Finding a Dynamical Model of a Social Norm Physical Activity Intervention
Asanza, Victor; Martin, César A; Eslambolchilar, Parisa; van Woerden, Hugo C; Cajo, Ricardo; Salazar, Carlos
Published in: 2nd IEEE Ecuador Technical Chapters Meeting
Publication date: 2017

Link to author version on UHI Research Database

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the UHI Research Database are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights:
1) Users may download and print one copy of any publication from the UHI Research Database for the purpose of private study or research.
2) You may not further distribute the material or use it for any profit-making activity or commercial gain
3) You may freely distribute the URL identifying the publication in the UHI Research Database

Take down policy
If you believe that this document breaches copyright please contact us at RO@uhi.ac.uk providing details; we will remove access to the work immediately and investigate your claim.

Download date: 24. Sep. 2019
Finding a Dynamical Model of a Social Norm Physical Activity Intervention

Víctor Asanza¹, César A. Martí¹, Parisa Eslambolchilar², Hugo van Woerden³, Ricardo Cajo¹, Carlos Salazar¹
¹ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ingeniería en Electricidad y Computación, Campus Gustavo Galindo Km. 30.5 Via Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador {vasanza, cmartin, rcajo, asalazar}@espol.edu.ec
²Cardiff University, School of Computer Science and Informatics, Cardiff CF10 3AT, United Kingdom {eslambolchilarp}@cardiff.ac.uk
³University of the Highlands and Islands, Centre for Health Sciences, Inverness IV2 3JH, United Kingdom {Hugo.VanWoerden}@uhi.ac.uk

Abstract—Low levels of physical activity in sedentary individuals constitute a major concern in public health. Physical activity interventions can be designed relying on mobile technologies such as smartphones. The purpose of this work is to find a dynamical model of a social norm physical activity intervention relying on Social Cognitive Theory, and using a data set obtained from a previous experiment. The model will serve as a framework for the design of future optimized interventions. To obtain model parameters, two strategies are developed: first, an algorithm is proposed that randomly varies the values of each model parameter around initial guesses. The second approach utilizes traditional system identification concepts to obtain model parameters relying on semi-physical identification routines. For both cases the obtained model is assessed through the computation of percentage fits to a validation data set, and by the development of a correlation analysis.

Keywords — Behavioral interventions; Dynamic modeling; System identification.

1. INTRODUCTION

Sedentary lifestyle has been identified as the fourth risk factor for overall mortality, and this type of lifestyle is on the rise around the globe. Sedentary lifestyle has significant implications on the public health (i.e., the prevalence of non-communicable diseases such as cardiovascular diseases, diabetes, cancer and increasing other risk factors such as: blood pressure, blood sugar and obesity). It is estimated that sedentary lifestyle is the main cause of approximately 21-25% of breast and colon cancer, 27% of diabetes and approximately 30% of the burden of ischemic heart disease [1]. In addition, non-communicable diseases now account for almost half of the global burden of global morbidity and it is estimated that out of every 10 deaths, 6 are attributable to non-communicable conditions [2].

One way to increase levels of physical activity is active walking; this is one of the most widely available types of physical activity and is related to lower mortality rates [3]. It does not require special skills, places or equipment; it is often a natural part of domestic and work routines and is described by most people as pleasant and relaxing [4]. Walking is very beneficial for health [5, 6]. It can prevent or ameliorate long-term conditions such as obesity, type-2 diabetes and cardiovascular disease [7, 8]; helps reduce depression and anxiety, may improve self-esteem [9] and has been shown to reduce cognitive impairment [10].

The study conducted in [11] used a randomized controlled trial to assess the impacts on physical activity of healthy male adults by using a smartphone app called bActive to provide conventional and social norms of feedback in their daily walking. The purpose of the study was to test three hypotheses: H1 – receipt of social feedback generates higher step-counts than receipt of no feedback; H2 – receipt of social feedback generates higher step-counts than only receiving feedback on one’s own walking; H3 – receipt of feedback on one’s own walking generates higher step-counts than no feedback.

