Desenvolvimento de um Módulo de *Data Webhouse* para analisar informações do comportamento dos consumidores durante o uso de *sites* na *Internet*

André Zapala
Universidade Nove de Julho
Diretoria dos Cursos de Informática – Ciência da Computação
Av. Francisco Matarazzo, 612. São Paulo – SP
familia.zapala@gmail.com

Rafael Lacerda
Universidade Nove de Julho
Diretoria dos Cursos de Informática – Ciência da Computação
Av. Francisco Matarazzo, 612. São Paulo – SP
faleiros.lacerda@gmail.com

Daniel Lemos Santos
Universidade Nove de Julho
Diretoria dos Cursos de Informática – Ciência da Computação
Av. Francisco Matarazzo, 612. São Paulo – SP
Daniel.lemos.santos@gmail.com

Renato José Sassi
Universidade Nove de Julho
Programa de Mestrado em Engenharia de Produção
Av. Francisco Matarazzo, 612. São Paulo – SP
renato.sassi@ieee.org

RESUMO

A partir de 1995, quando se iniciou a massificação da *Internet*, os pesquisadores da área de *Data Warehouse* passaram a estudar as possibilidades de transmissão, de manipulação dos dados e, conseqüentemente, da informação contida no *Data Warehouse* através da *Internet*. Com isso, surgiu uma nova área de pesquisa denominada *Data Webhouse*, que possui os mesmos princípios do *Data Warehouse*, mas com uma arquitetura diferenciada por ter a presença da *Internet*. O *Data Webhouse* tem como foco central as operações do negócio voltado para a *Internet*. Mais especificamente, um *Data Webhouse* acumula dados detalhados da seqüência de cliques (*clickstream*) que identificam o comportamento dos consumidores ao navegar no *site* de uma empresa. Através da utilização eficaz desses dados é possível traçar o perfil individual de cada consumidor e personalizar o *site* de acordo as necessidades individuais de cada um deles. Assim, os dados comportamentais dos consumidores gerados pela *Internet* alimentam o *Data Webhouse* permitindo analisar detalhadamente esse comportamento, obtendo informações a cada clique, a cada gesto e de acordo com a trajetória realizada no *site*, além da possibilidade de responder perguntas como: Qual o perfil de navegação de um consumidor cadastrado? Qual o perfil de navegação de um consumidor que cancela o serviço? Quanto tempo os consumidores gastam no *site* da empresa? Com o *Data Webhouse* as empresas possuem uma excelente oportunidade de coletar dados valiosos sobre os seus consumidores na *Internet* e, com isso, auxiliar na criação de melhores serviços e no aprimoramento das vendas. As empresas podem, ainda, utilizar esses dados para determinar os hábitos de compra, fornecer aos consumidores recomendações sobre novos produtos e obter outras informações relevantes para o processo de descoberta de padrões de comportamento dos consumidores. Uma vez alimentado o *Data Webhouse*, são criados vários módulos que tem como objetivo dar suporte à análise das informações

sobre o comportamento dos consumidores durante a utilização de *sites* na *Internet*. Assim, o objetivo deste trabalho foi criar um módulo para ler os dados de uma base de dados de consumidores na *Internet* gerando gráficos para facilitar o entendimento do comportamento dos consumidores e comparar os dados obtidos recentemente com dados anteriores. Para a criação do módulo foi utilizada a linguagem PHP/Ajax, o banco de dados utilizado foi o MySQL e a API dos gráficos foi o Google *Chart*. A base de dados utilizada contém dados de consumo de 1.968 consumidores com 48 atributos, sendo 47 atributos condicionais e um atributo de decisão dividido em duas classes (M = masculino; F = feminino). Os resultados confirmam que a utilização do módulo para dar suporte à análise do comportamento dos consumidores durante o uso de *sites* na *Internet* pode apoiar uma tomada de decisão com maior precisão.

PALAVRAS-CHAVE: *Data Webhouse*, Comportamento dos Consumidores, *Internet*.

Development of a Module of the DataWebhouse to analyze the behavior of consumers during the use of Internet sites

ABSTRACT

Since 1995, when started the mass access to the Internet, the researchers in Date Warehouse began to study the possibilities of transmission, data manipulation and consequently, the information contained in the Data Warehouse via the Internet. Thus, a new area of research emerged, the Data Webhouse, which has the same principles of the Data Warehouse, but with a differentiated architecture by have the presence of the Internet. The Data Webhouse is focused on the central business operations facing the Internet. More specifically, a Data Webhouse accumulates data detailed click sequence (clickstream) that identifies the behavior of consumers to browse the site for the company. Through the effective use of these data is possible to trace the individual profile of each customer and customize the website according to individual needs. Thus, the behavioral data generated by consumer Internet feed the Data Webhouse allowing detailed analysis of this behavior, providing information at every click, every gesture and the trajectory carried out on the site, plus the ability to answer questions such as: What is the profile of a consumer registered on the website? What is the browsing profile of a consumer that cancels services? How much time these consumers spend on the company website? Using Data Webhouse the companies have an excellent opportunity to collect valuable data about their consumers on the Internet and thereby help create better services and improving sales. The Companies can also use these data to determine the buying habits, provide consumers with recommendations on new products and obtain others relevant information to the process of discovering patterns in consumer behavior. Once fed the Data Webhouse, several modules are created that aims to support the analysis of information about the customer behavior when using Internet sites. The objective of this work was to create a module to read data from a customer database on the Internet generating graphics to facilitate understanding of consumer behavior and compare the data with previous data obtained recently. To create the module was used PHP / Ajax, the database used was MySQL and graphics API was Google Chart. The database used contains data from 1,968 consumers with consumption of 48 attributes, 47 conditional attributes and one decision attribute divided into two classes (M = male F = female). The results confirm that the use of the module to support analysis of consumer behavior when using Internet sites can support decision making with precision.

KEYWORDS: Data Webhouse, Consumer Behavior, Internet.