Special Section on

Data Mining in Pervasive Environments

Guest Editors:

Dr. Nirmalya Roy, University of Maryland Baltimore County, (nroy@umbc.com)
Dr. Parisa Rashidi, University of Florida, (parisa.rashidi@ufl.edu)
Dr. Larry Holder, Washington State University, (holder@wsu.edu)
Dr. Liming Chen, De Montfort University, (liming.chen@dmu.ac.uk)

Purpose and Significance:

The prevalence of sensors and mobile phones and their interaction with human life funnel a phenomenal amount of data through pervasive computing environments. The data from these sensors including ambient sensors such as motion sensors, smart phone sensors like accelerometers and GPS and object sensors such as RFID tags have to be carefully analyzed to extract meaningful and actionable information. The sheer volume of sensor data, as well as its dynamic and distributed nature, poses many challenges to the data mining, mobile sensing and knowledge discovery communities. Analyzing these data trails can support different applications in a novel way. The applications vary from personal and community healthcare (smart home independent living, fitness and exercising), green computing (building energy management, environment monitoring), urban sensing (intelligent transportation system, natural resource management), marketing industry (advertisement, consumer shopping habits) and after all social networking.

Papers were solicited from researchers working on data mining in pervasive environments to highlight current challenges, and to showcase the latest research results. The results demonstrate how current data mining, mobile sensing and knowledge discovery methods are extended to mining solutions for dealing with challenging real world problems. The unique nature of sensor data in pervasive environments demands novel data mining, mobile sensing and knowledge discovery methods that can handle large, multi-modal, heterogeneous and distributed streams of data. It is expected that the published results of this special issue not only can be beneficial for the pervasive computing community, but the resulting algorithms and solutions will also be adapted by researchers in various other application fields.

We received 22 papers in response to the solicitation, of which we eventually accepted 7 papers for publication. These seven papers cover multiple different areas relating to data mining in pervasive environments, namely energy usage predictability, behavior analysis, understanding the social aspects of pilgrimage, and activity recognition.

The first paper “A System for Collecting Activity Annotations for Home Energy Management” by Sami Rollins, Nilanjana Banerjee, Lazeeb Choudhury, and David Lachut proposes a system to enable the user to draw connections between activities and energy consumption by attaching contextual labels to energy events. They build a system for collecting in situ annotations using a smart phone application coupled with an off-the-shelf home energy measurement infrastructure. The paper considers a power profiling approach to identify important energy consumption events and solicits contextual annotations from the users via a push notification sent to a smart phone.

The second paper “Understanding Aspects of Pilgrimage using Social Networks derived from Smartphones” by Amir Muaremi, Agon Bexheti, Franz Gravenhorst, Julia Seiter, Sebastian Feese, Bert Arnrich, and Gerhard Troster investigates the behaviors of pilgrims using social network analysis, and
spatial proximity information as obtained from ANT+ and Bluetooth system on a smartphone. They consider a unified framework and integrate auxiliary proximity information from social rules of pilgrims, existing wearable devices and GPS location information. The proximity system has been used to identify pilgrims’ social networks around the prayers followed by the social network analysis to differentiate leaders from followers among the pilgrims. The system has been validated using data from 41 participants during an 8 day pilgrimage.

The third paper “Behavior Analysis of Elderly using Topic Models” by Kristin Rieping, Gwenn Englebienne, and Ben Krose describes topic models for behavior analysis of elderly people. The paper considers a novel way to cluster activity data into topics, jointly optimizing the sensor data and the topic distributions over a day. They introduce two novel models of the sensor data, namely LDA-Gaussian and LDA-Poisson. The paper also describes experimental setup using real-world, binary sensor data, gathered from different houses of elderly people.

The fourth paper “b-COELM: A Fast, Lightweight and Accurate Activity Recognition Model for Mini-Wearable Devices” by Lisha Hu, Yiqiang Chen, Shuangquan Wang, and Zhenyu Chen examines the computational challenges of activity recognition on small devices. To meet the low-computational-complexity, lightweight implementation constraints, this paper considers an activity recognition model based on a proximal support vector machine and extreme learning to devise novel multi-class classification problems. The paper also provides a validation of the ideas using real activity data collected from a set of wearable devices.

The fifth paper “Discovery of Activity Composites Using Topic Models: An Analysis of Unsupervised Methods” by Julia Seiter, Oliver Amft, Mirco Rossi, and Gerhard Troster articulates unsupervised activity discovery approaches using three topic model (TM) approaches, based on Latent Dirichlet Allocation (LDA), ngram TM (NTM), and correlated TM (CTM). The authors introduce NTM to incorporate frequently occurring primitive sequence similarity information into the activity discovery process and CTM to deal with co-occurring topics. The paper compares TM approaches with others including k-means clustering and supports model selection under ideal and noisy primitive input conditions.

The sixth paper “An Automatic Data Mining Method to Detect Abnormal Human Behaviour using Physical Activity Measurements” by Juan Luis Carúns Candás, Victor Pelaez, Miguel Angel Fernandez, Gloria Lopez, Eduardo Alvarez, and Gabriel Díaz investigates the problem of detecting abnormal behavior in a pervasive computing environment. The paper builds a human behavior model according to the measured physical activity levels and detects abnormalities when the measured behavior is not expected according to the historical measured behaviors. The methods have been validated using real users data traces.

The last paper “Using Unlabeled Data in a Sparse-coding Framework for Human Activity Recognition” by Sourav Bhattacharya, Petteri Nurmi, Nils Hammerla, and Thomas Ploetz postulates a sparse-coding framework for activity recognition. The paper describes the derivation of a compact, sparse and meaningful feature representation of sensor data that does not rely on prior expert knowledge and generalizes well across domain boundaries. The authors also consider exploiting unlabeled sample data for bootstrapping effective activity recognizers. The paper shows the practical potential of the new approach by successfully evaluating its generalization capabilities across both domain and sensor modalities by considering the popular Opportunity dataset.

We believe that the research presented in this special issue is of considerable interest to the PMC readership concerned with the role of data mining and knowledge discovery in pervasive computing and that these articles spawn further research on this important subject.

We would like to thank the authors of the paper for their interesting work and the submission to this special section and their effort in revising the papers – sometimes multiple times – in response to the reviews. We would also like to thank them for their patience with the review process. We would like to thank the reviewers for doing an excellent job of reviewing the papers. Finally, we would like to thank
the PMC special issues editor, Dr. Behrooz Shirazi for his guidance and support during the long process of review and paper selection.

Nirmalya Roy, University of Maryland Baltimore County
Parisa Rashidi, University of Florida
Larry Holder, Washington State University
Liming Chen, De Montfort University