Smartphone app usage as a predictor of perceived stress levels at workplace

Raihana Ferdous, Venet Osmani and Oscar Mayora

CREATE-NET, Trento, Italy

{raihana.ferdous, venet.osmani, oscar.mayora}@create-net.org

Abstract—Explosion of number of smartphone apps and their diversity has created a fertile ground to study behaviour of smartphone users. Patterns of app usage, specifically types of apps and their duration are influenced by the state of the user and this information can be correlated with the self-reported state of the users. The work in this paper is along the line of understanding patterns of app usage and investigating relationship of these patterns with the perceived stress level within the workplace context. Our results show that using a subject-centric behaviour model we can predict stress levels based on smartphone app usage. The results we have achieved, of average accuracy of 75% and precision of 85.7%, can be used as an indicator of overall stress levels in work environments and in turn inform stress-reduction organisational policies, especially when considering interrelation between stress and productivity of workers.

keywords: occupational stress, monitoring, smartphones

I. Introduction

Studies have established that moderate levels of stress can be helpful while, excessive, chronic, and repeated exposure to stress has been associated with chronic diseases, depression, and immune disorders [1], [2]. Prolonged exposure to stress at work has a significant negative effect on work productivity [3], eventually leading to burnout [4], which reduces the motivation and effectiveness of workers. Considering the intertwined relationship between stress, burnout and productivity; monitoring stress at work and understanding related factors is becoming more and more important. Even though a number of research studies [5], [6] have investigated causes and consequences of chronic stress at work through self-reported data, there is far less work focusing on automatic detection of objective behaviour changes. In this regard, mobile technologies can play a significant role in monitoring behaviour of subjects [7]–[9], helped by familiarity of users with these devices in one hand and their multi-modal sensing capabilities on the other.

In this paper we investigate whether it is possible to predict the stress level of the participants considering their app usage patterns. Although there have been studies investigating the usage pattern of mobile apps to predict the next app usage of subjects [10]–[14], there has been no study focused on investigating the relationship between the pattern of mobile apps usage of the users at work and their perceived stress levels.

The results presented in this paper are based on a study of 28 workers, monitored over 6 weeks through their smartphones during their daily, real-world behaviour. Each participant was provided with a smartphone that had our data collection app installed and there were no restrictions placed upon the subjects as to particular usage or position of the phone. The purpose of the app was two fold: i) recording app usage; which is done through a continuous process of storing the active apps, their starting and ending timestamps, and their duration. ii) capturing perceived stress levels; which was done through implementation of a questionnaire on the smartphone, where subjects were prompted three times a day to answer a question in order to record their perceived stress level. Initial analysis have revealed that the most active apps, for majority of participants, are email, calendar and browsers; while, the least used apps are games apps, as one would expect in work environments. In total there were 128 unique apps recorded with an average of around 12 apps per participant and the standard deviation is 6.45. While some common characteristics for majority of subjects exist however, the app usage behaviour varies across subjects and it is this variance that we have captured and investigated its relationship with perceived stress levels. In addition, individual differences between subjects have prompted the necessity of devising individual behaviour model for each subject, rather than establishment of a group behaviour model. It should be noted that in our tests, group behaviour model resulted in a low average accuracy of 54%, reported in Section IV-B. As such, in order to efficiently predict the perceived stress level of the participants, we have focused on developing subject-centric app usage models to characterise the app usage behaviour of each of the participant. A classifier is trained on individual subject data and prediction performance is measured for each subject. The classification results are presented in Section IV-C and they show that on average, the subjects' stress level can be predicted with 75% accuracy. The results of the study could be used as an indicator of the overall level of stress in work environments and in turn inform stress-reduction policies, especially when considering interrelation between stress and productivity.

The rest of this paper is structured as follows: Section II covers a discussion on related work of this field, while description of the dataset, feature selection and model development is outlined in Section III. The experimental results are shown in Section IV and Section V summarises the results.

