InCense IoT: A Collective Sensing System for Behavior Data in Shared Spaces †

Jesus Ramos-Monteon 1, Luis A. Castro 1,*, Luis-Felipe Rodriguez 1 and Oresti Banos 2

1 Department of Computing and Design, Sonora Institute of Technology (ITSON), Ciudad Obregon, 85000 Sonora, Mexico; jramos.monteon@gmail.com (J.R.-M.); luis.rodriguez@itson.edu.mx (L.-F.R.)
2 Department of Computer Architecture & Computer Technology, University of Granada, 18071 Granada, Spain; oresti@ugr.es
* Correspondence: luis.castro@acm.org; Tel.: +52-644-410-9000 (ext. 1523)
† Presented at the 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018), Punta Cana, Dominican Republic, 4–7 December 2018.
Published: 23 October 2018

Abstract: Behavioral sensing systems collect data from smartphones, wearables, and other devices with the aim of analyzing and making sense of them. In this work, we present InCense IoT, a collective sensing system which uses mobile and ubiquitous sensors for collecting behavior data of groups of participants in shared spaces. This paper describes the concept of collective sensing, an implementation onto InCense called InCense IoT, innovative features, advantages over individual-centric sensing systems. Finally, this paper presents results of a use case using it in monitoring behaviors in mother-child interactions.

Keywords: collective sensing; mobile sensing; behavioral sensing system

1. Introduction

In recent years, the Internet of Things (IoT) has become popular due to the implementation of useful applications in different areas (e.g., activity recognition). Researchers typically relate this term with technologies such as sensors, actuators, and mobile devices, which combine efforts to solve problems of daily life. Due to the advancement and popularity of IoT, there is interest in using IoT-based systems in the industry such as security, transportation, environmental monitoring, and many others.

Researchers have created various tools which can be used for creating mobile sensing campaigns. These tools, which we refer to as Behavioral Sensing Systems (BSS), are responsible for monitoring human subjects using smartphones, wearables, and other devices. These BSSs have been applied in several domains such as health care. For instance, in [1] implemented a BSS as an alert mechanism for in-hospital emergencies. Also, there are specialized platforms in health monitoring for patients with chronic diseases for home care, hospital or in travel environments [2]. BSSs are used in transport domain to measure and locate delays in public roads and reroute the user [3]. Also, BSSs have been used to identify the user’s transport method [4,5].

Although BSSs have been previously proposed, such as [6–10], they have several limitations, such as the following: (1) BSSs are not standardized nor very flexible, so each time an investigation or sensing campaign is planned, a sensing system is typically created or extended; (2) Apart from battery limitations, mobile phones, wearables are heterogeneous, meaning that the model and brand determine the types of sensors included, the quality of them, as well as the quality of data they can collect; (3) Finally, the majority of BSSs are individual-centric [6,7,10–12], therefore some group context can be difficult to infer. We will explain this in the following sections.
In scenarios in which studying family members and their context is important, an integral sensing system is required so that it can collect data from the family members’ mobile phones and sensing devices positioned in locations where the family gathers such as the kitchen, the living room or the backyard. Beyond that, from a research point of view, the researcher must be capable of configuring the sensing system to collect and pre-process selective data without too much hassle. Although sensing platforms can indeed collect data from fixed sensors, they mainly focus on collecting continuous streams of data regardless of their significance or user privacy [13,14] or portraying objects as data/service providers [15,16]. Having a platform that can be configured to collect individual (e.g., mobile phone sensor data) and group context (e.g., sensor data from a fixed device in the living room) from the researcher’s desk at pre-deployment or at runtime are particularly rare. This is desirable since sensing campaigns (i.e., data collection protocols) requirements can vary across time as they can be running for months.

Conventional BSSs mainly focused on the sensors placed in the mobile phone. Using fixed devices positioned at particular locations in the home setting can open the possibility of using much more specialized sensors than those used in mobile devices (e.g., indoor temperature, carbon dioxide levels). Needless to say, in particular scenarios or studies, this can enrich context to the extent that making sense of collected data would be otherwise difficult. Furthermore, since BSSs focus typically on inferring human activities, using individual and collective context can provide a better understanding of them. For instance, some studies have shown that there is a link between mood and outside weather [17]. Studying similar variables, and family dynamics, can help explain, and perhaps infer, much more complex variables such as mood changes in patient with bipolar disorder being monitored. Again, several sensors that may be of interest are not typically included in off-the-shelf devices such as smartphones.

