rTREFEX: Reweighting norms for detecting changes in the response of MOX gas sensors

Sepideh Pashami[*], Achim J. Lilienthal, Erik Schaffernicht and
Marco Trincavelli

Centre for Applied Autonomous Sensor Systems, Örebro University, SE-70182. Orebro, Sweden.

[*] Corresponding author

Email: sepideh.pashami@oru.se

Abstract - The detection of changes in the response of metal oxide (MOX) gas sensors deployed in an open sampling system is a hard problem. It is relevant for applications such as gas leak detection in mines or large-scale pollution monitoring where it is impractical to continuously store or transfer sensor readings and reliable calibration is hard to achieve. Under these circumstances, it is desirable to detect points in the signal where a change indicates a significant event, e.g. the presence of gas or a sudden change of concentration. The key idea behind the proposed change detection approach is that a change in the emission modality of a gas source appears locally as an exponential function in the response of MOX sensors due to their long response and recovery times. The algorithm proposed in this paper, rTREFEX, is an extension of the previously proposed TREFEX algorithm. rTREFEX interprets the sensor response by fitting piecewise exponential functions with different time constants for the response and recovery phase. The number of exponentials, which has to be kept as low as possible, is determined automatically using an iterative approach that solves a sequence of convex optimization problems based on l_1-norm. The algorithm is evaluated with an experimental setup where a gas source changes in intensity, compound, and mixture ratio, and the gas source is delivered to the sensors exploiting natural advection and turbulence mechanisms. rTREFEX is compared against the previously proposed TREFEX, which already proved superior to other algorithms.

Keywords - MOX Sensor, Open Sampling System, Change Point Detection, Reweighted Norm Minimization.

1   Introduction

Detecting changes in the response of gas sensors directly exposed to the environment in an Open Sampling System (OSS) is beneficial for many applications like environmental monitoring [1], odour impact assessment [2], and gas leak detection in mines [3], [4]. Change points can be due to changes of the concentration of the analyte to which the sensors are exposed, changes of the compound to which the sensors are exposed, or changes in the mixture ratio of two or more compounds. Once identified, change points can then be used to reduce the amount of transferred data in wireless sensor network (transfer only the significant parts of the signal), and to ease the tasks of algorithms for gas discrimination [5] and quantification (segmenting the sensor signal into homogeneous parts). In OSS applications, detecting changes in the response of the gas sensor, which entails a change in the compound, gas concentration or mixture of compounds to which the sensor are exposed, is complicated by a variety of reasons. First, sensor calibration for OSS is very hard and the raw sensor response is loosely correlated to the absolute gas concentration level. Moreover, due to gas dispersal being dominated by turbulence, the sensor response is almost never stable, but continuously fluctuates. Therefore, it is infeasible to detect changes just setting a threshold on the sensor response.

Several methods have been proposed for solving the change detection problem in generic time series, which are inspired on a variety of disciplines. The simplest methods are threshold-based algorithms, which are often applied to quality control applications [6]. Statistical approaches, among which the Generalized Likelihood Ratio (GLR) test [7], Marginalized Likelihood Ratio (MLR) [8], CUmulative SUM (CUSUM) algorithm [7], are commonly used. More recently, the machine learning community adapted kernel methods like SVMs to be able to deal with the change point detection problem [9], [10]. Another category, to which the algorithm presented in this paper belongs, is methods based on trend detection, where change points are declared when the trend of the signal changes. Methods for discovering piecewise constant [11] and piecewise linear [12], [13] trends have been proposed.

rTREFEX, the algorithm presented in this paper, is an improvement on the previously proposed TREFEX algorithm [14]. The TREFEX algorithm starts from the observation that a sudden change in the exposition of the sensors appears as an exponential function in the response of MOX sensors due to their long response and recovery times. Therefore, the change point detection problem can be seen as the problem of fitting the *minimum number* of exponential functions to the sensor response and considering the kinks between exponentials as change points. However, this problem, since it requires to find the minimum number of exponential, is a cardinality problem, which is well known to be NP-hard. In order to obviate this, TREFEX approximates the l_0-norm involved in cardinality problems with an l_1-norm, that makes the problem convex and therefore tractable from the computational viewpoint. In order to limit the number of false alarms, TREFEX refines the solution doing a post-processing step, which depends on two parameters that have to be arbitrarily set. Here we propose to use a closer approximation to the l_0-norm than the l_1-norm, in order to improve the quality of the solution [15]. Since the proposed penalty function would make the problem non-convex, we solve the problem by iteratively solving a sequence of convex problems. The algorithm is evaluated with a set of experiments where a gas source is changing in intensity, emitted compound or mixture ratio between two compounds. This algorithm is compared with the original TREFEX algorithm which have already been demonstrated to outperform GLR in [14].

