CONSIGLIO NAZIONALE DELLE RICERCHE

ISTITUTO DI SCIENZA E TECNOLOGIA
DELL'INFORMAZIONE

LABORATORIO SEGNALI E IMMAGINI

Criticality of human breath detection with a portable device II: Data Analysis

D. Germanese (ISTI- CNR)
M. Righi (ISTI- CNR)
M. D'Acunto (ISTI- CNR)
M. Guidi (IGG- CNR)
M. Magrini (ISTI- CNR)
P. Paradisi (ISTI- CNR)

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

EtOH	Ethanol
GC	Gas Chromatography
HME	Heat and Exchange Moisturizers
KNN	K nearest voting rule
min	minute
ml	milliliter
PC	Personal Computer
PCA	Principal Component Analysis
Ppm	parts per million
RH	Relative Humidity
s	seconds
SRAM	Static Random Access Memory
TCP	Transmission Control Protocol
USB	Universal Serial Bus
VOC	Volatile Organic Compounds
Vol	Volume
WS	Wize Sniffer

TABLE OF FIGURES

TABLE LEGENDS

1. INTRODUCTION

Human breath is largely composed of oxygen, carbon dioxide, water vapor, nitric oxide, and numerous volatile organic compounds (VOCs) [1, 2]. Changes in the concentration of the molecules in VOCs could suggest various diseases or at least changes in the metabolism. Indeed, breath gases are recognized to be excellent indicators of the presence of diseases and clinical conditions. Such gases have been identified as biomarkers using accurate but expensive benchtop instrumentations such as gas chromatography (GC) or electronic nose (e-nose) [1]. As a consequence, in recent years, it has been stimulated the necessity to develop a portable device for breath analysis, easy to use, and feasible for patients living far from medical structures or physicians.

In the framework of SEMEOTICONS (SEMEiotic Oriented Technology for Idividual's CardiOmetabolic risk self-assessmeNt and Self-monitoring) European Project, we developed a low cost, portable, easy-to- use device for the analysis of breath composition: the Wize Sniffer (WS). The WS captures breath samples, the chemical selective sensors sense the sample and accordingly form a sort of *odor-print* of healthy people or patients with known and specific diseases, in order to evaluate the *well-being* state of a human subject [3, 4]. It should be noted that does not exist a general definition of "well-being state", rather some indices for well-being that can be correlated to cardio-metabolic risk, which is representing the leading cause of worldwide mortality [3, 5].

The first prototype of such device is based on commercial, semiconductor-based gas sensors.

This type of gas sensors is very robust and easy to integrate.

Nevertheless, they are non- selective at all. This leads to several issues for data analysis. In this report we focus our attention on the different strategies for data analysis, evaluating also their performances and outcomes.

2. WIZE SNIFFER AND WELL-BEING STATE: VOCS TO BE DETECTED

The breath molecules detected by the Wize Sniffer are related to those noxious habits for cardio-metabolic risk, principally alcohol intake and smoking. Here is reported a brief overview about these VOCs.

Carbon Monoxide (CO): CO is present in cigarette smoke (it is the major component, 75,95%). Mean carbon monoxide concentration in exhaled breath is about 3,5 ppm. Increasing levels of exhaled carbon monoxide can be detected in smoking subjects: 13.8 - 29 ppm.

Carbon Dioxide (CO_2) and Oxygen (O2): Exhaled air has a decreased amount of oxygen and an increased amount of CO2. These amounts show how much O2 is retained within the body for use by the cells and how much CO2 is produced as a by-product of cellular metabolism. CO2 is also one of constituents of tobacco smoke (13%). Exhaled O2 amount is about 13.6%-16% while mean CO2 concentration in exhaled breath is about 4% (= 40000ppm). Lower values may be due to respiration disorders.

Ethanol: Exogenous Ethanol comes from alcoholic drink. It is important to note that it is recognized that ethanol breakdown leads to an accumulation of free radicals into the cells, a clear example of oxidative stress. Ethanol may cause arrhythmias and depresses the contractility of cardiac muscle.

Nitrogen (di)Oxide (NOx): it is a vasodilator and it modulates inflammatory response (operating in combination with CO and Hydrogen Sulfide). It is also a good indicator for asthma diseases.

Hydrogen Sulfide (H_2S): it is a vascular relaxant agent, and has a therapeutic effect in various cardiovascular diseases (myocardial injury, hypertension). In general, H_2S could have therapeutic effect against oxidative stress due to its capability to neutralize the action of free radicals.

Ammonia (NH_3): Mean baseline levels of ammonia in exhaled gas are about 0.42ppm. Elevated breath ammonia usually could be due to liver disease, such also to kidney disease. Ammonia is also one of the major compounds, together with CO, of tobacco fumes (approximately 22,15%).

Hydrogen (H2): The production of hydrogen is a metabolic consequence of carbohydrate fermentation by anaerobic bacteria. This hydrogen traverses the gut wall, and is transported via the circulation to the lungs, where is execrated in exhaled breath. Not only, a certain amount of exhaled hydrogen is the result of fermentation by oropharingeal bacteria. As a consequence, increased values of breath hydrogen may be due to overweight problems, intestinal diseases, improper life-style. The baseline value is about 9.1ppm.

3. THE WIZE SNIFFER: AN OVERVIEW ABOUT THE COMMERCIAL GAS SENSORS USED

The general scheme of the Wize Sniffer and its final configuration are shown in Figure 1.

