Our Actions, Ourselves: How Unconscious Actions Become a Productivity Indicator.

Manuel Rodrigues, Ricardo Santos
CIICESI, ESTG-IPP - School of Technology and Management
Polytechnic of Porto
Felgueiras, Portugal
{mfsr, rjs}@estg.ipp.pt

Paulo Novais
Informatics Department/ Computer Science and Tech. Centre
University of Minho
Braga, Portugal
pjon@di.uminho.pt

Abstract— Productivity always was, and still is, the main goal of organizations, that being economic, governmental, military or educational. Having the means to control, detect and monitor features that have impact on productivity is a major issue, and subject to various investigation. Considering that most of the times, if not always, unconscious actions play a very important role in the way we work, study, socialize, and even in the way we have fun, the high significance of that factors becomes very clear. Monitoring unconscious actions, selecting those of them that do play a role regarding productivity, and trying to proactively take measures to improve processes, is then the goal of this work. Specifically, we are concerned about using computers peripherals to non-intrusively monitor user´s actions. The term non-intrusively assumes greater importance, as we are concerned with unconscious actions, thus we need to strongly ensure that no entropy is derived by the way this process is done. Peripherals such as mouse, keyboard, touch screens, and possibly webcams and microphones can act as sensors, completely hidden from the user. As we use them daily, they somehow assume part of our life, and can be used to collect data that will be processed to get useful information regarding that particular user. We then can build a behavioral profile, for instance, that will provide a better insight of user´s actions. We can predict some possibly negative features, such as stress, fatigue, level of attention, for instance. If detected or predicted, they can greatly help to better manage all the information we need, in the right way. We can suggest that someone takes a coffee break, because she/he´s stressed. We can tell him/her to work/study in the morning, because the information we have collected suggests that´s the period of the day that´s more suitable to get better results, for that person. We can suggest postponing the following meeting, because the actual mood indicates that that person is more suitable to conflicts at that moment. In short, we aim to use computer peripherals and smartphones to collect data from the user, non-intrusively, aiming to detect or predict behavioral features (stress, fatigue, attention) and unconscious actions that will allow us to build a behavioral profile about the user, thus making it possible to improve productivity for instance. This is accomplished by monitoring mouse, keyboard and touch screen usage, non-intrusively, and in real time. Through the collected data, some inferences are made regarding the patterns of interaction that make it possible to detect variations regarding behavioral features and unconscious actions. With this information, we can have a detailed insight into those behavioral features, allowing us to proactively mitigate some of the potential problems that could arise. In this work a framework is proposed as a way to integrate all these features. We aim mainly to apply these concepts in learning contexts, as to improve student's outcomes.

Keywords— *Sensing, Stress, Fatigue, Attention, Deep Learning, User Profiles*

I. INTRODUCTION

Whatever the context is, we are always concerned in ways to improve the outcomes of any process, in any given activity. We live surrounded by technology, so computers, smartphones, tablets, do play an important role: that being in our work, in the way we learn, in how we socialize, in how we are medically treated, even in the way we relax. It seems naturally then, that technological devices that we use every day, can be used to monitor our actions, is such a way that we do not realize it, while acquiring and gathering data that will be used to characterize each user, possibly to build a model/profile of him. In possession of such information, the possibilities are huge, regarding what we can do with it. We can use that information to authenticate users randomly throughout the work, and not only at the beginning, improving security issues as in [1], we can detect stress and fatigue levels as in [2], allowing for a better work distribution throughout the day, thus improving outcomes, we can test for students attention and commitment to work, as in [3] allowing for better student success. Ultimately, with successive modelling of users, we can build and upgrade profiles that will help the way we interact with one person (user), that is, we can have available a model of the user, that will suggest us the best way/day/hour to make a request to that person, because it can predict that in that time, it will be more receptive. It can suggest to that particular user that it is time to take a break, because their monitored levels of attention, stress fatigue are getting near "dangerous" values. The possibilities are endless.

