A Heuristic Optimization Approach for HMM in Machinery Prognostics and Health Management

Wen-Zhu Liao †
Department of Industrial Engineering
Chongqing University, Chongqing, China
Tel: (+86) 13629771031, Email: liaowz@cqu.edu.cn

Shi-Hao Cui
Department of Industrial Engineering
Chongqing University, Chongqing, China
Tel: (+86) 13629771031, Email: 20150702070@cqu.edu.cn

Dan Li
Department of Industrial Engineering
Chongqing University, Chongqing, China
Tel: (+86) 13629771031, Email: ie_lidana@cqu.edu.cn

Abstract. This paper proposes a heuristic optimization approach for hidden Markov model (HMM) in machinery diagnosis based on simulated annealing (SA) algorithm and expectation maximization (EM) algorithm. As traditional HMM is sensitive to initial values and EM is easy to trap into a local optimization, SA is combined to improve HMM which can overcome local optimization searching problem. The proposed HMM has strong ability of global convergence, and optimizes the process of parameters estimation. Finally, through a case study, the computation results illustrate this SAEM-HMM has high efficiency and accuracy, which could help machinery diagnosis in practice.

Keywords: Hidden Markov model, Expectation maximization, Simulated annealing, Diagnosis

1. INTRODUCTION

Nowadays, HMM has been widely used in machinery prognostics and health management, due to its good performance about model construction and pattern classification for dynamic time sequence (Deng et al., 2013). However, previous HMM-based machinery diagnosis research focuses on applications of basic HMM theory mainly including two parts: improving HMM to better fit machine degradation in real production process, such as semi-HMM (Dong and He, 2007), multi branch-HMM (Thanh et al., 2014), etc; using feature extraction method for traditional HMM to achieve accurate diagnosis results (Xu et al., 2015; Zhou et al., 2015). However, these research cannot avoid that the traditional HMM is sensitive to initial value. Moreover, the traditional training algorithm is easy to trap into a local optimization.

Nowadays, some research has begun to consider the optimization of HMM training algorithm, which is focused on human behavior recognition, speech recognition (Huda et al., 2014), visual speech recognition (Lee and Park, 2006), etc. Some heuristic algorithms are studied to search global optimum solution such as genetic algorithm (GA) (Li and, 2013) and Tabu search (TS), but general application reflects their performance to a certain extent. Thus, HMM training algorithm needs to be further studied.

SA which originates from a physical annealing process
can solve the global optimization problem and get better solution. This paper is devoted to optimize HMM process by combining SA with EM. The local optimization problem of model parameters estimated by EM can be overcome by SA. Moreover, it can avoid the sensitivity of initial values selection.

2. HIDDEN MARKOV MODEL

HMM presents a double stochastic process based on general Markov chain. In general, it can be formulated as $\lambda = (N, M, \pi, A, B)$. Its parameters are defined as follow.  

1) $N$ is the states number of hidden Markov chain, where the states are $\theta_1, \theta_2, \ldots, \theta_N$. Suppose the state of hidden Markov chain at time $t$ is recorded as $q_t$.

2) $M$ is the possible number of observation values for each state, where the observation values are $v_1, v_2, \ldots, v_M$. Suppose the observation value at time $t$ is recorded as $o_t$.

3) $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ is the probability distribution of initial states, where

\[
\pi_i = P(q_1 = \theta_i) \sum_{i=1}^{N} \pi_i = 1 \quad 1 \leq i \leq N \quad (1)
\]

4) $A = (a_{ij})_{N \times N}$ is the probability matrix of states transition, and $a_{ij}$ represents the transition probability from state $i$ to state $j$, shown as

\[
a_{ij} = P(q_{t+1} = \theta_j | q_t = \theta_i) \sum_{i=1}^{N} a_{ij} = 1 \quad 1 \leq i \leq N \quad 1 \leq j \leq N \quad (2)
\]