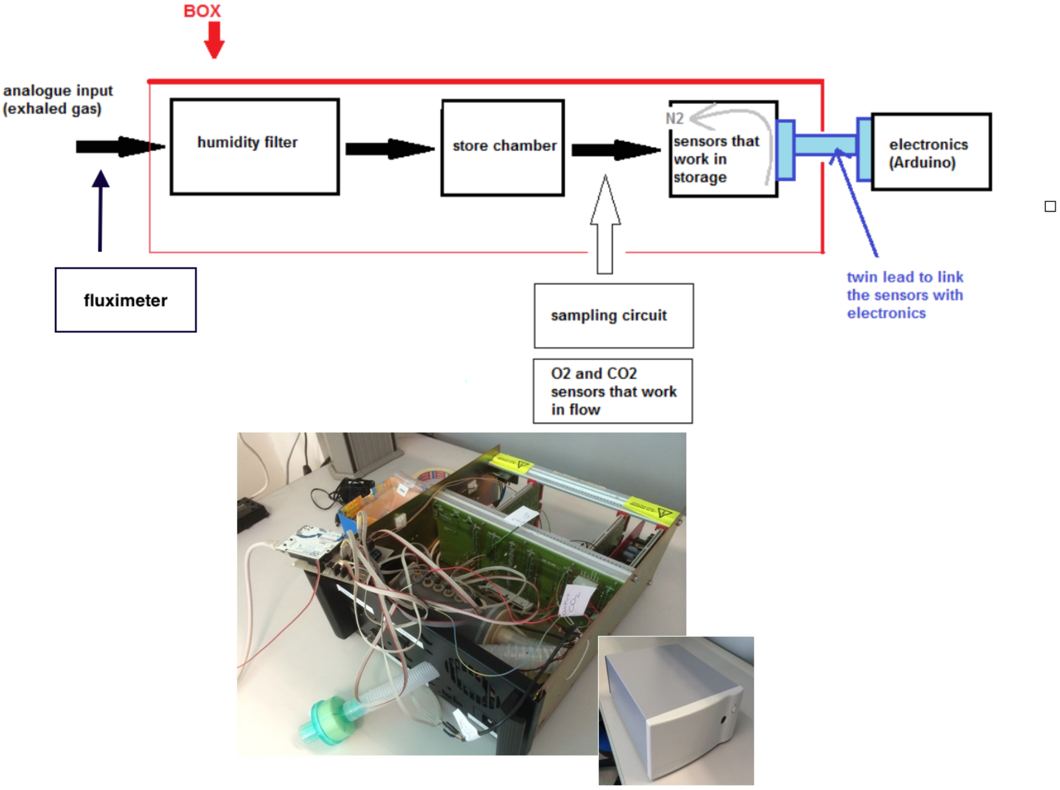


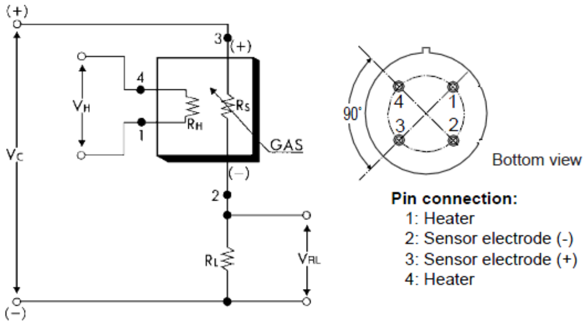
Figure 1. Schematic sketch of the WS system and its final configuration

The WS operates in three phases: gas collection, gas sampling, and data analysis (see Figure 2). It has a store chamber, a HME filter to absorb the water vapor from the breath, and a disposable mouthpiece. The gas sensors placed in the store chamber sense gas particles and generate measurable electronic signals, subsequently sent to the computer for further analysis. Temperature and humidity percentage are measured in situ in the store chamber. A flow-meter allows evaluating the exhaled breath volume. According to the block scheme in Figure 1, the individual breaths in a disposable mouthpiece at the tube's entrance. A heat and exchange moisturizers (HME) filter retains part of the humidity present in human exhaled breath. We provided for a corrugated tube to convey the exhaled gas in the store chamber, made of Delrin and ABS, where six gas sensors are placed. Two uni-directional valves allow the gases to enter the chamber and to be collected within it, preventing the mixing with ambient air. The store chamber's dimension was fixed at about 600ml, according to pulmonary capacity. The store chamber is provided also with a flushing pump which allows to purge the chamber itself, after each measuring. A sampling circuit injects, at a certain rate (=120 ml/sec, basing on [1]), the exhaled gas samples in a second chamber, where two gas sensors provided by COSMED s.r.l. (one sensitive to oxygen manufactured by CityTechnology, the other sensitive to carbon dioxide manufactured by Servomex, both working in flow regime) are placed. To link the gas sensors with other electronics, to send the collected data to Arduino micro controller board, and then to the computer, we thought about a kind of twin lead. An important aspect for the functionality of the WS is that the chamber embedding the sensors could be opened easily, in order to be able to change the sensors, if necessary. All the different steps during the design of the WS were considered at the light of this basic requirement.

Box dimensions provided for the WS 1.x family are: 30cm x 38cm x 14cm; its weight is 9 Kg. The device requires a voltage supply of 12V.

