
[1]Universidade Federal da Paraíba, João Pessoa – Paraíba – Brasil, jcapjunior@di.ufpb.br
[2]Universidade Federal da Paraíba, João Pessoa – Paraíba – Brasil, natasha@ci.ufpb.br
[3]Universidade Federal da Paraíba, João Pessoa – Paraíba – Brasil, ramon.nobrega@lavid.ufpb.br

RESUMO

Este artigo apresenta o *Mining Knowledge TV* - MKTV, abordando sua arquitetura e uma aplicação no contexto do projeto. *O MKTV é* um ambiente para mineração de dados inserido no contexto da plataforma *Knowledge TV* - KTV. Esta plataforma propõe a criação de uma camada semântica a ser incorporada no ambiente de TV Digital, melhorando entre outros, a forma de acessar os dados pelos desenvolvedores de aplicações de TV Digital Interativa.

PALAVRAS CHAVES

Mineração de Dados, TV Digital Interativa (TVDI), Personalização em TVDI, MKTV, KTV.

INTRODUÇÃO

A Televisão Digital Interativa [1] [2] é mais uma nova fase vivida pela TV, estágio esse que prima pela convergência de tecnologias digitais, através da sistêmica substituição de equipamentos analógicos por digitais, produzindo grandes mudanças em toda a cadeia produtiva e principalmente no consumo de mídias.

A visão da Web Semântica [3][5] tem sido amplamente difundida e apoiada por um grande número de iniciativas tanto da indústria como da academia. A Web Semântica proporciona a compreensão e o gerenciamento dos conteúdos publicados na internet tornando-os entendíveis.

A Mineração de Dados [6] é uma área multidisciplinar, que envolve as áreas de Banco de Dados (BD), Inteligência Artificial (IA), Estatística, etc. Ela tem como principais objetivos encontrar relacionamentos entre dados e fornecer alternativas para que se possam tomar decisões de forma eficiente.

Devido seu poder computacional a mineração de dados vem sendo trabalhada no contexto da TV Digital, não como parte de um contexto maior de um processo de *KDD*, mas utilizada através da aplicação de algoritmos estatísticos e com técnicas de IA.

Os Principais sistemas de TV Digital existentes, são: o ATSC - sistema Americano [17], DVB - sistema europeu [18], ISDB - sistema Japonês [19] e o mais recente SBTVD - Sistema brasileiro [20] que tem como *middleware* o Ginga [8] e será foco desse estudo.

O projeto *Knowledge TV* [4][7] propõe ser a camada semântica da TV Digital Brasileira incorporada ao *middleware* Ginga, e tem a finalidade de fornecer uma rica base de conhecimento contendo descrição de dados, recursos, serviços, aplicações, e relacionamentos entre estes, etc. Cada uma destas descrições é feita em uma linguagem padronizada formalmente, capaz de ser processada automaticamente por computadores.

Neste contexto, o presente artigo tem como objetivo apresentar a implementação do módulo de mineração de dados para a camada *Knowledge TV*, enfocando a integração entre a mineração de dados e os conceitos semânticos na TV Digital. Este módulo é chamado *Mining Knowledge TV*.

1. *KNOWLEDGE TV*

A plataforma computacional que dá suporte à TV Digital tem crescido em ritmo acelerado. Pesquisadores e desenvolvedores concentram-se tanto em questões de *hardware* e *software*, quanto na comunicação entre estes dois aspectos, que é solucionado com o desenvolvimento de uma camada intermediária entre ambos, denominada *middleware*. Um aspecto a se destacar é a forte ligação entre as aplicações e questões de *hardware*, inviabilizando ao desenvolvedor a utilização de abordagens mais abstratas para o desenvolvimento do *software*. Por razões como esta, existe uma necessidade de melhorias a serem desenvolvidas nesta área.

O projeto *Knowledge TV* - Uma Camada Semântica para Prover Serviços em TV Digital está inserido neste contexto e, para buscar soluções para estas falhas ainda presentes no modelo, procuram-se soluções baseadas na semântica das informações. Trata-se de uma camada semântica entre as camadas de aplicações e o *middleware* em uma plataforma para TV Digital.

Esta camada semântica avançará o estado da arte desta tecnologia por fornecer uma rica camada de descrição dos recursos e serviços obtidos por uma abordagem de modelagem semântica. Com isso, muitas questões complexas do *hardware* passam a ser abstratas aos desenvolvedores das aplicações, que fazem uso dos recursos e serviços da TV Digital.

2. MINING KNOWLEDGE TV - MKTV

O objetivo geral desse trabalho é a implementação de um ambiente de *KDD* (do inglês *Knowledge Discovery in Databases*), com ênfase na mineração de dados das informações provenientes da plataforma *Knowledge TV* (KTV). Este ambiente proporciona aos desenvolvedores de aplicação para TV Digital Interativa, que utilizam o *middleware* Ginga, padrão do SBTVD; dados não conhecidos, mas de grande relevância para solucionar problemas como sobrecarga de informação, personalização, propaganda direcionada, entre outros.

Os dados que serão minerados virão dos metadados contidos nas tabelas de informação de serviços (*Service Information* - SI) do padrão MPEG-2 usado no *middleware* Ginga, que são utilizados para representar as informações advindas dos programas de TV, serviços e interação multimídia existentes em toda a plataforma.

Uma das principais tabelas existente é a EIT, que dentre outras informações contêm o gênero do programa, subgênero, a data e hora, e a sinopse do programa, entre outras. Além de possuir estruturas descritoras que servem para introduzir novas informações.

Outra relevante tabela é a SDT que contém dados dos provedores de serviços, popularmente conhecidos como emissoras de TV.

Outra fonte a ser explorada é o comportamento do usuário, ou seja, qual canal assistiu, em que horário, e quantos minutos ele consumiu a programação fornecida através da televisão.

O processo de mineração de dados nessas inúmeras fontes trará informações novas que serão enriquecidas semanticamente por meio de uma ontologia, a *OntoMKTV*, também desenvolvida no contexto do projeto. A mesma torna possível uma melhor análise, explicitando ainda mais o significado presente nesse conjunto de novos dados que foi resultado da mineração de dados. Esses dados novos são então fornecidos como um serviço, que poderá ser utilizado principalmente para desenvolvedores de aplicativos NCL ou Java para TVDI, tornando mais ágil o desenvolvimento de inúmeras soluções.

2.1 Arquitetura

O *Mining Knowledge TV* (MKTV) é o componente da arquitetura da KTV responsável especificamente pelo tratamento e busca de conhecimento útil dos dados advindos de ambiente de TV digital interativa, incluindo o comportamento do usuário e também de outras fontes como a Web, caso necessário. Esses dados são armazenados inicialmente em um banco de dados relacional local, e progressivamente, é iniciado o processo de extração, transformação e carga (ETL – do inglês *Extraction, Transformation and Load*) das informações.

Após o processo de ETL os dados seguem para o módulo seguinte que é o de *Data Warehouse* (DW) [6].Esta técnica que é geralmente utilizada em conjunto com a Mineração de Dados [6], organiza os dados em *Data Marts* de acordo com o assunto a ser minerado, ex: personalização, *marketing*, business, concentrando dados históricos e integrados.

Com os dados historicamente organizados no DW, o módulo de Mineração de Dados aplica algoritmos, buscando e descobrindo padrões úteis e não conhecidos nos dados existentes do DW. De posse dos conhecimentos extraídos o MKTV os encapsula através de arquivos semânticos OWL (*Ontology Web Language*), que é o padrão de comunicação entre os módulos, e transporta os dados para os módulos integrados de modelagem semântica e base de conhecimento, responsáveis pelo conteúdo semântico no KTV. A Figura 1 ilustra a arquitetura do KTV com o módulo MKTV.

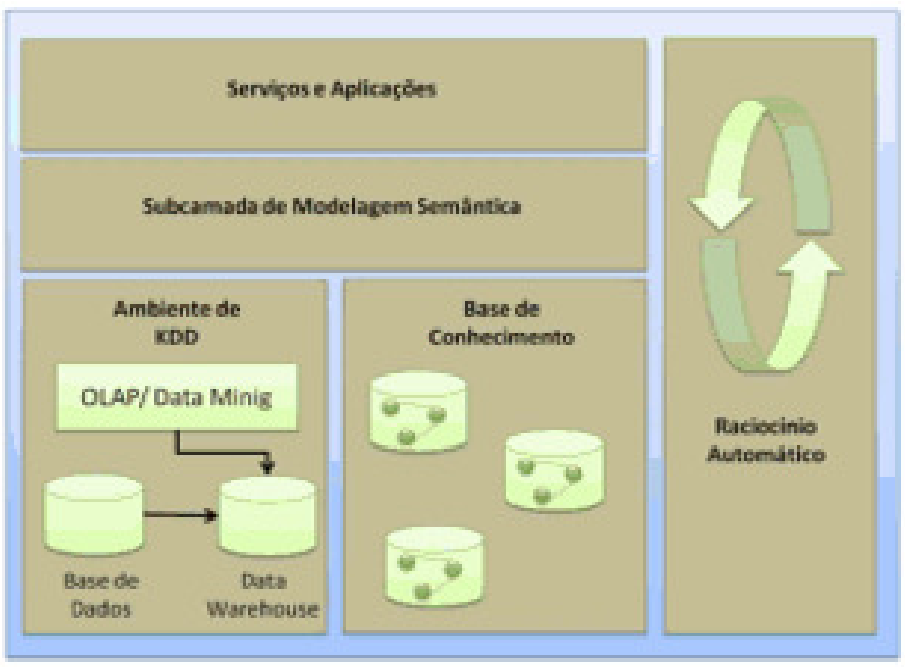

Figura 1 – Arquitetura do Módulo *Mining Knowledge TV* na KTV

O fluxo das informações no MKTV é detalhado na Figura 2. Os dados referentes às visualizações do usuário, programas, etc., são capturados pelo Componente de Monitoração de Interação do Usuário (A). Em seguida é enviado por ele, via canal de retorno, para o módulo Servidor do KTV(B). Os dados são recepcionados pelo componente de obtenção de

dados e armazenados em um banco de dados operacional local (C). Depois de armazenados, os esquemas de dados relacionais são descritos no componente de modelagem semântica e armazenados semanticamente no componente base de conhecimento (D). Em seguida as informações sofrem progressivamente o processo de extração, transformação e carga (ETL) das informações (E) para o de Data Warehouse – DW (F), onde os dados estão modelados multidimensionalmente em cubos de dados de acordo com o contexto e o propósito que são analisados.

Com os dados organizados no DW, o módulo de Mineração de Dados deverá aplicar algoritmos, buscando e descobrindo padrões úteis e não conhecidos nas informações existentes do DW (G) de acordo com o domínio do problema que se quer resolver. Por exemplo, caso o domínio seja entrega de propaganda personalizada, deve-se escolher a tarefa de Mineração de Dados mais adequada para solucionar o problema em questão, bem como outras primitivas de Mineração de Dados como algoritmo, limites de suporte e confiança, etc.

De posse dos conhecimentos extraídos, o MKTV os envia para a camada de modelagem semântica (I), onde ficam os modelos (ontologias) do KTV, especificados através de ontologias OWL e para a base de conhecimento (H), onde ficam as instâncias de acordo com os modelos (ontologias).

O componente de serviços e aplicações terá diversas aplicações, entre elas, uma de recomendação de programação ao usuário de TVDI. Essa aplicação pode se beneficiar do conhecimento descoberto, solicitando informações através de uma API diretamente à base de conhecimento (J) (K).

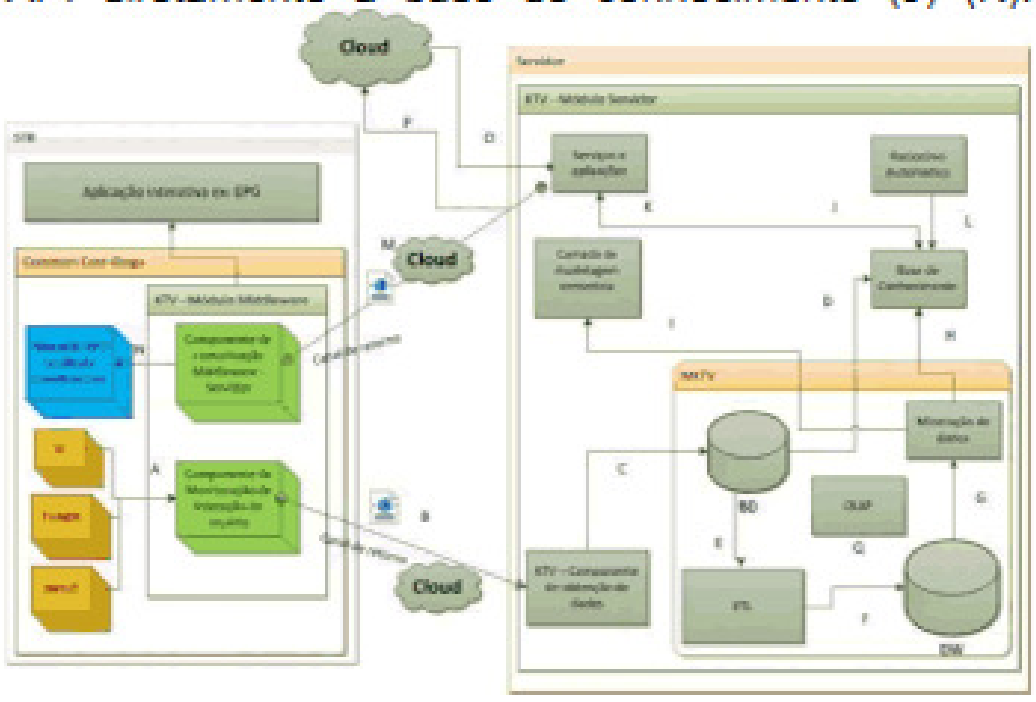

Figura 2 - Fluxo de dados MKTV

Outra possibilidade é solicitar algum serviço do MKTV via protocolo HTTP pela Web, ou seja, algum programa que deseje utilizar de serviços de mineração de dados poderá, via serviços Web, utilizar a API do KTV.

Para o componente OLAP (Q), não será do escopo desse trabalho a implementação de API ou ferramenta que o execute, ficando como trabalho futuro a implementação de recursos OLAP. Apesar de esse componente ser contemplado por meio do modelo multidimensional do DW, que possui em suas análises suporte a consultas OLAP, mesmo não tendo uma ferramenta específica para executá-las.

Os componentes de maior importância na arquitetura do MKTV são o *Data Warehouse*, a ontologia *OntoMKTV* e o de Mineração de Dados.

2.2 Data Warehouse – DW

O projeto KTV tem em seu escopo a criação de um *Data Warehouse* para dar suporte a serviços de recomendação, consulta semântica, Mineração de Dados, entre outros. O esquema usado para modelar o DW foi o esquema estrela [12], que consiste de uma tabela para cada dimensão e uma tabela de fatos relacionada com as mesmas.

Os conceitos utilizados para modelar o domínio de TVD no DW do MKTV vieram do modelo ontológico *CoreKTV* [13], que tomou como base as normas internacionais de telecomunicações J200 [14], J201 [15], J202 [16]. Desta forma, o modelo núcleo multidimensional do MKTV é composto pelas seguintes dimensões:

- Usuário – descreve informações relativas ao usuário, como localidade que reside, além de sexo e idade do mesmo;
- Programa – descrimina os dados acerca do conteúdo exibido pelas emissoras, ou seja, os programas de TV;
- Emissora – descreve informações sobre emissora de TV;
- Tempo – dimensão que descreve as informações sobre o tempo em que foi realizada determinada transação no modelo, por exemplo, meses, anos, data;
- Hora – determina os dados acerca de hora, como período do dia: manhã, tarde, noite e madrugada.

O DW do MKTV foi modelado para dar suporte a uma série de análises, além de auxiliar os serviços de consulta semântica e de recomendação de conteúdo, serviços estes sendo oferecidos inicialmente pela KTV. As principais consultas (análises) multidimensionais/OLAP suportadas através do modelo multidimensional núcleo do DW do MKTV são:

- Qual o canal com maior índice de audiência em determinada faixa de horário?
- Qual o programa de maior audiência por horário em determinado canal?
- Qual a classificação etária do programa de maior audiência por horário?
- Qual a duração do programa de maior audiência por horário?
- Qual o gênero de programa de maior frequência na programação semanal?
- Quantas aparições semanais determinado programa teve na grade de programação?
- Quais os turnos do dia (manhã, tarde, noite, madrugada) que um programa de classificação maior de 14 anos passa?
- Qual a audiência de programas com a classificação maior de 12,14,16,18, e livre por emissora?

2.3 *ONTOMKTV*

A *OntoMKTV* esta focada na representação do resultado da Mineração de Dados realizada pelos algoritmos da tarefa de regras de associação, feita nos dados do *Data Warehouse* provenientes do ambiente de convergência construído na *KTV* (TV Digital e WEB).

As principais vantagens da utilização de uma ontologia para o resultado da Mineração de Dados dizem respeito a: análise das informações mineradas, identificando similaridades e conceitos comuns; reusabilidade do conhecimento encontrado através da mineração de dados aplicável em outros domínios, como o de consultas semânticas, personalização de conteúdo; possibilidade de extensão do conhecimento descoberto através da Mineração de Dados, agregando novas informações; realizar inferências e suposições acerca das informações mineradas.

A *OntoMKTV* encapsula o conhecimento gerado pela tarefa de Mineração de Dados de regras de associação, isto é, para os algoritmos de regras de associação, podendo em um novo trabalho ser estendida para outras tarefas como a de clusterização, classificação.

2.4 Mineração de Dados no MKTV

O processo de Mineração de Dados no projeto MKTV se dará inicialmente nas informações de programa que são: nome, gênero, subgênero, faixa etária e a sinopse, vindas dos metadados contidos nas tabelas SI existentes no Middleware da TV Digital, juntamente com as informações de interação do usuário na TVD obtidas pelo ACIU (horário, o dia da semana, o provedor de conteúdo sintonizado, quantidade de tempo sintonizado no provedor de conteúdo).

Quando chegam ao servidor as informações são armazenadas no banco de dados, em seguida passam pelo processo de ETL e são armazenadas no DW. Após esse processo, as informações chegam ao módulo de Mineração de Dados onde são transformadas para o formato ARFF, padrão próprio da ferramenta de Mineração de Dados usada no projeto (WEKA) [9].

Em seguida são usados os algoritmos de regra de associação que utilizam técnicas de Inteligência Artificial e da área da Estatística para extrair o conhecimento implícito dos dados. Os principais algoritmos usados nesta primeira versão do MKTV foram: *Predictive Apriori* [10] *e Tertius* [11].

Os conhecimentos que os algoritmos de regras de associação encontram se concentram especificamente em traçar um perfil de uso do usuário (dispositivo set-top-box) da programação da TVD. Além de identificar preferências e padrões de utilização do usuário de TVD, como programas adequados aos gostos do usuário.

Esses conhecimentos descobertos ajudarão na implementação de vários produtos de software, como por exemplo, EPG's (do inglês *Electronic Programming Guide*) personalizados para TVD, aplicações que utilizam consulta semântica, etc.

