ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

A Review on Extensible Regular Humanistic Observable Impression Scooping Multivarious Physical Facts

Sneha D Patil[1], Dinesh D Patil[2]

ME Scholar, Department of Computer Science, SSGB College of Engineering and Technology, Bhusawal, India[1]

Professor, Department of Computer Science, SSGB College of Engineering and Technology, Bhusawal, India [2]

ABSTRACT: Despite the advent of wearable devices and the proliferation of smart phones, there still is no ideal platform that can continuously sense and precisely collect all available contextual information. Mobile sensing data collection approaches should deal with uncertainty and data loss originating from software and hardware restrictions. We have conducted life logging data collection experiments from many users and created a rich dataset (7.5 million records) to represent the real world deployment issues of mobile sensing systems. We create a novel approach to identify human behavioral motifs while considering the uncertainty of collect data objects. Our work benefits from combinations of sensors available on a device and identifies behavioral patterns with a temporal granularity similar to human time perception. Employing a combination of sensors rather than focusing on only one sensor can handle uncertainty by neglecting sensor data that is not available and focusing instead on available data. Moreover, we demonstrate that using a sliding window significantly improves the scalability of our analysis, which can be used by applications for small devices such as smart phones and wearable.

KEYWORDS: Frequent pattern mining, temporal granularity, Multivariate temporal data, Human centric data

I. INTRODUCTION

The proliferation of smart phones and, more recently, wearable devices such as fitness trackers and smart watches equipped with sensors, has led to a significant expansion of possibilities to study human behavior. Computing and networking capabilities of these devices within their multiple sensors makes them capable enough so we can easily observe and collect useful contextual information (mobile sensing). For instance, mobile health, which benefits from mobile sensing, offers the possibility of a shift from treatment to prevention in medical care systems. Researchers show that 69% of U.S. adults monitor and track their health status and 21% of them use technology for this purpose [8]. Unlike wearable devices, which are still quite new in the market, the smart phone platform has benefited from a significant amount of scientific work ranging from personal air pollution footprint trackers applications [15] to wellbeing [13]. Both wearable devices and smart phones are very capable of sensing and collecting basic patterns of human behavior and collecting contextual information.

While human behaviors are predictable, at least in aggregate [1], traditional approaches for detecting human behavioral patterns (which are not digital) are often difficult. However, the advent of these ubiquitous devices enables researchers to identify human behavior to an extent that was not previously possible. On one hand, this information collection paradigm should be moved from simple data collection tools to intelligent systems with cognition capabilities [4]. On the other hand, there is still a lack of wide acceptance of mobile sensing applications in real world settings.

There are several reasons for this mismatch of capability and acceptance. First is the resource limitation and lack of accuracy in the collected contextual data, especially with regard to the battery life [24]. The size of sensors that are dealing with radio frequency, i.e., Bluetooth, Wi-Fi and GPS, affects the quality of their data [22] (smaller devices have less accurate data). The next reason, which has been noted but has not been widely explored, is the proximity of the smart phone to users [5]. Smart watches and wearable are body mounted and thus the proximity problem has been resolved in those devices, but they still suffer from a lack of accuracy [12]. The third reason for this problem is operating system restrictions of mobile devices, which removes background services when the CPU is under a heavy

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

load in order to preserve the battery life. As a result, there is no ideal data collection approach that can sense and record individuals information 24/7 with no data loss. The uncertainty of these data objects is a major challenge that limits the applications that can benefit from them.

II. LITERATURE SURVEY

Device Analyzer: Large-scale Mobile Data Collection -D. Wagner et al.

In the Device Analyzer project we are building a dataset that captures real-world usage of Android smartphones. We have been collecting detailed usage information in the wild for nearly 2 years from 894 models of devices running 687 versions of Android. Over 12,500 users from 167 countries have installed a copy of the software from the Android market and consented to their data being collected. In total, our dataset covers over 1,450 phone-years of usage, with days of inactivity removed. 10,320 participants contributed for at least one day, 3,680 users contributed more than one month of usage information and over and 820 participated for at least six months. The dataset contains 53 billion data points.

In this paper we present our system architecture for collecting data from a large number of distributed sources that is resilient against failures of devices. There are six conceptual components to our system and we describe the operation of each and some of the problems that arise (Section 3). We explain how offline processing can recover wall clock time in situations where traditional collection methods would fail due to errors during collection or because of user interference (Section 4). We also show how this type of offline processing fits naturally into our architecture. We conclude (Section 5) with a brief discussion of some of the more general issues and lessons learned, which we think might also apply to other data collection and analysis projects.

