Information Sciences 330 (2016) 145–156

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A modular approach to landmark detection based on a Bayesian network and categorized context logs

Sungsoo Lim, Seung-Hyun Lee, Sung-Bae Cho*

Department of Computer Science, Yonsei University, 50 Yonsei-ro, Sudaemoon-gu, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 2 December 2013
Revised 25 August 2015
Accepted 2 October 2015
Available online 19 October 2015

Keywords:
Landmarks
Modular Bayesian network
Categorized context logs

ABSTRACT

Mobile context logs contain meaningful and private information about their owners that can be used to support users' human memory. However, it is difficult to efficiently retrieve the information because of the enormous amount of mobile context logs and the limitations of mobile devices in terms of power, memory capacity, and speed. To efficiently retrieve information, detection of important events or landmarks is required. In this paper, we propose a modular approach of a Bayesian network for landmark detection using categorized context logs. The proposed model consists of several modules of Bayesian networks used to reduce the time of inference and the size of memory used, and each module is learned using categorized context logs according to the days of the week in order to decrease learning time and increase accuracy. Our experiments on Nokia log data and our life-log data show that the modular approach is superior to a monolithic Bayesian network and confirm that using categorized context logs for learning enhances the inference performance.

1. Introduction

With advances in mobile device technology, storage sizes are increasing, sensors are becoming smaller and smaller, and prices are decreasing. Along with their diverse and useful services and their customization to reflect user preferences, mobile devices can also collect personal information relevant to various contexts. As people carry their mobile devices with them, the devices can continually and for long periods of time collect context logs containing meaningful and private user information. Users' locations can be gathered through GPS sensors, information such as temperature and humidity can be obtained by environmental sensors, and users' activities can be inferred from diverse sensors.

Several research groups are attempting to better store and manage the context logs of mobile devices, to provide users with a variety of smart services in the real world, such as intelligent calling services [27], messaging services [22], and analysis, collection and management of mobile logs [1,9,16,19,24,25,26,29]. The MyLifeBits project is one implementation of a personal record database system [8]. Personal information is collected by PC, SenseCam, and so on and stored in an MS SQL server. However, a user would have difficulty exploring and searching the contents because of the large amount of personal data. Eagle tried to develop a diary system based on log information collected from cellular phones [5]. This system showed raw information directly on a GUI, which made it difficult to intuitively understand the big picture of a given day. Nokia's Lifeblog service gives users a way to store and manage mobile contents.[1] It automatically collects all the photos, videos, and sound clips that the user creates

* Corresponding author. Tel.: +82 2 2123 2720; fax: +82 2 365 2579.
E-mail addresses: lss@sclab.yonsei.ac.kr (S. Lim), e2sh83@sclab.yonsei.ac.kr (S.-H. Lee), sbcho@cs.yonsei.ac.kr, sbcho@yonsei.ac.kr (S.-B. Cho).

[1] Nokia LifeBlog. <http://www.nokia.com/lifeblog>.

http://dx.doi.org/10.1016/j.ins.2015.10.017

on the mobile phone, as well as text messages and MMS messages sent and received, and it stores the log data and shares them with other people. However, this service does not include any abstraction or summarization methods.

Although the mobile context logs contain many kinds of information that can support users' human memory, it is difficult for users to efficiently retrieve it because of the enormous amount of data in mobile context logs and the limitations of mobile devices in terms of power, memory capacity, and speed. A key principle of human memory is that it is organized for episodic storage and retrieval. Our brains group related events as episodes and use landmark events to recall these episodes. Finding such landmark events can elicit recall of related items. The goal of this paper is to design a landmark detection model for use with the context logs in mobile environments.

Landmarks can be used to aid recall of specific functions [28]. Landmarks are derived from the concept of retrieval cues, which are events or experiences that facilitate retrieval of information from episodic memory because of their associations with that information. The context information is reconstructed in terms of episodic memories and used for searching for the user's personal information. Landmarks are made up of the location and time of the event, the people who experienced it, and the events that occurred before and after the event. These landmarks help users remember the episode. In AniDiary, Cho et al. used landmarks as a key aspect of summarizing a user's daily life in a cartoon-style diary [4].