3. EXPERIMENTO

Com o objetivo de validar a aplicabilidade do sistema MKTV, foram realizados experimentos para dar subsídios para avaliar de forma prática o projeto. Nesse sentido o projeto MKTV buscou dados reais de utilização de TVD, sendo atendido por meio da rede de televisão local - TV Arapuã.

A TV Arapuã (http://www.tvarapuan.com.br/), emissora de TV brasileira com sede em João Pessoa na Paraíba, forneceu para o projeto KTV, uma pesquisa de medição de audiência, desenvolvida pelo instituto brasileiro de pesquisa e medição de audiência (IBOPE), realizada no período de 11 a 17 de setembro de 2009, com alcance em todo o estado da Paraíba. A medição englobou a audiência das principais redes de TV aberta que atuam no estado nordestino, que são elas: Globo, Record, SBT, Rede TV, Bandeirantes, e TV Educativa.

As informações contempladas no relatório dizem respeito ao percentual de audiência de cada canal, em intervalos de uma hora, durante uma semana.

De posse dos dados de audiência por faixa de horário e por canal, ainda foi necessário buscar mais informações necessárias à mineração de dados. Neste caso, são as informações sobre quais programas eram executados em um determinado horário e canal. Então, como passo seguinte na aquisição dos dados foi coletado informações de um guia de programação de TV. Este informa, todos os programas exibidos pelo canal de TV e a faixa de horário que o mesmo é exibido.

A seguir, os dados foram inseridos no banco de dados operacional de acordo com o modelo relacional proposto e em seguida sofreram o processo de ETL. Durante o processo de Transformação, alguns atributos que são derivados de outros campos foram preenchidos, de acordo com o modelo multidimensional.

Como próximo passo, os dados foram carregados no DW, utilizando a ferramenta de ETL Kattle. Logo após, o processo de Mineração de Dados foi acionado gerando regras de associação sobre as informações coletadas.

Algumas regras que mostram conhecimento interessante são mostradas a seguir.

Algoritmo *Predictive Apriori* (Formato usando Lógica de 1º ordem):

- SE diasemana=segunda-feira E genero=Filme ==> ENTÃO canal=GLO tamanhoduracao=MEDIO
- SE programa=O Astro E classificacaoanoestreia=ATUAL ==> ENTÃO tamanhoduracao=CURTO

Sobre as regras do algoritmo *Predictive Apriori*, podemos inferir que a primeira regra informa que: "Quem assiste filmes as Segundas-feiras, assiste a programas de média duração no canal Globo".

A segunda regra implica que: "Quem assiste ao programa "O Astro" e a programas que estrearam a pouco tempo assistem programas de curta duração".

Depois de mineradas as regras foram encapsuladas, no formato ontológico, através da *OntoMKTV*, permitindo que o conhecimento descoberto possa ser utilizado através de outras aplicações, além do conhecimento poder ser reutilizado e estendido.

Há que se considerar que o conhecimento minerado através das informações de audiência de TV reflete um pouco da natureza da base de dados e de suas limitações. Isto se dá ao fato de que as informações mineradas provêm de uma pesquisa que abrange unicamente o estado da Paraíba, com informações de TV aberta, ou seja, 7 canais de TV. Além disso, os resultados obtidos demonstram a funcionalidade da MKTV no contexto do Projeto KTV.

4. CONCLUSÃO

Este trabalho apresentou a especificação e desenvolvimento do *Mining Knowledge TV*, uma arquitetura de descoberta de conhecimento em base de dados com técnicas semânticas integradas ao projeto *Knowledge TV*. Para a realização do MKTV foi necessário o cumprimento de algumas etapas.

Inicialmente realizou-se um estudo sobre a TV Digital Interativa e sobre a área de Descoberta de Conhecimento em Base de Dados e Mineração de Dados. Além disso, foram pesquisados os principais trabalhos que envolviam a utilização conjunta dessas áreas para identificar o estado da arte de Mineração de Dados em TVD.

Em seguida, foram realizadas pesquisas sobre a área de Representação do Conhecimento e o levantamento de requisitos para a construção do ambiente de descoberta de conhecimento em banco de dados do projeto *Knowledge TV*, o *Mining Knowledge TV* - MKTV.

Finalizada a pesquisa e o levantamento de requisitos, iniciou-se a especificação da arquitetura do MKTV, juntamente com a modelagem relacional e multidimensional do banco de dados operacional e do *Data Warehouse* utilizados no projeto. Como passo seguinte foi incorporado ao projeto o engenho de Mineração de Dados WEKA, necessário para a descoberta de conhecimento útil. Outra etapa importante, foi a criação da ontologia *OntoMKTV*, responsável pela representação semântica do conteúdo minerado.

Para validar o projeto foi necessário a busca por bases de dados reais no contexto da TV Digital como na pesquisa de audiência.

Desta forma, o projeto MKTV está inserido no contexto do Sistema Brasileiro de TV Digital – SBTVD junto ao *middleware* Ginga, fornecendo informações úteis para serem utilizadas em aplicações e serviços direcionadas ao consumo de usuários e provedores de conteúdo de TVD.

5. TRABALHOS FUTUROS

Como trabalho futuro, pode ser destacado:

- A utilização de outras bases de dados de TV, por exemplo, de TV a Cabo, para mais experimentos com o MKTV.

- A necessidade de se implantar algoritmos de diversas tarefas da Mineração de Dados, como por exemplo, classificação, clusterização, entre outros.
- Também poderá ser estudada a integração da solução da MKTV para *middlewares* de outros sistemas de TV digital seguindo as normas J200, J201 e J202.
- Poderá ser implantado algum recurso OLAP para visualização das consultas multidimensionais.

6. REFERÊNCIAS

[1] LEKAKOS, G.; CHORIANOPOULOS, K.; DOUKIDIS, G.. (2007). Interactive Digital Television: Technologies and Applications. IGI Publishing. EUA.

[2] LEITE, Luiz Eduardo Cunha ; BATISTA, Carlos Eduardo Coelho Freire ; SOUZA FILHO, G. L. ; KULESZA, Raoni ; ALVES, Luiz Gustavo Pacola ; BRESSAN, Graça ; RODRIGUES, Rogerio Ferreira ; SOARES, Luiz Fernando Gomes . FlexTV - Uma Proposta de Arquitetura de Middleware para o Sistema Brasileiro de TV. Revista de Engenharia de Computação e Sistemas Digitais (Impresso), São Paulo, v. 2, p. 29-49, 2005.

[3] W3C - World Wide Web Consortium. (2012). W3C Semantic Web Activity. Disponível em: <http://www.w3.org/2001/sw/>. Acesso em: Maio 2012.

[4] LINO, N. Q.; ARAÚJO, J.; ANABUKI, D.; PATRÍCIO JUNIOR, J. C. A.; BATISTA, M.; NOBREGA, R.; AMARO, M.; SIEBRA, C. Knowledge TV. In: European Conference on Interactive TV and Video – Euro ITV 2011. Lisboa -Portugal.2011.

[5] YU, H.; DIETZE, S.; BENN, NEIL. Semantic TV resources brokering towards future television. (2010) In 1st NoTube workshop on Future Television, EuroITV 2010. Jun 2010.

[6] Han, J. And Kamber, M. (2006). Data Mining Concepts and Techniques. 2a Edição, Editora Elsevier, Reino Unido.

[7] LINO, N.; ARAÚJO, J.; LEMOS, G. AND SIEBRA, C. (2010). Aspectos Semânticos e Convergência Digital (Web e TV Digital). Proceedings of 2a. Conferência Web W3C Brasil (W3C Web.br 2010), Belo Horizonte, Brasil.

[8] SOUZA FILHO, G. L.; LEITE, L. E. C.; BATISTA, C. E. C. F. (2007). 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.

[9] University of Waikato. (2012). Weka 3 – Machine Learning Software in Java. Disponível em <http://www.cs.waikato.ac.nz/ml/weka>. Acesso em Maio 2012.

[10] SCHEFFER, T. Finding association rules that trade support optimally against confidence. In: European conference on principles of data mining and knowledge discovery. 2004. pages. 424-435.

[11] FLACH, P.; LACHICHE, N. Confirmation-Guided Discovery of First-Order Rules with Tertius. In: Journal Machine Learning. Kluwer Academic. USA, 2001. pages. 61-95.

[12] KIMBALL, R. Data Warehouse Toolkit. 1998. São Paulo: Makron Books.

[13] ARAUJO, Jônatas Pereira Cabral De. CoreKTV - Uma infraestrutura baseada em conhecimento para TV Digital Interativa: um estudo de caso para o middleware Ginga. 2011. João Pessoa. Dissertação de Mestrado.

[14] ITU – International Telecommunication Union. ITU-T Recommendation J.200: Worldwide common core – Application environment for digital interactive television services. 2001. (a)

[15] ITU – International Telecommunication Union. ITU-T Recommendation J.201: Harmonization of declarative content format for interactive television applications. 2003.(b)

[16] ITU – International Telecommunication Union. ITU-T Recommendation J.202: Harmonization of procedural content formats for interactive TV applications. 2004.(c)

[17] ATSC - Advanced Television System Committee. (2012). Disponível em <http://www.atsc.org>. Acesso em Maio de 2012.

[18] DVB - Digital Video Broadcasting. (2012). Disponível em <http://www.dvb.org>. Acessado em Maio de 2012.

[19] ARIB - Association of Radio Industries and Business. (2012) Disponível em <http://www.arib.or.jp>. Acessado em Jan de 2012.

[20] SBTVD – Sistema Brasileiro de TV Digital. (2012). Disponível em <http://sbtvd.cpqd.com.br/>. Acessado em Maio de 2012.

Building and Monitoring Next-Generation Video Transport Infrastructure

Chin Chye Koh (ckoh@nevion.com)

BUILDING AND MONITORING NEXT-GENERATION VIDEO TRANSPORT INFRASTRUCTURES

Chin Chye Koh

Nevion USA, Inc., Oxnard, CA 93033, United States of America, ckoh@nevion.com

Abstract: Broadcasters building the infrastructure of tomorrow need to accommodate dynamic, flexible connections involving thousands of parameters such as IP addresses, node bandwidth capacities, and equipment firmware levels. Much of this detail should be hidden while making critical real-time information readily available. In addition, next-generation infrastructures are rarely created from a blank slate. More often than not, the new must interoperate with existing infrastructures and deliver services over a lengthy transition period. This can make it difficult to monitor QoS and ensure end-to-end signal protection. In addition, traditional monitoring tools oriented toward the network infrastructure rather than video transport services increase the difficulty in delivering high QoS. This paper explores how the right monitoring technology and managed connectivity can offer solutions to these challenges and help maximize legacy networks.

Keywords: Monitoring, Protection, Video Transport, Quality of Service.

1. INTRODUCTION

The nature of contribution media (audio, video, and data) transport is changing, as users demand more flexibility and efficient use of wide-area network resources. The industry is moving from the static, point-to-point reality of the past to a highly dynamic world where contribution services are set up and torn down on demand with changes occurring in minutes, all controlled directly by the end-user [1]. A new approach is required to deliver managed media services across fiber, SDH/SONET, or IP networks for this new reality.

Traditional network management is not sufficiently competitive within this landscape. Instead, the challenges of this new reality demand the use of fully managed services [2]. Service providers are also faced with new challenges to maximize utilization of both operational and networking resources while delivering media services with highly demanding performance guarantees to meet contracted service-levels.

The industry is quickly adopting IP technology for transporting media services. To realize the promise and flexibility of IP, IT based technologies must be leveraged to manage both the infrastructure and the services. To truly take advantage of this flexibility requires a fully integrated approach towards managing and monitoring the services delivered.

Efficient service provisioning is essential to optimal network utilization as it reduces the cost of delivering media services. In addition, service monitoring through analytics is equally important, as it reports exactly how delivered services are performing. Monitoring of all nodes through inventory management regardless if it is an edge or core element. As networks become more complex and hybrid infrastructure is built to achieve optimal efficiency, the health of every device is critical to ensuring that service providers stay on the path of providing the best service offering possible

2. SERVICE DELIVERY SYSTEM

To address these issues, a framework must be deployed that reduces the complex business of delivering media services over a variety of networks (fiber, SDH/SONET, IP) into fundamental elements. The primary directives include delivering managed services, consolidate data across the network, provide service perspective on operations, and provision occasional use or permanent services, allowing:

- Efficient utilization of operational resources
- Discovering potential service affecting problems
- Flexible service control by non-expert operators
- Maximizing use of infrastructure by adapting to dynamically changing needs

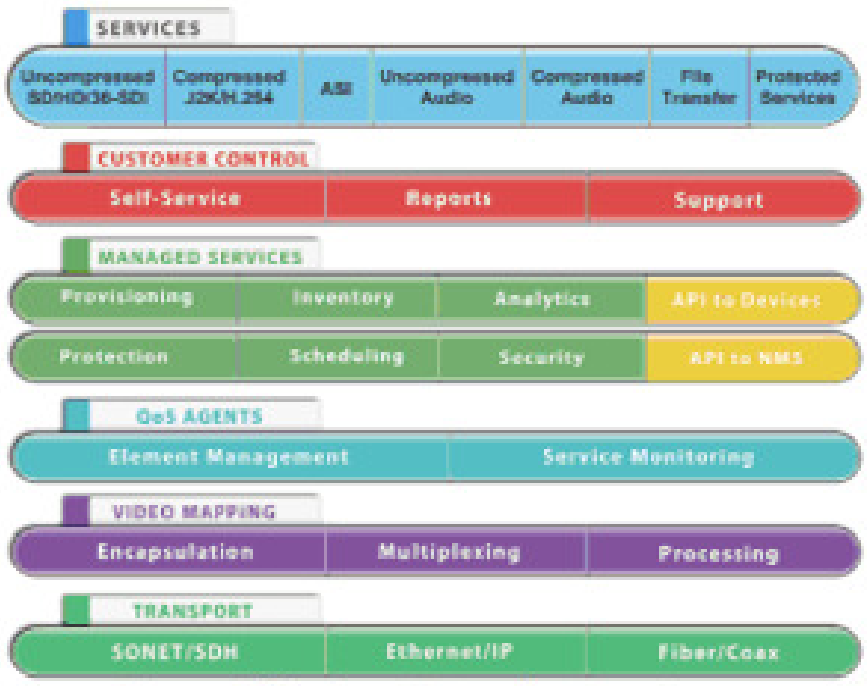

Figure 1: Service delivery framework.

In the framework presented (Figure 1), typical contribution services are presented on top, while the core infrastructure used to deliver these services are at the bottom. Everything in between are components necessary in delivering managed media services with the performance and quality users expect which in turn is often governed by service-level agreements. A

system for managing delivery of media services must be fully integrated to support external and internal workflows in order to minimize the cost of delivering services and includes critical components such as:

- Self-service interface to provision new and modify existing services with access to reports concerning performance and quality of delivered services
- Integrated provisioning and scheduling of the edge and core network devices in delivering services
- Protection through diverse paths for media flows ensuring maximum redundancy
- Security handled with special care in a multi-tenant environment to control content delivery
- Resource inventory/utilization for real-time planning for capacity control or to remove potential faults
- Open platforms that flexibly integrate with other systems and devices

To deliver managed media services, the level of operator intervention should ideally be zero. Users should be able to provision new or modify existing services as needs change. To further minimize operational cost, integrated service monitoring enables operators to take preventive action if contracted service-levels cannot be met.

Leading providers of media transport services are beginning to offer Web-based access to self-service applications. This enables end-users to provision new services without having to order them manually. While this increases flexibility for the user, it also assist service provider's maximize use of their infrastructure. Web based technology is well suited to business-to-business integration as it requires little integration while providing secure access to mission critical resources. In addition, deeper integration with automation or scheduling systems is also possible by utilizing Web service APIs for users that require a more integrated workflow.

3. MANAGING COMPLEXITY

IP/Ethernet has become the de-facto standard for data communication and has been shown to be capable of handling real-time applications such as VOIP and IPTV. Adoption for contribution quality video transport has rapidly increased with the scalability of video compression technology such as MPEG-2, H.264, JPEG2000, which allows every single bit of available bandwidth to be optimized. In addition, with the increasing availability of bandwidth, the deployment of uncompressed video is now growing.

Managing connectivity between media edge devices may appear simple. However, the complexity quickly grows when hundreds of end-points with dynamic connections are involved. This is further complicated when multipoint connectivity and redundancy mechanisms are considered [3]. As such, it is necessary to combine knowledge about the services that are to be delivered and the underlying infrastructure.

The fact that routing decisions in IP/Ethernet networks is largely the responsibility of the network itself works very well for data services. However, this non-deterministic behavior is not acceptable for live media contribution services as it leads to in-efficient use of network resources and increased buffering and latency to compensate for the behavior of the network. This lack of control is an important factor limiting adoption of IP/Ethernet for contribution media applications. Another factor is the migration path itself as broadcasters are moving from baseband to video over IP/Ethernet transport, which requires systems that can manage both worlds from ingress to egress while bridging the technology gap.

3.1. End-to-end provisioning of services

Proactive scheduling of transport paths for media services may be performed as a two-step process. In-advance resource scheduling is performed to reserve necessary resources during the duration of the services at the time of booking. This assumes ideal network conditions where all available links are operating normally according to the network design. Just-in-time scheduling is performed as part of the preparations for configuring a new service. This takes into account the current network conditions and attempts to optimize resource usage without breaking any in-advance resource reservations.

Avoiding network congestion caused by over-subscription requires bandwidth management of every link within the transport infrastructure. The transport path from source to destination(s) needs to be deterministic and managed by a provisioning system. In addition to bandwidth constraints, it is equally critical to ensure that path diversity exists with the appropriate failover mechanisms. The following multi-step approach supports path computation for any network (Figure 2):

- All network resources are modeled as a set of vertexes and interconnecting edges, representing a graph for the purpose of path calculation
- Availability of resources are considered to remove reserved resources from the graph
- Administrative and operational status of resources is considered to further reduce the graph.
- Constraints are applied to the graph, such as bandwidth and latency, before a shortest-path-first (SPF) algorithm is applied to the graph to find an optimal path from source to destination.
- The process is repeated for calculating redundant paths, taking into account resources reserved in the first path calculation.

The mechanisms for path calculation described above enables hybrid network architectures to be easily realized. Here video signals are switched at the

baseband, fiber, SONET/SDH, or IP/Ethernet layer to reach one or more destinations. This introduces a high level of flexibility in migrating from existing fiber/coax infrastructures to video over IP/Ethernet for metro and long-haul interconnections.

