The Editable Self: A Workbench for Personal Activity Data

Heather S. Packer
Web and Internet Science
University of Southampton
Southampton, UK
hp3@ecs.soton.ac.uk

Gustavo Buzogany
Web and Internet Science
University of Southampton
Southampton, UK
gbuzogany@gmail.com

Daniel Alexander Smith
Web and Internet Science
University of Southampton
Southampton, UK
ds@ecs.soton.ac.uk

Laura Dragan
Web and Internet Science
University of Southampton
Southampton, UK
lcd@ecs.soton.ac.uk

Max Van Kleek
Web and Internet Science
University of Southampton
Southampton, UK
emax@ecs.soton.ac.uk

Nigel R. Shadbolt
Web and Internet Science
University of Southampton
Southampton, UK
nrs@ecs.soton.ac.uk

CHI 2014, April 26–May 1, 2014, Toronto, ON, Canada.

ACM 978-1-4503-2474-8/14/04..$15.00.

Abstract
The many and varied *personal activity trackers* on the market have the potential to provide unprecedented detail and insight on our everyday activities. However, effective use and interpretation of data from them can be challenging due to common issues. Such issues include false readings due to sensing approaches taken, or missing data arising from a number of different causes. In order to understand user perceptions on this topic, we performed a preliminary survey, which found that users desired the ability to annotate, retroactively repair, and compare their data. Based on insights from this survey, we designed a direct-manipulation interface permitting the consolidated annotation and revision of activity data from multiple devices. A pilot study of this interface found that users understood readily how to use the features offered, and valued the ability to edit, yet preserve the provenance of their data.

Author Keywords
quantified self; self-tracking; health sensors; health monitoring; interfaces

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

General Terms

Interfaces; Interaction

Introduction

Tens of millions of wearable devices were sold last year and projections are set to reach close to 500 million in the next five years [1]. These devices allow people to record various dimensions of their daily routines, as well as provide interpretations through visualisations and statistics. While people personally use such devices for a wide range of reasons, the primary one has been towards achieving personal health goals, such as weight loss or increased fitness [9].

However, the data captured from these devices can be incorrect, due to missing data or misrepresented activities. For example, individuals may forget to wear or charge their device, resulting in gaps in data. Devices that provide activity inference capabilities can incorrectly identify these activities based on partial information, for example a user's movement while riding a bike might be misinterpreted as a rest period because of limited upper body movement [2]. Finally, the results from similar devices may also differ due to specific differences in sensing approaches, movement capture models, or body placement. Therefore, the data collected from different devices is seldom directly comparable, and thus fails to provide continuity when replacing or upgrading devices.

Despite these issues, industry trends suggest that such devices are increasing in popularity, and are achieving widespread adoption. In order to understand whether users of such consumer devices encountered the same sorts of problems previously documented in field studies of research systems, we conducted a survey. Its results showed that users desired the ability to compare and alter their data based on their perceived activity. Based on these results, we iteratively designed a user-friendly direct-manipulation activity data workbench to allow users to easily annotate activities, manipulate their captured data, and to compare data segments from different devices and activities. We then performed a pilot study to understand whether affordances provided could fulfil the needs identified in the pre-study.

We first briefly summarise related work, and the results of our pre-study survey. We then describe the designed interface and its capabilities. Before we conclude, we describe the pilot study evaluating the interface and discuss the results.

Related Work

The development, use and evaluation of wearable personal activity sensing technology has been the focus of numerous studies across a wide variety of research fields, spanning mobile and ubiquitous computing research, lifelogging, preventative medicine, and behavioural economics. Positive effects of the use of such sensors have also been documented extensively, for example, in weight loss [1, 3, 5], and monitoring physical activity [4]. However, to the best of our knowledge, previous work has not effectively addressed the issue of capturing data over long periods of time.

Additionally, several previous studies have looked at the errors arising from the use of personal activity sensing devices. Most notably, the UbiFit work [2] established seven categories of perceived device errors, consisting of errors in start time, duration, activity type, failure in detection and false positives. Their study revealed that

[1] https://intelligence.businessinsider.com/wearable-computing--adoption-happens-2013-4

Problem	Total
incorrect data	8
inaccurate gps	7
missing data	6
differences in data between devices	1

Table 1: Problems reported.

