Management of Physical, Mental and Environmental Stress at the Workplace

Božidara Cvetković, Martin Gjoreski, Jure Šorn, Martin Frešer, Mitja Luštrek
Department of Intelligent Systems
Jožef Stefan Institute
Ljubljana, Slovenia
boza.cvetkovic@ijs.si

Maciej Bogdański, Katarzyna Jackowska, Michał Kosiedowski, Aleksander Stroiński
Poznań Supercomputing and Networking Center
Poznań, Poland
kat@man.poznan.pl

***Abstract* — We present the Fit4Work system for monitoring and management of physical, mental and environmental stress at the workplace. The system was designed specifically for older workers who are subject to sedentary stressful work in an office environment. It uses commercially available devices and intelligent methods, which utilize machine-learning models to monitor the three aspects of the users' lifestyle, and provide recommendations for improving them. The results show that the system can adequately recognize the user's physical activities, estimate energy expenditure and detect mental stress, as well as recognize and reason about unhealthy environment. The system provides recommendations according to the monitoring results.**

Keywords— Physical activity monitoring, mental stress detection, environment quality management, wearable sensors, ambient sensors

I. INTRODUCTION

Unhealthy lifestyle characterized by insufficient physical activity and increased mental stress is a major issue in the developed countries. The lack of physical activity is mainly caused by a decline in physical labor, widely available motorized transportation and the fast pace life, which leaves little time and energy for leisure physical activity. Sedentary lifestyle contributes to the development of cardiovascular diseases, diabetes, obesity and other health problems, making it the fourth leading cause of death [1].

Mental stress is largely work-related, as the competitive labour market results in productivity pressure and job insecurity. While some stress increases productivity, a prolonged exposure to stress can result in chronical stress, which is a trigger for slower body recovery, poor mental health and vulnerability to infections due to a weakened immune system. The European Commission estimated the economic cost of work-related stress at €20 billion a year due to absence from work and decreased productivity [2].

Besides the physical and mental stress, unhealthy office environment can also negatively influence people by causing environmental stress. The term Sick Building Syndrome (SBS) is used for buildings (environments) that cause the occupants various health and comfort problems, and is most commonly linked with air quality. It has been shown that bad environmental conditions can cause decreased productivity. Inappropriate temperature decreases the productivity by 10%, humidity by 5% and air pollutants by 6–9%, which can in turn result in a higher level of mental stress due to productivity pressure [3][4].

Monitoring of physical activity is very popular in terms of the number of smartphone applications, dedicated devices and even smartwatch applications already available on the market [5][6]. Smartphone-only applications for activity monitoring use the smartphone accelerometer to estimate the burned calories and the amount of movement based on the number of steps the user takes over a day. Dedicated devices such as Microsoft Band 2 use simplified activity recognition and machine learning to improve the estimated calorie burn. Apple Watch [7], for example, goes a step further by featuring hourly reminders if the person is sedentary, and provides feedback and motivation to reach the daily goal. However, if the user wants to be monitored accurately they have to input the activity they are performing. The Fit4Work system offers similar functionality without the need for manual input of activities.

Monitoring mental stress using commercial and unobtrusive devices is a relatively new and challenging research topic. Healey and Picard [10] were the first to show the feasibility of stress detection using physiological sensors, which required intrusive wires and electrodes. Hovsepian et al. [11] proposed cStress – probably the most advanced stress-detection system for everyday life until now – which continuously monitors the stress level using an ECG sensor. For the future, they proposed replacing the somewhat uncomfortable ECG sensor with a wrist device as seen in the latest commercial products [8][9], and using information on

2472-7571/17 $31.00 © 2017 IEEE
DOI 10.1109/IE.2017.20

the user's context in the stress detection, which is what we do in the Fit4Work module for monitoring stress.

The monitoring and control of indoor environment parameters is a popular topic in smart–building research, which mainly focuses on the trade-off between the occupants' comfort and energy consumption [12]. However, smart buildings are equipped with many sensors as well as automated heating and ventilation, which most buildings do not possess. To develop a system that does not need expensive equipment, one must develop virtual sensors for the detection/estimation of parameters that are not sensed directly, and develop prediction models for the effect of human actions on environmental parameters. A correlation between the environmental values and the occupancy or window state was reported in previous research [13], and the number of occupants was successfully estimated from one or multiple parameters [14]. The prediction of the dynamics of the parameters is usually done with the predictive control technique [15], which requires the user to describe the environment in detail (e.g., the material of walls, windows, etc.). In the Fit4Work system, virtual sensors and prediction models are built with machine learning, and combined with an ontology-based recommendation module that suggests actions to improve the environmental parameters.