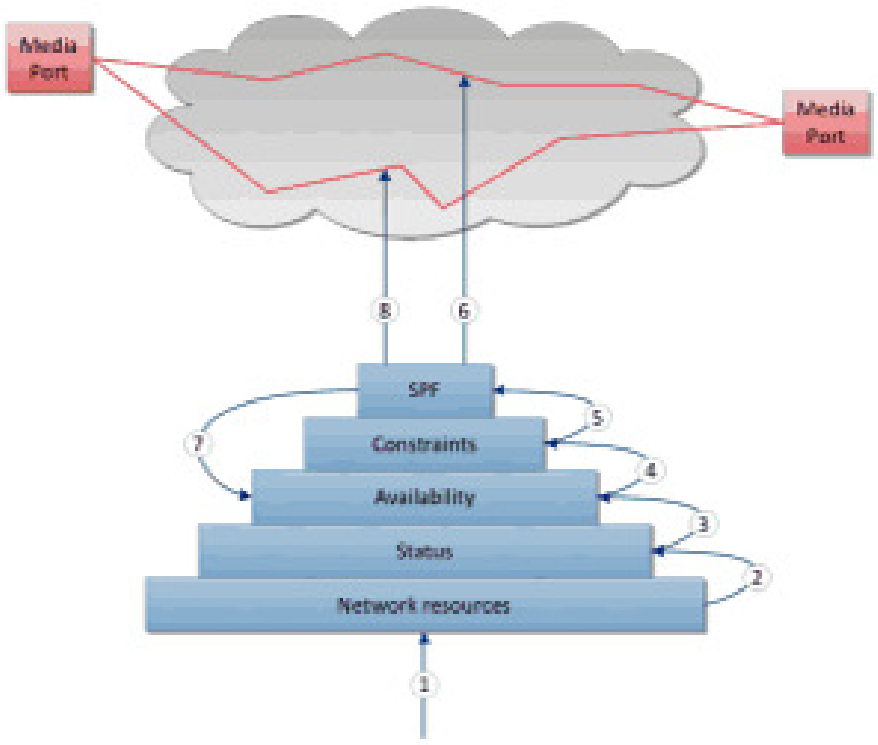

Figure 2: Path provisioning flow.

Such a system enables end-users to specify source and destination ports, handing of the complex process of optimal path selection to the system. To provision a path, the system utilizes a shortest path first algorithm to provision the least cost path from source to one or more destinations. Simultaneously, the system performs diverse path routing to support end-to-end redundancy through the network. In addition, the system keeps track of bandwidth allocation on each link to avoid over-provisioning, delivering predictable performance and the highest QoS possible.

3.2. Techniques for Provisioning Paths

There are various technologies that may be utilized when provisioning paths through the core network infrastructure [3]. A solution for IP core layer 3 networks can be deployed based on next-generation Multicast VPN (MVPN) router technology. This technique is based on configuring Label Switched Paths (LSP) at the provider edges that are then added to the appropriate MVPN. The provisioning of services across media edge and IP core devices from source to one or more destinations can be performed in less than 5 seconds.

In a layer 2 Ethernet environment, Virtual LAN (VLAN) technology may be used to dynamically create VLAN trunks for transport of media services throughout the network. A service is transmitted on a single VLAN that forms an end-to-end path through the network utilizing VLAN trunks from source to destination. This solution offers the benefit of well-defined paths that can be enabled and disabled in real-time based on network conditions. With VLANs, the entire network is segregated to form secure domains where no one service can affect the other. This approach provides a deterministic way to manage paths across any layer 2 Ethernet infrastructure.

In either of the above solutions, the management system is responsible for managing the bandwidth allocation on each path through the network. This minimizes propagation delay, while eliminating congestion, ensuring that the media reaches its intended destination in a timely manner.

In the real world, nearly all connections pass through a video or audio routing switcher, thus being able to perform path provisioning from end-to-end cannot be restricted to the intermediate data transport network. As such, path provisioning is not only related to networks, but in a fully integrated approach it is equally important to be able to provision paths at the baseband layer. The system needs to be flexible enough to manage crosspoints in video and audio routing switchers, and tie-lines between switchers and other baseband transport systems that offer switching capabilities. Performing path provisioning at the IP and baseband layers enables the system to take on a complete end-to-end view from the source to destination across any intermediate infrastructure.

This enables service providers to manage a hybrid environment and support a migration path from baseband video transport to IP/Ethernet transport.

4. SERVICE ANALYTICS

A system that quickly identifies problems allows operators to focus attention on those directly affecting users and prioritize the use of critical resources. These can be realized in the form of a set of essential information and performance indicators from the network elements and available services [4].

In-service probing is widely deployed today for video and audio transport services, but utilizing the information provided by these probes and correlating this with the actual services delivered is a challenging task. Once again an integrated approach is critical to make in-service performance data readily available for services that are delivered. In the past, service and network management have been separate functions but to deliver managed video and audio services it is key to combine the two.

Figure 3 illustrates how performance data from probes are correlated with equipment alarms and service data to enable the operator to easily identify what is or will be affected when appropriate action is taken to restore service or prevent potential problems.

When a new service is deployed there may also be requirements for in-service probing, and presenting evidence to the user that service objectives have been met. An integrated system can ensure that the service is not only provisioned from source to destination, but also that the appropriate probes are connected to this service to be able to provide service reports.

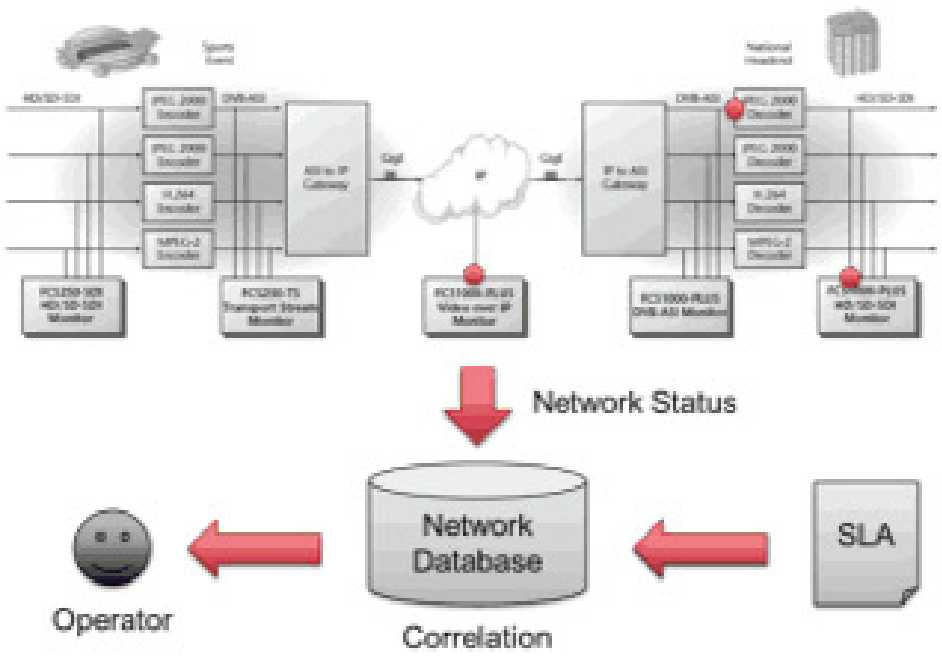

Figure 3: Service monitoring and data aggregation.

Service analytics also permit the system to proactively address potential problems, such as packet loss or packet jitter. By combining knowledge of services running across the network infrastructure with performance data it is possible to avoid such problems before they occur. This requires advanced correlation of service data with performance measurements from the network. The first step is to enable the operator to proactively utilize available data that previously have only been used for offline analysis. The next step is to enable the system to utilize the same data to take automatic action to avoid problems.

4.1. Monitoring: A Closer Look

Monitoring of media flows as it traverses the network is an important component in assuring Quality of Service (QoS). However, due to the complexity of handling multiple media types and the type of network, the need for monitoring can often be expensive and complex. This is especially true when you consider that it is often necessary to monitor every single element in the network.

As such, the goal here is to monitor a minimum number of systems or parameters in real-time, quickly collating the appropriate information in order to make decisions quickly and efficiently. This requirement eliminates complex test instruments that are predominantly targeted at engineering development.

To this end, there are various systems available that are capable of providing the 24x7 monitoring passive monitoring that is absolutely critical [4]. These are based on several standards, that when judiciously applied can be used to streamline the monitoring requirements to a manageable level

4.1.1. Video

Baseband video monitoring should be limited to parameters such as Error Detection Handling (EDH) [5] and Cyclic Redundancy Checks (CRC) [6]. Monitoring these basic parameters in a known environment provides sufficient information to characterize the signals and to make a determination of the health of these signals.

4.1.2. Audio

Audio can be presented independently or pre-embedded into video. However, the standards for audio in the contribution are complementary. The following standards are key to monitoring audio, and emphasis should be placed on detecting the presence of audio, if errors exist by examining the ECC parameters, and finally the audio levels [8-10].

4.1.3. Transport Streams

Video and Audio when carried in compressed formed, are typically encapsulated into Transport Streams. Transport Streams are widely used today and is the choice when deploying all the popular video codecs such as MPEG-2, H.264, and JPEG2000. It's systems architecture lends itself to efficient analysis and monitoring, allowing users to quickly determine if faults exist. The performance of Transport Streams is completely outlined by ETSI TR 101-290 [11].

Table 1: Priority levels in Transport Stream Analysis

Priority 1 : Errors that prevent decodability	Priority 2 : Errors that impair decodability	Priority 3 : Errors that do not affect decodability
TS synchronization loss	Transport error	NIT error
Synchronization byte error	CRC error	SI repetition error
PAT error	PCR error	Buffer error
Continuity Count error	PCR accuracy error	Unreferenced PID error
PMT error	PTS error	SDT error
PID error	CAT error	EIT error
		RST error
		TDT error
		Empty buffer error
		Data delay error

The standard specifies three different priority levels, Priority Level's 1, 2, and 3 (Table 1). For real-time monitoring purposes, only a subset is required which would consist of Priority Level's 1 and 2. Errors detected at these levels indicate that the media is not decodable or would suffer from severe artifacting after decoding.

4.1.4. IP Encapsulation

For media adaptation to an IP/Ethernet layer, existing standards leverage a wealth of knowledge within the Internet Engineering Task Force (IETF) domain. The pertinent standards can be found under SMPTE and EBU stewardship [12-16]. Note that the standards for uncompressed video over IP/Ethernet, SMPTE 2022-5 and -6 [14, 15], are not finalized yet. However, several vendors have already implemented the letter of the standard and participated in private and public inter-operability tests.

4.1.5. Network Performance

Once media is packetized and launched into an IP/Ethernet infrastructure, there must be a method of tracking its behavior as it traverses the network. To this end, the most important metrics from an end-end delivery perspective are packet jitter (packet delay

variation, PDV) and packet loss ratio (PLR) [17]. Other metrics such as latency and packet error ratio are not as significant primarily because latency does not affect the performance of the media itself while packet errors are usually converted to packet losses.

4.2. A Typical Monitoring Flow

In the end-to-end delivery of media, there are a number of different process as illustrated in Figure 4 which also includes the types of measurements that are taken at each stage.

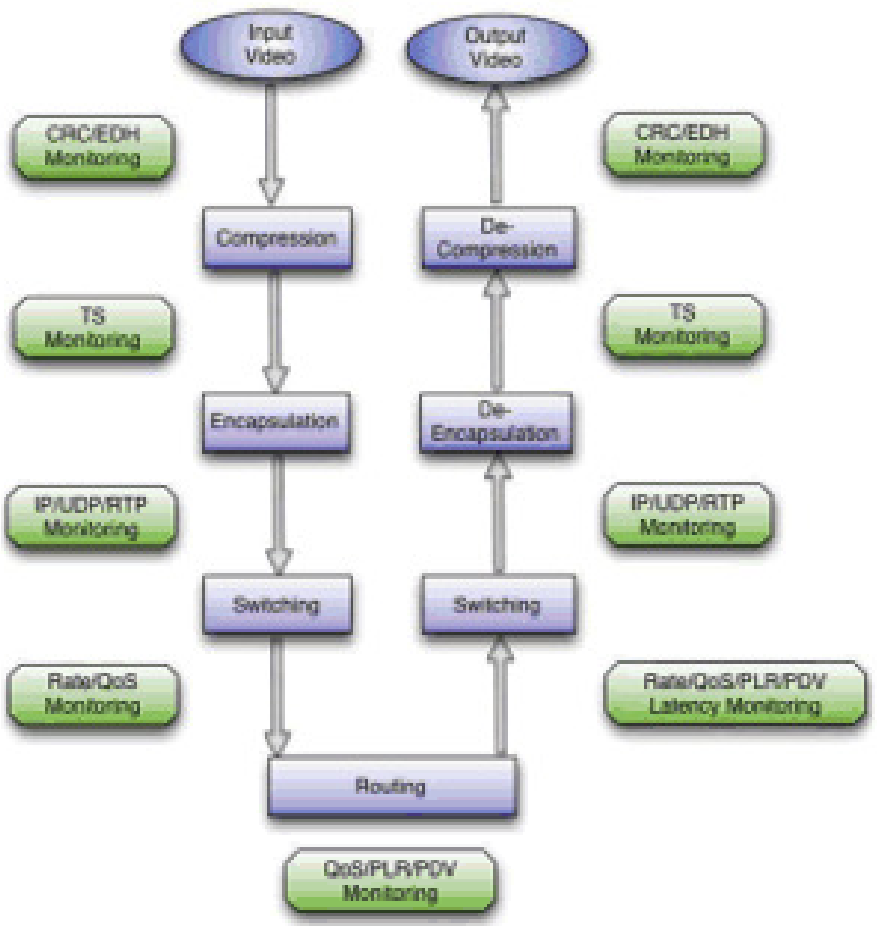

Figure 4: Typical monitoring flow.

At the baseband layer, a cursory inspection of the CRC/EDH values from the input and output signal provides a quick health assessment. Errors detected/signaled at this stage indicate poor health with higher occurrences indicating degraded signal quality. Based on thresholds, maintenance / troubleshooting services can be automatically triggered.

After compression, it is essential to ensure that the compression engine had completed its tasks. In general, examining the TR101-290 metrics yield an indication of the viability of the output. If errors are found, the severity can be quickly determined by comparison to preset thresholds with the appropriate actions taken immediately.

Once the media has been encapsulated the focus switches to the transport domain. Here, the IP/Ethernet headers are examined to ensure that no error has been made, or packets lost. At the routing and switching stages, the focus is on ensuring that the proper routes have been taken, that the packetized flow is error free with no packet losses, and that the packet jitter does not exceed tolerances. In essence, all the preceding stages provide confidence that the media is error free and that only a cursory examination of the IP layer is sufficient. In addition, monitoring of these metrics at the IP/Ethernet layer provide statistics towards the overall health and efficiency of the network that ties into the provisioning methods outlined previously.

5. PROTECTING MEDIA

As described previously, there are a number of ways of provisioning paths through a network. In a design that places emphasis on diversity, additional measures must be put into place to guarantee the highest possible QoS. Figure 5 provides examples of methods of transporting video [3].

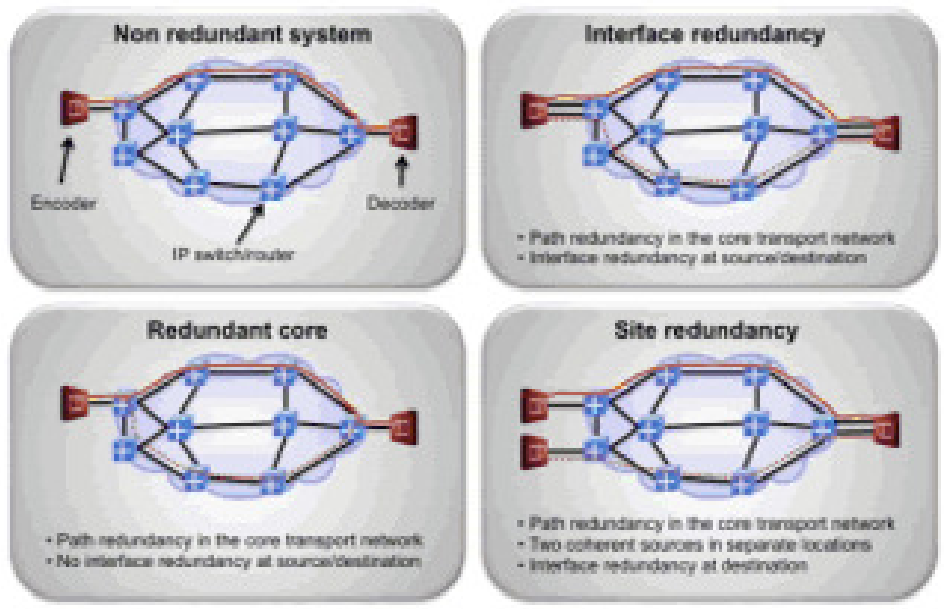

Figure 5: Media transport methods.

- Top-Left: Non-redundant system. The management system is responsible for detecting failures in the network. Once a failure is detected, an alternate path is automatically calculated and provisioned. In such a scenario, a video outage is unavoidable. In provisioning methods using both Layer 2 and Layer 3 protocols, end-end provisioning can be effected in 5 seconds with compressed video resynchronizing in less than 2 seconds. If the network has built-in automatic protection switching, packet flows can be restored in approximately 50msecs. As such, a real-world estimate indicates that loss of video for a properly designed network should never exceed 10 seconds.

- Bottom-Left: Redundant core. Packet flows are duplicated at the first ingress node in the core network itself and routed through diverse paths. At the receiver nodes, the egress node only outputs the primary feed. Upon detection of a failure in the core, the management system directs the egress node to output packets from the secondary feed. Since this is a failover at the packet level with no protection at the media adapter, the video service will be momentary interrupted. Depending on the content, codec, and switch over time, the glitch may be undetectable or detectable for a few frames.

- Top-Right: Interface Redundancy. The media adapter produces duplicate flows, routed through the core network over diverse paths. At the receiver, the

duplicate flows are buffered and aligned with the best packet selected for use. Such a scheme offers perfect/hitless switching with no impact on the output video in the event of a link or node failure. Furthermore, the provisioning methods outlined above would be automatically triggered to re-route the packets to ensure that two diverse paths always exists.

• Bottom-Right: Site Redundancy. Here, dual media adapters are synchronized and perform the duplicate media adaptation. As in the case of the interface redundancy, the duplicate packet flows are routed over diverse paths. In this case not only are link and hardware failures protected for, the media adaption layer is also protected.

6. SUMMARY

By adding intelligence to critical functions—service provisioning and analytics, network inventory, protection, and performance management—managed video services greatly reduce requirements for human intervention. When complexities of network configuration, maintenance, troubleshooting and problem resolution are hidden, it becomes possible to accomplish more with the same level of network infrastructure, while reducing human intervention. This optimized use of network resources helps broadcasters reduce costs, increase revenue, and monetize content.

The system must deliver all of the tools necessary to manage and monitor end-to-end connectivity, bringing automation and ease to the complexity of deploying media services. By adding a network abstraction layer, operators can provision and monitor services without having to consider detailed configuration settings in each network element. The system should auto-detect available endpoints and allows the user to set up connections by selecting the appropriate endpoints and service profile.

Using standard Web technologies, employing a graphical interface, and smart navigation, it is possible to provide a complete overview of scheduled and in-service connections over any infrastructure. Broadcasters can simplify connection scheduling based on service profiles, monitor services and bandwidth utilization, and access map and timetable views. Ultimately, such a system eliminates the need for manual configuration of each network element increasing efficiency. The connection management capability enables functional use of bandwidth—using only the amount of bandwidth required at any given time—which is essential to creating the most efficient and cost-effective network for video and audio contribution services.