In this work, we extend a mobile phone sensing platform [18–20] by including non-mobile sensors placed in commercial IoT devices such as Raspberry Pi or Intel Edison. By adding non-mobile sensors, and readily use them in a sensing campaign, altogether with smartphones, we are able to design a unique sensing campaign and collect behavior data from groups or collectivities. We also present how we extended the mobile phone sensing platform and how it was implemented in a semi-controlled setting to collect data from a dyad being monitored. Even when non-mobile sensors have been previously used in sensing campaigns, using a single platform to configure mobile phones and non-mobile devices can be useful for rapidly deploying sensing campaigns. This is indeed one of the advantages of the proposed approach.

This paper presents a novel approach for implementing sensing campaigns, using an extensible architecture of an existent sensing platform. In the following sections we describe the concept collective sensing, our implementation of a platform which supports collective sensing, its advantages over other similar platforms, and the main architecture and features of the platform. Also, we present a use case to illustrate how the platform can be deployed in such settings.

2. Collective Sensing

Collective sensing consists of using diverse mobile and non-mobile devices capable of selectively collecting group context when they interact with each other either in shared spaces or when they are not together. Also, collective sensing provides the ability to sense groups such as families, classroom or campus-wide studies, guest buildings, conferences, or any other type of groups or communities, which scientists can be interested in studying.

One of the appeals of collective sensing, as presented in this work, is not necessarily on full streams of raw data collected and aggregated on a central repository, which implies challenges of its own such as sensor stream synchronization or data fusion, but rather when those raw sensor data are collected and how are treated by the mobile and non-mobile devices. All this can be done at design stage of the sensing campaign from the researcher’s desk. That is, the effort put when deploying devices and preparing them for a sensing campaign can be minimal, but also in redeploying a sensing campaign at runtime since once connected to the network participating devices can be (re)programmed from the researcher’s desk through a web-based interface. The mobile and
non-mobile devices used run operating systems (OS) such as Android or Linux, which facilitate remote manipulation and configuration at runtime.

The advantages of collective sensing, when compared to most individual-centric sensing platforms, are (1) better understanding of group context, (2) greater coverage of the environment beyond the individual, (3) higher data richness due to specialized sensors typically lacking in mobile devices, and (4) greater control over the sensing platform at design stage and runtime through a single web-based interface through which configure and deploy sensing campaigns. The latter is particularly useful since researchers have typically little time or technical knowledge needed for reconfiguring or reprogramming smartphones or devices such as a Raspberry Pi.

We can envision several applications for collective sensing with the scope presented. First, studying the behavior of older adults in nursing homes, their affective state when they are together in a group, and the effects the environment may have on their wellness. Also, one could study dysfunctional families to better understand what and how family dynamics influence individual behavior, and the other way around. Another study of interest, such as the one of [21], can be students’ performance with respect to their environment and group coexistence. The latter is a less controlled environment since student life can involve several locations such as the home, university campus (e.g., library, classroom), a friend’s house, and public locations such as restaurants, which can be difficult to implement with some existing platforms.

Currently, there is no BSS that supports the requirements of collective sensing. As mentioned, most BSS are individual-centric, making use of mobile devices such as smartphones, typically leaving aside the group context. The implementation of collective sensing facilitates and enriches research that requires studying multiple participants who gather or cohabit in the same space. Important requirements for a collective sensing platform include: (1) flexible sensing campaign configuration at runtime, enabling the researcher to have multiple options for sensing (2) campaign editing at any time, in case the user has to make corrective changes on the fly, without delaying the re-deployment of devices, (3) automatic creation of sensing components, saving time of programming to the user, (4) automatic programming of non-mobile devices, and (5) security and privacy of collected data with isolated instances of relational databases and raw data preprocessed before leaving the participating devices.

There are several challenges associated with the implementation of collective sensing, among which are: (a) homogenization of collected data from different vendors; (b) aggregation of data for analysis; (c) web-based management of sensing campaigns; (d) support of both mobile (e.g., smartphones), and non-mobile devices (e.g., Raspberry Pi); (e) different energy consumption and uptime for mobile and non-mobile devices. Furthermore, in regards to formal research, there are also challenges associated with deploying a sensing campaign in scenarios like those discussed above. First, informed consent can be easier when dealing with individuals. Group members can surely sign informed consents, but it may be difficult to disaggregate group context when some of them do not consent to collect, say, temperature data from them.