II. Related Work

A number of authors [15]–[17] have investigated correlation between human behaviour and perceived stress at work by analysing the data derived from the smartphones data or sensors data. One of the conclusions is that there are behaviour differences before and during experience of stress, however these differences are distinct for each individual. Several studies [18]–[20] investigating the pattern of smartphone apps usage have noticed immense diversity

among smartphone users in different aspects, such as, per day user interaction, application use, network traffic and energy drain. The authors of [19] have suggested to learn and adapt user specific interaction pattern with smartphones in order to improve user experience and to predict the energy consumption more accurately. A number of research papers [10]–[14], [21]–[23] are found performing large scale analysis of the context related to mobile app use. Here, the authors have noticed strong dependencies between app usage and the contextual variables, such as, location, time and social aspects. They have shown that the rich contextual information related to the smartphone apps usage has the power to predict the next app that will be used by the user. Authors of [24] have shown that aggregated features obtained from smartphone usage data (e.g., app usage, phone call records, location information, etc) are indicators of the Big-Five personality traits (i.e., extraversion, agreeableness, conscientiousness, emotional stability and openness to experience). Analysis of application usage patterns as a mean to detect malware in smartphones is investigated in [25] and [26]. The authors show that identification of the behavioural differences between malware and human users is possible by developing a model which is trained with the dynamic app usage pattern from two major aspects: process state transitions and user operational patterns. Authors of [27] have reported that the changes in general behaviour of patients due to onset of a bipolar episode can be captured through the differences in their smartphone usage pattern. A study investigating the stress and strain of student life [28] reveals that students start the term with high positive affect and low stress, however, as the term progresses and the workload increases, stress appreciably rises.

III. METHODOLOGY

A. Collection of Data

Each of the 28 participants was provided with a smartphone, having our data collection application installed and were monitored over a period of 6 weeks, from November 2013 to December 2013. As our focus is to investigate the participants' behaviour during work, we limit the collection of app usage records only for the duration they spend at work. The timestamped records of start and end times of each active applications used by the participants are captured by our data collection app. In order to identify which applications are used by the participant, we considered those applications that are running within the time period when the screen status is on. In order to understand the subjective stress levels of the participants, a question ("what is your stress level?") was answered on a 5-point scale, where 1 indicates "very slightly or not at all" and 5 indicates "extremely". This questionnaire was implemented on the smartphones of the participants and was asked 3 times during the day, in the morning at the beginning of the work, around noon, and before leaving workplace.

B. Feature Extraction & Prediction of Stress

1) Categorisation of Apps: In order to get a more high level understanding of the applications used by the participants during the observation window, we have divided apps into 5 categories: (i) entertainment, (ii) social networking, (iii) utility, (iv) browser, and (v) game apps. Table I presents a number of examples of apps for each category.

TABLE I.   CATEGORISATION OF APPS USED BY THE PARTICIPANTS

Social networking service apps	Entertainment applications
Facebook, Twitter, google+, VoIP apps (e.g., LocalPhone, MobileVoIP, etc.), Dropbox, Pinterest, Flipboard, Whatsapp, gtalk, viber, skype	Audio & Video streaming apps, Online book reading and purchase, FM Radio, tv guide, Music instrument tuner & composition, Youtube, video player, tv streaming, News listening and reading apps
Utility applications	**Browser applications**
Calendar, map, navigator, clock, Weather app, note taking app, Voice to text, text to voice app, Trip advisor, QR Code Reader, Calculator, accounting app, Scanner, flashlight, voice recorder, Recipe apps, camera, Online shopping	Chrome Firefox Email google quick searchbox
Game applications	
Card games, puzzles, fun games (e.g., fruitninja, candycrushsaga, angry birds)	

TABLE II.   LIST OF FEATURES EXTRACTED FROM THE APP USAGE RECORDS OF A USER

Feature	Description
Frequency ent	Frequency of using apps belong to *entertainment* category per day
Time ent	Time (in second) of using apps belong to *entertainment* category per day
Frequency social	Frequency of using apps belong to *social networking service* apps category per day
Time social	Time (in second) of using apps belong to *social networking service* apps category per day
Frequency game	Frequency of using *game* apps per day
Time game	Time (in second) of using *game* apps per day
Frequency utility	Frequency of using apps belong to *utility* category per day
Time utility	Time (in second) of using apps belong to *utility* category per day
Frequency browser	Frequency of using apps belong to *browser* category per day
Time browser	Time (in second) of using apps belong to *browser* category per day
Number of unique app	Number of unique apps used per day

2) Feature Selection: In order to characterize the subject-centric app usage patterns, we have extracted a set of suitable features from the collected app usage records of the participants. Smartphone app usage behaviour is driven by factors such as popularity of apps (e.g., how many categories of apps a user visited and how is his/her attention spread), frequency and duration of app usage (e.g., how often does a user uses any application and how long is the app used), frequency of unique apps used (e.g., if the user keeps using only a few apps regularly or the list of his/her visited apps is large). Considering above criteria, a set of features (shown in Table II) is extracted from the app usage records of each participants that captures their usage pattern per day.