Ultimately, integrated monitoring and protection capabilities enhance the transport environment guaranteeing the highest possible Quality of Service. Such a system offers the most essential element in contribution media delivery, **Peace of MInd**.

REFERENCES

[1] GigaBooking, http://gigacontent.dk/loesninger/gigabooking.aspx

[2] J. I. Helgesen, "Leveraging video services management to enhance video transport," 2011 SMPTE Annual Technical Conference & Exhibition, Hollywood, USA, October 2011.

[3] C. C. Koh, "Media Networks – Overview of new techniques for real-time service protection and control, " Vidtrans2012, New Orleans, USA, February 2012.

[4] C. C. Koh, "Monitoring Next-generation Video Transport Infrastructure: Meeting SLA Requirements," BCA2012, Singapore, June 2012.

[5] SMPTE RP 165-1994: Error Detection Checkwords and Status Flags for Use in Bit-Serial Digital Interfaces for Television, 1994.

[6] SMPTE 292-2006: 1.5 Gb/s1 Signal/Data Serial Interface, 2006.

[7] SMPTE 424M-2006: 3 Gb/s Signal/Data Serial Interface, 2006.

[8] SMPTE 272M-2004: Formatting AES Audio and Auxiliary Data into Digital Video Ancillary Data Space, 2004

[9] SMPTE 299-1:2010: 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface, 2012

[10] SMPTE 299-2:2010: Extension of the 24-Bit Digital Audio Format to 32 Channels for 3 Gb/s Bit-Serial, 2010

[11] ETSI TR 101-290: Digital Video Broadcasting (DVB); Measurement guidelines for DVB systems

[12] SMPTE 2022-1-2007: Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks, 2007.

[13] SMPTE 2022-2-2007: Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks, 2007.

[14] SMPTE 2022-5: Forward Error Correction for High Bit Rate Media Transport over IP Networks, unpublished.

[15] SMPTE 2022-6: High Bit Rate Media Transport over IP Networks SMPTE2022-6, unpublished.

[16] EBU – TECH 3326: Audio Contribution over IP; Requirements for Interoperability

[17] ITU-T Y.1541: Global Information Infrastructure, Internet Protocol Aspects and Next-Generation Networks; Quality of Service and Network performance.

Impacto da Cobertura de TV Digital Brasileira na Prestação do Serviço de TV Por Assinatura

Mateus Cruz (mateus@cpgei.cefetpr.br)
Keiko Fonseca (keikoveronicaono@gmail.com)

IMPACTO DA COBERTURA DE TV DIGITAL BRASILEIRA NA PRESTAÇÃO DO SERVIÇO DE TV POR ASSINATURA

Keiko Verônica Ono Fonseca[1], *Mateus Cruz*[2]

[1] UTFPR, Curitiba, Brasil, keiko@utfpr.edu.br
[2] UTFPR, Curitiba, Brasil, mcruz@pop.com.br

Resumo: Este estudo visa discutir o impacto da cobertura da TV Digital Brasileira no novo modelo de negócio das empresas prestadoras de serviços de telecomunicações e especificamente as atuantes na prestação do serviço de TV por Assinatura, avaliando cenário, problemas e possíveis soluções ao problema apresentado.

Palavras chave: TV Digital, Sinal DTT, TV por assinatura e DTH.

1. INTRODUÇÃO

O atual cenário do negócio de telecomunicações brasileiro está passando por modificações no seu formato em virtude do crescimento da economia e o respectivo aumento do poder aquisitivo de consumidores das classes média e média baixa, das inovações tecnológicas na área de vídeo digital, da demanda por novos serviços, da maior competição no mercado nacional com novos participantes, e sobretudo no modelo de prestação de serviços de telecomunicações.

As prestadoras de serviços de telecomunicações atuantes em telefonia fixa, móvel e TV por assinatura estão em reformulando seus modelos de negócio e da forma de condução e entrega de seus serviços, impulsionadas pela abertura de mercado, a globalização do setor, da acirrada competição por otimização de custos, oferta de novos serviços e da busca por fidelização de clientes. As consequências deste cenário são notáveis e têm exigido modificações das práticas usuais destas empresas.

Recentemente, com a aprovação da PL29/2007 [1] transformada em Lei Ordinária 12485/2011 [2] abriu-se o mercado para a atuação das empresas de telefonia, como prestadoras de serviço de acesso condicionado para a distribuição de conteúdo televisivo, obedecidas os requisitos impostos pela referida lei.

Neste cenário, as prestadoras de serviços de telecomunicações, originalmente atuantes no setor de telefonia fixa e ou móvel, estão iniciando as suas operações para a prestação de serviços de TV por assinatura [3] [4] [5] [6] , oferecendo ao mercado novos formatos de entrega de serviços de telefonia fixa e/ou móvel, banda larga e TV por assinatura, conhecidos por pacotes triple-play (pacotes com prestação de três serviços) e ou também quadri-play (pacotes com prestação de quatro serviços) [7] [8].

Estes novos players têm por objetivo agregar valor aos produtos inicialmente oferecidos, e através destes otimizar os custos com maior e melhor aproveitamento da infraestrutura operacional, sistêmica e comercial atual, com um forte apelo de fidelização da sua base de clientes, criando oportunidades de redução de preços e maior competitividade, com a inclusão de novos serviços de telecomunicações e entrada e ou permanência no ciclo de consumo dos clientes. Busca-se assim garantir um acréscimo na receita e proteção contra a concorrência.

Associado a este novo cenário para as prestadoras de serviços de telecomunicações, o Governo Brasileiro implantou o Sistema Brasileiro de Televisão Digital Terrestre (SBTVD-T), e em 2006, através do Decreto 5.280 [9], estabeleceu as diretrizes para a transição do sistema de transmissão analógica para o sistema de transmissão digital do Serviço de Radiodifusão de Sons e Imagens e do Serviço de Retransmissão de Televisão [10] baseado no sistema de transmissão japonês, o ISDB-T (*Integrated Services Digital Broadcasting Terrestrial*), com ganhos de qualidade de som e imagem com conteúdo HDTV (TV em alta definição), multiprogramação, mobilidade, iteratividade, entre algumas das possibilidades de novos serviços.

O casamento da evolução do modelo de negócio das empresas do setor de telecomunicações associada com a nova tecnologia da TV Digital brasileira, geraram novas possibilidades e oportunidades de incremento na prestação de serviços de telecomunicações, mas também novos problemas nunca antes enfrentados. Este artigo apresenta o cenário como visto pelas operadoras de serviços de telefonia e TV por assinatura, as constatações obtidas da implementação de modelos de negócios incluindo conteúdo televisivo bem como uma discussão sobre possíveis soluções ou evolução do cenário descrito obtidos com a participação da academia.

2. CENÁRIO

As prestadoras de serviços de telecomunicações entrantes no mercado de TV por assinatura vislumbraram na evolução da tecnologia da TV Digital brasileira e na tendência crescente do mercado, uma nova oportunidade para buscar um diferencial no atendimento de seus clientes, ou seja, elas associaram a oportunidade de entrega do sinal digital de televisão terrestre (*Digital Terrestrial Television* -

DTT) [11] [12] em alta definição (*High Definition - HD*) em conjunto com a sua solução de sinal satelital doméstico (*Direct To Home - DTH*) [13] em seu serviço de TV por assinatura, realizando desta forma um incremento na quantidade de canais inicialmente entregues ao cliente em sua grade e com qualidade HD.

Conforme tabela 1, a distribuição por região do número de cidades cobertas com o sinal DTT brasileiro conta atualmente um total de 433 cidades já cobertas oficialmente [14].

Região	número de cidades cobertas
Sudeste	161
Nordeste	81
Sul	122
Norte	37
Centro Oeste	32
	433

Tabela 1 – Regiões com cobertura de TV Digital no Brasil [14]

Com este cenário de cobertura e de constante crescimento, as prestadoras de serviço de TV por assinatura passaram a adicionar e fornecer em seu serviço a entrega do sinal DTT e todos os canais HD associados e disponíveis em cada cidade de atuação como sendo um diferencial em fornecimento do produto com alta definição, garantindo que já no momento da instalação o cliente independente do pacote de serviços contratado de TV por assinatura, o fornecimento de canais HD da TV digital aberta gratuitamente [3] [4] [5] [6].No entanto, problemas foram detectados durante o processo de instalação do serviço em regiões urbanas de diversas cidades, problemas estes associados a visibilidade e a confirmação da intensidade do sinal DTT para um ou mais canais de sua região de cobertura. Conforme o trabalho de Vasco [15][16] a estimativa da atenuação por propagação é extremamente complexa e depende de inúmeras variáveis associadas, como condições climáticas, condições geográficas do terreno ao longo do percurso da propagação e os obstáculos da aglomeração urbana no contorno do receptor (NBR 15604 [10]).

O estudo comprova as dificuldades de se obter um modelo único de propagação [17] [18] e evidencia as dificuldades operacionais enfrentadas pelas empresas atuantes no setor na garantia da qualidade do sinal DTT, para todos os canais digitais ativos na região. A Figura 1 exemplifica e ilustra o impacto da aglomeração urbana da cidade de Curitiba versus a região de cobertura do sinal DTT de testes adaptado do sistema SIGAnatel [19].

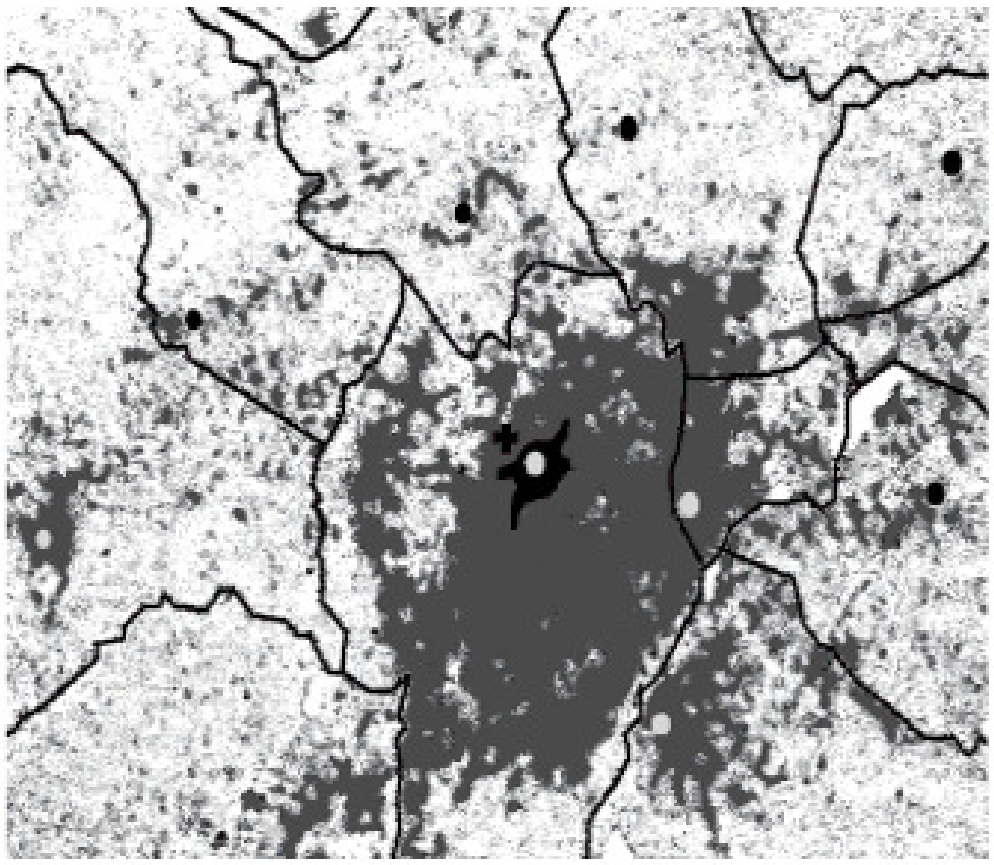

Figura 1. Exemplo cidade de Curitiba e aglomeração urbana versus região de cobertura do sinal digital adaptado do sistema SIGAnatel. [19]

3. CONSTATAÇÃO

A constatação operacional realizada em campo por empresas de acesso condicionado associada as diversas reclamações dos usuários do serviço coletadas nos exemplos disponíveis na internet [20] [21] [22] valida os dados do trabalho acadêmico citado onde se observa a qualidade da recepção dependente da região de cobertura dos canais digitais em serviço com sinal DTT. A tabela 2 [14] refere-se ao percentual da população por região com cobertura do sinal DTT: em média cerca de 45% da população é atendida por região.

Região	população total da região	população coberta na região	% de população na região com sinal de TVD
Sudeste	80.353.724	47.640.302	59%
Nordeste	53.078.137	16.523.755	31%
Sul	27.384.815	10.305.614	38%
Norte	15.866.533	5.171.014	33%
Centro Oeste	14.050.340	7.077.553	50%
	190.733.549	86.718.238	45%

Tabela 2 – Percentual da população por regiões com cobertura de TV Digital no Brasil

No entanto, há falta de informação sobre o nível de recepção do sinal DTT e sua região de cobertura, pode causa que o sinal DTT sofra do efeito precipício ou limiar (*Cliff effect*) [23]), ou seja, à medida que a relação sinal/ruído vai caindo, a imagem se mantém estável, até o ponto chamado de "*digital cliff*", ponto onde ocorre a queda abrupta do sinal DTT por

completo conforme Figura 2 [24], o *digital cliff* ocorre por causa da perda de bits do sinal digital, que em condições normais de recepção, o erro proveniente da correção da informação digital dever ser próximo de zero. À medida que a potência do sinal decresce, há um momento em que não é mais possível recuperar a perda de informação digital ou corrigi-la, e o sinal como um todo é cortado, e sua ocorrência implica em perda de qualidade na avaliação subjetiva do usuário do serviço de TV por assinatura.

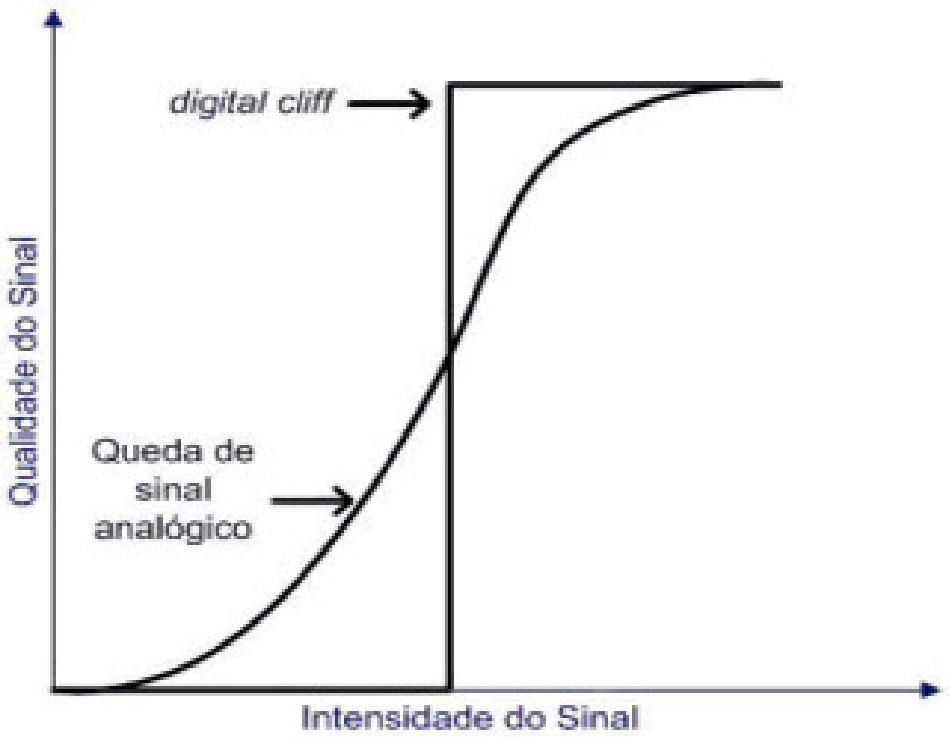

Figura 2 – Efeito precipício ou Cliff Effect [24]

A constante mudança de condições de propagação em regiões de grande variação no volume e altura de construções implica em mudanças no nível de recepção do sinal DTT na área de cobertura prevista no momento da concessão da licença para transmissão do sinal da TV aberta [15] [25]. Isto implica em custos de qualidade de serviço para as empresas prestadoras de TV por assinatura que se propõem a fornecer o sinal DTT aberto conforme seu modelo de negócio.

É interessante observar que conforme a Portaria MC nº 652/2006 a consignação de canais para TV aberta observará o estabelecido no PBTVD (art. 2º, § único), tal que o canal digital (art.10) deve "proporcionar a mesma cobertura que o atual canal utilizado para transmissão analógica observado o disposto no instrumento de outorga". [26].

No entanto, o usuário nem sempre está atento à questão da importância de antenas externas para recepção do sinal digital da TV aberta. Ele usualmente faz uso somente da antena interna, nem sempre adequada às condições de recepção com qualidade mínima (fora da zona do efeito precipício), fato constatável pela quantidade de dúvidas de leigos relacionadas à instalação de antenas nos foros da Internet [27] [28] [29].

Conforme [30], "70% dos assinantes veem mais canais abertos via TV paga do que os canais pagos de conteúdo específico", ou seja, usuários optam pelos canais abertos através do sistema de acesso condicionado por assinatura, o que mostra a força da penetração da TV aberta no Brasil.

Estas constatações reforçam a necessidade de uma aferição de cobertura no sistema de TV Digital Brasileiro, visando uma melhor qualidade de serviços que requeiram recepção de canais abertos (DTT). Vale frisar que a aferição obtida através de predição por métodos e modelos computacionais deve ser complementada por medições in loco [25].

4. CONCLUSÃO

A introdução de uma nova tecnologia nem sempre se faz em transições suaves, em especial quando considerada em modelos de negócios nunca antes testados. A entrada de novos players (operadoras de Telefonia) no mercado de distribuição de conteúdo televisivo associada a novos serviços de telecomunicações é um exemplo disto.

Este artigo procurou demonstrar a necessidade de aprofundar estudos sobre a cobertura do sistema de televisão digital brasileiro, com a finalidade de mapeamento do sinal DTT em âmbito nacional e prover um mapa de cobertura para todos os canais digitais visando orientar consumidores e prestadores de serviço em busca de uma melhor da qualidade de serviço, assim como a correta utilização da antena para captação do sinal DTT de forma a garantir melhor qualidade. Ávidos novos consumidores por conteúdo televisivo em formato digital podem associá-lo a baixa qualidade devido à falta de informações objetivas que permitam a implementação de serviços de entrega deste conteúdo em alta qualidade. A apresentação de orientações objetivas aos consumidores similares às iniciativas realizadas no Reino Unido em preparação ao desligamento da TV analógica (*switchover*) [26] [27] [28] é necessária para uma transição transparente e suave entre as tecnologias com garantia da Qualidade de Serviço.