Requested Features	Total
better accuracy	5
change GPS data	5
reminders devices are/aren't recording	2
a way to intuitively change the data	2
better access to data	2
a way to handle missing data	1

Table 2: Requested features.

users were sensitive to all seven types, and that these problems impact users' confidence in their fitness records.

Work that attempts to address missing data in personal fitness devices focuses on techniques that prolong battery life. For example, [6] attempts to work on data compression on mobile devices to improve battery life. Similarly, the work of [7] explored more accurate Energy Expenditure Estimates (EEE) of small wearable sensors. The limitations of personal fitness devices has been noted by [8], who state that "sensor measures are subject to various limitations like resolution errors, calibration variability, response time" and "precision depending on working conditions".

Preliminary survey: Personal Fitness Devices

In order to identify people's perceptions of problems with their activity tracking tools, we conducted a survey pre-study with active users of wearable tracking devices or smartphone tracking apps. This survey asked participants to list the devices or apps that they used, the reasons they used them, and problems they had encountered.

Twenty-two participants responded, nine who used only one device or app, but across all respondents an average of 2.23 tools were used. The most popular were FitBit and RunKeeper, followed by the Garmin GPS watch.

Most participants experienced problems with incorrect data when recording activities and GPS (see Table 1). They were concerned that their personal records could be based on erroneous data. Some participants reported that they had gaps in the data because they forgot their device, forgot to start recording their activity, or forgot to charge the battery. The participants used data from other devices to fill in missing activities, for example one participant said that "sometimes my Garmin watch runs out of battery and I use RunKeeper on my phone to replace it", another stated that they used their running partner's GPS data.

Table 2 lists the missing and desired features suggested by the participants as important to them. They mentioned that they would like intuitive ways to interact with their data, and they did not currently change it because it was "too fiddly". An important feature was a way to adjust the recorded values because they wanted to correct errors so that they could feel that the data matched their perceived energy expenditure.

Interface

Based on the insights gained from our pre-study survey, we designed a direct manipulation interface to visualise and annotate personal fitness data (see Figure 1). It allows users to:

- *Annotate* their data, because users found it difficult to track their activities using devices that continually capture readings and there is no delineation between activities.
- *Correct and replace* data, because the majority of users encountered problems with accuracy and wanted a way to intuitively change their data.
- *Compare* activities from different devices, because users found comparing and integrating data from different devices confusing and currently no interface supports this.

In order to enable users to label their data, we have created an annotation system where users can select a range of data, and enter their label. Different areas that are annotated with the same label become linked automatically, and aggregate statistics can be gleaned and compared against other labels immediately, see Figure 2.

To fill in missing data, we have introduced a drag-and-drop duplicate feature, where users can touch and hold an interval before dragging, to initiate a duplicate operation. A new interval will then be created with a duplicate of the selected interval and can be moved into the required place, see Figure 3. This is appropriate for use when an activity is missing, and the user wishes to use existing data from a similar activity (for example, a run on the same route).

To modify the magnitude of data (for example to correct over-reported of activity while driving), users can select an interval, click and select the central bar, and drag it up or down to modify its magnitude. A green box appears and must be selected to confirm this change, see Figure 4.

Pilot Study

In order to evaluate the features described above, we performed an in-lab pilot study with 5 expert users, recruited via word-of-mouth and a bulletin board within our lab, to evaluate the system's design and capabilities. Our goal with this evaluation was not to understand whether the interface would suit general users (which existing apps do already), but whether it could satisfy the needs of the most demanding kinds of users. We therefore recruited participants who used at least one personal activity tracking device daily, and who identified themselves as "data-geeks" or those interested in the potential for such tools to improve individual well-being. We asked participants to follow a think-aloud protocol during the study, which were recorded, transcribed, and thematically organised using an iterative-coding approach. The findings of the pilot study were then used to set the design goals for the next iteration of the system, as well as to lay the groundwork for a larger-scale field study.