A Probabilistic Approach to Mining Mobile Phone Data Sequences Katayoun Farrahi

We present a new approach to address the problem of large sequence mining from big data. The particular problem of interest is the effective mining of long sequences from large-scale location data to be practical for Reality Mining applications, which suffer from large amounts of noise and lack of ground truth. To address this complex data, we propose an unsupervised probabilistic topic model called the distant n-gram topic model (DNTM). The DNTM is based on Latent Dirichlet Allocation (LDA), which is extended to integrate sequential information. We define the generative process for the model, derive the inference procedure, and evaluate our model on both synthetic data and real mobile phone data. We consider two different mobile phone datasets containing natural human mobility patterns obtained by location sensing, the first considering GPS/wifi locations and the second considering cell tower connections.

The contributions of this paper are as follows: (1) we propose the distant n-gram topic model (DNTM) for sequence modeling; (2) we derive the inference process using Markov Chain Monte Carlo (MCMC) sampling [21]; (3) we generate a dataset of synthetic sequences and apply the DNTM to test the model under a controlled setting; (4) we apply the DNTM to two real large-scale mobile phone location datasets. The model discovers user location routines over several hour time intervals, corresponding to sequences, and these results are illustrated by differing means; (5) we also perform a comparative analysis with Latent Dirichlet Allocation (LDA) [4], showing that the DNTM performs better in predicting unseen data based on log-likelihood values. This paper is an extended version of the work originally presented at [11].

Lesson Learned from Collecting Quantified Self Information via Mobile and Wearable Devices Reza Rawassizadeh 1,*, Elaheh Momeni 2, Chelsea Dobbins 3, Pejman Mirza-Babaei 4 and Ramin Rahnamoun 5

In this article, we have provided our overview of the difficulties that researchers and developers may face while developing these systems. In doing so, we report on three lifelogging data collection studies that we have undertaken. Two of these studies utilized smartphones, where participants were required to install the lifelogging app *UbiqLog* [17] on their phones. These studies included fourteen different brands of phones, among 57 participants (three participants were repeated in both studies). The third study used a smartwatch, as an information collection tool to collect physical activity, location and self-reported mood.

This paper contributes by discussing the main findings of our studies as follows: (i) to reduce churn, it is useful if the developer can *minimize the need for manual intervention*, while continuously collecting information, even optional annotation. (ii) While mobile and wearable devices collect data, there is an element of *uncertainty and data loss* that

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

originates from manual sensor configuration changes (e.g., disabling WiFi to preserve battery) or sensor quality (e.g., geographical coordinates read from Cell ID). This should be considered while analyzing the collected data. (iii) There is a *lack of multivariate reflection methods* to analyze the collected daily life information, e.g., visualizing incoming calls based on the location and time of the day. Privacy issues [18] and battery limitations [19,20] are important but known issues, and thus we do not list them as our novel findings. Nevertheless, we have tackled them from another perspective, which is worth further explanation. In particular, we have summarized these challenges in a single report, which we think could benefit the community and further research in this area.
The remainder of this paper is organized as follows: the next section describes the related work in the field. This is then followed by a description of our study materials and methods. Afterwards, we explore the challenges that the area faces. This is followed by discussions of our findings, before concluding.

III. PROBLEM STATEMENT

We live in a spatiotemporal world and all of our behaviors occur in a specific location and time. Therefore, a digital system for quantifying human behavior should sense both time and location. Since location sensors such as GPS are not reliable (especially indoor) and it is not possible to collect their data 24/7, we can only use time to link different information together. Human behavior is composed of many daily activities that are distinctive and recurring. These types of activities have been called motifs (or life routines [6]) and our goal is to create a user profile that summarizes the behavioral motifs of a person.

IV. METHODOLOGY

The contributions of our work are listed as follows:

1. an algorithm for converting digital timestamp to a temporal representation similar to human temporal cognition;
2. a model that quantities human behavior based on sensor data; and
3. algorithms that will be used to exploit daily life behavioral motif from raw sensor data and their evaluation from three different perspectives, including when to run the algorithm.

