

GT-IMAV: Ferramentas para instrumentação e monitoração de aplicações de vídeo

Lucas Chigami
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
lchigami@larc.usp.br

Michael Hernandez
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
michael@larc.usp.br

Reinaldo Matushima
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
reinaldo@larc.usp.br

Adriano Adoryan
Escola de
Comunicação e Arte -
Universidade de São
Paulo
adriano@usp.br

Itana Stiubiener

Universidade Federal
do ABC

itana@ufabc.edu.br

Graça Bressan
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
gbressan@larc.usp.br

Regina Melo Silveira
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
regina@larc.usp.br

Wilson V. Ruggiero
Laboratório de
Arquitetura e Redes
de Computadores –
Escola Politécnica
Universidade de São
Paulo
wilson@larc.usp.br

Abstract. Forecast pointed that video will account for 90 % of network traffic by the end of the decade. This paper presents a set of monitoring tools that allow to get data of video consumption distributed over the Internet, related to the degree of consumer knowledge absorption, their satisfaction with the video and the content's quality and your behavior towards web pages presented, as compared with the consumption of resources used. For the tools' validation, they were integrated to the video on demand system (video@RNP) and evaluated within a scenario of video being used as a tool for teaching / learning and for cultural dissemination.

Resumo. Projeções indicam que vídeo será responsável por 90% do tráfego de rede até o final da década. Este artigo apresenta um conjunto de ferramentas de monitoração que permitem levantar dados de consumo de vídeo distribuído através da Internet, tanto com relação ao grau de absorção de conhecimento do consumidor, sua satisfação com relação à qualidade do vídeo e do conteúdo e seu comportamento com relação a página web que o apresenta, como com relação ao consumo de recursos utilizados. Para que estas ferramentas fossem validadas, elas foram integradas a uma instância do sistema de vídeo sob demanda da RNP (video@RNP) e as avaliamos dentro de um cenário de uso de vídeo como instrumento para o ensino/aprendizado e para divulgação cultural.

1. Introdução

Projeções recentes indicam que o tráfego de dados na Internet pode alcançar a soma de 767 exabytes, ou 767 bilhões de gigabytes de dados em 2014 [1]. Dentro deste montante, por volta de 91% são referentes a conteúdos relacionados a vídeos. Considerando o alto custo de produção do próprio conteúdo e,

principalmente, o alto consumo de recursos computacionais e de rede realizado por aplicações que a exploram, torna-se extremamente relevante monitorar o consumo desse tipo de conteúdo.

Considerando tal cenário, o projeto GT-IMAV¹ foi desenvolvido com o objetivo de desenvolver ferramentas para a instrumentação e monitoração de aplicações de vídeo. Por instrumentação deve-se entender ferramentas ou técnicas que possam ser agregadas a uma aplicação de vídeo para poder medir e avaliar a qualidade do consumo, tanto em termos de uso de recursos de rede como em relação ao grau de envolvimento do usuário com o conteúdo e a absorção do conhecimento. Para tanto, são considerados o monitoramento dos seguintes aspectos:

- Transmissão do vídeo: obter de maneira precisa quais são as solicitações atendidas e por quanto tempo. Ter conhecimento da qualidade de serviço (QoS) de transmissão [2], como atraso, variação de atraso e banda consumida torna-se relevante para a operação e manutenção da rede de transmissão. Todas estas situações foram previstas e mapeadas para se ter uma visão mais próxima da qualidade do que o usuário está recebendo.
- Reprodução do vídeo: avaliar a reprodução do fluxo de vídeo, que pode ser influenciada por diferentes fatores, proporcionando ao usuário diferentes QoE (Quality of Experience) [2], e conseqüentemente, níveis de envolvimento com o conteúdo (“viewer engagement”)[4][3] variável de uma pessoa para outra;

¹ Projeto financiado pela Rede Nacional de Ensino e Pesquisa (RNP).

• Absorção do conteúdo do vídeo: avaliar o grau de absorção de conhecimento com o consumo destas mídias e de envolvimento com o tema abordado.

A análise destes aspectos e as ferramentas elaboradas e desenvolvidas a partir desta, podem ser utilizadas para diferentes propósitos: (1) para auxiliar na operação, manutenção e evolução da rede de transmissão; (2) para estimar a aceitação de uma produção audiovisual e mensurar seu valor de interesse; (3) para verificar o grau de satisfação do usuário, permitindo melhorar a aplicação de vídeo utilizada, baseando-se no envolvimento do mesmo e (4) para mensurar a absorção de conhecimento em cenários específicos, como por exemplo, o de ensino/aprendizado.