Outro ponto importante para discussão e estudo futuro refere-se ao possível compartilhamento de infra-estrutura predial (por exemplo, nos moldes de *unbundling* [29] como em sistemas telefonicos) para atendimento do sistema de televisão por cabo e antenas únicas. Com a entrada de novos players fica evidente a limitação e muitas vezes a inexistência desta capacidade para atender a operação de diversas operadoras, causando aumento de custos, cancelamentos e insatisfação dos consumidores. Neste sentido, subsídios, renúncia fiscal ou negociação conjunta com operadoras interessadas na atualização de redes internas de condomínios trariam benefícios evidentes aos consumidores e empresas.

REFERÊNCIAS

[1] Brasil. Projeto de lei ordinária PL29/2007.Disponível em: http://www.camara.gov.br/proposicoesWeb/fichadetramitacao?idProposicao=339998, acesso em 27/05/2012

[2] Brasil. Lei Ordinária 12485/2011. Disponível em: http://www2.camara.gov.br/legin/fed/lei/2011/lei-12485-12-setembro-2011-611410-norma-pl.html, acesso em 27/05/2012

[3] GVT. Pacotes de TV por assinatura. Disponível em: https://www.gvt.com.br/portal/residencial/tv/. acesso em 27/05/2012.

[4] Claro TV. Pacotes de TV por assinatura. Disponível em: http://www.viaembratel.com.br/gradecanais.aspx acesso em 28/05/2012.

[5] OI TV. Pacotes de TV por assinatura Disponível em: http://www.oi.com.br/oi/oi-pra-voce/oi-tv acesso em 30/05/2012.

[6] Vivo TV, Pacotes TV por Assinatura : http://www.vivo.com.br/portalweb/appmanager/env/web?_nfpb=true&_nfls=false&_pageLabel=vcTvVivoTvBook acesso em 28/05/2012

[7] Triple-Play. Disponível em: http://pt.wikipedia.org/wiki/Triple_play acesso em 30/05/2012

[8] QuadriPlay. Disponível em: http://en.wikipedia.org/wiki/Quadruple_play acesso em 30/05/2012

[9] BRASIL. Decreto nº 5.280, de 29 de junho de 2006. Dispõe sobre a implantação do SBTVD-T. Diário Oficial da União, Brasília, 30 jun. 2006.

[10] Associação Brasileira De Normas Técnicas. NBR 15604: Televisão digital terrestre – Receptores. Rio de Janeiro, 2007.

[11] ANATEL (Brasil). Anexo II da Resolução Anatel nº 398, de 7 de abril de 2005. Recomendação ITU-R P.1546-1 - Método de previsões ponto-área para serviços terrestres na faixa de frequências de 30 a 3000 MHz. Diário Oficial da União, Brasília, 19 abr.2005.

[12] DTT. Disponível em: http://en.wikipedia.org/wiki/Digital_Terrestrial_Television. Acesso em 30/05/2012

[13] DTH. Disponível em: http://en.wikipedia.org/wiki/Satellite_television#Direct_to_Home_television acesso em 30/05/2012

[14] Brasil. DTV Disponível em: http://www.dtv.org.br/wp-content/uploads/2012/04/dados_sobre_a_tv_digital.pdf acesso em 24/05/2012

[15] Vasco, C. L. Modelo de propagação empírico para sinais de TV digital em Curitiba. 2009. 64f. Dissertação (Mestrado em Telemática) – CPGEI. Universidade Tecnológica Federal do Paraná. Curitiba, 2009.

[16] Vasco, C. L., Rehme, J. F., Martins, R. A., Agostinho, E., Pohl, A. A. P., Souza, R. D., Fonseca, K. V. O. e Pellenz, M. E. Ensaios de Transmissão e Recepção de TV Digital Terrestre na Região Metropolitana de Curitiba. Anais Congresso SET2008, São Paulo, 2008.

[17] Hata, M. Empirical Formula for Propagation Loss in Land Mobile Radio Services. IEEE Trans. On Vehicular Technology, vol.VT-29, nº 3, pp.317-325, Aug. 1980.

[18] ANATEL (Brasil). Anexo II da Resolução Anatel nº 398, de 7 de abril de 2005. Recomendação ITU-R P.1546-1 - Método de previsões ponto-área para serviços terrestres na faixa de frequências de 30 a 3000 MHz. Diário Oficial da União, Brasília, 19 abr.2005.

[19] ANATEL (Brasil). SIGAnatel - Sistema de Informações Geográficas da Anatel. 2008. Disponível em <http://sistemas.anatel.gov.br/siganatel>.

[20] Brasil. Reclamações de TV Digital. Disponível em: http://www.reclamao.com/reclamacao.asp?cod=41672 acesso em 30/05/2012.

[21] Brasil. Mapa Digital Disponível em: http://www.vejaisso.com/onde-pega-o-sinal-da-tv-digital-em-sao-paulo/ acesso em 30/05/2012

[22] Brasil. Reclamações de TV Digital. Disponível em: http://www.tvmagazine.com.br/talktv/readl.asp?ID=104839&pag=2 acesso em 3/05/2012

[23] ITU 2005. ITU Recommendation ITU-R BT.1735 Métodos de Evaluación Objetiva de la Calidad de la Cobertura de las Señales de Radiodifusión de Televisión Digital Terrenal del Sistema B especificado em la recomendación UIT-R BT.1306. Geneva: s.n., 2005.

[24] Cliff Effect. Disponível em: http://en.wikipedia.org/wiki/Cliff_effect acesso em 30/05/2012

[25] de Borba, M. A.C, Aferição de Cobertura no Sistema de TV Digital Terrestre brasileiro. Dissertação de mestrado. CPGEI. UTFPR, Curitiba, 2011.

[26] Brasil. Ministério das Comunicações. Portaria 625/2006. Disponível em:

http://www.mc.gov.br/legislacao/portarias/portaria-n-652-de-10-de-outubro-de-2006. Acesso em 27/05/2012

[27] Brasil. TV Digital Disponível em: http://www.simplificandotvdigital.com.br/ acesso em 28/05/12.

[28] Brasil. DTV. Disponível em: http://www.dtv.org.br/index.php/duvidas-frequentes/ acesso em 28/05/12.

[29] Brasil. Tecnologia. Disponível em: http://tecnologia.terra.com.br/noticias/0,,OI2115991-EI12882,00-Perguntas+e+respostas+sobre+a+TV+digital.html acesso em 28/05/12.

[30] Siqueira, E. Cotas absurdas na TV. Disponível em: http://ethevaldo.com.br/coluna/cotas-absurdas-na-tv/ . Acesso em 29/05/2012

[26] Desligamento da TV analógica. Notícias MC 2012. Disponível em: http://www.mc.gov.br/tema-radio-e-tv/noticias-radio-e-tv/25065-ministerio-conclui-primeira-versao-do-plano-de-desligamento-da-tv-analogica acesso em 28/05/2012

[27] Digital UK (analog TV switchover). Disponível em: http://www.digitaluk.co.uk/ acesso em 28/05/2012

[28] Brasil. MINICOM. Disponível em: http://www.mc.gov.br/todas-as-noticias/23827-131011-apagao-analogico-esta-mantido-para-2016-diz-bernardo, acesso em 27/05/2012

[29] Anatel. Despacho SPB 172/2004. Disponível em: http://www.anatel.gov.br/Portal/verificaDocumentos/documento.asp?numeroPublicacao=97760&assuntoPublicacao=Anatel%20anuncia%20condi%E7%F5es%20para%20desagrega%E7%E3o%20de%20redes%20fixas&caminhoRel=null&filtro=1&documentoPath=biblioteca/releases/2004/release_13_05_2004(4).pdf

TV Digital Interativa: Um Recurso Didático no Processo de Ensino e Aprendizagem da Educação Profissional e Tecnológica

Pricila Rodrigues de Souza (pricila.souza@am.senai.br)
Halison Ribeiro Maia (halison.maia@am.senai.br)
Marlos AndrÃ© Silva Rodrigues (marlos.rodrigues@am.senai.br)
Samir Pereira Batalha (samir.batalha@am.senai.br)
Daniel Albino Sandoval (daniel.sandoval@am.senai.br)
Orlewilson Maia (omaia@ufmg.br)

TV DIGITAL INTERATIVA: UM RECURSO DIDÁTICO NO PROCESSO DE ENSINO E APRENDIZAGEM DA EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA

Pricila Rodrigues de Souza [1], *Hálison Ribeiro Maia* [1], *Marlos André Silva Rodrigues* [1], *Samir Pereira Batalha* [1], *Daniel Albino Sandoval* [1], *Orlewilson Bentes Maia* [2]

[1] Núcleo de TV Digital Interativa – Serviço Nacional de Aprendizagem Industrial da Unidade Operacional Escola Antônio Simões, Manaus, Brasil
{pricila.souza, halison.maia, marlos.rodrigues, samir.batalha, daniel.sandoval}@am.senai.br
[2]Programa de Pós-Graduação em Engenharia Elétrica – Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
omaia@ufmg.br

Resumo: Este trabalho descreve um processo para a produção de uma aula interativa, utilizando a TV Digital Interativa como recurso didático na aprendizagem da educação profissional e tecnológica. A produção destas aulas seguiu algumas idéias relacionadas à aprendizagem significativa. Uma descrição detalhada das etapas deste processo, bem como os resultados alcançados, é apresentada no decorrer deste trabalho.

Palavras chave: TV Digital Interativa, Educação, *T-Learning*.

1. INTRODUÇÃO

A televisão é um dos meios de comunicação mais populares do Brasil, presente em 95% dos lares brasileiros segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) [1]. Desde os primeiros experimentos realizados com sucesso por engenheiros e pesquisadores conduzidos por John Baird (Inglaterra) e Kenjiro Takayanagi (Japão) na década de 1920, a televisão já passou por várias mudanças, tanto no campo tecnológico quanto no cotidiano dos telespectadores [2].

Atualmente, em vários palses do mundo, o sistema de televisão está migrando para uma nova tecnologia, denominada TV Digital (TVD), que além de proporcionar melhora na qualidade de imagem e de som, possibilita a interação com os conteúdos exibidos através de aplicações desenvolvidas para tal.

Quando falamos de interatividade no ambiente televisivo tradicional, podemos levar em consideração somente o ligar e desligar da televisão, mudar de canal e aumentar ou diminuir do volume. Entretanto, hoje falar de interatividade na TVD é fundir conceitos da TV tradicional com tecnologias de computação, de forma a permitir que o telespectador interaja no que está assistindo.

Diante deste cenário, observamos que a TV Digital Interativa (TVDI) tem um grande potencial para o desenvolvimento de serviços no âmbito educacional. A interatividade está relacionada à possibilidade de aliar conteúdo com aplicações, o que estende o papel da TV na relação com o telespectador. As capacidades computacionais dos receptores de TVD (*Set-Top Boxes* – STBs) permitem que os mesmos possam receber, juntamente com o conteúdo da emissora, aplicações.

A relação unidirecional emissora-telespectador é quebrada com a TVDI, uma vez que é possível realizar o caminho inverso da interação, ou seja, o telespectador é convidado a enviar informações à emissora, por meio do canal de interatividade ou canal de retorno. Dessa forma, ao usuário da TVDI pode ser disponibilizada uma vasta gama de serviços, desde aplicações típicas da internet (compra de produtos ou transações bancárias) até específicas do ambiente de TV (participar de uma votação em um programa).

É dentro da vasta gama de aplicações para TVDI que se inserem as aplicações educacionais. Desde a implantação do Sistema Brasileiro de TV Digital (SBTVD), a produção de unidades receptoras de baixo custo é incentivada, com o objetivo de usar o potencial dos STBs para promover serviços para locais remotos, principalmente serviços de governo e educação à distância. Na literatura, existe uma série de trabalhos que analisam o potencial da TVDI para aplicações educacionais [3-6].

Considerando as tecnologias disponíveis para a TVDI brasileira, o *middleware* Ginga, uma plataforma para a programação de serviços e aplicações interativas, está ganhando impulso, e suas características são bastante adequadas para a criação de aplicações educacionais. O Ginga, em sua versão mais estável Ginga-NCL, fornece um ambiente para a apresentação e sincronização de aplicações com conteúdo multimídia.

Para um ambiente educacional, a utilização de material multimídia como parte do processo de ensino, encontra na TVDI baseada em Ginga, uma aplicação bastante relevante: é possível, por exemplo, montar um ambiente de distribuição de aplicações interativas que avaliam o desempenho do aluno em determinado conteúdo de aula assistido, fornecendo resultados relevantes ao professor.

A proposta deste trabalho é explorar a TVDI brasileira para o desenvolvimento de um recurso pedagógico direcionado ao processo de ensino e aprendizagem de jovens e adultos da educação profissional e tecnológica, permitindo que o aluno entre em contato com uma gama de materiais multimídia relacionados às aulas.

O restante do trabalho está organizado da seguinte maneira: uma descrição do projeto, denominado SENAI/AM Interativo, é apresentada na Seção 2; o processo de elaboração do ambiente pedagógico para a criação de uma aula utilizando a TVDI é mostrado na Seção 3; e as considerações finais e perspectivas de trabalhos futuros são discutidas na Seção 4.

2. DESCRIÇÃO DO PROJETO

O Projeto SENAI/AM Interativo partiu da iniciativa de criar um grupo de docentes para explorar a interatividade da mídia TV. Ele tem como objetivo promover aos cursos técnicos da Rede SENAI uma alternativa inovadora de apoio ao processo de ensino e aprendizagem aos conteúdos ministrados em sala de aula. Esta iniciativa é apoiada pela instituição através do Programa de Inovação com Tecnologias Educacionais (PITE) do Serviço Nacional de Aprendizagem Industrial – Departamento Nacional.

O PITE formenta uma reflexão sobre as tecnologias educacionais, inspirando trajetórias para aprimorar a capacidade institucional de diagnóstico, sistematização, estruturação e incentivo às ações que promovam a inovação com uso de tecnologias na educação profissional. Para isto, o PITE oferece recursos para o uso de tecnologias na formação de novos profissionais, visando à aprendizagem, o conhecimento e à acessibilidade de novas tecnologias.

O projeto teve início com a aquisição de equipamentos para montar uma infraestrutura tecnológica, a qual possibilitasse a produção das aulas, o desenvolvimento do aplicativo interativo e a transmissão da aula interativa. Em paralelo a isto, o grupo de docentes teve capacitação profissional para o desenvolvimento do projeto.

Após estas atividades, o grupo elaborou o planejamento das atividades a serem desenvolvidas na sala de aula com o uso da TVDI como meio facilitador da aprendizagem dos alunos.

3. CRIANDO A AULA INTERATIVA

A produção das aulas interativas teve como ponto de partida a criação das mini-aulas interativas, conjunto de vídeos no padrão SBTVD sobre uma unidade curricular contida em um curso da Escola. O propósito destas mini-aulas é promover a interação do aluno com a unidade curricular ministrada pelo professor em sala de aula por meio da TVDI, possibilitando aos alunos uma visão macro de todo conteúdo.

Cada um dos conceitos apresentados na unidade curricular está mapeado através de vídeos interativos (mini-aulas), que são detalhados em capítulos. Esse detalhamento se dá em vídeos e aplicações interativas bem distintas, de forma que cada um dos conceitos abordados pela unidade curricular é tratado separadamente. Ou seja, para cada item da unidade curricular, é disponibilizado o conteúdo interativo correspondente.

Essa separação, que acontece durante aulas com o uso do recurso interativo, pode ajudar o aluno a dominar todos os tópicos abordados em uma unidade curricular nas suas minúcias, que é a meta almejada quando se está estudando cada um dos componentes e mecanismos de uma unidade curricular.

Cada aula, que utiliza esse recurso de interatividade, é baseada em pesquisa, revisão de assuntos abordados, intensa participação dos alunos, e, sobretudo, interação entre aluno, tecnologia e professor.

Trabalhar o mapeamento de uma unidade curricular através de vídeos dentro de um recurso de interatividade é uma forma de representar o conhecimento, sendo norteada pela proposta de Novak, que tem sua fundamentação teórica baseada na teoria de aprendizagem significativa de David Ausubel [7-8]. Na aprendizagem significativa, o professor estabelece uma relação do conteúdo que será ministrado com o conhecimento aprendido pelo aluno, que pode ser através de imagens, conceitos ou proposições [9]. Uma das formas de avaliar o processo de aprendizagem dos alunos é utilizando mapas conceituais, os quais representam a organização dos conceitos aprendidos graficamente, além de observar dificuldades de assuntos que não foram compreendidos pelo aluno [10].

Assim, os vídeos interativos representam os relacionamentos entre conceitos na forma de proposições, que são dois ou mais conceitos ligados por palavras em uma unidade semântica. Eles servem para tornar tarefa do aprendizado mais clara para os alunos e professores. Pode-se dizer que o mapeamento da unidade curricular através de vídeos e aplicações interativas podem tornar explícito o conhecimento interno de uma forma visual, que pode ser facilmente examinada e compartilhada.

O uso desses recursos multimídia, como técnica de ensino, tem o intuito de auxiliar na compreensão de como o conhecimento se organiza e se transforma na mente do aluno e como ele pode ser envolvido num processo cognitivo durante a aprendizagem. Realizar uma estruturação hierárquica de conceitos através da tecnologia da TVDI pode ajudar ao aluno e ao professor evidenciar os conceitos-chave ou as proposições que serão aprendidas, ao mesmo tempo em que sugerem conexões entre os novos conhecimentos tanto dos alunos quanto do professor.

O professor pode utilizar esse mapeamento para determinar que caminhos seguir para organizar significados para ele e para os alunos, assim como para apontar possíveis equívocos nessa organização, além de separar as informações significativas e contribuir na melhoria do processo avaliativo seguindo a metodologia por competências utilizada no SENAI, que baseia-se na constituição de perfis e currículos, na condição de estratégia viabilizadora de inserção profissional. Pressupõe ruptura com alguns conceitos e práticas educacionais que não favorecem uma aprendizagem verdadeiramente significativa.

Para a produção das aulas interativas, foi escolhida a unidade curricular Eletricidade Básica, pois a mesma está presente em quase todos os cursos técnicos ofertados na Escola SENAI Antônio Simões. Esta unidade foi dividida em dez aulas interativas, as quais seguiram as etapas: (1) gravação da aula; (2) desenvolvimento do aplicativo interativo; e (3) transmissão da aula interativa (ver Figura 1).

3.1. Gravação da Aula

A unidade curricular foi mapeada em capítulos. Cada capítulo foi designado como um vídeo. A partir disto, a equipe desenvolveu um roteiro para cada vídeo, gerando assim uma mini-aula para cada roteiro. Uma mini-aula consiste em saber apresentar ao telespectador um vídeo técnico de forma agradável, atrativa e interessante. Esse foi na verdade o primeiro passo da produção áudio-visual exigida no trabalho. Diante dessa tarefa, a equipe percebeu que seria necessário dividir etapas para a produção de um vídeo no padrão SBTVD. Também foram observadas que antes de começar as etapas de filmagem, a equipe deveria determinar com clareza a pesquisa e a roteirização.

A pesquisa refere-se sobre qual assunto cada mini-aula iria abordar. Assim, cada mini-aula representa, respectivamente, um dos tópicos de interesse da unidade curricular. Por exemplo, a primeira mini-aula tem como temática "Tensão Elétrica e Corrente Elétrica". Outro elemento importante foi a determinação do público que iria assistir essa mini-aula. Neste caso, o público seria uma turma do curso técnico em automação industrial, constituída de 20 alunos.