The system presented in this paper was developed as a part of the Fit4Work project [16], whose objective is to use wearable and environmental sensors to monitor the user's physical activity, mental stress and quality of the environment. The results of the monitoring are inputs for recommendations to the user. We focus on older workers, who are at greater risk due to generally less robust health. The Fit4Work system needs an average accelerometer-equipped smartphone, the Microsoft Band 2 [17] wristband, and the NetAtmo commercial weather station [18]. It interacts with the user via a smartphone application, which is a natural choice considering that by the end of 2017, over a third of the world population is projected to own a smartphone (2.6 billion users), and the number of devices a smartphone can connect to is increasing [19].

II. User Requirements and Architecture

Considering the large number of applications and systems functionally similar to Fit4Work, we designed the system with the following objectives: (i) tailor it to the target population and ensure good user experience through careful examination of the user requirements; (ii) provide clear feedback and recommendations regarding the users' behavior and environment to demonstrate the system's short-term benefits [20]; (iii) do not require the users to change the way they function dramatically [21] or alter their environment; and (iv) use affordable commercial devices.

A. User requirements

To develop the system that would fit the needs of the foreseen end users we collected the user requirements about two aspects of the system: (i) the most common habits and lifestyle parameters needed to design the intelligent modules for physical, mental and environmental stress monitoring, and (ii) the user interface to ensure good user experience.

The enquiry about the users' habits and lifestyle was conducted through a questionnaire with a number of questions related to health and fitness status of the target population, the characteristics of their occupation and the attitude towards technology. We gathered answers from 277 persons aged between 50 and 75 years living in The Netherlands, Poland, Romania, Slovenia and Spain. The key findings are:

- *Fit4Work should support and motivate the user to maintain good physical fitness level and detect prolonged sitting* – Although 92% of respondents rated their physical fitness as fair or good, 80% of them have difficulty staying physically active in their everyday life. 80% of respondents declared that they do not sweat very often at work and 77% perform their work while sitting.
- *Fit4Work should automatically detect the most popular physical activities* – The survey showed that the most popular activities are walking (23% find it a popular activity), gardening (18%) and cycling (17%).
- *Fit4Work should combine measures to monitor mental stress with helping the user to maintain their physical fitness* – 82% of stressed respondents declared that they are bothered by mental stress rather than physical stress
- *Fit4Work should use commercial non-intrusive devices or devices the users already have* – Smartphones were most often requested to be used for improvement of users' health (70% of respondents were willing to use smartphones). Wrist devices were found as most likely wearables to be accepted (65% of respondents).

The enquiry about the user interface design was conducted through focus groups in workshops where we discussed the functionality and design of the system. The first workshop was conducted with a group of nine Dutch older adults aged 55–75. The primary requirements resulting from this discussion were related to ensuring the system provides tangible health benefits and that it does not bother the end user too much. These conclusions served as our starting point to design the system and its user interface. The design we came up with (see Section VI) was then iteratively

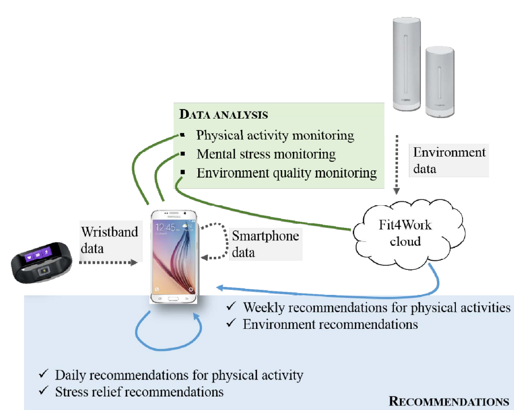

Fig. 1. Fit4Work system architecture.

evaluated through co-design techniques through a series of workshops with participation of a group of four Polish older adults aged 55–62 with various occupations (office worker, hairdresser, electronics technician and baker). We describe the methodology we used to evaluate the user interface and its usability in Section VII.

B. System architecture

The Fit4Work system relies on the two devices rated most acceptable by the users (a wristband and a smartphone), and a commercial weather station. The primary objective for selecting the devices was the number of embedded sensors. The chosen devices are the Microsoft Band 2 wristband and the Netatmo weather station, while most modern Android devices can serve as the smartphone. The smartphone plays a central role for interaction with the user through the user interface and for the analysis of the wearable sensor data, as well as for producing daily recommendations. The architecture is presented in Fig. 1.

The system is composed of three modules (physical activity module, mental stress module and environment quality module), each of which has its own component for data analysis (green box in Fig. 1): (i) the physical activity data analysis module utilizes data from the wristband and smartphone to recognize the current activity of the user and estimate the expended energy; (ii) the mental stress data analysis module utilizes data from the wristband and the results of the physical activity module to detect the level of mental stress; and (iii) the environment quality data analysis module utilizes data from the weather station to evaluate the current quality of the environment. The physical activity and mental stress data analysis is performed directly on the smartphone, whereas the environment quality data analysis is done remotely in the cloud. The results of the data analysis modules are evaluated and transformed into respective recommendations (blue box in Fig. 1), which are presented to the user through the user interface. The physical activity module produces two types of recommendations – daily and weekly; the mental stress module recommends stress relief exercises; and the environment quality module recommends the actions which will improve the quality of the environment. Each module is described in details in following sections.

