Active Chemical Sensing With Partially Observable Markov Decision Processes

Rakesh Gosangi and Ricardo GutierrezOsuna

Citation: AIP Conference Proceedings **1137**, 562 (2009); doi: 10.1063/1.3156617
View online: http://dx.doi.org/10.1063/1.3156617
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1137?ver=pdfcov
Published by the AIP Publishing

Active Chemical Sensing With Partially Observable Markov Decision Processes

Rakesh Gosangi and Ricardo Gutierrez-Osuna*

Department of Computer Science, Texas A & M University
{rakesh, rgutier}@cs.tamu.edu

Abstract. We present an active-perception strategy to optimize the temperature program of metal-oxide sensors in real time, as the sensor reacts with its environment. We model the problem as a partially observable Markov decision process (POMDP), where actions correspond to measurements at particular temperatures, and the agent is to find a temperature sequence that minimizes the Bayes risk. We validate the method on a binary classification problem with a simulated sensor. Our results show that the method provides a balance between classification rate and sensing costs.

Keywords: Active sensing, Chemical sensors, and Partially Observable Markov Decision Processes.
PACS: 07.07.Df

I. INTRODUCTION

Previous research has shown that modulating the working temperature of metal-oxide sensors can give rise to gas-specific temporal signatures that provide a wealth of discriminatory and quantitative information [1]. A number of empirical studies with various temperature waveforms (e.g. rectangular, sine, saw tooth, and triangular) and stimulus frequencies have been published [2-4], but only a handful of authors have approached the problem in a systematic fashion. Kunt et al. [5] developed a computational method to optimize the temperature profile in binary discrimination problems. The authors used a wavelet network to obtain a dynamic model of the sensor from experimental data, followed by an optimization procedure that found the temperature profile that maximized the distance between the two gas signatures. More recently, Vergara et al [6] proposed a system-identification method for optimizing temperature profiles. In their method, a pseudo-random binary sequence was used to drive the sensor heater while the sensors were exposed to various chemicals. The authors then estimated the frequency response of the sensor to each individual chemical, and selected a subset of the most informative frequencies. Both approaches, however, required that the temperature program be optimized off-line. Here we propose an active-sensing approach that can optimize the temperature profile on the fly, that is, as the sensor collects data from its environment. The method can also determine when sensing should be terminated in order to make a final classification; this is achieved by comparing the cost of measuring the sensor response at additional temperatures against the expected reduction in Bayes risk from those additional measurements. These capabilities are important not only to improve detection performance, but also to meet the increasing power constraints of real-time embedded applications as well as extend sensor lifetimes.

We model the problem as a decision-theoretic process, where the goal is to determine the next temperature pulse to be applied to the sensor based on information extracted from the sensor response to previous temperature pulses. Our method operates in two stages. First, we model the dynamic response of the chemical sensor to a sequence of temperature pulses as an Input-Output Hidden Markov Model (IOHMM) [7]. Then, we formulate the process of finding the ideal sequence of temperature pulses as a POMDP [8]. By assigning a cost to each temperature pulse and a cost for misclassifications, the POMDP is able to balance the total number of temperature pulses against the uncertainty of the classification decisions.

The paper is organized as follows. In section II we formulate the problem and show how IOHMMs can be used to model the dynamic response of a sensor. Section III describes the optimization of temperatures as an active sensing problem with POMDPs. Section IV provides experimental results on a dataset from a simulated metal-oxide sensor. The article concludes with a brief discussion and directions for future work.

CP1137, *Olfaction and Electronic Nose: Proceedings of the 13th International Symposium*, edited by M. Pardo and G. Sberveglieri

II. PROBLEM STATEMENT

Consider the problem of classifying an unknown gas sample into one of M known categories $\{\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(M)}\}$ using a metal-oxide sensor with D different operating temperatures $\{\rho_1, \rho_2, \ldots, \rho_D\}$. To solve this sensing problem, one typically measures the sensor's response at each of the D temperatures, and then analyzes the complete feature vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_D]^T$ with a pattern-recognition algorithm [9]. Though straightforward, this "passive" sensing approach is unlikely to be cost-effective because only a fraction of the measurements are generally necessary to classify the chemical sample. Instead, in active classification we seek to determine an optimal sequence of actions $\boldsymbol{a} = [a_1, a_2, \ldots, a_T]$, where each action corresponds to setting the sensor to one of the D possible temperatures (or terminating the process by assigning the sample to one of the M chemical classes). More importantly, we seek to select this sequence of actions dynamically, based on accumulating evidence. Our proposed solution to this problem is based on Ji and Carin [10].