Health behavior makes use of theories to guide research to prevent or treat diseases, promote health and / or improve well-being [14]. Some significant efforts have been made to integrate the principles of control engineering into health behavior [15, 16, 17]. Social Cognitive Theory (SCT) [22] is one of the most influential conceptual frameworks of health behavior and has been used as the conceptual basis of different types of behavioral interventions [18]. Martin et al. [21] proposed an SCT dynamic system model based on a fluid analogy. The purpose of this work is to find a validated model that describes a social norm feedback behavioral intervention on the basis of the SCT model developed in [21]. To find suitable model parameters two methods are tested: random search and regular system identification. The first one evaluates different values of model parameters that are taken from a set of random variations around predefined initial values. The second method relies on semi-physical identification routines that find model parameters by solving a quadratic optimization problem. The accuracy of both methods is evaluated through the computation of percentage fits to a validation data set. One reason of proposing the random search methods is to evaluate the possibility of obtaining model parameters in a simpler and more intuitive way, as a first step to evaluate computer science techniques applied to this area.

This paper is organized as follows. Section II presents a description of the behavioral problem and how the SCT model is adapted to represent the intervention. Section III describes the two methods for finding the model parameters. Section IV presents the obtained results. Finally, in the Section V we provide a summary of our conclusions and future work.
II. DESCRIPTION OF THE BEHAVIORAL PROBLEM

A. Problem description

Data used in this work were obtained from a controlled experiment that measured physical activity levels in adult participants aged 20 to 40 years. Measurements were performed using applications and sensors embedded into smartphones. The experiment lasted eight weeks, between October and December 2011, with a total of 165 enrolled participants. Personal information was gathered, and each participant signed a written consent to allow the use of data in scientific studies. Documents were approved by the research ethics committee of the University of Swansea [20]. Enrolled participants were randomly assigned to one of three groups (55 participants were allocated to each group):

1. Control group (no feedback was provided and there was no access to the application);
2. Individual feedback group (feedback was provided based on participant's own steps);
3. Social group (feedback was provided based on the participant's own steps and the average steps taken by others in his/her group).

Measurements were taken daily. The first two weeks were used to provide and measure baseline data. The set used for analysis considers the final six weeks (42 days). Many indicators were measured, however, for this work the following exogenous signals are considered:

- CountUsage: daily number of glances at the app.
- SumUsage: total length of time (in seconds) per day the app was glanced at.
- Weather: the environmental condition in which the subject is influenced to engage or not into physical activity. The values are as follows: sunny 10, half sunny 7.5, cloudy 5, cloudy with drizzle 2.5 and rainy 0.
- Weekend: indicates if the date of measurement is either in a weekend (1) or not (0).
- Steps: the performed daily steps for each participant.

Fig. 1 shows the average number of steps performed by participants in three groups over the 42 days of the study. Since one of the research goals is to analyze and contrast the impact of using a smartphone application with a group feedback [19], the modeling done in this work, is focused on the social group data.

B. SCT Model

Social Cognitive Theory (SCT) describes human behavior in terms of how individuals proactively self-reflect, self-regulate, and self-organize [22]. In this sense, the concept of triadic reciprocity is fundamental, since it describes how personal factors (cognitions, affect, and biology), environment, and behavior all co-interact and influence one another. SCT attempts to explain individual’s engagement in a targeted behavior, based on self and externally perceived factors and their interrelationships. SCT has been used as a basis for the design of physical activity interventions in behavioral settings [21].

A fluid analogy of SCT was developed in [21], where the principle of mass conservation was used to define the following set of equations that describes the system:

\[
\begin{align*}
\tau_1 \frac{d\theta_1}{dt} &= \gamma_1 \xi_1(t) + \beta_1 \xi_2(t) - \theta_1(t) \\
\tau_2 \frac{d\theta_2}{dt} &= \gamma_2 \xi_2(t) + \beta_2 \theta_1(t) + \beta_2 \xi_3(t) - \theta_2(t) \\
\tau_3 \frac{d\theta_3}{dt} &= \gamma_3 \xi_3(t) - \gamma_3 \xi_2(t) - \beta_3 \theta_1(t) + \beta_3 \xi_3(t) - \theta_3(t) \\
\tau_4 \frac{d\theta_4}{dt} &= \beta_4 \xi_4(t) + \beta_4 \theta_1(t) + \beta_4 \xi_2(t) + \beta_4 \xi_3(t) - \theta_4(t) \\
\tau_5 \frac{d\theta_5}{dt} &= \gamma_5 \xi_5(t) - \theta_5(t) \\
\tau_6 \frac{d\theta_6}{dt} &= -\eta(t)
\end{align*}
\]

