Designing Effective Movement Digital Biomarkers for Unobtrusive Emotional State Mobile Monitoring

Abhinav Mehrotra
University College London
The Alan Turing Institute
United Kingdom
a.mehrotra@ucl.ac.uk

Mirco Musolesi
University College London
The Alan Turing Institute
United Kingdom
m.musolesi@ucl.ac.uk

ABSTRACT

Mobile sensing technologies and machine learning techniques have been successfully exploited to build effective systems for mental health monitoring and intervention. Various approaches have recently been proposed to effectively exploit contextual information such as mobility, communication and mobile usage patterns for quantifying users' emotional states and wellbeing. In particular, it has been shown that location information collected by means of smartphones can be successfully used to monitor and predict depression levels, as measured by means of standard scores such as PHQ-8.

In this paper, we investigate the design of novel digital biomarkers based on the fine-grained characterization of the mobility patterns of a user, also considering the temporal dimension of their movements (e.g., sequence of places visited by them). We show that the proposed biomarkers have a statistically significant association with emotional states. We also demonstrate that emotional states have a stronger relationship with mobility patterns of weekdays compared to all days of a week. Finally, we discuss the challenges in using these biomarkers in the implementation of "emotion-aware" systems for digital health.

Keywords

Digital Biomarkers; Mobile Sensing; Mobility Analysis; Anticipatory Computing.

1. INTRODUCTION

Mobile phones today have become the most ubiquitous personal computing devices on the planet. These devices have transformed over a period of time from merely communication tools to smart and highly personal devices that are able to assist us in a variety of day-to-day situations. Besides being an indispensable part of our daily life and pervasive in nature, mobile phones come equipped with a plethora of sophisticated sensors with the capability to capture our physical contextual information such as location, movement, audio environment, proximity with other objects, collocation with other devices and many others [11]. Recent studies have shown the potential of exploiting mobile sensing data to learn and, potentially, predict the user's behavioral patterns such as physical activity [3] and mobile phone interaction [10, 8]. Moreover, mobile sensing data have been used to understand and predict mood [13, 7], well-being and mental health conditions such as depression [1, 2, 16, 21, 9].

Whilst sounding simple, understanding users' emotional states with respect to their physical contextual information is a complex task. Likamwa et al. proposed the use of mobile sensing and interaction logs (such as SMS, email, phone call, application usage, web browsing, and location) for predicting users' daily average mood [7]. In [20] Servia et al. present a longitudinal study based on data collected by means of a smartphone application investigating the relation between user's activity and sociability and a variety of psychological dimensions, such as perception of health, life satisfaction, and connectedness.

All of these studies show the potential of using mobile sensor data for inferring emotional states of users in real-time. In particular, in [2] Canzian et al. present an initial investigation of the possibility of using mobility traces for inferring users' depressive states, showing that indeed it is possible to use statistical measures of mobility in order to predict the mental health PHQ-8 depression score [6] of an individual. Starting from this work, we believe that there is a tremendous potential in crafting more refined and fine-grained digital biomarkers based on movement data.

In this paper, we propose a variety of novel digital biomarkers to capture users' daily mobility behavior and demonstrate the association of these biomarkers with their emotional states (including stress, happiness and activeness levels). Our purpose is to examine the predictability of users' emotional states through the analysis of mobility data. In order to test our hypotheses, we developed and deployed an application called MyTraces (Figure 1) that uses an experience sampling method (ESM) approach to collect users' emotional state levels reported by them during different times of the day and continuously logs sensor data. More specifically, the application collects information about three emotional states including activeness, happiness and stress levels on a 5-point Likert scale as well as different aspects of phone interaction and contextual information (such as location and physical activity).

The key contributions of this paper can be summarized as follows:

- we investigate the design of novel digital biomarkers to characterize user's mobility behavior using GPS traces collected by means of smartphones;
- we demonstrate that these novel digital biomarkers are associated with emotional states of the individuals, showing their potential exploitation for predicting their emotional states more accurately;

DigitalBioMarkers'17, June 23, 2017, Niagara Falls, NY, USA

DOI: http://dx.doi.org/10.1145/3089341.3089342

- finally, we also present a temporal analysis of the mobility patterns of users and their variations over different days of the week and their impact on the use of the proposed biomarkers.