A. Modeling the Sensor

Given a chemical from class $\omega^{(c)}$, we model the steady-state response of the sensor at temperature ρ_i with a Gaussian mixture:

$$p(x_i|\omega_c) = \sum_{m_i=1}^{M_i} \alpha_{i,m_i}^{(c)} N\left(x_i \middle| \mu_{i,m_i}^{(c)}, \Sigma_{i,m_i}^{(c)}\right) \quad (1)$$

where M_i is the number of mixture components, and $\alpha_{i,m_i}^{(c)}, \mu_{i,m_i}^{(c)}, \Sigma_{i,m_i}^{(c)}$ are the mixing coefficient, mean vector and covariance matrix of each mixture component for class $\omega^{(c)}$, respectively. Given a sequence of actions $[a_1, a_2, \ldots, a_T]$, we assume that the sensor transitions through a series of states $\boldsymbol{s} = [s_1, s_2, \ldots, s_T]$ to produce a corresponding observation sequence $\boldsymbol{o} = [o_1, o_2, \ldots, o_T]$. Each state s_i represents a mixture component in eq. (1) and is therefore hidden. Following Ji and Carin [10], we model the sensor dynamics with an IOHMM, a generalization of the traditional hidden Markov model (HMM) [11]. An IOHMM conditions the next state in a sequence not only on the previous state (as in a first-order HMM) but also on the current input to the sensor. In our case, this additional input consists of sensing actions (i.e. temperature steps).

Formally, an IOHMM can be defined as a 6-tuple $\{S, A, O, \pi, \tau, \phi\}$ where S is a finite set of states, each state corresponding to a mixture component in eq. (1), A is a finite set of discrete actions, each action corresponding to selecting one of D sensor temperatures, O is a set of observations, each corresponding to the sensor's response at a given temperature, $\pi(s)$ is the initial state distribution, $\tau(s'|s,a)$ is the state transition function, which describes the probability of transitioning from state s to state s' given action a, and $\phi(o|s)$ is the observation function, which describes the probability of making observation o at state s. We train a separate IOHMM for each individual chemical class, i.e. by driving the chemical detector with a random sequence of actions in the presence of the chemical, and recording the corresponding responses; for details see [7].

III. ACTIVE CHEMICAL SENSING AS A POMDP

We define a POMDP as a 7-tuple $\{S, A, O, b_0, T, \Omega, C\}$, where S, A, and O are the finite set of states, actions and observations from the IOHMMs respectively, $b_0(s)$ is an initial belief across states, $T(s'|s,a)$ is the probability of transitioning from state s to state s' given action a, $\Omega(o|s)$ is the probability of making observation o at state s, and $C(s,a)$ is the cost of executing action a at state s. These POMDP parameters can be obtained directly from the IOHMM as follows:

- Initial belief: $b_0(s) = p(\omega^{(c)})\pi^{(c)}(s)$; $s \in S^{(c)}$
- State transition: $T(s'|s,a) = \tau^{(u)}(s'|s,a)$; $s, s' \in S^{(u)}$; zero otherwise[1].
- Observation model: $\Omega(o|s) = \phi^{(c)}(o|s)$; $s \in S^{(c)}$

The POMDP stores information about the state of the system in a belief state $b_T(s)$, a probability distribution (across states from all the IOHMMs) *given* the initial belief $b_0(s)$ and the history of actions $[a_1 \ldots a_T]$ and observations $[o_1 \ldots o_T]$:

$$b_T(s) = p(s|o_1 \ldots o_T, a_1 \ldots a_T, b_0) = p(s|o_T, a_T, b_{T-1}) \quad (2)$$

