

Análise do comportamento do aluno através de técnicas de web analytics e sua proposta de representação interativa em um ambiente e-learning

Barbara Moissa
UDESC (Universidade do
Estado de Santa Catarina)
barbara.moissa@gmail.com

Lucas Simões de
Carvalho
UDESC
lucas.scarvalho@gmail.com

Isabela Gasparini
UDESC
isabela.gasparini@udesc.br

Avanilde Kemczinski
UDESC
avanilde@joinville.udesc.br

ABSTRACT

This paper describes a web analytics tool used in an e-learning environment to collect data about students in order to understand their behavior and find useful information about their interaction in disciplines. In order to provide a better analysis of the collected data, an information visualization tool is proposed. This tool goal is to allow teachers to analyze data in a visual and interactive way, helping them to understand students' behavior and navigational patterns.

Keywords

Web analytics, information visualization, e-learning.

RESUMO

Este artigo descreve uma ferramenta *web analytics* implementada em um ambiente *e-learning* para coletar dados sobre os estudantes de modo a entender o seu comportamento e encontrar informações úteis a respeito de sua interação nas disciplinas. Para fornecer uma melhor análise dos dados coletados, é proposta uma ferramenta de visualização da informação. Esta ferramenta tem como objetivo permitir que os professores analisem os dados de forma visual e interativa, ajudando-os a entender o comportamento e os padrões de navegação dos estudantes.

Palavras-chave

Web analytics, visualização da informação, *e-learning*.

1. INTRODUÇÃO

Em ambientes *e-learning*, uma das dificuldades dos professores é conhecer e entender os alunos de forma a adaptarem os materiais didáticos de acordo com suas necessidades. Além disso, o professor também pode desejar verificar a assiduidade dos alunos, seus tópicos de interesse, seus estilos e preferências navegacionais e outros fatores que podem influenciar no modo que o professor prepara sua disciplina. Para ajudar os professores que utilizam o ambiente AdaptWeb[®] (Ambiente de

Ensino-Aprendizagem Adaptativo na Web) nesta tarefa, os dados navegacionais dos alunos são coletados e armazenados, e métricas são calculadas para que os dados e padrões de comportamento dos alunos em uma disciplina possam ser acessos e apresentados aos professores.

No ambiente AdaptWeb[®], os dados são coletados através de uma ferramenta de *web analytics* e os professores podem analisar as métricas calculadas através de gráficos. Para que os professores possam explorar na totalidade os dados coletados durante a busca de informações sobre suas disciplinas e conseguir encontrar características e padrões de navegação são utilizadas técnicas de visualização da informação. Assim, estas análises podem fornecer algum tipo de filtro, ou até mesmo, meios de realizá-las a partir da interação com os dados, manipulando somente alguns dos dados apresentados e/ou agregando outras informações multifacetadas relacionadas a uma disciplina como dados relativos ao acesso ao sistema e seus recursos e também a realização de atividades dentro do ambiente. A partir da análise permitida pela ferramenta, o professor deve ser capaz de encontrar as informações que precisa e ter condições de adaptar a disciplina, possibilitando mudanças de comportamento dos alunos e no processo de ensino e aprendizagem.

2. ADAPTWEB[®]

O AdaptWeb[®] é um ambiente de ensino-aprendizagem adaptativo, com o objetivo de adaptar o conteúdo (a informação), a apresentação e a navegação em uma disciplina de acordo com o perfil do aluno. Essa adaptação é baseada em algumas características do aluno, entre elas o seu conhecimento sobre os conceitos da disciplina, suas preferências navegacionais e sua formação. Outra característica do ambiente é que uma disciplina pode ser oferecida para mais de um curso. Desta forma, a disciplina de Cálculo I pode ser oferecida para os cursos de Ciência da Computação, Engenharia Elétrica e Engenharia Civil com seu conteúdo adaptado de acordo com cada curso.

O aluno pode acessar três seções principais: o Ambiente de Aula, o Mural de Recados e o Fórum de Discussão. Cada uma destas áreas do AdaptWeb[®] e a ferramenta *web analytics* serão abordadas a seguir.