III. PHYSICAL ACTIVTIY MODULE

The physical activity module aims at recognizing the current activity of the user and estimating the energy expended by the user during the activity. It also provides information about the daily and weekly achievements as well as physical activity recommendations when needed.

A. Physical activity data analysis

The physical activity data analysis method is composed of six steps as presented in Fig. 2. The method's input is data from an accelerometer-equipped smartphone or the Microsoft Band 2 wristband capable of measuring acceleration and physiological signals (e.g., heart rate). The output are the recognized activity and the estimated energy expenditure in MET (1 MET is defined as the energy

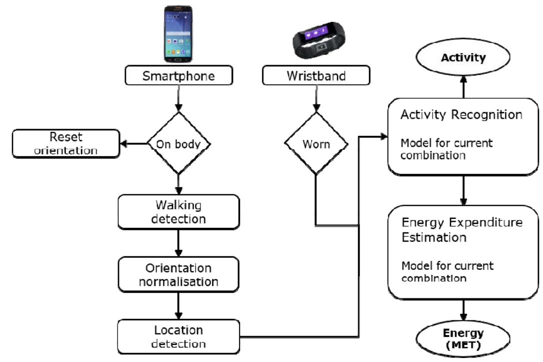

Fig. 2. Workflow of the physical activity data analysis.

expended at rest, while around over 20 MET is expended at extreme exertion).

The first step is the detection of the devices currently present on the user's body. The smartphone presence is detected with heuristics in which we utilize the smartphone's proximity sensor as well as the accelerometer to determine whether the device is being carried. The wristband self-reports its presence. If only the smartphone is present, the method anticipates a walking period of 10 seconds. Walking is detected with a machine-learning model in a location- and orientation-independent manner. The machine-learning model is trained to distinguish between walking and non-walking activity in two-second windows, and once the model recognizes the walking period (5 consecutive walkings recognized), the walking signal is used for normalizing the orientation (third step) of the smartphone, and for recognizing the location (fourth step) of the smartphone. The orientation is normalized under the assumption that the average acceleration during walking corresponds to the Earth's gravity. The average acceleration during the walking period is processed with the quaternion rotation transformation according to Tundo et al. [22], which gives us the orientation matrix to be used on further data. The normalized data is fed into the location detection machine-learning model, which is trained to recognize whether the smartphone is in the trousers pocket, jacket or a bag.

The present devices and the recognized location serve as a context for the selection of an appropriate machine-learning model for activity recognition. We trained four models: one for each location of the smartphone and one for the wristband. The recognized activity serves as one of the features in the feature vector of the energy expenditure estimation machine-learning model, which is again selected according to the present devices and the recognized location. For more information on the algorithm with a different device, the reader is referred to [23].

B. Physical activity recommendations

The recognized activity and the estimated energy expenditure aggregated in real time represents the amount of physical activity done so far (in a day, week, etc.). The estimated energy expenditure is transformed from MET into the more familiar kilo calories (kcal) according to Equation (1), as well as into three intensities (*energy* < 3 MET = light intensity, 3 MET < *energy* < 6 MET = moderate intensity, *energy* > 6 MET = vigorous intensity).

$$energy\ [\text{kcal}] = weight\ [\text{kg}] * energy\ [\text{MET}] * t\ [\text{hours}] \quad (1)$$

To monitor and motivate the user with goals and recommendations we have adopted the requirements for daily and weekly physical activity from the guidelines of the World Health Organization (WHO). The daily recommendations are that the user should (i) burn at least 200 active kcal (active kcal are burned during physical activities over 2.5 MET), (ii) walk at least 10 minutes continuously, and (iii) stand up and walk once per hour. The weekly recommendation is that the user should be engaged into 150 minutes of moderate-intensity physical activity or at least 75 minutes of vigorous-intensity physical activity (or an equivalent combination of both).

These recommendations are combined into daily and weekly goals. The daily goal is achieved if the user succeeds in burning 200 or more active kcal ($AC_{\text{recomm}} = 1$ in this case, and proportionally less otherwise), and they walk for 10 minutes continuously ($Walk_{\text{recomm}} = 1$ in this case, and proportionally less otherwise). The percentage of the achieved daily goal is calculated with Equation (2). The weekly goal is achieved if the user satisfies the need for moderate or vigorous activity; the percentage of weekly goal is calculated with Equation (3). The user is constantly presented with understandable information about the achievements (e.g., you have achieved 60% of your daily goal).

$$dailyGoal = 0.5 * AC_{\text{recomm}} + 0.5 * Walk_{\text{recomm}} \quad (2)$$

$$weeklyGoal = (t_{\text{moderate}}\ [\text{min}] + t_{\text{vigorous}}\ [\text{min}] * 2) / 150\ \text{min} \quad (3)$$

IV. MENTAL STRESS MODULE

The stress-monitoring module continuously monitors the user's stress, provides an overview of stressful events and suggests appropriate relaxation exercises when needed.

