Towards Recognizing the Emotions of Developers Using Biometrics: The Design of a Field Study

Daniela Girardi♦, Filippo Lanubile♦, Nicole Novielli♦, Luigi Quaranta♦, Alexander Serebrenik*
♦University of Bari Aldo Moro, Italy
*Eindhoven University of Technology, The Netherlands
{daniela.girardi, filippo.lanubile, nicole.novielli}@uniba.it, l.quaranta6@studenti.uniba.it, a.serebrenik@tue.nl

Abstract— **During their daily working activities, developers experience a wide range of emotions that are known to impact their personal wellbeing and, consequently, their work performance. As such, being aware of own and collaborators' emotions is crucial to enhance the collaborative development process. In this paper we present the design of a field study aimed at i) assessing the feasibility of emotion detection using non-invasive biometric sensors and ii) investigating the correlation between daily working activities and positive/negative emotions experienced by software developers. The long-term goal of our research is to provide recommendations to improve developers' mental well-being and productivity based on the emotions they experience.**

Keywords— Emotion detection, biometric sensors, field study, empirical software engineering

I. Introduction

The creator of the Linux kernel, Linus Torvalds, recently communicated to the Linux kernel developers he was willing to "take time off and get some assistance on how to understand people's emotions and respond appropriately" [17]. These words, emphasized by the role of their author, are just an example that reveals the importance of emotions in the workplace and the potential for emotion awareness. This is confirmed by research studies showing that emotions have a big impact on developers' perceived productivity [10][21] and social interactions [1]. Graziotin et al. [8] report about a correlation between developers' positive emotions and productivity. Conversely, unhappiness is reported to induce troubles when traying to stay focused on the problem at hand, causing drops in self-esteem and lose of motivation [9]. Emotion awareness can also be beneficial in project teams, where dynamics of encouragement (i.e., the ability to instill joy and hope among all the team members) and playfulness (i.e., the ability to create a context that encourages experimentation and tolerate errors) come into play [1].

In this paper, we present the design of a field study aiming at improving emotion awareness of developers through wearable devices. Specifically, we will investigate whether it is possible to automatically detect developers' emotions from skin and heart measurements recorded by a non-invasive wristband, Empatica E4, with the final goal to give some suggestions (e.g. "take a break") when negative emotions are detected. Moreover, we will examine whether there are some software engineering activities in which negative emotions prevail. This information can be used in further research to better understand the causes of negative emotions and the situations in which they are likely to occur, allowing targeted actions to improve developers' wellbeing.

We have thus formulated the following research questions:

RQ1: Can we identify emotions experienced by software developers using biometric sensors' during their working day?

RQ2: Is there a connection between developers' activities and emotions?

As this is an initial pilot study, we will involve one participant from a small software development company to validate the experimental protocol and examine the comfort of the chosen device over different days. In future, we plan to involve more subjects to give more general and robust results to the scientific community.

The rest of the paper is organized as follows. In Section II, we provide background on emotion detection and use of biometrics in software engineering. In Section III, we report the preparation of the study, whose design is reported in detail in Section IV. Finally, we conclude in Section V.

II. Background

A. The Circumplex Model of Affect

Consistently with prior research on emotion awareness in software engineering [12][18][19][21], we refer to the Russell's model [24], which represents emotions along a bi-dimensional representation schema, including valence and arousal in the horizontal and vertical axes, respectively. *Valence* describes the pleasantness of an emotional state. Pleasant emotional states, such as joy or amusement, are associated to high valence, while unpleasant ones, such as sadness or fear, are associated to low valence. *Arousal* describes the level of activation of the emotional state ranging from inactive (low arousal), as in boredom or sleepiness, to active (high arousal) as in excitement or anger.

B. Sensor-based Emotion Detection

Affective computing studies have largely investigated emotion recognition from several physiological signals, either alone or in combination [4][7][13][26][27], thus confirming the link between emotion and physiological feedback. In literature several physiological measures have been exploited to detect emotions, such as the electrical activity of the brain (EEG), electrical activity of the skin (EDA), electrical activity of the contracting muscles (EMG) and blood volume pulse (BVP), from which heart rate (HR) and heart rate variability (HRV) can be derived.