2. Trabalhos Relacionados

O Youtube Analytics² é uma ferramenta que ajuda os usuários que publicam conteúdo através da plataforma do Youtube a analisar os pontos altos e baixos do seu canal, através de métricas colhidas pelas visitas feitas pelos usuários. Informações como número de visualizações, quantidade de repetições de acessos a um mesmo vídeo, entre outros, são cruzadas com informações do perfil do usuário, como idade, sexo e localização.

Detalhes de como o usuário interagiu com o vídeo também são monitorados, mostrando por exemplo, se um vídeo atraiu a atenção em um determinado trecho, ou se existem muitos abandonos de um outro determinado vídeo. O engajamento também é medido através de quais vídeos os usuários realizam a inscrição do canal, além de possibilitar mapear os cliques em anotações dos vídeos.

Na área acadêmica poucos trabalhos sobre monitoramento e avaliação do consumo de vídeo foram publicados até o momento. Os trabalhos encontrados, e que serão discutidos a seguir, podem ser classificados em: estudos sobre engajamento cognitivo na audiência de TV, engajamento do usuário ao consumo de vídeos na Web e uso de vídeos no método de ensino/aprendizado.

Na área televisiva, Smith [5] faz um estudo do envolvimento e interesse do espectador com relação ao dinamismo da cena, mostrando que o aumento da frequência de mudança das cenas e associação com uma semântica humorística ou elementos inesperados, ativa o engajamento cognitivo. Já Liebes [4] categoriza o envolvimento do espectador com as telenovelas, e conclui que a identificação com algum personagem da trama faz com que este tenha um interesse maior pela novela. Outro trabalho que propõe avaliar o uso da TV, utiliza câmeras para detecção da fixação do olhar afim de mensurar o engajamento do telespectador com o conteúdo oferecido e correlacioná-lo com a absorção do conhecimento [6].

Milliken [7] estuda o potencial da ferramenta UGOV (User-Generated online Video) que permite a geração de vídeo. Tais vídeos são postados no YouTube para verificar a contribuição de seu consumo na identidade regional, nível de envolvimento e engajamento dos internautas. O engajamento neste caso é verificado a partir de um questionário que inclui, por exemplo, perguntas sobre a frequência de uso do aplicativo, quanto aprendeu, quanto um determinado vídeo mudou sua opinião sobre um determinado assunto e quantos comentários inseridos em redes sociais estavam relacionados ao vídeo em questão.

Em outro trabalho, conduzido por Bardzell [3], é feito uso de medidas fisiológicas, batimento cardíaco e respiração por minuto, para mensurar as emoções causadas pelos vídeos mais populares. São utilizadas também estratégias como tagging emocional (usuário indica seu estado emocional através da escolha de emoticons), espaço para a inclusão de um texto de review e log das ações dos usuários com relação ao vídeo (pausar, parar, navegar na linha do tempo do vídeo, etc.).

Em resumo pudemos verificar que algumas pesquisas fazem uso de instrumentos intrusivos como detectores de batimento cardíaco e respirações por minuto. Outra possibilidade não intrusiva, mas que requer dispositivo adicional (câmera) é a verificação de posição dos olhos (eyes tracking). Muitas análises são feitas baseadas em informações cedidas pelo usuário, onde este deve preencher questionários sobre seu estado emocional, seu interesse e opinião sobre o que assistiu ou ainda para demonstrar se absorveu ou não o conteúdo.

O objetivo do conjunto de ferramentas apresentado aqui é agregar as principais funcionalidades das ferramentas apresentadas com outros parâmetros mais técnicos relacionados à rede de distribuição e transmissão do vídeo, agregando valor não só ao público que é responsável pela criação do conteúdo, como também aos responsáveis pela administração do conjunto de servidores responsáveis pela transmissão da aplicação de vídeo. O potencial de correlacionamento entre estes parâmetros traz riqueza nas análises e nos relatórios que podem ser gerados.

3. Arquitetura Proposta

O projeto desenvolveu ferramentas capazes de mensurar o real uso dos fluxos de vídeo através do monitoramento de aspectos levantados dentro de uma aplicação de vídeo como:

- Comportamento do usuário diante do vídeo: através da captura da interação do usuário com o reprodutor do vídeo (player), verificando se houve pausa, retorno ou aceleração na reprodução do vídeo;
- Comportamento do usuário diante da página web: através da captura do chaveamento entre janelas e abas do browser, e movimento do mouse;
- Captura dos recursos consumidos: através da captura de informações como banda consumida, tempo de

² <https://www.youtube.com/yt/playbook/yt-analytics.html>

conexão, número de bytes transmitido pelo servidor, número de sessões simultânea no servidor;

- Percepção do usuário com relação ao vídeo: através de enquetes sobre a qualidade do vídeo, sobre a qualidade do conteúdo, e com questões sobre as informações absorvidas;

- Interação do usuário com redes sociais: capturando se o usuário compartilhou aquele conteúdo através de rede social.