3.1 THE COMMERCIAL GAS SENSORS

The choice of gas sensors has been made taking into account the breath compounds to detect and the working principle of the sensors. The aim of WS is to detect those compounds, present in human exhaled gas, correlated to cardio-metabolic risk. As a consequence, our interest, during a first phase of the work, was focused on: oxygen, carbon dioxide, carbon monoxide, ammonia, hydrogen, ethanol, hydrogen sulfide. Criteria for choosing sensors' working principle, was based on their ease of use. So, we selected metal oxide semiconductor gas sensors. Their functioning is based on a variation of resistance from R_0 to Rs when the sensor senses the gas particles. An analogue voltage is measured as indirect measure of change in resistance (see Figure 2):



$$R_S = \frac{V_C - V_{RL}}{V_{RL}} \times R_L$$

Figure 2 The inner electrical circuit

Furthermore, basing on the study of Guo et al. [1], we chose chemical gas sensors manufactured by Figaro Engineering, because such kind of sensors are very robust, sensitive, resistant to humidity and ageing. We selected the following 6 gas sensors:

- **TGS 2602**: sensitive to hydrogen, ammonia, ethanol, hydrogen sulfide, toluene, cigarette smoke;
- **TGS 2620**: sensitive to hydrogen, carbon monoxide, ethanol, methane, isobutane;
- **TGS 4161 and its module AM-4-4161**: sensitive to carbon dioxide;
- **TGS 2442**: sensitive to carbon monoxide;
- **TGS 2444**: sensitive to ammonia;
- **TGS 821**: sensitive to hydrogen.

3.2 TEMPERATURE AND HUMIDITY SENSOR

The sensitivity characteristic of each sensor varies according to humidity percentage present in the chamber or, generally, in the environment where the sensors are placed (see Figure 3A). For this purpose, a temperature and humidity sensor, Sensirion SHT11 was integrated into the store chamber (Figure 3B).

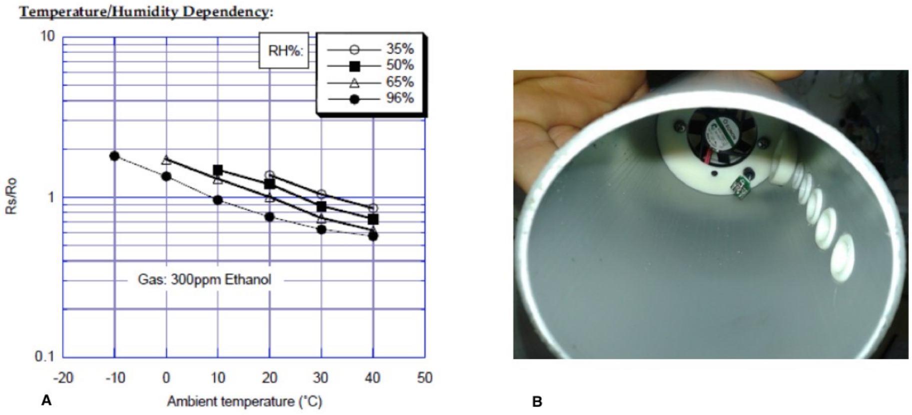


Figure 3A. - An example of the effects of temperature and humidity on gas sensors' sensitivity (TGS2602 in this case). Figure 3B- Sensirion SHT11 temperature sensor inside the store chamber.

Sensirion SHT11 has a humidity accuracy of about +/-3.5% and a temperature accuracy of about +/-0.5˚C@25˚C. It has a digital output, in contrast to other sensors (the gas ones) which have an analogue output.

3.3 CITYTECHNOLOGY AND SERVOMEX GAS SENSORS

Two additional sensors, MOX20 sensitive to **oxygen and** manufactured by CityTechnology, IR1507 sensitive to **carbon dioxide and** manufactured by Servomex, were inserted in our device. These sensors work in *flowing-regime*: they sense gases' stream that is injected at a constant rate (120ml/s) by the sampling pump. This electro-mechanic pump takes gas samples from the store chamber, injects them where the O2 and CO2 sensors are placed. Then, gases are brought back to store chamber.
About oxygen and carbon dioxide sensors' outputs, COSMED s.r.l. reported us the following informations:

- Warm-up time: 30min (for both sensors)
- Oxygen sensor's output: from 0 to 4V+/-0.01V @20.93% O2
- Carbon dioxide sensor's output:
- 0.02V +/-0.01V @ 0.03%
- 0.67V +/-0.01V @ 0.70%
- 1.75V +/-0.01V @ 2.00%
- 3.5V +/-0.01V @ 5.00%

Here, a table that summarizes the commercial sensors used for WS 1.0 and WS 1.1 and the molecules to be detected (Table 1).

Table 1. Sensors, detected molecules and optimal detection range.

