Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Hidden Markov Model Based Approach for Alarm Rationalization

Joshiba Ariamuthu Venkidasalapathy[a,b]. Costas Kravaris[a]

[a]Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77840, USA (Tel: +1 979-402-9823, e-mail: joshvenkat03@tamu.edu)
[b]Mary Kay O'Connor Process Safety Center, Texas A&M University, College Station, TX 77840, USA

Abstract: Alarm rationalization is a key element in ISA 18.2 alarm management lifecycle. During an abnormal event, alarms are generated in the control room to alert the operator of the affected regions in the process. An important objective of rationalization is to guide the operator troubleshoot quickly and help take necessary correction actions to restore normal operation. In this paper, the idea of representing the fault propagation path by the sequence of alarms generated by it, is explored. A model-based approach based on Hidden Markov Model (HMM) is proposed to predict the most likely cause of alarms using the alarm sequence generated. The probabilistic framework of HMM paves way to account for stochastic features of real plant operations that may arise due to random noises in sensors as well as the effect of fault magnitudes on sequences. The approach was applied to an industrial case study: Vinyl Acetate Monomer production process. The results show that the proposed approach was successful in predicting the probable cause of alarms generated with high accuracy. The model was able to predict the cause with reasonable accuracy even when tested with short alarm sub-sequences. This allows for early identification of faults, providing more time to the operator.

Keywords: Alarm systems, Markov models, Probabilistic models, Fault diagnosis, Fault isolation, Classification, Sequences

1. INTRODUCTION

Alarm systems play an important role in ensuring safe and efficient operation in process industries. Alarm rationalization is an element in ISA 18.2 alarm management lifecycle(ANSI, 2009). Guiding the operator is crucial especially when there is a high alarm rate, a situation also known as alarm flooding. The primary objective of this paper is to develop a model-based approach to determine the potential cause of an alarm sequence.

Alarm rationalization is typically done with the help of cause-effect based techniques such as HAZOPs and HAZIDs to determine a list of possible causes and corrective actions for each alarm. These methods are not sufficient to capture the complex interactions between process variables. Recently, data-driven methods that utilize historical alarm logs have been proposed(Lucke et al., 2019, Dorgo et al., 2018). Although these methods diagnose the possible cause of the alarm sequence with reasonable accuracy, lack of interpretability associated with these approaches is a limitation in its application as an operator guiding tool.

The idea of representing the sequence of alarms generated as the fault propagation path is exploited in this paper to diagnose the possible cause of the alarm sequence. The alarm data analysis can be broadly classified into time series analysis and sequence mining methods(Lucke et al., 2019). Feature based approach, sequence distance based approach and model based approach are the three common techniques available for sequence classification(Xing et al., 2010). Feature based approaches transform a sequence into a feature vector and then apply conventional classification methods such as k-nearest neighbours, neural networks or support vector machines(Dorgo et al., 2018). The sequence distance-based approaches aim to capture the similarity between sequences to determine the class. The model-based approach assumes sequences in a class are generated by an underlying model(Ahmed et al., 2013, Charbonnier et al., 2016). It models the probability distributions of sequences in each class. Naive Bayes sequence classifier, Markov Model, Hidden Markov Model (HMM) and other statistical models are used for this purpose. In this paper, an HMM based approach is proposed for alarm sequence classification for the following reasons: (i). The alarm sequences generated for faults depend on the magnitude of fault and sensor noises. There is a need for incorporating stochastic features to account for the uncertainties and randomness in processes and (ii). The model-based approach is relatively easier to interpret unlike other purely data driven methods.

The paper is organized as follows: in section 2 a brief overview of HMM and the approach proposed for application in alarm rationalization is outlined. The approach is applied to vinyl acetate monomer production process and the results are discussed in section 3. Lastly, the conclusions and recommendations are presented in section 4.

