Towards Multi-modal Anticipatory Monitoring of Depressive States through the Analysis of Human-Smartphone Interaction

Abhinav Mehrotra
University of Birmingham
University College London
United Kingdom
a.mehrotra@cs.bham.ac.uk

Robert Hendley
University of Birmingham
United Kingdom
r.j.hendley@cs.bham.ac.uk

Mirco Musolesi
University College London
United Kingdom
m.musolesi@ucl.ac.uk

Ubicomp/ISWC'16 Adjunct, September 12-16, 2016, Heidelberg, Germany.
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968299

Abstract
Remarkable advances in smartphone technology, especially in terms of passive sensing, have enabled researchers to passively monitor user behavior in real-time and at a granularity that was not possible just a few years ago. Recently, different approaches have been proposed to investigate the use of different sensing and phone interaction features, including location, call, SMS and overall application usage logs, to infer the depressive state of users. In this paper, we propose an approach for monitoring of depressive states using multi-modal sensing via smartphones. Through a brief literature review we show the sensing modalities that have been exploited in the past studies for monitoring depression. We then present the initial results of an ongoing study to demonstrate the association of depressive states with the smartphone interaction features. Finally, we discuss the challenges in predicting depression through multi-modal mobile sensing.

Author Keywords
Mobile Sensing; Depression; Anticipatory Computing; Behaviour Change Interventions.

ACM Classification Keywords
H.1.2. [Models and Principles]: User/Machine Systems; J.4. [Computer Applications]: Social and Behavioral Sciences

Introduction

Today's smartphone comes with an array of sensors and high-performance computing power. They are also carried by their owners all the time. These characteristics have not only enabled researchers to build very effective systems for passively monitoring numerous physical-context modalities such as users' location [7], physical activity [5, 6], mobile phone interaction [14, 15, 20, 17] and surrounding sound [13], but also cognitive context [22], such as mood and well-being states.

However, cognitive context is inferred mostly by employing ESM techniques, according to which users are prompted with a series of questions that are required to be responded repeatedly [16]. Past studies have shown the potential of exploiting mobile sensing data to learn and, potentially, predict the user's cognitive context [1, 2, 3, 4, 11, 8, 12]. For example, Canzian et al. have used mobility data to monitor depressive states [4] and Alvarez-Lozano at al. have exploited application usage logs to monitor patients affected by bipolar disorder [2].

The key limitation of these approaches is the fact that they rely on a single data source for monitoring depressive states. Instead, in this paper, we argue that depressive states should be monitored via *multi-modal sensing* as different modalities might collectively be more informative: this might contribute to the improvement of the performance of the machine learning algorithms. A similar approach has been employed by LiKamWa et al. [12] for monitoring user's behavior, such as happiness and activeness levels. The authors use location, application usage, calls, SMS and Web usage data together as inputs of a machine learning tool for predicting the current mood of a user. However, the authors of this work focus on predicting mood levels using the Circumplex mood model [21], which essentially aims at measuring two dimensions: the pleasure dimension (measuring how positive or negative one feels) and the activeness dimension (measuring how one is likely to take an action under the mood state).

Our focus instead is on monitoring *depressive states* that are quantified by means of a PHQ-8 questionnaire [9, 10] asked over a period of 14 days. Moreover, the authors of [12] do not consider the role of *micro-interactions* such as clicks, scrolls and many others, and reactions to notifications.

Finally, in this work, we will not only discuss the issues related to *multi-modal sensing* but also the aspects related to the design of *anticipatory* techniques for predicting future depressive episodes [19] and the *development of behavior interventions* based on these predictions [18].

Measuring Mobile Phone Interaction and Depression Scores

In order to study the association of users' depressive states on their mobile interaction behavior, we conducted a longitudinal field study. More specifically, we developed an Android app and collected smartphone interaction data from 25 participants for a time period of 30 days. The collected data include logs for notification handling and phone usage.

We use several smartphone interaction metrics to analyze the collected data, capturing various dimensions, from application usage to the number of clicks on the screen. The metrics are summarized in Table 1.