A. Mental stress data analysis

The mental stress data analysis is performed using the machine-learning method presented in Fig. 3. The method is applied on the data collected via the Microsoft Band 2. It consists of two machine-learning modules: a laboratory stress detector and a context-based stress detector.

The laboratory stress detector is a machine-learning model trained on laboratory data to distinguish stressful vs. non-stressful events in 4-minute data windows with a 2-minute overlap. As input it uses features computed from the physiological signals (blood volume pulse, heart rate, beat-to-beat intervals, skin temperature and electrodermal activity), and it outputs predictions of possible stress events.

In real life, there are many situations that induce a similar physiological arousal to stress (e.g., exercise, eating, hot weather, etc.), so the laboratory stress detector is inaccurate.

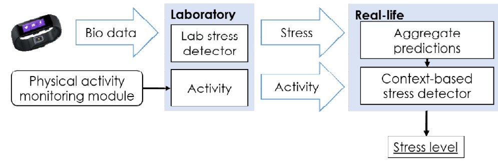

Fig. 3. Method for mental stress detection

Therefore we introduce a context-based stress detector to include real-life circumstances of potential stress events, thus improve the detection performance. This detector is a machine-learning model that uses as input the predictions of the laboratory stress detector, the information on the physical activity from the respective module, and other information such as the time of the day and a short history of predictions, to provide a final prediction every 20 minutes. The interval of 20 minutes was chosen empirically. [24]

B. Mental stress recommendations

The stress-monitoring module monitors the user's stress level continuously. The overall stress at time t (S_t) is calculated using the Equation (4), which combines the current stress at time t (C_t) as predicted by the stress-monitoring module, and the stress earlier in the day. The parameter a is set to 0.6 according to expert opinion, unless the user performed a stress-relief exercise, in which case it is 0.6 * 0.75 to provide an immediate positive feedback. The parameter b is set to 0.4. By subtracting $1 - S_t$, we obtain the relaxation score, which is a "positive" view of the stress presented in the user interface.

$$S_t = a * C_t + b * \sum_{i=0}^{t-1} C_i \quad (4)$$

Once the stress level (S_t) is increased for a prolonged period of time (e.g., 5 minutes), the Fit4Work system proposes relaxation exercises to the user. There are two types of exercises: breathing exercises and muscle relaxation exercises. The proposed exercise depends on the time of the day and the previous choices of the user. E.g., breathing exercises are more suitable during the day and muscle relaxation exercises are more suitable in the evening. In addition, the system tracks the previously chosen exercises by the user in order to make personalized suggestions.

V. QUALITY OF THE ENVIRONEMNT MODULE

The quality of the environment module aims at monitoring the environmental parameters (temperature, humidity, CO_2, noise, light) at the workplace and recommend appropriate actions if needed (open/close windows, turn on/off humidifier, air conditioning etc.).

A. Quality of the environment data analysis

The data analysis module is composed of two components: the monitoring component and the ontology, as presented in Fig. 5. The task of the monitoring component is to retrieve data from hardware sensors and from virtual sensors. The hardware sensors in the NetAtmo indoor and outdoor modules sense the indoor (T_{in}) and outdoor temperature (T_{out}), indoor (H_{in}) and outdoor humidity (H_{out}), indoor concentration of CO_2 and noise. The virtual sensors use machine-learning models to estimate the parameters such as the state of the office windows and the number of occupants, which cannot be measured directly. The machine-learning models are trained on features extracted from data from the hardware sensors. All machine-learning models return their classification/regression result in 5 minute windows (with features considering historic data up to 40 minutes), which we consider one time step.

The virtual sensor for the estimation of the number of occupants utilizes two machine-learning models in two steps. First, a binary classification model is used to distinguish between an occupied and non-occupied office. If the classifier returns occupied, a regression model is used to estimates the number of occupants. While the relation between the occupancy and CO_2 concentration is pretty straightforward (more occupants → CO_2 concentration increases), the relation between the window state and the T_{in}, H_{in} and CO_2 parameters is more complex. In order to detect the window state, we use a single classification model trained to distinguish between open and closed windows. A window is considered open if it is detected as such in two consecutive time steps. The same approach is used to detect if the window is closed.

The task of the ontology component is to use the parameters sensed by the hardware and virtual sensors to infer actions that can improve the quality of the environment if required. We encoded the devices, actions, and domain expert knowledge in an ontology using the Web ontology language (OWL) using Protége [25]. The reasoning is done with the descriptive logic reasoner Pellet. The representation of the knowledge with an ontology enables us to adapt the ontology to the specific office without any additional software development due to the simplicity of adding and removing new devices, actions, parameters and relations.