V. CONCLUSION

In this paper, we have proposed a scalable approach for daily behavioral pattern mining from multiple information sources. This work benefits from a realistic dataset and users who use different smart phone brands. We use a novel temporal granularity transformation algorithm that makes changes on timestamps to mirror the human perception of time. Our behavioral motif detection approach is generic and not dependent on a single source of information; therefore, we reduce the risk of uncertainty by relying on a combination of sensors to identify behavioral motifs and patterns. We investigate the efficiency of our work by evaluating it from three different perspectives: the execution time performance, the effect of threshold changes on motif detection, and the validity of the identified behavior from a temporal perspective. This approach is scalable enough to be used in several types of applications such as mobile health, context-aware recommendations and other quantified-self applications.

REFERENCES

1. [S. Foell et al.] " Micro-navigation for Urban Bus Passengers: Using the Internet of Things to Improve the Public Transport Experience," Urb-Iot '14.
2. [R. Dobbins and R. Rawassizadeh] "Clustering of Physical Activities for Quantified Self and mHealth Applications," in IUCC 15.
3. [A. Campbell and T. Choudhury] "From Smart to Cognitive Phones," IEEE Pervasive Computing '12, vol. 11, no. 3, pp. 7–11.
4. [D. Ferreira et al.] "Understanding Human-Smartphone Concerns: a Study of Battery Life," Pervasive Computing '11.
5. [R. Rawassizadeh et al.] "Wearables: Has the Age of Smartwatches Finally Arrived ?" Communications of the ACM '15, vol. 58, no. 1, pp. 45–47.
6. [A. Dey et al.] "Getting Closer: An Empirical Investigation of the Proximity of User to Their Smart Phones," UbiComp '11.
7. [N. Eagle and A. Pentland] "Reality Mining: Sensing Complex Social Systems," Personal and Ubiquitous Computing '06, vol. 10, no. 4, pp. 255–268.
8. [N. Kiukkonen et al.] "Towards Rich Mobile Phone Datasets: Lausanne Data Collection Campaign," ICPS '10.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

9. [S. Nath] “ACE: Exploiting Correlation for Energy-Efficient and Continuous Context Sensing,” MobiSys ’12.
10. [H. Ma et al.] “A Habit Mining Approach for Discovering Similar Mobile Users,” WWW ’12.
11. [R. Rawassizadeh et al.] “Ubiqlog: a generic mobile phone-based life-log framework,” Personal and Ubiquitous Computing ’13, vol. 17, no. 4, pp. 621–637.
12. [D. Wagner et al.] “Device Analyzer: Large-scale Mobile Data Collection,” SIGMETRICS Perform. Eval. Rev. ’14, vol. 41, no. 4, pp. 53–56.
13. [R. Rawassizadeh and A. Tjoa] “Securing Shareable Life-Logs,” SocialCom ’10.
14. [R. Rawassizadeh et al.] “Lesson Learned from Collecting Quantified Self Information via Mobile and Wearable Devices,” Journal of Sensor and Actuator Networks, vol. 4, no. 4, p. 315, 2015.
15. [J. Paek et al.] “Energy-efficient Positioning for Smartphones Using Cell-ID Sequence Matching,” in MobiSys ’11.
16. [K. Farrahi and D. Gatica-Perez] “A Probabilistic Approach to Mining Mobile Phone Data Sequences,” Personal and Ubiquitous Computing, vol. 18, no. 1, pp. 223–238, 2014.
17. [J. Zheng and L. Ni] “An unsupervised framework for sensing individual and cluster behavior patterns from human mobile data,” UbiComp ’12.
18. [T. Bhattacharya et al.] “Automatically Recognizing Places of Interest from Unreliable GPS Data Using Spatio-temporal Density Estimation and Line Intersections,” Pervasive and Mobile Computing ’14.
19. [N. Deblauwe and P. Ruppel] “Combining GPS and GSM Cell-ID Positioning for Proactive Location-based Services,” in MobiQuitous ’07.
20. [C. Zhou et al.] “Discovering Personal Paths from Sparse GPS Traces,” in JCIS ’05.
21. [N. Mamoulis et al.] “Mining, Indexing, and Querying Historical Spatiotemporal Data,” in KDD ’04.
22. [Y. Li et al.] “Mining Probabilistic Frequent Spatio-Temporal Sequential Patterns with Gap Constraints from Uncertain Databases,” in ICDM ’13.
23. [M. Ye et al.] “On the Semantic Annotation of Places in Locationbased Social Networks,” in KDD ’11.
24. [D.Wang et al.] “Human Mobility, Social Ties, and Link Prediction,” in KDD ’11.
25. [S. Isaacman et al.] “Identifying Important Places in Peoples Lives from Cellular Network Data,” in Pervasive Computing ’11, pp. 133–151.