A Novel Framework for the Holistic Monitoring and Analysis of Human Behaviour [†]

Kostas Konsolakis [1,*], Hermie Hermens [1,2] and Oresti Banos [1,3]

[1] Telemedicine Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 7522 NB Enschede , The Netherlands; h.j.hermens@utwente.nl (H.H.); oresti@ugr.es (O.B.)

[2] Roessingh Research and Development, 7522 AH Enschede, The Netherlands;

[3] Research Center for Information and Communication Technologies, University of Granada, E-18071 Granada, Spain

* Correspondence: k.konsolakis@utwente.nl

† Presented at the 13th International Conference on Ubiquitous Computing and Ambient Intelligence UCAmI 2019, Toledo, Spain, 2–5 December 2019.

Abstract: Recent technological advances have enabled the continuous and unobtrusive monitoring of human behaviour. However, most of the existing studies focus on detecting human behaviour under the limitation of one behavioural aspect, such as physical behaviour and not addressing human behaviour in a broad sense. For this reason, we propose a novel framework that will serve as the principal generator of knowledge on the user's behaviour. The proposed framework moves beyond the current trends in automatic behaviour analysis by detecting and inferring human behaviour automatically, based on multimodal sensor data. In particular, the framework analyses human behaviour in a holistic approach, focusing on different behavioural aspects at the same time; namely physical, social, emotional and cognitive behaviour. Furthermore, the suggested framework investigates user's behaviour over different periods, introducing the concept of short-term and long-term behaviours and how these change over time.

1. Introduction

Human behaviour analysis through passive sensing systems has gained significant attention in healthcare. Wearable technology with embedded sensor systems, such as smartphones and wrist-worn devices (e.g., Fitbit, smartwatch, etc.), has been increasingly used in healthcare and behavioural research in order to collect health data, monitor patient's vital signs and deliver comprehensive healthcare information to practitioners, researchers and patients.

Having a clear understanding of the user's behaviour is essential to characterize patient progress, make treatment decisions and elicit effective and relevant coaching actions. For this reason, there have been many ongoing studies focusing on the sensing and quantification of human behaviour. Despite the significant progress in this field, the vast majority of existing works infer human behaviour under the limitation of detecting a specific domain only [1]. For instance, they focus on detecting physical activity-related parameters, such as the steps count, calories burned and indoor mobility and not on detecting physical, social, cognitive and emotional behaviour at the same time.

In light of these limitations, some studies have moved towards more comprehensive and cross-domain monitoring of human behaviour. Gaggioli et al. [2] presented a platform to gather users'

psychological, physiological and activity information for describing mental health. Banos et al. [3] proposed a machine-learning-based framework to recognize behaviour elements, such as physical activities, emotions and locations, which are further combined through ontological methods into more abstract representations of user context. Furthermore, there is an increasing number of commercial frameworks, such as Apple HealthKit, GoogleFit and Samsung SAMI, that use different types of behaviour-related contextual information, while they rely on third-party applications and systems for inferring the behaviour information [3]. However, most of the existing solutions are domain-specific with a narrow application scope.

This work describes an innovative system to detect and infer human behaviour automatically and in a holistic manner, through the continuous and unobtrusive collection of multimodal sensor data. The proposed framework moves beyond the current trends in automatic behaviour analysis, by introducing the concept of using heterogenous sensor data in order to analyse different behavioural aspects, namely physical, social, emotional and cognitive, and in a holistic approach. Furthermore, the suggested framework serves as the principal generator of knowledge on the user's behaviour over different periods, introducing the concept of short-term and long-term behaviours and how these change over time. Our main goal is to develop a multimodal and unobtrusive system that can get insights into the user's context and provide the relevant background information for the development of healthcare applications.

The succeeding sections are structured in the following way. Section 2 introduces our innovative approach for detecting human behaviour and explains the aim of our research. Section 3 presents the architecture and Section 4 the implementation of the platform. Finally, we outline our conclusions and remarks in Section 5.