2. APPROACH PROPOSED FOR ALARM RATIONALIZATION

2.1 Brief overview of HMM

HMM is a probabilistic model that assumes the states follow the Markov property and are unobservable(Kouemou and Dymarski, 2011). The outputs that depend on the state are observable. The state transition probabilities and emission probabilities are estimated in HMM. HMM is used to predict the hidden state sequence given the measured output sequence. Their applications include speech recognition, facial expression identification and bioinformatics.

The model parameters include the following: the number of distinct states N, the state transition probability matrix A, the number of measurement variables at each state M, the emission probability matrix B and the probability distribution of the initial state $\prod$. The N states can be denoted by $S = \{S_1, S_2, \ldots, S_N\}$, the observed measurements from states by $V = \{v_1, v_2, \ldots, v_M\}$. The hidden state and the measured output at time t are represented by x_t and y_t respectively. The state transition probability matrix $A_{ij} = \{a_{ij}\}$, where a_{ij} is the probability that the state evolves from state S_i to state S_j.

$$a_{ij} = p\{x_t = S_j | x_{t-1} = S_i\}, (1 \le i, j \le N) \qquad (1)$$

The emission probability distribution in each state $B = \{b_j(k)\}$ is the probability of observing output v_k from state S_j.

$$b_j(k) = p\{y_t = v_k | x_t = S_j\}, (1 \le j \le N, 1 \le k \le M) \qquad (2)$$

The initial state distribution $\prod_i$ is the probability that the model is at state S_i at t = 0. It is given by,

$$\textstyle\prod_i = p\{x_0 = S_i\}, (1 \le i \le N) \qquad (3)$$

2.2 Summary of algorithms for HMM learning and decoding of hidden state sequence

In this section, algorithms available for model learning and decoding hidden state sequence is briefly summarized. The Baum-Welch algorithm is commonly used for learning the HMM parameters. The Viterbi algorithm is used for identifying the most likely hidden state sequence given an observation sequence.

The Baum-Welch algorithm is also known as the forward-backward algorithm. In the initial step, the parameters A, B and $\prod$ are initialized. They are randomly assigned if there is no prior knowledge. Given the initial model parameters and the observed sequence $Y_{0:T}$, the following steps are performed. In the forward step, $\alpha(x_k)$, which is the joint probability distribution of states at time k given the observables sequence up to time k, is calculated. In the backward step, $\beta(x_k)$, which is the conditional probability of the observed data from time k+1 given the state at time k, is calculated. The forward and backward steps are combined to calculate the joint probability distribution of a state at time k given $Y_{0:T}$. The joint probability of observing two consecutive states given $Y_{0:T}$ is also calculated. Finally, in the update step, the model parameters are updated to maximize the probability of obtaining the observed sequence. Viterbi algorithm helps make an inference based on a trained model and the observation sequence. It uses a recursive algorithm to identify the hidden state sequence.

2.3 Approach proposed for alarm rationalization

A safety review technique needs to be performed to identify possible faults (hazardous scenarios). Using the plant historic data or mathematical model of the plant, the alarm sequences can be extracted for the identified faults. If using a closed loop simulator to generate data, a range of fault magnitudes can be simulated to capture different possible fault propagation paths. Modelling random measurement noises will help capture the realistic stochastic aspects of plant operation. Model faults as states and the alarms as observable outputs. Split data into training and test datasets to learn and validate the model respectively. Software such as MATLAB or R can be used for training and predicting hidden state sequences.

3. CASE STUDY

3.1 Process description and faults characterization

The proposed approach for alarm rationalization is applied to the industrial case study: the Vinyl Acetate Monomer (VAM) process that was presented by Luyben and Tyreus (Luyben and Tyréus, 1998). A nonlinear public domain model of the VAM process based on Luyben's multi-loop control structure (Luyben et al., 1999), developed in C language and implemented in MATLAB environment (Chen et al., 2003), is used to obtain the alarm sequences data required for training and testing the HMM. This process consists of 246 states, 26 manipulated variables and 43 measurements.

