Data-driven assessments for sensor measurements of eating behavior

Christos Diou[1], Ioannis Sarafis[1], Ioannis Ioakimidis[2] and Anastasios Delopoulos[1]

***Abstract*— Two major challenges in sensor-based measurement and assessment of healthy eating behavior are (a) choosing the behavioral indicators to be measured, and (b) interpreting the measured values. While much of the work towards solving these problems belongs in the domain of behavioral science, there are several areas where technology can help. This paper outlines an approach for representing and interpreting eating and activity behavior based on sensor measurements and data available from a reference population. The main idea is to assess the "similarity" of an individual's behavior to previous data recordings of a relevant reference population. Thus, by appropriate selection of the indicators and reference data it is possible to perform comparative behavioral evaluation and support decisions, even in cases where no clear medical guidelines for the indicator values exist. We examine the simple, univariate case (one indicator) and then extend these ideas to the multivariate problem (several indicators) using one-class SVM to measure the distance from the reference population.**

I. INTRODUCTION

Use of portable and wearable sensors for measuring behavior introduces several challenges compared to usage of sensors for sports applications or even health applications in well-studied medical domains.

One of the problems is choosing *what* to measure. The raw signal can be processed to provide different information types about the individual's behavior, but the relevance of each information type is not always clear. For example, using triaxial accelerometer measurements for the assessment of physical activity, one can estimate (i) energy expenditure (METs, steps, etc), (ii) type of activity (categorical), (iii) intensity of specific physical activities, or even (iv) aggregate measures of physical activity level over a period of time. The relevance of each of these metrics will depend on the application. This problem is aggravated for complex human activities which are measured using several sensors, such as eating, as we will discuss in this paper.

Of equal importance is the problem of *interpreting* behavioral measurements. For example, assuming we use a device[1] to measure how fast one eats (in g/s), how do we know how fast is *too* fast? Again, this example can become even more complicated if we consider that eating behavior depends on food type [1], on gender [2], or even on social factors [3].

Ideally, studies of human behavior and medical knowledge would provide the answers to both questions; however this requires large-scale, longitudinal studies which in turn require significant resources and time. As new sensors and measurement technologies are developed every year (or perhaps faster [4]), solutions are needed in the present, even if the relevant medical knowledge is scarce or incomplete.

This paper describes the process that was followed towards solving these problems in the development of the decision support modules of the SPLENDID system [5], which aims at the prevention of obesity (OB) and eating disorders (ED) through eating and activity behavior monitoring and modification. A summary of the proposed steps is as follows: (i) Domain experts and technology developers compile a list of metrics that can be directly extracted from the raw signal data, denoted *behavioral indicators* or simply indicators. (ii) Sample data is collected from the target population. In our examples, the collected data corresponds to healthy individuals who may be at risk for developing OB or ED. (iii) An informal process is carried out for characterizing subject behavior by domain experts. (iv) Data is analyzed to select the indicators which are most relevant and discriminative for the domain experts. (v) The reference population data is used to build a model for measuring the "distance" from the reference population. (vi) Indicators are then interpreted on the basis of their distance to the reference population, either independently (univariate case) or jointly (multivariate case). These steps are discussed in more detail in the rest of the paper.

II. INDICATORS OF EATING AND ACTIVITY BEHAVIOR

A. Sensors and behavioral indicators

To measure in-meal eating behavior, SPLENDID uses the Mandometer®, developed by the Karolinska Institutet and Mando clinics[2]. Mandometer® is a medical device used successfully to treat eating disorders [6] and obesity [7]. Additionally, the Mandometer® has seen wide use as a research tool for the analysis of human eating behavior during single meals (e.g., [8]). SPLENDID also uses a chewing detection system to identify eating occurrences, developed by CSEM[3] and the Aristotle University of Thessaloniki [9]. It consists of an in-ear microphone to record chewing sounds and a photoplethysmography (PPG) sensor which detects movements of the jaw bone which correspond to chewing. These are placed at the ear using an ear-hook and are

*The work leading to these results has received funding from the European Community's ICT Programme SPLENDID under Grant Agreement No. 610746, 01/10/2013 - 30/09/2016 (splendid-program.eu).