B. Quality of the environment recommendations

The reasoning on the ontology provides a wide set of actions that can improve the quality of the environmental parameters. Each action can influence a single or multiple parameters, and it is not immediately obvious which action is best. Therefore we simulate and evaluate the effect of the entire set of actions. The simulator is composed of prediction models and the quality rating (Q-rating). The prediction models utilize machine-learning models trained to predict the environmental parameter values for T_{in}, H_{in} and CO_2 for all combinations of actions returned by the ontology. For each parameter we trained four prediction models to estimate its value in four time frames (15, 20, 25 and 30 minutes). The results of each action are evaluated with the Q-rating as good, medium or bad based on workplace regulations. The overall quality rating is composed of the ratings of the individual parameters, which we treat as equally important as follows. Good parameters are assigned the value of 1, medium parameters are assigned values between 0 and 1, and bad parameters are assigned values between 0 and –1, using linear interpolation both in the medium and bad range. The overall quality is the average of all three parameter values and can range from –1 to 1, which are scaled to the range from 0 to 1. The action resulting in the best Q-rating is recommended by the system. The reader is referred to [26] for more details. We also consider the Q-rating as the indicator of the current quality of the environment, which is presented to the user through the user interface.

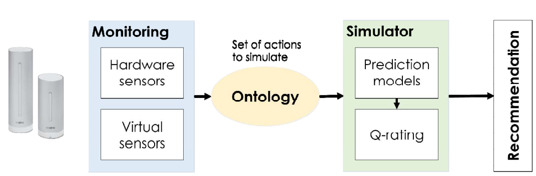

Fig. 5. Method for monitoring and management of the quality of the environment.

VI. User interface

The user interface is composed of the current status section, recommendation section, performance section and extended function section as seen on the left side of Fig. 7.

The current status section combines information related to the three areas of monitoring in one holistic view. Each area of monitoring is presented in a visual form (body – physical, head – mental and the circle – environmental state) with colors representing the current state (from green = good to red = bad). The value for each state represents the daily goal (Section III), relaxation score (Section IV) and Q-rating (Section V) respectively.

The Recommendations section presents the outcomes of the recommender system in an understandable way, using icons enabling easy interpretation (A in Fig. 7.). Tapping on each icon displays a short explanation of the recommendation (B in Fig. 7.). It is important to note that the recommendations are only displayed to the user and are available for them to see when they decide to look at the application, no sound or other measure is taken to alert the user to a new recommendation. The Performance section of the user interface presents the performance of the user versus the daily and weekly achievements (C in Fig. 7.). Extended functions section is the entry point for the functionalities such as functional fitness exercises (not discussed in this paper) and mental stress relief exercises (Section IV), both intended to help the user to stay in good health. It also

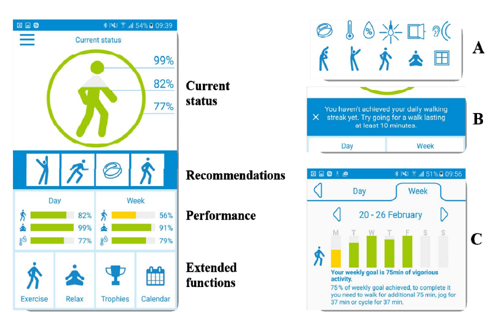

Fig. 4. Elements of the smartphone-based user interface of the Fit4Work system

contains a monthly calendar view of achievement so the user can self-evaluate his/her performance over a longer period.

VII. Evaluation and Results

A. Datasets

Three datasets were used for training and evaluation.

1) Physical activities dataset: To train and evaluate the machine-learning models of the physical activities module we utilized the dataset of ten healthy volunteers equipped with four smartphones (one in each trousers pocket, one in a jacket and one in a bag), a wristband that was used for heart-rate monitoring, and the Cosmed indirect calorimeter [27], which was used to label the real expended energy of the person. The volunteers were instructed to perform a predefined scenario (1 hour 45 minutes) containing rest activities, household chores and common types of exercises such as walking and cycling and running [28]. The dataset was collected in a laboratory at the Faculty of Sports, University of Ljubljana.

2) Mental stress dataset: To develop the stress-monitoring module, two datasets were recorded, one in a laboratory for training and evaluation of the laboratory stress detector, and one in real life for training and evaluation of the context-based stress detector. The laboratory data was collected with a standardized stress-inducing experiment [29] where we monitored 21 subjects while solving mathematical tasks under time and evaluation pressure. The baseline (no-stress) data was recorded on a separate day when subjects were relaxed. The subjects filled short STAI-Y questionnaires during the experiments, which were used to provide labels for the laboratory dataset. *The real-life dataset was collected* during everyday life. Five subjects were monitored using a wristband equipped with physiological sensors. The data was labeled with the combination of stress log and Ecological Momentary Assessment (EMA) prompts on the smartphone. The EMA prompts are questionnaires displayed at a random period of the day. The subjects had to answer 4-6 EMA prompts per day (with at least 2 hours between consecutive prompts), and in the case of a stressful situation, they logged the start, the duration and the level of stress on a scale from 1 to 5 (1 – no stress, 2 – low stress, 3 to 5 – high stress). Overall, the dataset contains 73k minutes of no-stress data, and 6.7k minutes of stress data per sensor.

