Issues inherent in controlling the interpretation of the Physiological Cloud

Kiel Gilleade
School of Natural Sciences and Psychology
Liverpool John Moores University
Liverpool, UK
K.M.Gilleade@ljmu.ac.uk

Kevin Lee
School of Information Technology
Murdoch University
Murdoch 6150, Western Australia
Kevin.Lee@murdoch.edu.au

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
This paper discusses the potential issues in controlling the interpretation of physiological data exposed to the public using Internet technologies. It identifies a range of issues and discusses potential solutions and their implications. These issues are highlighted and discussed using The Body Blogger project which exposes an individual user's physiological data on the Internet in real-time.

Keywords
Body Blogging, Physiological Computing, Cloud Computing, Privacy, Ethics

ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies. K.4.1 Public Policy Issues: Ethics and privacy.

General Terms
Design, Experimental, Legal Aspects

Introduction
With the rise in cheap physiological sensor technology there is a plethora of information individuals can record about themselves. There is a trend to automate data collection and integrate it into Internet applications: we refer to this as the Physiological Cloud. For example, a user's weight can be shared in a public space via Twitter over Wifi [1], physical activity can be shared

between authorised users via a web interface [2], and blood pressure can be shared in both a public and private setting [3].

By providing physiological data on the Internet, users can benefit from sharing this information with others [5]. For example by sharing physical activity logs users can motivate each other to maintain their level of fitness and improve it (e.g. EA Sports Active 2 [8]). This data can also be traded with others and integrated with different datasets and services, and exposed in different mash-ups providing valve added services [6]. For example, users can combine their pedometer and GPS data with a topological Internet service [7] to provide an overlay of their exercise regimes over a geographical location that subsequently would allow them to identify their ideal training areas.

As with many new services, whist there are many obvious advantages there are potential issues that need to be addressed when designing systems that expose data on the Internet. This is especially true of the Physiological Cloud as this is inherently personal information. There is a need to consider how this information is disseminated amongst the various stakeholders (e.g. users and service providers) interested in the data as there are a variety of privacy and ethical concerns. The remainder of this paper discusses potential issues and solutions with exposing physiological data in the Physiological Cloud. The Body Blogger project [4] is used as a real world case study to illustrate this.

Potential Issues

Exposing an archive of an individual's physiological data on the Internet has potential issues which we discuss in this section. These issues can be split into two groups; those focused on a single individual's dataset and those when the data is aggregated with other datasets.

There are three main issues concerning the usage and interpretation of a single individual's dataset:

- Data is used to support a decision impacting the individual. For example an insurance company uses recreational data in a medical context (e.g. heartbeat data collected by an exercise program, to deny health insurance).
- Data is interpreted outside the owner's preferred context. For example, if an individual is tracking their heartbeat rate, a high reading may be interpreted by another stakeholder to infer the wrong state, such as an employer inferring physical activity as work stress.
- Automation of data collection exposes intimate details about an individual in real-time. This can be used to track aspects of the user's behaviour. For example, real-time heartbeat data can be used to track when someone is asleep thereby posing a personal security risk.

There are two main ways combining an individual's data with another source can be an issue; combining with either another individual or another source type:

- Data is aggregated with data streams from a variety of sources. For example, using temperature sensors from a variety of individuals collocated in a specific area to monitor for flu out-breaks.
- Data being used in an unauthorised mash-up thereby removing an individual's control of the presentation of their data. For example, a mash-up combining a physiological stream and user context

might allow socialisation patterns previously hidden to become obvious.

Potential Solutions

The issues discussed in the previous section range in severity, but all may be of concern to different stakeholders of the Physiological Cloud. There are various solutions that can be implemented to alleviate their impact, discussed as follows:

- Access to the physiological data feed can be controlled. For example, data can be restricted to certain services or presented in a machine unreadable format (e.g. an image) to prevent screen scraping (automated extracting of data from a website).
- The fidelity of the physiological data can be set to an appropriate level for the service being provided. For example, in a medical context, an ECG trace would be required by the stakeholders involved (a doctor), however in an interactions project a derivative of heart activity (average heartbeat rate) would be acceptable.
- Allow the user to control the context (or the associated context) in which their data is presented. For example, only presenting data in the context in which it is intended, or presenting it in a different visual format to obscure any other potential interpretations of the data. For example, instead of representing heartbeat rate data as an absolute numeric value it can be represented as an interpretative visual.
- Allow the user to decide the granularity of the dataset. For example, if an individual's heartbeat rate is continuously updated to the Physiological Cloud, less data points should be uploaded at night to reduce the security risks involved in advertising sleep patterns.