[1]Multimedia Understanding Group, Information Processing Laboratory, Aristotle University of Thessaloniki, Greece

[2]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[1]https://mando.se/en/mandometer-method/the-mandometer-device/

[2]http://www.mando.se

[3]http://www.csem.ch

978-1-5090-4179-4/17/$31.00 ©2017 IEEE

(a) (b)

Fig. 1: The different sensors used by SPLENDID. (a) The Mandometer®. (b) Data logger with chewing sensors (audio and PPG) and accelerometer

TABLE I: A small subset (one third) of the eating-related indicators of SPLENDID

Name	**Units**
Meal duration	s
Total food intake	g
Average food intake rate	g/s
Initial food intake rate	g/s
Food intake deceleration	g/s^2
Average bite size	g
Bite size standard deviation	g
Average chewing rate	chews$/s$
Chewing rate standard deviation	chews

connected via wire to a data logger. Finally, the data logger includes a triaxial accelerometer which is used for assessing physical activity. The list of sensors used by SPLENDID is illustrated in Figure 1.

Signal processing and machine learning algorithms are used to automatically extract behavioral indicators (or simply indicators), i.e. the metrics that are used to quantify behavior. For the Mandometer®, the algorithms determine the food intake curve, i.e. the amount of food consumed at each time during the course of the meal and are described at [10]. For the chewing detection system, the algorithms aim at identifying snacks and meals during the day, and are described in [9]. Finally, the accelerometer is used to determine the physical activity level during different times of the day using state of the art algorithms [11].

A subset of in-meal eating behavior indicators is listed in Table I. Initially, a long list of 51 indicators was compiled for SPLENDID (27 out of which related to eating behavior). These were determined based on (i) what can possibly be measured with the sensors used and (ii) previous knowledge of the domain experts (e.g. the model of in-meal eating behavior of [8]).

B. *Selecting eating behavior indicators*

Individual indicator values can be used independently to evaluate specific behavioral aspects (e.g. for research purposes). However for health applications, such as measuring behavior for the prevention of obesity, not all indicators are equally important. In order to select the most relevant ones, we treated the problem of choosing indicators as a feature selection problem.

Initially, we recorded meals of 120 normal BMI individuals (summarized in Section IV). Then, an expert provided subjective assessments of the recorded meals based on his experience working with OB and ED patients. Meals were characterized in a scale from -2 (meals which resemble meals of ED patients) to $+2$ (as meals which resemble meals of OB patients). A value of 0 indicates that the meal does not resemble patient eating behavior. It is important to note that we did *not* use the meals to assess the health of the individuals, but to model the relative importance of indicators for the domain expert.

Based on these meal characterizations our goal is to identify which indicators are the most relevant for the discrimination between the "zero" meals (i.e. those that received a value of 0) and "non-zero" meals (i.e. those that received a non-zero assessment). To this end, we applied the minimum-Redundancy Maximum Relevance (mRMR) algorithm [12] for feature selection on the data. Based on the results, the four most relevant indicators proved to be "Food intake deceleration", "Initial food intake rate", "Total food intake" and "Average food intake rate" also shown in Table I. These indicators are later used for comparing individual behavior to the reference population (see Section III-B).

Given that the process of characterizing individual meals is subjective, it is important to emphasize that (i) the judgments provided by the expert are only used for estimating the relevance of individual indicators and not for health assessments and (ii) if data from clinically diagnosed patients of OB and ED were available, these could be used instead to contrast normal BMI behavior with OB or ED patient behavior without any need for subjective information.

III. Interpreting indicators

A. *Univariate case*

Interpretation of individual indicators ideally requires an estimate of the risk $P(D|X = x)$ that the subject *will* develop disease D (OB or ED in this case), given that the behavioral indicator X (treated as a random variable) has a value x. However, the precise relationship among eating related indicators and the risk of OB/ED development is not clear, mainly due to the multi-factorial nature of the problem [13].