However, all these must be made non-intrusively. The user must not be aware that he/she is being monitored, as we believe this is the best way to gather data, that isn´t contaminated by the user knowing what is happening. A data collection tool was developed, tested and used in [4], but the domain of usage was to detect stress. Using that same technology, we aim to extend that framework making it possible to build user models that can be used in many more ways and not only in stress detection. Smartphone´s various sensors will be used, and web cams too, in order to get the maximum data from user's interactions. We aim to build a user profile, dynamically updated throughout the day, making it possible to know the current state of that user. Here, by state we mean his/her level of engagement in his/hers tasks, stress levels, attention, his/her current mood: motivated, happy, sad, angry… [5],[6].

With these information in hands, important actions can be made such as: recommend breaks to restore stress levels, reschedule tasks, for instance, if the system knows that a particular student (for instance) gets less stressed and less fatigued in the morning, it can suggest that the most difficult or more urgent tasks, be made in that period, so that the probability of success is higher. The ultimate goal would be to have a virtual representation of the user, in that moment, enabling better interaction and better results, thus improving his/her productivity whatever the context is (learning, working,, …).

In this work we propose a framework to address all these previously referred issues. This framework is an improvement of an earlier work that has already been tested and implemented. With the new improvements, we aim to obtain newer and better results, in various fields, and not only in stress detection.

II. BACKGROUND ISSUES

To accomplish what we propose to in this work, several technologies need to be addressed. In such a system, the following stages need to be implemented: the data is collected using everyday devices that we use, as sensors (keyboard, mouse, touch screen, webcam); That raw data is pre-processed in the device and uploaded to a webserver; The pre-processed data is processed and stored in a database; With the analyses of that data, some features are detected and or predicted, and user profiles are derived, and stored in a database. Several techniques can be used, namely deep learning to accomplish these tasks. Information regarding those features (stress, fatigue, and burnout) can then be sent to the user, acting as recommendations to improve processes. All these steps are briefly mentioned next.

A. Sensing Data

Several studies exist on using keyboard dynamics to recognize emotions on users[7]. With the widespread of smartphone usage, various sensors provided by those equipment's can also be used as in [8], where they try to answer questions such as: Why do some students do better than others? Under similar conditions, why do some individuals excel while others fail? Why do student's burnout, drop classes, even drop out of college? What is the impact of stress, mood, workload, sociability, sleep and mental wellbeing on educational performance? Also, various input channels can be considered while trying to achieve an emotion recognition tool. Using webcam´s we can detect emotions through facial expression analysis as in [9]; Multimodal emotion recognition it's possible using audio, as in [10]; Analysing texts, for instance posts on social networks, allows to identify emotional characteristics as done in [11]; Physiological signals can be used to detect emotions as in [12], (but to do so, some intrusive devices must be used); Finally a multi-modal approach can also be used as in [13].

It becomes clear than, that it is possible to obtain such useful information while monitoring non-intrusively user's actions, when they use computers, smartphones or tablets

B. Data Processing and Storage

To obtain the information we need it is necessary to collect and label data samples, define and calculate some characteristic features regarding that user, and at the end, develop and test profiles is necessary. As so, we realize that big amounts of unstructured data are collected, and it is necessary to pre-process and store them. Due to its unstructured data, relational databases do not seem to be the best choice. MongoDB stands as a possible choice due to its MapReduce and Aggregation pipeline tools. Both the Aggregation pipeline and MapReduce can operate on a shared collection (partitioned over many machines, horizontal scaling). These are powerful tools for performing analytics and statistical analysis in real-time, which are useful for ad-hoc querying, pre-aggregated reports, and more. MongoDB provides a rich set of aggregation operations that process data records and return computed results; using these operations simplifies application code and limits resource requirements [14], [15].

C. Machine Learning

Machine learning can be: Supervised- all data is labelled, and the algorithms learn to predict the output from the input data; Unsupervised - all data is unlabelled, and the algorithms learn to inherent structure from the input data; Semi-supervised- some data is labelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used.