SENSOR	DETECTED MOLECULE(S) / PHYSICAL QUANTITY	OPTIMAL DETECTION RANGE
FIGARO TGS2602	hydrogen, ammonia, ethanol, hydrogen sulfide	1-10 ppm
FIGARO TGS821	hydrogen	1-1000 ppm
FIGARO TGS2620	hydrogen, carbon monoxide, ethanol	50-5000 ppm
FIGARO TGS2442	ammonia	10-1000ppm
FIGARO TGS2444	carbon monoxide	10-100ppm
FIGARO TGS4161	carbon dioxide	350-10000ppm
Servomex IR1507	carbon dioxide	0%-5%
MOX20 CityTechnology	oxygen	20.93%+/-10%
SENSIRION SHT11	temperature and humidity	-40°C-123°C 0%-100%

4. WIZE SNIFFER PERFORMANCE TEST AND DATA ANALYSIS

The first prototype of the Wize Sniffer is based on commercial, semiconductor-based gas sensors.
This type of gas sensors, as we said in the previous section, is very robust and easy to integrate. Nevertheless, they are non- selective at all. This leads to several issues for data analysis.
For this purpose, first we tested each gas sensor singularly, making it sense the several substances it is sensitive to; then, the whole Wize Sniffer was tested using a population containing individuals of different age, habits, body type.

4.1 SINGLE TESTS

The only aim of these tests was to evaluate sensors' response on a quality level. That is, only evaluate their change, their rise in voltage outputs when sensing that/those gas/gases they are sensitive to.
Each sensor was placed, together with the temperature and humidity sensor (Sensirion SHT11) in an airtight desiccator. The substances, which the sensors are sensitive to, were placed, one by one, in the desiccator. The data from sensors were registered and saved by means of a personal computer.
Here an example of outcome is described, relative to the test conducted on TGS2602 (sensitive to hydrogen, ammonia, ethanol, hydrogen sulfide, toluene). It should be noted how the voltage values in response to ethanol and toluene rose faster than in response to ammonia. Similarly, the peak values relative to ethanol and toluene were higher than the one corresponding to ammonia. This is in line with the slopes of bilogarithmic curves mapping the sensor sensitivity characteristics for the several substances. Temperature and humidity values were constant at about, respectively, 23°C and 70%.

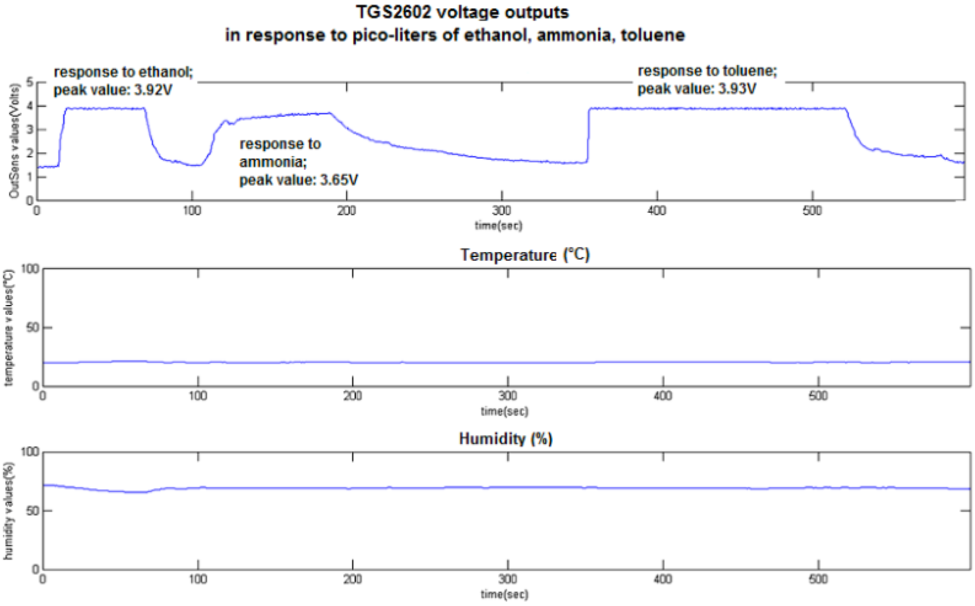


Figure 4. TGS2602 outputs in response to ethanol, ammonia, toluene concentrations on the order of pico-liters.

It should be noted how the voltage values in response to ethanol and toluene rose faster than in response to ammonia. Similarly, the peak values relative to ethanol and toluene were higher than the one corresponding to ammonia. This is in line with the slopes of bi-logarithmic curves mapping the sensor sensitivity characteristics for the several substances, as reported in the previous reports.

Temperature and humidity values were constant at about, respectively, 23 ° C and 70%.

4.2 TEST ON INDIVIDUALS, 1ST APPROACH FOR DATA ANALYSIS BASED ON A NON-LINEAR EQUATION MODEL

A second phase of our trials was focused on the evaluation of sensors' performances when used to sense human exhaled gases.

For this purpose, two important aspects were taken into account:

- the target population: the atherosclerotic illness develops insidiously, and clinical manifestations often become evident in its advanced stages. Altogether, the correlated diseases and complications occur between 40 and 60 years of age. As a consequence, since it's talking about prevention, we considered a population between 20 and 40 years of age, containing individuals with different habits (smokers/non smokers, sporty/sedentary lifestyle, healthy/unhealthy diet, teetotal/social drinker).
- the methodological issues about sampling procedure. In practice are often used three methods of sampling: "alveolar (end-tidal) sampling" (which corresponds to the plateau of the CO2 curve, that is the maximum value of exhaled CO2), "mixed expiratory air sampling" (which corresponds to a whole breath sample),"time-controlled sampling" (which corresponds to a part of exhaled air sampled after the start of expiration). Controlled alveolar sampling by means of expired CO2 concentrations is the method of choice if systemic volatile biomarkers are to be assessed, since only alveolar compounds are correlated to compounds in blood. The mixed expiratory air sampling, without controlled identification of the respiratory phases bears the risk of dilution with dead space air, but this method allows collecting also the compounds of exogenous origin. The last method shows large variations of compositions because of wide variations of individual dead space volumes and breathing maneuvers, and then it is less used in clinical practice. For our purposes, mixed expiratory air sampling method was chosen, since our interest was focused on endogenous biomarkers, but also to the compounds of exogenous origin (ethanol, for example). Additionally, since the composition of single breaths may vary considerably from each other, because of different modes and depth of breathing, in order to have, in this work phase, breath samples that were as reproducible as possible, we preferred a sampling of multiple breaths (three breaths).