In general, we believe that the results presented in this work represent an important step forward in understanding how sensor data, and in particular mobility information, can effectively be used for modelling and inferring human behavior and emotional states in a passive way and at a scale.

2. DATA COLLECTION

2.1 Overview

In order to investigate the relationship between emotional states and the mobility patterns of users, we conducted an *in-the-wild* study. More specifically, we developed an Android app called MyTraces (shown in Figure 1) that runs in the background to unobtrusively and continuously collect users' mobile phone interaction logs and context information (such as location and physical activity). Note that in this paper we analyze only mobility data and the analysis of other collected data features is not presented in this paper. Moreover, the application samples GPS data in an adaptive sensing fashion using the mechanisms described in [2].

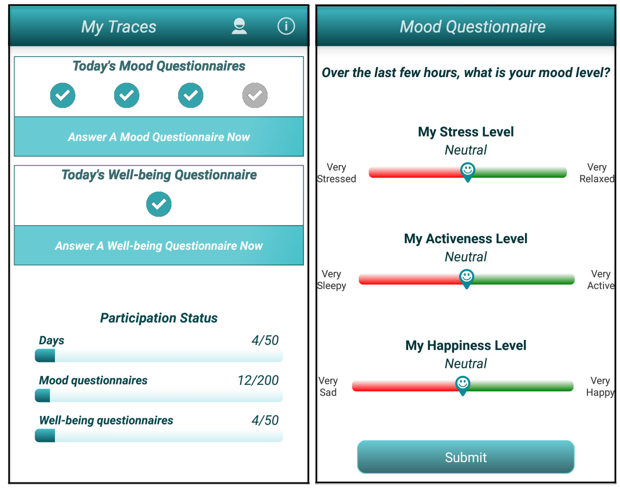
To acquire information about users' emotional states (activeness, happiness and stress level) throughout the day, we rely on the experience sampling method (ESM). As shown in Figure 1.b users can register their emotional states through a sliding bar. This bar uses a 5 point-based Likert scale where 1 indicates the lowest level and 5 the highest level. Every day a questionnaire is triggered once at a randomly selected time inside a 3-hour time window. We used 4 3-hour time windows between 8.00 am and 11.00 pm (in the local time zone of the user). We chose this time window so that the participants do not feel annoyed by being asked to respond to the surveys early in the morning and late at night. In case a questionnaire is dismissed or not responded to within 30 minutes from its arrival time, the application triggers another alert after 30 minutes.

Since the higher values of activeness and happiness levels indicate a positive emotion, we measured the stress level according to a negative scale that means lower value would indicate high level of stress. By doing that, we make the scale of all three indicators consistent (i.e., the lower values refer to negative emotion and higher values to the positive emotion). Therefore, we reverse the scale by subtracting each response value from 6. So, if for example the response is 5 (i.e., very low stress), we subtract it from 6 to rescale it to 1. Thus, with the reversed scale the lower value will refer to lower stress and the higher value would indicate higher stress.

2.2 Recruitment of the Participants and Ethics

The MyTraces application was published on Google Play Store and has been available to the general public for free since 4th January 2016. It was advertised through different channels including academic mailing lists, Twitter, Facebook and Reddit. In order to attract more participants for our study, we committed to give incentives to the participants for replying to the questionnaires for a minimum of 30 days. We committed to select (through a lottery) one winner of a Moto 360 Smartwatch and 20 winners of an Amazon voucher.

In order to ensure privacy compliance, the MyTraces application shows a list of information that is collected and asks for the user's consent. Additionally, the study was performed in accordance with our institution's ethical research procedure and the consent form itself for the data collection was reviewed by our institution's Ethics Review Board.



(a) (b)

Figure 1: MyTraces application: (a) main screen, (b) mood questionnaire.

2.3 Dataset

In this study we consider the data collected from 4th January 2016 to 1st July 2016. In this period the application was installed by 104 users. However, many users did not actively respond to the mood questionnaires and some uninstalled the application after a few days. Therefore, we selected a subset of the data for the analysis by considering only the users who ran the application and responded to questionnaires for at least 21 days in order to have a sufficiently large sample that can be statistically significant. Consequently, we obtained a set of 22 users who satisfied these constraints. Note that we do not have information about the demographics of these participants because this information was not collected during the study for privacy reasons.