5) $B = (b_{jk})_{N \times M}$ is the probability matrix of observation values, and $b_{jk}$ represents the probability of observation value $v_k$ at time $t$ for state $j$, shown as

\[
b_{jk} = P(o_t = v_k | q_t = \theta_j) \sum_{j=1}^{M} b_{jk} = 1 \quad 1 \leq j \leq N \quad 1 \leq k \leq M \quad (3)
\]

The above HMM can be abbreviated as $\lambda = (\pi, A, B)$ and EM algorithm is conventionally used to train HMM. The EM algorithm is used to search the maximum likelihood value of HMM parameters. The parameters of HMM can be reevaluated, shown as

\[
\pi_i = \gamma_i(t) \quad (4)
\]

\[
a_{ij} = \frac{\frac{\sum_{t=1}^{T-1} \xi(i, j)}{\sum_{i=1}^{T-1} \gamma_i(t)}}{\sum_{i=1}^{T-1} \gamma_i(t)} \quad (5)
\]

\[
\bar{b}_{jk} = \frac{\sum_{t=1}^{T-1} \gamma_j(t)}{\sum_{t=1}^{T-1} \gamma_j(t)} \quad (6)
\]

where $\pi_i$, $a_{ij}$, $\bar{b}_{jk}$ are the model parameters of $\lambda$. The forward probability variable $\alpha_i(t)$, the backward probability variable $\beta_i(t)$ and the probability variable $\gamma_i(t)$ are defined as follows:

\[
\alpha_i(t) = p(o_t, o_{t+1}, \ldots, o_N | q_t = \theta_i) \quad (7)
\]

\[
\beta_i(t) = p(o_t, o_{t+1}, \ldots, o_N | q_t = \theta_i, \lambda) \quad (8)
\]

\[
\gamma_i(t) = p(q_t = \theta_i | \lambda) \quad (9)
\]

where $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, N$.

And \( \xi(i, j) \) is the probability of machine state in $\theta_i$ at time $t+1$ while machine state is $\theta_j$ at time $t$, shown as

\[
\xi(i, j) = p(q_{t+1} = \theta_j | q_t = \theta_i, q_{t+1} = \theta_j | O, \lambda) \quad (10)
\]

Hence, there are,

\[
\alpha_i(t) = \left\{ \begin{array}{ll} \pi_i b_{i1} & , \quad t = 1, 1 \leq i \leq N \\ \sum_{j=1}^{N} a_{ij} \alpha_{j,t-1}(j) b_{j,t+1} & , \quad 2 \leq t \leq T-1, 1 \leq i \leq N \end{array} \right. \quad (11)
\]

\[
\beta_i(t) = \left\{ \begin{array}{ll} 1 & , \quad t = T, 1 \leq i \leq N \\ \sum_{j=1}^{N} a_{ji} b_{j,t+1} \beta_{j,t+1}(j) & , \quad 1 \leq t \leq T-1, 1 \leq i \leq N \end{array} \right. \quad (12)
\]

\[
P(O | \lambda) = \sum_{t=1}^{T} \sum_{j=1}^{N} \alpha_i(t) a_{ij} b_{j,t+1} \beta_{j,t+1}(j) \quad (13)
\]

\[
\xi(i, j) = \frac{\alpha_i(t) a_{ij} b_{j,t+1} \beta_{j,t+1}(j)}{\sum_{j=1}^{N} \sum_{i=1}^{N} \alpha_i(t) a_{ij} b_{j,t+1} \beta_{j,t+1}(j)} \quad (14)
\]

The new model is reconstructed to be $\tilde{\lambda} = (\tilde{\pi}, \tilde{A}, \tilde{B})$. Through new iterations until $p(O | \tilde{\lambda}) - p(O | \lambda) < \delta$, the optimal $\tilde{\lambda}_{EM}$ is obtained, where $\delta$ (which generally is set to $10^{-4}$) is predetermined convergent condition.