The individuals were divided in two age-groups: [20-30 years old] and [30-40 years old]).

The tests were run following subjects' habits and lifestyle.

Basing on these considerations, the test allowed us to:

- Evaluate if the sensors were able to discriminate between two different alcoholic grades
- Evaluate if the sensors were able to discriminate different amounts of the same drink
- Assess if the sensors were able to follow the trend in time of the alcohol disposal
- Assess if the sensors were able to discriminate between non smoker's exhaled gas and smoker's exhaled gas
- Evaluate sensors' output in response to a subject's exhaled breath after smoking
- Evaluate sensors' output in response to a subject's exhaled breath after smoking and drinking alcoholic drink
- Evaluate sensors' output in response to:
- overweight subject's exhaled breath;
- normal body type subject's exhaled breath;
- subject's exhaled breath after exercise.

The sensors were placed in a bottle of 1,5lt.

The voltages output from the gas sensors, were sent to pc by means of a TCP/IP communication protocol and saved in text files (.txt).

The first approach for WS data analysis was based on a software, implemented on Matlab®, to calculate substances concentrations, basing on the sensors' input-output relation. The pseudocode is reported in the previous reports.

In particular, for each sensor the variation of its internal resistance is calculated, and according to the following non linear equation system (2) the concentration of each detected breath molecule is evaluated:

$$\begin{cases} (\frac{Rs}{Ro})_{2442} = A_{2442_{\mathrm{CO}}} [\mathrm{CO}]^{\alpha_{2442_{\mathrm{CO}}}} \\ (\frac{Rs}{Ro})_{821} = A_{821_{\mathrm{H_2}}} [\mathrm{H_2}]^{\alpha_{821_{\mathrm{H_2}}}} \\ (\frac{Rs}{Ro})_{2444} = A_{2444_{\mathrm{NH_3}}} [\mathrm{NH_3}]^{\alpha_{2444_{\mathrm{NH_3}}}} \\ (\frac{Rs}{Ro})_{2620} = A_{2620_{Eth}} [Eth]^{\alpha_{2620_{Eth}}} + A_{2620_{\mathrm{H_2}}} [\mathrm{H_2}]^{\alpha_{2620_{\mathrm{H_2}}}} + A_{2620_{\mathrm{CO}}} [\mathrm{CO}]^{\alpha_{2620_{\mathrm{CO}}}} \\ (\frac{Rs}{Ro})_{2602} = A_{2602_{Eth}} [Eth]^{\alpha_{2602_{Eth}}} + A_{2602_{\mathrm{H_2}}} [\mathrm{H_2}]^{\alpha_{2602_{\mathrm{H_2}}}} + A_{2602_{\mathrm{H_2S}}} [\mathrm{H_2S}]^{\alpha_{2602_{\mathrm{H_2S}}}} + A_{2602_{\mathrm{NH_3}}} [\mathrm{NH_3}]^{\alpha_{2602_{\mathrm{NH_3}}}} \end{cases}$$

(2) For more details, see the previous reports

This system resulted from three hypothesis:

for the "multi-sensing" sensors (TGS2620 and TGS2602), we made the hypothesis that the several contributions from the several substances add together linearly;

- TGS2620 is sensitive to methane and isobutane, too. Since the sensitivity range for methane and isobutane (50-5000ppm)is higher than the amounts of these substances present in our breath (less than 2 ppm), we made the hypothesis that their contributions were null;
- The same for TGS2602 which is sensitive to toluene, too. TGS2602's sensitivity range for toluene is 1-10ppm, while in our breath this compound is present in lower amounts (less than 1 ppm); so (for our purposes) we can considered null its contribution.

The assessment of the CO2 concentration, detected by TGS4161, results from the equation (3):

[CO2] = 1000 x Vout (3)

For CO2 and O2 sensors, the concentrations are assessed by means of the Table 3.

Table 2. Sensitivity of the CO2 sensor.

CO2	O2
0.02V +/-0.01V@0.03%	0 to 4V+/-0.01V@20.93%
1.75V +/-0.01V@2.00%	
3.5V +/-0.01V@5.00%	

The most meaningful results are reported in more detail in the previous reports, and the sensors' performances are discussed as well.

On a quality level, the commercial gas sensors gave reliable outputs in response to the several test conditions.

They were able to detect exhaled ethanol and discriminate among different alcoholic grades. Similarly, they were able to follow the trend in time of alcohol disposal.

Reliable outputs were registered also during the test for smoking habit: the sensors were able to distinguish between smoker subjects and non-smoker subjects.

Interesting results were collected also for exhaled hydrogen concentration values: they were in line with bibliography, showing an increase in case of subjects suffering overweight problems or after smoking, and a decrease in case of exercise or after drinking alcohol.