## 2. Novel Framework for Detecting Human Behaviour

The current work is a part of the 'Council of Coaches' project, which introduces a radically new concept of virtual coaching. This project consists of a number of virtual characters, each one specialised in their own domain, that interact with each other and with the user in order to inform, motivate and discuss about health and well-being related issues (e.g., physical, cognitive, mental and social well-being). The individual coaches will listen, ask questions, inform, discuss between themselves, jointly set personal goals and inspire the users to take control of their health. Any combination of specialised council members collaboratively covers a broad spectrum of lifestyle interventions, with a substantial probability of positive impact on outcomes related to chronic diseases and healthy ageing [4].

Three of the main innovative tasks of the 'Council of Coaches' project are related to sensing and profiling, dialogue management and user interaction. The suggested framework focuses on sensing and profiling user's context by combining smart multimodal sensing technologies in order to seamlessly and opportunistically measure and model the user behaviour in a comprehensive way. In particular, the framework will serve as the principal generator of knowledge on the user's behaviours over different periods and in a holistic approach, through modelling, inferring and combining human behaviour, including physical, emotional, cognitive and social aspects. This holistic sensing and modelling approach not only aims at registering, analysing and inferring each determinant of behaviour in a user-centric manner but also mining the interactions among users and with their physical and virtual environment. Consequently, the proposed framework will provide the necessary information for improving the user-coach interaction and supporting the dialogue and argumentation process with objective and continuous behaviour-related contents.

3. Architecture

The suggested framework consists of four main layers; the *Multimodal Sensor Data*, the *Sensor Platform*, the *Behaviour Analysis* and the *Service API* layer. The high-level architecture of the proposed platform is depicted in Figure 1. The *Multimodal Sensor Data* layer focuses on receiving different types of

data from on-body and off-body devices. On-body devices are considered devices that are light-weight, unobtrusive and can be carried by the user throughout a day. These devices can be smartphones, smartwatches, activity trackers and other commercially available wearable devices. On the other hand, off-body devices include all the smart devices that can be used in the user's environment, such as weight scales, TV, webcams, voice assistants and so forth.
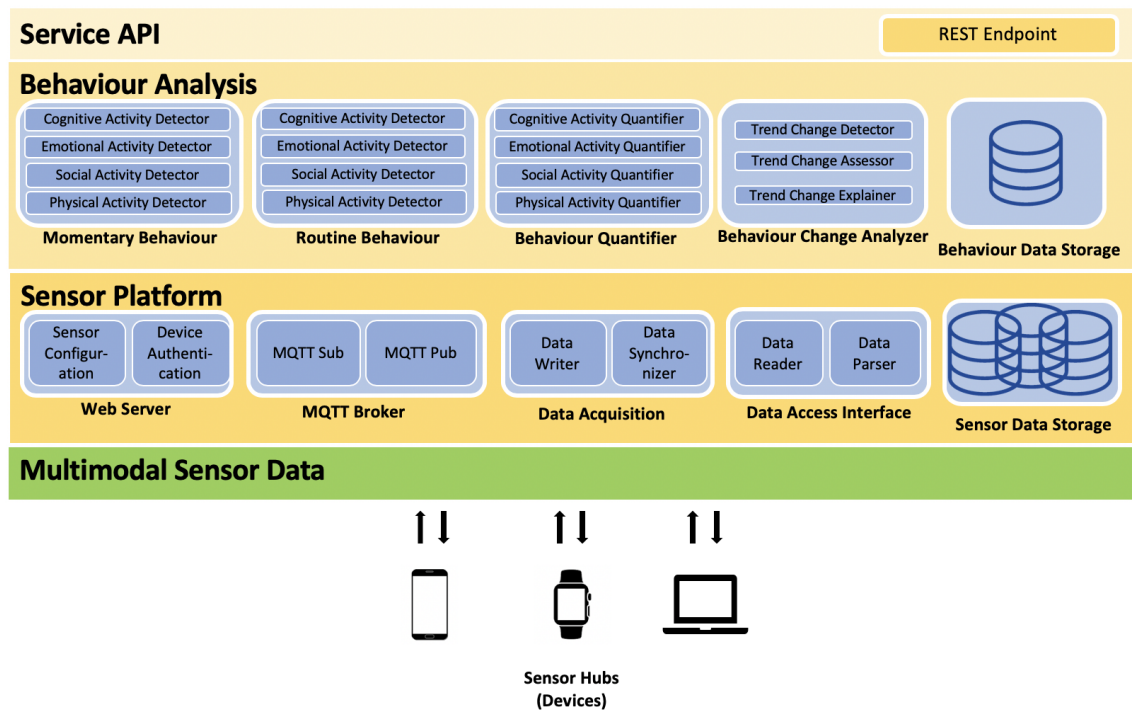
Figure 1. Overview of the architecture; a proposed framework for human behaviour analysis though multimodal on-body and off-body sensor data.