Machine learning classification builds a predictive model based on the available data. Typically, the process consists of getting a feature set (input) which the machine learning algorithm must correctly label (output). Regarding the collected data previously mentioned, as it is unlabelled, it falls into the unsupervised category. The goal is to model the underlying structure or distribution in the data in order to learn more about the data itself, i.e., building user profiles. Typically, unsupervised learning problems can be further grouped into clustering (discover the inherent groupings in the data, such as grouping customers by purchasing behaviour) and association problems (discover rules that describe large portions of your data, such as people that buy X also tend to buy Y). Some popular examples of unsupervised learning algorithms are: k-means for clustering problems and Apriori algorithm for association rule learning problems.

Machine learning techniques, can be applied to predict user behaviour as in [16]. Also, Deep learning emerges as a new promising field to be applied in this area. Deep learning refers to a wide class of machine learning techniques and architectures, using many layers of non-linear processing. The concept of deep learning originated from artificial neural network research. In the Australian Nacional University, a project is being developed to extract user profiles using Deep Learning techniques [17].

III. FRAMEWORK – DYNAMIC BEHAVIOURAL PROFILER SYSTEM

In order to achieve what we have been mentioning, the following improved framework is presented (figure 1). It aims to constantly monitor user's actions, deriving useful information from the collected data that can be used to determine users stress, fatigue, productivity, current mood and other helpful information. Building a user behavioral profile that represents the current state of that user is also in mind.

A. Client Side – Data Acquisition and Pre –Processing Module

The proposed system works with the client-server paradigm. On the client side, users are continuously monitored, throughout the devices available that act as sensors, and that can be used non-intrusively: keyboard, mouse, webcam and smartphones sensors such as accelerometer, motion detector... The raw data is collected, by the Data Acquisition Module. The Data Pre-Processing Module is responsible for applying redundancy elimination techniques (such as data segmentation, aggregation and measurements), that filter the data to be sent to the server. At a configurable rate, the client connects to the server, and sends the pre-processed data. This Pre-processing is very important to prevent big amounts of unwanted and useless data to be sent for processing at the server side. As previously referred, a data collection tool was developed in [2], [4], to collect data, using keyboard and mouse. In this work we are extending that framework, to accommodate new ways of sensing data, to improve the gathering of information about a particular user. As the use of smartphones is rising, such devices will be also used to acquire data. Several studies have been made regarding this usage,[5],[6]. Gathering accelerometer data for instance, can tell us if the user is moving, or not, and if there are sudden movements, that could possible indicate some angry state (this can be accomplished using Google ActivityRecognitionAPI). As all sensed data has a timestamp, we can associate it with a period of the day (morning, afternoon, night). Using microphone, it can be established the normal intensity of speaking for that user, thus allowing to detect variations that can indicate some abnormal state. Users' privacy is a major concern. In order to protect users' personal information, each participant's identity
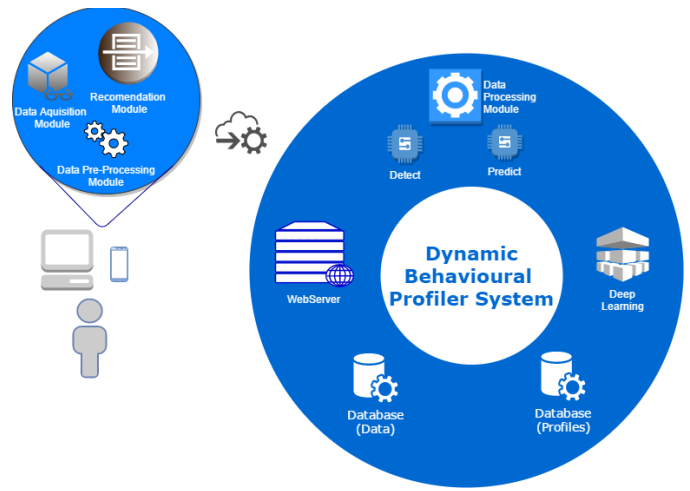

Fig. 1. The Proposed Framework

will be kept anonymous with a random user id and kept the user id map separate from all other project data so that the data cannot be traced back to individuals. Table 1 shows some possible data collected (only keyboard and mouse).