The second equality above reflects the fact that $b_T(s)$ is a *sufficient statistic* for the history of the system, which allows us to update $b_T(s)$ incrementally from its previous estimate $b_{T-1}(s)$ by incorporating the latest action a_T and observation o_T:

$$b_T(s') = \frac{p(o_T|s', a_T)\sum_S p(s'|a_T, s) b_{T-1}(s)}{p(o_T|a_T, b_{T-1})} = \frac{\Omega(o|s')\sum_S T(s'|s,a) b_{T-1}(s)}{\mu} \quad (3)$$

where the denominator $\mu = p(o_T|a_T, b_{T-1})$ can be treated as a normalization term to ensure that $b_T(s')$ sums up to 1, and all terms in the numerator are known from the POMDP/IOHMM model.

[1] This ensures that transitions from the IOHMM of one class onto another class are not allowed, since we assume that the chemical stimulus does not change over time.

Using this POMDP formulation, the problem becomes one of finding a policy that maps belief states into actions so as to minimize the expected cost of sensing. We consider two types of actions:

- Sensing actions ($a = \rho_i$), which correspond to setting the sensor to temperature ρ_i. Sensing actions have a cost of $c(s, a = \rho_i) = c_i$, which reflects the fact that certain temperatures may be more expensive (e.g. draw more power).
- Classification actions ($a = \hat{\rho}_c$), which assign the sample to a particular class. Classification actions are terminal; their cost is $(s, a = \hat{\rho}_u) = c_{uv}$ $(\forall s \in S^{(v)})$, which represents a misclassification penalty whenever $u \neq v$.

A. Finding the Sensing Policy

Unfortunately, the problem of finding an exact solution for a POMDP policy is P-SPACE complete and therefore intractable for most problems. Moreover, a standard POMDP solution allows repeated actions (measuring the response of the sensor at the same temperature multiple times), which is undesirable in our case. For these reasons, we employ a myopic policy [10] that only takes sensing action if the cost of sensing (c_i) is lower than the expected future reduction in Bayes risk. Given belief state $b_T(s)$, the expected risk of a classification action is:

$$R_C\big(b_T(s)\big) = \min_u \sum_v c_{uv} \sum_{s \in S^{(v)}} b_T(s) \quad (4)$$

where u corresponds to the class with minimum Bayes risk $(\sum_v c_{uv} \sum_{s \in S^{(v)}} b_T(s))$. In turn, the expected risk of a sensing action is:

$$R_S(b_T(s), a) = \sum_{\forall o} \min_u \Big(\sum_v c_{uv} \sum_{s' \in S^{(v)}} \sum_s p(o|s', a) p(s'|s, a)\, b_T(s)\Big) \quad (5)$$

which averages the minimum Bayes risk over all observations that may result from the action. Hence, the utility of sensing action a can be computed as:

$$U(b_T(s), a) = \big[R_C\big(b_T(s)\big) - R_S(b_T(s), a)\big] - c_a \quad (6)$$

If $U(b_T(s), a) < 0$ for all sensing actions, then the sensing costs exceeds the expected reduction in risk $[R_C(\cdot) - R_S(\cdot)]$, and a classification action is taken. Otherwise, the action with maximum utility is taken.

IV. EXPERIMENTAL RESULTS

We validated the method on a synthetic dataset of metal-oxide sensor responses. Following [12], we modeled the temperature-conductance response using a Gaussian function. We also modeled the sensor dynamics with a first-order linear filter, resulting in:

$$G(T(t)) = \alpha G(T(t-1)) + (1-\alpha)\Big(k_1 e^{-\left(\frac{T(t)-T_0}{\sigma}\right)^2} + k_2 T(t)\Big) \quad (7)$$