As regards the EEG, previous research has demonstrated the link between the brain activity and the cognitive [6] and emotional states [22][26]. To record EEG, a variety of headsets have been used, with a different level of intrusiveness. Some studies [13][26] used headsets with 64 or 32 electrodes, which are very precise but uncomfortable to wear. As such, research is now focusing on investigating the feasibility of emotion detection with non-intrusive and low cost devices, equipped with a smaller set of electrodes (i.e. up to five)[7] [21].

Concerning EDA, studies in psychology show that it varies considerably with respect to changes in emotional intensity, especially for emotions with high arousal [3]. As

Fig.1 Empatica E4 wristband

such, EDA has been widely employed in emotion recognition [4][13][23]. EDA can be measured using electrodes positioned on the palmar surface of the fingers or trough wearable devices, such as wristbands.

As for EMG, has been found that muscle activity increases during emotions with negative valence [11]. To capture EMG, sensors are usually placed over the face [13], since facial muscles clearly reflect changes in emotional states [5].

Finally, BVP is the volume of blood that passes through the tissues in a localized area with each beat (pulse) of the heart and can be used for calculating i) the heart rate (HR) and ii) the variation in time interval between heartbeats, that is the heart rate variability (HRV). Heart-related measurement are usually captured by a plethysmograph and employed for recognition of cognitive and emotional states [4][6].

C. *Use of biometrics in SE research*

The software engineering community has investigated the use of biometrics with different goals. For example, Fritz et al. [6] have used a light weight headset, a wristband and an eye tracker to assess the difficulty in code comprehension tasks in order to stop developers before they introduce a bug. In another study, Muller and Fritz [21] used the same set of sensors to predict emotions experienced by developers during programming tasks. Indeed, negative emotions are symptoms that a developer is stuck and is not making progress, so recommendation can be provided to unblock the situation and restore developers' mental wellbeing. Moreover, Zueger et al. [28] have used a chest strap to record the heart rate activity and a wrist-worn activity tracker to predict developers interruptibility, because of this kind of information can be used to optimize the timing of interruptions and minimize disruption. The use of eye tracker has also been explored in the field of software development [25]. Indeed, it can be used as an assessment tool for software artifacts, tools, and techniques. Moreover, eye tracking metrics can contribute to several software traceability tasks.

III. Preparation

To conduct our study, we contacted a small company, named Apuliasoft [1] due to the company's interest in monitoring emotions during working days. Apuliasoft offers software development, consulting and training services and is specialized in web and mobile development. The team is composed from eight developers, plus the CEO, the project manager, the R&D manager and the chief happiness officer. We will conduct the study with the project manager firstly since he is involved in several different activities, including development.

To understand the kind of activities and the schedule of the envisaged participants, the first author of this paper spent a day in the software company. Once arrived, the observer was welcomed and took a seat at one desk in the same room where the participant worked, that was positioned in the front of the participant desk. Other two developers were in the same place. Before starting the working day, she asked to the participant to wear Empatica, in order to test the comfort of the device. Then, both the participant and the observer started their own work. The observer took notes about the activities that the developer performed (e.g., when a colleague asked for help).

She noticed that the participant was involved in several different activities, including coding, reading and writing e-mails, helping colleagues and networking (e.g. socializing with colleagues during a break). In addition to these, the subject revealed that during the week he also usually i) deals with bug fixing, testing, designing new solutions, ii) writes documentation, participates to meetings, iii) is involved in training activities and iv) carries out administrative tasks. Starting from these observations and the taxonomy of activities by Latoza et al.[16], we have derived the categories of activities to be used in the self-report to assess the emotional state during working activities.

At the end of the day, the participant reported that he felt comfortable wearing Empatica.

IV. Field Study Protocol

A. *Instrumentation*

Empatica. We have chosen to use Empatica E4[2] wristband (Figure 1) since it is an unobtrusive wearable device. Non-invasiveness, comfort and ease of use are crucial criteria in the choice of the devices to employ for biofeedback acquisition, Indeed, sensors should be comfortably worn by developers' without causing additional distress or negative emotions. Differently from laboratory settings, where sensors must be worn for a limited period and the participant do not meet other people, in real context it is essential to propose devices that could not embarrass the subject, such as more or less intrusive headsets. Empatica E4 is equipped with EDA and BVP sensors.

