Process Mining to Analyze the Behaviour of Specific Users

Laura Măruşter
University of Groningen, The Netherlands

Niels R. Faber
University of Groningen, The Netherlands

INTRODUCTION

As the on-line services and Web-based information systems proliferate in many domains of activities, it has become increasingly important to model user behaviour and personalization, so that these systems will appropriately address user characteristics. In this sense, particular topics are addressed by research in human-computer interaction (HCI), such as the discovering of user behaviour or navigation styles (Balajinath & Raghavan, 2001; Herder & Juvina, 2005; Juvina & Herder, 2005b; Mensalvas et al., 2003; Obendorf, et al., 2007), developing metrics involved in modelling and assessing web navigation (Herder, 2002; McEneany, 2001; Spiliopoulou & Pohle, 2001), and cognitive models for improving the redesign of information systems (Bollini, 2003; Ernst & Story, 2005; Juvina & Herder, 2005a; Juvina et al., 2005b; Lee & Lee, 2003).

Various methods have been developed to model web navigation in case of generic users (Eirinaki &Vazirgiannis, 2003). The existence of systems and/or interfaces neglecting specific user groups results into low performance of these systems, which requires further redesign. By investigating navigational patterns of specific user groups, and combining with their specific characteristics, the (re)design of the systems can be made more effectively. In this chapter, farmers have been considered as a specific user group. However, the methodology discussed in this chapter can be used also in case of other specific user groups.

Focusing on farmers as a specific IT user group, becomes an important research issue (Thysen, 2000). Farmers show a low rate of management software adoption (Alvarez & Nuthall, 2006). Different projects have been initiated to support farmers, to pursue their decision-making activities with the aid of Information Systems (see Fountas, Wulfsohn, Blackmore, Jacobsen, & Pederson, 2006; US North Central Research in Farm Information Systems, 2000). Kuhlmann & Brodersen (2001) and Hayman (2003) express their pessimism about the fast diffusion of complex information technology tools and decision support systems (DSS) among farmers. Various studies aimed to find factors that hamper adoption of DSSs in agriculture (Faber, Jorna, Van Haren, & Maruster, 2007; Kerr, 2004; Kuhlmann & Brodersen, 2001). Alvarez & Nuthall (2006) conclude that "software developers must work with farmers, both in design, and training and support, and the system must be configurable to suit a range of farmer characteristics". Therefore, there seems to be a real need to personalize these systems, such that they address farmers' characteristics.

Personalization of website design that supports DSS systems, to incorporate user characteristics, enhances effectiveness and usage of these systems. "The goal of personalization is to provide users with what they want or need without requiring them to ask for it explicitly" (Mulvenna, Anand, & Buchner, 2000).

The enhancement of website effectiveness is especially relevant in case a website offers access to underlying systems that aim to provide support and advice to its users. For instance, Jensen (2001) analyzed the usage of a web-based information system for variety selection in field crops. He compared four user groups by constructing measures based on logged information. This analysis reveals interesting similarities and differences in behaviour concerning the four groups. However, no insights could be given about the most typical sequence of behaviour/navigation patterns, such that it could support the redesign of the system.

In this chapter, a methodology of investigating user behaviour/navigation patterns by employing process mining technique is presented (Maruster & Faber, 2007). This methodology consists of (i) obtaining user groups considering some performance criteria, (ii) determining navigational patterns for all users and

for each user group, and (iii) deriving implications for redesign. In the following section, this methodology is presented, using an example involving starch potato farmers. We conclude this chapter with future trends and conclusions.

PROCESS MINING FOR PERSONALIZATION

Web Usage Mining by exploiting web access logs targets General Access Pattern Tracking and Customized Usage Tracking. *General Access Pattern Tracking* is using KDD (Knowledge Discovery in Data) techniques to understand general access patterns and trends of web usage, while *Customized Usage Tracking* analyses access patterns of each user at a time (see Zaïane, 2007). Customized Usage Tracking can be used to provide recommendations to web users via personalization. Personalization can be done by distinguishing between different users or group of users, which is called *User profiling* (Eirinaki &Vazirgiannis, 2003).

Developing User Groups

Farmers are considered as a specific group of professional users, characterized by a modest use of IT. In the context of using applications (e.g. decision support systems) developed on Web-service platforms, personalization decisions are not straightforward and have to be based on knowledge about the user. Applications developed for farmers such as decision support systems (DSS) cannot consider them either as generic users (they have low experience with IT, lack of interest to use IT), or individual users (because DSS are complex systems that are hardly customizable for individual users).

Clustering is a common technique used in user profiling (Eirinaki &Vazirgiannis, 2003). Using two-step clustering method[1], Dutch starch potato farmers have been split up into four groups based on performance characteristics, where quantity and quality of a farmer's yield are the used dimensions for cluster analysis, both averaged over the last three years (Faber, Peters & Jorna, 2006). The found clusters have been labelled

Figure 1. Farmer clusters