Na roteirização, preocupou-se em estabelecer com o seu leitor uma relação mais parecida possível com a relação do seu aluno assistindo uma mini-aula. Os roteiros seguiram alguns princípios básicos, mas não foi esquecido que as mini-aulas nada mais são do que vídeos técnicos, cujo roteiro deve possuir precisão de informações, pois só assim darão credibilidade ao produto final. Dessa forma, foram estipuladas algumas regras básicas para a confecção dos roteiros:

- Apresentar as idéias de forma original e criativa.
- Redigir o roteiro de forma simples, breve e eficiente.
- Transmitir a mensagem no vídeo por meio de conteúdos (idéias ou conceitos que você quer transmitir) e formas (símbolos que representam as idéias – palavras e ilustrações).
- Em cada mensagem, sempre tratar de uma idéia por vez.
- Ao introduzir uma expressão técnica, fazer da seguinte maneira: repetir a expressão com frequência, estabelecendo paralelos e citando exemplos.
- Não introduzir com palavras complicadas. Não abreviar nada no roteiro.
- Não usar palavras com significado impreciso, como quase, bastante, muito, grande.
- Repetir as idéias importantes.
- Usar a ordem lógica da frase: sujeito-verbo-objeto.
- Sempre que possível, dirigir-se diretamente ao público.

Os roteiros foram separados em elementos textuais: divisão de cenas claramente indicadas; narração de toda a ação do vídeo; breve descrição dos personagens, cenários quando eles aparecem pela primeira vez; falas (diálogos e textos de narração) completas e destacadas do restante do texto; e rubricas ou indicações para os atores durante as falas. Na Figura 2 é ilustrado um fragmento de um dos roteiros elaborados.

O PROFESSOR fala: "Já sabemos que a Eletricidade só se torna útil se executamos algum trabalho com ela. Mas, para isso a eletricidade tem que ser controlada e convertida em outras formas de energia"

(APARECE A IMAGEM DE UMA PESSOA TOCANDO UMA CAMPAINHA, UMA IMAGEM DE UMA PESSOA LIGANDO UM APARELHO DE DVD PARA ASSISTIR UM FILME E POR FIM UMA IMAGEM DE UMA PESSOA LIGANDO PARA O HOSPITAL E DE REPENTE APARECE NA RUA UMA AMBULÂNCIA COM UM BARULHO BEM ALTO DE SIRENE. ESTE BARULHO VAI DIMINUINDO E VOLTA PARA O PROFESSOR).

Fig. 2. Fragmento de um dos roteiros elaborados

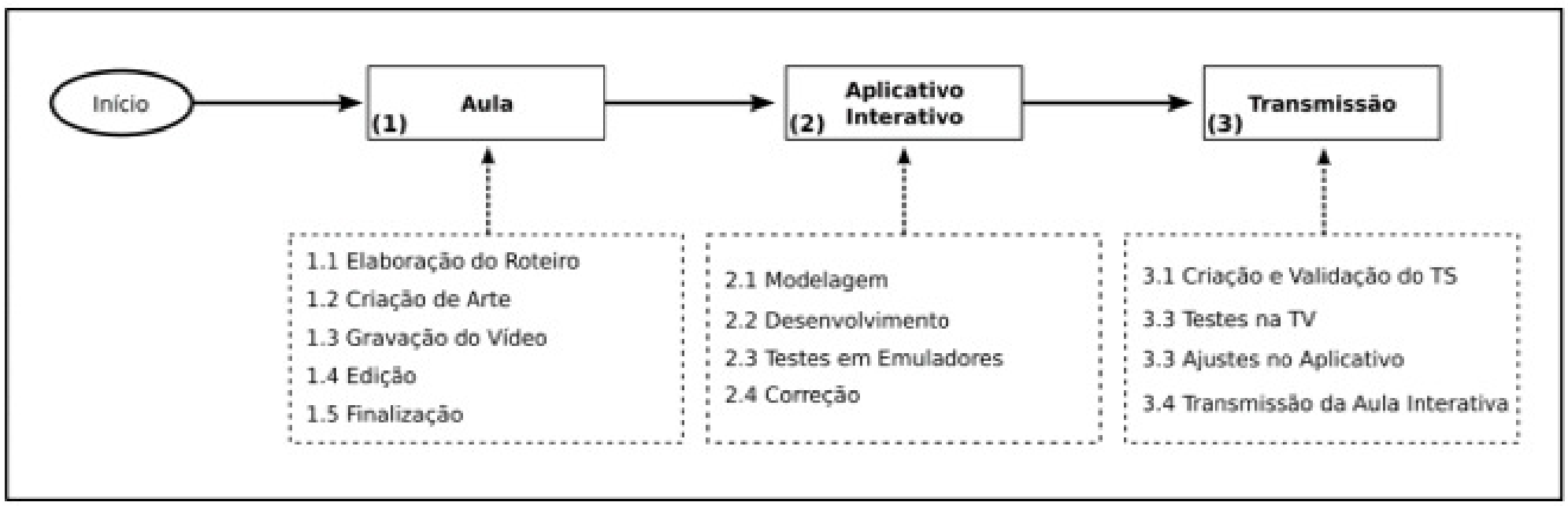

Fig. 1. Metodologia para a criação da aula interativa

Após a elaboração dos roteiros, avançou-se para a criação da arte tanto para o aplicativo interativo quanto para o vídeo. A criação da arte para o aplicativo interativo levou em consideração o uso de uma interface simples e intuitiva em que o aluno aprendesse a relacionar os botões do controle remoto da TVDI com o uso das cores da tela. A relação entre as telas, cores e posições foram estudadas para causar o mínimo de interferência em relação à aula, com base nas dicas de usabilidade para TVDI apresentadas em Hansen [11]. Ao mesmo tempo, foi pensado para ter um acréscimo de interesse, que convida o aluno a se fazer presente. A última versão da arte do aplicativo interativo é mostrada na Figura 3.

Fig. 3. Versão final da interface do aplicativo interativo

Depois disto, seguiu-se para a gravação e edição dos vídeos. Em cada gravação, foram realizadas as seguintes atividades: (1) ajustes no estúdio de gravação, tais como cenário, iluminação, ajuste de imagem na câmera (balanço de cores, foco, enquadramento/zoom) sonorização, ajuste no *teleprompter* e maquiagem; (2) filmagem; (3) validação da filmagem; (4) refilmagem; (5) edição; (6) captura dos vídeos; (7) organização de acordo com etapa de roteirização; (8) edição do vídeo de acordo com a linha do tempo; e (9) cortes das cenas;

Por fim, na pós-produção foram corrigidas as cores, utilizados *motion graphics*, aplicados efeitos especiais, realizados o Chroma e finalizados a renderização dos vídeos com as correções. Na Figura 4 é mostrada uma mini-aula após a etapa de pós-produção.

Fig. 4. Mini-aula após a etapa de pós-produção

3.2. Desenvolvimento do Aplicativo Interativo

Inicialmente, realizou-se uma dinâmica entre os integrantes do grupo para definir as funcionalidades do aplicativo interativo. Nesta dinâmica, foram levadas em consideração as peculiaridades das linguagens *Nested Context Language* (NCL) e Lua, como a sincronização do assunto apresentado na mini-aula com os conteúdos adicionais encontrados no aplicativo interativo. Como resultado desta dinâmica, foram definidos quatro tipos de conteúdos: Apresentação, Fique Ligado, Resumão e Quiz.

Na Apresentação é mostrada uma mensagem de boas vindas ao aluno, informando a unidade curricular que ele está assistindo e qual o assunto da aula (referência de acordo com Capítulo). No Fique Ligado são apresentados textos complementares e curiosidades sobre a aula. Um resumo do conteúdo abordado tanto em sala de aula quanto na mini-aula é exibido no Resumão.

Ao final de cada mini-aula, o aluno é convidado a participar de um Quiz, um teste de conhecimento sobre o assunto ministrado. O Quiz é comporto por quatro perguntas com três opções, onde somente uma é a correta. As respostas de cada aluno são armazenadas em uma base de dados localizada em um servidor. A partir destas informações, o professor pode fazer uma avaliação a cerca do conhecimento aprendido pela turma.

O aplicativo interativo é exibido desde o início da mini-aula, através de uma imagem com o texto conteúdo interativo localizada no canto superior direito da tela. Ao ser pressionado o botão vermelho do controle remoto, a tela de tela de boas vindas é apresentada ao aluno. Logo abaixo desta tela, estão presentes quatro ícones referentes à, respectivamente, Quiz (azul), Resumão (amarelo), Fique Ligado (verde) e Sair (vermelho), que podem ser acionados pelos botões de cores correspondentes do controle remoto.

A sincronia entre o vídeo da mini-aula e o aplicativo interativo foi definida no arquivo NCL, assim como a interação entre os botões coloridos e numéricos do controle remoto com os tipos de conteúdos. A navegação entre os conteúdos também foram descritos no arquivo NCL.

A linguagem Lua foi utilizada para exibir os textos dos conteúdos, os quais foram armazenados em um arquivo em formato *eXtensible Markup Language* (XML), utilizando a biblioteca LuaXML [12]. No Quiz, a linguagem Lua foi empregada para validar as respostas dos alunos. Estas respostas são enviadas para o servidor através da biblioteca TCP Lua [13].

Uma vez desenvolvido o aplicativo interativo, com seus conteúdos e interações, o mesmo foi testado na máquina virtual do Ginga-NCL [14], a qual simula um ambiente encontrado em um STB que contém o *middleware* Ginga. Nestes testes, foi verificada a conexão dos botões do controle remoto com cada conteúdo, além de averiguar o posicionamento e tamanho dos textos na interface.

Os erros encontrados na interação dos conteúdos, bem como as falhas de posicionamento dos textos, eram anotados e enviados para a pessoa responsável pelo desenvolvimento do aplicativo interativo. Assim que os erros eram corrigidos, o aplicativo era testado novamente na máquina virtual. Na Figura 5 é ilustrado o aplicativo interativo sendo testado na máquina virtual Ginga-NCL.

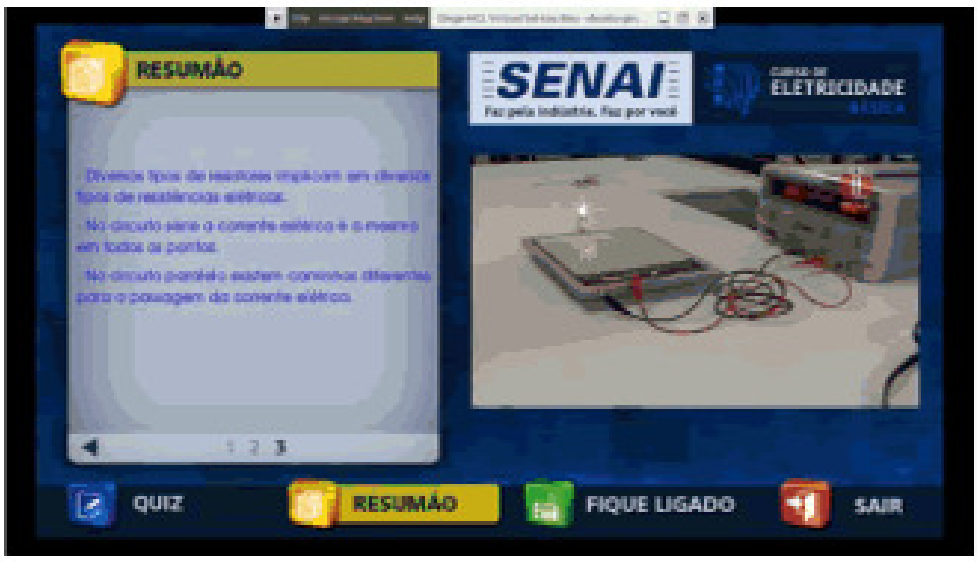

Fig. 5. Aplicativo interativo em testes na máquina virtual

3.3. Transmissão da Aula Interativa

Após testes e correções no aplicativo interativo, o mesmo foi entregue para a transmissão via *Broadcast*. Nesta fase, para cada aplicação interativa, existe uma mini-aula correspondente. Por exemplo, a Aplicação Interativa 1 é associada à mini-aula do Capítulo 1.

Assim, as mídias (áudio e vídeo) referentes a cada mini-aula foram codificadas no padrão MPEG-2. Em seguida, as mídias foram multiplexadas em um fluxo comum denominado *Transport Stream* (TS). Ao fim deste processo, o TS foi armazenado na estação de transmissão, representado pelo EITV Playout [15], a fim de verificar a sua integridade. Uma vez validado o TS, o mesmo fica disponível para a transmissão da mini-aula.

O próximo passo foi a integração do aplicativo interativo com o vídeo da mini-aula. Para isso, todas as mídias (figuras e arquivos NCL/Lua) do aplicativo foram inseridas em um arquivo compactado. Após a identificação do arquivo NCL principal, o arquivo compactado foi armazenado na estação e identificado como um serviço, o qual foi associado ao TS correspondente.

Ao final destas etapas, foi realizado um teste de transmissão pela estação da aula interativa para um STB, representado pelo EITV Developer Box [16]. Durante esta recepção, observam-se possíveis problemas relacionados exibição dos conteúdos e, caso sejam encontradas, são realizados ajustes finais tanto no aplicativo interativo quanto na mini-aula. Esta verificação é de extrema importância, pois as cores da interface ou os textos podem sofrer alterações quando são mostrados em uma televisão. Além disso, pode-se analisar a usabilidade do aplicativo interativo por meio da interação com o controle remoto. Por fim, a versão aula interativa é transmitida (ver Figura 6).

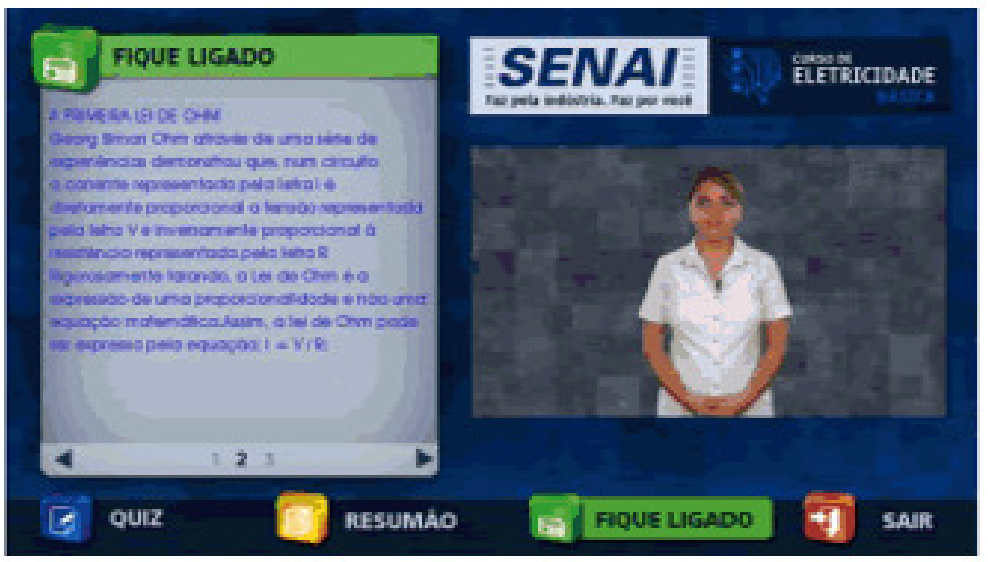

Fig. 6. Aula interativa transmitida para a TV

4. CONSIDERAÇÕES FINAIS

Neste trabalho foi descrita uma metodologia para a criação de uma aula interativa, utilizando a TVDI como um recurso para o ensino e a aprendizagem de jovens e adultos da educação profissional e tecnológica.

A produção destas aulas interativas partiu de conceitos relacionados à aprendizagem significativa, onde foi definida uma relação de vídeos interativos com os tópicos de uma unidade curricular para representar o conhecimento a ser compreendido pelos alunos.

Durante cada etapa da produção, foram estipulados objetivos a serem alcançados, desde a criação dos roteiros para cada mini-aula, passando pela criação das interfaces e desenvolvimento do aplicativo interativo, até a transmissão da aula interativa.

Como trabalho futuro, espera-se em médio prazo avaliar este processo de criação da aula interativa em uma turma piloto de Eletricidade Básica do curso técnico de Automação Industrial da Escola SENAI Antônio Simões do Departamento Regional do Amazonas, além de fornecer parâmetros para analisar como a TVDI influencia no processo de ensino e aprendizagem tanto para professores quanto para alunos.

AGRADECIMENTOS

Agradecemos ao Serviço Nacional de Aprendizagem Industrial – Departamento Nacional e Departamento Regional do Amazonas, pelo apoio financeiro através do Programa de Inovação com Tecnologias Educacionais (número 844 da unidade organizacional 20100206 – DR – AMAZONAS).

REFERÊNCIAS

[1] IBGE. Pesquisa Nacional por Amostras de Domicílios – Síntese de Indicadores de 2009. Disponível em: <www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2009/pnad_sintese_2009.pdf>. Acesso em: 15 maio 2012.

[2] MONTEZ, C.; BECKER, V. TV Digital Interativa: conceitos, desafios e perspectivas para o Brasil. 2 ed. Florianópolis: Ed. da UFSC, 2005. 201 p.

[3] D. S. MARTINS, L. S. OLIVEIRA. M. G. C. PIMENTEL. "Designing the user experience in iTV-based interactive learning objects". Anais: Proceedings of the 28th ACM International Conference on Design of Communication (SIGDOC '10). São Paulo: Universidade de São Carlos, 2010.

[4] PRATA, A. N.; CHAMBEL, T. Going beyond iTV: designing flexible video-based crossmedia interactive services as informal learning contexts. Anais: Proceddings of the 9th international interactive conference on Interactive television (EuroITV '11). Lisboa: Universidade de Lisboa, 2011.

[5] VERAS, D.; IBERT, I.; BARROS, H.; SILVA, M.; COSTA, E. A solution for personalized t-learning applications integrated with a web educational platform. Anais: Proceedings of the 2011 ACM Symposium on Applied Computing (SAC '11). TaiChung: Tunghai University, 2011.

[6] SANTOS, M. C.; VALENTE, V. C. P. Modelo de aplicação interativa para middleware ginga de televisão digital: ambiente virtual de aprendizagem para educação a distância. Revista de Radiodifusão, v.5, n.5, p. 54-61, 2011.

[7] AUSUBEL, David P.; NOVAK, Joseph D.; HANESIAN, Helen. Psicologia educacional. 2. ed. Rio de Janeiro: Interamericana, 1980. 623 p.

[8] AUSUBEL, David P. Aquisição e retenção de conhecimentos: uma perspectiva cognitiva. Lisboa: Plátano Editora, 2003. 219 p.

[9] RONCA, A. C. C. Teorias de ensino: a contribuição de David Ausubel. Temas em Psicologia, Ribeirão Preto, v. 2, n. 3, p. 91-95, Dezembro 1994.

