



# Pattern-based causal relationships discovery from event sequences for modeling behavioral user profile in ubiquitous environments



Belkacem Chikhaoui<sup>a,\*</sup>, Shengrui Wang<sup>a</sup>, Tengke Xiong<sup>a</sup>, H el ene Pigot<sup>b</sup>

<sup>a</sup> Prospectus Laboratory, Faculty of Science, University of Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

<sup>b</sup> Domus Laboratory, Faculty of Science, University of Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

## ARTICLE INFO

### Article history:

Received 28 February 2013

Received in revised form 31 March 2014

Accepted 21 June 2014

Available online 28 June 2014

### Keywords:

User profile

Causal relationship

Transfer entropy

Significant pattern

Ubiquitous computing

Behavior sequence

## ABSTRACT

This paper presents a novel and practical model for behavioral user profile modeling using causal relationships. In this model, causal relationships, which represent the influence among variables, are discovered from event sequences representing users behaviors, and used for modeling behavioral user profiles. Our model first discovers significant patterns using probabilistic suffix trees, and then discovers pattern correlations using a new sequence clustering algorithm and a modified version of the normalized mutual information (NMI) measure. Causal relationships between the significant patterns are then discovered using the transfer entropy approach. These relationships are used to construct the causal graphs of activities to generate user profiles. Through extensive experiments over a variety of datasets, we empirically demonstrate that these causality-based profiles lead to significant improvement of performance in activity prediction and user identification. We also show that our proposed model is generic and effective in constructing individual user profiles and common profiles for groups of users, in indoor and outdoor environments.

  2014 Elsevier Inc. All rights reserved.

## 1. Introduction

User profiling is a particularly active and challenging research area. It plays a central role in many application domains, such as healthcare, security, e-business, finance, and social media, including World Wide Web access and social networking [48,64]. In these domains, personalized services are increasingly required to satisfy a wide variety of user needs. However, tailoring services to individual users' needs cannot be achieved without studying individual user profiles.

According to [15], several definitions exist for a user profile. It can be defined as the description of the user's interests, characteristics, behaviors and preferences that can be obtained statically using questionnaires and interviews, or dynamically using machine learning and data mining techniques. The former type is called a knowledge-based profile, and the latter, a behavior-based profile [48]. The two types of profiles are complementary. However, collecting sufficient information via questionnaires and interviews is a difficult, obtrusive, and time-consuming task, which makes the knowledge-based profile difficult to implement. Behavior-based profiling aims to automatically discover user information, and employs behavioral logging to extract useful patterns in the user's behavior [48]. This makes the behavior-based approach more attractive for user profile construction.

\* Corresponding author. Tel.: +1 8198218000.

E-mail address: [belkacem.chikhaoui@usherbrooke.ca](mailto:belkacem.chikhaoui@usherbrooke.ca) (B. Chikhaoui).

The behavior-based profile is most commonly used in applications such as recommender systems. For instance, the profile can be constructed by mining the transaction history and ratings of booked items in order to personalize services, improve access and recommend items [9]. Behavior-based profiling is very important in healthcare-related services. For instance, doctors and caregivers are particularly interested in how impaired people perform their daily activities, in order to assist them and intervene when necessary. In this situation, therefore, in addition to preferences and interests, the user profile includes the way people perform activities. Similarly, an important goal for researchers in smart homes is to be able to provide assistance for people with special needs, and analyze the influence and effect of this assistance and the extent of feedback from users, in order to ensure and improve their well-being. User behaviors in this case might thus be caused by the kind of assistance provided.

User behavior is composed of performed actions. Most often, these actions take the form of streams or sequences of events ordered in time. Each sequence is associated with some activities. Because of physical, mental, cultural, and lifestyle differences [65], activities can be performed differently by users. The variations in the users' behavior are reflected by variations of patterns in the action sequences. By discovering these patterns and their underlying associations, it becomes possible to characterize users' behavior and understand the causes that generate it, which helps in predicting actions and providing users with personalized services.

In the existing approaches, user profile construction remains simplistic, taking the form of gathering preferences and interests by using data mining and machine learning techniques on vectorial [69,47] or graphical [48,19,1,7,4,12] representations. Several researchers have investigated graphical models such as Bayesian networks for constructing user profiles [68,71,33]. However, the proposed approaches in the literature have limitations and fail to model some phenomena such as bidirectionality between variables. For example, according to [11], social support has been shown to decrease emotional stress. But emotional distress has also been shown to decrease social support. Similarly, in a recent study conducted by [66], researchers discover a bidirectional relationship between schizophrenia and epilepsy. This study reports that patients with epilepsy are more likely to develop schizophrenia and those with schizophrenia are more likely to develop epilepsy. These facts cannot be modeled using Bayesian networks, which are directed acyclic graphs. This motivates us to choose causal modeling as a powerful statistical approach that overcomes these limitations and is capable of modeling more complex situations. To the best of our knowledge, no work has been reported on constructing user profiles comprising causal relationships between activities that represent behaviors or effects, such as purchases in the case of recommender systems, or abnormal behaviors in the case of monitoring and assisting impaired people.

Discovering meaningful causal relationships from datasets is a challenging and important task. Because causal modeling has proved so useful in simulation applications [14], it has been widely used in fields like biology [29], economics [16], neuroscience [62], etc. The goal in analyzing the cause-effect relationships in human behavior is to identify and understand the reasons why certain activities happen. Studying the user profile from the cause-effect perspective is a novel approach that opens up new research directions for user and profile modeling. It is receiving a great deal of attention from the areas of finance, e-business and recommender systems, in order to predict market prices, risks and user purchases [39]. This demonstrates the importance of studying causal relationships to analyze and interpret different situations.

The purpose of this paper is to address the problem of building a behavioral user profile by using causal relationships discovered from activity sequences. This problem involves two main issues:

1. Discovering user behavioral patterns from sequences.
2. Discovering causal relationships between the discovered behavioral patterns and constructing causal graphs constituting user profiles.

The sequences studied in this paper are composed of symbols that can represent events, user logs, transactions, GPS locations, or biological data. Given the variability of user behaviors, a user can have different profiles. Each user profile can be represented as a causal graph. Therefore, a user can have different causal graphs. Our work is motivated mainly by the need to process, understand and utilize the behavioral data available in many practical areas such as healthcare, smart homes and social networks. We will show that causality analysis allows the development of innovative and practical solutions to user profiling.

In order to effectively construct user profiles, several issues and challenges need to be addressed: (1) Heterogeneity of datasets and users: A user profile should be generic and take into account different types of datasets. (2) Data representation and dimensionality: Datasets are characterized by a large number of variables related to different sources, which need to be filtered and reduced to only those that are of interest.

In this paper, we propose a new user profile model based on causality analysis in order to address the above challenges, resulting in the following primary novel contributions:

- Proposing a generic and practical model for behavioral user profile construction, applicable in indoor and outdoor environments.
- Introducing a new method for optimizing the probabilistic suffix tree depth and detecting significant behavioral patterns in sequences.
- Introducing a new clustering method based on significant behavioral patterns in sequences.
- Detecting pattern dynamics and directionality of influence between patterns by means of transfer entropy.
- Conducting extensive experiments with real datasets to demonstrate the effectiveness and suitability of our model.

Download English Version:

<https://daneshyari.com/en/article/392660>

Download Persian Version:

<https://daneshyari.com/article/392660>

[Daneshyari.com](https://daneshyari.com)