Experiences from a Long Run with a Virtual Personal Trainer

Paolo Pilloni, Lucio Davide Spano, Fabrizio Mulas,
Gianni Fenu, and Salvatore Carta

Dipartimento di Matematica e Informatica, University of Cagliari
via Ospedale 72, 09124 Cagliari, Italy
{paolo.pilloni,davide.spano,fabrizio.mulas,fenu,salvatore}@unica.it

Abstract. In this paper, we report on our two-years experience with the commercial application *Everywhere Run!*, a mobile app that allows people to self monitor their running sessions and stay motivated in pursuing a wellbeing life-style. We consider a time interval of two-years, taking as breakpoint the first release of the application that improved the Virtual Personal Trainer presentation. The quantitative data we report comes from a remote logging of the app usage, while the qualitative data comes from the application reviews on the Google Play Store.

Keywords: self-monitoring, wellbeing monitoring, virtual trainer, running, Design for Quality of Life Technologies.

1 Introduction

The pervasive adoption of mobile technologies improved different aspects of our everyday life, providing different ways for supporting different tasks on the go, tracking or logging data that can help us in maintaining or changing our habits.

In the last years, such kind of technologies have been progressively adopted to support people in their daily physical routines. Many research studies provide solid evidence that peoples motivations may benefit from the guidance provided by mobile persuasive systems.

In this paper, we describe our experience with *Everywhere Run!*, a mobile application that allows self-monitoring during running sessions. *Everywhere Run!* differs from other similar apps by means of a Virtual Personal Trainer, able to guide users during a workout as if a real trainer was there with them. The virtual coach analyses the live input provided by the device's sensors, such as the covered distance and the elapsed time, in order to provide users with a broad set of performance metrics. The system provides a real time guidance and support during a workout by comparing a planned pace (i.e., the time to cover a certain distance) with the actual pace kept by the user.

In particular, we discuss the improvement we registered when radically re-designed the UI for presenting more effectively the monitored training parameters and their comparison with the planned values. We analyze the effect of the aforementioned redesign from both a quantitative and qualitative point of view, through a long run experience in the market.

2 Related Work

This section, with no claim of being exhaustive, puts forward some research studies and some technological systems relevant to the field of mobile persuasive technologies designed to promote an active lifestyle.

Berkovsky et al. [1] present a new game design to motivate people in changing their sedentary playing habits. They suggest to stimulate users' motivations by means of virtual rewards users can gain only after performing some physical activity for real during a game. For this purpose, they recommend to modify some native components of traditional video games in addition to the interaction model between players and the game environment. For example, they proposed to introduce virtual rewards to motivate physical activity, to provide an external activity monitor to users in order to calculate the proper virtual reward, and to play the game by means of a modified game controller specifically designed to favour the active gaming experience. Their proposals have been implemented and evaluated through an empirical test on 180 participants in an exergame called Neverball. Obtained results revealed that the proposed design stimulated users to be more active by augmenting the total amount of active gaming.

Hoysniemi in [2] presents the results of a study conducted on users of a dance-based exergame called Dance Dance Revolution. She sets up an online survey to investigate, among others, users' playing styles, motivational, social, and perceived effects of active gaming. The responses of more than 500 users allowed the author to conclude that playing the game have had a positive outcome on users' motivations, physical health, and social life.

In [3], authors propose an exergame called Jogging over a Distance to conduct a series of experiments investigating how social and exertion interactions can affect users while training at a distance. The application exploits a wireless heart rate monitor that allows users to measure their heart rate in real time and also to input a desired heart rate intensity in order for them to specify a physical effort (in terms of that intensity) best suited for their physical skills. By means of a remote server the system elaborates the heart rate data collected in order to modify the 2D position of the audio sent to the jogging partner according to their current performance. If the two runners maintain the heart rate at the level they had previously specified, they can hear the audio of the partner as if they were jogging side-by-side. The results obtained from a qualitative study put emphasis on the importance of social interactions for exertion activities in addition to provide some guidance for researchers that aim at taking advantage of social factors to favor an active gaming experience.

In [4], authors present a mobile system called UbiFit Garden designed to push sedentary people to an active lifestyle. The software consists of three main components: a *fitness device* that automatically records and transmits the data of the physical activity performed both to the *glanceable display* and to the *interactive application*. The application stores all the information of the users' activities and provides a journal where users can consult and edit the data of their activities. The display is placed on the background of the device's screen and it shows a stylized representation of user's physical activities and attained

goals in order to stimulate her motivation. The current level of physical activity is represented through the metaphor of a garden that becomes gradually more and more luxuriant as the user performs physical activities throughout the week. Results from a three-week field trial, revealed that users were positively surprised by the novelties the application introduced and their responses helped authors to derive some guidelines for the design and the evaluation of this kind of systems.