Table 1 - Example of collected data (keyboard and mouse)

Event	Description
MOV	Mouse movement, in a given time, to coordinates (posX, posY)
MOUSE_DOWN	Mouse button pressed down (left/right), in a given time and position (posX, posY)
MOUSE_UP	An event similar to the previous one, but describing the second part of the click, when the mouse button is released
MOUSE_WHEEL	This event describes a mouse wheel scroll of a given amount, in a given time
KEY_DOWN	The event of pressing down a given key, at a given time
KEY_UP	The event of releasing a given key, at a given time

B. Client Side – Recommendation Module

Also, at a regular timestamp, information is sent to the users, by the webserver, which aims to improve user's productivity. That information is based on the data that was previously sent and processed. With that processing, features like stress level, fatigue level, attention, burnout possibility, can be estimated and communicated to the user. Also, a user profile is being updated/created in the process, allowing for a full characterization of that user. With this user profile, we aim also to group users according to their similar features.

C. Server Side

The server side consists of the core element of our proposed framework. It will be formed by the following modules, each one dedicated to a specific task:

1) Server Side – Webserver

Responsible for attending client's requests. Its main goal is to receive the data from the clients and store it in the database. Upon processing, it will interact with the client's

recommendation module, to send useful information and alerts to the user.

2) *Server Side – Database (Data, Profile)*

As it was presented before, conventional databases are limited regarding flexibility (capacity to deal with semi-structured or unstructured data) issues. Also, the relational model shortcomings on scalability frequently cause bottlenecks on big data projects. Keeping this in mind, it was decided to use MongoDB as the storage engine. It´s capabilities and data processing features seem to be adequate to the problem in hands. This database will be responsible for storing the received data, and will be accessed by the other modules in order to extract all the features mentioned previously, through the processing of the data. The ability to operate on a shared collection (partitioned over many machines, horizontal scaling) is also a very important feature regarding performance. Also, it will keep users profile information (in a separate database).

3) *Server side – Data Processing Module, Deep Learning Module*

These modules will process the stored data, aiming to extract all the features previously mentioned. They constitute an analytic layer whose goal is to provide powerful tools for performing analytics and statistics analysis in real-time and building user behavioral profiles. It is important that in this process some values are filtered to eliminate possible negative effects on the analysis. The system calculates, at regular intervals, an estimation of the general level of each of the features we are dealing with: stress, performance, attention, fatigue, performance and attention of each student. The user behavioural profile is continuously updated, so based on the current and past values, recommendations can be sent to the users. These process needs a sufficient large dataset and a learning process. As previously mentioned, this work as already been done for stress detection in (Rodrigues et al, 2013), (Pimenta et al, 2015). We are now extending that work to new features and incorporating deep learning techniques to improve the results. Thus the learning algorithm is the core of this system, using deep learning techniques, the goal is to use the information collected about the characterization of human-computer interactions, and then improve the classifier accuracy on the task of estimating the parameters previously referred. As the flow of data obtained from the user's interaction is constant, there is an opportunity for continuously improving the results of the classification task, i.e., continuously learn from users interactions. A user's behavioural profile database is also created, allowing for longer time frame analyses.

IV. CONCLUSIONS

In our previous work, we could establish that it was possible to detect and predict stress levels on computer users, using keyboard and mouse interaction data (non –intrusively). Taking this as a starting point, we aim to extend it to detect and predict some new features such as fatigue, attention, performance, globally improving user's productivity, whatever the context is. To do this, a framework is proposed, describing the architecture of such a system. We plan to use more devices as sensors, always non-intrusively. Smartphones and tablets provide a new range of possibilities to acquire data from user's

interactions throughout their several sensors. Computer vision is also in our mind, but that´s an issue to be carefully addressed since not everyone likes, or evens allows to be filmed while working/studying/socializing. Deep learning techniques arises as the logical move to improve the system. The continuous flow of sensed data constitutes an amazing field to apply these techniques in order to extract the features we want and build users behavioral profiles.

