www.elsevier.com/locate/ijhcs

Understanding the Potential of Human-Machine Crowdsourcing for Weather Data

Evangelos Niforatos, Athanasios Vourvopoulos, Marc Langheinrich

PII: S1071-5819(16)30134-3
DOI: http://dx.doi.org/10.1016/j.ijhcs.2016.10.002
Reference: YIJHC2078

To appear in: *Journal of Human Computer Studies*

Received date: 22 February 2016
Revised date: 2 October 2016
Accepted date: 9 October 2016

Cite this article as: Evangelos Niforatos, Athanasios Vourvopoulos and Marc Langheinrich, Understanding the Potential of Human-Machine Crowdsourcing for Weather Data, *Journal of Human Computer Studies* http://dx.doi.org/10.1016/j.ijhcs.2016.10.002

Understanding the Potential of Human-Machine Crowdsourcing for Weather Data

Evangelos Niforatos[a*], Athanasios Vourvopoulos[b], Marc Langheinrich[a]

[a]Università della Svizzera italiana (USI), Lugano 6900, Switzerland
[b]Universidade da Madeira (UMa), Funchal 9000-390, Portugal

[*]Corresponding author. evangelos.niforatos@usi.ch.

Abstract

Reliable weather estimation traditionally requires a dense network of meteorological measurement stations. The concept of participatory sensing promises to alleviate this requirement by crowdsourcing weather data from an ideally very large set of participating users instead. Participation may involve nothing more than downloading a corresponding app to enable the collection of such data, given that modern smartphones contain a plethora of weather-related sensors. To understand the potential of participatory sensing for weather estimation, and how humans can be put "in the loop" to further improve such sensing, we created Atmos – a crowdsourcing weather app that not only periodically samples smartphones' sensors for weather measurements, but also allows users to enter their own estimates of both current and future weather conditions. We present the results of a 32-month public deployment of Atmos on the Google Play Store, showing that a combination of both types of "sensing" results in accurate temperature estimates, featuring an average error rate of 2.7 °C, whereas when using only user inputs, the average error rate drops to 1.86 °C.

Keywords: Sensor networks; smart cities; crowdsourcing; mobile sensing

1. Introduction

Today's smartphones come with an increasing number of sensors. Apart from collecting user-generated activity data (e.g., location, acceleration, orientation) these also often support measuring environmental data such as noise levels, ambient temperature, and barometric pressure. This abundance of sophisticated sensory hardware on modern mobile devices was a key driver in the development of the so-called "Participatory Sensing" paradigm. Participatory sensing involves the idea of community-driven sensor-based data collection for generating knowledge about the environment (Burke et al., 2006). Typically, one can contribute by simply downloading a mobile app that runs on one's mobile device which then collects relevant information in the background. By transferring such information automatically or semi-automatically to a central site for analysis, a mobile user effectively becomes a "human sensor". This type of data can be used for various reasons, such as monitoring public health or for crisis management (Boulos et al., 2011). As such, participatory sensing can be seen as a special case of "crowdsourcing," a geographically distributed sensing task (Kanhere, 2011).

In general, we can group participatory sensing into two broad categories of sensing with respect to the user: *personal* and *environmental*. Personal sensing applications such as *DietSense* (Burke et al., 2007) and *CenceMe* (Miluzzo et al., 2008) enable users to collect, track and share information about their personal nutrition and activity levels, respectively. Analyzing such information can provide users with important information about their habits, preferences, and wellbeing, which they might otherwise be unable to realize. An individual's measurements can also provide value for other users. The *Mappiness* app (MacKerron and Mourato, 2010), for example, polls a user's happiness levels in situ and displays it in aggregated form on a map. Soon, more approaches emerged as data aggregators for real-time assessment of participants' habits, health and behavior. As such, the *AndWellness* system includes an Android application and a server software component for collecting participants' input via Experience Sampling (ESM) and passive logging of on-board sensors (Hicks et al., 2010). The *AWARE* framework is a similar but open-source approach that facilitates the development of context-aware mobile applications and data-logging tools for research purposes (Ferreira et al., 2015).