3) Quality of the environment dataset: We equipped three offices – A (43 m2), B (27 m2), and C (20 m2) – for data collection and real-time validation of the recommendations. The offices were equipped with the NetAtmo indoor and outdoor modules to measure the environmental parameters, humidifier, window sensors to detect the window state (open, closed) and a smartphone application to self-report the occupancy and the state of other devices, and to receive the recommendations. We collected 67 days of data which was divided into three time periods: 1st period (2016-01-16 to 2016-02-26), when the occupants were allowed to manipulate room devices freely; 2nd period (2016-02-26 to 2016-03-23) when the office A was given the recommendation system, while the occupants of the other offices continued using the devices freely; and 3rd period (2016-03-23 to 2016-03-30), when the offices A and B were given the recommendation system, while the office C stayed as a control. The average number of occupants per office was 2.6 ± 1.5 (max 9) in A, 2.0 ± 0.9 (max 7) in B and 1.6 ± 0.9 (max 7) in C.

B. Evaluation of the physical activity module

We used the dataset presented in Section VII for the evaluation of the activity recognition (AR), in which we tried to recognize seven activities (lying, walking, running, standing, sitting, cycling, mixed), and the evaluation of the energy expenditure estimation (EEE). The results for the AR are presented in terms of the classification accuracy, and the results for the EEE in terms of the mean absolute error (MAE), where a smaller error means better performance. All the experiments were performed in the leave-one-subject-out (LOSO) manner, which means that we trained the models on data of nine people and tested on the remaining one (ten times, once for each person). The results are presented in Table I. The results for the EEE are also compared against the commercial device Bodymedia [30], one of the most accurate consumer devices for EEE. We can observe that we achieved an average accuracy of 84% for the AR and a better EEE compared to Bodymedia, except when the smartphone was carried in a bag. We assume that we outperformed Bodymedia because it is developed for sports activities and not everyday activities such as cleaning, cooking, etc.

C. Evaluation of the mental stress module

The results of the evaluation of the mental stress module are presented in Table II. It can be seen that the accuracy of the method for detecting stress events is 92%. In addition, the table provides performance comparison between stress detection with and without context. The no-context method is the laboratory stress detector applied directly on real-life data. For both experiments we used the LOSO evaluation. The context-based method performs significantly better than the no-context method. For example, the context-based classifier achieved a mean F-score of 80% (the mean value of the stress and no-stress F-scores), while the no-context classifier achieved a mean F-score of 62%. Additionally, the confusion matrix "with context" shows that the precision (73%) of the model is higher than the recall (55%) by 18 percentage points. This means that the model detects (recalls) 55% of the stress events with a precision of 73%.

D. Evaluation of the quality of the environment module

The evaluation was performed on the dataset presented in Section VII. We evaluated the performance of the developed machine-learning models (virtual sensors and prediction models) and the objective and subjective performance of the recommendations. The results of the developed machine-learning models are presented in Table III, where the window state is evaluated in terms of the classification accuracy (open/closed), and regression models in terms of the MAE and root mean squared error (RMSE). We can observe that we achieved 91% accuracy for the window state detection, and we miss-estimated the number of occupants by 0.6 person. The results for the prediction of the parameters shows that we mis-predicted the temperature by 0.4°C, humidity by 0.6% and CO_2 by 55 ppm on average. These errors are small enough that they are unlikely to result in the selection of grossly inappropriate actions, which is the final objective of the virtual sensors and prediction models.

We also evaluated the recommendations over the three periods. The results in terms of Q-rating, which can be interpreted as comfort, are presented in Table IV. We can observe that in first period all the offices had a comparable overall comfort. Office B was the best at keeping CO_2 at a good level due to frequently opening the windows, which resulted in a worse temperature. In the second period, the occupants of office A were using the recommendations and consequently the per-parameter quality and overall comfort increased, while the comfort of offices B and C stayed similar to the first period. In the third period, both offices A and B were using the recommendations and their comfort increased to the similar level, while the comfort of office C did not change significantly. These results prove that using the system objectively improves the comfort in the offices where it is used.

E. *Evaluation of the user interface*

The Fit4Work system was designed within an iterative process of usability evaluation where test groups were gathered in line with the suggestions of Nielsen and Landauer who argued that the best results of information system usability evaluation come from testing through running a series of small tests [31]. The working sessions were constructed according to the Concurrent Think-Aloud (CTA) methodology [32]. In a CTA session, the users were asked to use the system and to express their thoughts. The observations about the execution and stated problems are noted and taken into account before next iteration.