REFERENCES

[1] Pilsung K. Sungzoon C., (2015) Keystroke dynamics-based user authentication using long and free text strings from various input devicespges 72-93

[2] Rodrigues M., Gonçalves S., Carneiro D., Novais P., Fdez-Riverola F.,(2013) Keystrokes and Clicks: Measuring Stress on E-learning Students, Management Intelligent Systems, Second International Symposium, Jorge Casillas, Francisco J., Martínez-López, Rosa Vicari and Fernando De la Prieta (Eds.), Springer - Series Advances in Intelligent and Soft Computing Vol 220, ISBN 978-3-319-00568-3, pp 119-126, http://dx.doi.org/10.1007/978-3-319-00569-0_15

[3] Pimenta, A., et al., (2015), "Detection of distraction and fatigue in groups through the analysis of interaction patterns with computers." Intelligent Distributed Computing VIII. Springer International Publishing, 29-39.

[4] Gonçalves S.,Rodrigues M, Carneiro D, Fdez-Riverola F, Novais P, (2015) Boosting Learning: Non-intrusive Monitoring of Student's Efficiency, Methodologies and Intelligent Systems for Technology Enhanced Learning Advances in Intelligent Systems and Computing Volume 374, pp 73-80

[5] Miluzzo E., Lane N. D., Fodor K., Peterson R., Lu H., Musolesi M., Eisenman S. B., Zheng X., and Campbell A. T., (2008), Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application. In Proc. of SenSys.

[6] Choudhury T., Consolvo S., Harrison B., Hightower J., LaMarca A., LeGrand L., Rahimi A., Rea A., Bordello G., Hemingway B., et al. (2008) The mobile sensing platform: An embedded activity recognition system. Pervasive Computing, IEEE, 7(2):32–41

[7] Kołakowska A.,(2013) "A review of emotion recognition methods based on keystroke dynamics and mouse movements," in Proc. 6th International Conference on Human System Interaction, Gdańsk, doi: 10.1109/HSI.2013.6577879.

[8] Wang, Rui, et al., (2014) "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones." Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM.

[9] Ioannou S. V., Raouzaiou A. T., Tzouvaras V. A., Mailis T. P., Karpouzis K. C., Kollias S. D.,(2005), "Emotion recognition through facial expression analysis based on a neurofuzzy network," Neural Networks, vol. 18(4), pp. 423-435, doi: 10.1145/ 1980022.1980177.

[10] Schuller B., Lang M., Rigoll G.,(2002), "Multimodal emotion recognition in audiovisual communication", in Proc. IEEE Int. Conference on Multimedia and Expo, ICME, Lausanne, doi: 10.1109/ ICME.2002.1035889.

[11] Gill A. J., French R. M., Gergle D., Oberlander J.,(2008), "Identifying Emotional Characteristics from Short Blog Texts," in Proc. 30th Annual Conference of the Cognitive Science Society, pp. 2237- 2242.

[12] Szwoch W.,(2013), "Using Physiological Signals for Emotion Recognition," in Proc. 6th International Conference on Human System Interaction, Gdańsk, doi: 10.1109/HSI.2013.6577880.

[13] Landowska A.,(2015), "Emotion monitor – concept, construction and lessons learned," in Proc. Federated Conference on Computer Science and Information Systems, pp. 75-80, doi: 10.15439/2015F384.

[14] Van der Veen, Jan Sipke, Bram Van der Waaij, and Robert J. Meijer., (2012) "Sensor data storage performance: SQL or NoSQL, physical or virtual." Cloud Computing (CLOUD), IEEE 5th International Conference on. IEEE.

[15] Kang, Yong-Shin, et al., (2016), "MongoDB-based repository design for IoT-generated RFID/sensor big data." IEEE Sensors Journal 16.2 (2016): 485-497.

[16] Vieira, A. (2015). Predicting online user behaviour using deep learning algorithms. arXiv preprint arXiv:1511.06247

[17] A.N.University, (2017), "Extracting User profile with Deep Learning Techniques" https://rsise.anu.edu.au/projects/pid/0000001123