A Scenario for Collective Sensing

“Andrew is a psychologist who wants to study dysfunctional families to a different level. He has been working for years on this topic, but would really like to have a breakthrough in his area. He just heard about a new way in which mobile technologies can be used. As a trial, he has invited a mid-class family who live in the suburbs, the Johnston family, to collaborate in the study. A few devices have been setup in their home, and everyone agreed to install an app in their mobile phones. Andrew was interested in the time they spent together, and types of places the members of the family were at when they were not together. The devices at home enabled Andrew to have an idea of when they were having conversations, and through their mobiles, he knew exactly who were talking to whom. In addition, Andrew had information about specific aspects of the house such as interior temperature per room, luminosity, motion sensors, and other aspects that haven been reported in the literature to have effects on day-to-day mood. Since he has a psychological background of each member, he knew that John, the youngest one, was particularly prone to detachment which affected
Julia, his mother, and this in turn affected Pedro, the father, who at the time was unemployed. Pedro was having episodes of substance abuse, which made him more verbally aggressive toward Julia and Mary, one of the oldest children. Julia felt overly neglected and had depressive symptoms. Mary, on the other hand, did not know if she was to blame for her father’s behavior. After 16 weeks of collecting data, Andrew found that the more time they spent together in the kitchen, the less likely they displayed aggressive behaviors. He also found an interesting connection between the weather and mood changes of John. Andrew was very satisfied with the technology he acquired, since he did not consider himself to be tech-savvy”


We extended InCense [7,19,20] a mobile sensing platform to include IoT devices. InCense is a mobile sensing platform running on Android powered devices (e.g., mobile phones, Google Glass) with which several investigations have been carried out. Some of the core features are: (a) Dynamic reconfiguration, (b) Data condensation, (c) Data transmission, and (d) Graphical User Interface (GUI) for configuring sensing campaigns [7].

We extended the physical architecture of InCense, which is now composed of different hardware and software components, we named this extension InCense IoT. The InCense server directs the flow of the configuration of the sensing campaigns. A web platform runs on the InCense server for configuring and setting up the sensing campaigns. An Application Programming Interface (API) is used for receiving the data coming from the IoT devices. Also, an encoder component is now responsible for generating source code for the IoT devices (e.g., Edison, Raspberry Pi) and the database, which at the moment is located on the InCense server.

In the extended architecture, mainly text-based data, mainly text-based are collected in JavaScript Object Notation (JSON) representations. Audio streams are collected in audio files, and then processed for extracting features of interest, usually without the captured audio file leaving the participating device for privacy and network performance. Once on the InCense server, preprocessed data are stored in the data structure (see Figure 1). For the scientist’s convenience, the platform provides credentials for the direct management of the database, in case this is required. The owner of each campaign can provide data access to collaborators, according to the role they have in the study (e.g., participant, principal scientist, student).

<table>
<thead>
<tr>
<th>devices</th>
<th>name_sensor</th>
<th>name_sensor_sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>id</td>
<td>id</td>
</tr>
<tr>
<td>uid</td>
<td>device_id</td>
<td>sensor_id</td>
</tr>
<tr>
<td>model</td>
<td>type</td>
<td>value</td>
</tr>
<tr>
<td>location_description</td>
<td>measuring_range</td>
<td>timestamp</td>
</tr>
<tr>
<td>location_altitude</td>
<td>sampling_rate</td>
<td></td>
</tr>
<tr>
<td>location_latitude</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Figure 1. Basic data structure of the extended architecture of InCense.](image)

The data structure is based on a relational database diagram, in which each device has a one-to-many relationship with the sensors, and the sensors a one-to-many relationship with the collected samples. In this structure, the device’s universally unique identifier (UUID), model (i.e., Intel Edison), human-readable location, and geo-coordinates (i.e., latitude, longitude) of the device.

The data stored for each sensor varies depending on its capabilities (e.g., celsius_accuracy for temperature sensor). A table was defined in the structure (schema) for each type of sample associated with a sensor. A value and a timestamp fields are defined for each sample. Using a data structure has some advantages such as the scientist has a greater understanding of the characteristics of the samples and sensors. It also allows making use of queries using database engines.
The scientist running the campaign can create several sensing campaigns at the same time. Each campaign is assigned an individual database to manage the collected data in an isolated and safe way. Only the owner of the campaign has full access to the collected data, the owner’s collaborators have access to the features the owner wants, such as campaign editing, graphs visualization. Database credentials are provided only to the owner. Currently, the researcher cannot manipulate data through the web-based interface; the only way to manipulate the collected data is using standard SQL language using the credentials. Its storage efficiency is the same as a relational database with the pros and cons they currently have. To start the campaign, it is necessary to upload a configuration JSON file to the system through the web-based GUI. This file can be generated by the InCense platform or manually created by a programmer (if desired), and it includes the configuration file for the sensing campaign. At the moment, the GUI and the background routines translate this campaign into source code (e.g., Python) for the IoT devices. Once the campaign is configured, it can be modified by the InCense IoT platform, if necessary. Sensing campaigns metadata can also be modified: collaborators, devices, sensors, and the name, description and status of the sensing campaign (see Figure 2).