Chittaro et al. [5] propose a location-based exergame based on the famous Snake mobile game in which the snake is guided by users' movements. Authors set out to encourage users to walk more frequently, to introduce new evaluation metrics for this kind of persuasive systems, and finally, to demonstrate the effectiveness of the proposed solution by means of standard questionnaires mainly used in the medical field. Obtained results point out how users' behaviour and lifestyle can be positively influenced by the enjoyment resulting from the game.

In addition to the academic research studies and prototypes listed above, there exist several other commercial applications specifically designed for mobile devices.

One of the most famous application is Nike+ [6]. Some of its main strengths are: the advanced vocal cues and music management system, the availability of several features to favour social interactions with social networks, the support of a web community where users can create their training plans and socialize with other people, in addition to the possibility to enrich the training experience through ad-hoc devices. Along with Nike+, there exist several other commercial examples of such a kind of applications. Among the most popular ones: Adidas miCoach, Endomondo, MapMyFitness, RunKeeper, and Runtastic (see Section "Bibliography" for the references). All the applications share roughly the same core features: position tracking, statistical reports of users' activities, and support for social interactions through social networks and ad-hoc communities.

The experiences reported throughout this work are based on statistics from real users that chose our application freely for their training. With respect to the state-of-the-art proposals, our case study is the first example of mobile persuasive software born as an academic project and soon became a product with some characteristics similar to the most popular commercial products. Differently from the academic studies listed above¹, our application is allowing us to carry out a series of long-term experimentations, thus limiting the biasing factors typical of the most part of research studies on this field.

3 Experiences from a Long Run

In this section, we set out the main changes in design that have been made during the two years period under consideration. For the purposes of this work, we provide a comparison of the old and the new design of some of the most important screens and features of the application as a key to integrate and facilitate understanding of the data, the statistics, and all the remarks presented

¹ Please note that there are no publicly available statistics or results of experimentations for commercial products.

in Section 4. The new version of the application interface was introduced in February 2013.

Let us start with the workout creation screen for the old and the new design, respectively (see Figure 1 and Figure 2). The menu gives users the possibility to create relatively complex workout plans like the one, called “Monday”, we show in the two figures: the training consists of several “sessions”, called “traits”, defined in terms of a distance and a pace/speed to keep. For example, the “trait 1” (see Figure 1 or Figure 2) is a training session of 2 km that the user should run at a pace of 5 minutes per kilometer (please note that runners generally express speed as the time to run one kilometer or mile). “Trait 1” is followed by “trait 2” where the user plans to run 10 km at a slightly higher pace than before. The new screen allows users to design a training plan with the same flexibility and complexity as before, albeit it is completely different in terms of usability. In particular, we decided to introduce a bar in the topmost part of the screen to make available to users both global application settings and some screen specific options in order to favour a more comfortable and effective use before, during, and after a training session (please note that these considerations hold for all the new design screens). In addition, we removed the set of buttons on the bottom part of the interface, in order to better support the primary task in the training creation, through a direct selection of the existing traits and automatically saving the edited values.

Fig. 1. Workout creation UI (old design)

Fig. 2. Workout creation UI (new design)

As previously pointed up, the most important feature of the application is its ability to act like a real trainer. In short, this functionality sets out to support and motivate the runner during the whole workout in order for her to adhere to the goals set by means of the workout creation screen (see Figure 1 or Figure 2). Figure 3 and Figure 4 show the ongoing workout screen and how the virtual personal trainer feature works.

Figure 3 shows the virtual trainer represented by the orange stick man in the left center of the screen. The virtual trainer acts like a pacemaker (a pacemaker is a runner that leads the race to keep the pace for other participants) so that the

Fig. 3. Virtual personal trainer UI (old design)

Fig. 4. Virtual personal trainer UI (new design)

user (the green stick man in the center of the screen) has just to keep the pace of the trainer focusing solely on the run. In both designs, the horizontal bar in the topmost part of the screen provides a global overview of the whole workout (please note that the length of the workout has already been established during the workout creation) with the actual position of the runner with respect to the virtual trainer. Just below that bar there is a dashboard that reports current speed values, distances, and times regarding both the current trait and the whole training session. The two buttons in the bottommost part of the screen (only for the old design) allow to zoom in/out the part of the whole workout showed in the central part of the screen.