We also collect the responses to the PHQ-8 questionnaire from the users via an ESM approach. This data is then used to compute the depressive scores for each users. Finally, we compute the Kendall's Rank correlation coefficients to analyze the relationship between the severity of

Group	Metric	Description
Notifications	Count	Total number of notifications clicked.
	Acceptance %	Percentage of notifications clicked out of total arrived.
	% Handled (Other Device)	Percentage of notifications that are not handled on phone out of total notifications arrived.
	Average Seen Time (ST)	Average of the seen time of all notifications. Here, seen time is the time from the notification arrival until the time the notification was seen by the user.
	Average Decision Time (DT)	Average of the decision time of all notifications. Here, decision time is the time from the moment a user saw a notification until the time they acted upon it (by clicking, launching its corresponding app or swiping to dismiss).
	Average Response Time (RT)	Average of the response time of all notifications. Here, response time is the sum of seen and decision times.
Phone Usage	Launch Count	Number of times applications are launched.
	App Count	Number of applications launched.
	App Usage Time	Time duration for which applications were used.
	Sig Launch Count	Number of times significant applications are launched.
	Sig App Unique Count	Number of significant applications launched.
	Sig App Usage Time	Time duration for which applications were used.
	Non-Sig Launch Count	Number of times non-significant applications are launched.
	Non-Sig App Count	Number of non-significant applications launched.
	Non-Sig App Usage Time	Time duration for which applications were used.
	Phone Usage Time	Time duration for which phone was used.
	Click Count	Number of clicks on the phone screen.
	Long Click Count	Number of long clicks on the phone screen.
	Unlock Count	Number of times the phone was unlocked.

Table 1: Description of phone interaction metrics.

the depressive state and phone interaction metrics (based on notification, application and phone usage).

Preliminary Results

Past studies have used mobility, activity, application usage and communication data for inferring depressive state of users [8, 2, 4]. We hypothesize that there are additional features (which can be captured via mobile phones) that are associated with the changes in the user's depressive state. Therefore, in this paper we present the initial findings of our ongoing study to investigate the impact of depressive state on the micro-interaction data including notification and phone interaction data.

Interpretation of Correlation Plots

The correlation results are presented as a plot of the correlation matrix. In this matrix the x-axis indicates the notification and phone interaction metrics and y-axis indicates the days for which the metrics are computed. For instance, in Figure 1 the box in the first column (*Count*) and first row

(-*14*) presents the absolute value of the coefficient for the correlation of the depression score with the acceptance percentage of notifications, computed by using data from the current day to 14 days before the current day. Here, the current day refers to the day in which users reported their depressive state via PHQ-8 questionnaire. Moreover, we set the significance level α for the correlation results to 0.001 and non-significant correlation coefficients are indicated by the white boxes in the correlation plots.

Depressive State and Notifications
In Figure 1 we show the correlation coefficients that are computed to assess the relationship between depression score and notification metrics. The results show that users' depressive state moderately correlates with all metrics that are computed by using the past 14 days of data. The correlation results are the same for the metrics that are computed with past 7 days of data, except that the average DT has a weak correlation. On the other hand, users' depressive state does not correlate with the average ST and DT of notifications arriving on the current day when the user responded to the PHQ-8 questionnaire. Moreover, other metrics computed with the current day's data have a weak correlation with the depression score.

Depressive State and Phone Usage
In order to quantify the association between users' depressive state and their phone usage pattern we compute the correlation coefficients and present the results in Figure 2. The results show that users' depression score moderately correlates with all the metrics that are computed by using the past 14 days of data. On the other hand, users' depressive state weakly correlates with most of the metrics computed with the data of past 7 days and there is non-significant correlation with the metrics computed with the data of the current day.

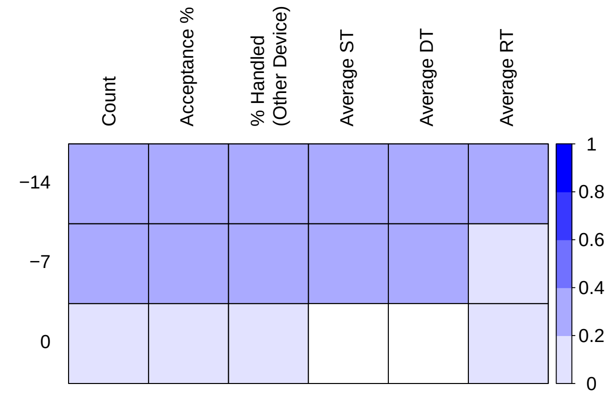

Figure 1: Results for correlation between depressive state and notification metrics.