The *Sensor Platform* layer aims to synchronize and connect the different devices that a user can have, giving access only to authorized devices and users to upload data through the *MQTT Broker* component, which is based on the Message Queuing Telemetry Transport communication protocol. In particular, the *Web Server* component allows a user to get access and link the devices to the system, where a unique and encrypted identification code is generated for both the user and the device. The *Data Acquisition* component receives packets of multimodal data that consist of raw values and metadata, such as data modality or type, timestamps, user ID and device ID. Furthermore, the collected data are stored in the *Sensor Data Storage* and are accessible through the *Data Access Interface* component.

The *Behaviour Analysis* layer is the core entity of the suggested framework, where raw sensor data are converted to meaningful representations of human behaviour, including physical, social, emotional and cognitive aspects. This layer consists of the *Momentary Behaviour*, *Routine Behaviour*, *Behaviour Quantifier* and *Behaviour Change Analyzer* components, which generate information that is stored in the *Behaviour Data Storage* component. In particular, the *Momentary Behaviour* component processes raw data in order to extract relevant features that can be used to detect physical, social, emotional and cognitive short-term behaviours. Short-term or momentary behaviours are behaviours that last less than a certain period of time (e.g., minutes, hours). For instance, the number of steps (physical behaviour), the level of interaction/isolation of the user with other users (social behaviour), the level of engagement loss (cognitive behaviour) and the mood (emotional behaviour), all per minute. Thus, this component, combined with the available multimodal sensor data, can be used to obtain information about user's lifestyles with respect to user's mobility patterns (i.e., using accelerometer and GPS data to track user's location), the sociability (i.e., using social media apps, calls and text messages to analyse the interaction with other users), the cognitive behaviour (i.e., using self-reported

questionnaires to monitor user's skills in different fields such as occupational and academic) and the emotional behaviour (i.e., using physiological signals, audio content and device usage patterns to investigate user's mood).

The *Routine Behaviour* component makes use of the *Momentary Behaviour* component in order to detect long-term behaviours. Long-term behaviours or routines refer to physical, social, emotional or cognitive short-term behaviours that take place over long periods of time (e.g., days, weeks, months). In contrast to the short-term behaviours, long-term behaviours represent more meaningful and long-lasting activities that can be used to derive information about users' lifestyles. Depending on the behaviour type and the interpretation of behaviour patterns, the period of time can range from days, weeks, months or even years. For instance, long-term physical behaviours contain information such as the amount of sedentariness during the last month. Long-term behaviours can uncover certain markers that characterize a person's or a group's behaviour over time and identify different types of lifestyles [5]. Additionally, long-term behaviours can reveal trends or patterns that restrain users from following a normative or healthy lifestyle.