3) Prediction of stress level of a user: User-centric model of app usage is developed by training a SVM classifier with individual app usage behaviour (features extracted from the app usage record) and participants' perceived stress levels as ground truth. The model thus trained with the user-specific app usage information is successfully able to predict the stress level of that user during the testing period. Before training each classifier, a suitable model (i.e., kernel function and the SVM parameters) is selected. Selection of kernel and tuning of the parameters of SVM are performed on the training set through the widely used *cross-validation technique*. This test was performed for three kernel functions: linear, RBF, and polynomial – and for various values of their relevant parameters (for example, SVM parameter *soft margin constant* C to controls the trade-off between maximizing the margin and

minimizing the training error, RBF kernel parameter γ, and, the degree d for polynomial kernel). Cross-validation method is applied for each set of parameters. Finally, the parameters with the best cross-validation index were selected.

IV. ANALYSIS AND EXPERIMENTAL RESULTS

A. General pattern of smartphone app usage

A number of common characteristics are observed across all the participants regarding their app usage history. Figure 1 presents the relationship between average usage time and frequency for applications of each category. It can be seen that apps belonging to the *browser* category have high usage frequencies but lower durations per use across all participants. Apps belong to *utility* category are also used frequently and for a longer duration. This is something that is expected in a working environment where use of apps such as email, calendar, browser, is common. Interestingly, in case of games apps, the average duration is high, however they are used few times only. This may be due to the nature of games, where once the users begin a game they typically stay engaged for longer duration, in comparison to say calendar, browser or email apps. In the context of perceived stress at work, the relationship
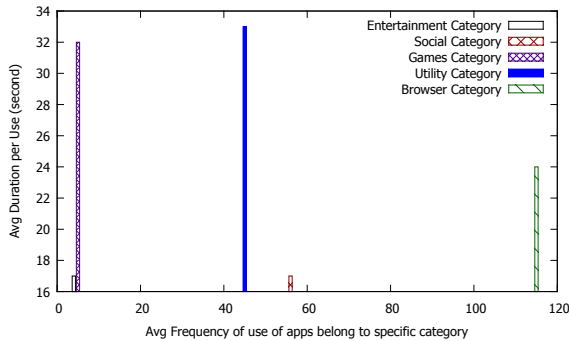

Fig. 1.   Average number of times apps of specific categories are used vs duration per use

between the app usage patterns and perceived stress also varies for different participants. For instance, Figure 2 shows the diversity of using apps belonging to *browser* category and stress level of two randomly chosen participants.

B. Generic model for stress level prediction

Considering the common characteristics of app usage observed for majority of the participants, first we have focused on developing a generic group behaviour model to predict the stress level. To this end, a combined dataset is prepared by integrating the features extracted from the app usage records for all the participants. The model is developed by training a SVM classifier with a subset of this dataset defined as *training* data, while the remaining subset is used for testing. Here, k-fold (k=10) is used, which indicates that each subset is tested (validated) using the model trained on the remaining $(k-1)$ subsets with the right choice of kernel function and SVM parameters. The k results are averaged to obtain a single index. In this context, the average accuracy obtained from the generic classifier is 54%. This result indicates the limitations of using group behaviour model in predicting stress level of the participants considering their app usage history.
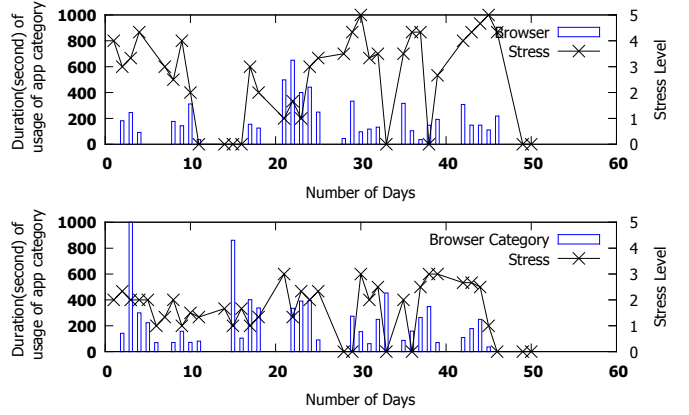