[10] ARAÚJO, A. M. T.; MENEZES, C. S.; CURY, D. Apoio automatizado à avaliação da aprendizagem utilizando mapas conceituais. Anais: XIV Simpósio Brasileiro de Informática na Educação. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2003.

[11] HANSEN, V. Designing for interactive television. Disponível em: <www.mhp.org/docs/itv design_v1.pdf>. Acesso em: 15 maio 2012.

[12] CHAKRAVARTI, P. LuaXML – a native XML parser for Lua. Disponível em: <http://viremo.eludi.net/LuaXML/>. Acesso em: 28 maio 2012.

[13] SANT'ANNA, F. Biblioteca para realização de conexões TCP. Disponível em: <www.telemidia.puc-rio.br/~francisco/nclua>. Acesso em: 28 maio 2012.

[14] GINGA-NCL VIRTUAL STB. Disponível em: <www.gingancl.org.br/sites/gingancl.org.br/files/ferramentas/ubuntu-server10.10-ginga-v.0.12.4 i386.zip>. Acesso em: 28 maio 2012.

[15] EITV Playout. Disponível em: <www.eitv.com.br/playout.php>. Acesso em: 28 maio 2012.

[16] EITV Developer Box. Disponível em: <www.eitv.com.br/devbox.php>. Acesso em: 28 maio 2012.

MAPS-TV: Uma solução de apoio para a geração dinâmica de conteúdo para a TV Digital Interativa Brasileira

Izequiel Norões (izequiel@gmail.com)
Elizabeth Furtado (elizabeth@unifor.br)
Vasco Furtado (vasco@unifor.br)

MAPS-TV: Uma solução de apoio para a geração dinâmica de conteúdo para a TV Digital Interativa Brasileira.

Izequiel Norões [1], *Elizabeth Furtado* [2], *Vasco Furtado* [3]

[1] Laboratório de Engenharia do Conhecimento - UNIFOR, Fortaleza, Brasil, Izequiel@gmail.com
[2] Laboratório de Usabilidade e Qualidade de Software - UNIFOR, Fortaleza, Brasil, elizabeth@unifor.br
[3] Laboratório de Engenharia do Conhecimento - UNIFOR, Fortaleza, Brasil, vasco@unifor.br

Resumo: Este trabalho tem como principal objetivo propor uma arquitetura de software para desenvolvimento e uso de mapas colaborativos aplicados ao Modelo do Sistema Brasileiro de Televisão Digital (SBTVD), chamada Maps-TV. Serão descritos os módulos da arquitetura e como foi estendido o modelo de representação de conteúdo TV-Anytime para permitir o uso mapas criados na Web. Para demonstrar o processo proposto na arquitetura será apresentada uma forma de criação de mapas colaborativos sem que haja a necessidade da participação de desenvolvedores de software.

Palavras chave: mapas colaborativos, convergência, TV digital interativa, geração dinâmica de aplicações.

1. INTRODUÇÃO

A TV Digital Brasileira foi um marco tecnológico, que vai além da imagem de alta definição, sendo a interatividade uma das vantagens que este novo padrão dispõe. Esta nova maneira de interação possibilita o uso de software embarcado na TV e com isso a criação de um novo mercado e possibilidades de desenvolvimento de aplicações. A TV Digital Interativa (TVDi) desponta como uma nova mídia de interação e produção de conteúdo.

Paralelo a isso, temos a explosão do uso de conteúdo na Internet, onde a exploração dos mais diversos meios de interação para acesso/divulgação de conteúdo se multiplicam. Passam a existir aplicações focadas em compartilhamento de informações (como Orkut, Instagram, Facebook, Foursquare, G+, entre outros), proporcionando cada vez mais uma maior participação de seus usuários por meio da colaboração.

Podemos categorizar ou mesmo segmentar as opções oferecidas por aplicações colaborativas, em funcionalidades que suportam a participação das pessoas por meio do uso de mapas. Em [3], os autores classificam as funcionalidades de ferramentas de criação de Mapas Colaborativos [3] como de interação e de colaboração. Para os autores, a noção de interação se refere a como as conexões são formadas por indivíduos com interesses, valores e objetivos comuns para o compartilhamento de informações, daí vem a necessidade de colaboração de informações diante da propagação dos meios que existem na atualidade e do aproveitamento de dados de geolocalização, para mapear tais informações.

Voltando ao contexto da TVDi, o SBTVD [16] com o uso do Ginga [8], se apresenta como um potencial um espaço de interação que pode fomentar o processo de colaboração da web social, suportando assim o conceito de TV-Social.

Neste artigo resolvemos viabilizar o uso de mapas colaborativos na TVD em programas televisivos, os quais podem ser manipulados pelos telespectadores que assistem a tais programas. Definimos uma arquitetura para apoiar o desenvolvimento de aplicações para o padrão SBTVD, considerando que os mapas podem ser gerados por alguma ferramenta de criação de Mapas Colaborativos já existente. Essa arquitetura, chamada MAPS-TV, define um processo que ocorre pela dissociação e modularização de atividades de desenvolvimento de software e de produção de conteúdo dos mapas. A dissociação dessas atividades nos obrigou a definir um modelo para representar as características inerentes ao conteúdo de mapas que são comunicadas entre os diferentes módulos da arquitetura. Esta representação nos permitiu propor uma extensão do modelo utilizado no padrão TV-Anytime (TVA) [22], com o objetivo de promover a ligação das informações entre as aplicações envolvidas na arquitetura.

Para o propósito de mostrar a viabilidade da nossa proposta, a arquitetura MAPS-TV inclui novos serviços para comunicação com a ferramenta WikiMapps [3] de criação de mapas colaborativos na web. O objetivo é demonstrar como MAPS-TV facilita a criação de uma aplicação colaborativa de mapas para TVDi usando uma estrutura semelhante a um sistema de gerenciamento de conteúdo na web. A contribuição desejada é a criação de aplicações TVDi sem haver necessidade de um usuário programador de software.

2. O ECOSSISTEMA DE DESENVOLVIMENTO DE SOFTWARE PARA A TV DIGITAL

Quando o padrão SBTVD foi definido em 2007, uma janela de oportunidade foi aberta para os desenvolvedores, à luz da possibilidade de criar aplicativos baseados no Ginga [2] para utilizar plenamente o recurso de interatividade apresentado nesse padrão recém-adotado. Aplicações TVDi têm uma chance para o uso popular por causa das possibilidades do modelo brasileiro, que tem uma gama de alcance, considerando que quase todos os domicílios brasileiros têm um aparelho de TV.

2.1. O desenvolvimento de Aplicações TVDi

Empresas e universidades têm mantido seu progresso em relação ao desenvolvimento de aplicações TVDi, resultando na geração de aplicativos para fins específicos, como a aplicação T-Gov, da empresa estatal Dataprev, com seu projeto chamado "TV Digital - Social" [18]. Há também plataformas de interatividade comerciais baseadas em Ginga (middleware AstroTV), como o StickerCenter [11], abrindo o leque de aplicações T-Commerce, que além das aplicações transmitidas pelas empresas de radiodifusão, permite o download de aplicações e traz um grau de personalização para aplicações TVDi, através do uso de IPTV combinado ao Ginga.

Mesmo considerando os avanços ocorridos no desenvolvimento de aplicações TVDi, há que se considerar a capacidade de ambientes para trabalhar em qualquer padrão da TV digital e a solução encontrada para comunicar o conteúdo criado entre plataformas, como internet e TV digital. Foram pesquisadas as seguintes ferramentas para edição de aplicações TVDi: Cuíca [20], Jame Autor [7], Icareus Suite [10], Compositor [9], iTV Project [13], Ginga Contextual [4], T-Vilo Creator [15] e Astro [11] (ver Tabela 1). Apenas as ferramentas [4], [9], e [15] são específicas para o padrão brasileiro. A maioria delas usa arquivos XML para comunicar um conteúdo criado entre plataformas. Em [15], por exemplo, todas as relações entre os meios de comunicação são representados em um arquivo XML, tais como ordem e execução dos meios de comunicação, sincronização e atributos temporários, entre outros. Mas nenhuma faz uso de elementos de descrição de dados ou elementos de interface (para uso na geração de aplicações), inviabilizando o uso de mapas na TVDi gerados por qualquer ferramenta de criação de Mapas Colaborativo já existente na Web (ver coluna suporte à interoperabilidade na Tabela 1). Outros recursos também foram avaliados.

Tabela 1. Comparação entre ferramentas de desenvolvimento para aplicações TVDi

	Geração dinâmica de interface	Necessita de especialista	Edição com recursos visuais	Suporte a interope-rabilidade
Cuica	NÃO	NÃO	NÃO	NÃO
Jame Author	SIM	SIM	NÃO	NÃO
Icareus Suite	SIM	NÃO	SIM	NÃO
Composer	NÃO	SIM	SIM	NÃO
iTV Project	NÃO	NÃO	SIM	NÃO
Contextual Ginga	NÃO	NÃO	SIM	NÃO
T-Vilo Maker	SIM	SIM	SIM	NÃO
Astro	SIM	SIM	SIM	NÃO

De uma forma geral, nenhuma das ferramentas de desenvolvimento de aplicações TVDi estudadas satisfaz as necessidades apresentadas, tornando assim necessário o desenvolvimento de uma arquitetura para promover a geração de aplicações com maior facilidade, especialmente para usuários não especialistas em desenvolvimento de software.

3. ENTENDENDO A REPRESENTAÇÃO DE CONTEÚDO MULTIMEDIA PARA TV DIGITAL

Uma das vantagens da TV Digital é que ao invés de uso do espectro de transmissão apenas para transmitir o vídeo e o áudio, pode haver também o envio de dados e de aplicações TVDi, Essas aplicações podem ou não fazer o uso de dados para sua composição, ou ainda de dados complementares à programação, como a grade de programação ou mesmo a descrição de um programa televisivo. Por este motivo é interessante o uso de elementos descritores de serviços e descrição de conteúdos multimídia por meio de metadados.

Todos os sistemas atuais de TVD aberta adotam modelos de metadados que não possuem extensibilidade, como é o caso do XMLTV e as Tabelas de Informação e Serviço. Entretanto, pesquisas realizadas principalmente pelos europeus, estudam a incorporação dos modelos flexíveis (extensíveis), como é o caso dos padrões de metadados MPEG-7, MPEG-21 e TV-Anytime. Contudo esses modelos não são especificados ainda pelo Ginga e não constituem parte do conjunto de metadados e funcionalidades disponíveis para o desenvolvimento de aplicações para o SBTVD. Mas podem ser usados como referência ou base para implementação de um modelo que, uma vez adaptado, poderá ser aplicado principalmente na comunicação de dados entre aplicações WEB e TVD. É neste contexto, que este artigo se insere visando a resolver o problema da interoperabilidade já citada.

A extensão que propomos aproveita as características de descrição de elementos multimídia do TVA, mas a extensibilidade deste padrão esta restrita a elementos que ele já possua, com o acréscimo de novas características. Desta forma tomamos o elemento tabelas de informações de segmento (SegmentInformationTable), e nele adicionamos as características dos dados para geoprocessamento e colaboração por meio de mapas, como posição, usuários, comentários e etc. Essas informações fazem parte do modelo para a representação dos conteúdos, chamado MapsTVXML.

4. CRIAÇÃO DE CONTEÚDO PARA MAPAS COLABORATIVOS NA

Como dito, falta um modelo para a representação da informação que flui entre os elementos que participam do processo de geração de conteúdo em padrões para TV digital. No caso desta proposta, apresentamos a arquitetura MAPS-TV usada para permitir a geração automática de aplicações TVDi para manipular mapas via TV digital, como mostrado a seguir.

4.1. Arquitetura Básica da Aplicação

A arquitetura MAPS-TV sugere um processo de geração de aplicações dos mapas para TVDi. Para um melhor entendimento desse processo, fazemos uso do seguinte cenário: Considere a existência de um mapa de uma determinada cidade, contendo registros de poluição sonora (que caracterizam as ocorrências no mapa). Esse mapa é acessado via web pelo aplicativo WikiMappsBarulho (ver Figura 1) e foi gerado pela ferramenta WikiMapps, que neste trabalho é usado como um exemplo de um provedor de conteúdo.

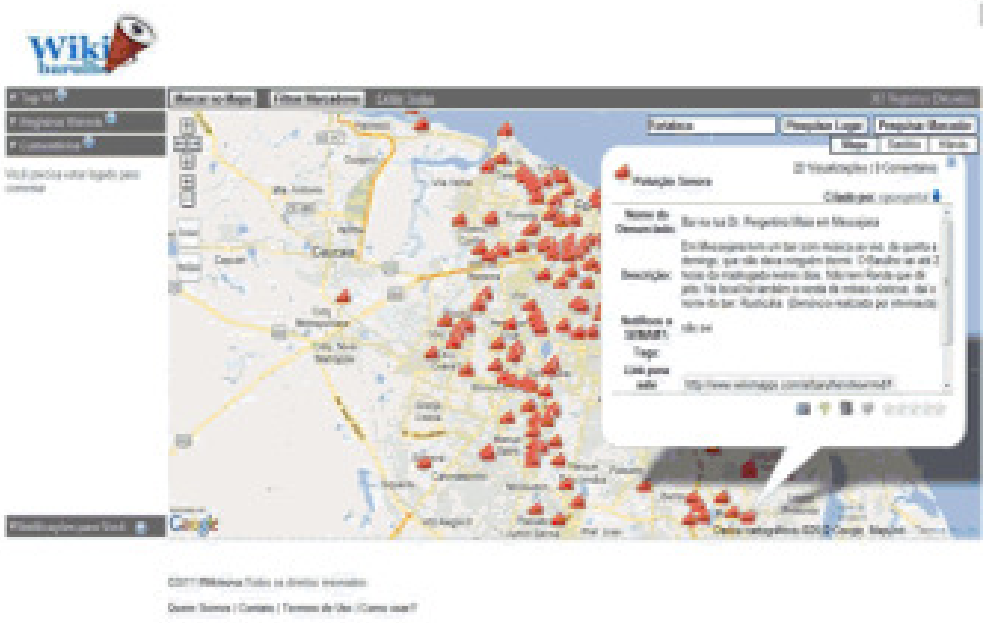

Figura 1. Ferramenta WikimappsBarulho

Uma ocorrência é configurada como um marcador na ferramenta WikiMapps e os dados transmitidos sobre a ocorrência seriam um evento associado a esse marcador. Nesse cenário o uso desse mapa em TV digital seria para referenciá-lo durante uma notícia de um telejornal local, quando o âncora do noticiário relatar algo sobre perturbação na vizinhança ou barulhos de casas de shows, por exemplo. Os telespectadores que assistem ao noticiário seriam convidados para ver mais informações sobre o fato relatado e também para participar de forma interativa com programa. Eles interagem com o controle remoto usando o aplicativo TVDi, que iria apresentar as ocorrências a partir do mapa WikiMappsBarulho.

A arquitetura proposta faz uso da base contendo as informações dos mapas e de serviços de produção e de geração de conteúdo ligado aos mapas, e de geração de uma aplicação para TVDi (usando templates - moldes pré-estabelecidos). Tais serviços foram desenvolvidos neste trabalho. Geramos uma API e a aplicamos na ferramenta WikiMapps, estendendo-a desta forma. O principal elemento de acesso a essa arquitetura é o Provedor de Serviço Interativo (PSI) [12], que é responsável por oferecer serviços aos profissionais televisivos (neste cenário o produtor do noticiário). Um serviço poderia ser a lista de mapas previamente criados pela ferramenta WikiMapps. Assim se o produtor de conteúdo quiser um mapa colaborativo para ilustrar o seu noticiário, ele deve consultar um PSI para saber da existência ou não de um tipo de mapa (barulho, de pontos da dengue, buracos, etc). O PSI lhe fornece a lista via a interface da ferramenta estendida e ilustrada na figura 2. Esta requisição é feita pelo PSI através de uma consulta aos serviços da web relacionados a este contexto. Essa figura também apresenta uma opção de pré-visualização da aplicação na TVD antes mesmo da sua transmissão.

Esses serviços comunicam-se com a ferramenta WikiMapps através da utilização de metadados gerados automaticamente a partir das informações de

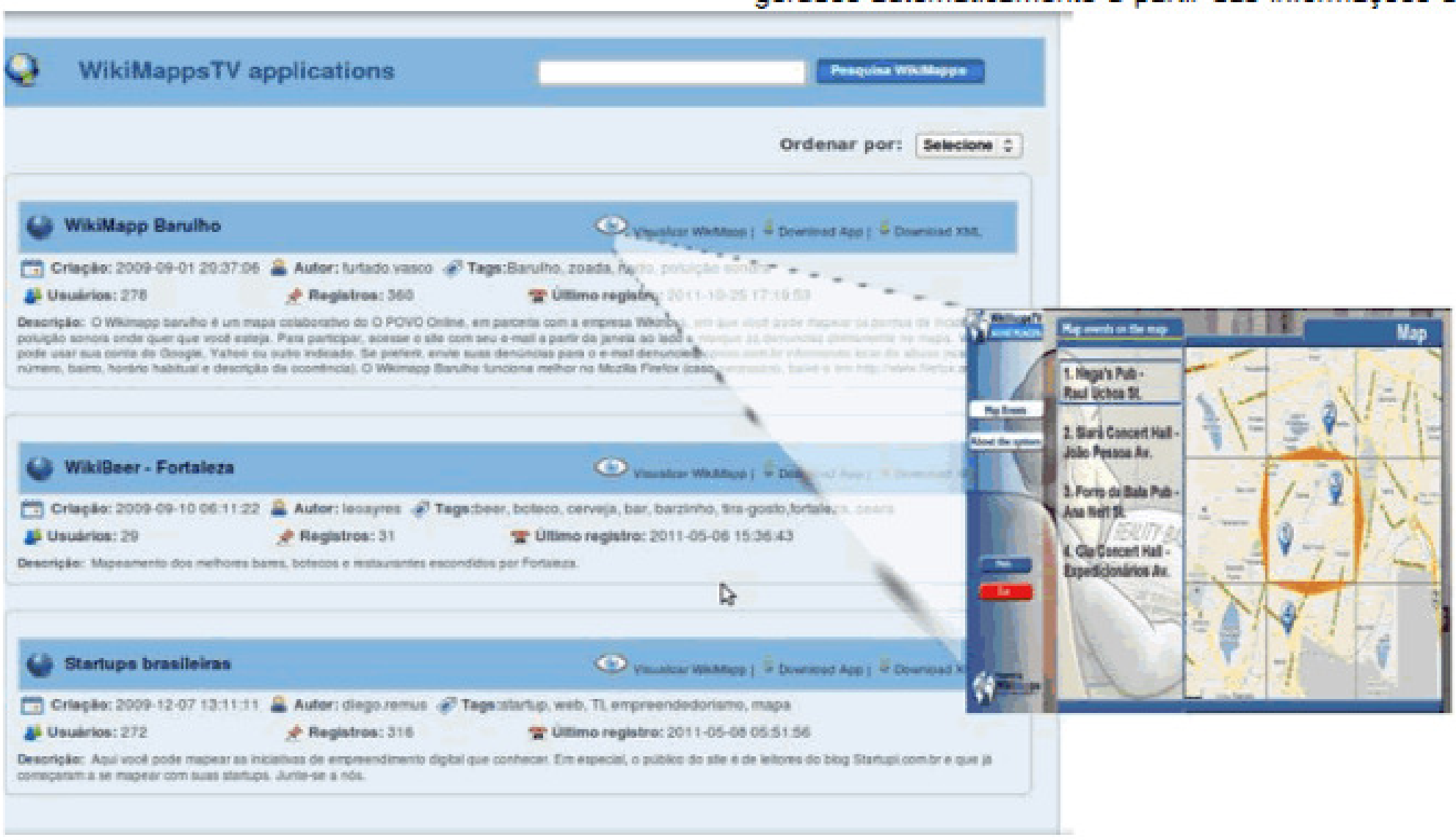

Figura 2. Interface da Ferramenta Wikimapps estendida e usada pelo produtor de conteúdo