2   Algorithm

The problem of fitting a signal with the *minimum number* of exponential functions can be cast as a cardinality problem, i.e. a problem involving l_0-norm. It is well known that cardinality problems are NP-hard [16], and therefore computationally intractable. A common approach to deal with this kind of problems is to use proxies of the l_0-norm in order to make the problem tractable. The most commonly used relaxation of the l_0-norm is the l_1-norm, that results in convex relaxations of the original cardinality problem [12]. An l_1-norm based

algorithm (TREFEX) for piecewise exponential trend filtering to detect change points in the response of a MOX sensor (or an array of them) has been presented in [14]. The problem has been formulated as follows:

$$\underset{x,d_+}{\text{minimize}} \quad \|x - y\|_2^2 + \lambda\|(I + \tau_+D)d_+\|_1 + \lambda\|(I + \tau_-D)(Dx - d_+)\|_1 \tag{1}$$
$$\text{subject to} \quad d_+ \geq Dx$$
$$d_+ \geq 0$$

where y is the sensor response, x is the underlying trend to be estimated, D is the matrix operator that calculates first order differences. $\lambda \geq 0$ is a regularization parameter used to control the trade-off between the magnitude of the residuals $\|x - y\|_2^2$ and the smoothness of the signal encoded by $\|(I + \tau_+D)d_+\|_1 + \|(I + \tau_-D)(Dx - d_+)\|_1$. The parameters τ_+ and τ_- are the time constants of the response and recovery phases of the signal. The variable d_+, and the corresponding inequality constraints, are introduced to model the derivative of the trend differently for the response and recovery phase. The kinks between subsequent exponentials, i.e. the points where the l_1-norm regularizer is different from zero, are interpreted as change point candidates. Along the lines of [15], we propose in this paper an enhancement to the TREFEX algorithm, where we use a penalty function that approximates better the l_0-norm than the l_1-norm. This is beneficial for sharply identifying the change points. Figure 1 presents the penalty functions associated with the l_0-norm ($f_0 = 1|t \neq 0$), l_1-norm ($f_1 = |t|$), and the penalty function that will be used in the extension proposed here ($f_{log,\epsilon} = \log(1 + |t|/\epsilon)$). The penalty function f_0 penalizes equally all the values $t \neq 0$, while the function f_1 applies a penalty which is proportional to the magnitude of t, which is undesirable when solving cardinality problems. To alleviate this problem we propose here to use the $f_{log,\epsilon}$ as a better proxy of the l_0-norm than the l_1-norm. The $f_{log,\epsilon}$ function penalizes sharply already small values of t, while the penalty for large values is less than for l_1-norm. However, the $f_{log,\epsilon}$ penalty is not convex and therefore the resulting optimization problem is non-convex. In general, the solution of non-convex problems is difficult and computationally demanding. This problem can be addressed, as proposed in [15], by solving a (small) number of convex optimization problems. This method belongs to the family of Majorization-Minimization (MM) [17] algorithms, where the minimization of a generic function is performed by iteratively minimizing a convex majorizer function. In order to achieve this, the TREFEX algorithm (Eq. (1) is modified in the following way:

$$\underset{x,d_+}{\text{minimize}} \quad \|x - y\|_2^2 + \lambda\|W(I + \tau_+D)d_+\|_1 + \lambda\|W(I + \tau_-D)(Dx - d_+)\|_1 \tag{2}$$
$$\text{subject to} \quad d_+ \geq Dx$$
$$d_+ \geq 0$$

where W is a diagonal matrix that contains weights that are initially set to 1 and then updated at each iteration as follows:

$$W_{i,i}^{(l+1)} = \frac{1}{|k| + \epsilon} \quad i = 1,..,n \tag{3}$$

where k is the kink vector obtained for solution of the problem in the previous iteration:

$$k = \left\|W^{(l)}(I + \tau_+D)d_+\right\|_1 + \left\|W^{(l)}(I + \tau_-D)(Dx - d_+)\right\|_1 \tag{4}$$

and ϵ is a small value used to modulate the shape of the $f_{log,\epsilon}$ penalty. According to an heuristic presented in [15] we set $\epsilon = \max\{0.1 \times \sigma_{k>10^{-3}}, 10^{-3}\}$. At each iteration, after the update, the weight matrix $W^{(l)}$ is normalized so that the weights sum up to the number of samples to avoid changing the magnitude of the regularization term.