2.1 Ambiente do aluno

O ambiente do aluno, monitorado pela ferramenta *web analytics*, possui três seções: Ambiente de Aula, Mural de Recados e Fórum de Discussão. Cada uma destas três seções faz parte de uma disciplina oferecida a um determinado curso, ou seja, um

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

aluno de Cálculo I no curso de Ciência da Computação verá, em cada uma destas três áreas, apenas o conteúdo disponível para os alunos de Ciência da Computação, enquanto um aluno de Engenharia Elétrica verá apenas o conteúdo disponível para alunos de Engenharia Elétrica.

O aluno pode acessar os conteúdos didáticos no Ambiente de Aula por dois modos de navegação, o modo livre ou o modo tutorial, e obter acesso a conceitos, exemplos, exercícios e materiais complementares. No modo tutorial, o aluno verá o conteúdo de acordo com os pré-requisitos definidos pelo professor. Por exemplo, na disciplina de Matemática Básica o aluno verá o conteúdo sobre exponenciação apenas se já tiver conhecimento sobre o conteúdo sobre multiplicação. Já no modo livre, os pré-requisitos definidos pelo professor são desconsiderados, o aluno tem acesso a todo o conteúdo de forma livre, podendo explorar os conteúdos da forma que achar melhor.

No Mural de Recados o aluno pode ver os recados recebidos nos últimos 30 dias. Nesta área do ambiente do aluno, ele ainda pode enviar um recado para todos os integrantes da turma — alunos e professor — ou ainda apenas para o professor. Da mesma forma, o professor pode deixar um recado para um aluno específico (somente ele irá ver a mensagem deixada pelo professor), para um grupo de alunos ou para toda a turma.

O Fórum de Discussão permite que o aluno interaja com os demais alunos de sua turma. Ele pode criar discussões sobre algum assunto referente ao conteúdo da disciplina e assim, conseguir auxílio para realizar os exercícios ou compreender algum conceito.

2.2 Coleta de dados navegacionais

A coleta dos dados navegacionais do aluno no AdaptWeb® é realizada através de uma ferramenta *web analytics* integrada a ele. Esta ferramenta coleta os dados para calcular métricas e permitir a análise dos *Key Performance Indicators* (KPI) definidos, ou seja, um conjunto de métricas que juntas permitem analisar algum objetivo do sistema monitorado.

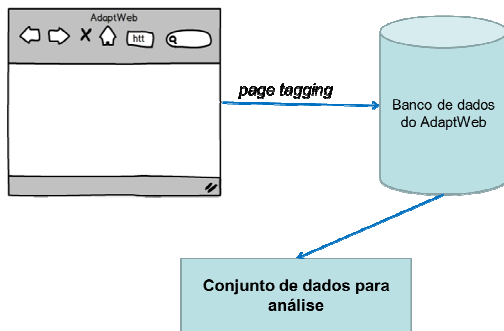


Figura 4. Modelo proposto para coleta de dados

Como o objetivo desta ferramenta é fornecer meios de conhecer os alunos que utilizam o ambiente, os dados são coletados nas três áreas do ambiente do aluno: Ambiente de Aula, Mural de Recados e Fórum de Discussão. Esta coleta ocorre através de *logs* das requisições ao servidor e também através de *page tagging*. Enquanto os *logs* são salvos pelo servidor ao processar a requisição, a técnica de *page tagging* captura outros dados através de um código *javascript* interpretado pelo navegador do aluno e os envia para o servidor, que os salva em sua base de

dados, conforme ilustrado na Figura 4. Os dados coletados são organizados de forma que possam ser enviados aos professores que os solicitam de acordo com métricas, período de tempo e aluno ajustados.

2.2.1 Key Performance Indicators

A partir dos dados coletados pela ferramenta, algumas metas do sistema podem ser mensuradas. Esta mensuração é realizada através dos *Key Performance Indicators*, conhecidos por KPIs. Um KPI é uma variável utilizada para visualizar se o *website* está atingindo ou não os seus objetivos. Por isto, cada KPI deve ter ações ou iniciativas diretamente relacionadas a ele para atingir estas metas [1].

A verificação dos KPIs é feita através das métricas do sistema. As métricas utilizadas são relacionadas ao conteúdo, a navegação e ao cumprimento de metas dentro do ambiente. Após uma análise das métricas de *web analytics* utilizadas comumente em diversos sistemas, criou-se um conjunto de métricas e objetivos a serem utilizados especificamente em ambientes *e-learning*.