2.4 Emotional States

Most of the previous studies [13, 7] have considered the Circumplex mood model with two dimensions as valence and arousal [14]. However, as discussed in [17], Schimmack and Rainer proposed that the arousal state can be split into two dimensions: tense arousal and energetic arousal. The authors justified this split with the fact that the energetic arousal is influenced by a circadian rhythm (i.e., it corresponds to activity in brain cells that regulate organisms' sleep-wake cycle), whereas tense arousal does not show a similar circadian rhythm. Therefore, in our study we split "arousal" into tense arousal (stressed-relaxed) and energetic arousal (sleepy-active).

Consequently, we consider three aspects of emotional states that are captured during the day:

- **activeness**: a state of being aroused and physiological readiness to respond [12];
- **happiness**: a state of positiveness and joy that is derived from external and momentary pleasures [18];
- **stress**: a negative state of being under high mental pressure [19].

The levels of these emotional states are computed on a 5 point-based Likert scale, where 1 indicates the lowest level and 5 the highest level.

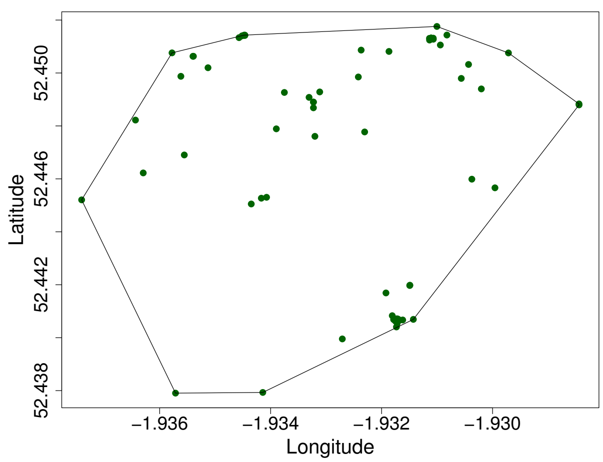


Figure 2: Visualization of *spatial coverage by convex hull* metric with a user's mobility data for a day.

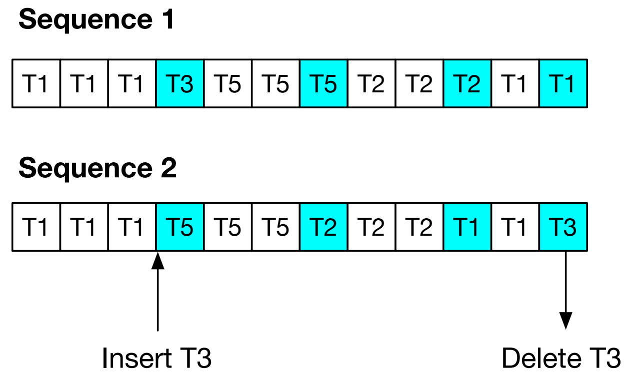


Figure 3: An example to demonstrate the process of computing the difference between two tile sequences. Here, we use tile sequences of length 12 meters due to space limitations. In this case, the sequence difference will be 2 as we need to (i) insert T3 at fourth position; and (ii) delete T3 from last position.

3. MOBILITY-BASED BIOMARKERS

We now introduce five mobility-based biomarkers (i.e., mobility metrics) capturing some key characteristics of users' mobility patterns. In order to compute these metrics we rely on the location data collected passively through the MyTraces app (discussed in the previous section). These digital biomarkers represent the basis of our correlation analysis that we will present in the following section. All metrics are computed for each user on a daily basis.

3.1 Spatial Coverage by Tiles Approximation

This metric indicates the places (i.e., tiles) visited by a user. In order to compute this we first group location points into square tiles of each side as t_{length}. We then provide unique identifier to each tile and map all location points to their matching tile identifiers. Finally, for each user we compute the number of unique tiles visited in a day by finding the count of unique tile identifiers for that day. It is worth noting that we optimize the value of $t_{length} \in [50, 100]$ with a step of 50 (meters). This optimization is performed based on the correlations with each emotional states. It is worth noting that in this paper we considered squared areas, however, different shapes might be considered. We plan to investigate different types of shapes as planned future work.