Open Research Challenges

In order to implement robust solutions for monitoring and predicting users' depression states in real-world settings, there are various challenges that have to be addressed. The general challenge of improving prediction accuracy could indeed be addressed by employing multi-modal sensing approaches as discussed in this paper. In this section, we concentrate on additional key aspects that are of fundamental importance for realizing this vision.

Large-scale Evaluation. Most of the previous studies are performed with a small number of participants [2, 3, 4, 11, 12]. This is due to the fact that the task of recruiting participants for such kind of studies is not just difficult but there is also non negligible ethical implications, given the type of condition(s) that might affect participants. For example, reminding a participant that is affected by depression about his/her condition might have a negative impact on him/her. Continuous assessment of the participants by the

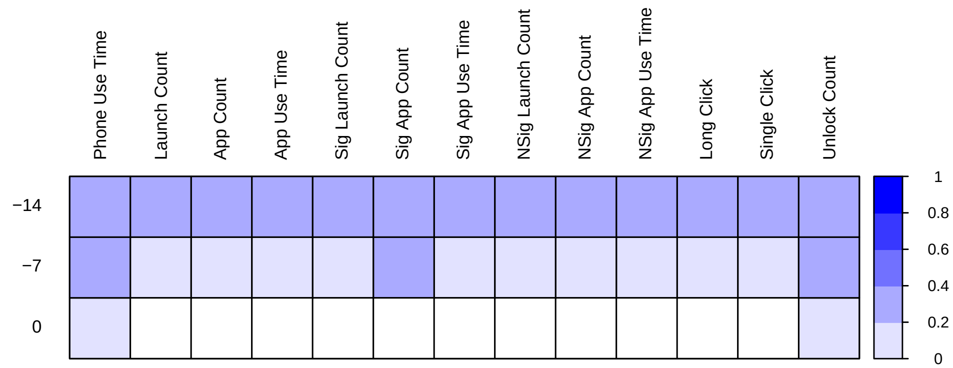

Figure 2: Results for correlation between depressive state and phone usage metrics.

experimenters is then needed and this might limit the size of deployments. Automatic assessment solutions might be adopted. This is an emerging research area and standard “good practices" have not been established yet by the community. It is worth noting that the medical research community has developed well-defined protocols for conducting these types of studies: we believe there is a need to adapt them to this new emerging technological scenario.

Validation of Results. Almost all previous studies are based on a first phase of data collection and then on a second phase focused on the development and validation of prediction models, which is usually performed off-line [4, 12]. This might be acceptable for monitoring mental health conditions and well-being of users, but not for the design of behavior intervention solutions. In these cases, *in-the-wild* evaluations are essential. Moreover, we believe that it is fundamental to test such systems on both patients affected by depression and the general public. This approach can give us a clearer picture about the robustness of the prediction models and their ecological validity.

Quantifying Causality. Existing studies on monitoring users’ well-being and depressive states via smartphones rely on correlation analysis to understand the association of well-being with different aspects of context modalities and users’ interactions with smartphones. A key challenge is to uncover the *causal links* between the depressive states and the smartphone interactions and context modalities. Understanding the causal effects between them would enable ubiquitous computing researchers as well as doctors to build more effective behavior interventions. We are currently investigating the causal impact of depressive states on smartphone interactions and context modalities. We plan to exploit the quasi-experimental approach-based framework for quantifying causality as proposed by Tsapeli and Musolesi in [23].

Summary

In this paper, we have proposed a multi-modal sensing approach to monitor and predict user’s cognitive context. We have presented the initial findings of our ongoing study concerning the impact of depressive state on user’s interaction with a smartphone. The results are based on the data collected from 25 participants for a period of 30 days. Our results suggest that by using sensed data (including both notification and phone interaction data) of the past 14 days, it is possible to improve the accuracy of the prediction of the current day's depression score of a user. Finally, we have discussed the key challenges in developing robust techniques for monitoring and predicting depressive states through multi-modal mobile sensing.

Acknowledgements
This work was supported by EPSRC through the grants "MACACO: Mobile context-Adaptive CAching for Content-Centric networking" (EP/L018829/2) and "Trajectories of Depression: Investigating the Correlation between Human Mobility Patterns and Mental Health Problems by means of Smartphones" (EP/L006340/1).