Case Study: The Body Blogger Project

The Body Blogger project involves the real-time collection of an individual user's physiological data and display on the Internet in a continuous fashion (24hrs a day, 7 days a week). This system is used to support a variety of interaction projects including a real-time heartbeat feed on Twitter and a physiological controlled website (i.e. the website graphics change colour according to different physiological readings). The aim of this system is to show the general public how an individual's physiology changes over a long period of time, and provide means for viewers to interact with this data.

In designing this system, potential privacy and ethical issues had to be addressed as it involves exposing an individual's physiology to the public for an extended period of time (one year and ongoing). In the following section we discuss the decisions made in the design of the Body Blogger system in respect to the potential issues and solutions discussed in the previous sections.

For this system, heart activity is provided as an averaged heartbeat rate over a specified period of time making it simple to interpret over a large dataset (e.g. several days worth). While an ECG trace provides a high level of fidelity, this data is unsuitable for viewing by the general public. The granularity of the dataset was originally set very high (e.g. second to a minute updates), however concerns were raised about the personal safety of the individual being tracked as long term monitoring exposed when they were asleep to the public.

Where physiological data is provided as a time series (e.g. on Twitter) it is provided without the context in

which to infer the individual's physiological state. Without such context, stakeholders have been known to infer incorrect interpretations based on their own understanding of physiological activity or their experience of other interaction projects. For example the Moody Web (http://physiologicalcomputing.net) uses fixed heartbeat rates to infer different activities based on prior data recordings. A stakeholder using these states to infer activity on the time series stream would need to be aware that physiological changes vary with the health of the individual and so any associations cannot be guaranteed to be valid. In this example, if the user assigns their sleep state to be 50 BPM or less, which after a health issue increases to 60 BPM someone may infer the individual fails to sleep continuously.

Discussion

The previous sections discussed the potential issues and solutions of exposing physiological data in the Physiological Cloud. The case study demonstrated how these issues can be considered in a real world project. Based on this discussion we propose that any system involved in exposing physiological data from individuals on the Internet should consider a number of factors, discussed as follows.

The designer of a Physiological Cloud application needs to decide who has control of the physiological data and to what extent; will the system allow the user to set the format of the data that mash-ups have access to. For example, providing low fidelity data to mash-ups from a high fidelity source.

Similarly, the level of user access to the data needs to be defined and this should be appropriate for the application and dataset. For example, in a medical context, only a trained professional should be able to see and interpret physiological data.

In order to infer the meaning behind a physiological change it must be interpreted within a relevant context (e.g. pairing a heartbeat monitor with an accelerometer will allow physical activity to be differentiated from cognitive or emotional affects on the heart). It is therefore important that a suitable context is provided with physiological data to prevent stakeholders from misinterpreting the dataset from what was intended. Therefore, the decision of what context, if any, to provide alongside the physiological data is important and should be appropriate for what the system or monitored individual wishes to convey to others.

References

[1] Witherings, BodyScale http://www.withings.com/en/bodyscale, accessed 12/01/11.

[2] Philips Direct Life, http://www.directlife.philips.com, , accessed 12/01/11.

[3] Blood Pressure Chart, http://bp-chart.com, accessed 12/01/11.

[4] The Body Blogger, http://www.physiologicalcomputing.net/wordpress/?page_id=461, accessed 12/01/11.

[5] Gilleade, K., Fairclough, S. H. Physiology as XP – Bodyblogging to Victory. *BioS-Play Workshop at Fun and Games 2010*. Leuven, Belgium.

[6] Merrill D., *Mash-ups: The new breed of Web app*, IBM developerWorks, http://www.ibm.com/developerworks/xml/library/x-mash-ups.html, 8 Aug 2006.

[7] Nike +, http://nikerunning.nike.com, accessed 12/01/11.