Visando validar este conjunto de métricas e objetivos, foi proposto um questionário aplicado a vinte e dois professores que utilizam ambientes *e-learning* em suas disciplinas, pertencentes a catorze universidades brasileiras. O objetivo principal foi verificar se as métricas identificadas realmente auxiliam os professores a averiguar se os objetivos educacionais estão sendo atingidos. Outro objetivo foi analisar se as métricas estão compatíveis com os recursos disponíveis nos ambientes utilizados por estes professores. Mais informações sobre esta pesquisa pode ser vista em [2]. As métricas, organizadas em três seções — Visitas, Conteúdo e Conversão, são as seguintes:

- Visitas: quantidade de alunos que acessam o sistema, o tempo médio de permanência dos alunos no sistema, o tempo de acesso individual na disciplina, a frequência de acesso dos alunos ao sistema, alunos conectados em tempo real, dados do usuário (localização, sistema operacional, acesso por *desktop* ou dispositivo móvel...);
- Conteúdo: materiais mais acessados pelos alunos, uso do sistema de busca (quantidade e palavras-chave), tempo de acesso em cada seção do sistema, caminho percorrido dentro do ambiente de aula e o caminho percorrido pelo aluno dentro de todo o ambiente;
- Conversão: estudo dos materiais complementares, participação nos fóruns de discussão, interação através de mensagens e realização dos exercícios.

3. VISUALIZAÇÃO DA INFORMAÇÃO

A visualização da informação (VI) é uma área de processamento gráfico que tem como objetivo auxiliar a análise de um grande conjunto de dados através de representações gráficas manipuláveis [3]. Desta forma, através das técnicas de VI um grande conjunto de dados brutos pode ser representado graficamente de forma a facilitar sua análise. Além de representar estes dados de forma visual, as técnicas de VI permitem que o usuário manipule os dados representados e a representação destes para encontrar as informações que procura.

Alguns modelos de referência estão presentes na literatura, tais como os de Haber e McNabb [4], Campo, Orosko e Teyseyre

[5], Chi e Riedl [6] e Card, Mackinlay e Shneiderman [7] para explicar o funcionamento de ferramentas que utilizam estas técnicas. Todos estes modelos possuem características em comum, pois em todos eles, os dados brutos sofrem alterações de forma a serem mapeados em uma estrutura visual que será então mostrada ao usuário.

Após mapear os dados para a estrutura visual, esta é mostrada ao usuário. O usuário então irá analisar os dados e interagir com a visualização. Estas interações podem apenas alterar características na própria visualização ou ainda solicitar que os dados brutos sejam outros — exigindo que o processo volte ao início, ou então trocar a representação visual utilizada, exigindo que os dados transformados sejam mapeados novamente.

A partir deste modelo é possível identificar a necessidade de utilizar dois tipos de técnicas, as de representação e as de interação, explicadas a seguir.

3.1 Técnicas de representação

As técnicas de representação são utilizadas para representar visualmente o conjunto de dados em análise pelo usuário para que ele possa compreendê-los mais facilmente e extrair informações. Ou seja, estas técnicas são utilizadas nos gráficos nos quais o usuário terá os dados representados.

Estas técnicas podem ser classificadas em quatro categorias:

- **Ortgonais:** São utilizadas quando as variáveis analisadas são no máximo três. Esta categoria contém gráficos de barras, linhas, área, dispersão, entre outros.
- **Multidimensionais:** São utilizadas quando as variáveis analisadas são no mínimo três. Esta categoria contém o gráfico de coordenadas paralelas.
- **Árvores:** São utilizadas quando os itens representados possuem um relacionamento hierárquico entre si. É o caso das árvores genealógicas e dos diretórios de arquivos. Estas técnicas podem ser baseadas em cinco tipos de representação: por nós e arestas, aninhamentos, adjacências, identificações e matrizes.
- **Redes:** São técnicas utilizadas para representar relacionamentos entre os itens, porém não necessariamente hierárquicos. Esta categoria engloba os grafos, sejam estes direcionados ou não.

Em conjunto com estas técnicas, são exploradas propriedades visuais como cor, tamanho e formato para indicar elementos diferentes. De acordo com Yi *et al* [8], estas características são importantes pois podem afetar a cognição humana e, principalmente, influenciar o modo como os usuários entendem os relacionamentos e a distribuição dos dados.

3.2 Técnicas de interação

As técnicas de interação permitem que o usuário interaja com a ferramenta de forma a alterar a representação visual dos dados e até mesmo os próprios dados.