The *Behaviour Quantifier* component contains statistical features that are extracted from the physical, social, emotional and cognitive short-term and long-term behaviours and can be used to identify potential descriptors of user's behaviour through time series analysis. For instance, the average number of steps per day, the average duration of sitting during the weekends and so forth.

The *Behaviour Change Analyzer* component processes the long-term behaviours in order to detect any changes that occur over different periods of time (e.g., seasons, holidays, etc.). For instance, this component could reveal information about a set of patterns that can help in distinguishing a 'working lifestyle' from a 'lifestyle after retirement' (modelling between-person variations) but could also characterize subjective behavioural patterns that force individuals to have a sedentary lifestyle on specific weekdays (modelling within-person variations).

Finally, the *Service API* layer gives access and pushes the processed information that is stored in the framework to the 'Council of Coaches' system or to any potential authorized third party application for further analysis.

4. Implementation

This section describes the framework components that have been implemented so far for the modelling of short-term and long-term behaviours using multimodal data from on-body and off-body devices. Furthermore, the software libraries and applications that are used in the framework will be mentioned.

4.1. Short-Term Behaviours Inference

The inference of short-term behaviour is based on acquired data from on-body and off-body sensors. On-body sensors refer here to both hardware and software sensors readily available in most commercial smartphones (using the AWARE app [6]) in order to analyse short-term physical and social behaviour. In particular, a model for detecting steps count, predicting the performed activities (walking, tilting, sitting, cycling and taking the bus) and detecting the level of user's social interaction and isolation has been developed. Smartphone data consist of accelerometer and GPS signals for tracking physical activity, while Bluetooth, ambient noise, location and phone usage metrics (e.g., number of incoming/outgoing calls, number of received/sent text messages) have been used to detect the user's level of being socially active. The aforementioned types of data are presented in the following table (see Table 1).

Off-body sensors refer to acquired data through devices that are not placed directly on the user's body. For instance, data collected from questionnaires and surveys or data from the laptop webcam or Kinect sensor. This type of sensor data is used to analyse short-term emotional and cognitive behaviour.

Table 1. Overview of on-body sensor data.

Accelerometer	The accelerometer sensor measures acceleration over time along three axes; axis x (double_values_0), axis y (double_values_1) and axis z (double_values_2).
Bluetooth	The Bluetooth sensor detects surrounding Bluetooth-enabled and visible devices with respective RSSI values at specified intervals (default is 60 seconds). The signal gets stronger when the RSSI values are closer to zero. RSSI is expressed in decibels relative to a milliwatt (dBm) from zero to -120 dBm.
Calls	The calls sensor logs incoming and outgoing call events, performed by or received by the user. This sensor does not record personal information, such as phone numbers or contact information, but uses a unique ID based on SHA-1 encryption.
Messages	The messages sensor logs received and sent SMS texts, performed by or received by the user. This sensor does not record personal information, such as phone numbers or contact information, but uses a unique ID based on SHA-1 encryption.
Locations	The locations sensor uses the GPS coordinates in order to provide an estimation of the user's current location, automatically.
Activity Recognition API	The Activity Recognition plugin uses the Google Location API, which is based on combining accelerometer and GPS data, in order to detect user's movement and the mode of transportation. The plugin can monitor the following activities: walking, cycling, using a vehicle (e.g., taking the bus), tilting and being still.
Ambient Noise	The ambient noise plugin measures the ambient noise based on the double_frequency (Hz), the double_decibels (dB) and the double_rms (Hz). An environment is considered noisy when the perceived loudness gets louder and above a certain decibel or RMS sound value. The plugin uses audio from the microphone sensor and detects if the environment is noisy or silent for the last five minutes.

4.1.1. Physical Short-Term Behaviours

The model for detecting short-term physical behaviour has been divided into two parts. For the first part, we focus on detecting steps based on accelerometer data, while for the second part we use the Google Recognition API [7] for detecting everyday activities (walking, being in a vehicle, biking, tilting or remaining still) based on accelerometer and GPS data.