Measurement noises for the sensors of type liquid level gauge, pressure gauge and composition analyzers were introduced in the MATLAB model as uniformly distributed random numbers of magnitude 0.1 %, 0.25 % and 0.5 % of the steady state value respectively. The amplitude of noise for temperature gauges were fixed at 2 K based on commercially available sensors. For this case study, an alarm is set to trigger when the measured variables go beyond 4 standard deviations from the mean value measured at normal operating conditions.

A list of seven faults identified for the VAM process are shown in Table 1. Faults are identified based on valve stiction and sensor malfunctions (faults 1-7). The high and low alarm limits for measurements are distinguished to enhance the classification accuracy. For modeling purposes, the high alarm limits of the measurements are numbered 1- 43 and the corresponding low alarm limits are numbered 44-86. The nonlinear dynamic model was used to simulate these scenarios for different fault magnitudes. The observed alarm sequences for a fault varied in the order in which the alarms were generated as well as the length of sequences. A total of 36 fault scenarios were simulated for learning the HMM parameters. In addition to this, 21 different fault scenarios were simulated for evaluating the performance of HMM.

Table 1. Faults list along with their type, number, and length of sequences in training dataset

#	Fault description	Type	No. of seqs.	Length of seqs.
1	Absorber level gauge drift	Sensor	7	13-16
2	Vaporizer level gauge drift	Sensor	3	32-34
3	Column fifth tray temp gauge drift	Sensor	5	16-18
4	Reactor exit temperature gauge drift	Sensor	6	31-32
5	Steam drum pressure set point increases	Actuat or	3	31-32
6	Column reboiler duty stuck	Actuat or	8	11-15
7	CO2 removal unit- Purge flow valve stuck	Actuat or	4	20-25

3.2 Results and discussions

The faults are modelled as states and the alarm sequences are modelled as the observed output sequence for the HMM. HMM parameters were trained using the MATLAB Statistics and Machine Learning Toolbox that uses the Baum-Welch algorithm. All states are assumed to be equally likely for being the initial state. The state transition probability and the emission probability matrices are pictorially shown in figure 1 and figure 2 respectively.

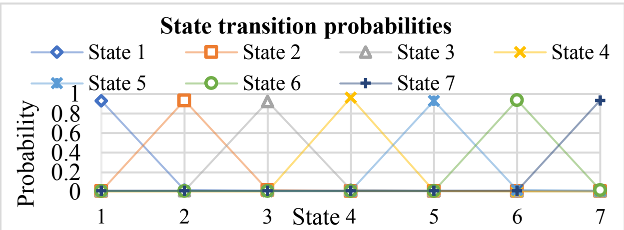

Fig. 1. State transition probabilities

It is seen from the state transition matrix in figure 1 that the probability of staying in state S_1 is close to one. This can be justified since the faults occur one at a time. The emission probability matrix in figure 2 shows the probability of measurements generating an alarm given the fault.

The Viterbi algorithm in the MATLAB toolbox was used to predict the hidden state sequence. Given that the simulations were generated for one fault at a time, it is reasonable to define the state corresponding to the given alarm sequence as the state with maximum frequency in the predicted state sequence.

The HMM performance is evaluated on its efficiency in predicting the fault correctly as early as possible. The first p elements of the sequences from the test dataset was used to identify the most likely state sequence, and subsequently the corresponding fault. The prediction accuracies are plotted as a function of p first elements of the sequence for p = {1, 2, … 10} in figure 3. As p increases, the prediction accuracy increases up to ~86%.

The true and estimated state for the 21 test sequences for p = 10 is shown in figure 4. It is seen that true state and predicted state matched for all states except for state 3, where state 3 is predicted as state 6. From the fault list in table 1 it is seen that faults 3 and 6 are the sensor and actuator faults originating from the same control loop. This suggests a problem of distinguishability between faults with similar propagation paths. To improve the prediction accuracy, the following modifications were done to the HMM:

- Fault 6 and fault 3 together will be modelled as state 3
- Fault 4 and fault 5 together will be modelled as state 4 as they originate from the same control loop

Grouping the faults originating from the same control loop together helps narrow down the faulty region precisely, rather than trying to predict the specific fault ambiguously. The reduced 5 state model was trained and tested. The prediction accuracy results are shown in figure 5. It is seen that the prediction accuracy increases to 100% for p > 7. An accuracy of 90% for p = 5 suggests that the proposed approach can be used for early prediction of faults with high accuracy.