Yi *et al* [8], com base em outras classificações, descrevem sete categorias de modificações/ações que o usuário pode realizar:

- **Seleção:** São técnicas que destacam/diferenciam itens de interesse para facilitar a análise, ou seja, artifícios como

alterar a cor ou o tamanho de um elemento selecionado de forma que o usuário possa distingui-lo dos demais.

- **Exploração:** São técnicas que permitem que o usuário analise subconjuntos de itens que não estão visíveis no momento. Arrastar uma imagem para ver outras partes dela ou então utilizar uma barra de rolagem são exemplos desta categoria.
- **Reconfiguração:** São técnicas que reorganizam os itens mostrados, ou seja, quando o usuário altera a raiz da árvore representada ou ordena alfabeticamente os itens de uma tabela de acordo com a coluna “Nome”, ele está reconfigurando a visualização. Estas ações permitem que o usuário veja os dados através de outra perspectiva, possibilitando a descoberta de novas informações que poderiam não ter sido encontradas com a configuração anterior.
- **Codificação:** São técnicas que permitem que o usuário altere as propriedades visuais dos itens representados como a aparência (cor, tamanho, formato) e até mesmo a técnica de representação utilizada.
- **Abstração/Elaboração:** São técnicas que permitem que usuário veja mais ou menos detalhes dos itens. Os exemplos mais comuns destas técnicas são *tooltips* com informações adicionais e o *zoom in* e *zoom out*.
- **Filtragem:** São técnicas que permitem que o usuário altere os dados representados com base em alguma condição definida por ele.
- **Conexão:** São técnicas que destacam associações/relações entre itens já representados ou então mostram itens escondidos relevantes a um determinado item.

3.3 FERRAMENTA PROPOSTA

Devido a grande quantidade de dados coletados no AdaptWeb[®] e a falta de representações visuais interativas fornecidas pela ferramenta *web analytics*, é proposta uma ferramenta que utiliza técnicas de visualização da informação para representar os dados coletados.

Os dados brutos, coletados pela ferramenta *web analytics* integrada ao ambiente, serão transformados e mapeados de acordo uma técnica de representação. Após mapear estes dados, a visualização dos dados estará disponível para o professor.

Por sua vez, o professor poderá interagir com a visualização resultante deste processo e encontrar as informações que procura, como particularidades dos alunos, padrões navegacionais, entre outros. Espera-se que a partir dos resultados da análise permitida por esta ferramenta, o professor consiga compreender os alunos e adaptar-se às suas necessidades de aprendizagem.

3.4 Representação dos dados

A ferramenta proposta baseia as representações dos dados nas métricas calculadas pela ferramenta *web analytics* para permitir a análise das KPIs. Estas análises serão realizadas através das técnicas relacionadas na Tabela 1. Assim, cada uma das métricas calculadas pela ferramenta *web analytics* poderá ser analisada individualmente.

Métricas de Análise	Técnicas de representação
Tempo de utilização de cada seção do sistema	Gráfico de setores
Alunos conectados em tempo real	
Tempo de utilização individual no Ambiente de Aula	
Tempo de utilização individual no Ambiente de Aula comparado com o tempo total de acesso	
Média de tempo <i>online</i>	
Frequência de acesso	Gráfico de barras
	Gráfico de coordenadas paralelas
Quantidade de alunos que acessaram o sistema	Gráfico de barras
Realização dos exercícios	
Caminho percorrido	Grafo
Materiais complementares mais acessados	<i>Sunburst</i>

Tabela 1. Métricas e respectivas técnicas de interação

3.5 Exploração dos dados

Para permitir que o professor explore os dados representados, foram propostas técnicas de interação como arrastar nós, remover itens, *panning*, detalhes, *zoom*, filtragem e também a exportação. A Figura 6 ilustra quais técnicas de interação (em cinza) serão utilizadas com cada uma das técnicas de representação propostas (em branco).