Concerning numerical results, that is, exhaled substances concentration values, some considerations have to be made. The observed concentration values corresponding to exhaled ethanol were generally in line with the ones output from the alcohol tester (Alcolino), even though we cannot conclude that there was a complete overlapping; on the contrary, there was not an overlapping with GAS Chromatography outcomes. Nevertheless, this is reasonable considering the different resolutions of the two approaches. Gas cromatography-mass spectrometry is the gold-standard for gas analysis, having very high resolution and sensitivity. Our sensors are affected by cross-sensitivity, and, if having multi-sensing sensors may be advantageous on one hand, because it allows detecting a wider set of analytes, on the other hand this affects the selectivity of the sensor system.

The comparison between the results obtained with our gas sensors and the ones output from piCO+ Smokerlyzer were rather satisfying. A relevant overlapping between the two approaches was observed, due to the specific transduction function of both the systems.

4.3 TEST ON INDIVIDUALS, 2ND APPROACH FOR DATA ANALYSIS BASED ON A NON-LINEAR EQUATION MODEL- COEFFICIENTS CALCULATED BY FACTOR ANALYSIS

The tests on individuals were repeated using the gas sensors placed in the Wize Sniffer's store chamber (made of, as reported before, ABS and Delrin).

In this case, the first approach used for WS data analysis (reported in the section 4.2) gave non-reliable results, probably because of:

- the change of measurement conditions;
- the cross-sensitivity (typical of semiconductor gas sensors) which affects the way the molecules add together;
- the different measurement conditions with respect to the ones in the sensors' datasheet.

As a consequence, we updated the non linear equation system (3) using *weights* calculated by factor analysis. Such factor analysis was implemented in order to calculate the weight that each input (that is, each molecule) has on each *"multi-substance"* sensor's output (that is, the change in resistance). This was implemented using R. Multiple R-squared and dusted R-squared values were always higher than 75%.

Here (table 4) are reported the *weights*, calculated by means of R, for TGS2602 and TGS2620. Thus, we can *numerically* estimate, for example, how the several molecules (hydrogen, ammonia, and ethanol) influence TGS2602's output.

Table 3. The weights, calculated by means of R, for TGS2602 and TGS2620.

TGS2602	TGS2620
W_H2 = 0.053	W_Eth= 6.5
W_NH3 = 0.092	W_CO= 0.161
W_Eth = 1.893	W_H2= 5.62
b= 0 intercetta nulla	b=0 intercetta nulla

By means of this approach we obtained results very similar to the ones obtained placing the sensors in a bottle. Nevertheless, a missing validation of these results, lead us to evaluate the use of another approach, described in the following section (section 4.4).

4.4 SEMEOTICONS 2ND ACQUISITION CAMPAIGN, 3RD APPROACH FOR DATA ANALYSIS BASED ON CLUSTERING AND CLASSIFICATION

During SEMEOTICONS 2nd acquisition campaign the Wize Sniffer 1.1 was tested on a population containing 26 subjects, male and female, with different habits. Their age belong to a range [30-60] .The tests followed subjects' habits and lifestyle.

In this case a new approach for data analysis was used.

First, a synthetic database containing WS data was created. In particular, this database contained:

21 non smoker, teetotal (or light drinker) subjects;

- 19 heavy drinker subjects;
- 15 very heavy drinker subjects;
- 19 light smoker subjects;
- 10 heavy smoker subjects;
- 30 subjects, with habits which were combinations of the previous four.

The new approach was based on:

- the study of the synthetic database by means of Principal Component Analysis, to evaluate if it was possible to reduce the dimensionality of the problem and to evaluate the presence of clusters;
- the use of clustering and classification algorithms, among which the K-Nearest Neighbor one.

Then, plotting the results of the analysis of the synthetic database by the PCA (Figure 4 and Figure 5), we saw several clusters, highlighted by means of an algorithm based on KNN (Figure 6).We saw the cluster plotted correctly according to increasing value of alcohol intake (that is, increasing voltages for TGS2602 and TGS2620), and according to increasing value of smoking habits (that is, increasing voltages for TGS2442).