This weight update strategy corresponds to perform a first order Taylor expansion of $f_{log,\epsilon}$ at the solution of the previous iteration. From a practical perspective, Eq. (3, updates the weights such that entries in the kink vectors which are close to zero receive a higher weight and large entries in the kink vector receive a smaller weight. Hence, the importance of the magnitude of the entries in the kink vector k decreases with the iterations, and the result using the reweighted l_1-norm will be more similar to the l_0-norm than the l_1-norm.

4

3   Results

The proposed change point detection algorithm is evaluated on 54 indoor experiments where a gas source was placed 0.5 m upwind an array of 11 commercial MOX gas sensors. In these experiments, the gas source emits ethanol and/or 2-propanol. Given the distance of the gas source from the sensor, the gas is transported to the sensor by natural dispersion mechanisms, dominated by turbulence and advection. The experiments include different characteristic gas emission profiles with changes in concentration, compound and mixture. More details on the experimental setup can be found in [14]. As already mentioned, the kinks between subsequent exponentials are the *alarms*, and the real position of the change points can be inferred from the control signal of the gas source. This enables the calculation of *precision*, *recall* and *F-measure* for evaluating the proposed changed detection algorithm. Precision is a ratio of the number of true alarms to the number of change points. Recall is a ratio of the number of true alarms to the number of alarms. F-measure is the harmonic mean of precision and recall, and it is a commonly used measure for combining precision and recall. For different values of the regularization parameter λ, the maximum F-measure indicates a good trade-off point between precision and recall. An additional evaluation criterion we used is based on how close (in time) the true alarms are to the change points.

Figure 2 shows an example experiment where the sensors were exposed to changes in compound (ethanol to 2-propanol and vice versa) with the trend estimated by rTREFEX after several iterations. The trend estimated after the first iteration corresponds to the trend estimated by the TREFEX algorithm. It is possible to observe that the trend obtained using the reweighted l_1-norm estimates more accurately the kink points among the exponential trends, since it tends to favour sharper changes rather than multiple small changes. This reflects into a much higher precision, with a comparable recall.

3.1   Selection of the regularization parameter

The estimated trend clearly depends on the choice of the regularization parameter λ. In [14] an unsupervised strategy for λ selection (that does not rely on the knowledge of the true change points) has been proposed. We found that the λ selected with that strategy is a good value also in case of reweighted l1-norm. This means that the regularization parameter can be selected just performing one iteration (no reweighting) and then the subsequent iterations will be performed only for the selected λ, yielding to a significant save in computational resources.

3.2   The number of iterations

Figure 3 shows the precision-recall curves (average over all the experiment) for sensor MiCS5135 attained by the proposed algorithm after each iteration. It is clear that most of the improvement is obtained during the first 3-4 iterations, while after that the performance stabilizes. This fast convergence is most probably due to the fact that already the result of the first iteration (l_1-norm regularized problem) is a good starting point for searching the solution of the $f_{log,\epsilon}$ penalized problem.

Table 1 presents the maximum F-measure for all the sensors in the array w.r.t. the iteration number, and confirms the fact that after 3-4 iteration the performance of the algorithm does not improve significantly.

3.3   Overall performance

The overall performance of the proposed algorithm has been compared with the TREFEX algorithm [14], of which the proposed method can be considered a refinement. Table 2 shows the maximum F-measure for all sensors attained by both algorithms, showing a clear advantage for the reweighted l_1-norm minimization. It is worth noting that the ranking of the sensors according to the maximum F-measure is roughly the same. The time difference (in seconds) between the true alarms and the change points does not show any significant difference. It is also worth noting that rTREFEX does not need the post-processing step that was needed in TREFEX to reduce the number of false alarms. The TREFEX post-processing step has two parameters to be set arbitrarily, and therefore removing it has the beneficial effect of removing two parameters from the algorithm.

5

4   Conclusions

In this paper we introduced rTREFEX, an improvement to the TREFEX algorithm which is based on reweighted norm minimization. The idea behind the proposed algorithm is to use a non-convex penalty function which closely approximates the l_0-norm. The l_0-norm is the norm that can perfectly express cardinality problems, but it entails NP-hard optimization problems. The proposed algorithm was demonstrated to outperform the TREFEX algorithm, at the expense of a limited amount of additional computations. Indeed, the TREFEX algorithm has to be run for few iterations (~ 10) already for selecting the regularization parameter, and the reweighting just requires to run the minimization 2-3 more times. On the other hand, the introduction of reweighted minimization made redundant a post-processing step (which involves the arbitrary choice of two thresholds) that was needed in TREFEX for limiting the number of false alarms. It is worth noting that, despite the proposed algorithm has been designed explicitly for MOX sensors, it is potentially applicable to any sensor that can be modelled as a first order system. The rTREFEX is a batch algorithm, i.e. it requires a sliding window of input data prior to the algorithm's execution. However, the algorithm is computationally very efficient. Indeed its runtime for input windows less than 1000 samples is less than 250 ms (which is the sampling time in our experiments) and the algorithm can therefore be run online. Future work will be devoted to applications of change detection, since change detection provides a signal segmentation that can be used to improve estimation of gas concentration or classification of the gases (gas discrimination).