The state space representation of the system has the following structure:

\[
x_p(t) = A(\theta_p)x_p(t) + B(\theta_p)u_p(t) + Ke(t) \\
y_p(t) = Cx_p(t) + v(t)
\]

where

- \(x_p(t)=[\eta_1...\eta_n]\) denotes a vector of \(n=6\) state variables,
- \(u_p(t)=[x_1, x_2, x_3, x_4]\) denotes a vector of \(m=4\) input variables,
Fig. 2. SCT model subsystem used to describe the social norm physical activity intervention.

\[ y_p(t) = [\eta_4] \] denotes a vector of \( p = 1 \) output variable,

\[ A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{m \times n} \] are the state matrices,

\[ \theta_p \in \mathbb{R}^{np} \] denotes a vector of \( np = 21 \) unknown model parameters,

\( e(t) \) and \( v(t) \) are uncertainties associated to each one of the states and outputs.

A subset of the SCT model is used to describe the levels of physical activity among male adults according to the randomized controlled trial conducted in [19]. Fig. 2 shows the input and output signals of the model, that are related to those from the experiment as:

- CountUsage \( \leftrightarrow \) Skills Training.
- SumUsage \( \leftrightarrow \) Observed Behavior.
- Weather \( \leftrightarrow \) Perceived Barriers.
- Weekend \( \leftrightarrow \) Environmental Context.
- Steps \( \leftrightarrow \) Behavior.

Based on the SCT structure and the actual signals, the unknown model parameters are:

\[ \theta_p = [r_1 \ r_2 \ r_3 \ r_4 \ r_5 \ r_6 \ \beta_{24} \ \beta_{31} \ \beta_{34} \ \beta_{32} \ \beta_{33} \ \beta_{35} \ \beta_{36} \ \beta_{34} \ \gamma_{11} \ \gamma_{22} \ \gamma_{33} \ \gamma_{57}]^T \]  

and the state matrices are:

\[
A(\theta_p) =
\begin{pmatrix}
1 & 0 & 0 & \beta_{31} \\
r_1 & r_2 & r_3 & 0 \\
0 & 0 & r_4 & \beta_{34} \\
r_5 & 0 & r_6 & 0
\end{pmatrix}
\]

(4)

\[
B(\theta_p) =
\begin{pmatrix}
\gamma_{11} \\
r_1 \\
0 & \gamma_{22} \\
r_2 & 0 & \gamma_{33} \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
r_3 & 0 & 0
\end{pmatrix}
\]

(5)

\[
C = (0 \ 0 \ 0 \ 1 \ 0 \ 0)
\]

(6)

III. PARAMETER SEARCH METHODS

Two approaches are tested to find the appropriate values for the model parameters \( \theta_p \) presented in equation (3). The first one is an intuitive random search that starts from an initial guess of parameters, and then performs random variations of each parameter around the initial values. The best model is the one with the lowest estimation error. The second method uses a semi-physical system identification algorithm that considers
the SCT structure model. For both cases the accuracy of the model is determined via the percentage of fit of the $\eta_4$ output model as:

$$% \text{fit} = 100 \left(1 - \frac{\|y_p - \hat{y}_p\|}{\|y_p - \text{mean}(y_p)\|}\right)$$  \hspace{1cm} (7)

Where $\|\cdot\|$ stands for two-norm, that for a vector $r$ is equal to $\|r\| = \sqrt{r^T r}$.

### A. Random Search

As a first step in utilizing computer science approaches to find model parameters, a simpler search algorithm is proposed. The algorithm is described in Fig. 3. During $Max_i$ times the initial random parameters are generated and the initial $% \text{fit}$ is calculated; then, these data are added as one row of the $MaxData$ matrix. During $Max_j$ times new parameters are generated through random variations around the initial parameters, and only those that exhibit $% \text{fit}$ greater than the initial one will be added as one more row of the $MaxData$ matrix. Before writing $MaxData$ as a csv file, all rows are sorted according to the $\text{fit}$ from lowest to highest.

1. Initialize variables: Maxi=100, Maxj=50, Array[].
2. For $i=1$ to $Max_i$:
   - Initialize_Parameters=Random (0 to 15); Generate random parameters
   - FitDaily=EvalModel (Parameters) ; Calculate the new parameters
   - $A = \text{Array}[\text{Initial_Parameters}, \text{FitDaily}]$; Save Initial parameters; with their respective calculated Fit
3. For $j=1$ to $Max_j$:
   - New_Parameters= Random (Initial_Parameters);
   - New_FitDaily= EvalModel (New_Parameters);
   - Calculate the; new Fit with current parameters
4. If New_FitDaily > FitDailyy then:
5. Parameters= New_Parameters; Save new Parameters

Fig. 3. Random Search Algorithm.