ESM. We intend to apply the experience sampling method (ESM) to elicit systematic self-reports at multiple occasions during working hours [15]. Software engineering researchers have previously used ESM to study the interruptibility of

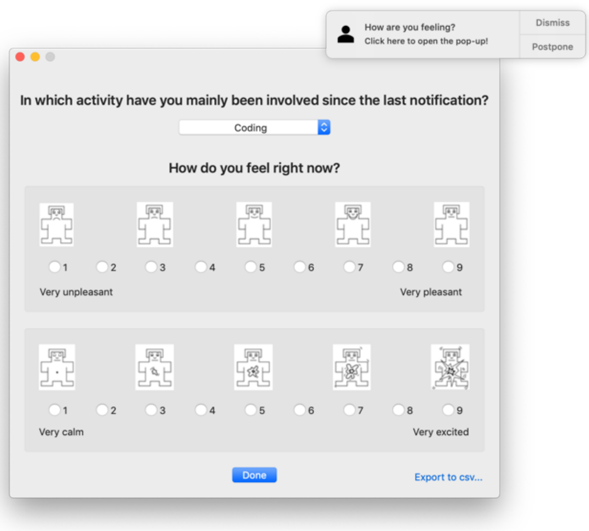

Fig. 2. Screenshot of the pop-up for emotion assessment

[1] https://www.apuliasoft.com/

[2] https://www.empatica.com/en-eu/research/e4/

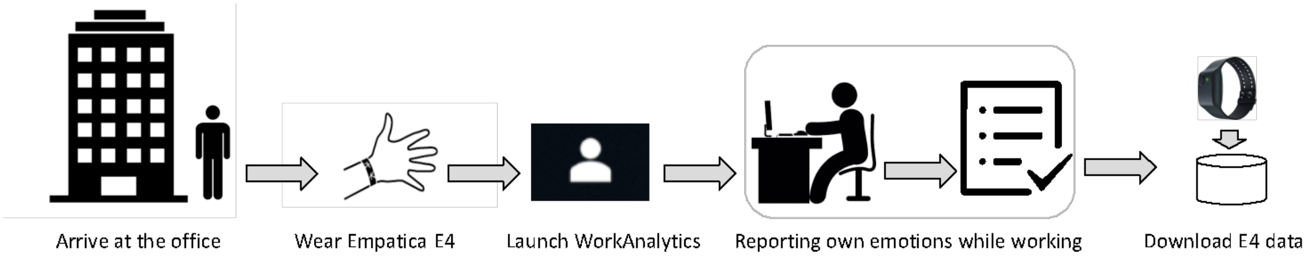

Fig.3. Workflow of the experimental protocol

developers [28] as well as the link between their emotions and the active contribution to software repositories [14].

WorkAnalytics. The ESM approach will be used together with *WorkAnalytics* [20], a tool for monitoring the perceived productivity of developers by means of a pop-up periodically appearing on their screens. In addition, *WorkAnalytics* captures information about various individual aspects of the software development work, including application usage, document accessed, development project worked on, websites visited and collaborative behaviors (i.e. attending a meeting or using email, instant messaging and code review tools). As such, we decided to use *WorkAnalytics* to gather background information on the context in which the emotion is experienced and reported, i.e. the activity performed by the developer when the pop-up is shown. The data collected are stored locally by the application, in order to avoid dealing with privacy issues. The log of activities can be summarized and visualized to provide users with clear and actionable insights on the way they spend their time at work.

Since we intend to assess the emotional state of a developer during his activities, we integrated a new pop-up window into WorkAnalytics[3] (See Figure 2). This pop-up can be used as an alternative to the original one, allowing perceived productivity self-assessment. It is meant to briefly ask information about the last activity carried on by the user and his current emotional state. The first question requires the developer to pick a task category from the following list: 1) coding, 2) reading/writing e-mails, 3) helping colleagues, 4) networking; categories have been derived from the activities observed on site at the company and validated with the pilot participant himself. The second question is answered in two different parts, each aimed at assessing one of the two dimensions of Russell's circumplex model [24]. In particular, valence and arousal rates are collected using a nine-point Likert scale, through a Self-Assessment Manikin (SAM) [2].