For the steps counter model, we calculate the magnitude value of 3-axial accelerometer raw data (50Hz sampling frequency) in order to make the model orientation independent. Then, we apply a lowpass filter with a cutoff frequency equal to 3.667 Hz. The filter passes signals lower than the cutoff frequency and attenuates signals higher than the cutoff frequency. After removing short-term fluctuations, we count the peaks of the magnitude signal above a certain threshold, where each peak equals to one step. This threshold is equal to 12 $(m/s)^2$ and has been selected among other values as the most optimal one for this steps counter model. It is also worth mentioning that an additional or alternative way to monitor user's steps is to use a wrist-worn activity tracker, such as Fitbit.

4.1.2. Social Short-Term Behaviours

The model for social behaviour detects if the user interacts with other individuals, trying to answer if the user has a socially active or isolated life. For this reason, different types of smartphone data are processed in order to estimate the level of social activity. The signal strength of Bluetooth visible devices is translated into social interaction. The model estimates that if a user is surrounded closely by Bluetooth devices, there might be individuals to interact with. The condition for that is the RSSI (Received Signal Strength Indicator) values to range from -70 to 0 dBm. Furthermore, the number of performed incoming and outgoing phone calls is also used to measure social interaction. The model checks every minute if at least one call is performed for more than 10 seconds (social interaction). Similarly, the number of received and sent text SMS messages is converted to social interaction. The model checks every minute if at least one SMS is received and one SMS is sent (social interaction). Moreover, ambient noise is recorded to evaluate if a user has a conversation with others or if the user is in a noisy environment. The condition for that is the decibel values to be more than 30 dB and the RMS (Root Mean Square) to be more than 150 Hz. Finally, the model detects the location of the user through the Google API activity recognition and estimates that if a user is taking the bus, there will be an interaction with the bus driver or with other passengers. For any other case, the model estimates

that the user is socially isolated. Consequently, if there is at least one sensor that indicates that the user is socially active per minute, then the label for this minute is 'socially active'. The approach of the social behaviour model is illustrated in Figure 2.
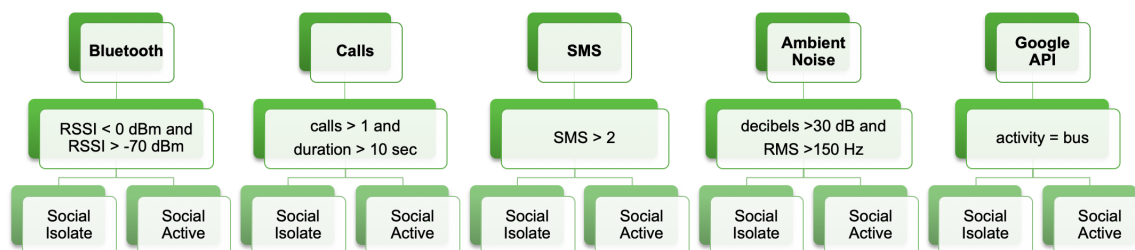

Figure 2. Layered social behaviour model.

4.1.3. Emotional & Cognitive Short-Term Behaviours

The model for detecting emotional and cognitive behaviours is mainly based on surveys. In particular, ESM methods and Likert scale questionnaires are used to measure user's valence and arousal that can provide information about user's affective states [8,9]. These states can be described in terms of discrete emotional categories (i.e., happiness, sadness, fear, anger, disgust and surprise) using the dimensional descriptor of Russell's circumplex model [10]. Valence is defined as the self-perceived level of pleasure, ranging from negative to positive affectivity, while arousal is the self-perceived level of calmness and excitement, ranging from low to high affectivity. Furthermore, visual cues (such as laptop's camera or Kinect sensor mounted on a screen) that monitor user's head, gaze and body posture can also be used to record facial expressions (emotional behaviour) and monitor the engagement level of the user (cognitive behaviour).