3. SAEM TRAINING ALGORITHM

As HMM trained by EM may trap into local optimization due to its initial values and SA can help to solve global optimization problem by using Metropolis criterion
and controlling the deterioration process of temperature
(Huda et al., 2009), a novel method combing SA and EM is
developed in this paper. Metropolis criterion means when the
parameter $\lambda_{\text{old}}$ changes into $\lambda_{\text{new}}$, the system energy should
be changed from $E(\lambda_{\text{old}})$ to $E(\lambda_{\text{new}})$ respectively. The
acceptance probability of $p$ is
\[
p = \begin{cases} 
1, & \text{if } E(\lambda_{\text{new}}) < E(\lambda_{\text{old}}) \\
\exp\left(\frac{E(\lambda_{\text{old}}) - E(\lambda_{\text{new}})}{T_0}\right), & \text{if } E(\lambda_{\text{new}}) > E(\lambda_{\text{old}}) 
\end{cases}
\]  
(15)
where $T_0$ is initial temperature given by empirical formula.

And $E(\lambda) = -\log P(O|\lambda)$ is selected to be energy
function, thus the optimal parameter $\lambda_{\text{final}}$ of HMM can be
obtained by minimizing $E(\lambda)$. The training steps of this
proposed heuristic algorithm are given as follows:

(i) Set initial values of the parameter $\lambda$ for HMM and
assum $T_0$ to be the standard deviation of selected samples. Assume the times
of temperature annealing is $s_{\text{num}}=1$ and the times of
iteration is $\text{iter}_\text{num}=1$.

(ii) Take the parameter $\lambda$ as $\lambda_{\text{old}}$. According to Eqs.(11)-(13), calculate
$P(O|\lambda_{\text{old}})$. Then, $E(\lambda_{\text{old}})$ can be obtained by reversing its logarithm.

(iii) Take the reevaluated parameter $\lambda$ obtained by Eqs.(4)-(6) as $\lambda_{\text{new}}$. According to Eqs.(11)-(13), calculate
$P(O|\lambda_{\text{new}})$. Then, $E(\lambda_{\text{new}})$ can be obtained by reversing its logarithm.

(iv) If $E(\lambda_{\text{new}}) < E(\lambda_{\text{old}})$, the new solution is accepted and
the times of temperature annealing is $s_{\text{num}}=s_{\text{num}}+1$. Otherwise, go to the next step.

(v) Calculate $p$ by Eq.(15). And use Rand() function to generate a constant $c$ less than 1 randomly. If $p > c$, the new solution is accepted. Otherwise, reject the new
solution and the algorithm ends.

(vi) Address the annealing method by Eq.(16). If the
terminal condition is satisfied, go to the next step.
Otherwise, assume the times of iteration is
$\text{iter}_\text{num} = \text{iter}_\text{num} + 1$, then go to step (iii).

(vii) Return the final model $\lambda_{\text{final}}$, End the algorithm.

4. MACHINERY DIAGNOSIS BASED ON
PROPOSED HMM

Machinery diagnosis mainly includes three processes:
data analysis, HMM construction and state evaluation. The
observation sequence for HMM construction training could
be processed by monitoring data of the machine. Based on
optimal HMM parameters determined, the current state of the
machine could be estimated by inputing real-time
observation data.

As machine deterioration process is hard to describe, this
paper classifies its health state into five levels: healthy, good,
normal, unhealthy and broken, and then maps the relationship
between health state and machine performance data so as to
train and evaluate HMM. This paper optimizes HMM by
using the SAEM algorithm. The optimal $\lambda_{\text{final}}$ of the model
could be obtained by the proposed optimization approach.

Then the observation sequence given by the real-time
collected data of the machine can be inputted into this
evaluation model to estimate machine health state. The
probability that the observation sequence is in state $\theta_i$ at
time $t$ can be calculated as
\[
\overline{y}_i(t) = \sum_{j=1}^{N} \overline{y}(t, j) = \frac{\overline{y}(i, j)}{P(O|\lambda_{\text{final}})} 
\]  
(17)

According to maximum membership principle, machine
health state at time $t$ calculated by the Eq.(17), which can be
described by a column map. Therefore, one state curve
used to describe machine health state at different time $t$
could well depict machine degradation.