Fig. 2.   Usage pattern of apps belonging to *Browser* category and stress level (1-5) per day for two randomly chosen participants [stress level, 0 = no data available]

C. User-centric model for stress level prediction

Considering the diversity of individual in the context of app usage, we have focused on developing user-centric model to find the correlation between app usage and perceived stress level. Here, each model is trained with the individual app usage features and his/her perceived stress level. To this end, the app usage features of each user is divided into two sets: (i) training set consists of 70% data, and (ii) test set containing the remaining 30% data. User specific models are selected by using k-fold cross-validation technique on the training sets. The validation of models is performed on the test sets. Table III shows the stress prediction accuracy obtained through the 10-fold cross-validation during model selection process from 22 models that correspond to 22 active participants (out of 28 participants that took part in the study). While the highest accuracy is 99% and lowest accuracy is 69.2%, the table shows that an average accuracy is 88.1%. The third column of the Table III shows the stress prediction accuracy obtained from the user specific models during the validation process. While the highest accuracy is 95% and lowest accuracy is 50%, the table shows that an average accuracy is 75%.

V. CONCLUSION

The results presented in this paper have shown that patterns of smartphone app usage are highly correlated with the self-reported stress level of the users and consequently can be used to predict stress levels within the workplace context. While there is some variation across participants, on average the results we have achieved can be used as solid indicators of individual and organisation-wide stress level, without requiring any intervention from the users. Furthermore, monitoring and understanding stress level has wide-reaching implications in informing organisation of workplaces.

REFERENCES

[1] A. Rosengren, S. Hawken, S. Ôunpuu, K. Sliwa, M. Zubaid, W. A. Almahmeed, K. N. Blackett, C. Sitthi-amorn, H. Sato, and S. Yusuf, "Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study," *The Lancet*, vol. 364, no. 9438, pp. 953–962, 2004.

User ID	Model Selection Accuracy (Cross-Validation)	Accuracy (Validation)	Precision	Recall
84616	74.194%	60%	72.727%	88.889%
94433	81.481%	66.67%	85.714%	60.000%
88187	77.273%	57.14%	72.727%	80.000%
89532	86.957%	57%	88.889%	80.000%
94441	96.552%	80%	100.000%	92.308%
95646	91.304%	80%	84.615%	85.714%
14446	83.871%	60%	84.615%	78.571%
96479	69.231%	50%	58.333%	70.000%
94516	75.000%	58%	69.565%	100.000%
94813	99.000%	90%	98.000%	96.000%
94615	93.750%	80%	50.000%	100.000%
95414	94.444%	91%	100.000%	50.000%
95596	95.000%	88%	91.435%	93.000%
95448	75.758%	70%	53.333%	88.889%
87676	88.000%	87.50%	88.374%	74.000%
93401	96.154%	75%	100.000%	83.333%
95216	94.737%	83.33%	100.000%	50.000%
87684	98.000%	95%	96.000%	92.000%
94714	86.667%	55%	100.000%	60.000%
89953	85.714%	71.42%	100.000%	63.333%
96040	98.000%	95%	95.000%	92.545%
95521	98.000%	93%	96.000%	97.594%
Average	**88.140%**	**75%**	**85.697%**	**80.735%**

[2] M. G. Marmot, H. Bosma, H. Hemingway, E. Brunner, and S. Stansfeld, "Contribution of job control and other risk factors to social variations in coronary heart disease incidence," *The Lancet*, vol. 350, no. 9073, pp. 235–239, 1997.

[3] P. K. Wilke, W. H. Gmelch, and N. P. Lovrich Jr, "Stress and productivity: Evidence of the inverted U function," *Public Productivity Review*, pp. 342–356, 1985.

[4] C. Maslach, W. B. Schaufeli, and M. P. Leiter, "Job burnout," *Annual review of psychology*, vol. 52, no. 1, pp. 397–422, 2001.

[5] C. Kyriacou, "Teacher stress and burnout: An international review," *Educational research*, vol. 29, no. 2, pp. 146–152, 1987.

[6] S. B. Bacharach, P. Bamberger, and S. Conley, "Work-home conflict among nurses and engineers: Mediating the impact of role stress on burnout and satisfaction at work," *Journal of organizational Behavior*, vol. 12, no. 1, pp. 39–53, 1991.