4.2. Long-Term Behaviours Inference

The model for detecting long-term behaviours is based on investigating the short-term physical, social, emotional and cognitive behaviours over longer periods (e.g., days, weeks, months). More specifically, a number of statistical features (i.e., sum, mean, standard deviation) is extracted from the short-term behaviour data in order to identify potential descriptors that might reveal information about behaviour routines. These features are used to perform time series analysis over different times (e.g., morning, afternoon, evening, weekdays, weekends, working hours), from which more meaningful and long-lasting events are identified in order to derive information about users' lifestyles.

For instance, the total number of steps per day can give further insights into the activity level of the user; below 5000 steps is considered sedentary, from 5000 to 7499 is lightly active, from 7500 to 9999 is moderately active and above 10,000 is considered vigorously active [11]. In Figure 3, the total number of steps is depicted for a single user and over four months, through swarm plots representations. Trying to identify the days of the week that are related to a sedentary lifestyle, we depict the number of steps grouped by each day, comparing the total number of steps with the steps performed during the morning, the afternoon and the evening. It can be seen that for almost every day of the week, the user has performed more than 10,000 steps per day, with some exceptions on Tuesdays and Wednesdays. Another example of estimating the long-term physical behaviour is through the average duration of the activity siting, which can reveal information about the sedentary lifestyle of the user. Furthermore, the total duration of being socially active versus isolated can be used to monitor long-term social behaviour. Similarly, the average level of user's valence/arousal/engagement per day, week or month, can define the user's long-term emotional and cognitive behaviours.
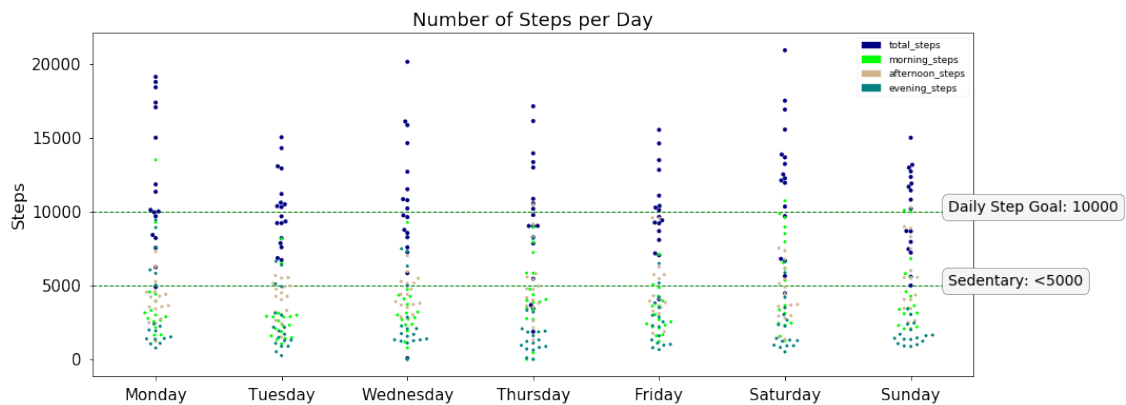
Figure 3. SwarmPlot of the steps per day comparing the total number of steps, with steps performed during the morning (from 8 am to 12 pm), during the afternoon (from 12 pm to 5 pm) and during the evening (from 5 pm to 12 am).