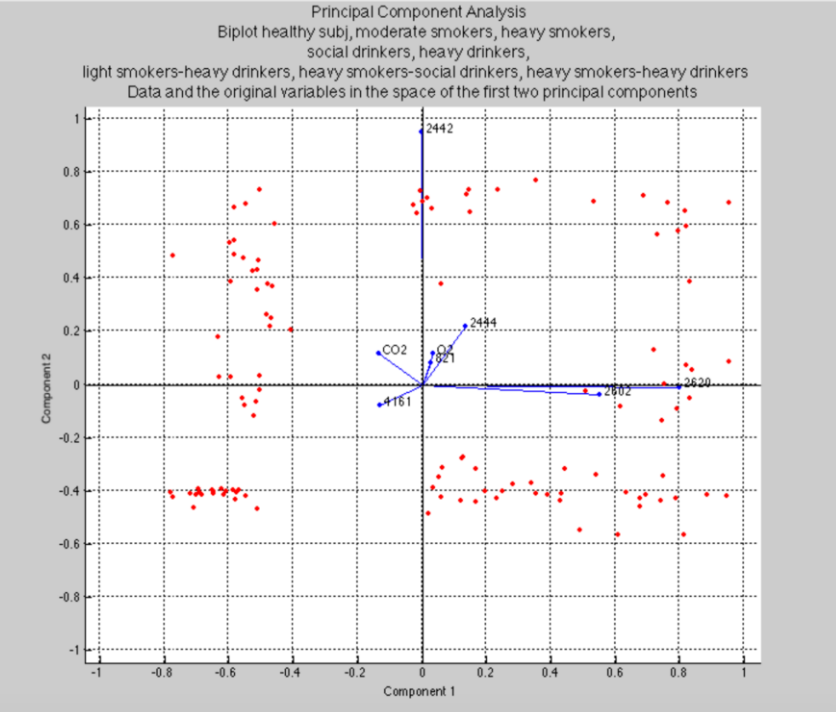


Figure 4- Biplot of the synthetic database analyzed by Principal Component Analysis

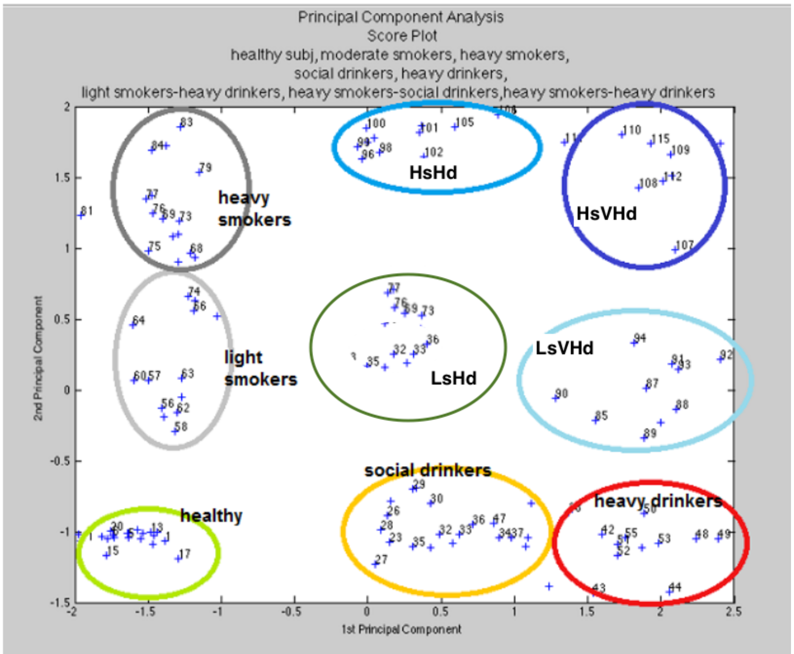


Figure 5. score plot of the synthetic database analyzed by Principal Component Analysis; LsHd = Light Smokers, heavy drinkers; LsVHd = Light Smokers, very heavy drinkers; HsHd= Heavy Smokers, heavy drinkers; HsVHd = Heavy Smokers, very heavy drinkers.

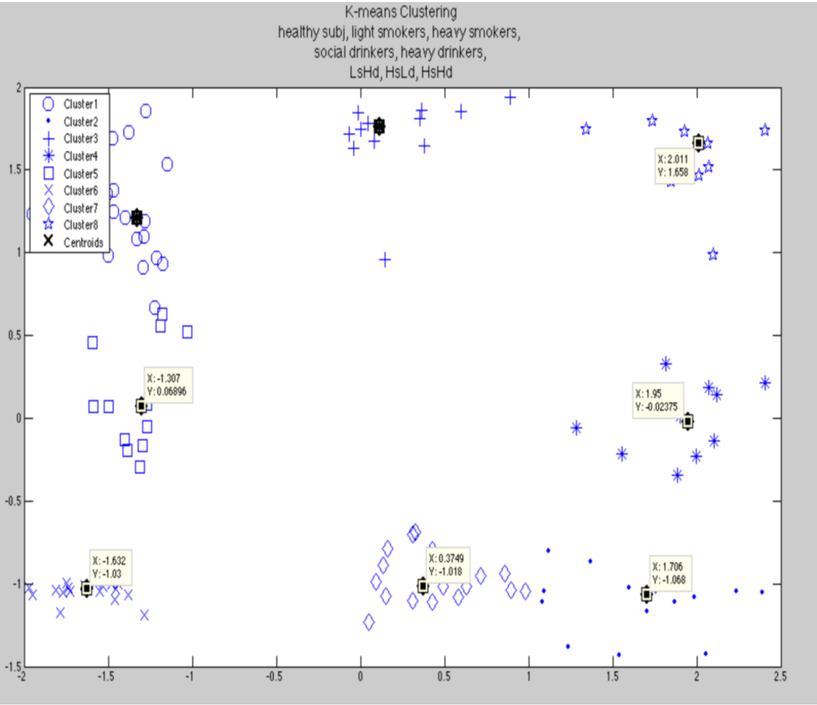


Figure 6. The several clusters plotted using KNN algorithm; we can see also the centers of the several centroids

Then, the data of the synthetic database were used as training set of KNN classifier, and the data collected during the SEMEOTICONS acquisition campaign as sampling set.
The results of the KNN classifier are shown in Table 5.

Table 4. WS data from SEMEOTICONS 2nd acquisition campaign classified by KNN classifier; "euclidean" and "cityblock" methods are compared. In the 4th column are reported subject's habits. The 5th column reports the comments (if the classifier is right/wrong).