An alternative option (weaker, but feasible in real life data collection conditions) is to quantify how much the indicator value deviates from the values measured in healthy individuals. To achieve this, we first estimate the probability density function $f_X(x|H)$ of indicator X from a reference population (where H indicates that the subject is healthy *now*). Then we use the cumulative density function $F_X(x|H) = P(X \leq x|H)$ to define the similarity measure

$$s_X(x) = \begin{cases} \frac{F_X(x|H)}{F_X(x_0|H)} & x \leq x_0 \\ \\ \frac{1-F_X(x|H)}{1-F_X(x_0|H)} & x > x_0 \end{cases} \quad (1)$$

where the value x_0 is chosen depending on the indicator. If we want to measure two-sided deviations from the behavior of a healthy population it is set to the mean, i.e. $x_0 = E(X)$.

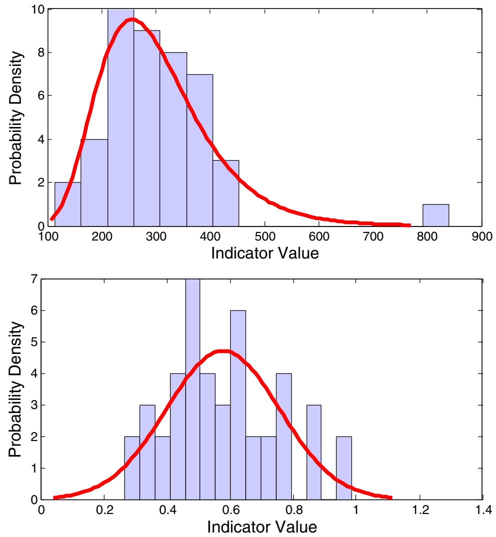

Fig. 2: Probability distribution functions of total food intake (top) and average food intake rate (bottom). The figures show the histogram of the data (blue bars) as well as the estimated PDF (red line).

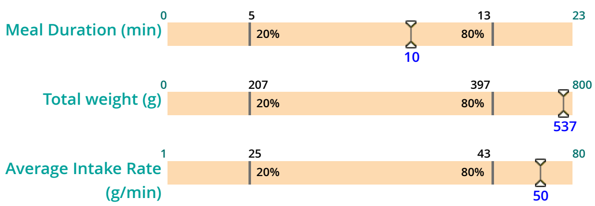

Fig. 3: The indicator values, as they are presented through the SPLENDID web tool. The position and percentile values on the bars correspond to $F_X(x)$ for each indicator X

Thus, a low value of $s_X(x)$ indicates that the subject's behavior deviates from the behavior of the reference group (subjects with normal BMI in our case).

The probability density function $f_X(x|H)$ is estimated based on a reference population of subjects with normal BMI and similar demographics with the target population. Depending on the indicator type, different families of probability density functions are used, such as (i) Gaussian (for symmetric PDFs), (ii) log-normal (when $x > 0$ and the PDF is non-symmetric), (iii) histograms (for categorical x), etc.

Figure 2 illustrates some example distributions which were derived from meals recorded from a reference population of normal BMI individuals at the Karolinska Institutet (dataset DT1 of Table II which provides details on the datasets used). Figure 3 illustrates how this approach is used to present the indicator value in the SPLENDID web tool.

B. Multivariate case

The procedure described in Section III-A assumes that each indicator provides sufficient information to assess the behavior of the individual. This can be true, for example, when assessing the average daily energy expenditure, where a single value (e.g. average daily METs or steps) is sufficient to determine how active the subject is. However these assumptions do not directly apply in the case of eating behavior, which is described by multiple dependent variables (Table I).

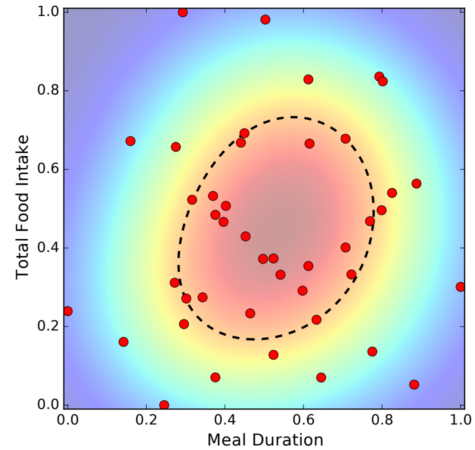

Fig. 4: Projection in two dimensions of the one-class SVM boundary learned from meals of a reference dataset of normal BMI individuals. Red color indicates high similarity to the reference population. The dashed line is the boundary learned by one-class SVM.