Estas ferramentas são compostas por: agentes instalados junto à página web e o player que apresentam o vídeo, agentes instalados no servidor de vídeo, e um software que permite gerar enquetes, perguntas e testes. Tais ferramentas foram implementadas de acordo com requisitos funcionais e técnicos de modo a conseguirem alcançar um nível de desacoplamento que permita o uso das ferramentas em diversas aplicações de vídeo. O objetivo é que as ferramentas possam ser integradas a qualquer serviço de distribuição de vídeo, não limitando seu escopo de uso.

3.1 Módulos

Os transmissores de vídeo fornecem dados que permitem obter de maneira precisa quais são as solicitações atendidas e por quanto tempo. Também consideramos informações tal como atraso, variação de atraso e banda consumida que podem ser relevantes para a operação e manutenção da rede de transmissão. Do lado do consumidor de vídeo, obtemos informação do comportamento do mesmo durante este consumo.

Concluída a parte de captura, torna-se necessário juntar todas essas informações e armazená-las de alguma forma. Para isso, foi desenvolvido um ambiente que pudesse receber esse volume de dados, além de prover um nível de segurança e autenticação das ferramentas de coleta.

A Figura 1 apresenta uma visão dos elementos que compõem o serviço de monitoração. Os elementos que coletam as informações ficam junto aos elementos monitorados, não fazendo parte da infraestrutura específica do serviço. Dentro da infraestrutura do serviço ficam os elementos de recepção e análise de dados. Detalhando os elementos que ficariam na infraestrutura do serviço temos:



Figura 1. Visão geral da infraestrutura

1. **Serviço de recepção (aplicação web em Python):** responsável por receber as informações coletadas pelos agentes operando junto às

aplicações e persistir localmente em uma base de dados. A arquitetura prevê um ou vários serviços de recepção (dependendo do volume de informações monitoradas);

2. **Base de dados de coleta (Postgresql):** onde são persistidos os dados brutos coletados. Demanda grande desempenho, suportando a persistência dos dados recebidos pelos múltiplos receptores. Não demanda muito espaço por ser uma base de passagem;
3. **Ferramenta de tratamento dos dados (aplicação standalone – ETL (Extract Transform Load):** ferramenta que processa os dados brutos presentes na base de dados de coleta e transforma as informações, salvando-as na base de dados de mineração. Opera periodicamente de forma agendada;
4. **Base de dados de análise (Postgresql):** responsável por armazenar as informações que são acessadas pela ferramenta de análise de dados. Pode demandar suporte a grande volume de armazenamento (dependência da política de retenção das informações);
5. **Ferramenta de análise de dados (aplicação web em Java):** responsável por cruzar as informações da base de dados de mineração, extraindo os relatórios criados.

As possibilidades de correlações entre os parâmetros monitorados são inúmeras, e por esse motivo, a ferramenta de análise de dados apresenta um conjunto mínimo de relatórios e gráficos dos resultados dos principais parâmetros relevantes a uma aplicação de vídeo, possibilitando a elaboração de novas correlações, de acordo com a necessidade de cada cenário de uso das ferramentas.

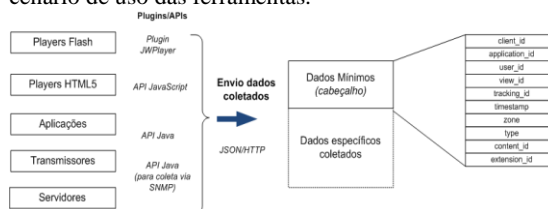


Figura 2 - Coleta e envio das informações de monitoramento

A Figura 2 ilustra o processo de coleta e envio dos dados. Nela destacamos o pacote com os dados coletados que apresenta o cabeçalho de controle, com os dados mínimos definidos e o campo com dados específicos, que permite personalização das ferramentas para diferentes cenários. Uma vez persistidos os dados, inicia-se um processo de preparação e configuração dos dados obtidos para a realização das análises sobre o uso das aplicações de vídeos. Para esta etapa foi utilizado o software livre de mineração de dados denominado Pentaho³. O Pentaho

³ www.pentaho.com

permite fazer o cruzamento dos dados obtidos e gerar gráficos e relatórios.

4. Resultados obtidos

Para a validação das ferramentas estas foram integradas a uma instância do serviço de vídeo sob demanda da Rede Nacional de Ensino e Pesquisa (RNP), o Vídeo@RNP⁴. Desta maneira foi possível elaborar diferentes cenários de uso, tanto para ensino como para divulgação cultural.