In the old design, the two big arrows in the right side of the screen suggested to the user if she has to slow down or to speed up. In between the two arrows is reported the current distance gap of the runner to the trainer. Moreover, the arrows will be alternatively filled proportionally to signal the need of slowing down or speeding up. All that can be observed in Figure 3. It depicts a runner just 3 meters behind the virtual coach, thus none of the arrows is filled signalling that the user is running with the planned pace.

At this point we discuss some other differences between the two designs: in addition to the aforementioned bar in the topmost of the screen, we concentrated to redesign the screen area that depicts the user and the virtual coach during a workout (i.e., the “personal trainer” area in the lower half of Figure 4). We decided, inspired also by some advices and comments from real users, to simplify some of the components originally present in the early design. First, in addition to a general graphic improvement, we changed the orientation of the “personal trainer” area from vertical to horizontal to be consistent with the whole workout perspective. In addition, we removed the zoom in/out buttons and one of the two arrows, and finally, we moved the distance gap indicator near the icon representing the user, on its left.

In this way, we simplified the overall appearance of the screen for an immediate readability, which is crucial for this kind of application: while running, the user should be able to read and understand the monitored parameters without

spending too much time, since her primary task is to keep running. The new layout is more effective in providing such kind of information, since it facilitates the recognition of specific UI elements which are hidden as soon as they are not needed any more (e.g. the arrow for increasing the speed). The first version instead relayed on short term memory for finding differences between two states of the same element. This change allows users to receive immediately a coarse grained information at a first look, and to focus more on the UI only when more precise information is needed, enhancing the efficiency of the self-monitoring task.

4 Application Usage

From the first releases of the application, we monitored the size and the composition of our user set. By now, we counted about 54000 trainings performed by our 12500 users. As we detail better in the following sections, most of the users (76.5%) are returning ones, which means that once they tried the application once, they continued using the application for further training sessions. Such numbers highlight an overall success of the application in satisfying our target users needs.

In our analysis, we report three different metrics we gathered using Google Analytics [7], which has been designed for collecting statistics about the usage of a website. We associated the end of a training with the visit to a particular URL. In this way, we were able to collect the following metrics:

1. *Number of trainings*, which is the total number of trainings registered in a given time unit (e.g. day, week or month). It corresponds to the definition of *Visit* in [7], which is a group of interactions that take place in the application (or a website in general) within a given time frame. Since a visit may consist of different accesses to the application, it finishes either after 30 minutes of inactivity or at midnight. Provided that we record through Google Analytics only one interaction per training, the term visit can be interchanged with training in our case.
2. *Returning user trainings*, which is the number of trainings created by users that have registered at least another training before the current one [7].
3. *New user trainings*, which is the number of trainings created by first-time users [7]

In the following sections we discuss the data about these metrics in detail.

4.1 Number of Trainings

Figure 5 shows the total number of trainings, starting from the 1st of October 2011 and ending to the 30th of November 2013. The data is reported considering a week as time unit, so each point on the graph represents the number of trainings recorded during the considered week (e.g. the one starting at the 11th of August 2012).

The plot shows an increasing trend for the number of trainings, starting from the first launch of the application: we fitted the data with a linear model and

Fig. 5. The number of trainings created with *Everywhere Run!* from the 2nd October 2011 to the 30th November 2013

we obtained $a = 9.99$ as slope and $b = -99.37$ and intercept ($R_{adj}^2 = 0.94$, $F_{(1,112)} = 2035$, $p < 2 \cdot 10^{-16}$), which means an increase of about 10 trainings per week.

In addition, we are interested here in establishing whether such trend can be considered as a constant during the whole application life-cycle, or if it possible to identify different phases during the examined interval. In order to identify possible slope changes, we exploited the algorithm described in [8,9], which is able to fit the relationship between the response and the explanatory variable (respectively the number of trainings and the week in our case) through a piecewise linear model, estimating the position of the breakpoints.

We started considering the whole time interval (from week 1 to week 114), and we proceeded iteratively excluding the interval higher or lower than the week selected as breakpoint. Figure 6 shows the result of the three iterations:

1. At the first iteration, the algorithm found a breakpoint at week 36. It corresponds to the 2nd of June, 2012.
2. At the second iteration, we considered the interval between week 36 and week 114. This time the algorithm found a breakpoint at week 98, corresponding to the 10th of August, 2013.
3. The last iteration considered the interval between week 37 and week 98. The algorithm found a breakpoint at week 70, corresponding to the 26th of January 2013. After this iteration, the algorithm was not able to find more breakpoints.