3.2 Spatial Coverage by Convex Hull Approximation

This metric represents the spatial area covered by the user. We use the convex hull algorithm (devised by Ronald Graham [5]) to find the smallest Euclidean space that contains the given set of location points. In Figure 2 we show the visualization of the *spatial coverage by convex hull* with a user's mobility data for a day.

3.3 Tiles Sequence

This metric indicates the similarity between the sequence of tiles visited in the current and previous day. In order to compute the similarity between these sequences we use the string-edit distance that gives us the minimum number of insertions, deletions, and substitutions required to transform one string into the other [15]. In our case tile identifiers are used as characters to form strings representing the sequences of tiles visited in two days. In Figure 3 we show an example of computing the difference between two tile sequences. It is worth noting that we optimize the value of tile sides (i.e., t_{length}) as $t_{length} \in [50, 100]$ with a step of 50 (meters) and the optimization is performed based on the correlations with each emotional states. In order to compute a sequence of tiles we need to transform the location data into continuous time series of locations at a consistent interval of 10 minutes. In theory it is possible to optimize both tile size and sampling rate. We plan to explore the selection of the sampling interval as future work.

3.4 Place Sequence

This metric indicates the similarity between the sequence of places visited in the current and previous day. This metric is computed in the same way as tile sequence metric. However, the tile sequence is replaced with the sequence of place identifiers. Here, place identifiers indicates unique significant places (i.e, clusters of locations visited by a user). To compute these significant places we firstly discard the location samples with more than 50 meters accuracy so that the estimated location clusters are of better quality. We then find the location samples that were collected while users were moving and we also discard them. So we use only location points where users stopped. In order to infer such location points, we compute the speed of the user by using the distance and the time between the last and the current location points. If the speed is less than a certain threshold (i.e., 5 km per hour) we consider that location reading was collected when the user was not moving (i.e., stopped). Now, we iterate over all remaining location samples and for each location point we create a new cluster only if the distance of this location from the centroid of each existing cluster is more than 50 meters. Otherwise, we add this location to the corresponding cluster and update its centroid. Finally, we consider all centroids as significant places.

3.5 Displacement Entropy

We construct this metric to capture the level of predictability in the daily movement patterns of a user. In order to compute this metric we first create a time series of displacements at a time interval of t_{window} (i.e., distance travelled by the user in continuous time windows of t_{window} during the day). This sequence of displacement values for a day is used to compute the Shannon entropy. It is worth noting that we find the optimal value of $t_{window} \in [10, 90]$ with a step of 10 (minutes) that results in the best correlation of the *displacement entropy* with each emotional states.

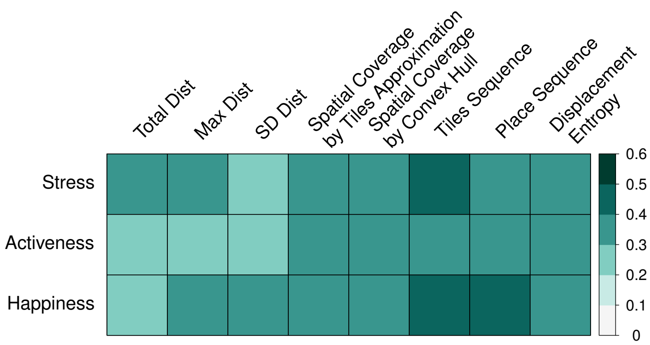


Figure 4: Association between digital biomarkers and average daily emotional states by considering all days of a week.

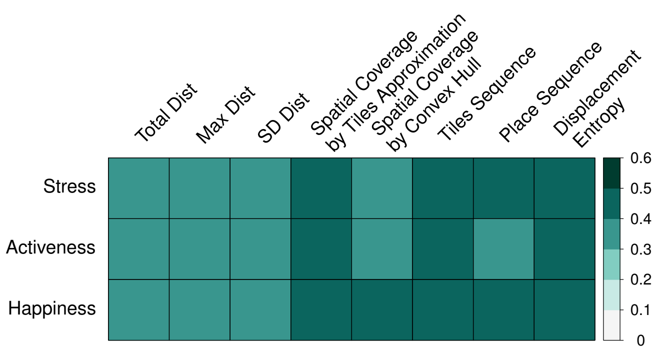


Figure 5: Association between digital biomarkers and average daily emotional states by considering only weekdays.