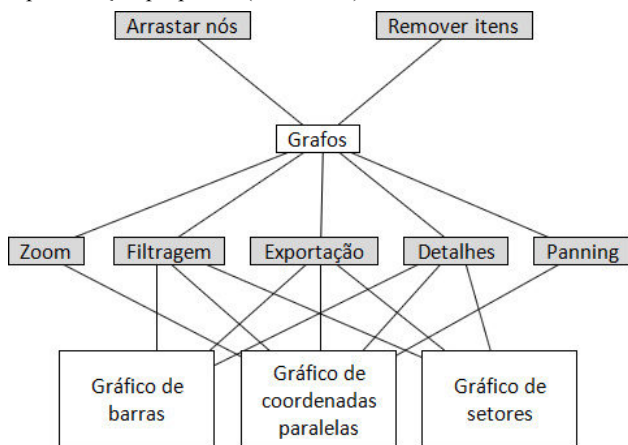


Figura 6. Técnicas de representação e respectivas técnicas de interação

A partir do filtro, o professor pode analisar os dados coletados de acordo com alguns parâmetros como: curso, disciplina, aluno e período. Com isto, será possível comparar diferentes grupos de alunos e verificar se o comportamento navegacional, de interação com outros alunos e de exploração dos materiais é similar entre os diversos grupos. Outra possibilidade será comparar o comportamento dos alunos em períodos diferentes, permitindo identificar quando os alunos se dedicaram mais à disciplina ou então comparar o comportamento de alunos da turma atual com o de alunos de turmas anteriores.

4. Próximas etapas

Com as técnicas de representação e interação definidas, a próxima etapa consiste no desenvolvimento da ferramenta.

Assim como a ferramenta *web analytics* descrita na seção 2.2 deste artigo, a ferramenta de visualização da informação será integrada ao AdaptWeb®.

Após desenvolver a ferramenta serão realizados alguns experimentos com professores que utilizam o ambiente para identificar deficiências tanto na utilização da ferramenta como nas representações definidas para cada uma das métricas.

5. CONSIDERAÇÕES FINAIS

Este artigo mostrou uma ferramenta de *web analytics* que coleta dados navegacionais dos alunos do ambiente *e-learning* AdaptWeb®. As métricas utilizadas para analisar os KPIs, foram definidas de acordo com a necessidade dos professores que utilizam este ambiente.

Para melhorar a análise dos dados coletados por esta ferramenta, este trabalho apresenta uma proposta de utilização das técnicas de visualização da informação para permitir a exploração visual destes dados. Com esta ferramenta espera-se que os professores consigam entender melhor o comportamento dos alunos e identificar problemas de aprendizado de maneira fácil e intuitiva.

As próximas etapas no desenvolvimento desta ferramenta consistem na implementação da ferramenta e em experimentos com a mesma. Estes experimentos serão realizados com professores que utilizam o ambiente AdaptWeb® para verificar se a ferramenta atingiu seus objetivos de análise, assim como identificar deficiências na utilização da mesma.

REFERÊNCIAS

- [1] Waisberg, D.; Kaushik, A. *Web Analytics 2.0: empowering customer centricity*. Em SEM.org Journal, v.2, n.1, 2009.
- [2] Carvalho, L. S. de; Gasparini, I. Compreensão dos objetivos do professor nos ambientes *e-learning* para análise de métricas *web analytics*. *Conferência Ibero-Americana WWW/Internet (CIAWI)*, Porto Alegre, 2013 (to appear).
- [3] Freitas, C. M. D. S.; Chubachi, O. M.; Luzzardi, P. R. G.; Cava, R. A. *Introdução à Visualização de Informações*. RITA, Porto Alegre, RS, Brasil, 2001.
- [4] Haber, R. B.; McNabb, D. A. *Visualization Idioms: a conceptual model for scientific visualization systems*. *Visualization in Scientific Computing* (1990), 74-93.
- [5] Campo, M.; Orosco, R.; Teyseyre, A. *Automatic Abstraction Management in Information Visualization Systems*. Em *Proceedings of the International Visualization Conference* (1997), 50-56.
- [6] Chi, E. H.; Riedl, J. T. *An operator Interaction Framework for Visualization Spreadsheets*. Em *Proceedings of IEEE Information Visualization Symposium* (1998), 63-70.
- [7] Card, S. K.; Mackinlay, J. D.; Shneiderman, B. *Readings in Information Visualization: using vision to think*. Morgan Kaufman (1999).
- [8] Yi, J. S.; Kang, Y.; Stasko, J. T.; Jacko, J. A. *Toward a Deeper Understanding of the Role of Interaction in Information Visualization*. *IEEE Transactions on Visualization and Computer Graphics* (2007), 1224-1231.