Our evaluation test was composed of 24 tasks in which the participants used functions of various sections and screens. The tasks included (i) understandability of the main screen, e.g., what is current state of the physical activity, mental stress, and environment, what recommendations do you see, etc., (ii) the interpretation of achievements, e.g., how well are you doing today/this week for each state, (iii) the flow of the user interface (navigate to today view, how to check yesterday achievements, how many relaxation exercises did you perform this week, etc.), and (iv) the use of extended functions (e.g., how would you start relaxation exercises, etc.).

At the end of each session we additionally used a Retrospective Probing technique to find out any more general issues with the user interface and the system as a whole [33]. The questions asked to the users included requesting them to suggest any modifications they would make to the system, listing three elements they liked and three they disliked, and provide any additional feedback. Using the methodology as described above we iteratively designed the user interface as presented in Fig. 7 to be used in long-term field trials.

TABLE I. RESULTS OF THE ACTIVITY RECOGNITION (CLASSIFICATION ACCURACY) AND ESTIMATION OF ENERGY EXPENDITURE (MEAN ABSOLUTE ERROR).

Module	**Wristband**	**Smartphone**			
		Trousers	***Jacket***	***Bags***	***Bodymedia***
AR [%]	88.4	87.9	78.9	79.5	/
EEE [MET]	0.87	0.87	0.92	1.15	1.0

TABLE II. CONFUSION MATRIX FOR THE MENTAL STRESS DETECTION MODULE.

	No Context		**With context**	
	No Stress	***Stress***	***No Stress***	***Stress***
No Stress	638	175	790	23
Stress	44	70	51	63
Recall (%)	78	61	97	55
Precision (%)	94	29	94	73
F1-score (%)	85	39	96	63
Accuracy (%)	76		92	

TABLE III. RESULTS OF THE EVALUATION OF VIRTUAL SENSORS (WINDOW STATE AND NUMBER OF OCCUPANTS) AND OF THE PREDICTION OF ENVIRONMENTAL PARAMETERS.

	ACC	***MAE***	***RMSE***
Window state [%]	91	✗	✗
No. of occupants	✗	0.6	1.2
Predict T_{in} [°C]	✗	0.4	0.5
Predict H_{in} [%]	✗	0.6	0.9
Predict CO_2 [ppm]	✗	55	104

TABLE IV. EVALUATION OF RECOMMENDATIONS. THE VALUES PRESENT COMFORT FROM 0 = BAD TO 1 = GOOD.

	Experiment period								
	1st			***2nd***			***3rd***		
Office	***A***	***B***	***C***	***A***	***B***	***C***	***A***	***B***	***C***
Recomm.	✗	✗	✗	✓	✗	✗	✓	✓	✗
Temp.	.65	.59	.81	.76	.48	.73	.97	.95	.93
Hum.	.29	.35	.38	.50	.41	.29	.29	.35	.26
CO2	.83	.94	.72	.93	.93	.78	.92	.91	.72
Comfort	**.59**	**.62**	**.64**	**.73**	**.61**	**.60**	**.73**	**.74**	**.64**

VIII. CONCLUSION

We presented a system for the management of physical, mental and environmental stress. The system collects data with wearable and environmental sensors, interprets it and provides recommendations to the user through a smartphone application. The system is composed of three intelligent modules: physical activity monitoring, mental stress monitoring and quality of the environment monitoring module. Each module uses machine learning to interpret the sensor data for its respective task, and expert knowledge to provide recommendations with the goal to improve user's physical, mental and environmental well-being. The evaluation show that the system can adequately detect sedentary, stressful and unhealthy environment and provide recommendations to the user.

The presented system was developed as part of the Fit4Work project [16], which aims to develop an affordable system that can be used in any environment without the need to renovate or automate the building. In the next phase of the project, the developed system will be validated in pilots with end-users. This will provide an opportunity to evaluate the performance of the algorithms in real life, as well as collect long-term user feedback of recommended actions, and acceptability and understandability of the user interface.

Future work includes: improved physical activity monitoring module to utilize data from both devices simultaneously; personalized mental stress monitoring and personalized stress-relief recommendations; enhanced quality of the environment monitoring module with additional machine-learning models to detect current state of other devices such as air conditioning, humidifier, light, etc.; and improved user interface based on real-life user feedback received through a long-term user experience assessment.

ACKNOWLEDGMENT

The research was done in the Fit4Work project co-financed by the AAL Program (AAL-2013-6-060).

REFERENCES

[1] WHO, http://www.who.int/

[2] Work-related stress, http://www.eurofound.europa.eu/

[3] H. W. Kohl, C. L. Craig, E. V. Lambert, S. Inoue, J. R. Alkandari, G. Leetongin, S. Kahlmeier, “The pandemic of physical inactivity: global action for public health,” The Lancet, vol. 380, no. 9838, pp. 294–305, 2012.

