Designing Realistic Mental Health Monitoring and Support System

Elena Vildjiounaite  Johanna Kallio  Satu-Marja Mäkelä  Marja Harjumaa
VTT Technical Research Centre of Finland
Kaitoväylä 1, 90570 Oulu
firstname.lastname@vtt.fi

ABSTRACT
This paper argues that it is not sufficient to provide mental support on just individual level (e.g., by offering mindfulness courses) and proposes to design mHealth applications at individual and group levels simultaneously. This paper also argues that design of mHealth applications should consider interactions with both parts of these applications: mental health monitoring and mental support. Recently developed environmental and wearable sensors and algorithms to analyse sensor data, on the one hand, enable thorough self-monitoring of mental conditions, but on the other hand, they determine interaction requirements. Interaction requirements should depend on two aspects of data analysis algorithms: first, these algorithms usually require input from end users to start working, and efforts to provide this input is an important design choice. Second, ways to present data analysis results to the user should depend on accuracy and granularity of these results. This paper also argues that it would be easier to understand users’ motivations and needs with a working monitoring system at hand, and suggests two unobtrusive stress assessment approaches to facilitate such studies.

Author Keywords
Design of mental health support; Stress monitoring; Sensor-based stress detection

ACM Classification Keywords
Ubiquitous and mobile computing; Artificial intelligence

INTRODUCTION
Work-related accidents and illnesses cost globally €2680 and the EU €476 billion every year [1], and in many countries the main reason for sick leaves is stress, depression, anxiety and other mental issues. In the UK, 1.3 million workers suffered from work-related health problems in 2016/2017, and the most common reason for these problems was stress, depression and anxiety (40%) [2]. In Finland 152,900 people received disability pension in 2016, and mental/behavioural disorders constituted the largest disability group (42%) [3].

Unfortunately, this problem cannot be solved on individual level. A recent review of interventions to promote work participation of older workers concluded that only combination of health services and work modifications improves work participation; there is not enough evidence to recommend health services alone [4]. A large international network of consulting and other services proposed that mental health problems should be tackled at several levels: society, employers and employees [5]:

- employees should be actively engaged in their own health and wellbeing (e.g., participate in dedicated programmes) and should support their colleagues;
- employers should proactively manage mental wellbeing of the employees;
- society and the state should encourage employers to take charge of employees’ mental health and promote actions that improve mental wellbeing.

The current mHealth monitoring solutions, however, only recognise and/or treat mental problems at individual level. Recent advances in wireless sensors allow to obtain various data about users and their environments, not available earlier, and recent advances in data platforms enable integration of data from different sources and sending results to the users. Thus, it is possible to integrate near real-time data from large groups of people - to provide an overview at group level.

Therefore, we propose to design mHealth applications at both individual and group level and discuss related research challenges. Furthermore, we suggest two unobtrusive approaches to detect stress of an individual, to facilitate studies into designing interactions with stress coping systems (we focus on stress detection because stress is an unhealthy mental state and a major risk factor for long-term mental illnesses [6]).

We suggest these approaches because our experiences in developing supportive systems in health and wellbeing domain convinced us that presenting a working system to end users allows to understand their needs better than mock-ups. This is especially the case with systems that monitor groups of persons and report to their supervisors, because the latter are usually overwhelmed with work and tend to avoid extra mental efforts like analysing mock-ups. For example, when we studied with mock-ups how to present information regarding conditions of multiple elderly persons to a caregiver of these persons, the caregivers declared that simple green (“no problems”) vs. red (“requires attention”) icon against each person name would suffice. However, when we installed monitoring system and presented its results in more details (temporal trends, current and recent physical activities etc.), the caregivers were surprised that the system output matches their observations, and this surprise (“oh, it really works!”) stimulated the caregivers to discuss possible use cases and desired interfaces in more detail. Thus we believe that developers of mental support systems would be able to better understand users’ needs and motivations if
they could experiment with realistic (i.e., convenient for the users) mental health assessment system.

**REALISTIC MENTAL HEALTH SUPPORT APPROACH**

A structure of the proposed mental health support system is presented in Figure 1. Below we discuss design of the three first layers because society and state layer is not a part of a personal mHealth application.

![Figure 1. Design of mental health support system. Dark blocks present aspects, rarely addressed by existing systems.](image)

**Mental health assessment layer**

Design of a mental health application starts from the choice of mental health assessment technology. Proactive mental health support requires a realistic monitoring system, to act e.g. when a person is experiencing long-term stress or when his/ her depression worsened. It is also important to detect individuals in higher risk of developing mental issues, e.g., highly stress-prone ones. Here, by “realistic monitoring” we mean (1) unobtrusive and (2) reasonably accurate. Hence the main trade-off here is a trade-off between accuracy of mental health assessment and system obtrusiveness/ privacy threatening (including user effort to launch the system). Both accuracy and obtrusiveness of the system depends on choice of sensors and on choice of algorithms to analyse sensor data.