We defined a piecewise linear regression model for the data, starting from the breakpoints, the slopes and the intercepts identified for each segment by the algorithm. The resulting model is depicted in figure 6, bottom-right part ($R_{adj}^2 = 0.97$, $F_{(7,106)} = 447$, $p < 2.2 \cdot 10^{-16}$).

Fig. 6. Breakpoints and piecewise linear regression model of the total number of trainings from the 1st October 2011 to the 30th November 2013

While the straight-line model prediction of the values can be considered already good, the piecewise linear regression is more precise. We compared the two linear models with an ANOVA analysis and they resulted significantly different ($p < 4.3 \cdot 10^{-9}$). In addition, the piecewise regression provides some interesting insights for discussing the history of the application, especially if we compare the breakpoint positions with the dates of the application versions published on the Google Play Store.

The first segment shows the increment of trainings recorded in the months after the first application release, which can be quantified roughly in 5 more trainings for each week.

The first breakpoint (about the 3rd of June 2012) corresponds with the launch of the 2.0 version of the application, which happened in the 30th of May 2012. The version introduced the monitoring features, with the first UI version (see section 3). The update caused a steep increment of the training count, since the increase rate changed from 5 to 10 trainings per week. This means that the new features were able to better satisfy the users' needs, and the changes were considered useful for our audience.

The second breakpoint sets an additional increment of the slope, which is about 15 in the third segment. The breakpoint date is the 27th of January 2013, which is close to the release of the version 2.1 of the application (the beginning of February 2013). In that version, as already discussed in section 3, the

self-monitoring features included in the application have been organised in a clearer and more effective way.

Finally, the last segment of the piecewise regression shows that at the beginning of August the increasing effect of the improvement finished, but we were able to maintain almost constant (with a slightly increasing trend) the number of trainings recorded.

4.2 Returning User Trainings

The number of trainings registered by returning users follows the same trend discussed for the total number of trainings, which is shown in the left part of figure 7. We recall that a user is defined as returning if she registered at least another training through the application in the past [7].

Following the same methodology we used in section 4.1, we found the same number of breakpoints:

1. Week 35, corresponding to the 27th of May 2012
2. Week 102, corresponding to the 1st of September 2013
3. Week 68, corresponding to the 13th of January 2013.

Fig. 7. The number of trainings created by returning users (left) and its piecewise linear regression (right)

Figure 7, right part shows the corresponding piecewise regression model ($R_{adj}^2 = 0.96$, $F_{(7,106)} = 432.6$, $p < 2.2 \cdot 10^{-16}$). Provided that the breakpoints are located more or less in the same weeks, we can associate the changes of slope in the model to the same events we identified for the total number of trainings. Again, the ANOVA analysis found a significant difference between the simple straight-line regression and the proposed piecewise model ($p < 2.2 \cdot 10^{-16}$).

As a consequence, the trainings created by the returning users follow the same trend of the total number of trainings. This fact, together with the trend identified for the trainings created by new users discussed in section 4.3, allows

us to conclude that our application has been able to satisfy the new users that tried *Everywhere Run!*, transforming most of them in returning ones.

This is confirmed also by a more deep insight on the composition of our user base. If we consider the time span between June 2012 (corresponding to the first breakpoint) and August 2013 (corresponding to the third breakpoint), we have that in the first period, one third of our 4300 users (30%) tried the application only once, while the other two thirds (70%) were returning ones. Our system registered a total of 7800 hours of running, with an average of 30 minutes per workout. Each user had 0.41 trainings per month (calculated excluding the one-time users).

After introducing the improvements discussed in this paper, we registered a consistent increment of the application usage. We doubled the number of users (7900) and, in addition, we increased the number of returning users reaching the 80% and the average training rate to 0.56 per month. The total running time increased from 7800 to 21700 hours (almost tripled), with the average value passed from 30 to 36 minutes per workout.

4.3 New User Trainings

We report the data on the trainings created by new users (which do not have registered any training before the current one) in figure 8 (left part).

We applied again the same iterative method discussed in the previous sections, and this time we got only one breakpoint at week 82 (the 21th of April, 2013). The piecewise model is significant ($R_{adj}^2 = 0.86$, $F_{(3,108)} = 225.7$, $p < 2.2 \cdot 10^{-16}$) and it is more precise with respect to a straight-line regression (ANOVA $p < 2 \cdot 10^{-7}$).

