Data-driven Modeling of Metal-oxide Sensors with Dynamic Bayesian Networks

Rakesh Gosangi[a] and Ricardo Gutierrez-Osuna[a]

[a] Department of Computer Science and Engineering, Texas A&M University, College Station, TX
{rakesh,rgutier}@cse.tamu.edu

Abstract. We present a data-driven probabilistic framework to model the transient response of MOX sensors modulated with a sequence of voltage steps. Analytical models of MOX sensors are usually built based on the physico-chemical properties of the sensing materials. Although building these models provides an insight into the sensor behavior, they also require a thorough understanding of the underlying operating principles. Here we propose a data-driven approach to characterize the dynamical relationship between sensor inputs and outputs. Namely, we use dynamic Bayesian networks (DBNs), probabilistic models that represent temporal relations between a set of random variables. We identify a set of control variables that influence the sensor responses, create a graphical representation that captures the causal relations between these variables, and finally train the model with experimental data. We validated the approach on experimental data in terms of predictive accuracy and classification performance. Our results show that DBNs can accurately predict the dynamic response of MOX sensors, as well as capture the discriminatory information present in the sensor transients.

Keywords: Metal-oxide sensors, Dynamic Bayesian networks, Sensor modeling.
PACS: 02.50.-r, 07.07.Df

METHODS AND RESULTS

Bayesian networks (BNs) are directed graphical models that represent conditional dependencies among a set of random variables. Dynamic Bayesian networks (DBNs) are an extension of BNs specifically aimed at modeling temporal relations between these random variables. Our sensor model uses four random variables (I, T, H, O), where I is the voltage step applied to a MOX sensor heater, T is the time elapsed from the edge of a voltage step, H is an unobserved variable that captures the underlying state of the sensor, and O is the sensor response. FIGURE 1 (a) shows a graphical representation of the DBN. We validated this model with a Figaro TGS 2600 MOX sensor exposed to three analytes: acetone, ammonia, and isopropyl alcohol. We serially diluted these chemicals to identify concentrations at which the isothermal responses were similar; this ensured that the three chemicals could not be trivially distinguished. We collected sensor responses (sampled at 1Hz) to 150 random voltage sequences, 50 sequences per analyte, and eight voltage steps per sequence. We build three DBN models (one per chemical) using 60 voltage sequences (20 per chemical) as training data, and tested on the remaining 90 sequences. We estimated the performance in terms of predictive accuracy (i.e., model response to a voltage sequence vs. actual sensor response) and classification rate (maximum likelihood

Olfaction and Electronic Nose: Proceedings of the 14th International Symposium on Olfaction and Electronic Nose
AIP Conf. Proc. 1362, 135-136 (2011); doi: 10.1063/1.3626334

criterion) on test sequences. To establish the influence of step duration, we repeated this experiment for τ sec, 2τ sec and 4τ sec, where τ = 5sec is three times the sensor's time constant. FIGURE 1 (b) shows a comparison between the signal predicted by the model and the original response of the sensor. FIGURE 2 (a) shows the average absolute errors in prediction along the 20 sec transient. This result indicates that model predictions are less accurate at the beginning of the voltage steps. Our preliminary analyses also suggest that errors are higher when the sensor is heated to a higher temperature as opposed to when it is cooled down (data not shown). Finally, we also tested the classification performance of the models as a function of the duration and number of voltage steps (FIGURE 2 (b)). As expected, the classification rate increases with increasing number of steps as well as with step duration. In particular, there appears to be a significant difference in performance between 10s and 20s steps, suggesting that steps of at least 15 second duration are needed to fully extract information from the sensor transients.

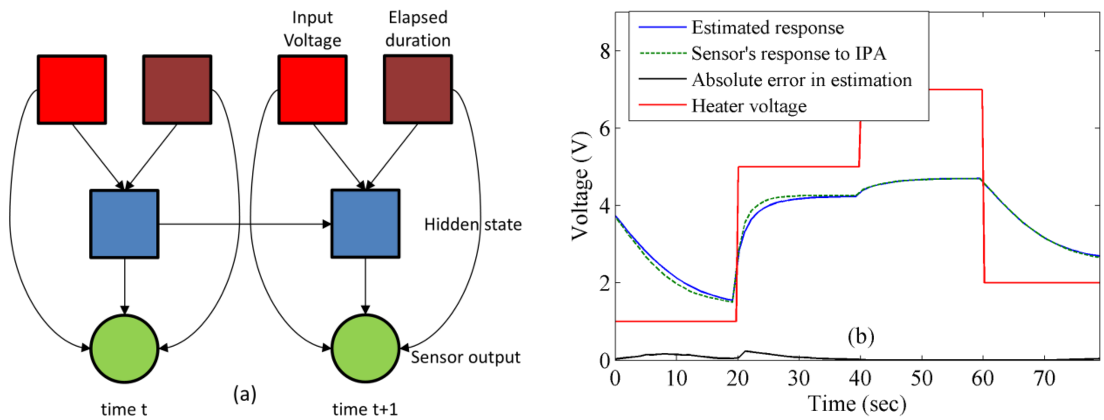

FIGURE 1. (a) Graphical representation of the DBN showing two time slices. **(b)** DBN model predictions vs. actual sensor response to a random sequence in the presence of IPA (SNR= 37 dB).

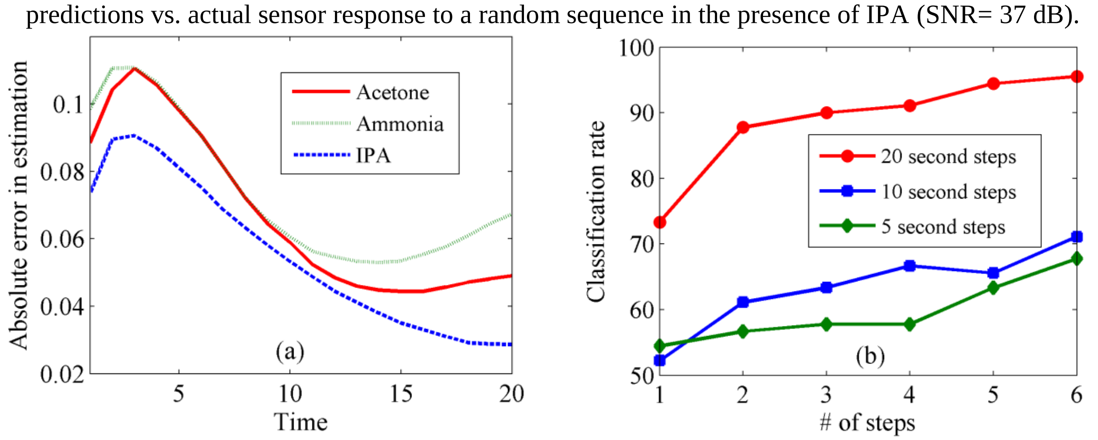

FIGURE 2. (a) Average absolute prediction errors along a 20 sec step. Errors are the highest in the early part of the transient, when the sensor is rapidly heating or cooling down. **(b)** Classification rate obtained by the DBNs on a three-class problem as a function of the number of steps and step durations.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Award Number 1002028.