[7] A. Grunerbl, A. Muaremi, V. Osmani, G. Bahle, S. Ohler, G. Troster, O. Mayora, C. Haring, and P. Lukowicz, "Smartphone-based recognition of states and state changes in bipolar disorder patients," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, no. 1, pp. 140–148, Jan 2015.

[8] A. Matic, V. Osmani, and O. Mayora-Ibarra, "Mobile monitoring of formal and informal social interactions at workplace," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 2014, pp. 1035–1044.

[9] A. Matic, V. Osmani, A. Maxhuni, and O. Mayora, "Multi-Modal Mobile Sensing of Social Interactions," *IEEE Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare*, pp. 105–114, 2012.

[10] T. M. T. Do, J. Blom, and D. Gatica-Perez, "Smartphone usage in the wild: a large-scale analysis of applications and context," in *Proceedings of the 13th international conference on multimodal interfaces*. ACM, 2011, pp. 353–360.

[11] Y. Xu, M. Lin, H. Lu, G. Cardone, N. Lane, Z. Chen, A. Campbell, and T. Choudhury, "Preference, context and communities: A multi-faceted approach to predicting smartphone app usage patterns," in *Proceedings of the 17th annual international symposium on International symposium on wearable computers*. ACM, 2013, pp. 69–76.

[12] Z.-X. Liao, P.-R. Lei, T.-J. Shen, S.-C. Li, and W.-C. Peng, "Mining temporal profiles of mobile applications for usage prediction," in *Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 890–893.

[13] Z.-X. Liao, S.-C. Li, W.-C. Peng, P. S. Yu, and T.-C. Liu, "On the Feature Discovery for app usage prediction in smartphones," in *Data Mining (ICDM), 2013 IEEE 13th International Conference on*. IEEE, 2013, pp. 1127–1132.

[14] C. Shin, J.-H. Hong, and A. K. Dey, "Understanding and prediction of mobile application usage for smart phones," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 173–182.

[15] G. Bauer and P. Lukowicz, "Can smartphones detect stress-related changes in the behaviour of individuals?" in *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE, 2012, pp. 423–426.

[16] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland, "Pervasive stress recognition for sustainable living," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*. IEEE, 2014, pp. 345–350.

[17] A. Muaremi, B. Arnrich, and G. Tröster, "Towards Measuring Stress with Smartphones and Wearable Devices During Workday and Sleep," *BioNanoScience*, vol. 3, no. 2, pp. 172–183, 2013.

[18] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer, "Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. ACM, 2011, pp. 47–56.

[19] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 2010, pp. 179–194.

[20] Q. Xu, J. Erman, A. Gerber, Z. Mao, J. Pang, and S. Venkataraman, "Identifying diverse usage behaviors of smartphone apps," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 329–344.

[21] K. Huang, C. Zhang, X. Ma, and G. Chen, "Predicting mobile application usage using contextual information," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 1059–1065.

[22] M. Böhmer, L. Ganev, and A. Krüger, "Appfunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications," in *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 2013, pp. 267–276.

[23] A. Gruenerbl, V. Osmani, G. Bahle, J. C. Carrasco, S. Oehler, O. Mayora, C. Haring, and P. Lukowicz, "Using smart phone mobility traces for the diagnosis of depressive and manic episodes in bipolar patients," in *Proceedings of the 5th Augmented Human International Conference on - AH '14*. New York, New York, USA: ACM Press, Mar. 2014, pp. 1–8. [Online]. Available: http://dl.acm.org/citation.cfm?id=2582051.2582089

[24] G. Chittaranjan, J. Blom, and D. Gatica-Perez, "Mining large-scale smartphone data for personality studies," *Personal and Ubiquitous Computing*, vol. 17, no. 3, pp. 433–450, 2013.

[25] L. Xie, X. Zhang, J.-P. Seifert, and S. Zhu, "pbmds: a behavior-based malware detection system for cellphone devices," in *Proceedings of the third ACM conference on Wireless network security*. ACM, 2010, pp. 37–48.

[26] I. Burguera, U. Zurutuza, and S. Nadjm-Tehrani, "Crowdroid: behavior-based malware detection system for android," in *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*. ACM, 2011, pp. 15–26.

[27] 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," in *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '14*. New York, New York, USA: ACM Press, May 2014, pp. 1–7.

[28] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 3–14.