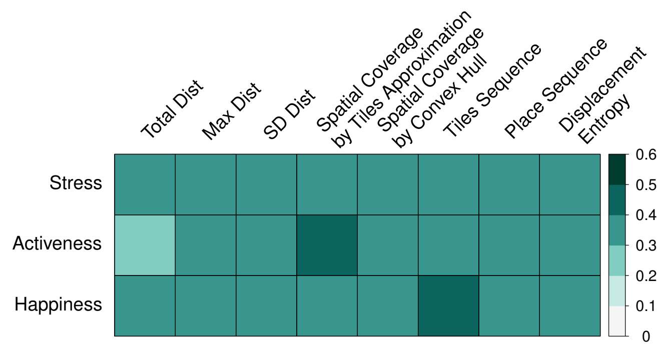


Figure 6: Association between digital biomarkers and strongest (i.e., highest or lowest) emotional states of a day by considering all days of a week.

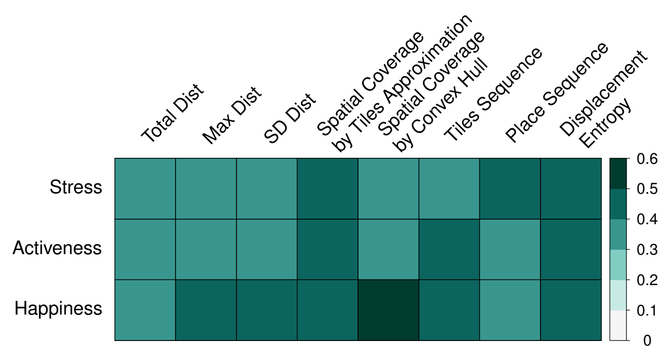


Figure 7: Association between digital biomarkers and strongest (i.e., highest or lowest) emotional states of a day by considering only weekdays.

4. METHODOLOGY

In this section we present our approach for analyzing the association between users' emotional states and their mobility metrics (presented in Section 3).

4.1 Computing Daily Average and Strongest Emotional States

Since previous studies have shown that the average daily emotional state scores do not has much variance, which indicates that a strong deviation of motion at one instance of a day could not be captured through the average score. For example, let us consider the happiness scores of a day are reported as 1, 3, 3 and 4. Here, 1 indicates that the user was very sad, but this information is unrevealed if we take the average that will be 2.75 indicating the user's happiness score is close to neutral for the day. Furthermore, we believe that the strong emotional state at a some moment of a day might have a big impact on the mobility. Therefore, we compute the emotional states in two ways (i.e., by taking the average and the strongest emotional state of a day) in order to investigate the link between emotional states and mobility patterns.

In order to compute the daily average emotional states, we take the mean of all emotional states reported in a day. On the other hand, we compute the strongest emotional state of a day by first transforming the emotional states from the scale of 1 to 5 into -2 to 2 scale. We now take all emotional states reported in a day to find which one has the highest absolute value and the corresponding emotional state value is used as the strongest emotional state of the day. For instance, if four reports of happiness are 1, 3, 3 and 4, they are converted to -2, -1,-1 and 1, and we use -2 as the strongest emotional state of the day. Finally, we merge all values in the range of -1 to 1 in order to rescale the strongest emotional state of the day into -1 to 1 scale.

4.2 Analysis for Weekdays and Entire Week

Since our routine for weekdays is different from weekends, we believe that our mobility follow a specific routine during the weekdays but it follows an uncertain routine during weekends. This could potentially be explained by the fact that we work during specific hours on weekdays but weekends plans could vary. Due to this reason, divergence in emotional states could reflect as the divergence in mobility patterns during weekdays. However, since people do not follow a specific routine, the divergence in their mobility during weekends could be associated with other variables rather than emotional states. Due to the limited size of the data, we have data on only 6 weekend days. So we could not directly compare differences by analyzing weekdays and weekends data separately. Instead, in order to understand this difference between, we can compare the analysis by considering the data of weekdays and all days of a week (i.e., weekdays and weekends included).