The configuration of the sensing campaign includes several sections. The Main section provides access to the following data: status, name, and description. The Collaborators section: A collaborator can be added by entering the email and the permissions (info campaign, edit campaign). The Device section: model, UUID, location description, location latitude, location longitude. The Sensors section: associated device, sensor type, sampling rate, and specific fields for each type of sensor. Once the sensing campaign is set up, the non-mobile devices are programmed automatically to collect data with the chosen configuration.

The GUI displays the information using time-based data graphs. In this way, the researcher has an overview of the data before analysis (Figure 3). This visualization allows monitoring the sensing campaign at any time with a dynamic interface, with the option of selecting the device, the sensor, and the time interval in which the data were collected. Finally, the geographical location of the associated device is shown on a map. This location is based on the Google Maps API, thus the granularity for indoor positioning can be still widely unavailable.

Following Figure 4, The Encoder interprets the configuration of the sensing campaign in JSON format to create the necessary scripts for sensing in non-mobile devices i.e., InCense IoT. What it does is that it takes the code template in Python (Figure 5) associated with the sensor, and replaces the indicated lines of code with the specifications of the campaign. This template file is manipulated through a JavaScript library. Based on this template, for each sensor configured a new file is created and stored in a Git repository, accessible to the non-mobile devices. Also, there is a file which is configured in the creation of the sensing campaign which is executed by the CRON (Daemon) of the OS. This python file executes all the scripts in a multithread way, which begins the sensing process transparently for the IoT device.
In the IoT device i.e., non-mobile device, data are stored in JSON representations. The devices are then synchronized to send the files to the server at a specific time of the day, usually at night. This configuration is executed by the CRON routines of the OS. Then, data are received by RESTful methods implemented on the InCense server, and are stored in the database associated with the...
sensing campaign. Data can be accessed in three different ways: (1) a web-based visualization system, (2) accessing the database with the corresponding credentials, and (3) downloading the SQL file with the schema and data collected, that is, the SQL dump files.

```
MAIN ()
1: temperature <- grove->GroveTemp(__PIN__)
2: WHILE true THEN
3: begin <- time()
4: data->append([time(), __VALUE__])
5: cont ++
6: IF cont == __FILE__ THEN
7: finish <- time()
8: json->add(data)
9: cont <- 0
10: END IF
11: wait(__SAMPLE__)
12: END WHILE
```

Figure 5. Pseudocode Script Example from the Encoder.

4. Use Case: Piloting InCense IoT and Collective Sensing

In order to illustrate how the InCense can be used by researchers from the social sciences, we include a use case in which we collected data from a mother-child in a semi-controlled environment. The purpose of this use case is two-fold: (1) illustrate how InCense IoT can be deployed to collect the data of interest, and (2) provide social scientists with a relatively simple use case that can help them envision the potential usefulness in their research.

This data collection protocol was designed by therapists of children with disabilities who were interested in studying how mothers behave when their children are faced with a mildly-challenging task. In particular, therapists were interested in mothers’ directive behaviors, which are important since they can have several implications for child’s self-management and self-determination.

For the data collection protocol, the task was defined to be putting a puzzle together, for which we implemented InCense IoT. The project was approved by an IRB, and we obtained an informed consent from all mothers.

For directive behaviors, we collected physical proximity and direct intervention, and voice directions or instructions by the parents. Through the therapists, we recruited 12 mother-child dyads. All children are individuals with Down syndrome. The sessions briefly consisted of a child putting 3 puzzles together in direct supervision of the mother (see Figure 6b). The child received three boxes, each containing a puzzle with increasing number of pieces, 4, 9 and 21 pieces, respectively. Each child received one box at a time, the ones with fewer pieces first.

(a)  (b)

Figure 6. Collective sensing use case.

For this use case, we used a sensing campaign configured with the implemented platform creating components that could help us monitor parents’ behavior (Figure 7). InCense Mobile was
running on an Android-based smartphone, using a mic headset. This was used for detecting the mother’s voice directions. Audio data was treated in discrete samples of 1000ms. We used standard pitch-based algorithms for detecting when the mother was speaking. For inferring mother’s intervention, we used an accelerometer sensor for the smartphone, and the ultrasonic sensor in the IoT device to detect when she approached her child. These two components can be combined to monitor mothers’ directive behaviors using rule-based inferences or other approaches such as fuzzy logic or neural networks.