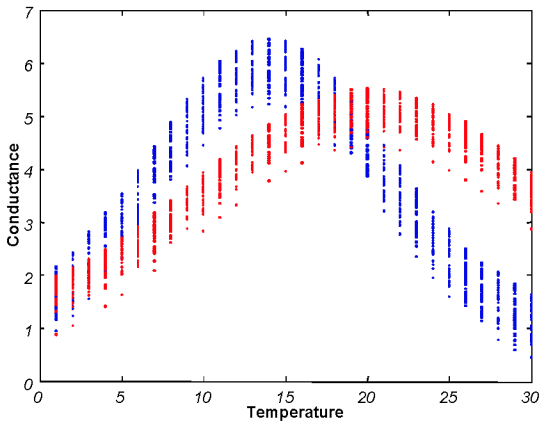

FIGURE 1. Conductance versus temperature for the two chemical classes.

where $T(t)$ is the sensor temperature at time t, and $G(T(t))$ is the conductance of the sensor at temperature $T(t)$, T_0 is the temperature at which the sensor conductivity is maximum, k_1, k_2 and σ are parameters that capture the steady-state properties of the sensor, and α captures history effects.

We evaluated the method on a problem with two chemicals and a sensor with 30 different temperatures. Sensor parameters were as follows: $\alpha = 0.2$, $k_1 = 6.0$ and $k_2 = 0.2$ for both classes; $\sigma = 10$ and $T_0 = 14$ for $\omega^{(1)}$; $\sigma = 15$ and $T_0 = 20$ for $\omega^{(2)}$. The temperature-dependent response of the sensor to the two chemicals is shown in FIGURE 1. These results were obtained by running the sensor with a random temperature sequence and recording the corresponding responses; thus, the spread at each temperature illustrates the effect of the sensor dynamics.

Training data for each analyte consisted of 40 random temperature sequences, 60 temperature pulses per sequence. Two IOHMMs (one per class) were trained; the number of Gaussian components in eq. (1) was set to $M_i = 4$. FIGURE 2 shows IOHMM predictions against the sensor response in eq. (7). These results show that the IOHMM can capture the temperature dependence and dynamics of the sensor.

The model was tested on 80 samples, 40 from each class. Each sample was generated by randomly selecting an initial temperature $T(0)$ unknown to the POMDP, and initializing the sensor response to $G(T(0)) = \big(k_1 e^{-((T(0)-T_0)/\sigma)^2} + k_2 T(0)\big)$. Classification costs c_{uv} were assumed uniform ($c_{uv} = 1$ if $u \neq v$; 0 otherwise). FIGURE 3 shows classification rate and average length of the temperature sequence as a function of feature acquisition costs c_i. For $c_i = 0.025$, the system achieves 100% classification rate with an average sequence length of 2.9 temperatures. For $c_i = 0.5$ the system performs at chance level (50%), and essentially produces a classification after measuring the response at a single temperature –this

happens because sensing costs become too high compared to misclassification costs. Between these two extremes, the POMDP provides a balance between sequence length and classification rate: as feature acquisition costs increase relative to misclassification costs, the POMDP selects increasingly shorter sequences at the expense of classification rates.

V. DISCUSSION AND CONCLUSION

We have presented an active sensing approach for metal-oxide sensors that is capable of selecting operating temperatures in real-time. The problem is formulated as one of sequential decision making under uncertainty, and is solved by means of a POMDP. We have validated the method on a binary classification problem using synthetic data from a computational model of metal-oxide sensors that captures their temperature-selectivity dependence and history effects. Our results show that the POMDP is able to balance sensing costs and classification accuracy: higher classification rates can be achieved by increasing the length of the temperature sequence. The method also appears to be robust to the particular choice of sensing and classification costs, since classification rates degrade smoothly as a function of these parameters. Future work will validate the method using experimental data. The results presented here assumed uniform sensing costs, but the method can also be used to penalize high temperatures, and as a result reduce power consumption and increase sensor lifetime.