5. RESULTS

In this section we present the results of our analysis regarding the potential association between users' emotional states and the mobility-based digital biomarkers discussed in the previous section. In order to quantify this relationship, we compute the Spearman's rank-order correlation coefficients. The analysis is performed for each user separately. We consider the absolute values of these coefficients because we are interested in the strength of the relationships between the variables. We then compute the average of these coefficient values. We rely on Fisher's method [4] for combining the p-values of individual-based correlation analysis.

As discussed in Section 4, we compute emotional states in two ways: (i) average emotional state of a day; and (ii) strongest emotional state of a day. The results of the association between the proposed digital biomarkers and the emotional states computed in both ways are presented below.

5.1 Association with Daily Average Emotional States

We first present the results for the association between the digital biomarkers and the average emotional states. In order to compare the difference in the biomarkers effectiveness during different days of the week, we perform this analysis in two ways: (i) by using the data of all days of a week; and (ii) by using the data of only weekdays. The results of these analyses are presented in Figure 4 and Figure 5.

The results demonstrate that all biomarkers have a statistically significant correlation ($\rho \in [0.3, 0.5]$ with $p < 0.05$) with average emotional states. However, our biomarkers outperform all basic mobility features that are used in previous studies (for example, in [2]). At the same time, some of our biomarkers such as spatial coverage by tile approximation, tile sequence, and displacement entropy always show a very strong association ($\rho \in [0.4, 0.5]$), which is consistently stronger than previously used feature that obtain the $\rho \in [0.2, 0.4]$.

Moreover, the comparison between the results obtained considering the data of all days of a week and only weekdays indicates that biomarkers built on weekdays data shows a stronger correlation. Since people might not have a specific routine on weekends, biomarkers calculated using weekends data are more noisy and, consequently, the correlation is lower.

5.2 Association with Daily Strongest Emotional States

We first present the results concerning the association between the digital biomarkers and the strongest emotional states of each day. As discussed earlier, we perform this analysis in two ways: (i) by using the data of all days of a week; and (ii) by using the data of only weekdays, in order to compare the difference in the biomarkers' effectiveness during different days of a week. The results of these analyses are presented in Figure 6 and Figure 7.

The results demonstrate that all biomarkers have more marked association with strongest emotional states compared to average ones. At the same time, we observe a statistically significant correlation ($\rho \in [0.3, 0.6]$ with $p < 0.05$) of emotional states with our biomarkers, which is much greater than the results of all basic mobility features (i.e., $\rho \in [0.2, 0.5]$). Moreover, we note that the *spatial coverage by convex hull* biomarker shows a very strong association ($\rho = 0.52$) with *happiness*.

On the other hand, the comparison between the results with the data of all days of a week and only weekdays indicates that biomarkers built on weekdays data shows a stronger correlation, as also observed for the case of the average emotional states discussed above.

6. LIMITATIONS

Our results show the effectiveness of the proposed biomarkers for capturing users' mobility behavior and to infer their emotional states. In particular, we have showed that more refined mobility-based biomarkers, which exploit both the temporal and the spatial dimensions, outperform those presented in [2].

The present work has also some limitations. We believe that there are some limitations in our study that must be overcome to deploy our biomarkers-based system for inferring *in-the-wild* emotional states. First, there is a need of developing biomarkers that take into consideration the variability of behavior over different days of the week. One possibility is to train biomarkers for specific days, but this will require a longer training period and this might not be feasible in practice. Second, this study was performed with a small number of participants and over a short duration of time. It is important to test these results with a larger number of participants with different demographics and over a longer period of time, also possibly in various periods of the year. Finally, almost all previous studies are based on a first phase of data collection and then on a second phase focused on the offline analysis. We believe that in order to further validate the robustness and ecological validity of the biomarkers or models based on these biomarkers, there is a need to perform *in-the-wild* evaluations of these techniques.