Existing approaches to monitoring of stress, mood and depression fall into three major groups [7]: (1) relying on wearable physiological sensors; (2) obtaining physiological data via video cameras and (3) analysing behavioural data, obtained from mobile phones or environmental sensors.

Physiological data allows most accurate stress assessment in lab studies, but unfortunately, existing wearable devices with physiological sensors provoke too much discomfort for everyday use [7] because they require tight attachment and frequent charging, and they are too expensive for screening everybody to detect high-risk individuals. Hence launching of such systems is expensive, and maintenance requires a lot of user efforts. Furthermore, we are not aware of any long-term real-life studies using physiological sensors for mental health assessment, which probably means that data quality in real life is not as high as in lab studies. Physiological parameters can be obtained also by monitoring of faces via video cameras and hyperspectral cameras, but these approaches work only if the monitored person is facing the camera, and not everybody would like to have cameras constantly analysing their faces. Such systems do not require maintenance efforts from end users, but they require to install one camera per user and are not so well accepted because of potential privacy threats. To the best of our knowledge, these systems were not tested in long-term real life use either.

Thus we believe that analysis of behavioural data is the most realistic approach for long-term real life applications because behavioural data can be acquired unobtrusively via mobile phones (e.g., usage of phone applications, location data etc.) and environmental sensors (e.g., infrared sensors, computer mice, pressure-sensitive chairs etc.). Our belief is supported by multiple existing works, using these data sources for real life assessment of stress and other mental health issues [7, 8].

Such systems require little installation and maintenance efforts from the end users because mobile phones are widespread nowadays, whereas environmental sensors in workplaces are installed and maintained by the employers. Such systems may require notable launching efforts from the end users, though: it depends on the choice of mental health assessment algorithms. As behavioural patterns strongly depend on individual (what is normal for one user, can be abnormal for another one), mental health assessment on the basis of behavioural data is usually performed in person-specific ways, i.e., the system learns to distinguish between normal and mentally unhealthy (stressed, depressed etc.) conditions of each user separately. Unfortunately, the majority of existing methods to learn the difference between healthy and unhealthy behaviours employ supervised machine learning approaches, i.e., these methods require data labels. The labels are usually acquired by asking each end user to provide reports of the kind “now I am OK” vs. “not OK” - typically, 100 or more self-reports are required from each person. This is a notable effort, and human beings tend to avoid it. One study into stress detection reported that on average end users submitted less than one third of reports they were asked to submit [9]. Thus with this approach either the system will not get enough training data and hence will be working inaccurately, or collection of required number of training data samples will take long time and annoy the users, so that they will abandon the system altogether.

Thus we believe that a realistic mental health management application should employ unsupervised or partially supervised learning methods: the latter usually achieve higher accuracies than the former, but at the cost of requiring certain (relatively small) number of self-reports from each end user. Hence the main design trade-off here is a trade-off between accuracy and granularity of mental health assessment and efforts of end users, required to achieve this accuracy. For example, various works attempted at stress detection of the following granularity: (1) detection of stress occurrences (in the morning vs. in the afternoon or on which
day [7]); (2) detection of prolonged stress periods [9]; (3) classification of individuals as high-risk vs. low-risk [8, 9].

**Existing unobtrusive stress detection systems**

Surprisingly, not so many works studied methods to develop realistically unobtrusive mental health assessment systems. One unobtrusive system [10] detects stress of each user via analysis of his/ her mobile phone usage patterns. It evaluates each day as stressful vs. normal, and in the tests on the dataset, containing 200 days of two test subjects, it achieved 72% accuracy, which is comparable with accuracies of more obtrusive systems, reported by other studies. However, not everybody agrees to install an application, collecting his/ her phone data - because of privacy concerns and because phone battery drains quicker when data collection is activated.

Another unobtrusive system [9] employs environmental sensors (depth cameras) to detect stress of each office worker. Launching and maintenance of this system does not require any effort from end users. As this system does not recognise individuals and logs only anonymous motion trajectories in the offices, it is not perceived as privacy-threatening, and so far we have not heard any objections from potential test subjects against installation of this system in their offices. This system is so unobtrusive that it allowed us to conduct 10-month long trial, which is, to the best of our knowledge, the longest stress assessment study, conducted so far (duration of other studies ranged from a few days up to two months, and only one study lasted four months). However, greater unobtrusiveness in this case implies lower accuracy: this system was not able (not yet, at least) to classify every day as stressful vs. normal; instead, it detected long-lasting stress periods and classified individuals into stress-prone vs. stress-resistant classes. Figure 2 presents results of truly unobtrusive detection of stressful periods. It shows that the system missed only one period (of Subject 1 after day 40; it happened because this person spent very little time in the office during this period) and estimated condition of Subject 3 as stressful for too long time (period around day 60). However, the system clearly separated more stress-prone Subjects 3 and 4 from less stress-prone Subjects 1 and 2, which is good enough result to take preventive measures.