In the case of eating, the direct extension of Eq. (1) requires an estimate of the multivariate probability density function $f_{\mathbf{X}}(\mathbf{X} = \mathbf{x}|H)$ where $\mathbf{X} = [X_1, \ldots, X_N]^T$ and $\mathbf{x} = [x_1, \ldots, x_N]^T$. However such a PDF is hard to estimate, since the number of samples required increases exponentially with the number of dimensions N (i.e. the curse of dimensionality).

An alternative option is to use one-class Support Vector Machines (SVM) which computes a function $y(\mathbf{x}; \nu)$, $0 < \nu \leq 1$, that approximates a boundary of $f_{\mathbf{X}}(\mathbf{x}|H)$ containing at least probability mass $1-\nu$ (details on one-class SVM can be found in [14]). In other words, instead of computing the PDF (which is hard), we discover a boundary which contains a fraction of the probability mass and use this boundary to estimate the similarity $s_{\mathbf{X}}(\mathbf{x}) = y(\mathbf{x}; \nu)$, of a vector of indicator values $\mathbf{x}$ to the reference population. An example is provided in Figure 4 which illustrates sample data points and the selected boundary on the plane defined by two indicators (dataset DT1 of Table II, with values normalized to $[0, 1]$).

For indicators related to in-meal eating behavior, we restricted the models to use only the four most relevant indicators selected by the mRMR algorithm, as described in Section II-B.

IV. EXPERIMENTS

In order to demonstrate the value of the proposed approach in the multivariate case, we used the domain expert assessments described in Section II-B for evaluation purposes. The goal is to examine the agreement between the expert assessments and the assessments derived from the one-class SVM algorithm.

TABLE II: Datasets used for evaluation

Dataset	Description	Male/Female
DT1	44 meals from normal BMI healthy adults (age average: 22.84). Details in [15]	0/44
DT2	25 meals from adult individuals (age average: 30.01)	13/12
DT3	40 meals from Swedish adolescents collected at a high school students (age average: 16). Details in [16]	18/22
DT4	11 meals from adult females (age average: 22.8)	0/11

TABLE III: Results of evaluation of one-class SVM model trained on DT1

Test set	Accuracy	Sensitivity	Specificity	Precision	MAP
DT2	0.840	0.867	0.800	0.867	0.923
DT3	0.725	0.880	0.467	0.734	0.883
DT4	1.0	1.0	1.0	1.0	1.0

The datasets that were used in the experiment are summarized in Table II. For each recorded meal the domain expert examined the list of indicator values and the meal curve recorded by the Mandometer® and provided a subjective assessment of the meal (values from -2 to 2, as described in Section II-B). We formulate this problem as a binary classification problem, where the objective is to automatically identify which meals have "nonzero" values, i.e. which ones resemble meals eaten by patients with OB or ED, according to the expert.

We trained a one-class SVM model using dataset DT1 (44 meals) *without* taking into account the expert assessments, and applied the model in datasets DT2-4 (76 meals). Evaluation was based on the expert assessments for those meals. We used both classification metrics (accuracy, sensitivity, specificity, precision), where we just examined the sign of function $y(\mathbf{x};\nu)$, for $\nu = 0.5$, as well as mean average precision (MAP) to evaluate the ranking of results.

The results are shown in Table III, for each dataset. We can observe that for datasets with similar characteristics as the training dataset (young adult females, DT4) the results are excellent, while when the population type changes (high school males and females, DT3), performance drops. However it is important to highlight that even in these cases, MAP remains high, i.e. subjects characterized as "nonzero" by the expert are ranked higher than those characterized as "zero", even if the characterization based on $sign\left(y(\mathbf{x};\nu)\right)$ is incorrect. Thus, we can conclude that the proposed one-class SVM model reaches high agreement with the domain expert, despite the fact that it has been trained without any assessments on a dataset of normal BMI individuals.