For the audio data, a volume filter was applied combined with a rule-based for recording audio excerpts. It records an audio excerpt when it detects a pitch and it stops when it detects a 3-s silence. Also, a filter was applied to the distance data. The boundaries were configured 0–3 m distance between the parent and the child.

Apart from the technical aspects enabling the implementation of this data collection protocol, and the obvious limitations such as the reduced number of participants, it is important to highlight how a platform such this one can help social scientists, physicians, psychologists, or therapists. One of the most obvious ones is automatic labeling of events in semi-controlled or potentially naturalistic environments. Typically social scientists base their research on self-report through questionnaires or interviews. Self-report based on users’ accounts has been reported to be prone to unintended bias and is often unreliable [7]. Other research methods used in controlled or semi-controlled environments by behavioral scientists is direct or indirect observation (using typically two observers for unbiased analysis) based on video analysis or in situ observation, which is time consuming and exhausting not only to plan, and perform, but also to analyze. Using automatic or semi-automatic labeling of events of interest can speed up and scale up studies such as the ones shown.

```json
{
  "campaign": "Monitoring Parents’ Behaviors for Self-management in Children with Disabilities",
  "begin": "1524715078",
  "end": "1524721176",
  "mobile_devices": [{
    "model": "Android Device",
    "os_version": "8.0",
    "participant_id": "1",
    "sensors": [{
      "component": "voice_directions",
      "filter": "volume",
      "sampling_rate": "1000ms",
      "type": "microphone_sensor"
    },
    {"component": "hands_action",
      "sampling_rate": "500ms",
      "type": "accelerometer_sensor"
    }
  }],
  "IoT_devices": [{
    "location_description": "Room A",
    "location_latitude": "27.4790107",
    "location_longitude": "-109.93227",
    "model": "Intel Edison",
    "uuid": "4e835c7e-0000-4803-95d-859659c9e690",
    "sensors": [{
      "component": "physical_proximity",
      "pin": "A30 Pin 1",
      "sampling_rate": "1000ms",
      "type": "ultrasonic_sensor"
    }
  }]
}
```

Figure 7. JSON representation of the sensing campaign.
5. Discussion

Technology such as the one discussed in this paper can help social scientists, physicians, psychologists, or therapists to better understand scenarios where multiple people interact with each other, particularly in contexts wherein group interaction can be meaningful. Apart from recording human activity of interest, researchers can utilize this technology to automatically label events of interest. This in turn can be used to better tailor certain therapies or non-pharmacological interventions.

There are several challenges associated with this. One of them is disaggregating group context, in case this is required. Having the individual and the group as separate units for analysis can be definitely useful for advancing research. We believe that collective context can have several implications for advancing research in areas where collective and individual contexts may both matter, as one may be interested in understanding which one influences the other.

Our technical work has still several limitations, like knowing in advance which IoT device i.e., hardware will be deployed, and which sensors and in what pins they have been mounted. This is because at design stage sensing campaigns program devices to selectively collect context, which implies that not all data streams are considered at once. Although for some areas collecting all raw sensor data can be desirable and may be needed at the same time (e.g., machine learning), in practical applications such as a household this can be unfeasible for privacy and security (i.e., on-device processing), network usage, and costs associated with storage, retrieval, and computing of those data.

6. Conclusions

In this paper, we presented a BSS that can be used for collective sensing, a paradigm for sensing campaigns which augments mobile sensing campaigns. It provides a paradigm capable of studying groups combining sensors in fixed locations with mobile sensing. Also, it makes use of a GUI for creating sensing campaigns without deep technical aspects of programming, since the platform is responsible for making the necessary code for the devices. It stores the data collected in a schema previously defined in a relational database. The researcher is provided with credential for accessing collected raw data.

Our current implementation enables adding more researchers to the campaign, so they can contribute, the owner grants privileges through the platform (information, edit, visualization). The visualization of the data is a feature in which the researcher can select the time interval of the data to visualize as well as the specific device and sensor can be selected. We are planning several improvements for this visualization feature, like adding new graphs which can give more sense to the data, the graphs interactions have a lot of potential for the researchers in terms of data analysis and advanced visualization features.

As future work, we are going to provide support to several IoT devices and typical sensors for behavior analysis. The reason for adding more devices is to give researchers more than one alternative in terms of the devices they can use. Our initial device (i.e., Edison) is no longer being manufactured by Intel. However, we are abstracting the complexity by using high-level languages running on Linux-based open-source boards such as Raspberry Pi 3. In the long run, we plan to implement a web-based tool capable of analyzing collected data, with the features to make statistical operations and apply strategies for patterns recognition such as fuzzy logic or neural networks using standard Python libraries. We plan to release a public version of our platform soon.
References


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).