No cenário educacional, onde uma turma de alunos da UFABC (Universidade Federal do ABC) foi submetida ao monitoramento durante do uso de um material em vídeo relacionado ao tema explorado na disciplina “Redes de Computadores”. Um conjunto de testes de múltiplas escolhas presente na mesma página web em que o vídeo foi apresentado permitiu fazer uma primeira avaliação do conhecimento absorvido pelos alunos. Estes mesmos alunos também foram submetidos a uma avaliação subjetiva referente a qualidade da reprodução, assim como da qualidade do conteúdo abordado no vídeo. Uma outra turma, considerada de controle, não teve contato com o vídeo, mas foi exposto a conteúdo equivalente em aula expositiva [9].

O desempenho dos alunos da turma que assistiu o vídeo com relação ao tema explorado foi verificado em questões colocadas em uma prova (sem uso do computador), onde observou-se resultados positivos com relação a absorção do conhecimento da turma que assistiu o vídeo. Ao interpretarmos os resultados dos gráficos, vimos que a quantidade de alunos que estão no intervalo 1 é maior dentre os alunos que assistiram ao vídeo que tiveram um desempenho melhor que o outro grupo [8].

5. Demonstração

O ambiente de demonstração consiste em um portal de vídeo, onde estarão instaladas as ferramentas responsáveis pela coleta dos dados referentes ao player de vídeo e aos componentes da página, ou seja, os parâmetros coletados na máquina do usuário. Para a outra parte da coleta, realizada nos servidores responsáveis pela transmissão do vídeo, serão montadas máquinas virtuais que farão o papel do parque de máquinas que geralmente são encontrados em sistemas de distribuição de vídeo convencionais, onde as ferramentas de coleta também estarão instaladas e monitorando todo o tráfego de rede e uso de recursos computacionais.

A demonstração dessa ferramenta será feita de duas maneiras, na primeira será demonstrado o consumo de vídeo simulando ações dos usuários finais, e em um segundo momento serão mostrados gráficos e relatórios dos eventos monitorados do ambiente em

demonstração, ou seja, dados coletados através do consumo de vídeo colocados no portal para a demonstração.

6. Conclusão

O conjunto de ferramentas elaborado no projeto GT-IMAV permite monitorar o consumo de vídeo para vários cenários, e seu baixo acoplamento permite ser integrado a diversos serviços de distribuição de vídeo na Web.

Salientamos que as ferramentas desenvolvidas e que nos propomos demonstrar são para uso em ambiente de rede, instalados junto a servidores de vídeo e Web. Por este motivo as mesmas se encontram disponíveis para download e instalação. O site do projeto pode ser acessado através do endereço <http://gtimav.larc.usp.br>, onde se encontram os documentos do projeto e os manuais das ferramentas.

Uma versão da ferramenta está disponível em: <http://143.107.111.93> (portal com a ferramenta instalada) e demonstrações dos gráficos obtidos com o uso das ferramentas podem ser acessados através do endereço <http://143.107.111.92:9090> (Link do Pentaho com alguns relatórios - acessar utilizando as credenciais gtimav/102938). Um vídeo demonstrativo do uso das ferramentas e os resultados colhidos mostrados através da interface da ferramenta Pentaho podem ser visto em: http://143.107.111.93/portal/video/demo_gtimav.

7. Referências

- [1] Cisco; Entering the Zettabyte Era, White Paper, June 2011.
- [2] M. Sillar and J.C. Woods; “QoS arbitration for improving the QoE in Multimedia Transmission”; International Conference on Visual Information Engineering (VIE) 2003; pp. 238-241, June, 2003, ISSN: 0537-9989
- [3] Bardzell, S.; Bardzell, J.; Pace, T.; Understanding affective interaction: Emotion, engagement, and internet videos; 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009.
- [4] T. Liebes, “Notes on the struggle to define involvement in television viewing”; *Revue de la Recherche en Communication*, 1996, volume 4, n° 1, pp. 35-46.
- [5] Smith, M e Gevins, A.; “Attention and Brain Activity while watching television: Components of Viewer Engagement”. 2004. DOI: 10.1207/s1532785xmep0603_3
- [6] Sawahata, R e Khosla, R.; “Determining Comprehension and Quality of TV Programs using eye-gaze tracking”. Elsevier – Pattern Recognition, 2008.
- [7] Milliken, M., Kerri, G., O'Donnell, S. e Singer, J.; “User-generated Online Video and the Atlantic Canadian Public Sphere: A YouTube Study”. In International Communication Association Annual Conference, 2008, Canada.
- [8] Stiubiener, R.M. Silveira, R. Matushima, W.V. Ruggiero; “Monitoring the video use for learning support”. IEEE Frontiers in Education Conference, 2013.
- [9] Stiubiener, R.M. Silveira, R. Matushima, G. Bressan, W.V. Ruggiero; “Evaluating the effectiveness of the use of teaching materials in video format in distance learning environments”. IEEE Frontiers in Education Conference, 2012.

⁴ Acesso ao serviço através do endereço www.video.rnp.br