From this perspective, Eagle and Pentland analyzed the user's life pattern by using a hidden Markov model with the context log from a smartphone, and developed models about users with respect to specific locations [6]. They recognized social patterns in the user's daily activities, inferred relationships, identified socially significant locations, and modeled organizational rhythms. Hovitz, Dumais, and Koch attempted to reorganize personal information storage in desktop PCs in terms of an episodic style of memory [10]. For landmark detection, they built a single Bayesian network using an online calendar and user information data from the desktop environment. The calendar information, which consisted of 28 nodes, was learned by using Beckeman and Meek's learning method of Bayesian networks with local structure [3].

In this paper, we define the landmarks, composed of episodes, as key log data used to search for context information efficiently. We propose a modular approach based on Bayesian networks (BNs) for landmark detection using categorized context logs. The proposed method adopts a Bayesian probabilistic model to efficiently manage uncertainties in real-world environments, such as sensor inaccuracies and uncertain causal factors. To reduce the time of learning and inferring and the size of the memory used, the proposed model employs several modules of Bayesian networks, where each module learns using context logs categorized by the days of the week. We discuss how to infer the modular BNs (MBNs) using selective reasoning, and how the MBNs learn from the categorized context logs. We evaluate the efficiency of the proposed method in terms of time and accuracy through a comparative study using the Nokia log data.

2. Related works

There have been various attempts to analyze log data and to support expanded services by using a probabilistic approach. Li and Ji used a probabilistic model of detecting the user's active affective state [21]. They used a dynamic Bayesian network and utility theory to reason about fatigue, nervous, and confused states. They showed that the probabilistic approach was good for managing uncertain information like affect.

Ji et al. also used a dynamic Bayesian network model for real-time monitoring of human fatigue [15]. They considered and incorporated many conditions, such as light, heat, humidity, time, napping, anxiety, temperature, weather, and work type. Zhang and Ji proposed an active and dynamic information fusion method for multi-sensor systems with dynamic Bayesian networks [32]. They showed the usefulness of the Bayesian approach for information fusion. These works show that the Bayesian probabilistic approach is good at handling, reasoning about, and combining uncertain information [18].

Krause et al. clustered sensor and log data collected on mobile devices. They discovered a context classifier that reflected a given user's preferences, and they estimated the user's situation in order to provide smart services [19]. Their context classifier was constructed using the BN model, which was based on a general learning method for a small domain of classification subjects. Horvitz et al. proposed a method that detected and estimated landmarks from a person's cognitive activity model based on PC log data and the Bayesian approach [10]. Their method achieved good performance for recognizing and learning people's everyday PC life.

However, these methods are not suitable for mobile devices having limited capacity and power. For larger domains, the general BN and BN learning method require very heavy computation. This is a crucial problem when it comes to modeling everyday life situations with mobile devices. Thus, a more appropriate approach is required to reduce the complexity. Marengoni et al. tried to reduce the complexity of the BN model by dividing it into several multi-level modules and using procedural reasoning by the connected BNs (just like chain inference) [23]. However, this method required procedural and classified properties of the target functions.

Tu et al. proposed a hybrid BN model that enabled hierarchical hybridization of BNs and HMMs [31]. However, it supported links only from lower-level HMMs to higher-level BNs, without consideration of links between BNs of the same level. Hwang et al. proposed a hierarchical probabilistic graphical model that constructed a hierarchical and distributed BN structure using generated hidden nodes and the links between them [11]. However, this method is unsuited to mobile and private service environments because it leads to an increase in the number of nodes and does not retain an intuitive causal structure.

Gao et al. proposed an automatic landmark-ranking method to select the landmarks, and developed a travel guidance system W2Go to recognize and rank landmarks for travelers [7]. Ji et al. presented a system to mine landmarks from geographically