4.3. Technologies

Smartphone data are collected in a continuous and unobtrusive way though the smartphone application AWARE [6]. AWARE is an Android framework dedicated to instrument, infer, log and share mobile context information, for application developers, researchers and smartphone users (see Figure 4). AWARE captures hardware-, software- and human-based data. In particular, smartphone raw data such as accelerometer, GPS, Bluetooth, phone calls, SMS and ambient noise are recorded and stored into an Apache Web server running on a Linux system [12]. Then, scripts are developed in Python 3.5 [13] in order to process the raw data and detect short-term and long-term behaviours. At first, short-term behaviour models detect physical, social, emotional and cognitive behaviours, which are stored into a MySQL 5.7 database [14]. Then, the processed data are pushed to authorized platforms via a secured rest API connection [15], based on JSON web token authentication. Finally, the data are distributed through JSON format according to RFC 8259 [16] and ECMA-404 [17].
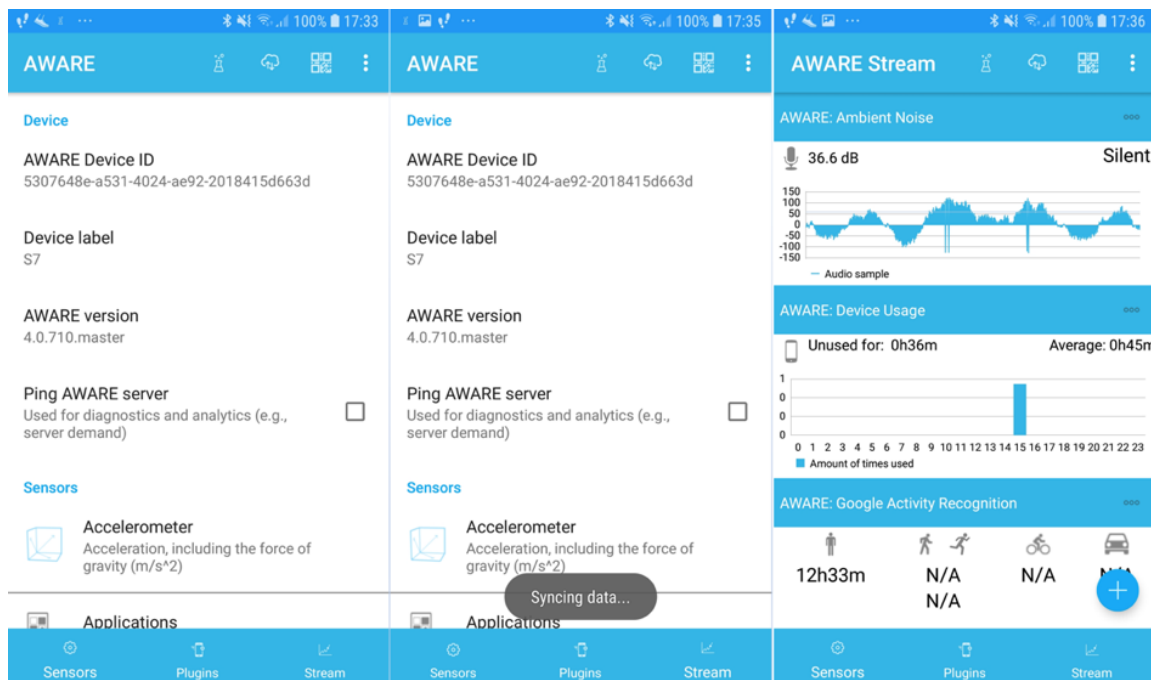
Figure 4. Snapshots of the AWARE application for smartphones [6].

5. Conclusions

This work presents a novel framework for sensing, identifying and quantifying different domains of human behaviour in a holistic approach. Even though the suggested framework has been designed under the scope of the 'Council of Coaches' project in order to nurture the dialogues, arguments and interactions between users and virtual coaches, it can be used in many healthcare application systems which require an automatic detection of human behaviour. The framework collects, stores and analyses raw multimodal on-body and off-body sensor data in order to extract useful information regarding the user's behaviour. More specifically, four principal components of human behaviour, namely physical, social, emotional and cognitive, are analysed over different periods of time. After presenting the framework architecture, we explained the implementation of short-term and long-term behaviours using heterogeneous data from on-body and off-body devices (e.g., smartphone data, laptop visual cues). In addition to the current devices, we plan to integrate more apparatuses into the system including, smartwatches, Fitbits, smart weight scales and other IoT devices. Our next step is to conduct a longitudinal case study, collecting multimodal sensor data from different users, in order to validate the current system but also to investigate changes in human behaviour that occur under certain events and restrain users from following a normative or healthy lifestyle.