![Figure 2. Unobtrusive detection of stress periods.](image)

**Individual awareness and control layer**

As discussed above, unobtrusive monitoring is not so accurate. Obtrusive approaches are not so accurate either because physical activities notably affect physiological data (skin conductivity, breathing rate, heart rate, etc.). For example, in [8] each user was wearing two (!) wrist devices for one month, and nevertheless collected data did not allow to classify every day as stressful vs. normal; it only allowed to classify subjects into two classes: highly stress-prone vs. not-so-stress-prone. As technologies will never be perfect, realistically “just-in-time-mental-support” means support for relatively long-lasting problems, such as prolonged stress or depression. Users with problems can be offered personalised help, while some other persons can be selected as potential volunteers to help their colleagues (if they have time and willingness to help). Potential supporters can be, for example, subjects with similar problems in the past.

Personalised support can be provided in a form of individual programmes, such as mindfulness exercises or online cognitive behaviour therapy, and/or in a form of group programmes, e.g., as an advice to talk to a relative/ colleague/ superior or to attend a course. The main challenge is to adapt form of support to needs and peculiarities of each person, because personalisation of content presentation and persuasive strategies result in higher adherence to therapeutic programmes and in longer lasting effects, especially in preventive mental health interventions [11].

It is also important to adapt granularity of presenting mental health assessment results to accuracy and confidence of results. For example, if the system is not highly confident in detecting certain stress period, it might be more beneficial (for gaining user trust) to ask the user “did you have tough time at work recently?” instead of showing exact days, classified by monitoring system as “bad days”. It might be also good to present temporal trends, to show the user that the proposed supporting methods indeed work; or, if they do not work, to suggest the user something else. It might be also very useful for the users to relate detected “bad days” or “bad periods” to user contexts, e.g., calendar or location data.

The main design challenge for individual control over data sharing is a trade-off between privacy and usefulness of data. People often find it difficult to discuss own stress and other mental health issues, e.g., one survey reported that 95% of people, who took time off due to workplace stress, were not able to give their employer the real reason [5]. An easy way to protect user privacy is anonymous data aggregation, but better support could be provided if the user would allow selected colleagues and/ or supervisors to know that he/ she has problems now (or has free time to help others). It will be interesting to study, how users would prefer to share such information: for example, they could inform everybody concerned about approaching deadlines, or send less detailed data like “don’t disturb now”, or request brainstorming. Potential volunteers may instead specify available times or own competences. In any case, such data sharing is more
likely to gain user acceptance if it is performed only when the user permits it, but then another trade-off should be considered: requesting permission each time would increase the burden of already overwhelmed persons, whereas setting data sharing preferences “once and forever” may result in sharing too much or too little data. In addition, some persons do not want to admit that they need help; hence, also control part of the interface may require persuasive messages.

**Organisational awareness and control layer**

Assessment of organisational units can be done in many ways: evaluating short-term teams (e.g., project team), long-term teams, locations (e.g., noise in some part of the building may cause troubles, or persons in a remote location may feel neglected) etc. It is worth noting that although we call this layer “organisational layer”, at least part of this layer should be present also in a personal mHealth application: for example, individuals may want to know situation of their unit and to compare it with other units.

This layer requires developing of algorithms to aggregate and visualise data of multiple persons, and we are not aware of any works, proposing them. For example, data visualisation can be performed in ways, similar to the ways to aggregate preferences of multiple users in recommender applications [12]. One way is to average states of all team members and to show it as a shade of green or red, depending on the result. Another way is to present separately (1) percentage of “best” (most mentally healthy) persons; (2) percentage of “worst” cases and (3) an average state of others. Presentation of temporal trends can be less detailed than presentation of a current unit condition. It would be interesting to study, which way is most useful and better accepted in different kinds of organisations and for different types of teams (e.g., whether data of short-term teams and location-based teams should be presented differently or not).

**CONCLUSION**

This paper proposes to advance monitoring and support of mental health from individual level to group (employer) level. Moreover, we suggest that design of mHealth applications should simultaneously consider both their sides (mental health assessment and reactions at the assessment results) because choice of health assessment methods determines interaction requirements. On the one hand, end users often need to submit a lot of data to launch mHealth applications, and such efforts might easily avert the users. On the other hand, monitoring technologies are not perfect nowadays, and are not likely to be perfect ever. Hence these imperfections should affect ways to present mental health assessment results to the users: for example, it does not make sense to present inaccurate results in high granularity. As we believe that experimenting with working monitoring system should facilitate studying of users’ motivations and needs for mental support, we also propose two realistic stress monitoring systems that require very little or no launching efforts from end users, to be used in such studies.

**ACKNOWLEDGMENTS**

The authors express their gratitude to VTT for providing financial support for this work.

**REFERENCES**