

Análise de logs do sistema Agritempo por meio do log do PHPNuke e WebAlizer

Thallita Justo Gomes Dias Thomaz¹

Luciana Alvim Santos Romani²

Nos últimos anos, a tecnologia da informação, em especial a web, tem provido mecanismos capazes de tornar disponível um grande volume de informações a baixo custo para um número expressivo de usuários (VASCONCELLOS-SILVA et al., 2003). Os *websites* fornecem uma ampla gama de conteúdos diferentes para todo os tipos de perfis de usuário. Diferentes formas de avaliação desse público têm sido propostas e usadas pelos administradores, a fim de adaptar e aperfeiçoar o conteúdo a seu público-alvo. Uma das formas de monitoramento é a análise de arquivos de *log* (ALENCAR et al., 2003).

Um arquivo de *log* é gerado pelo servidor web e contém informações relativas ao número de IP do cliente, data e hora do acesso, transcrição da solicitação efetuada pelo *browser* ao servidor, status da solicitação e quantidade de dados transferidos. Dependendo da configuração e da capacidade do servidor, é possível registrar ainda a página que levou o visitante ao website, o nome e a versão do *browser* utilizado (COI..., 2002). Esse tipo de análise é fundamental também para a constante manutenção do sistema, uma vez que essa análise permite verificar quais são as informações existentes de maior e menor interesse dos usuários. Alguns sistemas gratuitos para análise de *logs* estão disponíveis como o *Webalizer*, o *Google Analytics*, o *Piwiki*.

O foco deste trabalho é a análise dos *logs* do sistema Agritempo, um sistema de monitoramento agrometeorológico que disponibiliza informações meteorológicas e agrometeorológicas de diferentes regiões brasileiras

¹ FT/Unicamp, thallita.thomaz@colaborador.embrapa.br

² Embrapa Informática Agropecuária, luciana.romani@embrapa.br

gratuitamente na internet³. O Agritempo possui um amplo público-alvo: produtores, extensionistas, consultores, agentes do governo, estudantes e professores universitários, além da iniciativa privada.

Nesse trabalho, foram analisados os *logs* armazenados pelo servidor Apache por meio da ferramenta PHP Nuke⁴, que apresenta a quantidade de *Page Views* que houve no site em períodos de anos, meses, dias e horas. *Page View* é uma métrica utilizada para indicar a quantidade de acessos às páginas de um *website*. As estatísticas referentes a *Page Views* podem auxiliar na análise do perfil de acessos das páginas do *website*, em um determinado período, em determinadas horas do dia, entre outros critérios.

Outra ferramenta também avaliada nesse artigo é o *WebalizerDisponível* em:⁵, um aplicativo gratuito, que analisa os *logs* e gera estatísticas por meio de diversas variáveis como país de acesso, pastas e visitas. Nesse documento, as análises foram feitas considerando os *Hits* dos anos de 2004, 2007, 2009, 2010 e 2011 nos meses de setembro e outubro de cada um deles. *Hits* são quaisquer pedidos feitos pelo servidor que está conectado, ou seja, representa o número total de pedidos feitos em um determinado espaço de tempo já estabelecido por ele.

Estatísticas de *Page Views* do PHP Nuke

Os dados gerados pelo PHP Nuke estão disponíveis na interface principal do sistema, não são armazenados os dados de *Page Views* dos campos “Consultas”, “Sumário” e “Acesso Restrito” do sistema Agritempo. Com a análise dos logs de acesso do Agritempo, entre 2003 e 2010, pretende-se avaliar a usabilidade do sistema e também o seu conteúdo, uma vez que um baixo número de *Page Views* pode indicar problemas no conteúdo do website. A Figura 1 apresenta uma comparação dos acessos anuais,

³ Disponível em: <www.agritempo.gov.br>.

⁴ Disponível em: <<http://phpnuke.org/>>.

⁵ Disponível em: <http://www.webalizer.org/>.