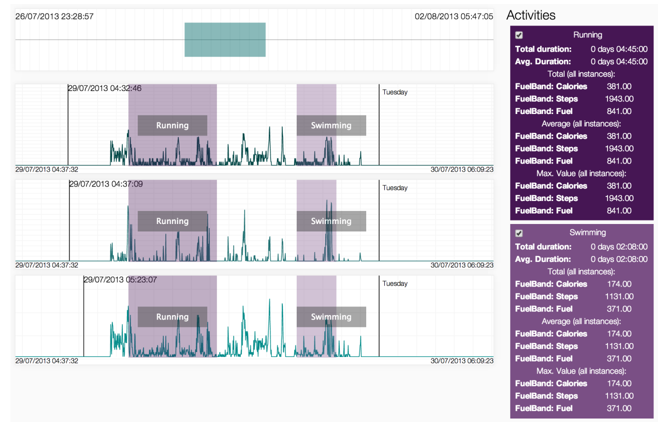

Figure 1: Interface screenshot showing multiple data sources in the center, a time viewport at the top, and multiple activities to the right.

We introduced participants to the interface, which had been pre-loaded with several weeks of data taken from a Nike+ FuelBand, a Jawbone UP and a Fitbit Ultra. We explained and demonstrated features of the interface, and asked participants for their initial thoughts. They could then interact with and explore the system. During the interaction, participants could reset the interface and edit the data as they wished. Afterwards, participants were asked about their thoughts on the system, including features liked and disliked. We also asked whether and how they would use the various features, and in each of these uses, what additional functionality they desired. We then had an unstructured discussion about the interface.

Our participants consisted of 3 males and 2 females, in the age range 25-35. All satisfied the criteria of *data*

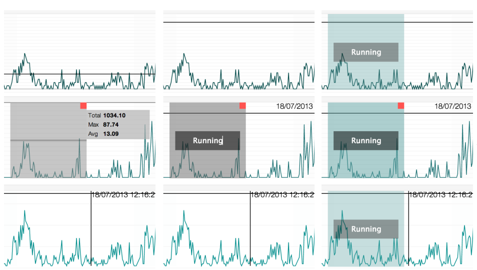

Figure 2: A screenshot of the labelling feature in use. First the user shift-clicks (or pinches) to create an interval. They then double tap to bring up a text label and type their label. When they press enter, the interval label is created for all data sources at that time period.

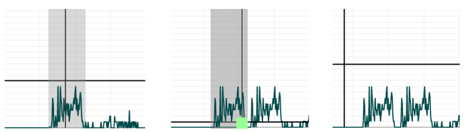

Figure 3: A screenshot of the drag and drop duplication feature in use. First the user click (or touches) and holds an interval, then drags it. Second, the user the drops the now duplicated interval to its new location.

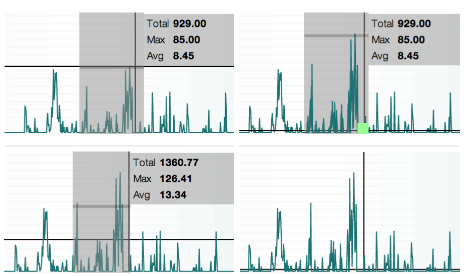

Figure 4: A screenshot of the magnitude modification feature in use. First the user selects an interval, second they click the central bar and drag it upwards to increase the magnitude, third they confirm the change by clicking the green box.

expert we were seeking. The key findings from our pilot user study are grouped by theme in:

Labelling activities: The participants commented that labelling their activities was very useful and something that they could not do in the interfaces that they use. One participant wanted the ability to allow the interface to interact with their calendar because that is where they recorded the type of activities they performed. Another participant suggested recognising patterns in the data so that it could automatically suggest activities for newly entered data using labels from past data.
Filling in missing activities: All of the participants said that they would want their current system to be able to fill in missing time periods for workouts that they regularly perform, and that they liked how easy it was to drag and drop old data to replace missing data.
Altering the magnitude of data: All but one participant said that they would regularly use the feature of altering the magnitude of the data for correcting activities like running up a hill or driving. One participant mentioned that it was similar to music editing software for repeating musical phrases during composition.
Comparing different activities: The comparison of activities from different devices was well received as it is also not available in existing interfaces.