7. CONCLUSIONS

In this paper, we have proposed a series of novel fine-grained mobility-based spatio-temporal biomarkers that can be used to effectively capture users' daily mobility behavior. We have evaluated the proposed biomarkers and we have showed they have a statistically significant correlation with users' emotional states, collected by means of a ESM mobile application and provide better performance in terms of strength of the association with respect to the state of the art. Additionally, we have noted that biomarkers built on weekdays data shows a stronger correlation with the emotional states taken into consideration compared to biomarkers built on the data of all days of a week. Finally, we have also demonstrated that daily strongest emotional states are strongly correlated with digital biomarkers compared to daily average emotional states. Our future research agenda includes the design of novel biomarkers based on other contextual modalities and the use these biomarkers for predicting individuals' cognitive context.

8. ACKNOWLEDGEMENTS

This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1 and at UCL through the EPSRC grants EP/L018829/2 and EP/P016278/1.

9. REFERENCES

[1] J. Alvarez-Lozano, V. Osmani, O. Mayora, M. Frost, J. Bardram, M. Faurholt-Jepsen, and L. V. Kessing. Tell me your apps and i will tell you your mood: correlation of apps usage with bipolar disorder state. In *PETRA'14*, May 2014.

[2] L. Canzian and M. Musolesi. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *UbiComp'15*, September 2015.

[3] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Borriello, B. Hemingway, P. Klasnja, K. Koscher, J. A. Landay, J. Lester, D. Wyatt, and D. Haehnel. The Mobile Sensing Platform: An Embedded Activity Recognition System. *IEEE Pervasive Computing*, 7:32–41, 2008.

[4] R. A. Fisher and F. Yates. *Statistical tables for biological, agricultural and medical research*. Longman, 1938.

[5] R. L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Information processing letters*, 1(4):132–133, 1972.

[6] K. Kroenke, T. W. Strine, R. L. Spitzer, J. B. W. Williams, J. T. Berry, and A. H. Mokdad. The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114(1):163–173, 2009.

[7] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong. Moodscope: Building a mood sensor from smartphone usage patterns. In *MobiSys'13*, June 2013.

[8] A. Mehrotra, R. Hendley, and M. Musolesi. PrefMiner: Mining User's Preferences for Intelligent Mobile Notification Management. In *Proceedings of UbiComp'16*, Heidelberg, Germany, September 2016.

[9] A. Mehrotra, R. Hendley, and M. Musolesi. Towards Multi-modal Anticipatory Monitoring of Depressive States through the Analysis of Human-Smartphone. In *Adjunct UbiComp'16*, Heidelberg, Germany, September 2016.

[10] A. Mehrotra, M. Musolesi, R. Hendley, and V. Pejovic. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *UbiComp'15*, September 2015.

[11] A. Mehrotra, V. Pejovic, and M. Musolesi. SenSocial: A Middleware for Integrating Online Social Networks and Mobile Sensing Data Streams. In *Middleware'14*, December 2014.

[12] K. H. Pribram and D. McGuinness. Arousal, activation, and effort in the control of attention. *Psychological Review*, 82(2):116–149, 1975.

[13] K. K. Rachuri, M. Musolesi, C. Mascolo, J. Rentfrow, C. Longworth, and A. Aucinas. EmotionSense: A Mobile Phones based Adaptive Platform for Experimental Social Psychology Research. In *UbiComp'10*, Copenhagen, Denmark, September 2010.

[14] J. Ressel. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–78, 1980.

[15] E. S. Ristad and P. N. Yianilos. Learning string-edit distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5):522–532, 1998.

[16] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, and D. C. Mohr. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *Journal of Medical Internet research*, 17(7):e175, 2015.

[17] U. Schimmack and R. Rainer. Experiencing activation: energetic arousal and tense arousal are not mixtures of valence and activation. *Emotion*, 2(4):412, 2002.

[18] M. E. Seligman. Can happiness be taught? *Daedalus*, 133(2):80–87, 2004.

[19] H. Selye. *The stress of life*. McGraw-Hill, 1956.

[20] S. Servia-Rodríguez, K. K. Rachuri, C. Mascolo, P. J. Rentfrow, N. Lathia, and G. M. Sandstrom. Mobile sensing at the service of mental well-being: a large-scale longitudinal study. In *WWW'17*, April 2017.

[21] Y. Suhara, Y. Xu, and A. Pentland. Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks. In *WWW'17*, April 2017.