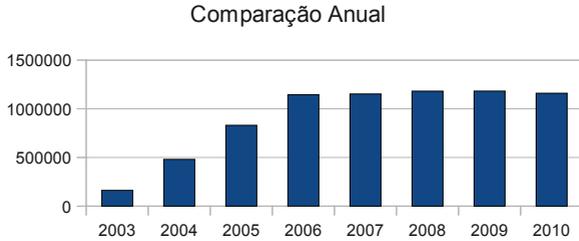


Figura 1. Comparação Anual de *Page Views*.

evidenciando a evolução da quantidade de páginas acessadas entre 2003 e 2010, indicando um acentuado crescimento.

A Figura 2 apresenta uma comparação das páginas visitadas no ano de 2007, indicando que no mês de agosto se dá o maior número de *Page Views*. Uma causa possível para esse perfil de acesso seria pelo fato de ser um mês de baixa taxa de umidade do ar e que precede ao mês de plantio, o que poderia gerar maior interesse nos usuários por consultas a informações agrometeorológicas e previsão de chuvas

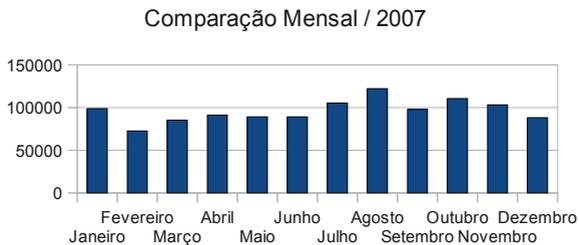


Figura 2. Comparação Mensal de *Page Views*.

A análise dos acessos diários do mês de agosto de 2007, apresentada na Figura 3, demonstra constância nos acessos diários ao site durante os dias úteis, com ligeira queda nos finais de semana.

Avaliando detalhadamente os acessos realizados em dias da semana nota-se que, embora haja acessos em todos os horários, o número de *Page Views* é maior em dias úteis do que nos finais de semana, como pode ser visto nas Figuras 4 e 5.

Comparação Diária / Agosto - 2007

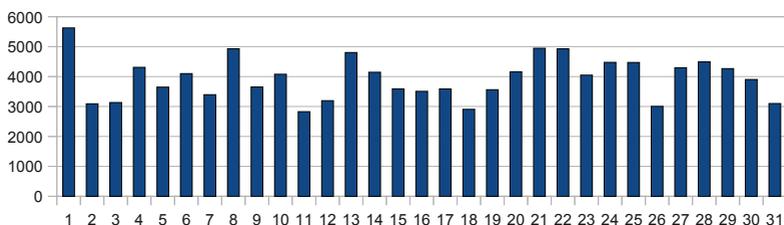


Figura 3. Comparação Diária de *Page Views*.

Comparação às quartas-feiras

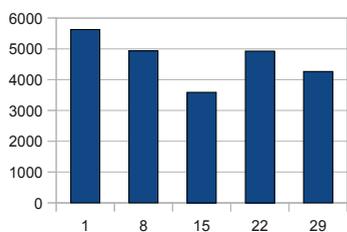


Figura 4. *Page Views* às quartas-feiras.