Discussion

Despite the small nature of the pilot study, we identified three common themes that the participants discussed. First, the **preservation of data** was important; most participants were happy to alter their data, although four out of five expressed concern about losing their original data. The participants agreed while there were problems with their data (notably missing data and activities incorrectly captured such driving, cycling, and running up hills) they wanted to be able to refer back to the original data in the interface. They said that they would like to be able to view both the corrected version and the original data in the same time series, and suggested using a toggling function to view both.

The **provenance of data** was also important, the users wanting to record how the original data was altered and, if using old data to fill in missing time periods, where it originated from. They again wished to be able to toggle this information via the interface. They felt it would allow them to distinguish which data had been altered, when reviewing their performance over a long time period.

Third, all participants requested to be able to **share their data** with others. Sharing was seen as a method of self-encouragement, as well as receiving support from others. Besides public sharing, one participant wished for the ability to share their data privately with their training partner, so they could copy data. This participant revealed that she had previously reused her running partners' GPS data because her device only captured her heart rate. The perception was, however, that reusing some data was acceptable (e.g. GPS data), while for other types of data it was not (e.g. heart rate).

These results will influence the future design of our interface. We plan to provide support for toggling between viewing original and revised data, and allow users to view and access provenance metadata for all streams. To support sharing, we will also add functionality to allow users to selectively share raw data streams with others, as well for sharing aggregate activity statistics among multiple channels. Finally, we plan to pursue the capability of automatically detecting activities to reduce the burden of manual annotation, using statistical models and via context from users' calendars.

For these features, the research challenges include showing the user the original data and provenance of the changes to data without cluttering the interface, and providing intuitive interactions. This capability would give the users more confidence in the system, and allow them the freedom of using the tool without having to worry about losing the original data. Sharing data raises privacy issues, such as sharing information that would expose a user's home address or running route which could leave them vulnerable.

Acknowledgements

This work was supported by the EPSRC Theory and Practice of Social Machines Programme Grant, EP/J017728/1.

References

[1] Brown, B., Chetty, M., Grimes, A., and Harmon, E. Reflecting on health: a system for students to monitor diet and exercise. In *Proc. of CHI Extended Abstracts on Human Factors in Computing Systems* (2006).

[2] Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proc. of SIGCHI Conf. on Human Factors in Computing Systems* (2008).

[3] Eastwood, C., Travis, L., Morgenstern, T., and Donaho, E. Weight and symptom diary for self-monitoring in heart failure clinic patient. In *J. of Cardio. Nursing* (2007).

[4] King, A. C., Ahn, D. K., Oliveira, B. M., Atienza, A. A., Castro, C. M., and Gardner, C. D. Promoting physical activity through hand-held computer technology. In *Am. J. of Preventive Medicine* (2008).

[5] Mattila, E., Parkka, J., Hermersdorf, M., Kaasinen, J., Vainio, J., Samposalo, K., Merilahti, J., Kolari, J., Kulju, M., Lappalainen, R., et al. Mobile diary for wellness managementresults on usage and usability in two user studies. *IEEE Trans. on Inf. Tech. in Biomed.* (2008).

[6] Pande, A., Baik, E., and Mohapatra, P. Efficient health data compression on mobile devices. In *ACM MobiHoc Workshop on Pervasive Wirel. Healthcare* (2013).

[7] Pande, A., Zeng, Y., Das, A. K., Mohapatra, P., Miyamoto, S., Seto, E., Henricson, E. K., and Han, J. J. Energy expenditure estimation with smartphone body sensors. In *Proc. of ACM Wirel. Health Conf.* (2013).

[8] Puentes, J., Montagner, J., Lecornu, L., and Lhteenmki, J. Quality analysis of sensors data for personal health records on mobile devices. In *Pervasive Health Knowl. Manag.* 2013.

[9] Swan, M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *Intl. J. of Environmental Research and Public Health* (2009).