### B. Semi-physical Identification

This approach uses semi-physical system identification ideas. To estimate $\theta_p$, the well-known prediction-error identification methods (PEM) [23] are used. The one-step ahead prediction error of the system is:

$$\epsilon(t, \theta_p) = y_p(t) - \hat{y}_p(t \mid t-1, \theta_p)$$  \hspace{1cm} (8)

where $\hat{y}_p(t \mid t-1, \theta_p)$ is the predicted output based on estimated models. Estimation is performed in MATLAB® using the commands idgrey and greyest from the System Identification Toolbox [24].

An idgrey model represents a system as a continuous-time or discrete-time state-space model with identifiable (estimable) coefficients. To estimate the unknown parameters of the idgrey model, the estimating function greyest is used.

greyest function is based on the Levenberg-Marquardt algorithm (LMA), which is a numerical search method. This algorithm is used to solve non-linear least squares problems.

### IV. RESULT ANALYSIS

One subject, participant number 115, is selected out of 55 participants from social feedback group. The selection criteria is based on the fact that there is no data loss during the study period.

The random parameter search is performed, and one hundred initial values are generated ($\text{max}_1 = 100$). From here fifty random searches ($\text{max}_j = 50$) are performed around the initial values attempting to improve the fit. Considering all these five thousand interactions, the best fit obtained was 12.35%. Fig.4 shows a graphic comparison between real data and results from the estimated model with the best fit. As expected obtained fits are not high enough. This method is very simple and gives the opportunity of improvement in many directions. However, the algorithm is not efficient enough since it requires too many iterations to obtain better fits.

Results obtained using the semi-physical system identification procedure are shown in Fig. 5. After nine iterations, a 30.5% fit was achieved. The advantage of using this method is the low number of interactions required to achieve a good fit and therefore a very short processing time.
An additional validation technique is applied to both models via correlation analysis. A prediction error vector according to equation (8) is computed, and statistical autocorrelation \( \rho_e(k) \) and cross-correlation \( \rho_{ue}(k) \) of residual errors to the different inputs are computed and plotted using:

\[
\rho_e(k) = \frac{\gamma_e(k)}{\sigma_e^2}; \quad \rho_{ue}(k) = \frac{\gamma_{ue}(k)}{\sqrt{\sigma_u^2 \sigma_e^2}}
\] (9)

If the obtained model is appropriate, autocorrelation must observe a value of 1 only at lag zero, and values below confidence regions for the other lags. On the other hand, cross-correlations to the inputs must be below confidence bounds for all the lags.

Results of the correlation analysis for the random search method are shown in Fig. 6. Cross-correlation values between the prediction error and the four inputs are within confidence bounds, indicating the error is not related to the nature of the inputs. However, error autocorrelation values are outside bounds what means the obtained error does not resemble white noise as assumed in the model structure.

Results for the semi-physical identification method are shown in Fig. 7. This time both, autocorrelation and cross-correlation of prediction error and inputs, are within confidence bounds. This confirms the quality of this model as the obtained fit suggested.

V. CONCLUSIONS AND FUTURE WORK

In this article, we adapted an SCT dynamical model structure to represent human behavior in terms of daily number of steps walked during a social norm physical activity intervention delivered using smartphones. To find numerical values for model parameters, we tested two search methods: Random Search and regular System Identification. The accuracy of the models is tested through the computation of percentage fits of predicted models to real measurements. Results of the random search algorithm were not acceptable; however, it serves as a good starting point. Acceptable results were obtained by using a semi-physical system identification algorithm; in general this method achieved a better fit with fewer interactions. Both approaches present better performance for participants that show a minimum of five hundred steps per day.

One of the advantages of the proposed approach to represent behavioral interventions is the capability to represent the time varying nature that is common in behavioral settings. Another advantage is the existence of a model structure that relies on a well known theoretical framework; this enables the possibility of performing improving actions over aspects that are more relevant according to the theory. This approach can be generalized to other contexts and datasets via the design of primary system identification experiments particularly selected to find the most important dynamic excitation modes of the system.

For future work, the intervention should be redesigned considering the obtained model and utilizing some control design approaches such as Model Predictive Control (MPC). An improved version of the parameter search algorithm will be also formulated, using machine learning computational intelligence approaches.

ACKNOWLEDGEMENT

The authors thank the EPSRC, UK, for supporting the RCT of smartphone app study [11].