Comparação aos domingos

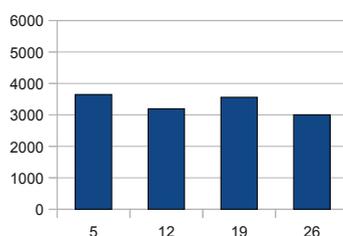
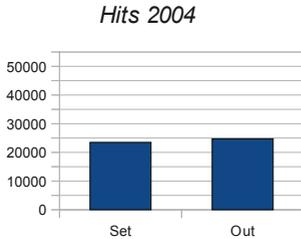
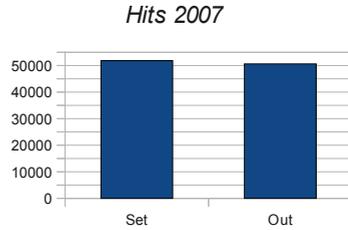
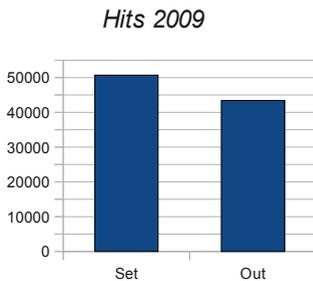
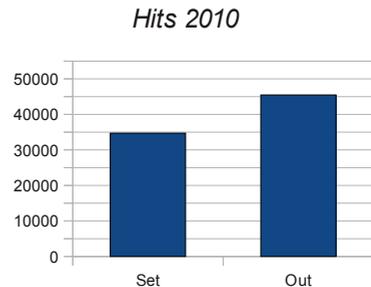
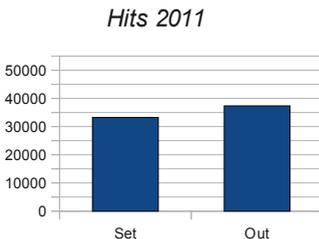


Figura 5. *Page Views* aos domingos.

Relatório de *Hits* do *Webalizer*

Os relatórios gerados pelo *Webalizer* mostram um aumento no número de *Hits* de 2004 para 2007, passando de 23500 em setembro de 2004 para 51800 no mesmo mês de 2007, como pode ser visto nas Figuras 6 e 7.

O aumento no número de *Hits* (Figuras 6 e 7) mostra um aumento no número de acessos ao site após 4 anos de sistema em produção. Entre os anos de 2009 e 2011 em contrapartida a análise feita usando o *Webalizer* indica uma queda no número de *Hits*. Durante esse período, os arquivos de *log* não foram armazenados de forma integral, o que prejudica uma análise mais detalhada dos acessos. Além disso, uma falta de atualização das páginas do sistema Agritempo também pode ter motivado um não aumento no número de visitas ao site.

**Figura 6.** Hits do ano de 2004.**Figura 7.** Hits do ano de 2007.**Figura 8.** Hits do ano de 2009.**Figura 9.** Hits do ano de 2010.**Figura 10.** Hits do ano de 2011.

Com isso, este trabalho de análise dos *logs* auxiliou a mostrar a necessidade de armazenamento dos *logs* do sistema de forma a possibilitar uma análise mais adequada e detalhada das consultas feitas pelo usuário. Além disso, corrobora a necessidade de atualização do sistema para uma versão mais moderna e alinhada com conceitos da web 2.0.

A análise dos *logs* do sistema Agritempo, por meio da ferramenta de análise do sistema PHP-Nuke e *Webalizer*, permitiu identificar tendências e padrões de uso do sistema pelos usuários do sistema Agritempo. Os resultados desse tipo de análise podem auxiliar na definição de melhorias

a serem incorporadas ao sistema e na proposição de novos recursos para melhor atender às necessidades do seu público-alvo.

Devido ao armazenamento incorreto dos logs no servidor, podemos concluir que a ferramenta mais eficaz na análise dos *logs* do sistema Agritempo atualmente é a realizada pelo PHP Nuke.

Referências

ALENCAR, M. C. F.; GOMES, E. G.; PIEROZZI Jr., I. Uso do web site da Embrapa Monitoramento por Satélite para educação: perfil de usuário e análise de conteúdo. **Contrapontos**, v. 3, n. 1, p. 27-38, 2003.

BARRET, B. L. **The Webalizer - A web server log file analysis tool**. 1997. Disponível em: <<ftp://ftp.mrunix.net/pub/webalizer/README>>. Acesso em: 15 out. 2012.

COI Website evaluation, usage and analysis, Cabinet Office UK: UK, 2002.

VASCONCELLOS-SILVA, P. R.; CASTIEL, L. D.; RIVERA, F. J. U. Assessing an internet health information site by using log analysis: the experience of the National Cancer Institute of Brazil. **Revista Panamericana Salud Pública**, v. 14, n. 2, 2003.