The first segment has a more gradual slope if compared with the other two metrics we considered (2.2). This increasing trend reached the equilibrium at the breakpoint (week 82). After that the number of new users has been approximately constant. Considering this, we can conclude that the growth of the total number of trainings recorded through the application cannot be explained with any particular increase on the number of new users, although it would be desirable to increase such rate for enlarging the audience. However, the analysis in section 4.2 shows that people usually create other trainings after the first one, therefore the application is able to convince people in using *Everywhere Run!* for further trainings after the first try.

4.4 Users' Feedback

We analysed the user reviews in the Google Play Store, in order to assess how our users feel and think about the app, and to see whether the difference between the two designs was reflected also in the ratings. In the first version of the UI, we had an average rate of 3.6, increased to 4.5 with the new version (5.0 is the maximum rate). Comparing such quantitative feedback with the qualitative comments of the app reviews, we collected the evidence that the key factor for

Fig. 8. The number of trainings created by new users from the 2nd October 2011 to the 30th November 2013

such improvements is the ability to guide step-by-step the user in their workouts in realtime, providing a precise position tracking together with different professional-level metrics (particularly important for high-end runners), which are shown to the users for a comparison with previous sessions.

5 Conclusion and Future Work

In this paper we discussed our experience with *Everywhere Run!*, a mobile application for self-monitoring during running sessions. The application provides a virtual trainer which allows the user to monitor her performance while running and to motivate her for reaching a planned result.

The analysis of the data, collected through Google Analytics in two years, shows an increase of trainings in correspondence to the introduction of the self-monitoring features, together with an additional increase when the UI has been enhanced. The trainings from returning users followed the same trend, showing that the application is able to fulfil our target user's needs.

For further improvements of the application, we plan to add the support for tracking the altitude and the calories burned. In addition, we are starting some experiments for using other external sensors, such as the heart-rate belts, for monitoring the runner's physical state.

Acknowledgments. This work is partially funded by Regione Sardegna under project SocialGlue, through PIA - Pacchetti Integrati di Agevolazione "Industria Artigianato e Servizi" (annualità 2010).

We also gratefully acknowledge Sardinia Regional Government for the financial support (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 - Axis IV Human Resources, Objective 1.3, Line of Activity 1.3.1 Avviso di chiamata per il finanziamento di Assegni di Ricerca).

References

1. Berkovsky, S., Coombe, M., Freyne, J., Bhandari, D., Baghaei, N.: Physical activity motivating games: virtual rewards for real activity. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 243–252. ACM (2010)
2. Hoysniemi, J.: International survey on the dance dance revolution game. *Computers in Entertainment (CIE)* 4(2), 8 (2006)
3. Mueller, F.F., Vetere, F., Gibbs, M.R., Agamanolis, S., Sheridan, J.: Jogging over a distance: the influence of design in parallel exertion games. In: *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, Sandbox 2010*, pp. 63–68. ACM, New York (2010)
4. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J.A.: Activity sensing in the wild: a field trial of ubifit garden. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2008*, pp. 1797–1806. ACM, New York (2008)
5. Chittaro, L., Sioni, R.: Turning the classic snake mobile game into a location-based exergame that encourages walking. In: Bang, M., Ragnemalm, E.L. (eds.) *PERSUASIVE 2012. LNCS*, vol. 7284, pp. 43–54. Springer, Heidelberg (2012)
6. Nike: Nike+ gps, nikerunning.nike.com (accessed January 29, 2014)
7. Google: Google Analytics, <http://www.google.com/analytics/> (accessed January 13, 2013)
8. Muggeo, V.M.: Estimating regression models with unknown break-points. *Statistics in Medicine* 22, 3055–3071 (2003)
9. Muggeo, V.M.: segmented: an r package to fit regression models with broken-line relationships. *R News* 8(1), 20–25 (2008)
10. Runtastic: Runtastic, <http://www.runtastic.com> (accessed January 29, 2014)
11. Endomondo: Endomondo, <http://www.endomondo.com> (accessed January 29, 2014)
12. RunKeeper: Runkeeper, <http://runkeeper.com> (accessed January 29, 2014)
13. MapMyFitness: Mapmyfitness, <http://www.mapmyfitness.com> (accessed January 29, 2014)
14. miCoach: micoach, <http://www.adidas.com/fi/micoach> (accessed January 29, 2014)