5 Bibliography

[1] S. Asadi, C. Badica, T. Comes, C. Conrado, V. Evers, F. Groen, S. Ilie, J. S. Jensen, A. Lilienthal, B. Milan, T. Neidhart, K. Nieuwenhuis, S. Pashami, G. Pavlin, J. Pehrsson, R. Pinchuk, M. Scafes, L. Schou-Jensen, F. Schultmann, and N. Wijngaards. In *Proceedings of the International Conference on Innovations in Sharing Environmental Observation and Information (EnviroInfo)* 4. Shaker Verlag **(2011)**, pp. 920–931.

[2] V. Naddeo, V. Belgiorno, and T. Zarra. *Odour Impact Assessment Handbook*. John Wiley & Sons, Inc., Hoboken, NJ, USA **(2012)**.

[3] X. R. Wang, J. T. Lizier, O. Obst, M. Prokopenko, and P. Wang. In *Wireless Sensor Networks*, LNCS, Springer, Berlin, Heidelberg **(2008)**, vol. 4913, pp. 90–105.

[4] M. Li, Y. Liu, and L. Chen. *ICDCS '07*. IEEE Computer Society, Washington, DC, USA **(2007)**, p. 9.

[5] M. Trincavelli, Künstliche Intelligenz **(2011)**, vol. 25, no. 4, pp. 351-354.

[6] L. A. Aroian and H. Levene, Journal of the American Statistical Association **(1950)**, vol. 45, no. 252, pp. 520-529.

[7] M. Basseville and I. V. Nikiforov, Detection of Abrupt Changes - Theory and Application, Upper Saddle River, NJ, USA: Prentice-Hall, Inc. **(1993)**.

[8] F. Gustafsson, IEEE Trans. Automatic Control **(1996)**, vol. 41, no. 1, pp. 66-78.

[9] F. Desobry, M. Davy and C. Doncarli, IEEE Transactions on Signal Processing **(2005)**, vol. 53, no. 8, pp. 2961-2974.

[10] Z. Harchaoui, F. Bach and A. Moulines, in *Proc. Neural Info. Proc. Systems*, Vancouver, British Columbia, Canada **(2008)**.

[11] Z. Harchaoui and C. Levy-Leduc, Journal of the American Statistical Association **(2010)**, vol. 105, no. 492, pp. 1480-1493.

[12] S.-J. Kim, K. Koh, S. Boyd and D. Gorinevsky, SIAM Review **(2009)**, vol. 51, no. 2, pp. 339-360.

[13] X. Huang, M. Matijas and J. Suykens, IEEE Transactions on Neural Networks and Learning Systems **(2013)**, vol. 24, no. 8, pp. 1279-1291.

[14] S. Pashami, A. J. Lilienthal, E. Schaffernicht and M. Trincavelli, Sensors **(2013)**, vol. 13, no. 6, pp. 7323-7344.

[15] E. J. Candes, M. B. Wakin and S. Boyd, Journal of Fourier Analysis and Applications **(2008)**, vol. 14, no. 5, pp. 877-905.

[16] B. K. Natarajan, SIAM J. Comput. **(1995)**, vol. 24, no. 2, pp. 227-234.

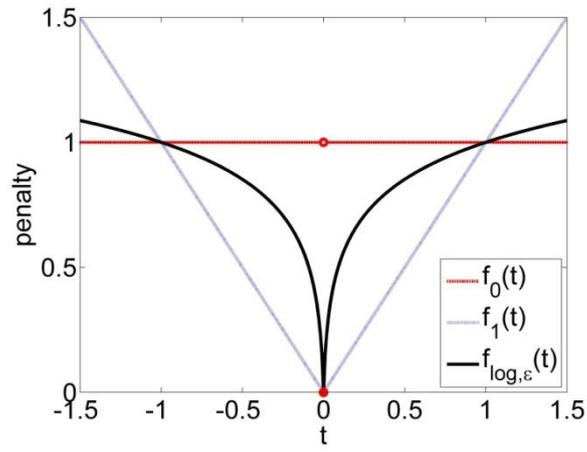[17] K. Lange, *Optimization*, Springer, New York, NY, USA **(2004)**.