[4] L. Lan, P. Wargocki, Z. Lian, “Optimal thermal environment improves performance of office work,” REHVA European HVAC Journal, pp. 12–17, January 2012.

[5] FitBit, https://www.fitbit.com

[6] Garmin, https://www.garmin.com

[7] Apple Watch, http://www.apple.com/watch/

[8] Feel, http://www.myfeel.co/

[9] Thync, http://www.thync.com/

[10] J. A. Healey, R. W. Picard, “Detecting Stress During Real-World Driving Tasks Using Physiological Sensors,” IEEE Trans. Intell. Transp. Syst., vol. 6, no. 2, pp. 156–166, 2005.

[11] K. Hovsepian, M. Absi, T., M. Nakajima, “cStress : Towards a Gold Standard for Continuous Stress Assessment in the Mobile Environment,” In the Proceedings of the 2015 ACM UbiComp, pp. 493-504, 2015.

[12] P. Kumar, C. Martani, L. Morawska, L. Norford, R. Choudhary, M. Bell, M. Leach, “Indoor air quality and energy management through real-time sensing in commercial buildings,” Energy and Buildings 111, 145–153, 2016.

[13] A. B. Konuah, Occupant window opening behaviour: the relative importance of temperature and carbon dioxide in university office buildings, Ph.D. Dissertation. University of Sheffield, 2015.

[14] L. M. Candanedo, V. Feldheim, “Accurate occupancy detection of an office room from light, temperature, humidity and CO 2 measurements using statistical learning models,” Energy and Buildings 112, 28–39, 2016.

[15] D. Kolokotsa, A. Pouliezos, G. Stavrakakis, C. Lazos, “ Predictive control techniques for energy and indoor environmental quality management in buildings,” Building and Environment 44, 9, 1850–1863, 2009.

[16] Fit4Work, http://www.fit4work-aal.eu/

[17] Microsoft Band 2, https://www.microsoft.com/

[18] NetAtmo, https://www.netatmo.com/

[19] Smartphone statistics, https://www.statista.com/

[20] H. Jimison, P. Gorman, S. Woods, P. Nygren, M. Walker, S. Norris, W. Hersh, “Barriers and Drivers of Health Information Technology Use for the Elderly, Chronically Ill, and Underserved,” In Evidence Reports/Technology Assessments, No. 175. AHRQ Publication, pp.1-63, 2008.

[21] T. Heart, E. Kalderon, “Older adults: Are they ready to adopt health-related ICT?,” In International Journal of Medical Informatics 82, pp. 209-231, 2013.

[22] M. D. Tundo, E. Lemaire and N. Baddour, “Correcting Smartphone orientation for accelerometer-based analysis,” In Proc. IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA), pp. 58-62, May 2013.

[23] B. Cvetković, V. Janko, M. Luštrek, “Demo abstract: Activity recognition and human energy expenditure estimation with a smartphone,” PerCom 2015, pp. 193-195, 2015.

[24] M. Gjoreski, H. Gjoreski, M. Luštrek, M. Gams, “Continuous stress detection using a wrist device: in laboratory and real life,” In Proceedings of the 2016 ACM UbiComp '16, 2016.

[25] H. Knublauch, R. W. Fergerson, N. F. Noy, M. A. Musen, “The Protégé OWL plugin: an open development environment for Semantic Web applications,” In Third International Conference on the Semantic Web, Hiroshima, Japan, pp. 229—243, 2004.

[26] M. Frešer, B. Cvetković, A. Gradišek, and M. Luštrek, ”Anticipatory system for T--H--C dynamics in room with real and virtual sensors,” In Proceedings of the 2016 ACM UbiComp '16, 2016

[27] Cosmed, http://www.cosmed.com/

[28] B. Cvetković, R. Milić and M. Luštrek, "Estimating Energy Expenditure With Multiple Models Using Different Wearable Sensors," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 4, pp. 1081-1087, July 2016.

[29] K. Dedovic, R. Renwick, N. K. Mahani, and V. Engert, “The Montreal Imaging Stress Task : using functional imaging to investigate the ...,” vol. 30, no. 5, pp. 319–325, 2005.

[30] Bodymedia, https://templehealthcare.wordpress.com/s

[31] J. Nielsen, and T.K. Landauer, "A mathematical model of the finding of usability problems," Proceedings of ACM INTERCHI'93 Conference (Amsterdam, The Netherlands, 24-29 April 1993), pp. 206-213

[32] E. L. Olmsted-Hawala, J. Romano Bergstorm, “Think-Aloud Protocols: Does Age Make a Difference?”, Proceedings of Society for Technical Communication Conference, Chicago, pp. 86-94, January 2012.

[33] J. H. Birns , K. A. Joffre , J. F. Leclerc , C. Andrews Paulsen, “Getting the Whole Picture: Collecting Usability Data Using Two Methods -- Concurrent Think Aloud and Retrospective Probing”, Presented at Eleventh Annual Meeting of the Usability Professionals’ Association, 2002.