SubjID	knn classify 'euclidean'	knn classify 'cityblock'	habits	comments
215	Healthy	Healthy	1 alcoholic unit/day	ok
218	Healthy	Healthy	1 alcoholic unit/2 times a week	ok
211	Healthy	Healthy	12-13 cigarettes/day; 1 alcoholic unit/ 4-5 times a week	wrong!
201	Healthy	Healthy	1 alcoholic unit/once a week	ok
207	Healthy	Healthy	1 alcoholic unit/1-2 times a week	ok
213	HD	Healthy	teetotal, non smoker	ok cityblock
208	Healthy	Healthy	1 alcoholic unit/ 3 times a week	ok
221	HD	HD	teetotal, non smoker	Wrong!
220	HD	HD	1 alcoholic unit/2 times a week	Wrong!
214	Healthy	Healthy	1 alcoholic unit < 1 times a week	ok
206	Healthy	Healthy	1 alcoholic unit/5-6 times a week	ok
205	Healthy	Healthy	1 alcoholic unit/ 3 times a week	ok
223	Healthy	Healthy	2 cigarettes/day; 1 alcoholic unit/ times a week	ok
212	Healthy	Healthy	1 alcoholic unit/day	ok
216	Healthy	Healthy	1 alcoholic unit/1 times a week	ok
217	Healthy	Healthy	teetotal, non smoker	ok
203	Healthy	Healthy	1 alcoholic unit < 1 times a week	ok
209	Healthy	Healthy	1 alcoholic unit < 1 times a week	ok
202	Healthy	Healthy	1 alcoholic unit/5-6 times a week	ok

210	Healthy	Healthy	1-2 alcoholic unit/ day	ok
225	Healthy	Healthy	4 cigarettes/day; 1 alcoholic unit < 1 times a week	Wrong!
224	Healthy	Healthy	1 alcoholic unit < 2 times a week	ok
204	HD	HD	1-2 cigarettes/day; 1 alcoholic unit < 2 times a week	Wrong!
219	Healthy	Healthy	1 alcoholic unit/2 times a week	ok
226	Healthy	Healthy	1 alcoholic unit/1 times a week	ok
222	Healthy	Healthy	10-15 cigarettes/day; 1 alcoholic unit/1 times a week	Wrong!

Among the volunteers for SEMEOTICONS 2nd acquisition campaign there were not heavy drinkers. Nevertheless there were 2 heavy smokers and 2 light smokers, but the device and, as a consequence, the classifier was not able to recognize them. This was due probably to the high detection range of TGS2442 (>50ppm, concentration typical for a subject used to smoke at least 20-30 cigarettes/day).

Anyway, the KNN classifier was able to classify correctly 20/26 subjects.

6. CONCLUSIONS

In this report, we have described the several approaches for Wize Sniffer data analysis, evaluating their performances and criticalities.

We can affirm that the first two analysis methods, although more innovative for such a device, present several problems related to:

- a lack of standardized procedures for sensors' characterization;
- a lack of results validation;
- sensors non- selectivity.

This impedes us to quantitatively calculate, for each detected molecule, its concentration.

For this reason, a more classical method has been used to process WS data, based on pattern recognition.

It is able to overcome the sensors non-selectivity from a quantitative point of view. In addition, being very simple, allows the integration of our device in a bigger multi-sensory platform.

REFERENCES

[1] Guo D.; Zhang D.; Li N.; Zhang L.; Yang J., 2007. *A Novel Breath Analysis System Based on Electronic Olfaction*, IEEE Transaction on Biomedical Engineering.

[2] Di Francesco F.; Fuoco R.; Trivella M.; Ceccarini A. 2005. *Breath Analysis: trends in techniques and clinical applications*, Micromechanical Journal, 79, 405-410.

[3] Lison M. W.; Carl H.; Rafael P.; Dan L.; David N.; Mant D., Glasziou P., 2010. *What are the basic self-monitoring components for cardiovascular risk management?* MC Medical Research Methodology, 10:105.

[4] Miekisch W.; Schubert J.; Nöeldge-Schomburg, G., 2004. *Breath analysis in critically ill patients: potential and limitations*, Expert Review Molecular Diagnosys , 4, 619-629.

[5] World Health Organization. The Global Burden of Disease: 2004 and update 2008.

[6] M. D'Acunto, A. Benassi, F. Chiellini, D. Germanese *et al.* Wize Sniffer: a new portable device designed for selective olfaction, SuperHeal Conference Proceedings, Angers, 4-6 March, 2014, pp. 577-582.

[7] D'Amico, A.; Di Natale C.; Paolesse R.; Macagnano A.; Martinelli E.; Pennazza G.; Santonico M.; Bernabei M.; Roscioni C.; Galluccio G. et al., 2007. *Olfactory systems for medical applications*, Sensors & Actuators: B-Chemical, 130, 458-465.

[8]D.R.Taylor, M.W.Pijnenburg, A.D.Smith, JCDe Jongste, 2006. Exhaled Nitric Oxide measurements: clinical application and interpretation. Thorax, 61:817-827.

[15] D.G. Thompson, P. Bin_eld, A. De Belder, J. O'Brien, S. Warren, 1985 *Extra intestinal influences on exhaled breath hydrogen measurements during the investigation of gastrointestinal disease*, Gut