Every hour since the application is run, the user receives a notification with the default system sound. If he clicks in the notification's area, the pop-up window appears, and the form with the two questions can be filled. The notification can also be dismissed or postponed. In the former case the pop-up must be launched manually next time, while in the latter case the developer can choose to see it back in one or two hours. When the pop-up eventually closes, a new notification is automatically scheduled an hour later.

The trackers in WorkAnalytics can be paused and resumed altogether. Our pop-up makes no exception: the developer can use the option available in the main menu of the application to suspend the sampling for a certain amount of time and then resume the acquisition process.

B. *Procedure*

Some preliminary steps will be needed. Specifically, we will explain the purpose of the study to our participant to motivate him in using the tool, describing the positive benefits that can be returned from keeping track of emotions during working days. Then, we will introduce Empatica wristband and show how it works. As a next step, we will assist the participants in installing and setting *WorkAnalytics*. Finally, before starting the experiment, we will ask the participant to sign the consent form for personal data treatment.

We plan to collect data for two weeks, during which each participant must wear Empatica E4, turning it on and off, respectively, when in the office or out of it. The participant must launch *WorkAnalytics* when starting to work and answer to the pop-up questions when the related notification appears on the screen. The pop-up will be prompted once per hour by default, but there is also the possibility to postpone it (see IV.A). Right after filling in the pop-up, before submitting, the participant must press the Empatica E4 event marker button to record the timestamp of the interruption. Each marker will delimit different time intervals from which we will extract the data for the analysis. At the end of the day, the participant must synchronize data collected on the device using the E4 Manager application. We report the workflow of the experimental protocol in Figure 3.

V. CONCLUSIONS

The main goal of this pilot study is to explore whether emotion awareness of software developers can be improved using machine learning models able to detect developers' emotions with data captured by a non-invasive and low-cost biometric sensor. We plan to conduct this initial study with just one participant, with the goal to expand our findings involving more subjects. Additionally, we will investigate which developers' activities increase the likelihood to feel negative emotions.

ACKNOWLEDGMENT

This work is partially funded by the project "E-SHELF", under the program INNONETWORK, POR Puglia FESR-FSE 2014-2020 and the project ``Emotion recognition based on biometrics" funded by the Department of Computer Science of the University of Bari.

REFERENCES

[1] A. E. Akgun, Halit Keskin, John C. Byrne, Ayse Gunsel. 2011. Antecedents and Results of Emotional Capability in Software Development Project Teams. Journal of Product Innovation Management. 5 (July 2011), 1-17.

[3] Our fork available at https://github.com/collab-uniba/PersonalAnalytics

[2] M. M. Bradley and P. J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavioral Therapy and Experimental Psychiatry, Volume 25, 49-59.

[3] M.M. Bradley, P.J. Lang. 2007. Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), Handbook of psychophysiology, pp. 581-607. New York, NY, US: Cambridge University Press.

[4] F. Canento, A. Fred, H. Silva, H. Gamboa and A. Lourenço. 2011. Multimodal biosignal sensor data handling for emotion recognition. SENSORS, 2011 IEEE, Limerick, 647-650.

[5] P. Ekman and W. Friesen. 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.

[6] T. Fritz, A. Begel, S. C. Müller, S. Yigit-Elliott, and M. Züger. 2014. Using psycho-physiological measures to assess task difficulty in software development. In Proceedings of the 36th International Conference on Software Engineering (ICSE 2014). ACM, New York, NY, USA, 402-413.

[7] D. Girardi, N. Novielli, F. Lanubile. 2017. Emotion Detection Using Noninvasive Low Cost Sensors. In Proceedings of the 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII 2017), 125-130.

[8] D. Graziotin, X. Wang, P. Abrahamsson. 2013. Are Happy Developers More Productive?. In Proceedings of PROFES 2013. Lecture Notes in Computer Science, vol 7983. Springer, Berlin, Heidelberg.

[9] D. Graziotin, F. Fagerholm, X. Wang, and P. Abrahamsson. 2017. Consequences of unhappiness while developing software. In Proceedings of the 2nd International Workshop on Emotion Awareness in Software Engineering (SEmotion '17). IEEE press, Piscataway, NJ, USA, 42-47.