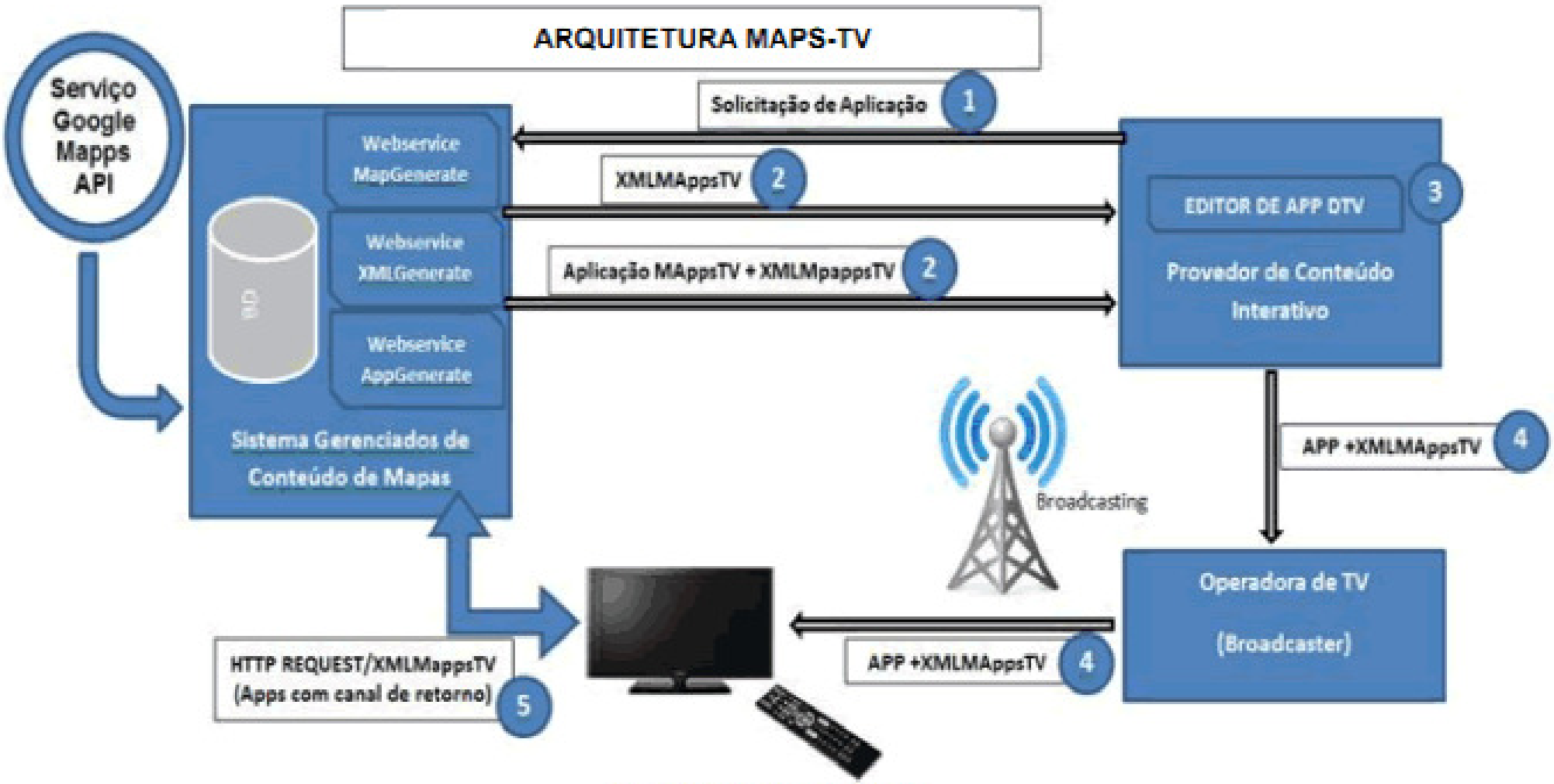

Fig. 3 Arquitetura Maps-TV

um mapa. Uma aplicação TVDi faz uso de metadados no processo de consumo e geração de informação, promovendo a interatividade e colaboração possível no que diz respeito aos mapas.

A Figura 3 ilustra os elementos que participam do processo de geração de aplicações dos mapas colaborativos para TVDi, e os fluxos que compõem tal processo (numeração de acordo com a arquitetura). Vale lembrar, que o usuário deste processo é o produtor de conteúdo.

O fluxo aqui apresentado, descreve como a arquitetura Maps-TV funciona, com o propósito de gerar aplicações TVDi, que servirão como consumidores e fornecedores de conteúdo para os mapas, indicando claramente o desacoplamento entre as etapas de geração de conteúdo e aplicativos para TVDi. Percebemos isso com base em fluxos 1 e 2 com o uso de informações localizadas no Gerenciador de Conteúdo de Mapas (chamado ferramenta de criação de Mapas Colaborativos na Figura 2).

O usuário, via o PSI, recebe o pacote de dados contendo os resultados dos serviços desse gerenciador (a aplicação para TVDi e os dados do mapa no formato XML) e solicita enviá-los para a emissora de TV. Se for necessário, ele pode personalizar a aplicação antes da transmissão (fluxo 3). Assim, o pacote de dados pode ser entregue às emissoras de TV, que farão a sua transmissão para os clientes, ou seja, para a TV dos telespectadores em seus receptores de sinal, como visto no fluxo 4 da Figura 3.

Ao final de todo processo, caso o telespectador disponha de acesso ao canal de retorno no seu aparelho receptor, poderá fazer uso de recursos de colaboração por meio da aplicação interativa, como realizar comentários sobre o assunto do mapa ou mesmo uma denúncia de incidente na área do mapa que estiver utilizando. Esta possibilidade permite além de tornar a aplicação um fornecedor de conteúdo (alimentar os mapas), expandir o uso da aplicação.

Vale salientar, que na geração da aplicação (na fonte do conteúdo) e no PSI, um programador de software perito não foi necessário. Os conteúdos do mapa foram criados usando a ferramenta WikiMapps e inseridos na aplicação TVDi, cujos possíveis formatos (templates de interface) foram previamente definidos por um programador. Esta inserção foi possível devido à aplicação compreender o modelo de conteúdo acoplado à aplicação, o qual está no formato XML, e é gerado com base no metadados proposto nesta arquitetura.

4.2. Representando conteúdo no modelo Maps-TV

Como dito, o metadados gerado na compilação dos elementos pela aplicação WikiMapps é passado para um arquivo XML, que contém as especificações do modelo XMLMapsTV. Esse arquivo representa a estrutura básica necessária para a montagem do aplicativo PSI interativo. Devido à restrição de espaço, esse metadados não está descrito neste artigo. Porém ele será apresentado na conferência internacional EuroiTV'2012 em Berlim, em julho. O modelo proposto XMLMapsTV é baseado no modelo TV-Anytime (TVA) [21], que entre os modelos existentes [17], tem uma maior extensibilidade e flexibilidade, podendo vir a ser adotado no padrão brasileiro, pelas deficiências apresentadas na seção de estado da arte.

4.3. Uso de metadados com as Ferramentas para Geração de Aplicações TVDI

Aplicações feitas para o ambiente de TV Digital usam as linguagens NCLua [19] e Java [6], e estes não são muito simples de usar por um editor de TV ou elaboradas de uma grade de programação da TV. Para que editores de TV sejam capazes de usar o contexto dos mapas de colaboração, ferramentas para geração de aplicações mapa para TVDi devem ter a capacidade para interpretar arquivos XML para obter dados para uma aplicação. Considerando que o modelo XML utilizado na proposta deste estudo pode ser adotado, essas ferramentas fazem uso de metadados, de modo a incorporar essa funcionalidade em seus recursos e podendo, portanto, aplicar-se à geração de suas aplicações.

5. VALIDAÇÃO DA PROPOSTA: GERAÇÃO DE UMA APLICAÇÃO USANDO A FERRAMENTA WIKIMAPPS

A seguir iremos descrever e ilustrar como uma aplicação foi implementada (gerada) para o padrão SBTVD, viabilizando a manipulação na TV de mapas gerados pela ferramenta WikiMapps.

5.1. Geração automática da aplicação TVDi

Nas aplicações web podemos fazer uso de Sistemas Gerenciadores de Conteúdo para proporcionar a geração de websites para as mais variadas finalidades, considerando o conteúdo e o uso de moldes (*templates*) que personalizam e caracterizam suas interfaces. Com base nesse modelo, aproveitamos o mesmo conceito e o aplicamos para a geração de aplicações de mapas para o contexto da TVDi, com a definição da API que estendeu a ferramenta WikiMapps. Pudemos gerar uma aplicação TVDi com o uso de moldes. Esses moldes são na realidade uma aplicação TVDi pré-elaborada com uma linguagem específica, neste caso NCL + Lua, que recebe os mapas gerados no provedor de conteúdo, bem como os dados relacionados ao assunto escolhido para o mapa, de acordo com a plataforma determinada para a exportação do serviço.

5.2. Usando Mapas Colaborativos na TV

Vamos retomar o cenário hipotético em que um programa de televisão de notícias local, discute problemas relacionados à perturbação sonora, como a ocorrência de barulho em clubes, casas de show ou veículos que possam estar causando problemas em determinadas ruas/áreas de uma cidade. O problema pode causar muita irritação e inquietude de várias pessoas, o que poderia tornar os cidadãos prejudicados irritados e ansiosos para relatar as ocorrências desses locais barulhentos. Durante a notícia, o mapa pode ser exibido para mostrar onde estão os locais mais barulhentos, bem como pedir aos usuários para informar novas ocorrências que eles tenham conhecimento.

A Figura 4 mostra um exemplo de como a interface exibe os eventos que ocorrem nos lugares. Cada evento vem da base de dados do provedor de conteúdo, o site WikiMapps, e pode ser selecionado pelo telespectador, para uma visão detalhada através da sua interação com o mapa. Assumindo-se também que o canal de retorno esteja disponível, outros tipos de interação podem estar disponíveis. O telespectador pode, por exemplo, traçar novos eventos no mapa, como locais próximos a sua residência e relatar a respeito dos problemas que tem conhecimento sobre o assunto.

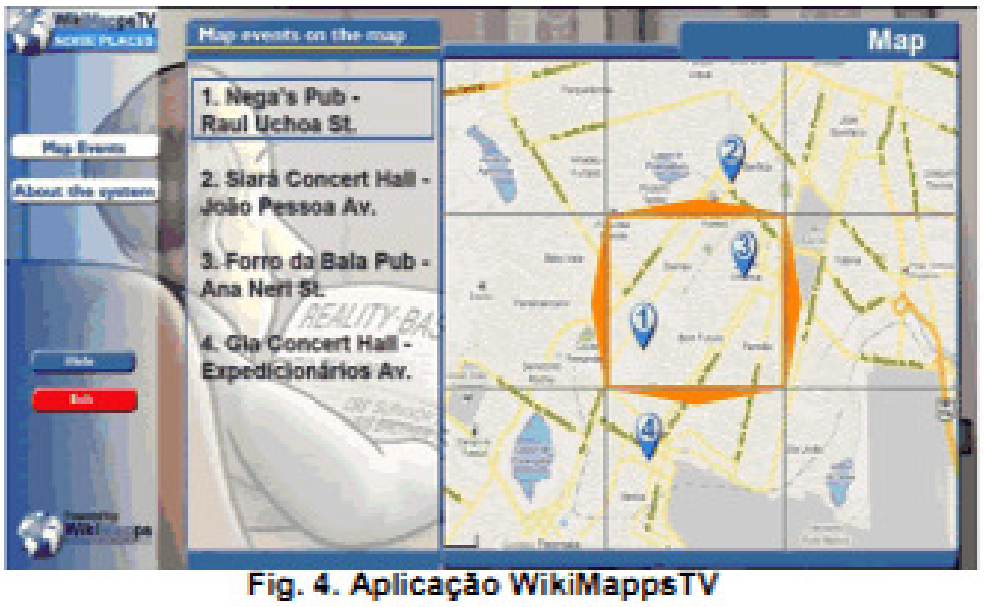

Fig. 4. Aplicação WikiMappsTV

Após a confirmação dos dados, a aplicação que está sendo executada no dispositivo receptor enviará, através de uma solicitação HTTP, o resultado da interação com o telespectador para o serviço web WikiMapps (ver fluxo 5 na Figura 3). Este resultado será armazenado e disponibilizado para o retorno informativo para o PSI ou até mesmo o operador de TV / emissora.

Vale lembrar, que a nossa proposta neste artigo é genérica para qualquer WikiMapps ou fonte de conteúdo relacionada a mapas, sendo assim possível criar aplicações de mapas para TV em diferentes domínios.

6. CONCLUSÃO

Neste artigo, propomos uma arquitetura, denominada MAPS-TV, para o desenvolvimento de um aplicativo de mapas colaborativos na TVDi, que tem como característica básica, a dissociação e a modularização das atividades de desenvolvimento e a geração de aplicações interativas por usuários não programadores de software. MAPS-TV propõe um modelo de dados para a transmissão de informações a partir dos elementos que participam na arquitetura. Ele permite a geração automática de aplicações TVDi através da ferramenta de WikiMappsTV, que oferece funcionalidades para melhorar a interação para esta finalidade.

REFERÊNCIAS

[1] Araújo, E. C. NCL Live: ambiente de autoria visual para prototipação de aplicações para TV Digital Interativa. Dissertação de mestrado. Universidade Federal do Maranhão. (2009).

[2] Barbosa, S. D. J. & Soares, L. F. G. TV digital no Brasil se faz com Ginga: Fundamentos, Padrões, Autoria Declarativa e Usabilidade. PUC RIO. Pg. 105–174. (2008).

[3] Caminha, C., Furtado, V., Vasconcelos, E.; Ayres, L. Uma ferramenta de autoria para criação de mapas colaborativos para aplicações em E-gov 2.0. In Proceedings of the XXX CSBC 2010 – WCGE, II Workshop de Computação Aplicada em Governo Eletrônico. Minas Gerais, Brasil, July, p. 20 – 23. (2010).

[4] Carvalho, A. P. B. A., Ferraz, C. A. G. 2010. Contextual Ginga: Uma ferramenta de autoria de Aplicações Interativas Sensíveis ao Contexto de TV digital para Ginga-NCL. In Proceedings of the WebMedia - Simpósio Brasileiro de Sistemas Multimídias e Web. Minas Gerais, Brazil, October, p. 5 – 8. (2010).

[5] Durand, G., Kazai, G., et al. A metadata model supporting scalable interactive TV services. Proceedings of the IEEE conference on Multi-Media Modeling. (2005).

[6] Filho, Guido Lemos S., Leite, Luiz Eduardo C., Batista, Carlos Eduardo C. F. Ginga-J: The Procedural Middleware for the Brazilian Digital TV System. J. Braz. Comp. Soc. pp. 47–56. (2007).

[7] Geuer, Olaf et al. JAME – System Concepts and New Technologies for iTV Services. Fraunhofer Institute for Media Communication IMK. (2009).

[8] Ginga Home Page. Retrieved April 11, 2012, from Official Site of Ginga Middleware: http://www.ginga.org.br.

[9] Guimarães, R. L., de Resende Costa, R. M., e Soares, L. F. G. Composer: Ambiente de autoria de aplicações declarativas para a tv digital interativa. WEBMEDIA: Brazilian Symposium on Multimedia and the Web. (2007).

[10] Icareus iTV Suite Author. Retrieved April 11, 2012, from Icareus Technology Home Page: http://www.icareus.com.

[11] Middleware AstroTV. Retrieved April 11, 2012, from TQTVD Home Page: http://www.tqtvd.com.

[12] Neves, Rogério Baldini. Explorando o canal de retorno em sistema de televisão digital interativa: uma abordagem centrada no suporte à comunicação entre aplicações e provedores de serviços. Pontifícia Universidade Católica de Minas Gerais. Belo Horizonte. (2010).

[13] Oliveira, M. R., Filho, C. B., and Fer, A. F. iTV project: An authoring tool for MHP and Ginga-J based on a web environment. In UXTV '08: Proceeding of the 1st international conference on designing interactive user experiences for TV and video, pages 179–182, New York, NY, USA. ACM. (2008).

[14] Rouse, L. J., Bergeron, S.J., Harris, T.M. "Participating in the geospatial web: collaborative mapping, social networks and participatory GIS," in Scharl, A., Tochtermann, K. (Eds.),The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, Springer, New York, NY, pp.153–158. (2007).

[15] SANTOS, Ítalo Matos Cavalcante. T-VILO TOOL: ferramenta orientada a serviços para interação e produção de conteúdos educativos baseada no sistema brasileiro de televisão digital / Ítalo Matos Cavalcante Santos. Dissertação de mestrado. Universidade de Fortaleza. (2011).

[16] SBTVD Forum Home Page. Retrieved April 11, 2012, from STBVD Forum: http://www.forumsbtvd.org.br.

[17] Silva, F. S. et al. Análise Comparativa de Metadados em TV Digital. In: 24º Simpósio Brasileiro de Redes de Computadores, SBRC, Curitiba. (2006).

[18] Silva, Marco A. M.; Castilhos, Edson L.; Eichler, Flávio A. V.; Nogueira, Roger F. 2010. T-government: O Projeto "TV Digital – Social" um "case" de Oferta de Serviços Interativos para os Cidadãos. In Proceddings of the Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Minas Gerais, Brazil, October, p. 5 – 8. (2010).

[19] Soares, Luiz Fernando Gomes; RODRIGUES, Rogério Ferreira & Moreno, Márcio Ferreira. Ginga-NCL: The Declarative Environment of the Brazilian Digital TV System. Journal of the Brazilian Computer Society. Vol. v12, pp. 37–46. (2007).

[20] Soares, P. H. S. CUICA - Uma Ferramenta para Suportar a convergência da TV Digital com a WEB através da criação de conteúdo digital. Dissertação de Mestrado. Universidade de Fortaleza. (2007).

[21] TVA. TV-Anytime Forum Home Page. Retrieved April 11, 2012, from: http://www.tv-anytime.org.

[22] Norões, Izequiel., Furtado, Vasco e Furtado, Elizabeth. Maps-TV: A Software Architecture to Support Development of Collaborative Maps in DTVi. Proceedings of the 10th European Interactive TV Conference (EuroiTV 2012), Berlim, Alemanha, julho de 2012.