REFERENCES

1. A. P. Lee and B. J. Reedy, "Temperature modulation in semiconductor gas sensing," *Sensor Actuator B,* vol. 60, pp. 35-42, 1999.
2. R. Gutierrez-Osuna, S. Korah, and A. Perera, "Multi-Frequency Temperature Modulation For Metal-Oxide Gas Sensors," in *Proc. ISOEN*. Washington, DC, 2001.
3. A. Ortega et al., "An intelligent detector based on temperature modulation of a gas sensor with a digital signal processor," *Sensor Actuator B,* vol. 78, pp. 32-39, 2001.
4. X. Huang et al., "Gas sensing behavior of a single tin dioxide sensor under dynamic temperature modulation," *Sensor Actuator B,* vol. 99, pp. 444-450, 2004.
5. T. A. Kunt et al., "Optimization of temperature programmed sensing for gas identification using micro-hotplate sensors," *Sensor Actuator B,* vol. 53, pp. 24-43, 1998.
6. A. Vergara et al., "Optimized temperature modulation of micro-hotplate gas sensors through pseudorandom binary sequences," IEEE *Sensors J,* vol. 5, pp. 1369-1378, 2005.
7. S. Fu et al., "Audio/visual mapping with cross-modal hidden Markov models," *IEEE Trans Multimedia,* vol. 7, pp. 243-252, 2005.
8. L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artif Int,* vol. 101, pp. 99-134, 1998.
9. R. Gutierrez-Osuna, "Pattern Analysis for Machine Olfaction: A Review," *IEEE Sensors J,* vol. 2, pp. 189-202, 2002 2002.
10. S. Ji and L. Carin, "Cost-sensitive feature acquisition and classification," *Pattern Recogn,* vol. 40, pp. 1474-1485, 2007.
11. L. Rabiner and B. Juang, "An introduction to hidden Markov models," *ASSP Magazine,* vol. 3, pp. 4-16, 1986.
12. S. Wlodek, K. Colbow, and F. Consadori, "Signal-shape analysis of a thermally cycled tin-oxide gas sensor," *Sensor Actuator B,* vol. 3, pp. 63-68, 1991.

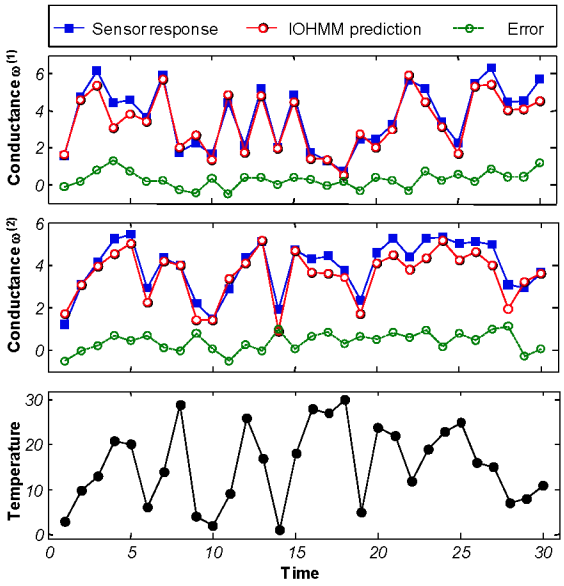

FIGURE 2. Simulated response of the sensor model in eq. (7) for two classes, IOHMM predictions and residuals. The same random temperature sequence was used in all cases for comparison purposes.

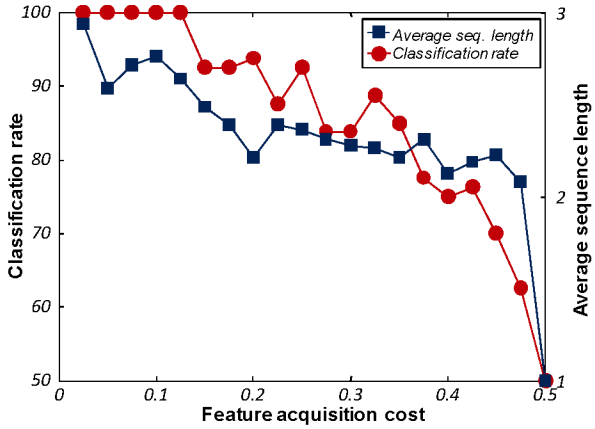

FIGURE 3. Classification performance and average sequence length as a function of feature acquisition costs. Misclassification cost $\boldsymbol{c_{uv}} = \mathbf{1}$ if $\boldsymbol{u} \neq \boldsymbol{v}$; zero otherwise.