V. Conclusions

We have outlined a methodology for selecting eating-related behavioral indicators and for producing meaningful measures of similarity to reference populations. This allows the measurement and interpretation of eating behaviors, whose relationship with the risk for developing obesity and eating disorders is not fully understood. In our experiments, the proposed data-driven method achieved high agreement with subjective assessments of a domain expert, which is an indication that it produces meaningful results. Given that the proposed methodology is not limited to the eating behavior domain, the next step is to evaluate its applicability and effectiveness for indicators related to other areas of human behavior.

References

[1] K. Kerameas, L. Vartanian, C. Herman, and J. Polivy, "The effect of portion size and unit size on food intake: Unit bias or segmentation effect?" *Health Psychology*, vol. 34, no. 6, 2015.

[2] B. Rolls, E. Morris, and R. L.S., "Portion size of food affects energy intake in normal-weight and overweight men and women," *The American journal of clinical nutrition*, vol. 76, no. 6, pp. 1207–13, Dec 2002.

[3] J. van den Boer and M. Mars, "Modeling of eating style and its effect on intake," *Appetite*, vol. 1, no. 86, pp. 3–18, Mar 2015.

[4] A. C. King, K. Glanz, and K. Patrick, "Technologies to measure and modify physical activity and eating environments," *American Journal of Preventive Medicine*, vol. 48, no. 5, pp. 630 – 638, 2015.

[5] C. Maramis, C. Diou, I. Ioakeimidis, I. Lekka, G. Dudnik, M. Mars, N. Maglaveras, C. Bergh, and A. Delopoulos, "Preventing obesity and eating disorders through behavioural modifications: The splendid vision," in *4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, Nov 2014, pp. 7–10.

[6] C. Bergh, M. Callmar, S. Danemar, M. Hölcke, S. Isberg, M. Leon, J. Lindgren, Å. Lundqvist, M. Niinimaa, B. Olofsson *et al.*, "Effective treatment of eating disorders: Results at multiple sites." *Behavioral neuroscience*, vol. 127, no. 6, p. 878, 2013.

[7] A. Ford, C. Bergh, P. Södersten, M. A. Sabin, S. Hollinghurst, L. Hunt, and J. Shield, "Treatment of childhood obesity by retraining eating behaviour: randomised controlled trial," *Bmj*, vol. 340, p. b5388, Jan 2010.

[8] I. Ioakimidis, M. Zandian, F. Ulbl, C. lund, C. Bergh, and P. Sdersten, "Food intake and chewing in women," *Neurocomputing*, vol. 84, pp. 31 – 38, 2012, from Neuron to Behaviour: Evidence from Behavioral Measurements.

[9] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos, "Robust dietary monitoring using a novel audio and ppg-based sensor," *IEEE Journal of Biomedical and Health Informatics*, 2016, (in press).

[10] V. Papapanagiotou, C. Diou, B. Langlet, I. Ioakimidis, and A. Delopoulos, "A parametric probabilistic context-free grammar for food intake analysis based on continuous meal weight measurements," in *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 7853–7856.

[11] A. Reiss, D. Stricker, and G. Hendeby, "Towards robust activity recognition for everyday life: Methods and evaluation," in *7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, May 2013, pp. 25–32.

[12] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, Aug 2005.

[13] L. Birch, J. Savage, and J. Fisher, "Right sizing prevention. Food portion size effects on children's eating and weight," *Appetite*, vol. 1, no. 88, pp. 11–16, May 2015.

[14] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[15] I. Ioakimidis, M. Zandian, C. Bergh, and P. Södersten, "A method for the control of eating rate: A potential intervention in eating disorders," *Behavior Research Methods*, vol. 41, no. 3, pp. 755–760, 2009.

[16] B. Langlet, A. Anvret, V. Papapanagiotou, C. Diou, I. Moulos, C. Maramis, E. Lekka, R. Heimeier, A. Delopoulos, and I. Ioakimidis, "Objective measures of behaviour in a Swedish high-school," in *Measuring Behaviour*, Dublin, Ireland, May 2016.