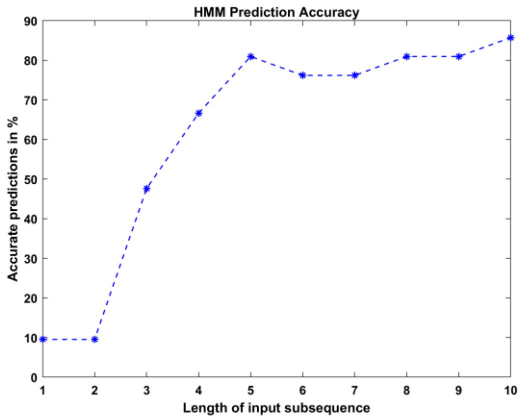

Fig. 2. HMM prediction accuracy for the test dataset as a function of the length of input subsequence

4. CONCLUSIONS

The HMM based approach for alarm rationalization is proposed to guide the operator for early troubleshooting of the abnormal scenario. The model-based approach that is proposed in this paper analyses the fault propagation path to identify the most probable cause. The probabilistic framework provides scope to include the effects of differences in propagation paths that arise because of different fault magnitudes and random sensor noises.

Application of the HMM approach on the Vinyl Acetate Monomer process showed that the model was successful in early prediction of the fault with reasonable accuracy. A distinguishability problem between faults were observed when they originated from the same control loop. The model was modified to lump faults from one control loop as one fault. The reduced model helped improve the prediction accuracy from 80% to 90% for a sequence length of 5. The HMM approach seems promising for application in alarm rationalization. Future work will focus on building model-based alarm systems, using the concept of dynamic safety sets proposed in (Venkidasalapathy and Kravaris, 2020) to characterize process safeness. An optimization formulation proposed by the authors for alarm identification in (Venkidasalapathy et al., 2018) will be extended to simultaneously optimize alarm identification and rationalization to aid the operator in quick troubleshooting without alarm overload.