REFERENCES
1. Saeed Abdullah, Mark Matthews, Ellen Frank, Gavin Doherty, Geri Gay, and Tanzeem Choudhury. 2016. Automatic Detection of Social Rhythms in Bipolar Disorder. *Journal of the American Medical Informatics Association* 23, 3 (2016), 538–543.
2. Jorge Alvarez-Lozano, Venet Osmani, Oscar Mayora, Mads Frost, Jakob Bardram, Maria Faurholt-Jepsen, and Lars Vedel Kessing. 2014. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. In *PETRA'14*.
3. Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. Designing Mobile Health Technology for Bipolar Disorder: a Field Trial of the Monarca System. In *CHI'13*.
4. Luca Canzian and Mirco Musolesi. 2015. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *UbiComp'15*.
5. Tanzeem Choudhury, Sunny Consolvo, Beverly Harrison, Jeffrey Hightower, Anthony LaMarca, Louis LeGrand, Ali Rahimi, Adam Rea, Gaetano Borriello, Bruce Hemingway, Predrag Klasnja, Karl Koscher, James A. Landay, Jonathan Lester, Danny Wyatt, and Dirk Haehnel. 2008. The Mobile Sensing Platform: An Embedded Activity Recognition System. *IEEE Pervasive Computing* 7, 2 (2008), 32–41.
6. Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, and others. 2008. Activity Sensing in the Wild: a Field Trial of UbiFit Garden. In *CHI'08*.
7. N. Eagle, A. Clauset, and J. Quinn. 2009. Location Segmentation, Inference and Prediction for Anticipatory Computing. In *AAAI'09*.
8. Agnes Grünerbl, Patricia Oleksy, Gernot Bahle, Christian Haring, Jens Weppner, and Paul Lukowicz. 2012. Towards smart phone based monitoring of bipolar disorder. In *mHealthSys'12*.
9. Kurt Kroenke, Robert L Spitzer, Janet B. W. Williams, and Bernd Löwe. 2010. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *General hospital psychiatry* 32, 4 (2010), 345–359.
10. Kurt Kroenke, Tara W. Strine, Robert L Spitzer, Janet B. W. Williams, Joyce T. Berry, and Ali H. Mokdad. 2009. The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders* 114, 1 (2009), 163–173.

11. Nicholas D. Lane, Mashfiqui Mohammod, Mu Lin, Xiaochao Yang, Hong Lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. 2011. Bewell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In *PervasiveHealth'11*.

12. Robert LiKamWa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. 2013. Moodscope: Building a mood sensor from smartphone usage patterns. In *MobiSys'13*.

13. Hong Lu, Wei Pan, Nicholas D. Lane, Tanzeem Choudhury, and Andrew T. Campbell. 2009. SoundSense: Scalable Sound Sensing for People-centric Applications on Mobile Phones. In *MobiSys'09*.

14. Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *UbiComp'15*.

15. Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My Phone and Me: Understanding User's Receptivity to Mobile Notifications. In *CHI'16*.

16. Abhinav Mehrotra, Jo Vermeulen, Veljko Pejovic, and Mirco Musolesi. 2015. Ask, But Don't Interrupt: The Case for Interruptibility-Aware Mobile Experience Sampling. In *UbiComp'15 Adjunct*.

17. Abhinav Parate, Matthias Böhmer, David Chu, Deepak Ganesan, and Benjamin M. Marlin. 2013. Practical prediction and prefetch for faster access to applications on mobile phones. In *UbiComp'13*.

18. Veljko Pejovic and Mirco Musolesi. 2014. Anticipatory Mobile Computing for Behaviour Change Interventions. In *UbiComp'14 Adjunct*.

19. Veljko Pejovic and Mirco Musolesi. 2015. Anticipatory Mobile Computing: A Survey of the State of the Art and Research Challenges. *ACM Computing Surveys* 47, 3 (2015), 1–47.

20. Martin Pielot, Rodrigo de Oliveira, Haewoon Kwak, and Nuria Oliver. 2014. Didn't you see my message?: predicting attentiveness to mobile instant messages. In *CHI'14*.

21. Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The Circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 3 (2005), 715–734.

22. Kiran K. Rachuri, Mirco Musolesi, Cecilia Mascolo, Jason Rentfrow, Chris Longworth, and Andrius Aucinas. 2010. EmotionSense: A Mobile Phones based Adaptive Platform for Experimental Social Psychology Research. In *UbiComp'10*.

23. Fani Tsapeli and Mirco Musolesi. 2015. Investigating causality in human behavior from smartphone sensor data: a quasi-experimental approach. *EPJ Data Science* 4, 1 (2015), 1–15.