Figure 1: Penalty functions f_0, f_1 and $f_{log,\epsilon}$ corresponding to the l_0-norm, l_1-norm and reweighted l_1-norm.
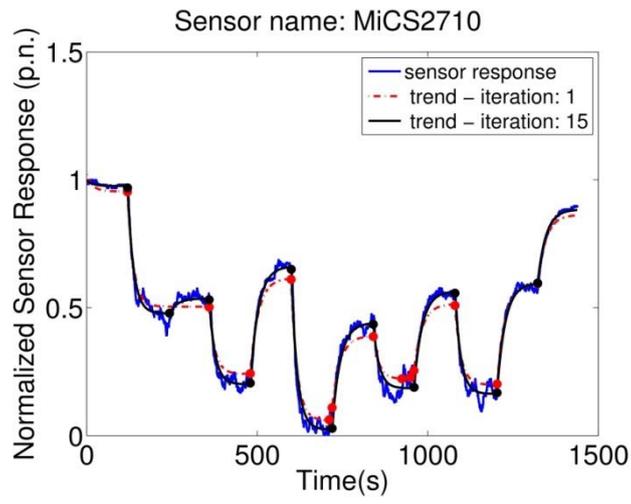
Figure 2: Comparing the estimated trend after the 1st iteration (TREFEX) to the trend estimated by rTREFEX after the 15th iteration. The regularization parameter is set to $\lambda = 16$. Alarms are marked with circles.
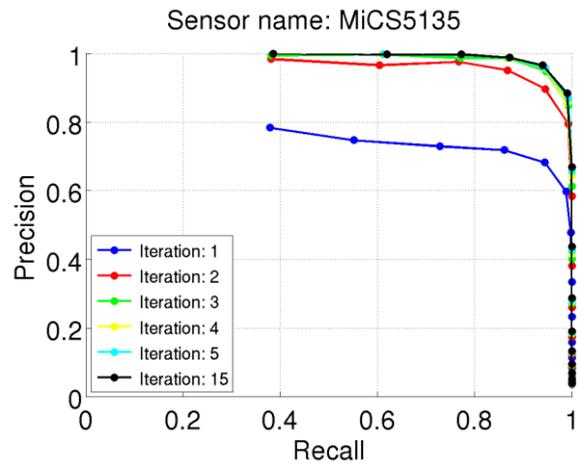
Figure 3: Precision-recall curve for the TREFEX algorithm (iteration 1) and for the rTREFEX at various iterations.

Table 1: Maximum F-measure at different iteration

Sensor	Iteration Number					
	1	**2**	**3**	**4**	**5**	**15**
MiCS2610	0.77	0.91	0.92	0.93	0.94	0.94
MiCS2710	0.81	0.91	0.95	0.94	0.96	0.96
MiCS5521-1	0.67	0.78	0.81	0.81	0.82	0.82
MiCS5121	0.77	0.87	0.90	0.91	0.91	0.93
MiCS5135	0.79	0.92	0.94	0.95	0.95	0.95
MiCS5521-2	0.68	0.82	0.84	0.85	0.86	0.87
TGS2600-1	0.74	0.84	0.87	0.87	0.87	0.88
TGS2611	0.75	0.85	0.87	0.88	0.89	0.90
TGS2620	0.72	0.82	0.85	0.85	0.85	0.86
TGS2600-2	0.75	0.83	0.86	0.87	0.87	0.87
TGS2602	0.70	0.79	0.83	0.82	0.83	0.82

Table 2: Comparing the maximum F-measure and time difference of rTREFEX to TREFEX

Sensor	rTREFEX		TREFEX	
	Max. F.	TD	Max. F.	TD
MiCS2610	0.94	3.75	0.90	3.72
MiCS2710	0.96	4.09	0.95	4.78
MiCS5521-1	0.82	5.50	0.77	5.59
MiCS5121	0.93	4.78	0.88	5.19
MiCS5135	0.95	3.67	0.92	4.17
MiCS5521-2	0.87	5.53	0.81	5.05
TGS2600-1	0.88	7.29	0.83	6.05
TGS2611	0.90	6.04	0.84	5.51
TGS2620	0.86	6.41	0.80	5.89
TGS2600-2	0.87	6.96	0.83	6.40
TGS2602	0.82	8.52	0.78	6.57