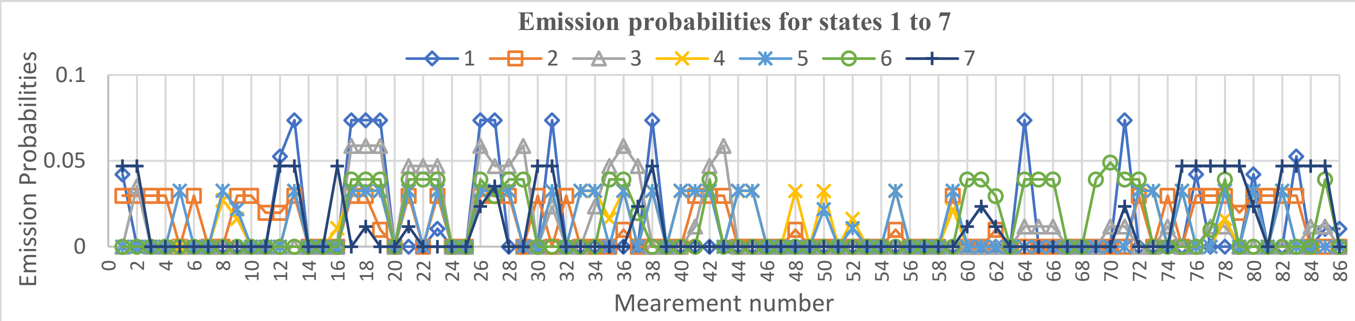

Fig. 3. Emission probabilities of 43 measurements from each fault

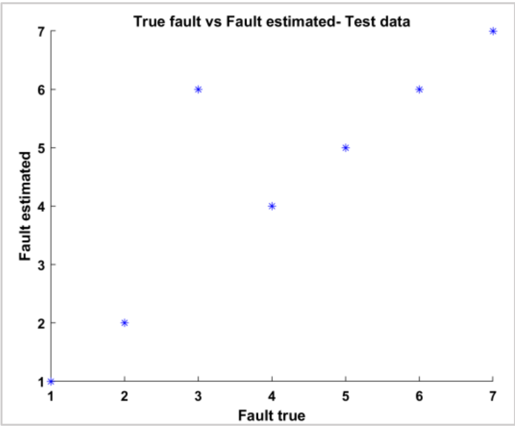

Fig. 4. True fault vs predicted fault with length of subsequence being 10

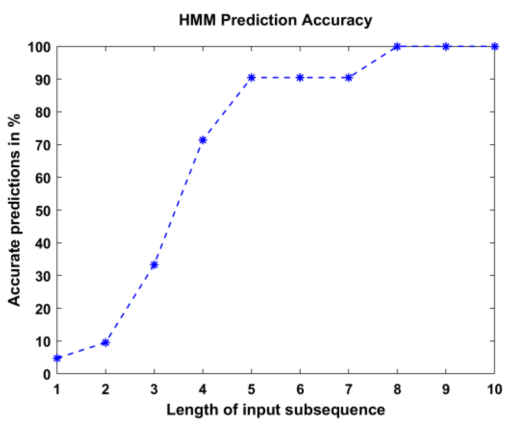

Fig. 5. HMM prediction accuracy for the test dataset for the reduced model

REFERENCES

AHMED, K., IZADI, I., CHEN, T., JOE, D. & BURTON, T. 2013. Similarity analysis of industrial alarm flood data. *IEEE Transactions on Automation Science and Engineering,* 10**,** 452-457.

ANSI, I. 2009. ISA-18.2-2009 Management of alarm systems for the process industries. *Durham, NC, USA: International Society of Automation*.

CHARBONNIER, S., BOUCHAIR, N. & GAYET, P. 2016. Fault template extraction to assist operators during industrial alarm floods. *Engineering Applications of Artificial Intelligence,* 50**,** 32-44.

CHEN, R., DAVE, K., MCAVOY, T. J. & LUYBEN, M. 2003. A nonlinear dynamic model of a vinyl acetate process. *Industrial & engineering chemistry research,* 42**,** 4478-4487.

DORGO, G., PIGLER, P. & ABONYI, J. 2018. Understanding the importance of process alarms based on the analysis of deep recurrent neural networks trained for fault isolation. *Journal of Chemometrics,* 32**,** e3006.

KOUEMOU, G. L. & DYMARSKI, D. P. 2011. History and theoretical basics of hidden Markov models. *Hidden Markov Models, Theory and Applications,* 1.

LUCKE, M., CHIOUA, M., GRIMHOLT, C., HOLLENDER, M. & THORNHILL, N. F. 2019. Advances in alarm data analysis with a practical application to online alarm flood classification. *Journal of Process Control,* 79**,** 56-71.

LUYBEN, M. L. & TYRÉUS, B. D. 1998. An industrial design/control study for the vinyl acetate monomer process. *Computers & Chemical Engineering,* 22**,** 867-877.

LUYBEN, W. L., TYRÉUS, B. D. & LUYBEN, M. L. 1999. *Plantwide process control*, McGraw-Hill.

VENKIDASALAPATHY, J. A. & KRAVARIS, C. 2020. Safety-centered process control design based on dynamic safe set. *Journal of Loss Prevention in the Process Industries*, 104126.

VENKIDASALAPATHY, J. A., MANNAN, M. S. & KRAVARIS, C. 2018. A quantitative approach for optimal alarm identification. *Journal of loss prevention in the process industries,* 55**,** 213-222.

XING, Z., PEI, J. & KEOGH, E. 2010. A brief survey on sequence classification. *ACM Sigkdd Explorations Newsletter,* 12**,** 40-48.