[10] D. Graziotin, X. W., and P. Abrahamsson. 2015. Do feelings matter? On the correlation of affects and the self-assessed productivity in software engineering. J. Softw. Evol. Process 27, 7, 467-487.

[11] K. Gouizi, C. Maaoui and F. Bereksi Reguig. 2014. Negative emotion detection using EMG signal. In Proceedings of International Conference on Control, Decision and Information Technologies (CoDIT), 690-695.

[12] M. R. Islam and M. F. Zibran. 2018. DEVA: sensing emotions in the valence arousal space in software engineering text. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC '18). ACM, New York, NY, USA, 1536-1543.

[13] S. Koelstra., C. Muhl, M. Soleymani, J. S. Lee, A.Yazdani, T. Ebrahimi, T. Pun, A.. Nijholt, and I. Patras. 2012. Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, Volume 3, 18-31.

[14] M. Kuutila, M. V. Mäntylä, M. Claes, M. Elovainio, and B. Adams. 2018. Using experience sampling to link software repositories with emotions and work well-being. In Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '18). ACM, New York, NY, USA, Article 29, 10 pages.

[15] R. Larson and M. Csikszentmihalyi. 1983. The Experience Sampling Method. New Directions for Methodology of Social & Behavioral Science, 15 (March 1983), 41-56.

[16] T. D. LaToza, G. Venolia, and R. DeLine. 2006. Maintaining mental models: a study of developer work habits. In Proceedings of the 28th international conference on Software engineering (ICSE '06). ACM, New York, NY, USA, 492-501.

[17] Linux 4.19-rc4 released, an apology, and a maintainership note. https://lkml.org/lkml/2018/9/16/167. Accessed: 2019- 01- 07

[18] M. V. Mäntylä, N. Novielli, F. Lanubile, M. Claes, and M. Kuutila. 2017. Bootstrapping a lexicon for emotional arousal in software engineering. In Proceedings of the 14th International Conference on Mining Software Repositories (MSR '17). IEEE Press, Piscataway, NJ, USA, 198-202.

[19] M. Mäntylä, B. Adams, G. Destefanis, D.l Graziotin, and M. Ortu. 2016. Mining valence, arousal, and dominance: possibilities for detecting burnout and productivity?. In Proceedings of the 13th International Conference on Mining Software Repositories (MSR '16). ACM, New York, NY, USA, 247-258.

[20] A. N. Meyer, Gail C. Murphy, Thomas Zimmermann, and Thomas Fritz. 2017. Design Recommendations for Self-Monitoring in the Workplace: Studies in Software Development. Proc. ACM Hum.-Comput. Interact. 1, CSCW, Article 79 (December 2017), 24 pages.

[21] S. C. Müller and T. Fritz. 2015. Stuck and frustrated or in flow and happy: sensing developers' emotions and progress. In Proceedings of the 37th International Conference on Software Engineering - Volume 1 (ICSE '15), Vol. 1. IEEE Press, Piscataway, NJ, USA, 688-699.

[22] M. Murugappan and S. Murugappan. 2013. Human emotion recognition through short time Electroencephalogram (EEG) signals using Fast Fourier Transform (FFT). IEEE 9th International Colloquium on Signal Processing and its Applications, 289-294.

[23] R. W. Picard, E. Vyzas and J. Healey. 2001. Toward machine emotional intelligence: analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 10, 1175-1191.

[24] J. A. Russell. 1980. A Circumplex Model of Affect. Journal of Personality and Social Psychology, Volume 33, 1161-1178

[25] B. Sharif and H. Kagdi. 2011. On the use of eye tracking in software traceability. In Proceedings of the 6th International Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE '11). ACM, New York, NY, USA, 67-70.

[26] M. Soleymani, M. Pantic and T. Pun. 2012. Multimodal Emotion Recognition in Response to Videos. IEEE Transactions on Affective Computing, vol. 3, no. 2, 211-223.

[27] G. Valenza and E.P. Scilingo. 2014. Autonomic Nervous System Dynamics for Mood and Emotional-State Recognition: Significant Advances in Data Acquisition. Springer, 2015.

[28] M. Züger, S. C. Müller, A. N. Meyer, and Thomas Fritz. 2018. Sensing Interruptibility in the Office: A Field Study on the Use of Biometric and Computer Interaction Sensors. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, New York, NY, USA.