References

1. Konsolakis, K.; Hermens, H.; Villalonga, C.; Vollenbroek-Hutten, M.; Banos, O. Human Behaviour Analysis through Smartphones. *Proceedings* **2018**, *2*, 1243.
2. Gaggioli, A.; Pioggia, G.; Tartarisco, G.; Baldus, G.; Corda, D.; Cipresso, P.; Riva, G. A mobile data collection platform for mental health research. *Pers. Ubiquitous Comput. J.* **2013**, *17*, 241–251.
3. Banos, O.; Villalonga, C.; Bang, J.; Hur, T.; Kang, D.; Park, S.; Hyunh-The, T.; Le-Ba, V.; Amin, M.B.; Razzaq, M.A.; et al. Human behavior analysis by means of multimodal context mining. *Sensors* **2016**, *16*, 1264.
4. op den Akker, H.; op den Akker, R.; Beinema, T.; Banos, O.; Heylen, D.; Bedsted, B.; Pease, A.; Pelachaud, C.; Traver-Salcedo, V.; Kyriazakos, S.; et al. Council of Coaches–A Novel Holistic Behavior Change Coaching Approach. *ICT4AWE* **2018**, *1*, 219–226.
5. Harari, G.; Lane, N.; Wang, R.; Crosier, B.; Campbell, A.; Gosling, S. Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. *Perspect. Psychol. Sci.* **2016**, *11*, 838–854.
6. Ferreira, D.; Kostakos, V.; Dey, A.K. AWARE: Mobile Context Instrumentation Framework. *Front. ICT* **2015**, *2*, 6.
7. AWARE: Google Activity Recognition. GitHub. Available online: https://github.com/denzilferreira/com.aware.plugin.google.activity_recognition (accessed on 12 July 2019).
8. Kanning, M.; Schlicht, W. Be Active and Become Happy: An Ecological Momentary Assessment of Physical Activity and Mood. *J. Sport Exerc. Psychol.* **2010**, *32*, 253–261.
9. Bailon, C.; Damas, M.; Pomares, H.; Sanabria, D.; Perakakis, P.; Goicoechea, C.; Banos, O. Intelligent Monitoring of Affective Factors Underlying Sport Performance by Means of Wearable and Mobile Technology. *Proceedings* **2018**, *2*, 1202.
10. Russell, J.A. A circumplex model of affect. *J. Personal. Soc. Psychol.* **1980**, *39*, 1161–1178,
11. Bassett, D.J.; Toth, L.; LaMunion, S.; Crouter, S. Step Counting: A Review of Measurement Considerations and Health-Related Applications. *Sports Med.* **2017**, *47*, 1303–1315.

12. Canonical. Ubuntu 14.04.6 LTS (Trusty Tahr). Available online: http://releases.ubuntu.com/trusty (accessed on 12 July 2019).

13. "Python 3.5" Python.org. Available online: https://www.python.org/downloads/release/python-350 (accessed on 12 July 2019).

14. "MySQL 5.7 Release Notes." MySQL. Available online: https://dev.mysql.com/doc/relnotes/mysql/5.7/en/ (accessed on 12 July 2019).

15. "Slim 3.9.0 Released." Slim Framework. Available online: https://www.slimframework.com/2017/11/04/slim-3.9.0.html (accessed on 12 July 2019).

16. RFC 8259—The JavaScript Object Notation (JSON) Data Interchange Format. Available online: https://datatracker.ietf.org/doc/rfc8259/?include_text=1 (accessed on 12 July 2019).

17. Standard ECMA-404. Available online: https://www.ecma-international.org/publications/standards/Ecma-404.htm (accessed on 12 July 2019).