5. A CASE STUDY

This paper takes a drilling machine as a case study to apply
SAEM-HMM in machinery diagnosis, in which its
spindle voltage signal is used to indicate machine
performance. According to the method of reference (Liao,
2011), 500 groups of spindle voltage signal are collected.
Based on the data, machine health state is divided into 5
levels (as Table 1). After processing, machine health index
can be obtained.

<table>
<thead>
<tr>
<th>Health state</th>
<th>Healthy</th>
<th>Good</th>
<th>Normal</th>
<th>Unhealthy</th>
<th>Broken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health index</td>
<td>0.85~1</td>
<td>0.7~0.85</td>
<td>0.5~0.7</td>
<td>0.3~0.5</td>
<td>0.3~0.3</td>
</tr>
</tbody>
</table>

As this proposed approach focuses on the sensitivity to
initial values of traditional Markov model, therefore model
parameters can be randomly given. Then machine health
index considered as the observation sequence is taken into EM algorithm and SAEM algorithm for training, \( \hat{\lambda}_{EM} = (\pi_{EM}, A_{EM}, B_{EM}) \) and \( \hat{\lambda}_{final} = (\pi_{final}, A_{final}, B_{final}) \) can be obtained respectively. In order to validate this proposed SAEM-HMM, 50 groups of data (seen in Table 2) are selected to calculate the probability while the machine in different states.

Then, machine health state could be estimated by maximum membership principle. The comparison of results is shown as Table 3, which can prove this proposed algorithm achieve much higher accuracy.

Table 2: The evaluation results of machine health states.

<table>
<thead>
<tr>
<th>State Index</th>
<th>Healthy</th>
<th>Good</th>
<th>Normal</th>
<th>Unhealthy</th>
<th>Broken</th>
<th>Healthy</th>
<th>Good</th>
<th>Normal</th>
<th>Unhealthy</th>
<th>Broken</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.89</td>
<td>0.79</td>
<td>0.18</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
<td>0.82</td>
<td>0.18</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.87</td>
<td>0.57</td>
<td>0.42</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0.63</td>
<td>0.37</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.85</td>
<td>0.51</td>
<td>0.47</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0.55</td>
<td>0.446</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.62</td>
<td>0.16</td>
<td>0.36</td>
<td>0.48</td>
<td>0.06</td>
<td>0</td>
<td>0.05</td>
<td>0.465</td>
<td>0.464</td>
<td>0.021</td>
<td>0</td>
</tr>
<tr>
<td>0.41</td>
<td>0</td>
<td>0</td>
<td>0.02</td>
<td>0.68</td>
<td>0.30</td>
<td>0</td>
<td>0.004</td>
<td>0.026</td>
<td>0.65</td>
<td>0.32</td>
</tr>
<tr>
<td>0.30</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.53</td>
<td>0.46</td>
<td>0</td>
<td>0.007</td>
<td>0.013</td>
<td>0.37</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Table 3 Comparison of results

<table>
<thead>
<tr>
<th>Health index</th>
<th>Evaluation results</th>
<th>EM algorithm</th>
<th>SAEM algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Right=1 &amp; Wrong=0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.89</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0.74</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0.62</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.53</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Rate of accuracy</td>
<td>47/50</td>
<td>49/50</td>
<td></td>
</tr>
</tbody>
</table>

6. CONCLUSION

In this paper, a heuristic algorithm considering SA and EA is developed to optimize HMM for machinery diagnosis. With the advantage of global convergence by SA, the local optimization searching problem of model parameters estimated by EM can be overcome. Moreover, it avoids the issue of initial values selection. Through a case study, the computation results indicate that this proposed algorithm can achieve high accuracy, which makes this new approach efficient and applicable in machinery diagnosis. Later, in order to achieve good machinery prognostics and health management, the future work will focus on the extension of this optimization model.
ACKNOWLEDGMENTS

The authors would like to thank anonymous referees for their remarkable comments and great support by National Natural Science Foundation of China (71301176) and Research Fund for the Doctoral Program of Higher Education of China (No. 20130191120001)

REFERENCES


