# A stochastic model of e-customer behavior

Mamata Jenamani[a], Pratap K.J. Mohapatra[a,*], Sujoy Ghose[b]

[a]*Department of Industrial Engineering and Management, Indian Institute of Technology, Kharagpur, India*
[b]*Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, India*

## Abstract

Web usage mining techniques are increasingly used today to understand e-customers' within-site behavior. We propose a data mining model that considers e-customers' activities as a discrete-time semi-Markov process and explains their behavior. An algorithm is proposed to compute transition probability matrix and holding time mass functions from the site navigation data. Finally, the model is used to explain customer behavior in an example site. A software agent, implemented in the site, collects and stores navigation data in the required form and thus helps to avoid data preprocessing. The model results helped to improve the site design and judge its performance.
© 2003 Elsevier Science B.V. All rights reserved.

*Keywords:* E-customer behavior; Web usage mining; Discrete-time semi-Markov process

## 1. Introduction

Websites having emerged as the first contact point with the customers for many organizations, the study of e-customer's behavior has become an essential field of research. In a traditional business environment, when a seller meets a buyer in person he can understand his behavior, like intention for purchase and choice of product, etc., from his expressions and verbal communication. The sales person accumulates this knowledge while dealing face to face with the customers. Subsequently, he uses this knowledge to create facilities to increase customer satisfaction.

But, in an online environment, no such knowledge can be directly obtained. Online feedback forms can help to get customer preferences. Customer preferences can be analyzed to understand e-customer behavior. But there is no way to ensure that a visitor fills up this form. An alternative is to conduct visitor surveys. However, unless properly designed, these surveys may involve samples from specific parts of the visitor population and may not depict a true picture of customer behavior due to the global nature of the ecommerce environment. In this scenario, the website visitors' page access pattern can be analyzed, using the Web usage mining techniques, to understand the e-customers' within-site behavior. Analysis of e-customer behavior can help in improving the content and design of the website, customizing the website, building stronger customer relationship, heightening communication with the customer, and in enhancing customer service.

*Corresponding author. Tel.: +91-322-255-221; fax: +91-322-255-303.

*E-mail addresses:* mamataj@vgsom.iitkgp.ernet.in (M. Jenamani), pratap@hijli.iitkgp.ernet.in (P.K.J. Mohapatra), sujoy@cse.iitkgp.ernet.in (S. Ghose).

The rest of the paper is organized as follows. First, we give a comprehensive survey of literature in the area of e-customer behavior. Thereafter we propose a model of e-customer behavior treating it as a semi-Markov process. Next, we present an algorithm that transforms the navigational data to a form that can be readily used by the model. Finally, we use the model to analyze the navigational data of an example site and use the results to find a few design defects in the site and predict the performance of the site

## 2. Literature survey

There are basically two approaches evolving in parallel for the study of e-customer behavior. The first approach involves conducting questionnaire surveys and experiments, and subsequently developing models based on these studies. The second approach is based on the application of data mining techniques to the data generated from the online transactions. The application of data mining techniques to Web usage data has generated a new terminology called Web Usage Mining (WUM). We divide the literature on e-customer behavior into non-WUM-based and WUM-based approaches.

### 2.1. Non-WUM-based approaches

Hoffman and Novak [1] have proposed a model of consumer navigational behavior. The model suggests that a visitor who becomes totally engrossed (enters into flow state) during navigation, is unaware of the time spent in navigation, perceives better, and retains more of what he perceives than one who does not experience such a state. The team at Georgia Institute of Technology's Graphic, Visualization, and Usability Center [2] conducted a series of online surveys to develop a generalized consumer profile of the Internet shoppers and to understand online consumers' Web usage pattern.

A survey by Jarvenpaa and Todd [3] among a sample of 220 Internet shoppers reveals that 31% of respondents are disappointed with the product variety and 80% have at least one negative comment about the customer service on the Web. Dholakia and Rego [4] have found that the factors that attract more traffic are: the updates made in the site and the external links from the site. Guttman and co-workers [5,6] have proposed a six-stage model of consumers' across site buying behavior in the context of automated bargaining. Liang and Huang [7] have found the dependence of online shopping adoption on the product type, the perceived risk, and on the consumer's experience. Lohse and Spiller [8] have measured 32 interface design features that influence the store traffic and sales. O'Keefe and McEachern [9] have modeled consumers' within-site decision process considering five operational stages: need recognition, information search, evaluation of alternative, purchase, and after-sales evaluation.

Bellman et al. [10] have found a typical consumer to be time starved. They recommend making websites more convenient to buy standard and repeat purchase items by providing customized interfaces. Gehrke and Turban [11] have indicated that the primary requirements for successful website design are: page loading speed, business content, navigation efficiency, security, and marketing/customer focus. An empirical study by Ho and Wu [12] suggests that besides the homepage presentation, logical support, technological characteristics (i.e., hardware and software), information characteristics and product characteristics are the factors for customer satisfaction. Hoque and Lohse [13] have observed that the ineffectiveness of the online advertisements is due to the long download times of the graphics related to the advertisements. Rhee and Riggins [14] have found that consumers with online purchasing experience believe the Web-based businesses to support all consumer activities. But the visitors who only seek information about products and services, do not believe the Web-based businesses to support all consumer activities. Salkino [15] finds two most important obstacles to online shopping to be the lack of security and the network reliability.

O'Connor and O'Keefe [16] have found that information search cost is low in case of Internet-based shopping due to the presence of news groups, and comparison shopping agents. They also found the word of mouth to spread rapidly over discussion groups, causing an accelerated diffusion curve. Jahng et al. [17] argue that the visitors are attracted to the sites that are interactive and personalized to cater to the visitor's individual need. Limayem et al. [18] have found that online shopping is facilitated by